ESAIM: COCV 24 (2018) 1849–1879
https://doi.org/10.1051/cocv/2017002

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

LINEAR QUADRATIC CONTROL PROBLEMS OF STOCHASTIC VOLTERRA INTEGRAL EQUATIONS *

Tianxiao Wang**

Abstract. This paper is concerned with linear quadratic control problems of stochastic differential equations (SDEs, in short) and stochastic Volterra integral equations (SVIEs, in short). Notice that for stochastic systems, the control weight in the cost functional is allowed to be indefinite. This feature is demonstrated here only by open-loop optimal controls but not limited to closed-loop optimal controls in the literature. As to linear quadratic problem of SDEs, some examples are given to point out the issues left by existing papers, and new characterizations of optimal controls are obtained in different manners. For the study of SVIEs with deterministic coefficients, a class of stochastic Fredholm−Volterra integral equations is introduced to replace conventional forward-backward SVIEs. Eventually, instead of using convex variation, we use spike variation to obtain some additional optimality conditions of linear quadratic problems for SVIEs.

Mathematics Subject Classification. 93E20, 49N10, 45D05.

Received January 27, 2016. Accepted January 5, 2017.

1. Introduction

In this paper, we investigate the controlled linear stochastic Volterra integral equation (SVIE, in short) of the following form

$$X(t)=\varphi(t)+\int_0^t[A_1(t,s)X(s)+B_1(t,s)u(s)]\mathrm{d}s+\int_0^t[A_2(t,s)X(s)+B_2(t,s)u(s)]\mathrm{d}W(s), \tag{1.1}$$

with a quadratic cost functional of the form

$$J(0,x;u(\cdot))=\frac{1}{2}\mathbb{E}\int_0^T\big[X(t)^\top Q(t)X(t)+2u(t)^\top S(t)X(t)+u(t)^\top R(t)u(t)\big]\mathrm{d}t+\frac{1}{2}\mathbb{E}X(T)^\top GX(T). \tag{1.2}$$

Keywords and phrases. Stochastic Volterra integral equations, stochastic Fredholm−Volterra integral equations, stochastic linear quadratic problems, spike variation.

* *This research is partially supported by the NNSF of China (Grant 11231007, 11301298, 11471231, 11401404, 11371226), China Postdoctoral Science Foundation (2014M562321). The author highly appreciate the anonymous referees for their constructive comments which led to this improved version.*

School of Mathematics, Sichuan University, Chengdu 610064, China.

** Corresponding author: xiaotian2008001@gmail.com

The problem of minimizing (1.2) subject to (1.1) is called a linear quadratic (LQ, in short) optimal control problem for SVIE (1.1). If φ, A_i, B_i do not depend on t, (1.1) becomes a controlled linear stochastic differential equation (SDE, in short), and the problem degenerates into the so-called linear quadratic optimal control problem for SDEs.

As an extension of SDE, SVIE of the form

$$X(t) = \varphi(t) + \int_0^t b(t, s, X(s))\mathrm{d}s + \int_0^t \sigma(t, s, X(s))\mathrm{d}W(s), \qquad t \in [0, T], \tag{1.3}$$

was studied in *e.g.* [3, 19]. Related topics along this include: SVIEs with anticipated stochastic integral ([1, 17]); SVIEs or VIEs in economic models ([9, 10, 12]); SVIEs in infinite dimensional spaces and related PDEs ([29]), etc. In this paper we will study the linear quadratic optimal control problems of SVIEs. Optimal control problems of deterministic VIEs were received a lot of attentions after the paper [23], see *e.g.* [2, 5, 12, 18]. The general stochastic case was recently studied in [21, 26, 27]. When dealing with optimal control problems of SDEs with control domain $\mathbb{R}^m$, spike variation is more advantageous than convex variation in the sense that the former approach can provide additional useful necessary optimality conditions. However, until now the spike variation and related tricks have not been spread out successfully in the SVIEs setting. Actually, for the corresponding study on SDEs, Itô formula and second-order adjoint equations play indispensable roles. However, their usefulness in the SVIEs setting becomes quite challenging due to the dependence of φ, A_i, B_i on parameter t. To our best, how to provide more information among different coefficients by spike variation is open even under the linear quadratic framework associated with (1.1) and (1.2).

In this paper, we study linear quadratic problem of SVIEs from the view of open-loop optimal controls, the reasons of which are based on the following two aspects. In contrast with SDEs case, for SVIEs there is an additional time variable t appearing in both drift and diffusion terms, see (1.3). This simple change makes one of the most important notions in control theory, feedback optimal controls, become absent in the SVIEs framework. For example, even for deterministic VIEs case, the usual feedback controls and related Riccati equations are replaced by *causal feedback controls* and some *Fredholm integral equations* (see [18]). Moreover, even under the particular SDEs setting, it is shown in this paper that optimal feedback controls become ill-posed while the control problems are solvable. In other words, for these control problems we have to turn to investigating optimal open-loop controls. The second motivation is concerned with the natural connections among (open-loop) optimal controls, necessary optimality conditions and spike variation. It is our belief that more deeper explorations in this topic, such as the seeking for counterparts of Itô formula and second-order adjoint equations, can help us a lot in treating the interesting problem mentioned in the last paragraph.

As to linear quadratic problems for SVIEs, there are indeed new features and obstacles arising which are essentially distinctive from the SDEs case. The first one is concerned with the indefiniteness of $R(\cdot)$. Actually, in the SDEs setting, when the diffusion terms depend on control variables, it was pointed out in [7] that the optimal control problem may still make sense even though $R(\cdot)$ is indefinite. The appearance of indefinite $R(\cdot)$ in solvable linear quadratic problems relies heavily on optimal feedback controls where stochastic Riccati equations can play substantial roles. However, as we mentioned before, even for LQ problem of deterministic VIEs, there is no traditional feedback optimal control, not to mention the appearance of Riccati equations. Hence how to demonstrate indefinite $R(\cdot)$ in LQ problem of SVIEs is a tricky question to answer. The second difficulty lies in the seeking of proper sufficiency conditions of optimal controls. More precisely, if we follow the conventional ideas along this (see *e.g.* [8,25,27]), one must encounter with the solvability of coupled forward-backward SVIEs, see *e.g.* (4.9) next. We emphasize that their unusual and complicated structures make some well applicable tools and methods in FBSDEs setting (*e.g.* [11, 15]) hard to be proposed or adjusted here. In addition, to represent optimal controls one has to suppose $R^{-1}(\cdot)$ to be well-defined, see *e.g.* (4.8) later, not to mention the complicated expressions of optimal controls.

Instead of following the existing routine, the aim of this paper is devoted to providing new ideas on linear quadratic problem of both SDEs and SVIEs. As to SDEs case, using the main result of [22] we give three examples where optimal closed-loop controls do not exist. Moreover, for these examples the classical open-loop representations *via* FBSDEs also fail to work. This motivates us to give some improvements of available conclusions in this topic. More precisely, we cook up some new equivalent criteria for the existence of optimal controls, the procedures of which are partially inspired by [6, 13]. Keeping the basic ideas in mind, we then continue to treat the SVIEs case. Since the general study with respect to (1.1) can not get around anticipated stochastic integral (see the beginning of Sect. 4) and is substantially much more involved, we then turn to a systematical investigation with deterministic A_1 and $A_2 = 0$. More precisely, we start with some motivations by analyzing the problems in the related literature. Then we give detailed discussions from both the random and deterministic coefficients views, point out the comparisons and improvements on existing results. At last, in Section 4.2 we obtain a new pointwise necessary optimality condition, partial tricks of which rely on the arguments in Section 4.1.

At this moment we point out some novelties of this work. In the first place, we provide some new characterizations of optimal controls in linear quadratic problem for SDEs. They are potentially useful when neither closed-loop optimal controls nor classical FBSDEs representations of optimal controls work well. Moreover, $R(\cdot)$ is allowed to be indefinite by usual open-loop optimal controls but not limited to particular closed-loop optimal controls. In the second place, by applying the fundamental ideas in the Section 3 we then obtain new conditions for the existence of optimal controls in linear quadratic problem of SVIEs. Compared with the existing conclusions, ours here seem more straightforward and briefer in terms of representation of optimal controls and associated forward-backward systems. In addition, the state process X is replaced by $\mathcal{X}$, two pairs of (P_Q, Λ_Q), (P_G, Λ_G) are introduced, and the powerfulness of Itô formula and the idea of second-order adjoint equation return back. To our best these tricks appear for the first time in the literature. In the third place, when the coefficients become deterministic, instead of using the forward-backward stochastic equations, a new class of stochastic Fredholm−Volterra integral equations (SFVIEs, in short) is introduced here. From this new viewpoint, the requirement of $R^{-1}(\cdot)$ and the complexity of coupled FBSVIEs can be avoided. In addition, it is shown by one example that $R(\cdot)$ is allowed to be negative, which does not happen in deterministic VIEs case. Last but not the least important, using the introduced terms $\mathcal{X}$, (P_Q, Λ_Q), (P_G, Λ_G) and spike variation, we derive new necessary optimality conditions in linear quadratic problems of SVIEs. In contrast with the convex variation tricks in [21, 26, 27], we provide additional connections among the involved coefficients. Moreover, ours here can cover the corresponding results in SDEs case.

The rest of this paper is organized as follows. In Section 2, we give some preliminary results. In Section 3, after some illustrative counterexamples in the SDEs setting, we give new equivalent conditions for the existence of optimal controls, and show the advantages of our method by some following-up remarks and examples. In Section 4, by pointing out some essential difficulties in general framework, we choose to discuss linear quadratic problem of SVIEs when $A_2 = 0$ and A_1 is deterministic. Even so, there are still some unsolved problems in the related literature, the point of which prompts us to investigate both random and deterministic coefficients cases. Eventually we obtain new necessary conditions of optimal controls in our linear quadratic problems.

2. Preliminaries

Suppose $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a complete filtered probability space on which $W(\cdot)$ is scalar-valued Brownian motion, $\mathbb{F} := (\mathcal{F}_t)_{0\leq t\leq T}$, $\mathcal{F}_t := \sigma\{W(s),\ s \leq t\}$. For $H := \mathbb{R}^m,\ \mathbb{R}^n,\ \mathbb{S}^n$, etc. let us give some spaces which are needed in the sequel.

$$
L^2_{\mathcal{F}_T}(\Omega; H) := \left\{ X : \Omega \to H \;\middle|\; X \text{ is } \mathcal{F}_T\text{-measurable random variable s.t. } \mathbb{E}|X|^2 < \infty \right\},
$$

$$
L^2_{\mathcal{F}}(0,T; H) := \left\{ X : [0,T]\times\Omega \to H \;\middle|\; X(\cdot) \text{ is } \mathbb{F}\text{-adapted, measurable process s.t. } \mathbb{E}\int_0^T |X(s)|^2 \mathrm{d}s < \infty \right\},
$$

$$L^2(0,T;L^2_{\mathcal{F}}(0,T;H)) := \left\{ X : [0,T]^2 \times \Omega \to H \;\middle|\; X(t,\cdot) \text{ is } \mathbb{F}\text{-adapted, measurable} \right.$$
$$\left. \text{for almost } t \in [0,T] \text{ s.t. } \mathbb{E}\int_0^T \int_0^T |X(t,s)|^2 \mathrm{d}s\mathrm{d}t < \infty \right\},$$

$$L^\infty(0,T;H) := \left\{ X : [0,T] \to H \;\middle|\; X \text{ is deterministic function s.t. } \sup_{t\in[0,T]} |X(t)| < \infty \right\},$$

$$C_{\mathcal{F}}([0,T];L^2(\Omega;H)) := \left\{ X : [0,T] \to L^2(\Omega;H) \;\middle|\; X(\cdot) \text{ is } \mathbb{F}\text{-adapted, measurable, and continuous} \right.$$
$$\left. \text{in } L^2(\Omega;H) \text{ s.t. } \sup_{r\in[0,T]} \mathbb{E}|X(r)|^2 < \infty \right\}.$$

Similarly one can define $L^\infty(0,T;L^\infty_{\mathcal{F}}(0,T;H))$, $L^2_{\mathcal{F}}(\Omega;C([0,T];H))$, etc. In the following we will show some preliminary results on the linear quadratic control problem for SVIE (1.1). We introduce the assumptions as follows.

(H1) Suppose $\varphi(\cdot) \in C_{\mathcal{F}}([0,T];L^2(\Omega;\mathbb{R}^n))$, $Q(\cdot)$, $S(\cdot)$, $R(\cdot)$ and G are bounded and random, $A_1(\cdot,\cdot)$, $B_1(\cdot,\cdot)$, $A_2(\cdot,\cdot)$, $B_2(\cdot,\cdot) \in L^\infty(0,T;L^\infty_{\mathcal{F}}(0,T;H))$ with $H := \mathbb{R}^{n\times m}$, $\mathbb{R}^{n\times n}$. Moreover, there exists a modulus of continuity $\rho : [0,\infty) \to [0,\infty)$ (*i.e.*, $\rho(\cdot)$ is continuous and strictly increasing with $\rho(0) = 0$) such that

$$|A_i(t,s) - A_i(t',s)| + |B_i(t,s) - B_i(t',s)| \le \rho(|t-t'|), \qquad t,t',s \in [0,T].$$

Under (H1) we know that there exists a unique solution of (1.1) such that $X(\cdot) \in C_{\mathcal{F}}([0,T];L^2(\Omega;\mathbb{R}^n))$. For any $x(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, let us define a bounded operator $\mathcal{A}$ from $L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ to itself,

$$(\mathcal{A}x)(t) = \int_0^t A_1(t,s)x(s)\mathrm{d}s + \int_0^t A_2(t,s)x(s)\mathrm{d}W(s), \quad t \in [0,T]. \tag{2.1}$$

At this moment, one can show that $(I-\mathcal{A})^{-1}$ is bounded under (H1) (see [8,25]). Similarly one can define $\mathcal{B}$ with A_1, A_2 in (2.1) replaced by B_1, B_2. Therefore the state equation can be rewritten as,

$$X(\cdot) = \varphi(\cdot) + (\mathcal{A}X)(\cdot) + (\mathcal{B}u)(\cdot). \tag{2.2}$$

To treat the terminal term, we need to define bounded and linear operator Δ_T from $L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ to $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$, *i.e.*, for any $p(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$,

$$\Delta_T p = \int_0^T A_1(T,s)p(s)\mathrm{d}s + \int_0^T A_2(T,s)p(s)\mathrm{d}W(s). \tag{2.3}$$

We can also define Λ_T with A_1, A_2 in (2.3) replaced by B_1, B_2. Like (2.2) we have

$$X(T) = \varphi(T) + (\Delta_T X) + (\Lambda_T u). \tag{2.4}$$

In what follows, $\mathcal{A}^*$ is denoted as the adjoint operator of $\mathcal{A}$. Substituting (2.2) and (2.4) into the cost functional we obtain,

$$J(0,x;u) = \langle \Theta u + 2\Theta_1\varphi, u\rangle_2 + \langle \Theta_2\varphi,\varphi\rangle_2 + 2\left\langle (I-\mathcal{A})^{-1}\varphi, \Delta_T^* G\varphi(T)\right\rangle_2 + \langle G\varphi(T),\varphi(T)\rangle_1,$$

where $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$ are the inner product of $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$, $L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ and

$$
\begin{aligned}
\Theta &:= \big(\mathcal{B}^*(I-\mathcal{A}^*)^{-1}\mathcal{Q}'+\mathcal{S}'\big)(I-\mathcal{A})^{-1}\mathcal{B}+\mathcal{B}^*(I-\mathcal{A}^*)^{-1}\mathcal{S}'^*+\mathcal{R}',\\
\Theta_1\varphi &:= \big(\mathcal{B}^*(I-\mathcal{A}^*)^{-1}\mathcal{Q}'+\mathcal{S}'\big)(I-\mathcal{A})^{-1}\varphi+\mathcal{B}^*(I-\mathcal{A}^*)^{-1}\Delta_T^*G\varphi(T)+\Gamma_T^*G\varphi(T),\\
\Theta_2 &:= (I-\mathcal{A}^*)^{-1}\mathcal{Q}'(I-\mathcal{A})^{-1},\\
\mathcal{Q}' &:= \mathcal{Q}+\Delta_T^*G\Delta_T,\quad \mathcal{S}':=\mathcal{S}+\Lambda_T^*G\Delta_T,\quad \mathcal{R}':=\mathcal{R}+\Lambda_T^*G\Lambda_T,\\
\big[\mathcal{Q}X\big](\cdot) &:= Q(\cdot)X(\cdot),\quad \big[\mathcal{S}X\big](\cdot):=S(\cdot)X(\cdot),\quad \big[\mathcal{R}X\big](\cdot):=R(\cdot)X(\cdot),\quad \forall X(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}).
\end{aligned}
\tag{2.5}
$$

The following result gives one sufficient and necessary condition for the existence of optimal control *via* above operators (see also [8, 16]).

Lemma 2.1. *Let* (H1) *hold and* $\varphi(\cdot)\in C_{\mathcal{F}}([0,T];L^2(\Omega;\mathbb{R}^n))$ *be given. Then the linear quadratic problem admits an optimal control* $\widehat{u}(\cdot)$ *if and only if the following conditions are true,*

$$\text{(i) } \Theta\geq 0,\qquad \text{(ii) } \Theta_1\varphi\in\mathcal{R}(\Theta)$$

where $\mathcal{R}(\Theta)$ *is the range of* Θ. *In this case any optimal control* $\widehat{u}(\cdot)$ *is a solution of* $\Theta\widehat{u}+\Theta_1\varphi=0$.

In order to represent the conclusion more explicitly, let us introduce,

$$
\begin{cases}
Y(t)=Q(t)X(t)+S(t)^\top u(t)+A_1(T,t)^\top GX(T)+A_2(T,t)^\top\theta(t)\\
\qquad\qquad +\displaystyle\int_t^T\big[A_1(s,t)^\top Y(s)+A_2(s,t)^\top Z(s,t)\big]\mathrm{d}s-\int_t^T Z(t,s)\mathrm{d}W(s),\\
\chi(t)=\mathbb{E}_t\displaystyle\int_t^T\big[B_1(s,t)^\top Y(s)+B_2(s,t)^\top Z(s,t)\big]\mathrm{d}s,\\
GX(T)=\mathbb{E}\big[GX(T)\big]+\displaystyle\int_0^T\theta(s)\mathrm{d}W(s),\quad Y(t)=\mathbb{E}Y(t)+\int_0^t Z(t,s)\mathrm{d}W(s),
\end{cases}
\tag{2.6}
$$

where here and subsequently, $\mathbb{E}_tX:=\mathbb{E}^{\mathcal{F}_t}X$ with random variable X. We emphasize that the first equation in (2.6) is a linear backward stochastic Volterra integral equation (BSVIE, in short). As to the study on BSVIEs, readers are referred to [14, 20, 24, 26, 27], etc. The following result is from [8] or [25].

Lemma 2.2. *Let* (H1) *hold,* $\varphi(\cdot)\in C_{\mathcal{F}}([0,T];L^2(\Omega;\mathbb{R}))$. *Then the linear quadratic problem of SVIEs is solvable with optimal pair* $(\widehat{X}(\cdot),\widehat{u}(\cdot))$ *if and only if the following hold:*

(i) *given* $X_0(\cdot)$, χ_0 *in* (1.1), (2.6) *associated with* $\varphi(\cdot)=0$, $X_0(\cdot)$, *respectively, we have*

$$\mathbb{E}\int_0^T u(s)^\top\big[\chi_0(s)+S(s)X_0(s)+R(s)u(s)+B_1(T,s)^\top GX_0(T)+B_2(T,s)^\top\theta_0(s)\big]\mathrm{d}s\geq 0,\tag{2.7}$$

(ii) *for* $(\widehat{\chi}(\cdot),\widehat{\theta}(\cdot))$ *in* (2.6) *associated with* $(\widehat{X}(\cdot),\widehat{u}(\cdot))$, *we have*

$$\widehat{\chi}(t)+S(t)\widehat{X}(t)+R(t)\widehat{u}(t)+B_1(T,t)^\top\mathbb{E}_t\big[G\widehat{X}(T)\big]+B_2(T,t)^\top\widehat{\theta}(t)=0,\quad t\in[0,T].\ \text{a.e.}\tag{2.8}$$

To conclude this section, let us point out some preliminary knowledge on pseudo-inverse. For any $M\in\mathbb{R}^{m\times n}$, there exists a unique matrix $M^\dagger\in\mathbb{R}^{n\times m}$, called the (Moore-Penrose) *pseudo-inverse* of M, satisfying

$$MM^\dagger M=M,\quad M^\dagger MM^\dagger=M^\dagger,\quad \big[MM^\dagger\big]^\top=MM^\dagger,\quad \big[M^\dagger M\big]^\top=M^\dagger M.$$

In addition, if $M\in\mathbb{S}^m$, then

$$M^\dagger=\big[M^\dagger\big]^\top,\quad MM^\dagger=M^\dagger M,\quad M\geq 0\Longleftrightarrow M^\dagger\geq 0.$$

3. Linear quadratic problem for SDEs

In this section, we give some discussions on linear quadratic problem of SDEs with deterministic coefficients. It seems that the results here are new in the literature, and some of them are sharper than existing counterparts. More importantly, the ideas developed here inspire us to study the SVIEs case in the next section.

If φ, A_i, B_i do not depend on t, then (1.1) becomes a SDE of

$$\mathrm{d}X(t)=\big[A_1(t)X(t)+B_1(t)u(t)\big]\mathrm{d}t+\big[A_2(t)X(t)+B_2(t)u(t)\big]\mathrm{d}W(t),\quad t\in[0,T], \tag{3.1}$$

with $X(0)=x$. The cost functional is defined as (1.2). For later usefulness, let us introduce the following equation for $(\mathcal{P}(\cdot),\mathcal{Q}(\cdot))$,

$$\begin{cases}\mathrm{d}\mathcal{P}(t)=-\big[A_1(t)^\top\mathcal{P}(t)+\mathcal{P}(t)A_1(t)+A_2(t)^\top\mathcal{P}(t)A_2(t)+Q(t)\big]\mathrm{d}t,\quad t\in[0,T],\\ \mathcal{P}(T)=G,\end{cases} \tag{3.2}$$

and two associated notations,

$$\mathcal{S}:=S+B_2^\top\mathcal{P}A_2+B_1^\top\mathcal{P},\quad \mathcal{R}:=R+B_2^\top\mathcal{P}B_2.\ \ a.e. \tag{3.3}$$

3.1. Some motivations

Before stating the main result, we give a revisit to the traditional study and illustrate the starting points/motivations for later arguments. At first, we discuss closed-loop optimal controls *via* stochastic Riccati equations under this framework. Before it, we recall the definitions of closed-loop optimal strategies and closed-loop optimal controls when the coefficients are deterministic (see [22]).

Definition 3.1. Suppose all the coefficients in (3.1) and (1.2) are deterministic. A pair $(\bar\Theta(\cdot),\bar v(\cdot))\in L^2(0,T;\mathbb{R}^{m\times n})\times L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ is called a closed-loop optimal strategy on $[0,T]$ if

$$J(0,x;\bar\Theta(\cdot)\bar X(\cdot)+\bar v(\cdot))\le J(0,x;u(\cdot)),\qquad x\in\mathbb{R}^n,\quad u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m).$$

In this case, the control variable

$$\bar u(\cdot):=\bar\Theta(\cdot)\bar X(\cdot)+\bar v(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R})$$

is called a closed-loop optimal control associated with initial pair $(0,x)$.

For $(\bar\Theta(\cdot),\bar v(\cdot))$, that are independent of initial state $x\in\mathbb{R}^n$, the authors in [22] obtained the following kind of equivalent condition,

Proposition 3.1. *Suppose* (H1) *holds true with deterministic coefficients. Then the linear quadratic problem associated with* (3.1) *and* (1.2) *admits a closed-loop optimal strategy* $(\bar\Theta,\bar v)$*/closed-loop optimal control* $\bar u(\cdot)$ *in the sense of Definition* 3.1 *if and only if the following Riccati equation admits a solution* $P(\cdot)\in C([0,T];\mathbb{S}^n)$,

$$\begin{cases}\dot P(t)+P(t)A_1(t)+A_1(t)^\top P(t)+A_2(t)^\top P(t)A_2(t)+Q(t)-\bar S(t)^\top\bar R(t)^\dagger\bar S(t)=0,\ a.e.\ \ t\in[0,T],\\ \mathcal{R}\big(\bar S(t)\big)\subseteq\mathcal{R}\big(\bar R(t)\big),\quad \bar R(t)\ge0,\quad a.e.\ \ t\in[0,T],\quad P(T)=G,\end{cases} \tag{3.4}$$

such that $\bar R^\dagger(\cdot)\bar S(\cdot)\in L^2(0,T;\mathbb{R}^m)$. *Here* $\mathcal{R}(\bar R(t))$ *represents the range of* $\bar R(t)$ *with* $t\in[0,T]$, *and*

$$\bar R(\cdot):=R(\cdot)+B_2(\cdot)^\top P(\cdot)B_2(\cdot),\quad \bar S(\cdot):=B_1(\cdot)^\top P(\cdot)+B_2(\cdot)^\top P(\cdot)A_2(\cdot)+S(\cdot). \tag{3.5}$$

In this case, any closed-loop optimal strategy $(\bar\Theta(\cdot),\bar v(\cdot))$ *can be represented as,*

$$\bar\Theta(t)=-\bar R(t)^\dagger\bar S(t)+\Big[I-\bar R(t)^\dagger\bar R(t)\Big]\theta(t),\quad \bar v(t)=\Big[I-\bar R(t)^\dagger\bar R(t)\Big]v(t),$$

for some $\theta(\cdot)\in L^2(0,T;\mathbb{R}^{m\times n})$ *and* $v(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$.

By means of Proposition 3.1 let us look at a few examples.

Example 3.1. Suppose

$$m = n = 1, \quad R(\cdot) = B_2(\cdot) = Q(\cdot) = S(\cdot) = 0, \quad A_1(\cdot) = B_1(\cdot) = A_2(\cdot) = G = 1.$$

Hence according to (3.4) and (3.5), $\bar{S}(t) = P(t)$ and $\mathcal{P}(t) = \mathrm{e}^{-3t}$ with $t \in [0, T]$. Moreover,

$$X(t) = \mathrm{e}^{\frac{t}{2}+W(t)}x + \int_0^t \mathrm{e}^{\frac{t-s}{2}+W(t)-W(s)}u(s)\mathrm{d}s, \qquad t \in [0, T].$$

In this case, it is easy to see that

$$\widehat{u}(s) = -\mathrm{e}^{\frac{s}{2}+W(s)}\frac{x}{T}, \qquad s \in [0, T],$$

is an optimal control such that $\widehat{X}(T) = 0$. If there exists an closed-loop optimal strategy/control, by Proposition 3.1 there exists $P(\cdot)$ satisfying

$$\mathcal{R}\big(\bar{S}(t)\big) \subseteq \mathcal{R}\big(\bar{R}(t)\big), \quad \bar{R}(t) \geq 0, \qquad t \in [0, T]. \ \ a.e. \tag{3.6}$$

For such $P(\cdot)$, by comparison theorem and the second condition in (3.6),

$$0 = R(\cdot) + B_2^2(\cdot)\mathcal{P}(\cdot) \geq \bar{R}(\cdot) = R(\cdot) + B_2^2(\cdot)P(\cdot) \geq 0$$

which implies that $\bar{R}(t) = 0$ and $\bar{S}(t) = \mathcal{P}(t) = P(t) = \mathrm{e}^{-3t} > 0$ with $t \in [0, T]$. However, this contradicts with the first condition in (3.6). Consequently, according to Proposition 3.1 the closed-loop optimal control does not exist. □

Above example shows that closed-loop optimal control may not exist when $B_2(\cdot) = 0$. As to $B_2(\cdot) \neq 0$, let us look at,

Example 3.2. Suppose $m = n = 2$, $R(\cdot) = Q(\cdot) = S(\cdot) = 0$, and for $H_1(\cdot) = A_1(\cdot),\ B_1(\cdot),\ A_2(\cdot)$,

$$H_1(\cdot) = \begin{bmatrix} 1, & 0 \\ 0, & 1 \end{bmatrix}, \quad B_2(\cdot) = \begin{bmatrix} 1, & 0 \\ 0, & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0, & 0 \\ 0, & 1 \end{bmatrix}.$$

In this case, for $X(\cdot) := (X_1(\cdot), X_2(\cdot))^\top$, $u(\cdot) := (u_1(\cdot), u_2(\cdot))^\top$, the linear quadratic problem is to find suitable $\widehat{u}(\cdot)$ such that $J(u(\cdot)) = \mathbb{E}|X_2(T)|^2$ is minimized where,

$$\mathrm{d}X_2(t) = \big[X_2(t) + u_2(t)\big]\mathrm{d}t + X_2(t)\mathrm{d}W(t), \qquad X_2(0) = x_2.$$

Similar as Example 3.1, it is easy to see that

$$\widehat{u}(t) = \Big(\widehat{u}_1(t), -\mathrm{e}^{\frac{t}{2}+W(t)}\frac{x_2}{T}\Big)^T, \qquad t \in [0, T],$$

is an optimal control. At this moment we claim that there is no closed-loop optimal control. Actually, if it is so, by Proposition 3.1 there exists $P(\cdot)$ satisfying

$$\mathcal{R}\big(\bar{S}(t)\big) \subseteq \mathcal{R}\big(\bar{R}(t)\big) = \mathcal{R}\big(R(t) + B_2(t)^\top\mathcal{P}(t)B_2(t)\big), \qquad t \in [0, T]. \ \ a.e. \tag{3.7}$$

For such $P(\cdot)$, by comparison theorem and some simple calculations,

$$P(\cdot) \leq \mathcal{P}(\cdot), \qquad 0 = R(\cdot) + B_2(\cdot)^\top\mathcal{P}(\cdot)B_2(\cdot) \geq \bar{R}(\cdot) \geq 0,$$

which implies that $\bar{R}(\cdot) = 0$ and $\mathcal{P}(\cdot) = P(\cdot)$. Moreover,

$$\bar{S}_{11}(t) = \bar{S}_{12}(t) = \bar{S}_{21}(t) = 0, \quad \bar{S}_{22}(t) = \mathrm{e}^{-3t}, \qquad \bar{S}(t) = \big[\bar{S}_{i,j}(t)\big]_{1\leq i,\ j\leq 2}.$$

Therefore the obtained results of $\bar{R}(\cdot)$ and $\bar{S}(\cdot)$ contradict with (3.7). □

Note that in above two examples, a common feature is $R(\cdot)+B_2(\cdot)^\top\mathcal{P}(\cdot)B_2(\cdot)=0$. The next example shows that even though $R(\cdot)+B_2(\cdot)^\top\mathcal{P}(\cdot)B_2(\cdot)>\delta>0$, the feedback control still may not exist.

Example 3.3. Suppose $m=n=1$, and

$$\begin{aligned}&A_1(\cdot)=A_2(\cdot)=0,\quad B_1(\cdot)=2,\quad B_2(\cdot)=1,\quad R(\cdot)=0,\\&Q(s)=2s-6s^2,\qquad S(s)=2s^2-4s^3,\quad G=2T^3.\end{aligned}$$

If the linear quadratic admits a closed-loop optimal control, then according to Proposition 3.1 the following equation admits a solution $P(\cdot)$

$$\dot{P}(s)+2s-6s^2-\big[2P(s)+2s^2-4s^3\big]^2P^\dagger(s)=0,\quad P(T)=2T^3. \tag{3.8}$$

such that the conditions in (3.4) are satisfied. Now let us look at the solvability of (3.8). To this end, since $P(T)=2T^3$, let us denote by T_0 the biggest point satisfying $P(T_0)=0$. It is easy to see that,

$$P(s)=2s^3,\quad \forall s\in[T_1,T],\quad \forall T_1\in(T_0,T].$$

By the arbitrariness of T_1 and the continuity of $P(\cdot)$, one has $T_0=0$ and $P(s)=2s^3$ with $s\in[0,T]$. In this case,

$$\bar{R}(s)=P(s)=2s^3,\quad \bar{S}(s)=2s^2\quad\Rightarrow\mathcal{R}(\bar{S}(s))\subseteq\mathcal{R}(\bar{R}(s)),\quad s\in[0,T].$$

However, $\bar{R}^\dagger(s)\bar{S}(s)=\frac{1}{s}$, with $s\in(0,T]$ which shows that $\bar{R}^\dagger(\cdot)\bar{S}(\cdot)$ does not belong to $L^2(0,T;\mathbb{R})$. By Proposition 3.1 the closed-loop optimal control does not exist. At last we have

$$R(t)+B_2^2(t)\mathcal{P}(t)=T^2+2t^3-t^2\geq\delta>0,\qquad t\in[0,T].$$

As to Example 3.3 it is natural to ask whether (open-loop) optimal controls make sense. Before checking it, let us derive a related result by Lemma 2.2. For SDE (3.1), condition (2.8) becomes,

$$B_1(t)^\top\mathbb{E}_t\int_t^T\widehat{Y}(s)\mathrm{d}s+B_2(t)^\top\int_t^T\widehat{Z}(s,t)\mathrm{d}s+S(t)\widehat{X}(t)+R(t)\widehat{u}(t)+B_1(t)^\top\mathbb{E}_t\big[G\widehat{X}(T)\big]+B_2(t)^\top\widehat{\theta}(t)=0, \tag{3.9}$$

where

$$\begin{aligned}\widehat{Y}(t)&=Q(t)\widehat{X}(t)+S(t)^\top\widehat{u}(t)+A_1(t)^\top G\widehat{X}(T)+\int_t^TA_1(t)^\top\widehat{Y}(s)\mathrm{d}s\\&\quad+A_2(t)^\top\widehat{\theta}(t)+A_2(t)^\top\int_t^T\widehat{Z}(s,t)\mathrm{d}s-\int_t^T\widehat{Z}(t,s)\mathrm{d}W(s),\quad t\in[0,T].\end{aligned}$$

By making the following convention,

$$\widehat{\mathcal{Y}}(t):=\mathbb{E}_t\left[G\widehat{X}(T)+\int_t^T\widehat{Y}(s)\mathrm{d}s\right],\quad \widehat{\mathcal{Z}}(t):=\widehat{\theta}(t)+\int_t^T\widehat{Z}(s,t)\mathrm{d}s,\quad t\in[0,T], \tag{3.10}$$

one can rewrite (3.9) as,

$$B_1(t)^\top\widehat{\mathcal{Y}}(t)+S(t)\widehat{X}(t)+R(t)\widehat{u}(t)+B_2(t)^\top\widehat{\mathcal{Z}}(t)=0, \tag{3.11}$$

where

$$\mathrm{d}\widehat{\mathcal{Y}}(t)=-\big[Q(t)\widehat{X}(t)+S(t)^\top\widehat{u}(t)+A_1(t)^\top\widehat{\mathcal{Y}}(s)+A_2(t)^\top\widehat{\mathcal{Z}}(t)\big]\mathrm{d}t-\widehat{\mathcal{Z}}(t)\mathrm{d}W(t),\quad \widehat{\mathcal{Y}}(T)=G\widehat{X}(T). \tag{3.12}$$

Similarly the condition $\Theta \geq 0$ or (2.7) becomes,

$$\mathbb{E}\int_0^T \langle\, B_1(t)^\top \mathcal{Y}_0(t) + S(t)X_0(t) + R(t)u(t) + B_2(t)^\top \mathcal{Z}_0(t), u(t)\,\rangle\,\mathrm{d}t \geq 0, \tag{3.13}$$

where $(\mathcal{Y}_0(\cdot), \mathcal{Z}_0(\cdot))$ satisfies (3.12) with $\widehat{u}(\cdot)$ replaced by $u(\cdot)$, and $x = 0$. To sum up,

Lemma 3.1. *Given (3.1) and (1.2), suppose the coefficients satisfy* (H1)*. Then the linear quadratic problem of SDEs is solvable with optimal pair $(\widehat{X}(\cdot), \widehat{u}(\cdot))$ if and only if (3.11) and (3.13) hold true.*

Remark 3.1. Note that Lemma 3.1 also appears in the literature (see *e.g.* [28]). Here we point it out for the reason of self-completeness. Now let us return back to Example 3.1, 3.2, 3.3. Unfortunately, it seems that one can not use Lemma 3.1 to represent optimal control $\widehat{u}(\cdot)$ and obtain corresponding coupled FBSDEs due to $R(\cdot) = 0$. □

To conclude this part, let us make some points here. As a particular class of control variables, closed-loop/feedback optimal controls are quite useful in both theoretical and practical aspects. Nevertheless, even in some simple cases (like Ex. 3.1, 3.2) this kind of ideal controls do not make sense any more. In other words, for these linear quadratic problems, we have to investigate the usual optimal controls. Unfortunately again, the existing criteria for optimal controls (*e.g.* Lem. 3.1) may fail to represent optimal controls. Therefore furthermore developments in this topic are needed.

3.2. Main results

Motivated by above arguments, we show some new results in the linear quadratic problem of SDEs. The ideas here not only enable us to improve the related present conclusions, but also can be extended and adapted into the SVIEs setting next. As a preparation, let us look at the following type of Fubini theorem between Lebesgue integral and Itô integral (see also Thm. 4.A. in p. 224 of [3]).

Lemma 3.2. *Suppose $f(\cdot,\cdot)$ is a measurable process on $[0,T]^2$ s.t. $f(t,s)$ is $\mathcal{F}_s$-measurable for $t \in [0,T]$ and*

$$\mathbb{E}\int_0^T \int_0^t |f(t,s)|^2 \mathrm{d}s\mathrm{d}t < \infty.$$

Then we have

$$\int_0^T \int_0^t f(t,s)\mathrm{d}W(s)\mathrm{d}t = \int_0^T \int_s^T f(t,s)\mathrm{d}t\mathrm{d}W(s).$$

Given the coefficients in (H1), $\mathcal{S}(\cdot)$, $\mathcal{R}(\cdot)$ in (3.3), and control variable $v(\cdot) \in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ $(p > 2)$, let us make the convention that,

$$\begin{cases} \mathrm{d}\Psi(t) = A_1(t)\Psi(t)\mathrm{d}t + A_2(t)\Psi(t)\mathrm{d}W(t), \qquad t \in [0,T], \quad \Psi(0) = I, \\ \mathcal{L}_1(t) := \mathcal{S}(t)\Psi(t); \quad \mathcal{L}_2(t) := \Psi^{-1}(t)\big[B_1(t) - A_2(t)B_2(t)\big]; \quad \mathcal{L}_3(t) := \Psi^{-1}(t)B_2(t), \quad t \in [0,T], \\ \mathcal{L}_4(t) := \mathcal{L}_1(t)x + \mathcal{L}_1(t)\displaystyle\int_0^t \mathcal{L}_2(s)\big[I - \mathcal{R}^\dagger(s)\mathcal{R}(s)\big]v(s)\mathrm{d}s + \mathcal{L}_1(t)\int_0^t \mathcal{L}_3(s)\big[I - \mathcal{R}^\dagger(s)\mathcal{R}(s)\big]v(s)\mathrm{d}W(s) \\ \qquad\qquad + \mathcal{L}_2(t)^\top \mathbb{E}_t \displaystyle\int_t^T \mathcal{L}_1(s)^\top \big[I - \mathcal{R}^\dagger(s)\mathcal{R}(s)\big]v(s)\mathrm{d}s + \mathcal{L}_3(t)^\top \int_t^T K_{\mathcal{L}_1,v}(s,t)\mathrm{d}s, \quad t \in [0,T], \end{cases} \tag{3.14}$$

and introduce the following equation,

$$\begin{aligned} \widehat{\lambda}(t) = &-\mathcal{L}_4(t) - \mathcal{L}_1(t)\int_0^t \mathcal{L}_2(s)\mathcal{R}^\dagger(s)\widehat{\lambda}(s)\mathrm{d}s - \mathcal{L}_1(t)\int_0^t \mathcal{L}_3(s)\mathcal{R}^\dagger(s)\widehat{\lambda}(s)\mathrm{d}W(s) \\ &-\mathcal{L}_2(t)^\top \mathbb{E}_t \int_t^T \mathcal{L}_1(s)^\top \mathcal{R}^\dagger(s)\widehat{\lambda}(s)\mathrm{d}s - \mathcal{L}_3(t)^\top \int_t^T K_{\mathcal{L}_1,\mathcal{R}\widehat{\lambda}}(s,t)\mathrm{d}s, \quad t \in [0,T], \end{aligned} \tag{3.15}$$

where $K_{\mathcal{L}_1,v}$, $K_{\mathcal{L}_1,\mathcal{R}\widehat{\lambda}}$ are determined by martingale representation arguments as follows,

$$\mathcal{L}_1(t)^\top g(t) = \mathbb{E}\big[\mathcal{L}_1(t)^\top g(t)\big] + \int_0^t K_{\mathcal{L}_1,g}(t,r)\mathrm{d}W(r), \quad t\in[0,T], \quad g(\cdot) := \mathcal{R}^\dagger(\cdot)\widehat{\lambda}(\cdot),\ v(\cdot). \tag{3.16}$$

For later convenience, for any admissible control $u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ we define $K_{\mathcal{L}_1,u}$ as follows,

$$\mathcal{L}_1(t)^\top u(t) = \mathbb{E}\big[\mathcal{L}_1(t)^\top u(t)\big] + \int_0^t K_{\mathcal{L}_1,u}(t,r)\mathrm{d}W(r), \quad t\in[0,T], \tag{3.17}$$

We discuss the case of $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ with $p>2$.

Theorem 3.1. *Suppose* (H1) *holds true with deterministic coefficients. Then the linear quadratic problem is solvable with optimal control $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ $(p>2)$ if and only if the following hold,*

(i) (3.15) *admits a solution $\widehat{\lambda}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$*
(ii) *for any $u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$, we have*

$$\mathbb{E}\int_0^T u(t)^\top\left[\mathcal{R}(t)u(t) + 2\mathcal{L}_2(t)^\top\int_t^T \mathcal{L}_1(s)^\top u(s)\mathrm{d}s + 2\mathcal{L}_3(t)^\top\int_t^T K_{\mathcal{L}_1,u}(s,t)\mathrm{d}s\right]\mathrm{d}t \geq 0, \tag{3.18}$$

where $K_{\mathcal{L}_1,u}(\cdot,\cdot)$ is defined in (3.17)*. In this case,*

$$\widehat{u}(t) = \mathcal{R}^\dagger(t)\widehat{\lambda}(t) + \big[I - \mathcal{R}^\dagger(t)\mathcal{R}(t)\big]v(t), \quad t\in[0,T], \quad \text{a.e.} \tag{3.19}$$

for some $v(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$.

Proof. In the following, we split the proof into three parts.

Step 1. By Lemma 2.1 and 3.1, it is easy to see that the linear quadratic problem is solvable with optimal control $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ if and only if (3.11), (3.13) are satisfied with $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ and any $u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$, respectively.

Step 2. In this step by exploring the potential relations between $(\widehat{\mathcal{Y}}(\cdot),\widehat{\mathcal{Z}}(\cdot))$ $((\mathcal{Y}_0(\cdot),\mathcal{Z}_0(\cdot))$ resp.) and $\widehat{u}(\cdot)$ ($u(\cdot)$ resp.), we obtain new equivalent versions of (3.11), (3.13).

Given $\mathcal{P}(\cdot)$ in (3.2), $\mathcal{S}(\cdot)$ in (3.3), we know that $\mathcal{S}(\cdot)$ is bounded. As a result, $\big[\mathcal{S}(\cdot)^\top u(\cdot)\big]\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ with $u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$. By the standard BSDE theory, the following equation admits a pair of processes $(\mathcal{M}(\cdot),\mathcal{N}(\cdot))\in L^2_{\mathcal{F}}(\Omega;C([0,T];\mathbb{R}^n))\times L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$,

$$\begin{cases} \mathrm{d}\mathcal{M}(t) = -\big[A_1(t)^\top\mathcal{M}(t) + A_2(t)^\top\mathcal{N}(t) + \mathcal{S}(t)^\top u(t)\big]\mathrm{d}t + \mathcal{N}(t)\mathrm{d}W(t), \quad t\in[0,T], \\ \mathcal{M}(T) = 0. \end{cases} \tag{3.20}$$

For later usefulness we point out the following fact,

$$\mathbb{E}\int_0^T u(t)^\top\big[B_1(t)^\top\mathcal{M}(t) + B_2(t)^\top\mathcal{N}(t)\big]\mathrm{d}t = \mathbb{E}\int_0^T u(t)^\top\mathcal{S}(t)X_0(t)\mathrm{d}t, \tag{3.21}$$

where $X_0(\cdot)$ satisfies (3.1) with $x=0$. Actually, using Itô formula to $\langle X_0(\cdot),\mathcal{M}(\cdot)\rangle$ on $[0,T]$, we can obtain that,

$$\begin{aligned} 0 &= \int_0^T \big[\langle B_1(t)u(t),\mathcal{M}(t)\rangle + \langle B_2(t)u(t),\mathcal{N}(t)\rangle - \langle X_0(t),\mathcal{S}(t)^\top u(t)\rangle\big]\mathrm{d}t \\ &\quad + \int_0^T \big[\langle A_2(t)X_0(t) + B_2(t)u(t),\mathcal{M}(t)\rangle + \langle X_0(t),\mathcal{N}(t)\rangle\big]\mathrm{d}W(t). \end{aligned} \tag{3.22}$$

Since

$$\mathbb{E}\left[\int_0^T \left[\langle A_2(t)X_0(t)+B_2(t)u(t),\mathcal{M}(t)\rangle+\langle X_0(t),\mathcal{N}(t)\rangle\right]^2 \mathrm{d}s\right]^{\frac{1}{2}}<\infty,$$

by taking expectation on both sides of (3.22) one then obtain (3.21). At this moment, we want to obtain one equivalent condition of (3.11). To this end, given $\widehat{u}(\cdot)$ and $\widehat{X}(\cdot)$, we use Itô formula to $\mathcal{P}(\cdot)\widehat{X}(\cdot)$ on $[0,T]$,

$$\mathrm{d}\big[\mathcal{P}\widehat{X}\big]=-\Big[\big[A_1^\top\mathcal{P}+A_2^\top\mathcal{P}A_2+A_2^\top\mathcal{Q}+Q\big]\widehat{X}+\big[\mathcal{P}B_1+\mathcal{Q}B_2\big]\widehat{u}\Big]\mathrm{d}t+\big[\mathcal{Q}\widehat{X}+\mathcal{P}A_2\widehat{X}+\mathcal{P}B_2\widehat{u}\big]\mathrm{d}W(t).$$

By denoting $(\widehat{\mathcal{M}}(\cdot),\widehat{\mathcal{N}}(\cdot))$ in (3.20) associated with $\widehat{u}(\cdot)$,

$$\begin{aligned}\mathrm{d}\big[\mathcal{P}(t)\widehat{X}(t)+\widehat{\mathcal{M}}(t)\big]=&-\Big[A_1(t)^\top\big[\mathcal{P}(t)\widehat{X}(t)+\widehat{\mathcal{M}}(t)\big]+A_2(t)^\top\big[\mathcal{P}(t)A_2(t)\widehat{X}(t)+\mathcal{Q}(t)\widehat{X}(t)\\&+\mathcal{P}(t)B_2(t)\widehat{u}(t)+\widehat{\mathcal{N}}(t)\big]+Q(t)\widehat{X}(t)+S(t)^\top\widehat{u}(t)\Big]\mathrm{d}t\\&+\big[\mathcal{P}(t)A_2(t)\widehat{X}(t)+\mathcal{Q}(t)\widehat{X}(t)+\mathcal{P}(t)B_2(t)\widehat{u}(t)+\widehat{\mathcal{N}}(t)\big]\mathrm{d}W(t).\end{aligned}$$

The uniqueness of BSDE (3.12) then indicates that,

$$\begin{aligned}&\widehat{\mathcal{Y}}(t)=\mathcal{P}(t)\widehat{X}(t)+\widehat{\mathcal{M}}(t),\qquad \forall t\in[0,T],\quad \text{a.s.}\\&\widehat{\mathcal{Z}}(t)=\big[\mathcal{Q}(t)+\mathcal{P}(t)A_2(t)\big]\widehat{X}(t)+\mathcal{P}(t)B_2(t)\widehat{u}(t)+\widehat{\mathcal{N}}(t),\qquad \text{a.s.}\quad t\in[0,T],\quad \text{a.e.}\end{aligned}\tag{3.23}$$

Therefore given $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$, condition (3.11) is equivalent with,

$$\mathcal{R}(t)\widehat{u}(t)+\mathcal{S}(t)\widehat{X}(t)+B_1(t)^\top\widehat{\mathcal{M}}(t)+B_2(t)^\top\widehat{\mathcal{N}}(t)=0,\quad t\in[0,T].\quad \text{a.e.}\tag{3.24}$$

Following above arguments, for any $u(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$, condition (3.13) can be rewritten as,

$$\mathbb{E}\int_0^T u(t)^\top\big[B_1(t)^\top\mathcal{M}(t)+\mathcal{S}(t)X_0(t)+\mathcal{R}(t)u(t)+B_2(t)^\top\mathcal{N}(t)\big]\mathrm{d}t\geq 0.\tag{3.25}$$

Thanks to (3.21), one can furthermore simplify (3.25) as,

$$\mathbb{E}\int_0^T \big[u(t)^\top\mathcal{R}(t)u(t)+2u(t)^\top\mathcal{S}(t)X_0(t)\big]\,\mathrm{d}t\geq 0.\tag{3.26}$$

Step 3. In this step by studying the relation between $(\widehat{\mathcal{M}}(\cdot),\widehat{\mathcal{N}}(\cdot))$ and $\widehat{u}(\cdot)$, let us continue to establish the equivalent link between (3.24), (3.26) and (3.15), (3.18) in the previous.

To this end, recalling $\Psi(\cdot)$ in (3.14), we use Itô formula to $\langle\Psi(\cdot),\mathcal{M}(\cdot)\rangle$ on $[0,T]$,

$$\Psi(t)^\top\mathcal{M}(t)=\int_t^T\Psi(s)^\top\mathcal{S}(s)^\top u(s)\mathrm{d}s-\int_t^T\Psi(s)^\top\big[A_2(s)^\top\mathcal{M}(s)+\mathcal{N}(s)\big]\mathrm{d}W(s),\quad t\in[0,T].\tag{3.27}$$

Given $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ with $p>2$, since $\mathcal{S}(\cdot)$ is bounded, for $\mathcal{L}_1(\cdot)$ in (3.14) we have

$$\mathbb{E}|\mathcal{L}_1(s)^\top\widehat{u}(s)|^2=\mathbb{E}|\Psi(s)^\top\mathcal{S}(s)^\top\widehat{u}(s)|^2<\infty,\qquad s\in[0,T],\quad \text{a.e.}\tag{3.28}$$

By martingale representation theorem there exists measurable process $K_{\mathcal{L}_1,\widehat{u}}(t,\cdot)\in L^2_{\mathcal{F}}(0,t;\mathbb{R}^n)$ such that

$$\mathcal{L}_1(t)^\top\widehat{u}(t)=\mathbb{E}\big[\mathcal{L}_1(t)^\top\widehat{u}(t)\big]+\int_0^t K_{\mathcal{L}_1,\widehat{u}}(t,r)\mathrm{d}W(r),\quad t\in[0,T],\quad \text{a.e.}$$

As to $K_{\mathcal{L}_1,\widehat{u}}(\cdot,\cdot)$, using Lemma 3.2 to BSDE (3.27) associated with $\widehat{u}(\cdot)$, we can obtain,

$$\int_s^T K_{\mathcal{L}_1,\widehat{u}}(r,s)\mathrm{d}r = \varPsi(s)^\top\big[A_2(s)^\top\widehat{\mathcal{M}}(s)+\widehat{\mathcal{N}}(s)\big], \qquad s\in[0,T].\ \text{ a.e.} \tag{3.29}$$

Therefore it then follows from (3.27), (3.29) that,

$$\begin{aligned}
\widehat{\mathcal{M}}(t) &= \big(\varPsi(t)^\top\big)^{-1}\mathbb{E}_t\int_t^T \varPsi(s)^\top\mathcal{S}(s)^\top\widehat{u}(s)\mathrm{d}s, \quad \forall t\in[0,T], \ \text{ a.s.}\\
\widehat{\mathcal{N}}(t) &= \big(\varPsi(t)^\top\big)^{-1}\int_t^T K_{\mathcal{L}_1,\widehat{u}}(s,t)\mathrm{d}s - A_2(t)^\top\big(\varPsi(t)^\top\big)^{-1}\mathbb{E}_t\int_t^T \varPsi(s)^\top\mathcal{S}(s)^\top\widehat{u}(s)\mathrm{d}s.\ \text{ a.e.}
\end{aligned} \tag{3.30}$$

Recalling $\varPsi(\cdot)$, $\mathcal{L}_i(\cdot)$ in (3.14), one can rewrite (3.1) as,

$$X(t) = \varPsi(t)x + \varPsi(t)\int_0^t \mathcal{L}_2(s)u(s)\mathrm{d}s + \varPsi(t)\int_0^t \mathcal{L}_3(s)u(s)\mathrm{d}W(s). \tag{3.31}$$

Plugging (3.30), (3.31) into (3.24), one then rewrite (3.24) as,

$$\begin{aligned}
&\mathcal{R}(t)\widehat{u}(t) + \mathcal{L}_1(t)\int_0^t \mathcal{L}_2(s)\widehat{u}(s)\mathrm{d}s + \mathcal{L}_1(t)\int_0^t \mathcal{L}_3(s)\widehat{u}(s)\mathrm{d}W(s) + \mathcal{L}_1(t)x\\
&\quad + \mathcal{L}_2(t)^\top\mathbb{E}_t\int_t^T \mathcal{L}_1(s)^\top\widehat{u}(s)\mathrm{d}s + \mathcal{L}_3(t)^\top\int_t^T K_{\mathcal{L}_1,\widehat{u}}(s,t)\mathrm{d}s = 0.
\end{aligned} \tag{3.32}$$

Let $\widehat{\lambda}(\cdot) := \mathcal{R}(\cdot)\widehat{u}(\cdot)$, one then obtain the representation (3.19) and equation (3.15). From (3.31),

$$\begin{aligned}
\mathbb{E}\int_0^T u(t)^\top\mathcal{S}(t)X_0(t)\mathrm{d}t &= \mathbb{E}\int_0^T u(t)^\top\mathcal{L}_1(t)\int_0^t \mathcal{L}_2(s)u(s)\mathrm{d}s\mathrm{d}t + \mathbb{E}\int_0^T u(t)^\top\mathcal{L}_1(t)\int_0^t \mathcal{L}_3(s)u(s)\mathrm{d}W(s)\mathrm{d}t\\
&= \mathbb{E}\int_0^T\int_s^T u(t)^\top\mathcal{L}_1(t)\mathrm{d}t\mathcal{L}_2(s)u(s)\mathrm{d}s + \mathbb{E}\int_0^T\int_s^T K_{\mathcal{L}_1,u}(t,s)^\top\mathrm{d}t\mathcal{L}_3(s)u(s)\mathrm{d}s.
\end{aligned}$$

Hence condition (3.26) is equivalent with (3.18). □

In order to obtain result (3.30) *via* Lemma 3.2, we need some square integrable condition in (3.28). However, due to the lack of boundedness of $\varPsi(\cdot)$, in Step 3 we have to assume that $\widehat{u}(\cdot)\in L^p_{\mathcal{F}}(0,T;\mathbb{R}^m)$ with $p>2$. Note that such situation can be improved if $A_2(\cdot)=0$.

Theorem 3.2. *Suppose* (H1) *is imposed with deterministic coefficients and* $A_2(\cdot)=0$. *Then the linear quadratic problem is solvable with optimal* $\widehat{u}(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ *if and only if the following hold,*

(i) (3.15) *admits a solution* $\widehat{\lambda}(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$,
(ii) (3.18) *is true for any* $u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$.

The proof is almost the same as above and we omit it here.

Remark 3.2. We emphasize that Theorem 3.1, Theorem 3.2 also work when the coefficients in (3.1), (1.2) are random. Actually, in this case we require $(\mathcal{P},\mathcal{Q})$ to be bounded where

$$\begin{cases}
\mathrm{d}\mathcal{P}(t) = -\big[A_1(t)^\top\mathcal{P}(t)+\mathcal{P}(t)A_1(t)+A_2(t)^\top\mathcal{P}(t)A_2(t)+A_2(t)^\top\mathcal{Q}(t)\\
\qquad\qquad + \mathcal{Q}(t)A_2(t)+Q(t)\big]\mathrm{d}t + \mathcal{Q}(t)\mathrm{d}W(t), \quad t\in[0,T],\\
\mathcal{P}(T) = G.
\end{cases} \tag{3.33}$$

Also we define $\mathcal{S}(\cdot)$, $\mathcal{R}(\cdot)$ as,

$$\mathcal{S} := S + B_2^\top \mathcal{P} A_2 + B_1^\top \mathcal{P} + B_2^\top \mathcal{Q}, \quad \mathcal{R} := R + B_2^\top \mathcal{P} B_2. \text{ a.e. a.s.} \tag{3.34}$$

Then one can easily repeat the arguments in Theorems 3.1 and 3.2 to obtain desired conclusions. To get more feelings about this point, let us give one explicit example. Suppose $A_1 = A_2 = 0$, $G = \sin W(T)$, $Q(\cdot) = \cos W(\cdot)$. Using Mallivain calculus one can obtain that

$$\mathcal{P}(t) := \mathbb{E}_t \left[\sin W(T) + \int_t^T \cos W(s)\mathrm{d}s\right], \quad \mathcal{Q}(t) := -\mathbb{E}_t \left[\cos W(T) - \int_t^T \sin W(r)\mathrm{d}r\right],$$

is the unique pair of solutions for (3.33). Of course, above $(\mathcal{P}, \mathcal{Q})$ are bounded processes.

Remark 3.3. To see the basic ideas above, let us make some comparisons with traditional approach. In the existing literature, under certain conditions one can firstly obtain a formal representation of $\widehat{u}(\cdot)$ by (3.11). The following-up steps are plugging $\widehat{u}(\cdot)$ into the state equation and adjoint equation respectively, and solving a fully coupled FBSDE afterward. In contrast, it seems that our method here is established in a essential different way. Firstly by using $\mathcal{P}$ and $(\mathcal{M}, \mathcal{N})$ one can simplify/transform equivalently (3.11), (3.13) into other forms (*i.e.* Step 2). By doing so, one can conveniently represent $(\mathcal{M}, \mathcal{N}, X)$ *via* $\widehat{u}(\cdot)$, plug the representations into (3.24), (3.26), and finally end up with a new equation for $\widehat{u}(\cdot)$, like (3.32). We emphasize that by these tricks we not only provide another way of dealing with optimal controls, but also demonstrate new advantages in dealing with the linear quadratic problem, see *e.g.* Remark 3.5 and Example 3.4 next. More importantly, the ideas here can be nicely applied into the SVIEs setting, which is one of the main aims of this paper.

Let us return back to Example 3.3.

Example 3.4. We study the existence of open-loop optimal control in Example 3.3. Given $P(\cdot)$ in (3.8) and $X_0(\cdot)$ being the state process with $x = 0$, by Itô formula to $P(\cdot)X_0^2(\cdot)$,

$$2\mathbb{E}|X_0(T)|^2 = \mathbb{E}\int_0^T \left[2s^3|u(s)|^2 + 8s^3 X_0(s)u(s) + 6s^2|X_0(s)|^2\right]\mathrm{d}s.$$

Plugging it into the corresponding cost functional, one has

$$J(u(\cdot)) = \mathbb{E}\int_0^T 2s\left[su(s) + X_0(s)\right]^2 \mathrm{d}s \geq 0,$$

from which $\Theta \geq 0$ or (3.18) is fulfilled. Now let us look at the other half. According to the coefficients in Example 3.3,

$$\mathcal{L}_1(t) = 2T^2, \quad \mathcal{L}_2(t) = 2, \quad \mathcal{L}_3(t) = 1, \quad \mathcal{R}(t) = T^2 + 2t^3 - t^2 > 0, \quad \mathcal{S}(t) = 2T^2.$$

In this case, equation (3.15) becomes,

$$\widehat{\lambda}(t) = -2T^2 x - 4T^2 \int_0^T \mathcal{R}^{-1}(s)\widehat{\lambda}(s)\mathrm{d}s - 2T^2 \int_0^t \mathcal{R}^{-1}(s)\widehat{\lambda}(s)\mathrm{d}W(s) - 2T^2 \int_t^T \mathcal{R}^{-1}(s)K_{\widehat{\lambda}}(s,t)\mathrm{d}s. \tag{3.35}$$

To see the solvability of (3.35), we use the idea of fixed point arguments. Given $\mu_i(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ with $i = 1, 2$, suppose $\widehat{\lambda}_i(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ satisfies

$$\widehat{\lambda}_i(t) = -2T^2 x - 4T^2 \int_0^T \mathcal{R}^{-1}(s)\mu_i(s)\mathrm{d}s - 2T^2 \int_0^t \mathcal{R}^{-1}(s)\mu_i(s)\mathrm{d}W(s) - 2T^2 \int_t^T \mathcal{R}^{-1}(s)K_{\mu_i}(s,t)\mathrm{d}s.$$

It follows from some computations that,

$$\mathbb{E}\int_0^T |\widehat{\lambda}_1(s)-\widehat{\lambda}_2(s)|^2 \mathrm{d}s \leq 32T^5\big[4T+1\big]\mathbb{E}\int_0^T \mathcal{R}^{-2}(s)|\mu_1(s)-\mu_2(s)|^2\mathrm{d}s.$$

As to $\mathcal{R}(\cdot)$, it is easy to check that $t\mapsto \mathcal{R}(t)$ is non-increasing in $[0,T]$. Moreover, $R(t)\geq 2T^3$ if $T\leq \frac{1}{3}$. To sum up, one can find a positive constant $\delta_0<1$ such that for any $T<\delta_0$, equation (3.35) is solvable. Therefore for such T, by Theorem 3.1 or 3.2 we can represent optimal control as $\widehat{u}(\cdot)=\mathcal{R}^{-1}(\cdot)\widehat{\lambda}(\cdot)$. □

Remark 3.4. We emphasize that (3.15) can be rewritten as

$$\widehat{\lambda}(t)=-\mathcal{L}_4(t)-\int_0^T H(t,s)\mathcal{R}^{\dagger}(s)\widehat{\lambda}(s)\mathrm{d}s-\int_0^t H_2(t,s)\mathcal{R}^{\dagger}(s)\widehat{\lambda}(s)\mathrm{d}W(s)-\int_t^T H_2(s,t)^{\top}\mathcal{R}^{\dagger}(s)K_{\widehat{\lambda}}(s,t)\mathrm{d}s, \quad (3.36)$$

where for $s,\ t\in[0,T]$, $H_i(\cdot,\cdot)$ and $K_\lambda(\cdot,\cdot)$ are determined by

$$\begin{aligned} &H(t,s)=H_1(t,s)I_{t\geq s}+H_1(s,t)^{\top}I_{s>t}, \qquad H_1(t,s)=\mathcal{L}_1(t)\mathcal{L}_2(s),\\ &H_2(t,s)=\mathcal{L}_1(t)\mathcal{L}_3(s), \quad \lambda(t)=\mathbb{E}\lambda(t)+\int_0^t K_\lambda(t,s)\mathrm{d}W(s), \quad t\in[0,T]. \ \text{ a.e.} \end{aligned} \quad (3.37)$$

We call (3.36) stochastic Fredholm−Volterra integral equation since it has both the features of Fredholm and Volterra integral equations. A slight general form also appears in the next section. □

Remark 3.5. Let us make some points on $\mathcal{R}(\cdot)$. In the first place, if G and $Q(\cdot)$ are indefinite, then $\mathcal{R}(\cdot)\geq R(\cdot)$. Hence the requirement of $\mathcal{R}^{-1}(\cdot)$ is weaker than that of $R^{-1}(\cdot)$. In the second place, the optimal control may exist (see Ex. 3.4) even though $R(\cdot)=0$. Moreover, from Theorem 3.1 or 3.2 the representations of optimal controls do not involve pseudo-inverse if $\mathcal{R}(\cdot)>0$. In the third place, the replacement of $R(\cdot)$ with $\mathcal{R}(\cdot)$ allows us to treat optimal control (not limited to feedback form in the literature) with potential negative $R(\cdot)$. In the fourth place, when $B_2(\cdot)=0$, by maximum principle a necessary optimality condition is $R(\cdot)\geq 0$. Hence the requirement of $R^{-1}(\cdot)$ being well-defined is reasonable. Likewise, for general $B_2(\cdot)$, $\mathcal{R}(\cdot)=R(\cdot)+B_2(\cdot)^{\top}\mathcal{P}(\cdot)B_2(\cdot)\geq 0$ is also necessary when $B_2(\cdot)\neq 0$. Therefore in this case it is more natural to use $\mathcal{R}^{-1}(\cdot)$ rather than $R^{-1}(\cdot)$. At last, the ideas of dealing with $\mathcal{R}(\cdot)$, as well as the necessity of $\mathcal{R}(\cdot)\geq 0$ in some sense inspire us to explore the counterparts under SVIEs setting, the procedures of which have not been discussed elsewhere to our best. □

Remark 3.6. Define $\alpha(\cdot):=B_1(\cdot)^{\top}\widehat{\mathcal{M}}(\cdot)+B_2(\cdot)^{\top}\widehat{\mathcal{N}}(\cdot)$. One can rewrite (3.31), (3.32) as another equivalent forward-backward system,

$$\begin{cases} \widehat{X}(t)=\Psi(t)x+\Psi(t)\displaystyle\int_0^t \mathcal{L}_2(s)\widehat{u}(s)\mathrm{d}s+\Psi(t)\int_0^t \mathcal{L}_3(s)\widehat{u}(s)\mathrm{d}W(s),\\ \alpha(t)=\mathcal{L}_2(t)^{\top}\displaystyle\int_t^T \mathcal{L}_1(s)^{\top}\widehat{u}(s)\mathrm{d}s+\mathcal{L}_3(t)^{\top}\int_t^T K_{\mathcal{L}_1,\widehat{u}}(s,t)\mathrm{d}s-\int_t^T \beta(t,s)\mathrm{d}W(s),\\ \mathcal{R}(t)\widehat{u}(t)+\mathcal{S}(t)\widehat{X}(t)+\alpha(t)=0, \end{cases} \quad (3.38)$$

where the related notations are defined in (3.14). We emphasize that the second equation is a linear backward stochastic Volterra integral equation. We will use this idea in the next section. □

4. The linear quadratic problem of SVIEs

In this section, we discuss linear quadratic optimal control problem of SVIE. Suppose $A_2=0$ and $A_1(\cdot,\cdot)$ is deterministic. As a result, the state equation becomes,

$$X(t)=\varphi(t)+\int_0^t \big[A_1(t,s)X(s)+B_1(t,s)u(s)\big]\mathrm{d}s+\int_0^t B_2(t,s)u(s)\mathrm{d}W(s). \quad (4.1)$$

Before going further, we illustrate the requirements on A_1 and A_2.

Since the coefficients in (1.1) depends on t, the corresponding study becomes much more involved than the SDEs case. To see this, let us recall $\Psi(\cdot)$ in (3.14), from which one can represent state process $X(\cdot)$ *via* $u(\cdot)$, see (3.31). We emphasize that this point is crucial in deriving Theorem 3.1. Now we return back to SVIEs case. Given $b(\cdot,\cdot)$, $\sigma(\cdot,\cdot)\in L^\infty(0,T;L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{n\times n}))$, $\varphi(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, suppose $X(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ is the unique solution of

$$X(t)=\varphi(t)+\int_0^t b(t,s)X(s)\mathrm{d}s+\int_0^t\sigma(t,s)X(s)\mathrm{d}W(s),\qquad t\in[0,T].$$

It was proved in [4] that the solution $X(\cdot)$ can be rewritten as,

$$X(t)=\varphi(t)+\int_0^t\widehat{b}(t,s)\varphi(s)\mathrm{d}s+\int_0^t\widehat{\sigma}(t,s)\varphi(s)\mathrm{d}W(s),\tag{4.2}$$

where the last term is one kind of anticipated stochastic integral (see [4]), and

$$\begin{cases} f_k(t,s)=\displaystyle\int_s^t\sigma(t,\tau)f_{k-1}(\tau,s)\mathrm{d}W(\tau)+\int_s^t b(t,\tau)f_{k-1}(\tau,s)\mathrm{d}\tau,\quad f:=b,\ \sigma,\\ \widehat{b}(t,s)=\displaystyle\sum_{k=1}^\infty b_k(t,s),\quad \widehat{\sigma}(t,s)=\sum_{k=1}^\infty\sigma_k(t,s),\quad b_1(t,s):=b(t,s),\quad \sigma_1(t,s):=\sigma(t,s).\end{cases}\tag{4.3}$$

We point out that the replacement of Itô stochastic integral by Skorohod anticipated stochastic integral is unavoidable, even in the following three special cases.

Case 1. For deterministic $b(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$,

$$\sigma_2(t,s)=\int_s^t b(t,\tau)\sigma(\tau,s)\mathrm{d}\tau+\int_s^t\sigma(t,\tau)\sigma(\tau,s)\mathrm{d}W(\tau),$$

which is of course $\mathcal{F}_t$-measurable. Therefore the corresponding stochastic integral does not make sense in the Itô integral framework. □

Case 2. If $b(\cdot,\cdot)$ is deterministic, $\sigma(t,\cdot)$ does not rely on t, then

$$\sigma_2(t,s)=\left[\int_s^t b(t,\tau)\mathrm{d}\tau+\int_s^t\sigma(\tau)\mathrm{d}W(\tau)\right]\sigma(s).$$

Like $b(\cdot,\cdot)$, here $\sigma_2(\cdot,\cdot)$ is also $\mathcal{F}_t$-measurable. □

Case 3. For random $b(\cdot,\cdot)$, $\sigma(\cdot,\cdot)=0$, we apply (4.2) to SVIE (1.1) with $b(\cdot,\cdot)\equiv A_1(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)\equiv A_2(\cdot,\cdot)=0$. To do so, we use Fubini arguments to the following expression,

$$\int_0^t\widehat{b}(t,s)\int_0^s B_2(s,r)u(r)\mathrm{d}W(r)\mathrm{d}s,\qquad t\in[0,T].$$

However, for random $b(\cdot,\cdot)$, the $\mathcal{F}_t$-measurability of $\widehat{b}(t,\cdot)$ makes this procedure by no means meaningful in Itô integral framework unless $\widehat{b}(\cdot,\cdot)$ is deterministic. □

To sum up, we keep our attention in the Itô integral setting by supposing $A_2\equiv 0$ and deterministic A_1, since the general anticipated stochastic integral reflects almost another different picture. In this case, we observe that by (4.2), we can rewrite (4.1) as,

$$X(t)=\widetilde{\varphi}(t)+\int_0^t\widetilde{B}_1(t,s)u(s)\mathrm{d}s+\int_0^t\widetilde{B}_2(t,s)u(s)\mathrm{d}W(s),$$

with bounded $\widetilde{\varphi}(\cdot)$, $\widetilde{B}_1(\cdot,\cdot)$, $\widetilde{B}_2(\cdot,\cdot)$. Therefore, without loss of generality, we can furthermore assume $A_1=0$.

On the other hand, for random $Q(\cdot)$ and G, using martingale representation arguments we introduce two pairs of processes $(P_Q(\cdot,\cdot), \varLambda_Q(\cdot,\cdot))$, $(P_G(\cdot), \varLambda_G(\cdot))$ as follows,

$$P_Q(t,s) = Q(t) - \int_s^t \varLambda_Q(t,r)\mathrm{d}W(r), \quad \forall s \in [0,t], \quad P_G(s) = G - \int_s^T \varLambda_G(r)\mathrm{d}W(r), \quad \forall s \in [0,T]. \tag{4.4}$$

At this moment, we impose the following assumption through out the whole section,

(H2) Suppose **(H1)** holds true with $A_1 = A_2 = 0$, and deterministic $\varphi(\cdot)$. Moreover, $\varLambda_Q(\cdot,\cdot)$, $\varLambda_G(\cdot)$ are bounded.

Under **(H2)** the state equation then becomes,

$$X(t) = \varphi(t) + \int_0^t B_1(t,s)u(s)\mathrm{d}s + \int_0^t B_2(t,s)u(s)\mathrm{d}W(s), \quad t \in [0,T], \tag{4.5}$$

and the following two processes are bounded,

$$\mathcal{M}_1(\cdot,\cdot) := P_Q(\cdot,\cdot)B_1(\cdot,\cdot) + \varLambda_Q(\cdot,\cdot)B_2(\cdot,\cdot), \qquad \mathcal{N}_1(\cdot) := P_G(\cdot)B_1(T,\cdot) + \varLambda_G(\cdot)B_2(T,\cdot). \tag{4.6}$$

For later usefulness, we define $K_{\mathcal{M}_1,u}(t,s,\cdot)$ and $K_{f,u}(t,\cdot)$ by martingale representation theorem as follows,

$$\begin{cases} \mathcal{M}_1(t,s)u(s) = \mathbb{E}\big[\mathcal{M}_1(t,s)u(s)\big] + \displaystyle\int_0^t K_{\mathcal{M}_1,u}(t,s,r)\mathrm{d}W(r), \quad t,s \in [0,T], \quad \text{a.e.} \\ f(t)u(t) = \mathbb{E}\big[f(t)u(t)\big] + \displaystyle\int_0^t K_{f,u}(t,r)\mathrm{d}W(r), \quad f := \mathcal{N}_1(\cdot),\ S(\cdot)^\top, \quad t \in [0,T]. \ \text{a.e.} \end{cases} \tag{4.7}$$

Remark 4.1. Observe that the general discussion with random $\varphi(\cdot)$ can be treated without involving additional essential difficulties. On the other hand, we give some special cases with bounded $\varLambda_Q$ and $\varLambda_G$. For example, for any $t \in [0,T]$, let $Q(t) := f(W(t))$ where function $f(\cdot)$, $f_x(\cdot)$ are bounded. In this case, from Malliavin calculus we know that $\varLambda_Q(t,r) = \mathbb{E}_r f_x(W(t))$ with $r \in [0,t]$. Obviously, $\varLambda_Q(\cdot,\cdot)$ is bounded. Similarly, let $G = g(W(T))$ where function $g(\cdot)$, $g_x(\cdot)$ are bounded. We can see the boundedness of $\varLambda_G(\cdot) = \mathbb{E}_\cdot g_x(W(T))$ by Malliavin calculus as well.

4.1. New characterizations of optimal controls

In this part, we adjust the previous ideas on SDEs into our SVIEs framework and give new equivalent conditions for the existence of optimal controls. Before it, let us point out two more theoretical motivations.

4.1.1. Some motivations

Firstly let us look at the conventional routine. By supposing $R^{-1}(\cdot)$ to be well-defined and bounded, according to Lemma 2.2, optimal control $\widehat{u}(\cdot)$ can be shown as (see [8,27]),

$$\widehat{u}(t) = -R^{-1}(t)\big[S(t)\widehat{X}(t) + \widehat{\chi}(t) + B_1(T,t)^\top \mathbb{E}_t[G\widehat{X}(T)] + B_2(T,t)^\top \widehat{\theta}(t)\big], \quad t \in [0,T], \quad \text{a.e.} \tag{4.8}$$

where the related forward-backward system is,

$$
\begin{cases}
\widehat{X}(t)=\varphi(t)-\int_0^t \big[B_1(t,s)R^{-1}(s)S(s))\widehat{X}(s)+B_1(t,s)R^{-1}(s)\widehat{\chi}(s)\big]\mathrm{d}s\\
\qquad\quad -\int_0^t B_1(t,s)R^{-1}(s)\big[B_1(T,s)^\top\mathbb{E}_s[G\widehat{X}(T)]+B_2(T,s)^\top\widehat{\theta}(s)\big]\mathrm{d}s\\
\qquad\quad -\int_0^t \big[B_2(t,s)R^{-1}(s)S(s))\widehat{X}(s)+B_2(t,s)R^{-1}(s)\widehat{\chi}(s)\big]\mathrm{d}W(s)\\
\qquad\quad -\int_0^t B_2(t,s)R^{-1}(s)\big[B_1(T,s)^\top\mathbb{E}_s[G\widehat{X}(T)]+B_2(T,s)^\top\widehat{\theta}(s)\big]\mathrm{d}W(s),\\
\widehat{Y}(t)=\big[Q(t)-S(t)R^{-1}(t)S(t)\big]\widehat{X}(t)-S(t)R^{-1}(t)\widehat{\chi}(t)\\
\qquad\quad -S(t)R^{-1}(t)\big[B_1(T,t)^\top\mathbb{E}_t[G\widehat{X}(T)]+B_2(T,t)^\top\widehat{\theta}(t)\big]-\int_t^T\widehat{Z}(t,s)\mathrm{d}W(s),\\
\widehat{\chi}(t)=\mathbb{E}_t\int_t^T\Big[B_1(s,t)^\top\widehat{Y}(s)+B_2(s,t)^\top\widehat{Z}(s,t)\Big]\mathrm{d}s,\quad t\in[0,T],\\
\widehat{Y}(t)=\mathbb{E}\widehat{Y}(t)+\int_0^t\widehat{Z}(t,s)\mathrm{d}W(s),\qquad G\widehat{X}(T)=\mathbb{E}\big[G\widehat{X}(T)\big]+\int_0^T\widehat{\theta}(s)\mathrm{d}W(s).
\end{cases}
\tag{4.9}
$$

However, it seems that the effectiveness of these ideas is questioning due to the following aspects, even when all the coefficients are deterministic. First of all, if optimal control exists, (4.9) and (4.8) may fail to work as long as $R(\cdot)=0$. In other words, above approach only works with well-defined $R^{-1}(\cdot)$. Second of all, the appearing of G not only makes (4.8) and (4.9) complicated, but also brings additional difficulties in solving (4.9) with fixed point argument. Third of all, (4.9) can be considerably simplified when $G=0$ and $S(\cdot)=0$. Even so, to guarantee its well-posedness by contraction method, $Q(\cdot)$ or T is required to be small enough. In conclusion, new ideas are demanded to be inject into this topic.

Our second motivation is concerned with the coefficients in the cost functional. As we know, by stochastic Riccati equations and the method of completion of square, the existence of optimal control can be ensured even when the control weight costs $R(\cdot)$ is indefinite. Nevertheless, as mentioned in the Introduction, the tricks involving Riccati equations do not work here. How to demonstrate the similar indefinite property of $R(\cdot)$ under SVIEs setting becomes an interesting problem. In the second place, it seems that the choice on the coefficients is important for condition (2.7) in Lemma 2.2. For example, this condition holds true if $R(\cdot)$, $Q(\cdot)$, G are nonnegative, $S(\cdot)=0$. However, it becomes complicated to check if one of $R(\cdot)$, $Q(\cdot)$, G is negative, or $S(\cdot)\neq 0$. How to go further along this line is another important problem to work on.

4.1.2. The general case with random coefficients

To begin with, we make proper change to $X(\cdot)$ by introducing $\mathcal{X}(\cdot,\cdot)$ as,

$$
\mathcal{X}(t,s)=\varphi(t)+\int_0^s B_1(t,r)u(r)\mathrm{d}r+\int_0^s B_2(t,r)u(r)\mathrm{d}W(r),\quad \forall s\in[0,t],\quad \forall t\in[0,T],
\tag{4.10}
$$

where $\mathcal{X}(t,t)=X(t)$ for $t\in[0,T]$. Similarly one can introduce $\mathcal{X}_0(\cdot,\cdot)$ with $\varphi=0$.

Lemma 4.1. *Suppose G, $Q(\cdot)$ are bounded, $X(\cdot)\in C_{\mathcal{F}}([0,T];L^2(\Omega;\mathbb{R}^n))$ satisfies (4.5), (P_Q,Λ_Q), (P_G,Λ_G), and $K_{\mathcal{M}_1,u}(\cdot,\cdot)$, $K_{\mathcal{N}_1,u}(\cdot,\cdot)$ are in (4.4) and (4.7) respectively. Then we have,*

$$
\begin{cases}
Q(t)X(t)=\mathbb{E}\big[Q(t)X(t)\big]+\int_0^t K_{Q,X}(t,s)\mathrm{d}W(s),\qquad t\in[0,T],\quad \text{a.e.}\\
GX(T)=\mathbb{E}\big[GX(T)\big]+\int_0^T K_{G,X}(s)\mathrm{d}W(s),
\end{cases}
$$

where $K_{Q,X}$, $K_{G,X}$ are defined by

$$\begin{cases} K_{Q,X}(t,s) := \int_s^t K_{\mathcal{M}_1,u}(t,r,s)\mathrm{d}r + \big[P_Q(t,s)B_2(t,s)u(s) + \varLambda_Q(t,s)\mathcal{X}(t,s)\big], \quad t \geq s, \text{ a.e.} \\ K_{G,X}(s) := \int_s^T K_{\mathcal{N}_1,u}(r,s)\mathrm{d}r + \big[P_G(s)B_2(T,s)u(s) + \varLambda_G(s)\mathcal{X}(T,s)\big], \quad s \in [0,T]. \text{ a.e.} \end{cases} \tag{4.11}$$

Proof. Given $t \in [0,T]$, let us use Itô formula to $P_Q(t,\cdot)\mathcal{X}(t,\cdot)$ on $[0,t]$,

$$\begin{aligned} Q(t)X(t) - P_Q(t,0)\varphi(t) = &\int_0^t \big[P_Q(t,s)B_1(t,s) + \varLambda_Q(t,s)B_2(t,s)\big]u(s)\mathrm{d}s \\ &+ \int_0^t \big[P_Q(t,s)B_2(t,s)u(s) + \varLambda_Q(t,s)\mathcal{X}(t,s)\big]\mathrm{d}W(s). \end{aligned} \tag{4.12}$$

Due to notation (4.6) and Lemma 3.2, the first term on the right hand of (4.12) can be rewritten as,

$$\int_0^t \mathcal{M}_1(t,s)u(s)\mathrm{d}s = \mathbb{E}\int_0^t \mathcal{M}_1(t,s)u(s)\mathrm{d}s + \int_0^t \int_r^t K_{\mathcal{M}_1,u}(t,s,r)\mathrm{d}s\mathrm{d}W(r). \tag{4.13}$$

Then the result of $Q(\cdot)X(\cdot)$ in (4.11) follows from (4.12) and (4.13). The case of $K_{G,X}(\cdot)$ can be obtained in a similar way. □

Given $t,\ r \in [0,T]$, $\mathcal{X}(\cdot,\cdot)$ in (4.10) and the notations in (4.4), (4.6), (4.7), we introduce

$$\begin{cases} \mathcal{R}(t) := \int_t^T B_2(s,t)^\top P_Q(s,t)B_2(s,t)\mathrm{d}s + B_2(T,t)^\top P_G(t)B_2(T,t) + R(t), \\ \mathcal{L}(r,t) := \int_r^T B_1(s,t)^\top \mathcal{M}_1(s,r)\mathrm{d}s + B_1(r,t)^\top S(r)^\top + B_1(T,t)^\top \mathcal{N}_1(r), \\ \mathcal{K}(r,t;u(\cdot)) := \int_r^T B_2(s,t)^\top K_{\mathcal{M}_1,u}(s,r,t)\mathrm{d}s + B_2(r,t)^\top K_{S^T,u}(r,t) + B_2(T,t)^\top K_{\mathcal{N}_1,u}(r,t), \\ \mathcal{G}_1(t,\mathcal{X}(\cdot,t)) := \mathbb{E}_t\int_t^T \mathcal{M}_1(s,t)^\top \mathcal{X}(s,t)\mathrm{d}s + \mathcal{N}_1(t)^\top \mathcal{X}(T,t) + S(t)\mathcal{X}(t,t). \end{cases} \tag{4.14}$$

Note that $\mathcal{R}(t)$ is $\mathcal{F}_t$-measurable, $\mathcal{R}(\cdot)$, $\mathcal{L}(\cdot,\cdot)$ only depends on given coefficients. We introduce the following forward-backward system,

$$\begin{cases} \widehat{\mathcal{X}}(t,s) = \varphi(t) + \int_0^s B_1(t,r)\widehat{u}(r)\mathrm{d}r + \int_0^s B_2(t,r)\widehat{u}(r)\mathrm{d}W(r), \quad s \in [0,t], \\ \widehat{\xi}(t) = \int_t^T \big[\mathcal{L}(r,t)\widehat{u}(r) + \mathcal{K}(r,t;\widehat{u}(\cdot))\big]\mathrm{d}r - \int_t^T \widehat{\zeta}(t,s)\mathrm{d}W(s), \\ \mathcal{R}(t)\widehat{u}(t) + \widehat{\xi}(t) + \mathcal{G}_1(t,\widehat{\mathcal{X}}(\cdot,t)) = 0, \quad t \in [0,T]. \end{cases} \tag{4.15}$$

Theorem 4.1. *Suppose* (H2) *holds true. Then the linear quadratic problem is solvable with optimal* $\widehat{u}(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ *if and only if the following hold,*

(i) *system* (4.15)*, admits a triple of* $(\widehat{\mathcal{X}}(\cdot,\cdot),\widehat{\xi}(\cdot),\widehat{\zeta}(\cdot,\cdot))$*;*
(ii) *for any* $u(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$*,*

$$2\mathbb{E}\int_0^T u(s)^\top \mathcal{G}_1(s,\mathcal{X}_0(\cdot,s))\mathrm{d}s + \mathbb{E}\int_0^T u(s)^\top \mathcal{R}(s)u(s)\mathrm{d}s \geq 0, \tag{4.16}$$

where $\mathcal{X}_0(\cdot,\cdot)$ *is in* (4.10) *with* $\varphi = 0$*. In this case,*

$$\widehat{u}(t) = -\mathcal{R}^\dagger(t)\big[\widehat{\xi}(t) + \mathcal{G}_1(t,\widehat{\mathcal{X}}(\cdot,t))\big] + \big[I - \mathcal{R}^\dagger(t)\mathcal{R}(t)\big]u'(t), \quad t\in[0,T], \quad \text{a.e.} \tag{4.17}$$

for some $u'(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$*.*

Proof. We point out that the third equality in (4.15) implies directly expression (4.17). As to the rest, we shall separate the proof into two parts.

Step 1. The equivalent link between (2.8) and (4.15).

Under (H2), the first backward equation in (2.6) associated with $\widehat{u}(\cdot)$ becomes,

$$\widehat{Y}(t) = Q(t)\widehat{X}(t) + S(t)^\top\widehat{u}(t), \quad \widehat{Z}(t,s) = 0, \quad 0\leq t\leq s\leq T. \tag{4.18}$$

From (4.12) in Lemma 4.1, for almost all $s\in[0,T]$,

$$\begin{aligned} Q(s)\widehat{X}(s) &= P_Q(s,t)\widehat{\mathcal{X}}(s,t) + \int_t^s \mathcal{M}_1(s,r)\widehat{u}(r)\mathrm{d}r \\ &\quad + \int_t^s \big[P_Q(s,r)B_2(s,r)\widehat{u}(r) + \varLambda_Q(s,r)\widehat{\mathcal{X}}(s,r)\big]\mathrm{d}W(r), \qquad t\in[0,s], \end{aligned}$$

where $\mathcal{M}_1(\cdot,\cdot)$ is defined in (4.6). Hence by (4.18) and Fubini theorem,

$$\begin{aligned} \mathbb{E}_t\int_t^T B_1(s,t)^\top\widehat{Y}(s)\mathrm{d}s &= \mathbb{E}_t\int_t^T B_1(s,t)^\top P_Q(s,t)\widehat{\mathcal{X}}(s,t)\mathrm{d}s \\ &\quad + \mathbb{E}_t\int_t^T\left[\int_r^T B_1(s,t)^\top\mathcal{M}_1(s,r)\mathrm{d}s + B_1(r,t)^\top S(r)^\top\right]\widehat{u}(r)\mathrm{d}r. \end{aligned} \tag{4.19}$$

Due to Lemma 4.1 and $\widehat{Y}(\cdot)$ in (4.18), we also have

$$\begin{aligned} \mathbb{E}_t\int_t^T B_2(s,t)^\top\widehat{Z}(s,t)\mathrm{d}s &= \mathbb{E}_t\int_t^T B_2(s,t)^\top\big[K_{Q,\widehat{X}}(s,t) + K_{S^\top,\widehat{u}}(s,t)\big]\mathrm{d}s \\ &= \mathbb{E}_t\int_t^T\int_r^T B_2(s,t)^\top K_{\mathcal{M}_1,\widehat{u}}(s,r,t)\mathrm{d}s\mathrm{d}r + \mathbb{E}_t\int_t^T B_2(s,t)^\top K_{S^\top,\widehat{u}}(s,t)\mathrm{d}s \\ &\quad + \left[\mathbb{E}_t\int_t^T B_2(s,t)^\top P_Q(s,t)B_2(s,t)\mathrm{d}s\right]\cdot\widehat{u}(t) + \mathbb{E}_t\int_t^T B_2(s,t)^\top\varLambda_Q(s,t)\widehat{\mathcal{X}}(s,t)\mathrm{d}s, \end{aligned} \tag{4.20}$$

where $K_{S^\top,\widehat{u}}(\cdot,\cdot)$ is defined in (4.7). Therefore it follows from (4.19) and (4.20) that

$$\begin{aligned} \widehat{\chi}(t) = {} & \mathbb{E}_t\int_t^T\big[B_1(s,t)^\top P_Q(s,t) + B_2(s,t)^\top\varLambda_Q(s,t)\big]\widehat{\mathcal{X}}(s,t)\mathrm{d}s \\ & + \mathbb{E}_t\int_t^T\left[\int_r^T B_1(s,t)^\top\mathcal{M}_1(s,r)\mathrm{d}s + B_1(r,t)^\top S(r)^\top\right]\widehat{u}(r)\mathrm{d}r + \mathbb{E}_t\int_t^T\int_r^T B_2(s,t)^\top K_{\mathcal{M}_1,\widehat{u}}(s,r,t)\mathrm{d}s\mathrm{d}r \\ & + \left[\mathbb{E}_t\int_t^T B_2(s,t)^\top P_Q(s,t)B_2(s,t)\mathrm{d}s\right]\widehat{u}(t) + \mathbb{E}_t\int_t^T B_2(s,t)^\top K_{S^\top,\widehat{u}}(s,t)\mathrm{d}s. \end{aligned} \tag{4.21}$$

As to $B_1(T,t)^\top \mathbb{E}_t G\widehat{X}(T) + B_2(T,t)^\top \widehat{\theta}(t)$, $t \in [0,T]$, a.e. recalling $\mathcal{X}(T,\cdot)$ in (4.10), we obtain the following by Itô formula

$$G\widehat{X}(T) = P_G(0)\widehat{\mathcal{X}}(T,0) + \int_0^T \mathcal{N}_1(r)\widehat{u}(r)\mathrm{d}r + \int_0^T \big[P_G(r)B_2(T,r) + \Lambda_G(r)\widehat{\mathcal{X}}(T,r)\big]\mathrm{d}W(r),$$

where $\mathcal{N}_1(\cdot)$ is defined in (4.7). Then for $t \in [0,T]$, a.e.

$$B_1(T,t)^\top \mathbb{E}_t\big[G\widehat{X}(T)\big] = B_1(T,t)^\top P_G(t)\widehat{\mathcal{X}}(T,t) + B_1(T,t)^\top \mathbb{E}_t \int_t^T \mathcal{N}_1(r)\widehat{u}(r)\mathrm{d}r. \tag{4.22}$$

Thanks to Lemma 4.1, for $t \in [0,T]$, a.e.

$$B_2(T,t)^\top \widehat{\theta}(t) = B_2(T,t)^\top \int_t^T K_{\mathcal{N}_1,\widehat{u}}(r,t)\mathrm{d}r + \big[B_2(T,t)^\top P_G(t)B_2(T,t)\big]\widehat{u}(t) + B_2(T,t)^\top \Lambda_G(t)\widehat{\mathcal{X}}(T,t). \tag{4.23}$$

Therefore from (4.22) and (4.23) we can deduce that

$$\begin{aligned} B_1(T,t)^\top \mathbb{E}_t\big[G\widehat{X}(T)\big] + B_2(T,t)^\top \widehat{\theta}(t) &= \mathbb{E}_t \int_t^T B_1(T,t)^\top \mathcal{N}_1(r)\widehat{u}(r)\mathrm{d}r + \big[B_2(T,t)^\top P_G(t)B_2(T,t)\big]\widehat{u}(t) \\ &\quad + \int_t^T B_2(T,t)^\top K_{\mathcal{N}_1,\widehat{u}}(r,t)\mathrm{d}r \\ &\quad + \big[B_1(T,t)^\top P_G(t) + B_2(T,t)^\top \Lambda_G(t)\big]\widehat{\mathcal{X}}(T,t). \end{aligned} \tag{4.24}$$

To sum up, recalling (4.14), we obtain the desired conclusion from (4.21), (4.24) and

$$\widehat{\chi}(t) + S(t)\widehat{X}(t) + R(t)\widehat{u}(t) + B_1(T,t)^\top \mathbb{E}_t\big[G\widehat{X}(T)\big] + B_2(T,t)^\top \widehat{\theta}(t) = \mathcal{G}_1(t,\widehat{\mathcal{X}}(\cdot,t)) + \widehat{\xi}(t) + \mathcal{R}(t)\widehat{u}(t).$$

Step 2. The equivalent relation between (2.7) and (4.16).

Following above arguments, one can also derive

$$\chi_0(t) + S(t)X_0(t) + R(t)u(t) + B_1(T,t)\mathbb{E}_t\big[GX_0(T)\big] + B_2(T,t)\theta_0(t) = \mathcal{G}_1(t,\mathcal{X}_0(\cdot,t)) + \xi(t) + \mathcal{R}(t)u(t), \tag{4.25}$$

where

$$\xi(t) = \mathbb{E}_t \int_t^T \big[\mathcal{L}(r,t)u(r) + \mathcal{K}(r,t;u(\cdot))\big]\mathrm{d}r. \tag{4.26}$$

Next we claim that

$$\mathbb{E}\int_0^T u(t)^\top \mathcal{G}_1(t,\mathcal{X}_0(\cdot,t))\mathrm{d}t = \mathbb{E}\int_0^T u(t)^\top \int_t^T \big[\mathcal{L}(r,t)u(r) + \mathcal{K}(r,t;u(\cdot))\big]\mathrm{d}r\mathrm{d}t. \tag{4.27}$$

Then by combining (2.7), (4.27) and (4.25), one can finish the procedures in this step. To prove (4.27), let us look at the following equalities,

$$\begin{aligned} &\mathbb{E}\int_0^T \left\langle \int_s^T \mathcal{M}_1(t,s)^\top \left[\int_0^s B_1(t,r)u(r)\mathrm{d}r\right]\mathrm{d}t, u(s)\right\rangle \mathrm{d}s \\ &= \mathbb{E}\int_0^T \int_s^T \int_0^s \langle B_1(t,r)u(r), \mathcal{M}_1(t,s)u(s)\rangle\, \mathrm{d}r\mathrm{d}t\mathrm{d}s \\ &= \mathbb{E}\int_0^T \int_0^s \int_s^T \langle u(r), B_1(t,r)^\top \mathcal{M}_1(t,s)u(s)\rangle\, \mathrm{d}t\mathrm{d}r\mathrm{d}s \\ &= \mathbb{E}\int_0^T \left\langle u(r), \int_r^T \int_s^T B_1(t,r)^\top \mathcal{M}_1(t,s)u(s)\mathrm{d}t\mathrm{d}s\right\rangle \mathrm{d}r, \end{aligned} \tag{4.28}$$

which is derived by classical Fubini theorem. Moreover, by applying Lemma 2.2, we also have,

$$
\begin{aligned}
&\mathbb{E}\int_0^T \left\langle u(s), \int_s^T \mathcal{M}_1(t,s)^\top \int_0^s B_2(t,r)u(r)\mathrm{d}W(r)\mathrm{d}t \right\rangle \mathrm{d}s \\
&= \mathbb{E}\int_0^T \int_s^T \left\langle \mathcal{M}_1(t,s)u(s), \int_0^s B_2(t,r)u(r)\mathrm{d}W(r) \right\rangle \mathrm{d}t\mathrm{d}s \\
&= \mathbb{E}\int_0^T \int_s^T \int_0^s \langle B_2(t,r)^\top K_{\mathcal{M}_1,u}(t,s,r), u(r) \rangle \mathrm{d}r\mathrm{d}t\mathrm{d}s \\
&= \mathbb{E}\int_0^T \left\langle u(r), \int_r^T \int_s^T B_2^{(}t,r)^\top K_{\mathcal{M}_1,u}(t,s,r)\mathrm{d}t\mathrm{d}s \right\rangle \mathrm{d}r.
\end{aligned}
\tag{4.29}
$$

As a result, by (4.28) and (4.29),

$$
\begin{aligned}
&\mathbb{E}\int_0^T \left\langle u(s), \int_s^T \mathcal{M}_1(t,s)^\top \mathcal{X}_0(t,s)\mathrm{d}t \right\rangle \mathrm{d}s \\
&= \mathbb{E}\int_0^T \left\langle u(s), \int_s^T \mathcal{M}_1(t,s)^\top \left[\int_0^s B_1(t,r)u(r)\mathrm{d}r\right] \mathrm{d}t \right\rangle \mathrm{d}s \\
&\quad +\mathbb{E}\int_0^T \left\langle u(s), \int_s^T \mathcal{M}_1(t,s)^\top \left[\int_0^s B_2(t,r)u(r)\mathrm{d}W(r)\right] \mathrm{d}t \right\rangle \mathrm{d}s \\
&= \mathbb{E}\int_0^T \left\langle u(r), \int_r^T \int_s^T B_1(t,r)^\top \mathcal{M}_1(t,s)u(s)\mathrm{d}t\mathrm{d}s \right\rangle \mathrm{d}r \\
&\quad +\mathbb{E}\int_0^T \left\langle u(r), \int_r^T \int_s^T B_2(t,r)^\top K_{\mathcal{M}_1,u}(t,s,r)\mathrm{d}t\mathrm{d}s \right\rangle \mathrm{d}r.
\end{aligned}
\tag{4.30}
$$

Similarly we have,

$$
\begin{cases}
\mathbb{E}\int_0^T \left\langle u(s), \mathcal{N}_1(s)^\top \mathcal{X}_0(T,s)\right\rangle \mathrm{d}s = \mathbb{E}\int_0^T \left\langle u(s), \int_s^T B_1(T,s)^\top \mathcal{N}_1(r)u(r)\mathrm{d}r \right\rangle \mathrm{d}s \\
\qquad +\mathbb{E}\int_0^T \left\langle u(r), \int_r^T B_2(T,r)^\top K_{\mathcal{N}_1,u}(s,r)\mathrm{d}s \right\rangle \mathrm{d}r, \\
\mathbb{E}\int_0^T \langle u(s), S(s)\mathcal{X}_0(s,s)\rangle \mathrm{d}s = \mathbb{E}\int_0^T \left\langle u(s), \int_s^T B_1(r,s)^\top S(r)^\top u(r)\mathrm{d}r \right\rangle \mathrm{d}s \\
\qquad +\mathbb{E}\int_0^T \left\langle u(r), \int_r^T B_2(s,r)^\top K_{S^\top,u}(s,r)\mathrm{d}s \right\rangle \mathrm{d}r.
\end{cases}
\tag{4.31}
$$

As a result, (4.27) follows from (4.30) and (4.31). □

Remark 4.2. In the above arguments we proposed two new ideas which fits the SVIEs framework well. Firstly, we introduce $(P_Q(\cdot,\cdot), \Lambda_Q(\cdot,\cdot))$ and $(P_G(\cdot,\cdot), \Lambda_G(\cdot,\cdot))$. Their connections with $(\mathcal{P}(\cdot), \mathcal{Q}(\cdot))$ of (3.33) can be found in next Remark 4.4. Secondly, instead of using $X(\cdot)$ as above, we prefer to slight different $\mathcal{X}(\cdot,\cdot)$ which makes Itô formula useful again. Also it is worthy mentioning that these two points helps us derive new pointwise necessary optimality conditions in the Section 4.2 next. □

Remark 4.3. In contrast with conventional results (4.8) and (4.9), ours here have advantages in several aspects. The first one is concerned with the replacement of $R(\cdot)$ by $\mathcal{R}(\cdot)$, which considerably relaxes the existing condition on $R(\cdot)$. Related details can be found in Remark 3.5. The second one lies in the simplicity of optimal controls. Notice that in (4.8), $\widehat{u}(\cdot)$ depends on $\widehat{X}(\cdot)$, $\widehat{X}(T)$, $\widehat{\chi}(\cdot)$ and $\widehat{\theta}(\cdot)$, while $\widehat{\chi}(\cdot)$ furthermore relies on $(\widehat{X}(\cdot), \widehat{Y}(\cdot), \widehat{Z}(\cdot,\cdot))$. In some sense these indexes are a little bit of scattered and superficial. How to find a proper way to reorganize and incorporate the coefficients becomes important. Here using some essential terms $\mathcal{X}(\cdot,\cdot)$ of (4.10) and $(P_Q, \varLambda_Q)$, $(P_G, \varLambda_G)$ of (4.4), we only require $\widehat{u}(\cdot)$ to depend on $\widehat{\mathcal{X}}(\cdot,t)$ and $\widehat{\xi}$, see (4.17). At last, when applying the ideas to deterministic coefficients case, one can obtain clearer pictures of optimal controls and related equations, see the arguments in the next Section 4.1.3. □

Remark 4.4. In Theorem 4.1 there are three crucial terms, $\mathcal{R}(\cdot)$, $\xi(\cdot)$ and $\mathcal{G}_1(\cdot, \mathcal{X}(\cdot,\cdot))$. We will look into their counterparts in the SDEs setting, and show some connections between Theorem 4.1 and Theorem 3.1. With $A_1 = A_2 = 0$, we rewrite above (3.33) as follows,

$$\mathcal{P}(t) = G + \int_t^T Q(s)\mathrm{d}s - \int_t^T \mathcal{Q}(s)\mathrm{d}W(s), \qquad t \in [0,T]. \tag{4.32}$$

By Lemma 3.2, for almost $t \in [0,T]$, we can express $(\mathcal{P}(\cdot), \mathcal{Q}(\cdot))$ as,

$$\mathcal{P}(t) = \mathbb{E}_t\left[G + \int_t^T Q(s)\mathrm{d}s\right] = P_G(t) + \int_t^T P_Q(s,t)\mathrm{d}s, \qquad \mathcal{Q}(t) = \varLambda_G(t) + \int_t^T \varLambda_Q(s,t)\mathrm{d}s, \tag{4.33}$$

where $(P_Q, \varLambda_Q)$, $(P_G, \varLambda_G)$ are defined in (4.4).

(1) Let us look at $\mathcal{R}(\cdot)$ with $B_2(t,\cdot)$ independent of t. In this setting, for $t \in [0,T]$,

$$\mathcal{R}(t) = B_2(t)^\top \int_t^T P_Q(s,t)\mathrm{d}s B_2(t) + B_2(t)^\top P_G(t) B_2(t) + R(t) = B_2(t)^\top \mathcal{P}(t) B_2(t) + R(t). \tag{4.34}$$

Thanks to the dependence of $B_2(t,\cdot)$ on t, the role of $(\mathcal{P}(\cdot), \mathcal{Q}(\cdot))$ is replaced by that of $(P_Q, \varLambda_Q)$, $(P_G, \varLambda_G)$. On the other hand, (4.34) also inspires us the following study in Section 4.2 since $R + B_2^\top \mathcal{P} B_2 \geq 0$ is a necessary optimality condition under SDEs framework.

(2) If $B_1(t,\cdot)$, $B_2(t,\cdot)$ do not rely on t, then for almost $t \in [0,T]$, and $s \in [0,t]$,

$$\mathcal{X}(t,s) = X(s), \quad \mathcal{M}_1(t,s) = P_Q(t,s)B_1(s) + \varLambda_Q(t,s)B_2(s), \quad \mathcal{N}_1(t) = P_G(t)B_1(t) + \varLambda_G(t)B_2(t),$$

and $\mathcal{G}_1(\cdot, \mathcal{X}(\cdot,\cdot))$ becomes,

$$\begin{aligned}\mathcal{G}_1(t, \mathcal{X}(\cdot,t)) &= \left[\int_t^T \left[B_1(t)^\top P_Q(s,t) + B_2(t)^\top \varLambda_Q(s,t)\right]\mathrm{d}s + B_1(t)^\top P_G(t) + B_2(t)^\top \varLambda_G(t) + S(t)\right] X(t) \\ &= \left[B_1(t)^\top \mathcal{P}(t) + B_2(t)^\top \mathcal{Q}(t) + S(t)\right] X(t) = \mathcal{S}(t) X(t),\end{aligned}$$

where $\mathcal{S}(\cdot)$ is defined in (3.34) of Remark 3.2. Hence the following term ($\mathcal{G}_1(\cdot, \mathcal{X}(\cdot,\cdot))$ resp.) can be seen an extension of $\mathcal{S}(\cdot)$ ($\mathcal{S}(\cdot)X(\cdot)$ reps.) into SVIEs framework,

$$\bar{\mathcal{S}}(t) := \int_t^T \mathcal{M}_1(s,t)^\top \mathrm{d}s + \mathcal{N}_1(t)^\top + S(t), \qquad t \in [0,T]. \ \text{ a.e.} \tag{4.35}$$

(3) If $B_i(t,\cdot)$ is independent of t, then

$$\begin{aligned}\mathcal{L}(r,t) &= B_1(t)^\top \left[\int_r^T \big[P_Q(s,r)B_1(r)+\varLambda_Q(s,r)B_2(r)\big]\mathrm{d}s + S(r)^\top + P_G(r)B_1(r)+\varLambda_G(r)B_2(r)\right]\\ &= B_1(t)^\top\left[\mathcal{P}(r)B_1(r)+\mathcal{Q}(r)B_2(r)+S(r)^\top\right] = B_1(t)^\top\mathcal{S}(r)^\top,\\ \mathcal{K}(r,t;u(\cdot)) &= B_2(t)^\top\left[\int_r^T K_{\mathcal{M}_1,u}(s,r,t)\mathrm{d}s + K_{\mathcal{N},u}(r,t)+K_{S^\top,u}(r,t)\right].\end{aligned} \tag{4.36}$$

According to (4.35),

$$\begin{aligned}\mathcal{S}(r)^\top u(r) &= \int_r^T \mathcal{M}_1(s,r)u(r)\mathrm{d}s + \mathcal{N}_1(r)u(r)+S(r)^\top u(r)\\ &= \mathbb{E}\int_r^T \mathcal{M}_1(s,r)u(r)\mathrm{d}s + \int_0^r\int_r^T K_{\mathcal{M}_1,u}(s,r,t)\mathrm{d}s\mathrm{d}W(t)\\ &\quad + \mathbb{E}\big[\mathcal{N}_1(r)u(r)\big] + \int_0^r K_{\mathcal{N},u}(r,t)\mathrm{d}W(t) + \mathbb{E}\big[S(r)^\top u(r)\big] + \int_0^r K_{S^\top,u}(r,t)\mathrm{d}W(t)\\ &= \mathbb{E}\big[\mathcal{S}(r)^\top u(r)\big] + \int_0^r\left[\int_r^T K_{\mathcal{M}_1,u}(s,r,t)\mathrm{d}s + K_{\mathcal{N},u}(r,t)+K_{S^\top,u}(r,t)\right]\mathrm{d}W(t).\end{aligned}$$

Hence if we denote by

$$\mathcal{M}(t) := \mathbb{E}_t\int_t^T \mathcal{S}(s)^\top u(s)\mathrm{d}s,\quad \mathcal{N}(t) := \int_t^T\left[\int_r^T K_{\mathcal{M}_1,u}(s,r,t)\mathrm{d}s + K_{\mathcal{N},u}(r,t)+K_{S^\top,u}(r,t)\right]\mathrm{d}r,$$

similar as (4.32)–(4.33), one can easily obtain that $(\mathcal{M}(\cdot),\mathcal{N}(\cdot))$ satisfies BSDE of

$$\mathcal{M}(t) = \int_t^T \mathcal{S}(s)^\top u(s)\mathrm{d}s - \int_t^T \mathcal{N}(s)\mathrm{d}W(s), \qquad t\in[0,T]. \tag{4.37}$$

Recall $\widehat{\xi}(\cdot)$ in (4.15), as well as (4.36), we then derive,

$$\xi(t) = B_1(t)^\top\mathcal{M}(t) + B_2(t)^\top\mathcal{N}(t), \qquad t\in[0,T].\ \text{ a.e.}$$

We emphasize that $(\mathcal{M}(\cdot),\mathcal{N}(\cdot))$ in (4.37) is just the pair of processes in the proof of Theorem 3.1 with $A_1 = A_2 = 0$ and deterministic coefficients. □

4.1.3. The deterministic coefficients case

In this part, we give detailed study with deterministic coefficients. To this end, one can simplify the aforementioned coefficients in (4.6), (4.14) as,

$$\begin{cases}\mathcal{M}_1(t,r) = Q(t)B_1(t,r),\quad \mathcal{N}_1(t) = GB_1(T,t),\quad \mathcal{K}(r,t;u(\cdot)) := \mathcal{B}(r,t)K_u(r,t),\\ \mathcal{L}(r,t) := \displaystyle\int_r^T B_1(s,t)^\top Q(s)B_1(s,r)\mathrm{d}s + B_1(r,t)^\top S(r) + B_1(T,t)^\top GB_1(T,r),\\ \mathcal{B}(r,t) := \displaystyle\int_r^T B_2(s,t)^\top Q(s)B_1(s,r)\mathrm{d}s + B_2(r,t)^\top S(r) + B_2(T,t)^\top GB_1(T,r).\end{cases} \tag{4.38}$$

In the following, with $\mathcal{L}(\cdot,\cdot)$ in (4.38), we define $\mathcal{H}(\cdot,\cdot)$ as

$$\mathcal{H}(r,t) := \mathcal{L}(r,t)I_{[r\geq t]} + \mathcal{L}(t,r)^\top I_{[t>r]},\quad t,\ r\in[0,T]. \tag{4.39}$$

When either $n = m = 1$ or $S(\cdot) = 0$, we have $\mathcal{L}(t,r) = \mathcal{L}(r,t)^\top$, and $\mathcal{H}(r,t) = \mathcal{L}(r \vee t, r \wedge t)$. For given $\widehat{v}(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$, we define $\widehat{\varphi}(\cdot)$ and $K_{\widehat{v}}(\cdot,\cdot)$ as,

$$\begin{aligned}\widehat{\varphi}(t) &:= \int_t^T \mathcal{M}_1(s,t)^\top \varphi(s)\mathrm{d}s + \mathcal{N}_1(t)^\top \varphi(T) + S(t)\varphi(t) + \int_0^t \mathcal{B}(t,r)^\top \big[I - \mathcal{R}^\dagger(r)\mathcal{R}(r)\big]\widehat{v}(r)\mathrm{d}W(r) \\ &\quad + \mathbb{E}_t \int_0^T \mathcal{H}(r,t)\big[I - \mathcal{R}^\dagger(r)\mathcal{R}(r)\big]\widehat{v}(r)\mathrm{d}r + \int_t^T \mathcal{B}(r,t)\big[I - \mathcal{R}^\dagger(r)\mathcal{R}(r)\big]K_{\widehat{v}}(r,t)\mathrm{d}r, \\ \widehat{v}(t) &= \mathbb{E}\widehat{v}(t) + \int_0^t K_{\widehat{v}}(t,s)\mathrm{d}W(s), \qquad t \in [0,T].\end{aligned} \tag{4.40}$$

Then we consider the following equation,

$$\widehat{\lambda}(t) = -\widehat{\varphi}(t) - \int_0^t \mathcal{B}(t,r)^\top \mathcal{R}^\dagger(r)\widehat{\lambda}(r)\mathrm{d}W(r) - \mathbb{E}_t \int_0^T \mathcal{H}(r,t)\mathcal{R}^\dagger(r)\widehat{\lambda}(r)\mathrm{d}r - \int_t^T \mathcal{B}(r,t)\mathcal{R}^\dagger(r)K_{\widehat{\lambda}}(r,t)\mathrm{d}r, \tag{4.41}$$

where $K_{\widehat{\lambda}}(\cdot,\cdot)$ is defined in a similar manner as $K_{\widehat{v}}(\cdot,\cdot)$ in (4.40).

Theorem 4.2. *Suppose* (H2) *holds true with deterministic coefficients. Then the linear quadratic problem is solvable with optimal pair* $(\widehat{X}(\cdot), \widehat{u}(\cdot))$ *if and only if the following hold,*

(i) (4.41) *admits a solution* $\widehat{\lambda}(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ *for some* $\widehat{v}(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$,
(ii) *for any* $u(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$,

$$\mathbb{E}\int_0^T \Big\langle u(t), \mathcal{R}(t)u(t)\,\Big\rangle \mathrm{d}t + 2\mathbb{E}\int_0^T \Big\langle u(t), \int_t^T \big[\mathcal{L}(r,t)u(r) + \mathcal{B}(r,t)K_u(r,t)\big]\mathrm{d}r \Big\rangle \mathrm{d}t \geq 0. \tag{4.42}$$

In this case,

$$\widehat{u}(t) = \mathcal{R}^\dagger(t)\widehat{\lambda}(t) + \big[I - \mathcal{R}^\dagger(t)\mathcal{R}(t)\big]\widehat{v}(t), \qquad t \in [0,T]. \text{ a.e.} \tag{4.43}$$

Proof. As to $\mathcal{G}_1(\cdot,\cdot)$ in (4.14), it follows by some calculations that,

$$\begin{aligned}\mathcal{G}_1(t,\widehat{\mathcal{X}}(\cdot,t)) &= \mathbb{E}_t \int_t^T \mathcal{M}_1(s,t)^\top \widehat{\mathcal{X}}(s,t)\mathrm{d}s + \mathcal{N}_1(t)^\top \widehat{\mathcal{X}}(T,t) + S(t)\widehat{\mathcal{X}}(t,t) \\ &= \mathcal{C}(t) + \int_0^t \mathcal{L}(t,r)^\top \widehat{u}(r)\mathrm{d}r + \int_0^t \mathcal{B}(t,r)^\top \widehat{u}(r)\mathrm{d}W(r), \qquad \text{a.e.}\end{aligned} \tag{4.44}$$

where

$$\mathcal{C}(t) := \int_t^T \mathcal{M}_1(s,t)^\top \varphi(s)\mathrm{d}s + \mathcal{N}_1(t)^\top \varphi(T) + S(t)\varphi(t), \qquad t \in [0,T]. \text{ a.e.}$$

On the other hand, from (4.38) and $\widehat{\xi}(\cdot)$ in (4.15), we can present $\widehat{\xi}(\cdot)$ as,

$$\widehat{\xi}(t) = \mathbb{E}_t \int_t^T \mathcal{B}(r,t)K_{\widehat{u}}(r,t)\mathrm{d}r + \mathbb{E}_t \int_t^T \mathcal{L}(r,t)\widehat{u}(r)\mathrm{d}r, \qquad t \in [0,T]. \text{ a.e.} \tag{4.45}$$

Plugging (4.44) and (4.45) into the third equality in (4.15), for $t \in [0,T]$, a.e., we then obtain

$$\mathcal{R}(t)\widehat{u}(t) + \mathcal{C}(t) + \int_0^t \mathcal{B}(t,r)^\top \widehat{u}(r)\mathrm{d}W(r) + \mathbb{E}_t \int_0^T \mathcal{H}(r,t)\widehat{u}(r)\mathrm{d}r + \int_t^T \mathcal{B}(r,t)K_{\widehat{u}}(r,t)\mathrm{d}r = 0,$$

where $\mathcal{H}(\cdot,\cdot)$ is defined in (4.39). As a result, we have (4.41) and (4.43) by letting $\widehat{\lambda}(\cdot) := \mathcal{R}(\cdot)\widehat{u}(\cdot)$. If we replace $(\widehat{u}(\cdot), x)$ by $(u(\cdot), 0)$, and recall $\chi_0(\cdot),\ X_0(\cdot),\ \theta_0(\cdot)$ in Lemma 2.2, we know in a similar manner that

$$\begin{aligned}&\chi_0(t) + S(t)X_0(t) + R(t)u(t) + B_1(T,t)^\top \mathbb{E}_t\big[GX_0(T)\big] + B_2(T,t)^\top \theta_0(t)\\ &= \mathcal{R}(t)u(t) + \int_0^t \mathcal{B}(t,r)^\top u(r)\mathrm{d}W(r) + \mathbb{E}_t\int_0^T \mathcal{H}(r,t)u(r)\mathrm{d}r + \int_t^T \mathcal{B}(r,t)K_u(r,t)\mathrm{d}r.\end{aligned}$$

Observe that

$$\begin{aligned}&\mathbb{E}\int_0^T \Big\langle u(t), \int_0^t \mathcal{B}(t,r)^\top u(r)\mathrm{d}W(r)\,\Big\rangle\,\mathrm{d}t = \mathbb{E}\int_0^T \Big\langle\, u(t), \int_t^T \mathcal{B}(r,t)K_u(r,t)\mathrm{d}r \Big\rangle\,\mathrm{d}t,\\ &\mathbb{E}\int_0^T \Big\langle u(t), \mathbb{E}_t\int_0^T \mathcal{H}(r,t)u(r)\mathrm{d}r\,\Big\rangle\,\mathrm{d}t = 2\mathbb{E}\int_0^T \Big\langle\, u(t), \int_t^T \mathcal{L}(r,t)u(r)\mathrm{d}r \Big\rangle\,\mathrm{d}t,\end{aligned}$$

the equivalent relation between (4.16) and (4.42) follows obviously. □

It is worthy mentioning that (4.41) is a linear stochastic Fredholm−Volterra integral equation, see also (3.36). In the rest of this part, to avoid the complexity and difficulties in forward-backward system approach, we will give some investigations on (4.41) instead. Given $\mathcal{H}(\cdot,\cdot)$, $\mathcal{B}(\cdot,\cdot)$ in (4.39), (4.38) respectively, define

$$L_1 := \left[\int_0^T\int_0^T |\mathcal{H}(t,s)|^2\mathrm{d}s\mathrm{d}t\right]^{\frac{1}{2}},\quad L_2 := \sup_{t\in[0,T]}\left[\int_t^T |\mathcal{B}(s,t)|^2\mathrm{d}s\right]^{\frac{1}{2}}. \tag{4.46}$$

Theorem 4.3. *Suppose* (H2) *holds true with deterministic coefficients, and*

$$\mathcal{R}(t)^\top\mathcal{R}(t) \geq kL := k[L_1 + 2L_2]^2 I_{[m\times m]},\qquad t\in[0,T],\quad \text{a.e.}\quad k>3. \tag{4.47}$$

Then the linear quadratic problem admits an optimal control represented by (4.43).

Proof. If $L_1 = L_2 = 0$, then $\widehat{\lambda}(\cdot) = -\widehat{\varphi}(\cdot)$ and for any $u(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$,

$$\mathbb{E}\int_0^T \langle\, u(t), \int_t^T \big[\mathcal{L}(r,t)u(r) + \mathcal{B}(r,t)K_u(r,t)\big]\mathrm{d}r\,\rangle\,\mathrm{d}t = 0.$$

Since $\mathcal{R}(\cdot)\geq 0$, inequality (4.42) can be fulfilled, and the conclusion follows from Theorem 4.2. Next suppose that at least one of L_i is positive. In this case $\mathcal{R}^{-1}(\cdot)$ exists, and the equation for $\widehat{\lambda}(\cdot)$ becomes,

$$\widehat{\lambda}(t) = -\widehat{\varphi}(t) - \int_0^t \mathcal{B}(t,r)^\top\mathcal{R}^{-1}(r)\widehat{\lambda}(r)\mathrm{d}W(r) - \mathbb{E}_t\int_0^T \mathcal{H}(r,t)\mathcal{R}^{-1}(r)\widehat{\lambda}(r)\mathrm{d}r - \int_t^T \mathcal{B}(r,t)\mathcal{R}^{-1}(r)K_{\widehat{\lambda}}(r,t)\mathrm{d}r, \tag{4.48}$$

with $\widehat{v}(\cdot) = 0$ a.s. in (4.40). As a result, the conclusion is lead by the following two steps.

Step 1. We discuss the existence and uniqueness of (4.48), the proof of which is based on fixed point arguments. For later convenience, let us define $\kappa(\cdot) := \mathcal{R}^{-1}(\cdot)\widehat{\lambda}(\cdot)$, and rewrite (4.48) as,

$$\mathcal{R}(t)\kappa(t) = -\widehat{\varphi}(t) - \int_0^t \mathcal{B}(t,r)^\top\kappa(r)\mathrm{d}W(r) - \mathbb{E}_t\int_0^T \mathcal{H}(r,t)\kappa(r)\mathrm{d}r - \int_t^T \mathcal{B}(r,t)K_\kappa(r,t)\mathrm{d}r. \tag{4.49}$$

For any $\alpha_i(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R})$, and $\beta_i(\cdot,\cdot)$ determined by $\alpha_i(s) = \mathbb{E}\alpha_i(s) + \int_0^t \beta_i(t,s)\mathrm{d}W(s)$ with $t \in [0,T]$, $(i = 1,2)$, define

$$\mathcal{R}(t)\kappa_i(t) = \widehat{\varphi}(t) - \mathbb{E}_t \int_0^T \mathcal{H}(t,s)\alpha_i(s)\mathrm{d}s - \int_0^t \mathcal{B}(t,s)^\top \alpha_i(s)\mathrm{d}W(s) - \mathbb{E}_t \int_t^T \mathcal{B}(s,t)\beta_i(s,t)\mathrm{d}s,$$

from which one can deduce that,

$$\begin{aligned}\mathcal{R}(t)\big[\kappa_1(t) - \kappa_2(t)\big] = & -\mathbb{E}_t \int_0^T \mathcal{H}(t,s)[\alpha_1(s) - \alpha_2(s)]\mathrm{d}s - \int_0^t \mathcal{B}(t,s)^\top [\alpha_1(s) - \alpha_2(s)]\mathrm{d}W(s) \\ & -\mathbb{E}_t \int_t^T \mathcal{B}(s,t)[\beta_1(s,t) - \beta_2(s,t)]\mathrm{d}s. \end{aligned} \tag{4.50}$$

As to the right hand of (4.50),

$$\begin{cases} \mathbb{E}\displaystyle\int_0^T \left|\mathbb{E}_t \int_0^T \mathcal{L}(t,s)[\alpha_1(s) - \alpha_2(s)]\mathrm{d}s\right|^2 \mathrm{d}t \leq L_1^2 \mathbb{E}\int_0^T |\alpha_1(s) - \alpha_2(s)|^2 \mathrm{d}s, \\ \mathbb{E}\displaystyle\int_0^T \left|\int_0^t \mathcal{B}(t,s)^\top [\alpha_1(s) - \alpha_2(s)]\mathrm{d}W(s)\right|^2 \mathrm{d}t \leq L_2^2 \mathbb{E}\int_0^T |\alpha_1(s) - \alpha_2(s)|^2 \mathrm{d}s, \\ \mathbb{E}\displaystyle\int_0^T \left|\mathbb{E}_t \int_t^T \mathcal{B}(s,t)[\beta_1(s,t) - \beta_2(s,t)]\mathrm{d}s\right|^2 \mathrm{d}t \leq L_2^2 \mathbb{E}\int_0^T |\alpha_1(s) - \alpha_2(s)|^2 \mathrm{d}s. \end{cases} \tag{4.51}$$

Consequently one has,

$$\mathbb{E}\int_0^T \big|\mathcal{R}(t)[\kappa_1(t) - \kappa_2(t)]\big|^2 \mathrm{d}t \leq 3(L_1^2 + 2L_2^2)\mathbb{E}\int_0^T |\alpha_1(t) - \alpha_2(t)|^2 \mathrm{d}t.$$

Eventually by the requirement of $\mathcal{R}(\cdot)$ one can obtain the well-posedness of (4.48).

Step 2. We prove the inequality (4.42) with $u(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$. Recalling (4.39), we can finish this step as long as,

$$\mathbb{E}\int_0^T \langle\, u(t), \mathcal{R}(t)u(t)\,\rangle\, \mathrm{d}t + \mathbb{E}\int_0^T \langle\, u(t), \int_0^T \mathcal{H}(r,t)u(r)\mathrm{d}r + 2\int_t^T B(r,t)K_u(r,t)\mathrm{d}r\,\rangle\, \mathrm{d}t \geq 0. \tag{4.52}$$

To do this let us look at the following results implied by Hölder inequality and Fubini theorem,

$$\begin{cases} \mathbb{E}\displaystyle\int_0^T \left\langle u(t), \int_0^T \mathcal{H}(r,t)u(r)\mathrm{d}r\right\rangle \mathrm{d}t \leq L_1 \mathbb{E}\int_0^T |u(t)|^2 \mathrm{d}t, \\ \mathbb{E}\displaystyle\int_0^T \left\langle u(t), \int_t^T \mathcal{B}(r,t)K_u(r,t)\mathrm{d}r\right\rangle \mathrm{d}t \leq L_2 \mathbb{E}\int_0^T |u(t)|^2 \mathrm{d}t. \end{cases} \tag{4.53}$$

Plugging (4.53) back to (4.52), it follows from the requirement on $\mathcal{R}(\cdot)$ that (4.52) is true. □

To conclude this part, let us make a few points. First of all, if $B_1(\cdot,\cdot) = 0$, then $L_i = 0$ $(i = 1,2)$, and condition (4.47) becomes $\mathcal{R}(t) \geq 0$. As we will prove in Section 4.2, this condition is also necessary. Second of all, it seems that the limitations of treating FBSVIE (4.9) aforementioned do not happen here. Actually, with our method the term G doesn't bring additional troubles in solving the SFVIE (4.41), and neither coefficients nor time duration $[0,T]$ are required to be small enough. Third of all, to ensure condition (4.42), a common assumption is $S(\cdot)=0$ and non-negative G, $Q(\cdot)$, $R(\cdot)$. However, here we provide a more general criteria which allows $S(\cdot)\neq 0$ or G, $Q(\cdot)$, $R(\cdot)$ to be indefinite. The appearance of negative definite $R(\cdot)$ in our stochastic setting is a distinctive feature from that of deterministic VIEs. To see this, let us look at the following example,

Example 4.1. Suppose $m=n=1$, $T>2\pi$, $S(\cdot)=Q(\cdot)=0$, $B_2(t,s)=\sin(t-s)+a$ with constant $a>1$. We define $M:=[\int_0^T|B_1(T,s)|^2\mathrm{d}s]^{\frac{1}{2}}$. Under this framework we have

$$\inf_{t\in[0,T]}B_2(T,t)=a-1,\qquad \sup_{t\in[0,T]}B_2(T,t)=a+1,\qquad \mathcal{L}(s,t)=\mathcal{B}(s,t)=B_1(T,t)GB_1(T,s).$$

Suppose $\mathcal{R}(\cdot)$ satisfies that,

$$R(t)+G|B_2(T,t)|^2\geq 2|G|\left[M^2+2\sup_{t\in[0,T]}\left[|B_2(T,t)|^2\int_t^T|B_1(T,s)|^2\mathrm{d}s\right]^{\frac{1}{2}}\right],\quad t\in[0,T]. \tag{4.54}$$

Hence (4.47) holds true, and the linear quadratic problem admits an optimal control by Theorem 4.3. To keep $R(\cdot)$ not larger than zero, suppose G is non-negative and $M\in(0,\frac{-a-1+\sqrt{2a^2+2}}{2})$, since in this case

$$|B_2(T,t)|^2>G(a-1)^2>4\big[M^2+(a+1)M\big]>2\left[M^2+2\sup_{t\in[0,T]}\left[|B_2(T,t)|^2\int_t^T|B_1(T,s)|^2\mathrm{d}s\right]^{\frac{1}{2}}\right]. \tag{4.55}$$

Moreover, to keep (4.54) true, G is also allowed to be negative with proper requirement on $R(\cdot)$.

4.2. New pointwise necessary conditions of optimal controls

In this part, we give new necessary optimality conditions under SVIEs setting. Before it, we point out two motivations.

If we consider linear quadratic problem associated with SDE (3.1) and cost functional (1.2), one can obtain necessary conditions (3.11) and $\mathcal{R}(\cdot)\geq 0$ with the help of spike variation, or only (3.11) by means of convex variation. In other words, spike variation can provide more information on the coefficients than the later. We return back to SVIEs case with optimal $\widehat{u}(\cdot)$. Using convex variation method (see [27]) one can derive

$$H_u(t,\widehat{X}(T),\widehat{Y}(\cdot),\widehat{Z}(\cdot,t),\widehat{X}(t),\widehat{u}(t))=0,\qquad t\in[0,T],\quad \text{a.s.} \tag{4.56}$$

where $H(\cdot)$ is called the Hamiltonian defined as,

$$\begin{cases} H(t,\widehat{X}(T),\widehat{Y}(\cdot),\widehat{Z}(\cdot,t),X(t),u):=\langle B_1(T,t)^\top\mathbb{E}_t\big[G\widehat{X}(T)\big]+B_2(T,t)^\top\widehat{\theta}(t)+\mathbb{E}_t\displaystyle\int_t^T B_1(s,t)^\top\widehat{Y}(s)\mathrm{d}s \\ \qquad\qquad +\displaystyle\int_t^T B_2(s,t)^\top\widehat{Z}(s,t)\mathrm{d}s,u\rangle+\langle S(t)X(t),u\rangle+\frac{1}{2}\langle R(t)u,u\rangle+\frac{1}{2}\langle Q(t)X(t),X(t)\rangle, \\ \widehat{Y}(t)=Q(t)\widehat{X}(t)+S(t)\widehat{u}(t),\qquad \widehat{Y}(t)=\mathbb{E}\widehat{Y}(t)+\displaystyle\int_0^t\widehat{Z}(t,s)\mathrm{d}W(s),\quad t\in[0,T]. \end{cases} \tag{4.57}$$

Consequently, a natural question is how to deduce more optimality conditions for linear quadratic problem of SVIEs *via* spike variational method.

Our second motivation is about $\mathcal{R}(\cdot)$ in (4.14). From Remark 4.4, its counterpart in the SDEs situation is $B_2(\cdot)^\top\mathcal{P}(\cdot)B_2(\cdot)+R(\cdot)$, the definiteness of which is necessary for the existence of $\widehat{u}(\cdot)$. As a result, whether this necessity still holds true for the involved SVIEs here is an interesting problem to explore.

Motivated by above arguments, we are aim to derive some new necessary conditions in linear quadratic problem of SVIEs. To this end, suppose that

(H3) The conditions in **(H1)** hold true with $A_1(\cdot,\cdot)=A_2(\cdot,\cdot)=0$, $\varphi(\cdot)=x$.

We point out that the assumption of $\varphi(\cdot)=x$ is imposed for simplicity. Suppose $\widehat{u}(\cdot)$ is an optimal control, define

$$u^{\varepsilon}(\cdot):=u\cdot I_{[\tau,\tau+\varepsilon]}(\cdot)+\widehat{u}(\cdot),\qquad \tau\in[0,T],\qquad \varepsilon>0,$$

where u is $\mathcal{F}_\tau$-measurable random variable such that $\mathbb{E}|u|^2<\infty$. We need the following variational equation for $X_1(\cdot)$,

$$X_1(t)=\int_0^t B_1(t,s)\big[u^{\varepsilon}(s)-\widehat{u}(s)\big]\mathrm{d}s+\int_0^t B_2(t,s)\big[u^{\varepsilon}(s)-\widehat{u}(s)\big]\mathrm{d}W(s),\quad t\in[0,T], \tag{4.58}$$

and the following-up equation for $\mathcal{X}_1(\cdot,\cdot)$, that is, for $t\in[0,T]$ and $s\in[0,t]$,

$$\mathcal{X}_1(t,s)=\int_0^s B_1(t,r)\big[u^{\varepsilon}(r)-\widehat{u}(r)\big]\mathrm{d}r+\int_0^s B_2(t,r)\big[u^{\varepsilon}(r)-\widehat{u}(r)\big]\mathrm{d}W(r). \tag{4.59}$$

It is a direct calculation that,

Lemma 4.2. *Given $X_1(\cdot)$ and $\mathcal{X}_1(\cdot,\cdot)$ in (4.58) and (4.59), suppose* (H3) *is imposed. Then*

$$\left[\sup_{t\in[0,T]}\mathbb{E}|X_1(t)|^2+\sup_{t\in[0,T]}\mathbb{E}\sup_{s\in[0,T]}|\mathcal{X}_1(t,s)|^2\right]\le C\varepsilon. \tag{4.60}$$

By means of Lemma 4.2, as well as Taylor expansion, we immediately have,

Lemma 4.3. *Suppose* (H3) *holds true, $(\widehat{X}(\cdot),\widehat{u}(\cdot))$ is an optimal pair. Then*

$$\begin{aligned}o(\varepsilon)\le{}&\mathbb{E}\int_0^T\langle Q(s)\widehat{X}(s)+S(s)\widehat{u}(s),X_1(s)\rangle\mathrm{d}s+\mathbb{E}\langle G\widehat{X}(T),X_1(T)\rangle+\frac12\mathbb{E}\langle GX_1(T),X_1(T)\rangle\\&+\frac12\mathbb{E}\int_0^T\langle Q(s)X_1(s),X_1(s)\rangle\mathrm{d}s+\mathbb{E}\int_0^T\big[\langle S(s)\widehat{X}(s),u^{\varepsilon}(s)-\widehat{u}(s)\rangle\\&+\frac12\langle R(s)u^{\varepsilon}(s),u^{\varepsilon}(s)\rangle-\frac12\langle R(s)\widehat{u}(s),\widehat{u}(s)\rangle\big]\mathrm{d}s.\end{aligned} \tag{4.61}$$

We give the main result of this part,

Theorem 4.4. *Suppose* (H3) *holds true, and that $(\widehat{X}(\cdot),\widehat{u}(\cdot))$ is an optimal pair of the linear quadratic problem. Then condition (4.56) holds and $\mathcal{R}(\cdot)\ge0$, where $\mathcal{R}(\cdot)$ is defined by (4.14).*

Proof. Since $A_i(\cdot,\cdot)=0$, it follows from (2.6) in Section 2 that,

$$\widehat{Y}(t)=\big[Q(s)\widehat{X}(s)+S(s)\widehat{u}(s)\big]-\int_t^T\widehat{Z}(t,s)\mathrm{d}W(s),\qquad t\in[0,T].\ \text{ a.e.}$$

Hence by the definition of $X_1(\cdot)$ in (4.58), as well as Fubini theorem, one has,

$$\begin{aligned}&\mathbb{E}\int_0^T\langle Q(s)\widehat{X}(s)+S(s)\widehat{u}(s),X_1(s)\rangle\mathrm{d}s+\mathbb{E}\langle G\widehat{X}(T),X_1(T)\rangle\\&=\mathbb{E}\int_0^T\langle B_1(T,s)^\top G\widehat{X}(T),\big[u^{\varepsilon}(s)-\widehat{u}(s)\big]\rangle\mathrm{d}s+\mathbb{E}\int_0^T\langle B_2(T,s)^\top\widehat{\theta}(s),\big[u^{\varepsilon}(s)-\widehat{u}(s)\big]\rangle\mathrm{d}s\\&\quad+\mathbb{E}\int_0^T\int_s^T\langle\big[B_1(t,s)^\top\widehat{Y}(t)+B_2(t,s)^\top\widehat{Z}(t,s)\big],\big[u^{\varepsilon}(s)-\widehat{u}(s)\big]\rangle\mathrm{d}t\mathrm{d}s.\end{aligned} \tag{4.62}$$

Given $P_Q(\cdot,\cdot)$, $P_G(\cdot)$ in (4.4), $\mathcal{M}_1(\cdot,\cdot)$, $\mathcal{N}_1(\cdot)$ in (4.6), we use Itô formula to $\mathcal{X}_1(t,\cdot)^\top P_Q(t,\cdot)\mathcal{X}_1(t,\cdot)$ on $[0,t]$. Recalling $u^\varepsilon(s)-\widehat{u}(s)=uI_{[\tau,\tau+\varepsilon]}(s)$ with $s\in[0,T]$, we have

$$\begin{aligned}\mathcal{X}_1(t,r)^\top P_Q(t,r)\mathcal{X}_1(t,r)=&\int_0^r\Big[2u^\top I_{[\tau,\tau+\varepsilon]}(s)\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)+u^\top B_2(t,s)^\top P_Q(t,s)B_2(t,s)uI_{[\tau,\tau+\varepsilon]}(s)\Big]\mathrm{d}s\\&+\int_0^r\Big[u^\top I_{[\tau,\tau+\varepsilon]}(s)B_2(t,s)^\top P_Q(t,s)\mathcal{X}_1(t,s)+\mathcal{X}_1(t,s)^\top\varLambda_Q(t,s)\mathcal{X}_1(t,s)\Big]\mathrm{d}W(s)\end{aligned}\tag{4.63}$$

For given $t\in[0,T]$ and $n\geq 0$, we define a sequence of stopping time $\delta_{n,t}$ as

$$\delta_{n,t}:=\inf\Big\{r\geq 0,\ \int_0^r\Big[u^\top I_{[\tau,\tau+\varepsilon]}(s)B_2(t,s)^\top P_Q(t,s)\mathcal{X}_1(t,s)+\mathcal{X}_1(t,s)^\top\varLambda_Q(t,s)\mathcal{X}_1(t,s)\Big]^2\mathrm{d}s\geq n\Big\}\wedge t,$$

from which $\delta_{n,t}\to t$. a.s. As a result, by taking expectation on both sides of (4.63) we know that

$$\mathbb{E}\mathcal{X}_1(t,\delta_{n,t})^\top P_Q(t,\delta_{n,t})\mathcal{X}_1(t,\delta_{n,t})=\mathbb{E}\int_0^{\delta_{n,t}}\Big[2u^\top\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)+u^\top B_2(t,s)^\top P_Q(t,s)B_2(t,s)u\Big]I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s.\tag{4.64}$$

Thanks to $(P_Q(\cdot,\cdot),\varLambda_Q(\cdot,\cdot))$, $\mathcal{M}_1(\cdot,\cdot)$ in (4.4), (4.6) respectively, we can derive that

$$\mathbb{E}_s\int_s^t|\varLambda_Q(t,r)|^2\mathrm{d}r+|P_Q(t,s)|^2=\mathbb{E}_s|Q(t)|^2,\qquad \mathbb{E}_s\int_s^t|\mathcal{M}_1(t,r)|^2\mathrm{d}r\leq K\mathbb{E}_s|Q(t)|^2,\quad\text{a.s.}\tag{4.65}$$

where K may depend on T and the upper bound of $B_i(\cdot,\cdot)$. Using estimate (4.60) in Lemma 4.2, the $\mathcal{F}_\tau$-measurability of u, and the second result in (4.65),

$$\begin{aligned}&\mathbb{E}\int_0^t|u^\top\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)I_{[\tau,\tau+\varepsilon]}(s)|\mathrm{d}s\\&\leq\mathbb{E}\sup_{s\in[0,T]}|\mathcal{X}_1(t,s)|^2+\mathbb{E}\int_0^t|u^T\mathcal{M}_1(t,s)^\top|^2I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s\\&=\mathbb{E}\sup_{s\in[0,T]}|\mathcal{X}_1(t,s)|^2+K\mathbb{E}\Big[|u^\top|^2\mathbb{E}_\tau\int_\tau^t|\mathcal{M}_1(t,s)^\top|^2\mathrm{d}sI_{[\tau,T]}(t)\Big]\\&\leq\mathbb{E}\sup_{s\in[0,T]}|\mathcal{X}_1(t,s)|^2+K\mathbb{E}\Big[|u^\top|^2\mathbb{E}_\tau|Q(t)|^2I_{[\tau,T]}(t)\Big]<\infty.\end{aligned}$$

Consequently the following result is derived by applying dominated convergence theorem to (4.64),

$$\mathbb{E}\big[X_1(t)^\top Q(t)X_1(t)\big]=\mathbb{E}\int_0^t\Big[2u^T\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)+u^\top B_2(t,s)^\top P_Q(t,s)B_2(t,s)u\Big]I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s.\tag{4.66}$$

Similarly one can obtain

$$\mathbb{E}\big[X_1(T)^\top GX_1(T)\big]=\mathbb{E}\int_0^T\Big[2u^\top\mathcal{N}_1(s)^\top\mathcal{X}_1(T,s)+u^\top B_2(T,s)^\top P_G(s)B_2(T,s)u\Big]I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s.\tag{4.67}$$

Combining (4.66) and (4.67) together, and recalling $\mathcal{R}(\cdot)$ in (4.14), we deduce that

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\big[X_1(T)^\top GX_1(T)\big]+\frac{1}{2}\mathbb{E}\int_0^TX_1(s)^\top Q(s)X_1(s)\mathrm{d}s\\&\quad=\mathbb{E}\int_0^Tu^\top I_{[\tau,\tau+\varepsilon]}(s)\int_s^T\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)\mathrm{d}t\mathrm{d}s+\mathbb{E}\int_0^Tu^\top I_{[\tau,\tau+\varepsilon]}(s)\mathcal{N}_1(s)^\top\mathcal{X}_1(T,s)\mathrm{d}s\\&\qquad+\frac{1}{2}\mathbb{E}\int_0^Tu^\top\big[\mathcal{R}(s)-R(s)\big]uI_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s.\end{aligned}\tag{4.68}$$

As to the first term on the right hand of (4.68), we claim that,

$$\mathbb{E}\int_0^T u^\top I_{[\tau,\tau+\varepsilon]}(s)\int_s^T \mathcal{M}_1(t,s)^\top \mathcal{X}_1(t,s)\mathrm{d}t\mathrm{d}s = \mathbb{E}\int_0^T\int_0^t u^\top I_{[\tau,\tau+\varepsilon]}(s)\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)\mathrm{d}s\mathrm{d}t = o(\varepsilon). \quad (4.69)$$

Actually, as to the term in (4.69), one has,

$$\begin{aligned}&\mathbb{E}\int_0^T\int_0^t u^\top I_{[\tau,\tau+\varepsilon]}(s)\mathcal{M}_1(t,s)^\top\mathcal{X}_1(t,s)\mathrm{d}s\mathrm{d}t\\ &\leq \varepsilon^{\frac{1}{2}}\left[\mathbb{E}\int_0^T \sup_{s\in[0,T]}|\mathcal{X}_1(t,s)|^2\mathrm{d}t\right]^{\frac{1}{2}}\left[\mathbb{E}\int_0^T |u|^2\int_0^t|\mathcal{M}_1(t,s)^\top|^2 I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s\mathrm{d}t\right]^{\frac{1}{2}}. \end{aligned} \quad (4.70)$$

We observe that the following equalities hold true with the help of estimates in (4.65) and dominated convergence theorem,

$$\lim_{\varepsilon\to 0}\mathbb{E}\int_0^T|u|^2\int_0^t|\mathcal{M}_1(t,s)^\top|^2 I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s\mathrm{d}t = \mathbb{E}\int_\tau^T|u|^2\left[\lim_{\varepsilon\to 0}\mathbb{E}_\tau\int_\tau^t|\mathcal{M}_1(t,s)^\top|^2 I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s\right]\mathrm{d}t = 0. \quad (4.71)$$

Hence estimate (4.69) then follows from (4.60), (4.70) and (4.71). Similarly one can obtain that,

$$\mathbb{E}\int_0^T u^\top I_{[\tau,\tau+\varepsilon]}(s)\mathcal{N}_1(s)^\top\mathcal{X}_1(T,s)\mathrm{d}s = o(\varepsilon). \quad (4.72)$$

Consequently due to (4.69) and (4.72),

$$\frac{1}{2}\mathbb{E}\big[X_1(T)^\top G X_1(T)\big] + \frac{1}{2}\mathbb{E}\int_0^T X_1(s)^\top Q(s)X_1(s)\mathrm{d}s = \frac{1}{2}\mathbb{E}\int_0^T u^\top\big[\mathcal{R}(s)-R(s)\big]uI_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s + o(\varepsilon). \quad (4.73)$$

As a result, we have the following results by (4.62), (4.73) and Lemma 4.3,

$$\begin{aligned}o(\varepsilon) \leq\ &\mathbb{E}\int_0^T u^\top B_1(T,s)^\top G\widehat{X}(T)I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s + \mathbb{E}\int_0^T u^\top B_2(T,s)^\top\widehat{\theta}(s)I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s\\ &+\mathbb{E}\int_0^T\int_s^T u^\top\Big[B_1(t,s)^\top\widehat{Y}(t) + B_2(t,s)^\top\widehat{Z}(t,s)\Big]I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}t\mathrm{d}s\\ &+\mathbb{E}\int_0^T\Big[u^\top S(s)\widehat{X}(s)I_{[\tau,\tau+\varepsilon]}(s) + \frac{1}{2}\big[u^\varepsilon(s)\big]^\top R(s)u^\varepsilon(s) - \frac{1}{2}\big[\widehat{u}(s)\big]^\top R(s)\widehat{u}(s)\Big]\mathrm{d}s\\ &+\frac{1}{2}\mathbb{E}\int_0^T u^\top\big[\mathcal{R}(s)-R(s)\big]uI_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s\\ =\ &\mathbb{E}\int_0^T\big[H(s,\widehat{X}(T),\widehat{Y}(\cdot),\widehat{Z}(\cdot,s),\widehat{X}(s),u^\varepsilon(s)) - H(s,\widehat{X}(T),\widehat{Y}(\cdot),\widehat{Z}(\cdot,s),\widehat{X}(s),\widehat{u}(s))\big]\mathrm{d}s\\ &+\frac{1}{2}\mathbb{E}\int_0^T u^\top\big[\mathcal{R}(s)-R(s)\big]uI_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}s.\end{aligned}$$

By the definition of $u^\varepsilon(\cdot)$ we eventually obtain that,

$$\begin{aligned}0\leq\ & u^\top\left[B_1(T,\tau)^\top\mathbb{E}_\tau[G\widehat{X}(T)] + B_2(T,\tau)^\top\widehat{\theta}(\tau) + \mathbb{E}_\tau\int_\tau^T\Big[B_1(t,\tau)^\top\widehat{Y}(t) + B_2(t,\tau)^\top\widehat{Z}(t,\tau)\Big]\mathrm{d}t\right]\\ &+u^\top S(\tau)\widehat{X}(\tau) + \frac{1}{2}[u+\widehat{u}(\tau)]^\top R(\tau)[u+\widehat{u}(\tau)] - \frac{1}{2}\big[\widehat{u}(\tau)\big]^\top R(\tau)\widehat{u}(\tau) + \frac{1}{2}u^\top\big[\mathcal{R}(\tau)-R(\tau)\big]u,\end{aligned}$$

which leads to the desired conclusion. □

To conclude this part, we make some observations to the previous arguments. In contrast with the classical convex variation method in [27], there is one additional necessary condition $\mathcal{R}(\cdot) \geq 0$ appearing here which shows the powerfulness of spike variation tricks. Moreover, our result also covers the corresponding SDEs case. Second of all, the previous procedures rely on the introducing of $\mathcal{X}_1(\cdot,\cdot)$ and the use of Itô formula to $\mathcal{X}_1(\cdot,\cdot)^\top P_Q(\cdot,\cdot)\mathcal{X}_1(\cdot,\cdot)$, $\mathcal{X}_1(T,\cdot)^\top P_G(\cdot)\mathcal{X}_1(T,\cdot)$, the tricks of which are new in the literature to our best. Third of all, the method above can be extended to nonlinear cases with non-convex control domain. However, the corresponding diffusion and drift term are required to be state independent. The more general case are still under consideration and we will show more related results in our future publications.

References

[1] E. Alòs and D. Nualart, Anticipating stochastic Volterra equations. *Stochastic Process. Appl.* **72** (1997) 73–95.

[2] S. Belbas and W. Schmidt, Optimal control of Volterra equations with impulses. *Appl. Math. Comput.* **166** (2005) 696–723.

[3] M. Berger and V. Mizel, Volterra equations with Itô integrals, I, II, *J. Integral Equ.* **2** (1980) 187–245, 319–337.

[4] M. Berger and V. Mizel, An extension of the stochastic integral. *Ann. Probab.* **10** (1982) 435–450.

[5] D. Carlson, An elementary proof of the maximum principle for optimal control problems governed by a Volterra integral equation. *J. Optim. Theory Appl.* **54** (1987) 43–61.

[6] S. Chen, Riccati equations arising in infinite dimensional optimal control problems. *Control Theory Appl.* **2** (1985) 64–72 (in Chinese).

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

[8] S. Chen and J. Yong, A linear quadratic optimal control problems for stochastic Volterra integral equations, in Control Theory and Related Topics: In Memory of Prof. Xunjing Li, edited by S. Tang and J. Yong. Word Scientific Publishing Company (2007) 44–66.

[9] D. Duffie and C. Huang, Stochastic production-exchange equilibria. Vol. 974. Research paper, Graduate School of Business, Stanford University, Stanford (1986).

[10] R. Hartl, Optimal dynamic advertising policies for hereditary processes. *J. Optim. Theory Appl.* **43** (1984) 51–72.

[11] Y. Hu and S. Peng, Solution of forward-backward stochastic differential equations. *Probab. Theory Related Fields* **103** (1995) 273–283.

[12] M. Kamien and E. Muller, Optimal control with integral state equations. *Rev. Econ. Stud.* **43** (1976) 469–473.

[13] X. Li and J. Yong, Optimal Control Theory for Infinite Dimensional Systems. Birkhaüser (1995).

[14] J. Lin, Adapted solution of backward stochastic nonlinear Volterra integral equation. *Stoch. Anal. Appl.* **20** (2002) 165–183.

[15] J. Ma, P. Protter and J. Yong, Solving forward-backward stochastic differential equations explicitly-a four step scheme. *Probab. Theory Related Fields* **98** (1994) 339–359.

[16] L. Mou and J. Yong, Two-person zero-sum linear quadratic stochastic differential games by a Hilbert space method. *J. Ind. Manag. Optim.* **2** (2006) 93–115.

[17] E. Pardoux and P. Protter, Stochastic Volterra equations with anticipating coefficients. *Ann. Probab.* **18** (1990) 1635–1655.

[18] A.J. Pritchard and Y. You, Causal feedback optimal control for Volterra integral equations. *SIAM J. Control Optim.* **34** (1996) 1874–1890.

[19] P. Protter, Volterra equations driven by semimartingales. *Ann. Probab.* **13** (1985) 519–530.

[20] Y. Shi, T. Wang and J. Yong, Mean field backward stochastic Volterra integral equations. *Discrete. Cont. Dyn. Ser-B.* **18** (2013) 1929–1967.

[21] Y. Shi, T. Wang and J. Yong, Optimal control problems of forward-backward stochastic Volterra integral equations. *Math. Control Relat. Fields* **5** (2015) 613–649.

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

[23] V. Vinokurov, Optimal control of processes described by integral equations, I, II, III, Izv. Vysš. Učebn. Zaved. *Matematika* **7** 21–33; **8** 16–23; **9** 16–25. (in Russian) English transl. in *SIAM J. Control* **7** (1967) 324–336, 337–345, 346–355.

[24] T. Wang and Y. Shi, Symmetrical solutions of backward stochastic Volterra integral equations and applications. *Discrete Contin. Dyn. Syst. B.* **14** (2010) 251–274.

[25] T. Wang and Y. Shi, Linear quadratic stochastic integral games and related topics. *Sci China Math.* **58** (2015) 2405–2420.

[26] J. Yong, Backward stochastic Volterra integral equations and some related problems. *Stochastic Process Appl.* **116** (2006) 779–795.

[27] J. Yong, Well-posedness and regularity of backward stochastic Volterra integral equation. *Probab. Theory Related Fields* **142** (2008) 21–77.

[28] J. Yong and X. Zhou, Stochastic Control: Hamiltonian Systems and HJB Equations. Springer Verlag, New York (1999).

[29] X. Zhang, Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *J. Funct. Anal.* **258** (2010) 1361–1425.