ESAIM: COCV 26 (2020) 16
https://doi.org/10.1051/cocv/2020001

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

NECESSARY CONDITIONS OF PONTRAYGIN'S TYPE FOR GENERAL CONTROLLED STOCHASTIC VOLTERRA INTEGRAL EQUATIONS*

Tianxiao Wang**

Abstract. This article is addressed to giving a solution to a unsolved problem, *i.e.*, to establish the necessary optimality conditions of Pontraygin's type for controlled stochastic Volterra integral equations (SVIEs) when the control region is non-convex and the control variable enters into the diffusion. This problem has been open since [J. Yong, *Stochastic Process Appl.* **116** (2006) 779–795] obtained the analogue result for the case of convex control region. The key is to introduce *a pair of* suitable second-order adjoint processes (SOAPs). It is found that the usual way of using *only one* SOAP in the maximum condition for the classical setting of controlled stochastic differential equations does not work here.

Mathematics Subject Classification. 93E20, 60H20, 49K45.

Received June 5, 2019. Accepted December 23, 2019.

1. Introduction

Suppose $T > 0$ and $n, m \in \mathbb{N}$ are given, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a complete probability space, $W(\cdot)$ is a one-dimensional Wiener process which generates the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0 \le t \le T}$. In this paper, we shall study an optimal control problem for stochastic Volterra integral equations (SVIEs), where the state equation is

$$X(t) = \varphi(t) + \int_0^t b(t,s,X(s),u(s))\mathrm{d}s + \int_0^t \sigma(t,s,X(s),u(s))\mathrm{d}W(s), \quad t \in [0,T], \tag{1.1}$$

and the cost functional is

$$J(u(\cdot)) = \mathbb{E}\Big[h(X(T)) + \int_0^T l(s,X(s),u(s))\mathrm{d}s\Big]. \tag{1.2}$$

*This work is supported in part by NSF of China (Grant 11401404, 11971332, 11931011) and the Fundamental Research Funds for the central Universities (YJ201605).

Keywords and phrases: Stochastic Volterra integral equations, maximum principles, second-order adjoint processes, non-convex control region.

School of Mathematics, Sichuan University, Chengdu 610065, PR China.

** Corresponding author: xiaotian2008001@gmail.com

Here $u(\cdot)$ is a *control process* taking values in $U \subset \mathbb{R}^m$, and $X(\cdot)$ is the corresponding *state process* in $\mathbb{R}^n$, $\varphi : [0,T] \times \Omega \to \mathbb{R}^n$, $b, \sigma : [0,T]^2 \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $h : \Omega \times \mathbb{R}^n \to \mathbb{R}$ and $l : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are given functions (satisfying suitable conditions to be given later). Under certain conditions (Protter [12]), for any $u(\cdot) \in \mathcal{U}_{ad}$ (to be defined later), (1.1) admits a unique solution $X(\cdot)$ such that $J(u(\cdot))$ is well-defined. The optimal control problem considered next is stated as follows:

Problem (SVIEs): Find $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ such that

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}_{ad}} J(u(\cdot)).$$

In the above, we call $\bar{u}$ an *optimal control*, $\bar{X}$ an *optimal state process*, and $(\bar{X}, \bar{u})$ an *optimal pair*.

Stochastic optimal control problems for SVIEs have many applications. Here we present a stochastic epidemic prevention model as an example. Suppose there is a population suffered from one infected disease during $[0,T]$, where T is the time when there is no infective individual, and 0 is the time when the infective group are separated and received medical treatment. At $t \in [0,T]$, let $X(t)$ be the population of infected people (including the ones who received vaccines before recovering), $u(t)$ the amount of vaccines provided by the local government. Suppose any infected individual is likely to die at a future time. Hence we define a random variable ξ_1 as his/her life length with density function $m_1(\cdot)$. For $s \in [0,t]$, the infective population at $t-s$ is $X(t-s)$. After s (*i.e.* at t) the dying population becomes $X(t-s)m_1(s)\mathrm{d}s$. Then, the total number between $[0,t]$ is

$$\int_0^t X(t-s)m_1(s)\mathrm{d}s = \int_0^t X(s)m_1(t-s)\mathrm{d}s.$$

We shall observe that the infected people in different stage (or time period) respond to the vaccine in distinctive ways. Therefore, it is reasonable to introduce the efficiency index $a(\cdot)$ depending on time. As a result, at t there are $a(t)u(t)$ amount of people whose scenarios become stable. For this group, we define a random variable ξ_2 as the duration of recovering completely with density $m_2(\cdot)$. At $t-s$, the amount of vaccines is $u(t-s)$, and the population with stable physical condition is $a(t-s)u(t-s)$. After s, there are $a(t-s)u(t-s)m_2(s)\mathrm{d}s$ amount of people healing from the disease. Thus the total number between $[0,t]$ is

$$\int_0^t a(t-s)u(t-s)m_2(s)\mathrm{d}s = \int_0^t a(s)u(s)m_2(t-s)\mathrm{d}s.$$

To sum up, the increment of infected individuals at time t is

$$\Delta X(t) = \Big[-\int_0^t m_1(t-s)X(s)\mathrm{d}s - \int_0^t m_2(t-s)a(s)u(s)\mathrm{d}s \Big]\Delta t.$$

Observe that the efficiency coefficient $a(\cdot)$ can be easily influenced by other random factors, like the individual's physical quality, the epidemic situation in global group, the improvement of the vaccine, etc. Consequently, we may replace $a(t)$ with $a(t) + \dot{W}(t)$, where $\dot{W}(\cdot)$ represents the white noise. Then for $t \in [0,T]$, $F_i(r) := \mathbb{P}\big\{\xi_i \le r\big\}$,

$$X(t) = x_0 - \int_0^t \big[F_1(t-r)X(r) + F_2(t-r)a(r)u(r)\big]\mathrm{d}r - \int_0^t F_2(t-r)u(r)\mathrm{d}W(r). \tag{1.3}$$

We emphasize that the introducing of functions F_1 and F_2 is inspired by Arrow [2]. Suppose the local government wants to find a suitable $\bar{u}(\cdot)$ to minimize

$$J(u(\cdot)) := \mathbb{E}\int_0^T \big(G_1(X(s)) + G_2(u(s))\big)\mathrm{d}s, \tag{1.4}$$

where $G_1(\cdot)$ is the daily cost of living for infected people, $G_2(\cdot)$ represents the research and development cost of vaccines. Hence we arrive at an optimal control problem for (1.3) with the cost functional (1.4).

Pontryagin's maximum principle for (deterministic) controlled ordinary differential equations (ODEs) was established in Boltyanski *et al.* [5]. Since then, maximum principles of optimal controls have been developed extensively. Later, people attempted to extend them into the stochastic (diffusion) case, *e.g.*, Bensoussan [3], Bismut [4], Kushner [11], and so on. When the control region is non-convex and the state equation is an Itô's stochastic differential equations (SDEs)

$$X(t)= x_0+\int_0^t b(s,X(s),u(s))\mathrm{d}s+\int_0^t \sigma(s,X(s),u(s))\mathrm{d}W(s), \quad t\in[0,T], \tag{1.5}$$

the corresponding maximum principle was established by Peng [13]. We refer to the monograph of Yong-Zhou [21] for more details. On the other hand, in the real world, there are many models that cannot be simply described by ODEs but by (deterministic) Volterra integral equations (VIEs), such as the optimal capital policy problems in ([2], Sect. 2, pp. 23–25). For the optimal control theory for VIEs, Friedman [7] was one of the earliest literature along this line. Later in 1969, Vinokurov [15] investigated the case with $\sigma = 0$ in (1.1) and constrained state processes. Recent works in the deterministic framework can also be found in Lin-Yong [9], and so on.

Clearly, optimal control problems for both SDEs and VIEs are special cases of our Problem (SVIEs). In this paper, we will establish a Pontryagin-type maximum principle for Problem (SVIEs) *via* the technique of spike variation when the control region U is non-convex and the control variable $u(\cdot)$ enters into the diffusion term. Such a problem has been *open* since Yong [19] obtained the corresponding result for the particular convex control region case in 2006.

In contrast with the existing literatures, there exist some essential obstacles to handle the general case. In fact, when the control region is convex, the corresponding maximum principles were established in Yong [19, 20] and Shi *et al.* [14]. For the case that U is open, Agram-Øksendal [1] presented some investigations by means of Malliavin calculus. Recently, Wang-Zhang [18] obtained some necessary optimality conditions for the case of closed control region by set-valued analysis. In this paper, the control region is allowed to be an arbitrary non-empty subset of $\mathbb{R}^m$. By the spike variation with a parameter $\varepsilon > 0$, we obtain the following *quadratic functional* $\mathscr{E}(\varepsilon)$ with respect to the solution X_1 of first-order variational equation (*i.e.* a linear SVIE),

$$\mathscr{E}(\varepsilon) := \frac{1}{2\varepsilon}\mathbb{E}\Big\{\int_0^T \mathrm{tr}\big[\bar{H}_{xx}(t)X_1(t)X_1(t)^\top\big]\mathrm{d}t + \mathrm{tr}\big[\bar{h}_{xx}(T)X_1(T)X_1(T)^\top\big]\Big\}, \tag{1.6}$$

where $\bar{H}_{xx}(\cdot)$, $\bar{h}_{xx}(T)$ are some functions (to be defined later). To the best of our knowledge, the above $\mathscr{E}(\varepsilon)$ appears in the literature for the first time. The duality ideas in [1, 14, 18–20] do not work here. On the other hand, $\mathscr{E}(\varepsilon)$ is also different from that of SDEs [13], where X_1 satisfies a linear SDE. Actually, to study the corresponding $\mathscr{E}(\varepsilon)$, Peng [13] first derived a linear SDE of $X_1(\cdot)X_1(\cdot)^\top$ *via* Itô's formula, introduced a suitable linear bounded operator, and obtained the analogue *second-order adjoint process* (SOAP), denoted by (P_2,Λ_2), by Riesz's representation theorem. Using again Itô's formula, Peng proposed a *second-order adjoint equation* (SOAE), *i.e.*, a linear backward stochastic differential equation (BSDE) for (P_2,Λ_2) (see Sect. 3.1 for the involved notations below):

$$\begin{cases} dP_2(t) = -\big[\bar{b}_x(t)^\top P_2(t) + P_2(t)\bar{b}_x(t) + \bar{\sigma}_x(t)^\top \Lambda_2(t) + \Lambda_2(t)\bar{\sigma}_x(t) \\ \qquad\qquad + \bar{H}_{xx}(t) + \bar{\sigma}_x(t)^\top P_2(t)\bar{\sigma}_x(t)\big]\mathrm{d}t + \Lambda_2(t)\mathrm{d}W(t), \quad t\in[0,T], \\ P_2(T) = \bar{h}_{xx}(T). \end{cases} \tag{1.7}$$

Notice that the aforementioned arguments *heavily* depend on Itô formula. However, due to the general *non-differential* structure of SVIEs, we cannot directly follow the above approach. For example, in our case it is impossible to obtain a linear SVIE for $X_1(\cdot)X_1(\cdot)^\top$ by the Itô formula, not to mention mimicing the following-up procedures in the SDE setting.

According to the previous analysis, we need to introduce new idea to solve Problem (SVIEs). To this end, let us carefully revisit the pioneer work [13]. According to the arguments in ([13], pp. 971–974), the pointwise maximum condition of optimal controls can be established *merely* by a pair of first-order adjoint processes (FOAP), and the first part of SOAP, *i.e.*, $P_2(\cdot)$. Similar phenomena has also been observed and used essentially to handle optimal control problems for stochastic evolution equations (*e.g.,Du-Meng-2013, Fuhrman-et-al-2013, Lv-Zhang-2015*). This reminds us to think about such a problem: *Can we directly introduce the analogue of FOAP and SOAP (without SOAE) for Problem (SVIEs)?*

We will give a positive answer to the above problem in this article. To obtain the desired SOAP, we observe that the classical approach in [13, 21] depends on Itô's formula and therefore, it does not work here. To see this, let us first return to the SDE case. For any $\tau\in[0,T]$, $\xi_i\in L^2_{\mathcal{F}_\tau}(\Omega;\mathbb{R}^n)$, $i=1,2$, $\bar{b}_x$, $\bar{\sigma}_x$, $\bar{H}_{xx}$, $\bar{h}_{xx}$ in Section 3.1, we define

$$\begin{cases} Y_i(t) = \xi_i + \displaystyle\int_\tau^t \bar{b}_x(s)Y_i(s)\mathrm{d}s + \int_\tau^t \bar{\sigma}_x(s)Y_i(s)\mathrm{d}W(s), \quad i=1,2, \\ J(\tau,\xi_1,\xi_2) := \mathbb{E}_\tau \displaystyle\int_\tau^T Y_1(s)^\top \bar{H}_{xx}(s)Y_2(s)\mathrm{d}s + \mathbb{E}_\tau\big[Y_1(T)^\top \bar{h}_{xx}(T)Y_2(T)\big]. \end{cases} \tag{1.8}$$

One can prove that there exists a measurable, continuous, $\mathbb{F}$-adapted, $\mathbb{R}^{n\times n}$-valued process $\mathscr{P}(\cdot)$ such that $J(\tau,\xi_1,\xi_2)=\xi_1^\top\mathscr{P}(\tau)\xi_2$. On the other hand, applying Itô's formula to $Y_1^\top P_2 Y_2$ on $[\tau,T]$, we see that $\xi_1^\top P_2(\tau)\xi_2 = J(\tau,\xi_1,\xi_2)$. The arbitrariness of ξ_i and the continuity of P_2 and $\mathscr{P}$ lead to

$$\mathbb{P}\big\{\mathscr{P}(\tau) = P_2(\tau),\ \forall \tau\in[0,T]\big\} = 1. \tag{1.9}$$

Unlike [13, 21], for our Problem (SVIE) we need to avoid the use of Itô's formula and the equation for $X_1(\cdot)X_1(\cdot)^\top$. We shall develop a new approach to derive the corresponding SOAP to deal with the above $\mathscr{E}(\varepsilon)$ in (1.6).

The rest of this paper is organized as follows. In Section 2, some notations and assumptions will be given. Section 3 includes four parts. The first part is to illustrate how to obtain (1.6), as well as its relation with maximum principles. The second part and third part are devoted to treating (1.6). In the fourth part, the desired maximum principle will be established and several special cases will be discussed. Section 4, some concluding remarks will be presented. Finally, a few key lemmas will be given in the Appendix.

2. Preliminaries

For $H := \mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n\times m}$, etc., $0 \le s < t \le T$, $p > 1$, we define

$$
\begin{aligned}
&L^p_{\mathcal{F}_t}(\Omega;H) := \Big\{ X : \Omega \to H \ \Big|\ X \text{ is } \mathcal{F}_t\text{-measurable, } \mathbb{E}|X|^p < \infty \Big\},\\
&C_{\mathbb{F}}\big([s,t];L^p(\Omega;H)\big) := \Big\{ X : [s,t]\times\Omega \to H \ \Big|\ X(\cdot) \text{ is continuous from } [s,t] \text{ to}\\
&\qquad\qquad L^2(\Omega;H), \text{and measurable, } \mathbb{F}\text{-adpated, } \sup_{r\in[s,t]} \mathbb{E}|X(r)|^p < \infty \Big\},\\
&L^p_{\mathcal{F}_T}(s,t;H) := \Big\{ X : [s,t]\times\Omega \to H \Big| X(\cdot) \text{ is measurable, } \mathbb{E}\int_s^t |X(r)|^p \mathrm{d}r < \infty \Big\}\\
&L^2(s,t;L^2_{\mathbb{F}}(s,t;H)) := \Big\{ Z : [s,t]^2\times\Omega \to H \,\Big|\, Z(\cdot,\cdot) \text{ is measurable, } Z(u,\cdot) \text{ is } \mathbb{F}\text{-adapted}\\
&\qquad\qquad u\in[s,t],\ \|Z(\cdot\,,\cdot)\|^2_{L^2_{\mathbb{F}}(s,t;L^2(s,t;H))} \equiv \mathbb{E}\int_s^t\int_s^t |Z(u,v)|^2 \mathrm{d}v\mathrm{d}u < \infty \Big\}.
\end{aligned}
$$

Let $L^p_{\mathbb{F}}(s,t;H)$ be the set of $\mathbb{F}$-adapted process in $L^p_{\mathcal{F}_T}(s,t;H)$, and $L^\infty(s,t;L^\infty_{\mathbb{F}}(s,t;H))$ be the set of bounded process in $L^2(s,t;L^2_{\mathbb{F}}(s,t;H))$. We introduce separable Banach space $\mathbb{B} := \mathbb{R}^n \times L^4(0,T;\mathbb{R}^n)$. Obviously there exists numerable dense subset $\mathbb{B}_0$ and

$$
\mathbb{B}_1 := \Big\{ \sum_{i=1}^n k_i\alpha^i;\ k_i\in\mathbb{Q},\ \alpha^i\in\mathbb{B}_0,\ n\in\mathbb{N} \Big\} \subset \mathbb{B},
$$

where $\mathbb{Q}$, $\mathbb{N}$ is respectively the set of rational number, integer number in $\mathbb{R}$. We denote $\mathbb{B}' := \mathbb{R}^n \times L^{\frac{4}{3}}(0,T;\mathbb{R}^n)$, $\mathcal{L}(H';H'')$ the space of linear bounded operators between Banach spaces H' and H'',

$$
\begin{aligned}
&\mathscr{L}_1 := \Big\{ \mathcal{A} : [0,T]\times\Omega \mapsto \mathcal{L}(\mathbb{B};L^{\frac{4}{3}}(0,T;\mathbb{R}^n)) \ \Big|\ \text{for any } b\in\mathbb{B},\ a(\cdot)\in L^4(0,T;\mathbb{R}^n),\\
&\qquad \int_0^T a(s)^\top\big[\mathcal{A}(\cdot,\cdot)b\big](s)\mathrm{d}s \in L^2_{\mathbb{F}}(0,T;\mathbb{R}),\ \sup_{t\in[0,T]}\mathbb{E}\|\mathcal{A}(t)\|^2_{\mathcal{L}(\mathbb{B};L^{\frac{4}{3}}(0,T;\mathbb{R}^n))} < \infty \Big\},\\
&\mathscr{L}_2 := \Big\{ \mathcal{A} : [0,T]\times\Omega \mapsto \mathcal{L}(\mathbb{B};\mathbb{R}^n) \ \Big|\ \sup_{t\in[0,T]}\mathbb{E}\|\mathcal{A}(t)\|^2_{\mathcal{L}(\mathbb{B};\mathbb{R}^n)} < \infty, \text{ and for any}\\
&\qquad a(\cdot)\in\mathbb{R}^n,\ b\in\mathbb{B},\ a^\top\big[\mathcal{A}(\cdot,\cdot)b\big] \in L^2_{\mathbb{F}}(0,T;\mathbb{R}) \Big\}.
\end{aligned}
\tag{2.1}
$$

In this paper, K is a generic constant which varies in different context. For FSVIE (1.1), we introduce the following assumptions.

(H1) Given $\kappa > 0$, suppose $\varphi(\cdot)\in C_{\mathbb{F}}([0,T];L^{4+\kappa}(\Omega;\mathbb{R}^n))$, and for $U\subset\mathbb{R}^m$, $u(\cdot)\in\mathcal{U}_{ad}$, where

$$
\mathcal{U}_{ad} := \Big\{ u(\cdot) : [0,T]\times\Omega \to U \text{ is measurable and } \mathbb{F}\text{-adapted, } \sup_{t\in[0,T]}\mathbb{E}|u(t)|^{4+\kappa} < \infty \Big\},
$$

$b, \sigma : [0,T]^2\times\mathbb{R}^n\times U\times\Omega \to \mathbb{R}^n$ are measurable such that $s\mapsto \big(b(t,s,x,u),\sigma(t,s,x,u)\big)$ is $\mathbb{F}$-adapted, b, σ are linear growth of x, u, twice continuously differentiable of x with bounded first, second order derivatives,

$$
|f(t,s,x,u) - f(t',s,x,u)| \le \rho(|t-t'|)\big[1+|x|+|u|\big],\ t,t',s\in[0,T],\ f := b,\sigma,b_x,\sigma_x,
$$

with continuous modulus $\rho(\cdot)$. Moreover, $(s,x,u)\mapsto b_{xx}(t,s,x,u),\sigma_{xx}(t,s,x,u)$ are continuous uniformly in $t\in[0,T]$.

Based on these assumptions, we give the well-posedness of (1.1), the proof of which is straightforward adaptation of the counterparts in *e.g.* [14, 17].

Lemma 2.1. *Let* (H1) *hold. Then there exists* $X(\cdot)\in C_{\mathbb{F}}([0,T];L^{4+\kappa}(\Omega;\mathbb{R}^n))$ *satisfying (1.1) such that for absolute constant* K,

$$\begin{aligned}\sup_{t\in[0,T]}\mathbb{E}|X(t)|^{4+\kappa}\leq K\Big[&\sup_{t\in[0,T]}\mathbb{E}|\varphi(t)|^{4+\kappa}+\sup_{t\in[0,T]}\mathbb{E}\int_0^t|b(t,s,0,u(s))|^{4+\kappa}\mathrm{d}s\\&+\sup_{t\in[0,T]}\mathbb{E}\int_0^t|\sigma(t,s,0,u(s))|^{4+\kappa}\mathrm{d}s\Big].\end{aligned}\tag{2.2}$$

For the involved functions h and l in (1.2), we make the following assumption.

(H2) Let $h:\mathbb{R}^n\times\Omega\to\mathbb{R}$, $l:[0,T]\times\mathbb{R}^n\times U\times\Omega\to\mathbb{R}$ be measurable such that $x\mapsto h(x)$, $(x,u)\mapsto l(s,x,u)$ are twice continuously differentiable with

$$\big[|h_x(x)|+|l_x(t,x,u)|\big]\leq L[1+|x|+|u|],\quad \big[|h_{xx}(x)|+|l_{xx}(t,x,u)|\big]\leq L.$$

3. Maximum principle for controlled SVIEs

This section is devoted to obtaining maximum principle of Problem (SVIEs) *via* spike variation. Without further statement, let $u\in U$, $\tau\in[0,T)$, $\varepsilon>0$ be sufficiently small such that $\tau+\varepsilon\leq T$,

$$E_{\tau,\varepsilon}:=[\tau,\tau+\varepsilon],\quad u^{\varepsilon}(\cdot):=uI_{E_{\tau,\varepsilon}}(\cdot)+\bar{u}(\cdot)I_{[0,T]/E_{\tau,\varepsilon}}(\cdot).\tag{3.1}$$

3.1. First-order variational equations and related quadratic functional

In this part, we derive one quadratic functional of the solutions for first-order variational equations. The appropriate procedures of treating this functional, which are of course not necessary with convex control region [14, 19, 20], are indispensable in establishing our maximum principles.

Inspired by the SDEs case in *e.g.* [13, 21], we introduce the first-order variational equations which include two linear SVIEs as follows,

$$\begin{cases}X_1(t)=\displaystyle\int_0^t\bar{b}_x(t,s)X_1(s)\mathrm{d}s+\int_0^t\big[\bar{\sigma}_x(t,s)X_1(s)+\delta\sigma(t,s)\big]\mathrm{d}W(s),\\ X_2(t)=\varphi_2(t)+\displaystyle\int_0^t\bar{b}_x(t,s)X_2(s)\mathrm{d}s+\int_0^t\bar{\sigma}_x(t,s)X_2(s)\mathrm{d}W(s).\end{cases}\tag{3.2}$$

Here $t\in[0,T]$, and for $f:=b,\ \sigma,\ \sigma_x$, we make the following conventions,

$$\begin{cases}\varphi_2(t):=\displaystyle\int_0^t\Big[\frac{1}{2}\bar{b}_{xx}(t,s)X_1^2(s)+\delta b(t,s)\Big]\mathrm{d}s+\int_0^t\Big[\frac{1}{2}\bar{\sigma}_{xx}(t,s)X_1^2(s)+\delta\sigma_x(t,s)X_1(s)\Big]\mathrm{d}W(s),\\ \bar{f}_x(t,s):=f_x(t,s,\bar{X}(s),\bar{u}(s)),\ \ \delta f(t,s):=f(t,s,\bar{X}(s),u^{\varepsilon}(s))-f(t,s,\bar{X}(s),\bar{u}(s)),\\ \bar{f}_{xx}(t,s)X_1^2(s):=\Big(\mathrm{tr}\{f^1_{xx}(t,s)X_1(s)X_1(s)^{\top}\},\dots,\mathrm{tr}\{f^n_{xx}(t,s)X_1(s)X_1(s)^{\top}\}\Big)^{\top}.\end{cases}\tag{3.3}$$

Inspired by Lemma 1 of [13], we give the following estimates. Since the proof is standard in SVIEs theory, we choose to give a sketch for readers' convenience.

Lemma 3.1. *Suppose (H1) holds true, $(\bar X,\bar u)$ is an optimal pair, X^ε is the state process associated with u^ε in (3.1). Then*

$$\sup_{t\in[0,T]}\mathbb{E}|X_1(t)|^2\leq K\varepsilon,\quad \sup_{t\in[0,T]}\mathbb{E}|X^\varepsilon(t)-\bar X(t)-X_1(t)-X_2(t)|^2\leq o(\varepsilon^2). \tag{3.4}$$

Proof. For notational simplicity, suppose $n=1$. The multidimensional case can be obtained similarly.

For any $k\in[2,4+\kappa]$, and $t\in[0,T]$, it is a direct calculation that

$$\mathbb{E}\Big|\int_0^t\delta\sigma(t,s)\mathrm{d}W(s)\Big|^k\leq K\Big[\sup_{t\in[0,T]}\mathbb{E}|\bar X(t)|^k+\mathbb{E}|\bar u(t)|^k+1\Big]\cdot\varepsilon^{\frac{k}{2}}.$$

As a result, using (2.2) in Lemma 2.1, we then derive the estimate of $X_1(\cdot)$ in (3.4).

To prove the second result in (3.4), we need two preparations. First, we observe that

$$\begin{aligned}X^\varepsilon(t)-\bar X(t)=&\int_0^t\big[\delta b(t,s)+\widetilde b_x(t,s)[X^\varepsilon(s)-\bar X(s)]\big]\mathrm{d}s\\&+\int_0^t\big[\delta\sigma(t,s)+\widetilde\sigma_x(t,s)[X^\varepsilon(s)-\bar X(s)]\big]\mathrm{d}W(s),\end{aligned}$$

where for example,

$$\widetilde b_x(t,s):=\int_0^1\lambda b_x(t,s,\lambda\bar X(s)+(1-\lambda)X^\varepsilon(s),u^\varepsilon(s))d\lambda.$$

For any $l\in[2,4+\kappa]$, by estimate (2.2) and the assumed integrability of $(\bar X(\cdot),\bar u(\cdot))$, we deduce that

$$\sup_{t\in[0,T]}\mathbb{E}|X^\varepsilon(t)-\bar X(t)|^l\leq K\Big\{\sup_{t\in[0,T]}\mathbb{E}\Big[\int_0^t|\delta b(t,s)|\mathrm{d}s\Big]^l+\sup_{t\in[0,T]}\mathbb{E}\Big[\int_0^t|\delta\sigma(t,s)|^2\mathrm{d}s\Big]^{\frac{l}{2}}\Big\}\leq K\varepsilon^{\frac{l}{2}}. \tag{3.5}$$

As to the second preparation, we need to estimate $\widehat{\mathcal{X}}_2(\cdot):=X^\varepsilon(\cdot)-\bar X(\cdot)-X_1(\cdot)$, where

$$\begin{aligned}\widehat{\mathcal{X}}_2(t)=&\psi_1(t)+\int_0^t\bar b_x(t,s)\widehat{\mathcal{X}}_2(s)\mathrm{d}s+\int_0^t\bar\sigma_x(t,s)\widehat{\mathcal{X}}_2(s)\mathrm{d}W(s)\\&+\int_0^t\big[\delta\sigma_x(t,s)[X^\varepsilon(s)-\bar X(s)]+\frac12\widetilde\sigma_{xx}(t,s)[X^\varepsilon(s)-\bar X(s)]^2\big]\mathrm{d}W(s).\end{aligned}$$

In the above,

$$\begin{aligned}\psi_1(t):=&\int_0^t\big[\delta b_x(t,s)[X^\varepsilon(s)-\bar X(s)]+\frac12\widetilde b_{xx}(t,s)[X^\varepsilon(s)-\bar X(s)]^2+\delta b(t,s)\big]\mathrm{d}s,\\ \widetilde f_{xx}(t,s):=&\,2\int_0^1\lambda f_{xx}(t,s,\lambda\bar X(s)+(1-\lambda)X^\varepsilon(s),u^\varepsilon(s))\mathrm{d}\lambda,\quad f:=b,\sigma.\end{aligned}$$

Since $\sigma_x(\cdot)$, $\widetilde{\sigma}_{xx}(\cdot)$ are bounded, for the last stochastic integral term in $\widehat{\mathcal{X}}_2(\cdot)$,

$$\sup_{t\in[0,T]}\mathbb{E}\Big|\int_0^t\big[\delta\sigma_x(t,s)[X^\varepsilon(s)-\bar{X}(s)]+\frac{1}{2}\widetilde{\sigma}_{xx}(t,s)[X^\varepsilon(s)-\bar{X}(s)]^2\big]\mathrm{d}W(s)\Big|^p\le K\varepsilon^p.$$

The boundedness of $b_x(\cdot)$, $\widetilde{b}_{xx}(\cdot)$ and the assumption of b in (H1) yield that

$$\sup_{t\in[0,T]}\mathbb{E}|\psi_1(t)|^p\le K\big[\varepsilon^p+\varepsilon^{\frac{3p}{2}}\big].$$

Consequently, with the additional help of Lemma 2.1, one immediately derives that

$$\sup_{t\in[0,T]}\mathbb{E}|\widehat{\mathcal{X}}_2(t)|^p\le K\varepsilon^p. \tag{3.6}$$

Now we return back to $\widehat{\mathcal{X}}(\cdot):=X^\varepsilon(\cdot)-\bar{X}(\cdot)-X_1(\cdot)-X_2(\cdot)$ which is described as

$$\widehat{\mathcal{X}}(t)=\int_0^t G_1(t,s)\mathrm{d}s+\int_0^t G_2(t,s)\mathrm{d}W(s)+\int_0^t\bar{b}_x(t,s)\widehat{\mathcal{X}}(s)\mathrm{d}s+\int_0^t\bar{\sigma}_x(t,s)\widehat{\mathcal{X}}(s)\mathrm{d}W(s).$$

Here

$$\begin{aligned}G_1(t,s)&:=\delta b_x(t,s)[X^\varepsilon(s)-\bar{X}(s)]+\frac{1}{2}\widetilde{b}_{xx}(t,s)[X^\varepsilon(s)-\bar{X}(s)]^2-\frac{1}{2}\bar{b}_{xx}(t,s)X_1^2(s),\\ G_2(t,s)&:=\delta\sigma_x(t,s)\widehat{\mathcal{X}}_2(s)+\frac{1}{2}\widetilde{\sigma}_{xx}(t,s)[X^\varepsilon(s)-\bar{X}(s)]^2-\frac{1}{2}\bar{\sigma}_{xx}(t,s)X_1^2(s).\end{aligned}$$

Since b_x, σ_x are bounded, for the first term in G_1, G_2, we respectively obtain that

$$\begin{aligned}&\sup_{t\in[0,T]}\mathbb{E}\Big[\int_0^t|\delta b_x(t,s)[X^\varepsilon(s)-\bar{X}(s)]|\mathrm{d}s\Big]^2\le K\varepsilon^{\frac{5}{2}},\\ &\sup_{t\in[0,T]}\mathbb{E}\int_0^t|\delta\sigma_x(t,s)[X^\varepsilon(s)-\bar{X}(s)-X_1(s)]|^2\mathrm{d}s\le K\varepsilon^3.\end{aligned}$$

By virtue of (3.5), (H1) and dominated convergence theorem, one obtains that

$$\sup_{t\in[0,T]}\frac{1}{\varepsilon^2}\mathbb{E}\Big[\int_0^t\big|\widetilde{b}_{xx}(t,s)-\bar{b}_{xx}(t,s)\big||X^\varepsilon(s)-\bar{X}(s)|^2\mathrm{d}s\Big]^2\to 0,\quad\varepsilon\to 0.$$

We observe that

$$|X^\varepsilon(\cdot)-\bar{X}(\cdot)|^2-X_1(\cdot)^2=[\widehat{\mathcal{X}}_2(\cdot)]^2+2X_1(\cdot)\widehat{\mathcal{X}}_2(\cdot).$$

Therefore by above (3.6) and the previous obtained estimate of $X_1(\cdot)$,

$$\sup_{t\in[0,T]}\mathbb{E}\Big[\int_0^t|\bar{b}_{xx}(t,s)|\big|(X^\varepsilon(s)-\bar{X}(s))^2-X_1^2(s)\big|\mathrm{d}s\Big]^2\le K[\varepsilon^3+\varepsilon^4].$$

One can use similar ideas to treat the terms of σ. At last, thanks to (2.2) in Lemma 2.1,

$$\sup_{t\in[0,T]}\mathbb{E}|\widehat{\mathcal{X}}(t)|^2\leq \sup_{t\in[0,T]}\mathbb{E}\Big[\int_0^t|G_1(t,s)|\mathrm{d}s\Big]^2+\sup_{t\in[0,T]}\mathbb{E}\Big|\int_0^t G_2(t,s)\mathrm{d}W(s)\Big|^2=o(\varepsilon^2).$$

□

Thanks to Lemma 3.1, by following the same procedures as Lemma 2 in [13], one sees that

$$\begin{aligned}o(\varepsilon)\leq&\ \mathbb{E}\int_0^T\bar{l}_x(s)^\top\big[X_1(s)+X_2(s)\big]\mathrm{d}s+\mathbb{E}\big[\bar{h}_x(T)^\top(X_1(T)+X_2(T))\big]\\&+\mathbb{E}\Big[\int_0^T\delta l(s)\mathrm{d}s+\frac{1}{2}\int_0^T X_1(s)^\top\bar{l}_{xx}(s)X_1(s)\mathrm{d}s+\frac{1}{2}X_1(T)^\top\bar{h}_{xx}(T)X_1(T)\Big],\end{aligned}\tag{3.7}$$

where for example,

$$\begin{aligned}&\bar{l}_x(\cdot):=l_x(\cdot,\bar{X}(\cdot),\bar{u}(\cdot)),\ \bar{l}_{xx}(\cdot):=l_{xx}(\cdot,\bar{X}(\cdot),\bar{u}(\cdot)),\ \bar{h}_x(T):=h_x(\bar{X}(T)),\\&\bar{h}_{xx}(T):=h_{xx}(\bar{X}(T)),\ \delta l(\cdot):=l(\cdot,\bar{X}(\cdot),u^\varepsilon(\cdot))-l(\cdot,\bar{X}(\cdot),\bar{u}(\cdot)).\end{aligned}\tag{3.8}$$

To treat the first two terms on the right hand of (3.7), we need two preparations. First, we introduce *first-order adjoint equation* of

$$\begin{aligned}\bar{Y}(t)=&\ \bar{l}_x(t)^\top+\bar{b}_x(T,t)^\top\bar{h}_x(T)+\bar{\sigma}_x(T,t)^\top\bar{\pi}(t)+\int_t^T\bar{b}_x(s,t)^\top\bar{Y}(s)\mathrm{d}s\\&+\int_t^T\bar{\sigma}_x(s,t)^\top\bar{Z}(s,t)\mathrm{d}s-\int_t^T\bar{Z}(t,s)\mathrm{d}W(s),\end{aligned}\tag{3.9}$$

and the Hamiltonian function

$$\begin{aligned}H(t,x,\bar{X}(\cdot),\bar{Y}(\cdot),\bar{Z}(\cdot,t),u):=&\ \mathbb{E}_t\big[b(T,t,x,u)^\top\bar{h}_x(T)+\sigma(T,t,x,u)^\top\bar{\pi}(t)\big]\\&+l(t,x,u)+\mathbb{E}_t\int_t^T b(s,t,x,u)^\top\bar{Y}(s)\mathrm{d}s+\int_t^T\sigma(s,t,x,u)^\top\bar{Z}(s,t)\mathrm{d}s.\end{aligned}\tag{3.10}$$

Here $\bar{h}_x(T)$, $\bar{\pi}(\cdot)$ of (3.9) satisfy the following relation,

$$\bar{h}_x(T)=\mathbb{E}\bar{h}_x(T)+\int_0^T\bar{\pi}(s)\mathrm{d}W(s).$$

The existence and uniqueness of $\bar{\pi}(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$ is guaranteed by martingale representation theorem. Observe that (3.9) is a linear backward stochastic Volterra integral equation (BSVIE) which admits a unique pair of solution $(\bar{Y}(\cdot),\bar{Z}(\cdot,\cdot))\in\mathcal{H}^2(0,T;\mathbb{R}^n)\equiv L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)\times L^2(0,T;L^2_{\mathbb{F}}(0,T;\mathbb{R}^n))$ such that

$$\bar{Y}(t)=\mathbb{E}_s\bar{Y}(t)+\int_s^t\bar{Z}(t,r)\mathrm{d}W(r),\quad s\in[0,t],\quad t\in[0,T].\ \ a.e.$$

The above $(\bar{Y},\bar{Z})$ is named as M-solutions of BSVIEs. We refer to *e.g.* [14, 17, 20] for more details.

As the second preparation, we need the dual principle established in [20]. For readers' convenience, we demonstrate it as follows.

Proposition 3.2. *Suppose* $A_i(\cdot\,,\cdot) \in L^\infty(0,T;L^\infty_{\mathbb{F}}(0,T;\mathbb{R}^{n\times n}))$, $(i=0,1)$, $\varphi(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$, *and* $\psi(\cdot) \in L^2_{\mathcal{F}_T}(0,T;\mathbb{R}^n)$. *Let* $X(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$ *be the solution of FSVIE:*

$$X(t) = \varphi(t) + \int_0^t A_0(t,s)X(s)\mathrm{d}s + \int_0^t A_1(t,s)X(s)\mathrm{d}W(s), \quad t \in [0,T], \tag{3.11}$$

and $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2(0,T;\mathbb{R}^n)$ *be the adapted M-solution to the following BSVIE:*

$$Y(t) = \psi(t) + \int_t^T \Big[A_0(s,t)^\top Y(s) + A_1(s,t)^\top Z(s,t)\Big]\mathrm{d}s - \int_t^T Z(t,s)\mathrm{d}W(s), \quad t \in [0,T]. \tag{3.12}$$

Then the following relation holds:

$$\mathbb{E}\int_0^T \langle\, X(t), \psi(t)\,\rangle\, \mathrm{d}t = \mathbb{E}\int_0^T \langle\, \varphi(t), Y(t)\,\rangle\, \mathrm{d}t. \tag{3.13}$$

Further, suppose $A_1(T,\cdot)$, $A_2(T,\cdot)$, $\varphi(T)$ *are well-defined,* $\varphi(T)$, $\eta \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$. *Let* $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2(0,T;\mathbb{R}^n)$ *be the adapted M-solution to the following BSVIE:*

$$\begin{aligned} Y(t) = \psi(t) &+ A_0(T,t)^\top \eta + A_1(T,t)^\top \zeta(t) + \int_t^T \big[A_0(s,t)^\top Y(s) + A_1(s,t)^\top Z(s,t)\big]\mathrm{d}s \\ &- \int_t^T Z(t,s)\mathrm{d}W(s), \quad t \in [0,T], \end{aligned} \tag{3.14}$$

where $\zeta(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ *is the unique process that satisfies*

$$\eta = \mathbb{E}\eta + \int_0^T \zeta(t)\mathrm{d}W(t). \tag{3.15}$$

Then the following holds:

$$\mathbb{E}\Big\{\langle\, X(T), \eta\,\rangle + \int_0^T \langle\, X(t), \psi(t)\,\rangle\, \mathrm{d}t\Big\} = \mathbb{E}\Big\{\langle\, \varphi(T), \eta\,\rangle + \int_0^T \langle\, \varphi(t), Y(t)\,\rangle\, \mathrm{d}t\Big\}. \tag{3.16}$$

Now we make the following conventions

$$\begin{aligned} &\varphi(t) := \int_0^t \big[\frac{1}{2}\bar{b}_{xx}(t,s)X_1^2(s) + \delta b(t,s)\big]\mathrm{d}s + \int_0^t \big[\frac{1}{2}\bar{\sigma}_{xx}(t,s)X_1^2(s) + \delta\sigma_x(t,s)X_1(s) + \delta\sigma(t,s)\big]\mathrm{d}W(s), \\ &A_0(t,s) := \bar{b}_x(t,s), \quad A_1(t,s) := \bar{\sigma}_x(t,s), \quad \psi(t) := \bar{l}_x(t), \quad \eta := \bar{h}_x(T), \quad \zeta(t) := \bar{\pi}(t), \end{aligned}$$

and apply Proposition 3.2 into our problem. Recall the notations in (3.3), we observe that

$$
\begin{aligned}
\mathbb{E}\,\langle\,\varphi(T),\eta\,\rangle &= \mathbb{E}\int_0^T \langle\,\frac{1}{2}\bar{\sigma}_{xx}(T,s)X_1^2(s)+\delta\sigma_x(T,s)X_1(s)+\delta\sigma(T,s),\mathbb{E}\bar{h}_x(T)+\int_0^T\bar{\pi}(r)\mathrm{d}W(r)\,\rangle\,\mathrm{d}W(s)\\
&\quad+\mathbb{E}\int_0^T\langle\,\frac{1}{2}\bar{b}_{xx}(T,s)X_1^2(s)+\delta b(T,s),\bar{h}_x(T)\,\rangle\,\mathrm{d}s\\
&=\mathbb{E}\int_0^T\frac{1}{2}\big[\,\langle\,\bar{\sigma}_{xx}(T,s)X_1^2(s),\bar{\pi}(s)\,\rangle+\langle\,\bar{b}_{xx}(T,s)X_1^2(s),\bar{h}_x(T)\,\rangle\,\big]\mathrm{d}s\\
&\quad+\mathbb{E}\int_0^T\langle\delta\sigma_x(T,s)X_1(s),\bar{\pi}(s)\,\rangle\,\mathrm{d}s+\mathbb{E}\int_0^T\big[\,\langle\,\delta\sigma(T,s),\bar{\pi}(s)\,\rangle+\langle\,\delta b(T,s),\bar{h}_x(T)\,\rangle\,\big]\mathrm{d}s.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathbb{E}\int_0^T\langle\,\varphi(t),Y(t)\,\rangle\,\mathrm{d}t &= \frac{1}{2}\mathbb{E}\int_0^T\int_s^T\big[\,\langle\,\bar{b}_{xx}(t,s)X_1^2(s),Y(t)\,\rangle+\langle\,\bar{\sigma}_{xx}(t,s)X_1^2(s),Z(t,s)\,\rangle\,\big]\mathrm{d}t\mathrm{d}s\\
&\quad+\mathbb{E}\int_0^T\int_s^T\langle\delta\sigma_x(t,s)X_1(s),Z(t,s)\,\rangle\,\mathrm{d}t\mathrm{d}s+\mathbb{E}\int_0^T\int_s^T\big[\,\langle\delta b(t,s),Y(t)\,\rangle+\langle\,\delta\sigma(t,s),Z(t,s)\,\rangle\,\big]\mathrm{d}t\mathrm{d}s.
\end{aligned}
$$

Therefore, using again (3.16), we have

$$
\begin{aligned}
&\mathbb{E}\int_0^T\bar{l}_x(s)^\top\big[X_1(s)+X_2(s)\big]\mathrm{d}s+\mathbb{E}\big[\bar{h}_x(T)^\top(X_1(T)+X_2(T))\big]\\
&=\frac{1}{2}\mathbb{E}\int_0^T\Big\{\,\langle\,\bar{\sigma}_{xx}(T,s)X_1^2(s),\bar{\pi}(s)\,\rangle+\langle\,\bar{b}_{xx}(T,s)X_1^2(s),\bar{h}_x(T)\,\rangle\\
&\qquad+\int_s^T\big[\,\langle\,\bar{b}_{xx}(r,s)X_1^2(s),Y(r)\,\rangle+\langle\,\bar{\sigma}_{xx}(r,s)X_1^2(s),Z(r,s)\,\rangle\,\big]\mathrm{d}r\Big\}\mathrm{d}s\\
&\qquad+\mathbb{E}\int_0^T\Big[\,\langle\,\delta\sigma(T,s),\bar{\pi}(s)\,\rangle+\langle\,\delta b(T,s),\bar{h}_x(T)\,\rangle+\int_s^T\big[\,\langle\delta b(r,s),Y(r)\,\rangle+\langle\,\delta\sigma(r,s),Z(r,s)\,\rangle\,\big]\mathrm{d}r\Big]\mathrm{d}s\\
&\qquad+\mathbb{E}\int_0^T\Big\{\,\langle\delta\sigma_x(T,s)X_1(s),\bar{\pi}(s)\,\rangle+\int_s^T\langle\delta\sigma_x(r,s)X_1(s),Z(r,s)\,\rangle\,\mathrm{d}r\Big\}\mathrm{d}s.
\end{aligned}
\tag{3.17}
$$

Because of Lemma 3.1, we have

$$
\mathbb{E}\int_0^T\Big\{\,\langle\delta\sigma_x(T,s)X_1(s),\bar{\pi}(s)\,\rangle+\int_s^T\langle\delta\sigma_x(r,s)X_1(s),Z(r,s)\,\rangle\,\mathrm{d}r\Big\}\mathrm{d}s=o(\varepsilon). \tag{3.18}
$$

Plugging (3.17) into the above (3.7), and taking consideration of (3.18), we arrive at

$$
o(1)\leq\frac{1}{\varepsilon}\mathbb{E}\int_0^T\Delta H^\varepsilon(t)\mathrm{d}t+\mathscr{E}(\varepsilon), \tag{3.19}
$$

where $\Delta H^\varepsilon(\cdot)$, $\mathscr{E}(\varepsilon)$ are defined as

$$\begin{aligned}
&\Delta H^\varepsilon(t) := H(t,\bar X(t),\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot,t),u^\varepsilon(t)) - H(t,\bar X(t),\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot,t),\bar u(t)),\\
&\mathscr{E}(\varepsilon) := \frac{1}{2\varepsilon}\mathbb{E}\Big\{\int_0^T \mathrm{tr}\big[\bar H_{xx}(t)X_1(t)X_1(t)^\top\big]\mathrm{d}t + \mathrm{tr}\big[\bar h_{xx}(T)X_1(T)X_1(T)^\top\big]\Big\},\\
&\bar H_{xx}(t) := H_{xx}(t,\bar X(t),\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot,t),\bar u(t)),\quad t\in[0,T].
\end{aligned}\tag{3.20}$$

To give the maximum principle, in the following two subsections we deal with $\mathscr{E}(\varepsilon)$ in (3.19).

3.2. The limit of $\mathscr{E}(\varepsilon)$

In order to obtain maximum principle, we first explore the limit of $\mathscr{E}(\varepsilon)$ as $\varepsilon\to 0$.

To begin with, given optimal pair $(\bar X(\cdot),\bar u(\cdot))$, $u\in U$, we define

$$\begin{aligned}
&\delta\bar\sigma(t,\tau) := \sigma(t,\tau,\bar X(\tau),u) - \sigma(t,\tau,\bar X(\tau),\bar u(\tau)),\\
&\Delta\bar\sigma(\cdot,\tau) := \big(\delta\bar\sigma(T,\tau),\delta\bar\sigma(\cdot,\tau)\big),\quad t,\ \tau\in[0,T].
\end{aligned}\tag{3.21}$$

Under (H1), one has $\delta\bar\sigma(\cdot,\tau)\in C_{\mathcal{F}_\tau}([0,T];L^{4+\kappa}(\Omega;\mathbb{R}^n))$.

For $\bar b_x(\cdot)$, $\bar\sigma_x(\cdot)$, $\bar h_{xx}(T)$, $\bar H_{xx}(\cdot)$ in (3.3), (3.8), (3.20), we introduce

$$\begin{cases}
F^{\delta\bar\sigma,\delta\bar\sigma}(\tau) := \mathbb{E}_\tau\displaystyle\int_\tau^T \mathbb{X}(s)^\top\bar H_{xx}(s)\mathbb{X}(s)\mathrm{d}s + \mathbb{E}_\tau\big[\mathbb{X}(T)^\top\bar h_{xx}(T)\mathbb{X}(T)\big],\\
\mathbb{X}(t)=\delta\bar\sigma(t,\tau)+\displaystyle\int_\tau^t \bar b_x(t,s)\mathbb{X}(s)\mathrm{d}s+\int_\tau^t\bar\sigma_x(t,s)\mathbb{X}(s)\mathrm{d}W(s),\quad\forall t\in[\tau,T].
\end{cases}\tag{3.22}$$

Thanks to (H1) and Lemma 2.1, (3.22) admits a unique $\mathbb{X}(\cdot)\in C_{\mathbb{F}}([\tau,T];L^{4+\kappa}(\Omega;\mathbb{R}^n))$.

In this section, we aim to prove $F^{\delta\bar\sigma,\delta\bar\sigma}(\tau)$ is the limit of $\mathscr{E}(\varepsilon_n)$ for some subsequence $\{\varepsilon_n\}$. To get more intuitive feelings, let $n=1$, $\bar b_x(\cdot,\cdot)=\bar\sigma_x(\cdot,\cdot)=0$. Given deterministic $\mathcal{Q}(\cdot)$, by Fubini theorem and $X_1(\cdot)=\int_\tau^\cdot \delta\bar\sigma(\cdot,s)I_{[\tau,\tau+\varepsilon]}(s)\mathrm{d}W(s)$,

$$\frac{1}{\varepsilon}\mathbb{E}\int_\tau^T\mathcal{Q}(t)\Big|X_1(t)\Big|^2\mathrm{d}t = \frac{1}{\varepsilon}\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_s^T\mathcal{Q}(t)|\delta\bar\sigma(t,s)|^2\mathrm{d}t\mathrm{d}s.$$

For $\tau\in[0,T)$, a.e., by Lebesgue differentiation theorem,

$$\lim_{\varepsilon\to0}\frac{1}{\varepsilon}\mathbb{E}\int_\tau^T\mathcal{Q}(t)\Big|X_1(t)\Big|^2\mathrm{d}t = \mathbb{E}\int_\tau^T\mathcal{Q}(t)|\delta\bar\sigma(t,\tau)|^2\mathrm{d}t = \mathbb{E}\int_\tau^T\mathcal{Q}(t)|\mathbb{X}(t)|^2\mathrm{d}t.\tag{3.23}$$

Above (3.23) indicates certain asymptotic connection between $X_1(\cdot)$ and $\mathbb{X}(\cdot)$. We use this basic idea in our framework and present the following result.

Lemma 3.3. *Suppose (H1), (H2) hold true, $(\bar{X}(\cdot),\bar{u}(\cdot))$ is optimal pair. Then there exists $\{\varepsilon_n\}_{n\geq 1}$ satisfying $\varepsilon_n\to 0$ as $n\to\infty$, such that for almost $\tau\in[0,T]$,*

$$\begin{aligned}&\lim_{\varepsilon_n\to 0}\frac{1}{\varepsilon_n}\Big[\mathbb{E}\int_0^T X_1(s)^\top \bar{H}_{xx}(s)X_1(s)\mathrm{d}s+\mathbb{E}\big[X_1(T)^\top\bar{h}_{xx}(T)X_1(T)\big]\Big]\\&=\mathbb{E}\int_\tau^T \mathbb{X}(s)^\top\bar{H}_{xx}(s)\mathbb{X}(s)\mathrm{d}s+\mathbb{E}\big[\mathbb{X}(T)^\top\bar{h}_{xx}(T)\mathbb{X}(T)\big]\equiv\mathbb{E}F^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau).\end{aligned}$$

To prove Lemma 3.3, we take a closer look at $X_1(\cdot)$ in (3.2). Actually, according to its definition, one can rewrite $X_1(\cdot)$ as,

$$X_1(t)=\begin{cases}0, \qquad t\in[0,\tau],\\ \displaystyle\int_\tau^t \bar{b}_x(t,s)X_1(s)\mathrm{d}s+\int_\tau^t\big[\bar{\sigma}_x(t,s)X_1(s)+\delta\bar{\sigma}(t,s)\big]\mathrm{d}W(s),\ t\in[\tau,\tau+\varepsilon],\\ \displaystyle\rho_1(t)+\int_{\tau+\varepsilon}^t \bar{b}_x(t,s)X_1(s)\mathrm{d}s+\int_\tau^t\bar{\sigma}_x(t,s)X_1(s)\mathrm{d}W(s),\ t\in[\tau+\varepsilon,T],\end{cases}\tag{3.24}$$

where $\rho_1(\cdot)\in C([\tau+\varepsilon,T];L^4(\Omega;\mathbb{R}))$ is $\mathcal{F}_{\tau+\varepsilon}$-measurable defined as,

$$\rho_1(t):=\int_\tau^{\tau+\varepsilon}\bar{b}_x(t,s)X_1(s)\mathrm{d}s+\int_\tau^{\tau+\varepsilon}\big[\bar{\sigma}_x(t,s)X_1(s)+\delta\bar{\sigma}(t,s)\big]\mathrm{d}W(s),\quad t\geq\tau+\varepsilon.$$

For small $\varepsilon>0$, we introduce $Y_1(\cdot)$ on $[\tau+\varepsilon,T]$,

$$Y_1(t)=\varrho_1(t)+\int_{\tau+\varepsilon}^t\bar{b}_x(t,s)Y_1(s)\mathrm{d}s+\int_{\tau+\varepsilon}^t\bar{\sigma}_x(t,s)Y_1(s)\mathrm{d}W(s),\tag{3.25}$$

where $\varrho_1(\cdot):=\varepsilon^{-\frac{1}{2}}\delta\bar{\sigma}(\cdot,\tau)\Big(W(\tau+\varepsilon)-W(\tau)\Big)$.

We claim that $\varrho_1(\cdot)\in C_{\mathcal{F}_{\tau+\varepsilon}}([\tau+\varepsilon,T],L^4(\Omega;\mathbb{R}^n))$. Therefore, thanks to Lemma 2.1 and (H1), (3.25) admits a unique solution $Y_1(\cdot)\in C_{\mathbb{F}}([\tau+\varepsilon,T],L^4(\Omega;\mathbb{R}^n))$. In fact, since $\bar{u}(\cdot)\in\mathcal{U}_{ad}$, from Lemma 2.1 and (H1), $\sup\limits_{\tau\in[0,T]}\mathbb{E}|\bar{X}(\tau)|^{4+\kappa}<\infty$, and

$$\sup_{t\in[0,T]}\mathbb{E}|\delta\bar{\sigma}(t,\tau)|^{4+\kappa}\leq L\Big[1+\sup_{\tau\in[0,T]}\mathbb{E}|\bar{X}(\tau)|^{4+\kappa}+\sup_{\tau\in[0,T]}\mathbb{E}|\bar{u}(\tau)|^{4+\kappa}\Big]<\infty.$$

Consequently, by virtue of Hölder inequality and Jensen's inequality of expectation,

$$\sup_{t\in[\tau+\varepsilon,T]}\mathbb{E}|\varrho_1(t)|^4\leq\sup_{t\in[\tau+\varepsilon,T]}\Big[\mathbb{E}|\delta\bar{\sigma}(t,\tau)|^{4+\kappa}\Big]^{\frac{4}{4+\kappa}}\cdot\Big[\frac{\big(4[p]+4\big)!}{2^{2[p]+2}\cdot\big(2[p]+2\big)!}\Big]^{\frac{p}{[p]+1}}<\infty.\tag{3.26}$$

Here $p:=\frac{4+\kappa}{\kappa}$, $[p]$ is the integer part of p, and we used one formula on Brownian motions: $\mathbb{E}|W(t)|^{2k}=\frac{(2k)!}{2^kk!}t^k,\ \ k\in\mathbb{N},\ \ t\in\mathbb{R}^+$. Moreover, for $t_0\in[\tau+\varepsilon,T]$, $\lim\limits_{t\to t_0}\mathbb{E}|\varrho_1(t)-\varrho_1(t_0)|^4=0$. Hence the conclusion of $\varrho_1(\cdot)$ is obvious.

For notational simplicity, we denote

$$\begin{cases} \mathbb{H}(\varepsilon, X_1) := \dfrac{1}{\varepsilon}\mathbb{E}\Big[\displaystyle\int_{\tau+\varepsilon}^{T} X_1(s)^\top \bar{H}_{xx}(s) X_1(s)\mathrm{d}s + X_1(T)^\top \bar{h}_{xx}(T) X_1(T)\Big], \\ \mathbb{H}(\varepsilon, Y_1) := \mathbb{E}\Big[\displaystyle\int_{\tau+\varepsilon}^{T} Y_1(s)^\top \bar{H}_{xx}(s) Y_1(s)\mathrm{d}s + Y_1(T)^\top \bar{h}_{xx}(T) Y_1(T)\Big]. \end{cases} \tag{3.27}$$

Lemma 3.4. *For $X_1(\cdot)$, $Y_1(\cdot)$, $\mathbb{H}(\varepsilon, X_1)$, $\mathbb{H}(\varepsilon, Y_1)$ in (3.2), (3.25), (3.27), there exists $\{\varepsilon_n\}_{n\geq 1}$ such that*

$$\lim_{n\to\infty} \varepsilon_n = 0, \quad \lim_{n\to\infty} \Big[\mathbb{H}(\varepsilon_n, X_1) - \mathbb{H}(\varepsilon_n, Y_1)\Big] = 0.$$

Proof. At first, for any $\varepsilon > 0$ we make the following observation

$$\begin{aligned} \mathbb{H}(\varepsilon, X_1) &= \mathbb{E}\int_{\tau+\varepsilon}^{T} \big[\frac{1}{\varepsilon} X_1(s)^\top \bar{H}_{xx}(s) X_1(s) - Y_1(s)^\top \bar{H}_{xx}(s) Y_1(s)\big]\mathrm{d}s \\ &\quad + \mathbb{H}(\varepsilon, Y_1) + \mathbb{E}\big[\frac{1}{\varepsilon} X_1(T)^\top \bar{h}_{xx}(T) X_1(T) - Y_1(T)^\top \bar{h}_{xx}(T) Y_1(T)\big] \\ &:= \mathbb{H}_1(\varepsilon) + \mathbb{H}(\varepsilon, Y_1) + \mathbb{H}_3(\varepsilon). \end{aligned} \tag{3.28}$$

We first treat $\mathbb{H}_1(\varepsilon)$. An easy calculation shows that

$$\Big|\mathbb{H}_1(\varepsilon)\Big| \leq K\Big[\mathbb{E}\int_{\tau+\varepsilon}^{T} |\varepsilon^{-\frac{1}{2}} X_1(s) - Y_1(s)|^4 \mathrm{d}s\Big]^{\frac{1}{4}} \Big[\mathbb{E}\int_{\tau+\varepsilon}^{T} \big[|\varepsilon^{-\frac{1}{2}} X_1(s)|^4 + |Y_1(s)|^4\big]\mathrm{d}s\Big]^{\frac{1}{4}},$$

where we use the fact: $|aa^\top - bb^\top| \leq |a-b|[|a|+|b|]$, with $a, b \in \mathbb{R}^n$. From (3.24), (3.25) and Lemma 2.1, we immediately have

$$\begin{aligned} &\sup_{t\in[\tau+\varepsilon,T]} \mathbb{E}\big|\varepsilon^{-\frac{1}{2}} X_1(t) - Y_1(t)\big|^4 \leq K \sup_{t\in[\tau+\varepsilon,T]} \mathbb{E}|\varepsilon^{-\frac{1}{2}}\rho_1(t) - \varrho_1(t)|^4 \\ &\leq K\varepsilon^{-2}\mathbb{E}\Big[\int_\tau^{\tau+\varepsilon} |X_1(s)|^2 \mathrm{d}s\Big]^2 + K\varepsilon^{-2} \sup_{t\in[\tau+\varepsilon,T]} \mathbb{E}\Big[\int_\tau^{\tau+\varepsilon} |\delta\bar{\sigma}(t,s) - \delta\bar{\sigma}(t,\tau)|^2 \mathrm{d}s\Big]^2. \end{aligned}$$

For the first term, denoted by $\mathcal{G}_1(\varepsilon)$, on the right hand, we obtain $\lim_{\varepsilon\to 0} \mathcal{G}_1(\varepsilon) = 0$ by Lemma 3.1. As to the second term, $\mathcal{G}_2(\varepsilon)$, thanks to Lemma 2.5 of [10], there exists $\{\varepsilon_n\}_{n\geq 1}$ such that

$$\mathcal{G}_2(\varepsilon_n) \leq \varepsilon_n^{-1}\mathbb{E}\int_\tau^{\varepsilon_n+\tau} \|\delta\bar{\sigma}(\cdot,s) - \delta\bar{\sigma}(\cdot,\tau)\|^4_{C([0,T];\mathbb{R}^n)} \mathrm{d}s \to 0, \qquad n\to\infty.$$

Consequently, for such $\{\varepsilon_n\}$, we conclude that

$$\lim_{n\to\infty} \sup_{t\in[\tau+\varepsilon_n,T]} \mathbb{E}\big|{\varepsilon_n}^{-\frac{1}{2}} X_1(t) - Y_1(t)\big|^4 = 0.$$

As a result, $\lim_{n\to\infty} \Big|\mathbb{H}_1(\varepsilon_n)\Big| = 0$. Similarly we prove that $\lim_{n\to\infty} \Big|\mathbb{H}_3(\varepsilon_n)\Big| = 0$. The conclusion is established *via* (3.28). □

To treat $Y_1(\cdot)$ in Lemma 3.4, for $\varrho_2(\cdot) := \delta\bar{\sigma}(\cdot,\tau)$, we need $Y_2(\cdot)$ on $[t+\varepsilon, T]$,

$$Y_2(t) = \varrho_2(t) + \int_{\tau+\varepsilon}^t \bar{b}_x(t,s)Y_2(s)\mathrm{d}s + \int_{\tau+\varepsilon}^t \bar{\sigma}_x(t,s)Y_2(s)\mathrm{d}W(s). \tag{3.29}$$

The solvability of $Y_2(\cdot) \in C_{\mathbb{F}}([\tau+\varepsilon, T]; L^{4+\kappa}(\Omega;\mathbb{R}^n))$ is followed by that of $Y_1(\cdot)$ in (3.25). Moreover, by the uniqueness of solutions in $C_{\mathbb{F}}([\tau+\varepsilon, T]; L^4(\Omega;\mathbb{R}))$, we see that

$$\begin{cases} Y_1(\cdot) = \varepsilon^{-\frac{1}{2}}(W(\tau+\varepsilon) - W(\tau))Y_2(\cdot), \\ \mathbb{H}(\varepsilon, Y_1) = \mathbb{E}\Big[\varepsilon^{-1}|W(\tau+\varepsilon) - W(\tau)|^2 \cdot F_1^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau+\varepsilon)\Big], \end{cases}$$

where

$$F_1^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau+\varepsilon) := \mathbb{E}_{\tau+\varepsilon}\Big[\int_{\tau+\varepsilon}^T Y_2(s)^\top \bar{H}_{xx}(s)Y_2(s)\mathrm{d}s + Y_2(T)^\top \bar{h}_{xx}(T)Y_2(T)\Big].$$

Lemma 3.5. *Given $\mathbb{X}(\cdot)$, $Y_1(\cdot)$, $\mathbb{H}(\varepsilon, Y_1)$ in (3.22), (3.25), (3.27), we have*

$$\lim_{\varepsilon\to 0} \mathbb{H}(\varepsilon, Y_1) = \mathbb{E}\int_\tau^T \mathbb{X}(s)^\top \bar{H}_{xx}(s)\mathbb{X}(s)\mathrm{d}s + \mathbb{E}\big[\mathbb{X}(T)^\top \bar{h}_{xx}(T)\mathbb{X}(T)\big]. \tag{3.30}$$

Proof. Recalling $F^{\delta\bar{\sigma},\delta\bar{\sigma}}(\cdot)$ in (3.22), we shall derive the conclusion if

$$\lim_{\varepsilon\to 0} \mathbb{E}\Big[\varepsilon^{-1}|W(\tau+\varepsilon) - W(\tau)|^2\big[F_1^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau+\varepsilon) - F^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau)\big]\Big] = 0.$$

To obtain this result, using similar ideas as in (3.26), we only prove

$$\lim_{\varepsilon\to 0} \mathbb{E}\big|F_1^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau+\varepsilon) - F^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau)\big|^{p_0} = 0, \quad p_0 = \frac{2(4+\kappa)}{8+\kappa} \in (1,2).$$

To this end, by defining

$$\begin{aligned} \Theta_1(\tau+\varepsilon) &:= \int_{\tau+\varepsilon}^T Y_2(s)^\top \bar{H}_{xx}(s)Y_2(s)\mathrm{d}s + Y_2(T)^\top \bar{h}_{xx}(T)Y_2(T), \\ \Theta_2(\tau) &:= \int_\tau^T \mathbb{X}(s)^\top \bar{H}_{xx}(s)\mathbb{X}(s)\mathrm{d}s + \mathbb{X}(T)^\top \bar{h}_{xx}(T)\mathbb{X}(T), \end{aligned}$$

we deduce that

$$\begin{aligned} &\mathbb{E}\big|F_1^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau+\varepsilon) - F^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau)\big|^{p_0} \equiv \mathbb{E}|\mathbb{E}_{\tau+\varepsilon}\Theta_1(\tau+\varepsilon) - \mathbb{E}_\tau\Theta_2(\tau)|^{p_0} \\ &\leq K\mathbb{E}\big|\Theta_1(\tau+\varepsilon) - \Theta_2(\tau)\big|^{p_0} + K\mathbb{E}\big|\mathbb{E}_{\tau+\varepsilon}\Theta_2(\tau) - \mathbb{E}_\tau\Theta_2(\tau)\big|^{p_0}. \end{aligned}$$

We prove the terms on right hand approach to zero as $\varepsilon \to 0$. We observe that

$$\begin{aligned}
&\mathbb{E}\big|\Theta_1(\tau+\varepsilon)-\Theta_2(\tau)\big|^{p_0}\\
&\leq K\mathbb{E}\Big\{\Big|\int_{\tau+\varepsilon}^T |Y_2(s)Y_2(s)^\top-\mathbb{X}(s)\mathbb{X}(s)^\top|^2\mathrm{d}s\Big|^{\frac{p_0}{2}}\Big[\int_{\tau+\varepsilon}^T|\bar{H}_{xx}(s)|^2\mathrm{d}s\Big]^{\frac{p_0}{2}}\Big\}\\
&\quad+K\mathbb{E}\Big\{\Big[\int_\tau^{\tau+\varepsilon}|\mathbb{X}(s)\mathbb{X}(s)^\top|^2\mathrm{d}s\Big]^{\frac{p_0}{2}}\Big[\int_\tau^{\tau+\varepsilon}|\bar{H}_{xx}(s)|^2\mathrm{d}s\Big]^{\frac{p_0}{2}}\Big\}\\
&\quad+K\mathbb{E}\Big\{|Y_2(T)Y_2(T)^\top-\mathbb{X}(T)\mathbb{X}(T)^\top|^{p_0}|\bar{h}_{xx}(T)|^{p_0}\Big\}\\
&\leq K\Big[\mathbb{E}\int_{\tau+\varepsilon}^T\big[|Y_2(s)|+|\mathbb{X}(s)|\big]^{4p_0^*}\mathrm{d}s\Big]^{\frac{2-p_0}{4}}\Big[\mathbb{E}\int_{\tau+\varepsilon}^T|Y_2(s)-\mathbb{X}(s)|^{4p_0^*}\mathrm{d}s\Big]^{\frac{2-p_0}{4}}\\
&\quad+\Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|\bar{H}_{xx}(s)|^2\mathrm{d}s\Big]^{\frac{p_0}{2}}\cdot\Big[\mathbb{E}\Big(\int_\tau^{\tau+\varepsilon}|\mathbb{X}(s)|^4\mathrm{d}s\Big)^{p_0^*}\Big]^{\frac{2-p_0}{2}}\\
&\quad+K\Big[\mathbb{E}|\bar{h}_{xx}(T)|^2\Big]^{\frac{p_0}{2}}\Big[\mathbb{E}\big[|Y_2(T)|+|\mathbb{X}(T)|\big]^{4p_0^*}\mathbb{E}|Y_2(T)-\mathbb{X}(T)|^{4p_0^*}\Big]^{\frac{2-p_0}{4}}\to 0.
\end{aligned}\tag{3.31}$$

In the above, $p_0^* := \frac{p_0}{2-p_0} > 1$, and we use the result of

$$\begin{aligned}
&\sup_{r\in[\tau+\varepsilon,T]}\mathbb{E}|Y_2(r)-\mathbb{X}(r)|^{4+\kappa}\\
&\leq K\sup_{t\in[\tau+\varepsilon,T]}\mathbb{E}\Big|\int_\tau^{\tau+\varepsilon}\bar{b}_x(t,s)\mathbb{X}(s)\mathrm{d}s+\int_\tau^{\tau+\varepsilon}\bar{\sigma}_x(t,s)\mathbb{X}(s)\mathrm{d}W(s)\Big|^{4+\kappa}\to 0,\quad \varepsilon\to 0.
\end{aligned}$$

Our remaining aim is $\lim\limits_{\varepsilon\to 0}\mathbb{E}\big|\mathbb{E}_{\tau+\varepsilon}\Theta_2(\tau)-\mathbb{E}_\tau\Theta_2(\tau)\big|^{p_0}=0$. By Lemma 2.1

$$\begin{aligned}
\mathbb{E}|\Theta_2(\tau)|^{p_0}&\leq K\mathbb{E}\Big[\int_\tau^T|\mathbb{X}(s)|^2|\bar{H}_{xx}(s)|\mathrm{d}s\Big]^{p_0}+K\mathbb{E}\Big[|\mathbb{X}(T)|^{2p_0}|\bar{h}_{xx}(T)|^{p_0}\Big]\\
&\leq K\Big[\mathbb{E}\int_\tau^T|\delta\bar{\sigma}(s,\tau)|^{\frac{4p_0}{2-p_0}}\mathrm{d}s\Big]^{\frac{2-p_0}{2}}+K\Big[\mathbb{E}|\delta\bar{\sigma}(T,\tau)|^{\frac{4p}{2-p_0}}\Big]^{\frac{2-p_0}{2}}<\infty.
\end{aligned}$$

Because $\mathbb{E}_{\tau+\varepsilon}\Theta_2(\tau)\to\mathbb{E}_\tau\Theta_2(\tau)$, a.s., $\varepsilon\to 0$, and for any $r\in[\tau,T]$,

$$\mathbb{E}|\mathbb{E}_r\Theta_2(\tau)|^{p_0}\leq\mathbb{E}\sup_{r\in[\tau,T]}\mathbb{E}_r|\Theta_2(\tau)|^{p_0}\leq\frac{p_0}{p_0-1}\mathbb{E}|\Theta_2(\tau)|^{p_0}<\infty.$$

One has the desired conclusion by dominated convergence theorem. □

Now it is time for us to show the proof of Lemma 3.3.

Proof. By above (3.24), (3.27), one has,

$$\begin{aligned}
&\frac{1}{\varepsilon}\Big[\mathbb{E}\int_0^T X_1(s)^\top\bar{H}_{xx}(s)X_1(s)\mathrm{d}s+\mathbb{E}\big[X_1(T)^\top\bar{h}_{xx}(T)X_1(T)\big]\Big]\\
&=\frac{1}{\varepsilon}\mathbb{E}\int_\tau^{\tau+\varepsilon}X_1(s)^\top\bar{H}_{xx}(s)X_1(s)\mathrm{d}s+\mathbb{H}(\varepsilon,X_1).
\end{aligned}$$

For the first term on right hand, denoted by $\mathcal{Q}(\varepsilon, X_1)$, from Lemma 3.1 we see that, $\lim_{\varepsilon\to 0}\mathcal{Q}(\varepsilon, X_1)=0$. We thus derive the conclusion by Lemma 3.4 and Lemma 3.5. □

3.3. Representations of some quadratic functionals

In this part, as a continuation of Lemma 3.3, we represent the above $F^{\delta\bar\sigma,\delta\bar\sigma}(\tau)$ by duality ideas.

First we discuss a special case of $\delta\bar\sigma(\cdot,\tau)\equiv\delta\bar\sigma(\tau)$. A sufficient condition to verify this assumption is given in Remark 3.13 later. For $\bar h_{xx}(T)$, $\bar H_{xx}(\cdot)$, we define $\lambda(\cdot):=\mathbb{E}_\cdot\big[\bar h_{xx}(T)\big]$, $\Xi(\cdot,\cdot):=\mathbb{E}_\cdot\big[\bar H_{xx}(\cdot)\big]$, which satisfy

$$\begin{cases}\bar h_{xx}(T)=\lambda(r)+\displaystyle\int_r^T\zeta_{\bar h}(s)\mathrm{d}W(s),\\ \bar H_{xx}(s)=\Xi(r,s)+\displaystyle\int_r^s\Pi_{\bar H}(s,\tau)\mathrm{d}W(\tau),\quad s\in[\tau,T].\end{cases}\tag{3.32}$$

With unit matrix $I_{n\times n}$, for any $t\in[\tau,T]$, $r\in[\tau,t]$, we introduce

$$\begin{cases}\mathscr{X}(t)=I_{n\times n}+\displaystyle\int_\tau^t\bar b_x(t,s)\mathscr{X}(s)\mathrm{d}s+\int_\tau^t\bar\sigma_x(t,s)\mathscr{X}(s)\mathrm{d}W(s),\\ \mathcal{X}(r,t)=I_{n\times n}+\displaystyle\int_\tau^r\bar b_x(t,s)\mathscr{X}(s)\mathrm{d}s+\int_\tau^r\bar\sigma_x(t,s)\mathscr{X}(s)\mathrm{d}W(s).\end{cases}\tag{3.33}$$

Due to (H1) and Lemma 2.1, there exists a unique $\mathscr{X}\in C_{\mathbb{F}}(0,T;L^{4+\kappa}(\Omega;\mathbb{R}^n))$. With above notations, we introduce process $\mathscr{P}$ as follows,

$$\mathscr{P}(\tau):=\mathbb{E}_\tau\int_\tau^T\bar H_{xx}(r)\mathrm{d}r+\lambda(\tau)+\mathbb{E}_\tau\int_\tau^T\Big[\mathscr{M}_2(r)+\int_r^T\mathscr{M}_1(s,r)\mathrm{d}s\Big]\mathrm{d}r,\tag{3.34}$$

where

$$\begin{cases}\mathscr{M}_1(r,s):=\mathcal{X}(r,s)^\top\Xi(r,s)\bar b_x(s,r)\mathscr{X}(r)+\mathscr{X}^\top(r)\bar b_x^\top(s,r)\Xi(r,s)\mathcal{X}(r,s)\\ \qquad\qquad+\mathcal{X}(r,s)^\top\Pi_{\bar H}(s,r)\bar\sigma_x(s,r)\mathscr{X}(r)+\mathscr{X}^\top(r)\bar\sigma_x^\top(s,r)\Pi_{\bar H}(s,r)\mathcal{X}(r,s)\\ \qquad\qquad+\mathscr{X}^\top(r)\bar\sigma_x^\top(s,r)\Xi(r,s)\bar\sigma_x(s,r)\mathscr{X}(r),\\ \mathscr{M}_2(r):=\mathcal{X}(r,T)^\top\lambda(r)\bar b_x(T,r)\mathscr{X}(r)+\mathscr{X}^\top(r)\bar b_x^\top(T,r)\lambda(r)\mathcal{X}(r,T)\\ \qquad\qquad+\mathcal{X}(r,T)^\top\zeta_{\bar h}(r)\bar\sigma_x(T,r)\mathscr{X}(r)+\mathscr{X}^\top(r)\bar\sigma_x^\top(T,r)\zeta_{\bar h}(r)\mathcal{X}(r,T)\\ \qquad\qquad+\mathscr{X}^\top(r)\bar\sigma_x^\top(T,r)\lambda(r)\bar\sigma_x(T,r)\mathscr{X}(r).\end{cases}$$

Notice that $\mathscr{P}$ is continuous and $\mathbb{F}$-adapted, and depends on optimal pair $(\bar X,\bar u)$.

Lemma 3.6. *Suppose (H1), (H2) hold true, $\delta\bar\sigma(\cdot,\tau)\equiv\delta\bar\sigma(\tau)$, $(\bar X(\cdot),\bar u(\cdot))$ is optimal. Then for $\mathscr{P}(\cdot)$ in (3.34), one has*

$$F^{\delta\bar\sigma,\delta\bar\sigma}(\tau)=\delta\bar\sigma(\tau)^\top\mathscr{P}(\tau)\delta\bar\sigma(\tau),\quad a.s.\quad\tau\in[0,T].\tag{3.35}$$

Proof. Recall that $\Xi(\cdot,\cdot):=\mathbb{E}_\cdot\big[\bar H_{xx}(\cdot)\big]$. By Itô's formula, we have

$$d\big[\mathcal{X}(r,s)^\top\Xi(r,s)\mathcal{X}(r,s)\big]=\mathscr{N}(s,r)\mathrm{d}W(r)+\mathscr{M}_1(r,s)\mathrm{d}r,\tag{3.36}$$

where

$$\mathscr{N}(r,s) := \Big[\mathcal{X}(r,s)^\top \Pi_{\bar{H}}(s,r)\mathcal{X}(r,s) + \mathcal{X}(r,s)^\top \Xi(r,s)\bar{\sigma}_x(s,r)\mathscr{X}(r) \\ + \mathscr{X}^\top(r)\bar{\sigma}_x^\top(s,r)\Xi(r,s)\mathcal{X}(r,s)\Big].$$

Therefore,

$$\mathbb{E}_\tau \int_\tau^T \big[\mathscr{X}(s)^\top \bar{H}_{xx}(s)\mathscr{X}(s)\big]\mathrm{d}s - \mathbb{E}_\tau \int_\tau^T \bar{H}_{xx}(s)\mathrm{d}s = \mathbb{E}_\tau \int_\tau^T \int_r^T \mathscr{M}_1(s,r)\mathrm{d}s\mathrm{d}r.$$

Similarly, one has

$$\mathbb{E}_\tau \Big[\mathscr{X}(T)^\top \bar{h}_{xx}(T)\mathscr{X}(T)\Big] - \lambda(\tau) = \mathbb{E}_\tau \int_\tau^T \mathscr{M}_2(r)\mathrm{d}r.$$

To sum up,

$$\mathbb{E}_\tau \int_\tau^T \big[\mathscr{X}(s)^\top \bar{H}_{xx}(s)\mathscr{X}(s)\big]\mathrm{d}s + \mathbb{E}_\tau \Big[\mathscr{X}(T)^\top \bar{h}_{xx}(T)\mathscr{X}(T)\Big] = \mathscr{P}(\tau).$$

Since $\mathbb{X}(t) = \mathscr{X}(t)\delta\bar{\sigma}(\tau)$, $\mathbb{X}(T) = \mathscr{X}(T)\delta\bar{\sigma}(\tau)$, it is easy to see the conclusion. □

Remark 3.7. Suppose b, σ are independent of t. Then $\mathscr{X}(\cdot) = \mathcal{X}(\cdot, t)$, and

$$\begin{cases} \mathscr{M}_1(r,s) := \mathscr{X}(r)^\top \Big\{\Xi(r,s)\bar{b}_x(r) + \bar{b}_x^\top(r)\Xi(r,s) + \Pi_{\bar{H}}(s,r)\bar{\sigma}_x(r) \\ \qquad + \bar{\sigma}_x^\top(r)\Pi_{\bar{H}}(s,r) + \bar{\sigma}_x^\top(r)\Xi(r,s)\bar{\sigma}_x(r)\mathrm{d}s\Big\}\mathscr{X}(r), \\ \mathscr{M}_2(r) := \mathscr{X}(r)^\top \Big\{\bar{\sigma}_x^\top(r)\lambda(r)\bar{\sigma}_x(r) + \lambda(r)\bar{b}_x(r) + \bar{b}_x^\top(r)\lambda(r) \\ \qquad + \zeta_{\bar{h}}(r)\bar{\sigma}_x(r) + \bar{\sigma}_x^\top(r)\zeta_{\bar{h}}(r)\Big\}\mathscr{X}(r). \end{cases}$$

If we define

$$P_1(r) := \lambda(r) + \int_r^T \Xi(r,s)\mathrm{d}s, \quad Q_1(r) := \zeta_{\bar{h}}(r) + \int_r^T \Pi_{\bar{H}}(s,r)\mathrm{d}s,$$

then by (1.9) and Lemma 3.6, we obtain a new representation of $P_2(\cdot)$ in (1.7) as follows,

$$P_2(\tau) = P_1(\tau) + \mathbb{E}_\tau \int_\tau^T \mathscr{X}(r)^\top \Big\{\bar{\sigma}_x(r)^\top P_1(r)\bar{\sigma}_x(r) + P_1(r)\bar{b}_x(r) \\ + \bar{b}_x(r)^\top P_1(r) + Q_1(r)\bar{\sigma}_x(r) + \bar{\sigma}_x(r)^\top Q_1(r)\Big\}\mathscr{X}(r)\mathrm{d}r. \tag{3.37}$$

In a nutshell, in the SDEs scenario, we have obtained two equivalent ways, *i.e.* equation (1.7) and representation (3.37), to describe $P_2(\cdot)$ that appears in maximum principle.

Next we treat the general case when assumption $\delta\bar\sigma(\cdot,\tau)\equiv\delta\bar\sigma(\tau)$ is dropped.

Lemma 3.8. *Suppose (H1), (H2) hold true, $(\bar X(\cdot),\bar u(\cdot))$ is optimal. Then*

$$F^{\delta\bar\sigma,\delta\bar\sigma}(\tau)=\delta\bar\sigma(T,\tau)^\top\big(\mathcal{B}_1(\tau)\Delta\bar\sigma(\cdot,\tau)\big)+\int_0^T\delta\bar\sigma(s,\tau)^\top\big(\mathcal{B}_2(\tau)\Delta\bar\sigma(\cdot,\tau)\big)(s)\mathrm{d}s, \tag{3.38}$$

where $\mathcal{B}_1\in\mathscr{L}_2$ and $\mathcal{B}_2\in\mathscr{L}_1$ satisfy (see (2.1) for the definitions of $\mathscr{L}_i$),

$$\begin{aligned}&\Big[\|\mathcal{B}_1(\tau)\|_{\mathcal{L}(\mathbb{B};\mathbb{R}^n)}+\|\mathcal{B}_2(\tau)\|_{\mathcal{L}(\mathbb{B};L^{\frac43}(0,T;\mathbb{R}^n))}\Big]\\&\le K\Big[\mathbb{E}_\tau\int_\tau^T|\bar H_{xx}(s)|^2\mathrm{d}s\Big]^{\frac12}+K\Big[\mathbb{E}_\tau|\bar h_{xx}(T)|^2\Big]^{\frac12},\quad a.s.\quad \tau\in[0,T].\end{aligned} \tag{3.39}$$

If there exists $(\widetilde{\mathcal{B}}_1,\widetilde{\mathcal{B}}_2)$ satisfying (3.38), (3.39), then

$$\mathbb{P}\big(\omega\in\Omega;\ \widetilde{\mathcal{B}}_1(\tau,\omega)=\mathcal{B}_1(\tau,\omega)\big)=\mathbb{P}\big(\omega\in\Omega;\ \widetilde{\mathcal{B}}_2(\tau,\omega)=\mathcal{B}_2(\tau,\omega)\big)=1. \tag{3.40}$$

Remark 3.9. For fixed τ, $\eta:=(\eta_1,\eta_2(\cdot))\in L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})$, if we define

$$\begin{cases}\Big[\mathcal{B}_1(\tau)\eta\Big]:=\Big\{\lambda(\tau)+\mathbb{E}_\tau\int_\tau^T\mathscr{M}_2(r)\mathrm{d}r\Big\}\eta_1,\\ \Big[\mathcal{B}_2(\tau)\eta\Big](\cdot):=\Big\{\mathbb{E}_\tau\bar H_{xx}(\cdot)+\mathbb{E}_\tau\int_\cdot^T\mathscr{M}_1(s,\cdot)\mathrm{d}s\Big\}\eta_1 I_{[\tau,T]}(\cdot),\end{cases}$$

then it is easy to $\mathcal{B}_1\in\mathscr{L}_2$, $\mathcal{B}_2\in\mathscr{L}_1$. In other words, (3.35) is indeed a special case of (3.38).

To prove Lemma 3.8, for $\alpha:=(\alpha_1,\alpha_2(\cdot))\in\mathbb{B}$, $\tau\in[0,T]$, $t\in[\tau,T]$, we consider

$$\begin{cases}X^\alpha(t)=\alpha_2(t)+\int_\tau^t A(t,s)X^\alpha(s)\mathrm{d}s+\int_\tau^t B(t,s)X^\alpha(s)\mathrm{d}W(s),\qquad a.e.\\ \mathcal{X}^\alpha(T)=\alpha_1+\int_\tau^T A(T,s)X^\alpha(s)\mathrm{d}s+\int_\tau^T B(T,s)X^\alpha(s)\mathrm{d}W(s).\end{cases} \tag{3.41}$$

Moreover, for $\bar\alpha,\ \widetilde\alpha\in\mathbb{B}$, $\tau\in[0,T]$, we define

$$f_1^{\bar\alpha,\widetilde\alpha}(\tau):=\mathbb{E}_\tau\int_\tau^T X^{\bar\alpha}(s)^\top Q(s)X^{\widetilde\alpha}(s)\mathrm{d}s+\mathbb{E}_\tau\big[\mathcal{X}^{\bar\alpha}(T)^\top G\mathcal{X}^{\widetilde\alpha}(T)\big].\ \ a.s. \tag{3.42}$$

(H3) $Q\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^{n\times n})$, $G\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^{n\times n})$, $A,\ B:[0,T]^2\times\Omega\mapsto\mathbb{R}^{n\times n}$ are bounded measurable processes such that for $t\in[0,T]$, $s\mapsto A(t,s),\ B(t,s)$ are $\mathbb{F}$-adapted, and with modulus function $\rho(\cdot)$,

$$\big[|A(t,s)|+|B(t,s)|\big]\le K,\quad |A(t,s)-A(t',s)|+|B(t,s)-B(t',s)|\le\rho(|t-t'|).$$

Remark 3.10. Under (H3), (3.41) is solvable with $X^\alpha(\cdot)\in L^4_{\mathbb{F}}(\tau,T;\mathbb{R}^n)$, $\mathcal{X}^\alpha(T)\in L^4_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$. It is meaningless to discuss $X^\alpha(T)$ if $\alpha_2(\cdot)\in L^4(0,T;\mathbb{R}^n)$. Hence we introduce $\mathcal{X}^\alpha(T)$ and $\alpha_1\in\mathbb{R}^n$ in (3.41). The appearance of both $a_2(\cdot)$ and α_1 explains the introducing of space $\mathbb{B}$.

In the following, for $\tau\in[0,T]$, we denote $L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})$ the set of $\mathcal{F}_\tau$-strongly measurable $\mathbb{B}$-valued random variable ξ such that $\mathbb{E}\|\xi\|^4_{\mathbb{B}}<\infty$, and

$$M^{Q,G}(\tau):=\Big[\mathbb{E}_\tau\int_\tau^T|Q(s)|^2\mathrm{d}s\Big]^{\frac{1}{2}}+\Big[\mathbb{E}_\tau|G|^2\Big]^{\frac{1}{2}},\quad a.s.\ \ \forall\tau\in[0,T]. \tag{3.43}$$

Lemma 3.11. *Suppose (H3) holds. Then for any $\tau\in[0,T]$, $\xi:=(\xi_1,\xi_2(\cdot)),\ \eta:=(\eta_1,\eta_2(\cdot))\in L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})$,*

$$f_1^{\xi,\eta}(\tau)=\xi_1^\top\big[\mathcal{B}_1(\tau)\eta\big]+\int_0^T\xi_2(s)^\top\big[\mathcal{B}_2(\tau)\eta\big](s)\mathrm{d}s,\qquad a.s. \tag{3.44}$$

where $\mathcal{B}_1\in\mathscr{L}_2$, $\mathcal{B}_2\in\mathscr{L}_1$, and for $(\tau,\omega)\in[0,T]\times\Omega$,

$$\Big[\|\mathcal{B}_1(\tau,\omega)\|_{\mathcal{L}(\mathbb{B};\mathbb{R}^n)}+\|\mathcal{B}_2(\tau,\omega)\|_{\mathcal{L}(\mathbb{B};L^{\frac{4}{3}}(0,T;\mathbb{R}^n))}\Big]\le KM^{Q,G}(\tau,\omega). \tag{3.45}$$

The uniqueness also holds in the similar spirit of (3.40).

Proof. For reader's convenience, we separate the proof into two steps.

Step 1: We prove the above conclusion when $\xi,\ \eta\in\mathbb{B}$.

At first, we prove the existence of $\mathcal{B}_1\in\mathscr{L}_1$, $\mathcal{B}_2\in\mathscr{L}_2$. For any $t\in[\tau,T]$, $\bar\alpha,\ \widetilde\alpha\in\mathbb{B}$, $M^{Q,G}(\cdot)$ in (3.43), it follows from (3.41) and Gronwall inequality that

$$|f_1^{\bar\alpha,\widetilde\alpha}(\tau)|\le KM^{Q,G}(\tau)\|\bar\alpha\|_{\mathbb{B}}\|\widetilde\alpha\|_{\mathbb{B}}.\qquad a.s. \tag{3.46}$$

For $\bar\alpha^i,\widetilde\alpha^i,\bar\alpha,\widetilde\alpha\in\mathbb{B}_1$, $k,l\in\mathbb{Q}$, we define $N\equiv N(\bar\alpha^i,\ \widetilde\alpha^i,\bar\alpha,\widetilde\alpha,k,l,Q,G)$ and $\mathcal{N}$ as,

$$\begin{cases} N:=\Big\{(\tau,\omega)\in[0,T]\times\Omega;\,|f_1^{\bar\alpha,\widetilde\alpha}(\tau,\omega)|\le KM^{Q,G}(\tau,\omega)\|\bar\alpha\|_{\mathbb{B}}\cdot\|\widetilde\alpha\|_{\mathbb{B}},\\ \qquad\qquad f_1^{k\bar\alpha^1+l\bar\alpha^2,\widetilde\alpha}(\tau,\omega)=kf_1^{\bar\alpha^1,\widetilde\alpha}(\tau,\omega)+lf_1^{\bar\alpha^2,\widetilde\alpha}(\tau,\omega),\\ \qquad\qquad f_1^{\bar\alpha,k\widetilde\alpha^1+l\widetilde\alpha^2}(\tau,\omega)=kf_1^{\bar\alpha,\widetilde\alpha^1}(\tau,\omega)+lf_1^{\bar\alpha,\widetilde\alpha^2}(\tau,\omega)\Big\},\\ \mathcal{N}\equiv\mathcal{N}(G,Q):=\bigcap\limits_{\bar\alpha,\widetilde\alpha\in\mathbb{B}_1}\ \bigcap\limits_{\bar\alpha^i,\widetilde\alpha^i\in\mathbb{B}_1}\ \bigcap\limits_{k,l\in\mathbb{Q}}N(\bar\alpha^i,\ \widetilde\alpha^i,\bar\alpha,\widetilde\alpha,k,l,Q,G).\end{cases} \tag{3.47}$$

Notice that $\big[\lambda\times\mathbb{P}\big](N)=T$, $\big[\lambda\times\mathbb{P}\big](\mathcal{N})=T$, where λ is the Lebesgue measure. In addition, by inequality (3.46), for any $\tau\in[0,T]$, one has $\mathbb{P}(\mathcal{N}_\tau)=1$ with $\mathcal{N}_\tau:=\big\{\omega\in\Omega;\ (\tau,\omega)\in\mathcal{N}\big\}$. For any $(\tau,\omega)\in\mathcal{N}$, it is easy to see that $f_1^{\cdot,\cdot}(\tau,\omega):\mathbb{B}_1\times\mathbb{B}_1\mapsto\mathbb{R}$ is a bounded bilinear map in the sense of (3.47). According to Lemma A.3, there exists a unique linear bounded functional $\mathcal{B}_{1,1}(\tau,\omega):\mathbb{B}\mapsto\mathbb{R}^n$ and a unique linear bounded operator $\mathcal{B}_{1,2}(\tau,\omega):\mathbb{B}\mapsto L^{\frac{4}{3}}(0,T;\mathbb{R}^n)$ such that for $\alpha,\ \beta\in\mathbb{B}_1$,

$$f_1^{\alpha,\beta}(\tau,\omega)=\alpha_1^\top\big(\mathcal{B}_{1,1}(\tau,\omega)\beta\big)+\int_0^T\alpha_2(t)^\top\big(\mathcal{B}_{1,2}(\tau,\omega)\beta\big)(t)\mathrm{d}t. \tag{3.48}$$

For any $\tau\in[0,T]$, recalling $\mathbb{P}(\mathcal{N}_\tau)=1$, we know that (3.48) holds almost surely.

Given $\alpha:=(\alpha_1,\alpha_2(\cdot)),\beta:=(\beta_1,\beta_2(\cdot))\in\mathbb{B}$, and any $(\tau,\omega)\in\mathcal{N}$, let

$$\mathscr{B}(\tau,\omega):=\alpha_1^\top[\mathcal{B}_{1,1}(\tau,\omega)\beta]+\int_0^T\alpha_2(t)^\top\big(\mathcal{B}_{1,2}(\tau,\omega)\beta\big)(t)\mathrm{d}t.$$

By (3.48), as well as the denseness of $\mathbb{B}_1$ in $\mathbb{B}$, we see that $\mathscr{B}(\cdot)$ is measurable, adapted process. At this moment, we define two processes $\mathcal{B}_i(\tau,\omega)$ on $[0,T]\times\Omega$ as

$$\mathcal{B}_1(\tau,\omega):=\mathcal{B}_{1,1}(\tau,\omega)I_{\mathcal{N}}(\tau,\omega),\ \ \mathcal{B}_2(\tau,\omega):=\mathcal{B}_{1,2}(\tau,\omega)I_{\mathcal{N}}(\tau,\omega). \tag{3.49}$$

For any $\alpha_1\in\mathbb{R}^n$, by choosing $\alpha:=(\alpha_1,0)$, it follows from (3.49) that $\alpha_1^\top\big[\mathcal{B}_1(\tau,\omega)\beta\big]=\mathscr{B}(\tau,\omega)I_{\mathcal{N}}(\tau,\omega)$. It is then easy to see the measurability of $(\tau,\omega)\mapsto\alpha_1^\top\big(\mathcal{B}_1(\tau,\omega)\beta\big)$, as well as the adaptness. Similarly one can obtain the case of $\mathcal{B}_2(\cdot,\cdot)$.

For any $\tau\in[0,T]$, we prove (3.44) when $\xi,\ \eta\in\mathbb{B}$. In fact, for any $\tau\in[0,T]$ and $\omega\in\mathcal{N}_\tau$, (3.48) holds true with any $\alpha_n,\ \beta_n\in\mathbb{B}_1$. Moreover,

$$\mathbb{E}_\tau\int_\tau^T|X^{f_n}(t)-X^f(t)|^4\mathrm{d}t+\mathbb{E}_\tau|\mathcal{X}^{f_n}(T)-\mathcal{X}^f(T)|^4\leq K\|f_n-f\|_{\mathbb{B}},\quad f:=\alpha,\ \beta.$$

As a result, $\lim\limits_{n\to\infty}\big|f_1^{\alpha,\beta}(\tau)-f_1^{\alpha_n,\beta_n}(\tau)\big|=0$. Consequently,

$$f_1^{\alpha,\beta}(\tau)=\mathscr{B}(\tau):=\alpha_1^\top\big(\mathcal{B}_{1,1}(\tau)\beta\big)+\int_0^T\alpha_2(t)^\top\big(\mathcal{B}_{1,2}(\tau)\beta\big)(t)\mathrm{d}t.\ \ a.s.$$

Our conclusion then follows from the relation between $\mathcal{B}_{1,i}(\cdot,\cdot)$ and $\mathcal{B}_i(\cdot,\cdot)$ in (3.49).

Eventually, we discuss the integrability and uniqueness of $\mathcal{B}_i$. For $(t,\omega)\in\mathcal{N}$, by Lemma A.3, we have

$$\Big[\|\mathcal{B}_{1,1}(\tau,\omega)\|_{\mathcal{L}(\mathbb{B};\mathbb{R}^n)}+\|\mathcal{B}_{1,2}(\tau,\omega)\|_{\mathcal{L}(\mathbb{B};L^{\frac{4}{3}}(0,T;\mathbb{R}^n))}\Big]\leq KM^{Q,G}(\tau,\omega).$$

Hence (3.45) is lead by (3.49). From (3.46), $\sup\limits_{\tau\in[0,T]}\mathbb{E}|f_1^{\alpha,\beta}(\tau)|^2<\infty$. Hence, let $\alpha:=(\alpha_1,0)\in\mathbb{B}$, we have $\sup\limits_{\tau\in[0,T]}\mathbb{E}\big|\alpha_1^\top\big(\mathcal{B}_1(\tau)\beta\big)\big|^2<\infty$. Similarly, we derive the case of $\mathcal{B}_2(\cdot)$. The uniqueness of $\mathcal{B}_i(\cdot)$ is obvious to obtain.

Step 2. We prove (3.44) when $\xi,\ \eta\in L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})$.

We begin with the simple random variable case where

$$\xi(\omega):=\sum_{i=1}^n x_iI_{A_i}(\omega),\quad\eta(\omega):=\sum_{j=1}^m y_jI_{B_j}(\omega),\quad\omega\in\Omega,\quad A_i,\ B_j\in\mathcal{F}_\tau,\quad x_i,\ y_j\in\mathbb{B}.$$

From Lemma 3.11, as well as the definitions of ξ, η, we conclude that,

$$f_1^{\xi,\eta}(\tau)=\sum_{i=1}^n\sum_{j=1}^m\Big(\mathbb{E}_\tau\int_\tau^T X^{x_i}(s)^\top Q(s)X^{x_j}(s)\mathrm{d}s+\mathbb{E}_\tau\big[\mathcal{X}^{x_i}(T)^\top G\mathcal{X}^{x_j}(T)\big]\Big)\cdot I_{A_i}I_{B_j}.$$

For the general case of $\xi,\ \eta \in L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})$, there exist $\{\xi_n\}_{n\geq 1},\ \{\eta_n\}_{n\geq 1}$ such that $\|k-k_n\|^4_{L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})} \to 0$, $n\to\infty$, $k:=\xi,\ \eta$. Therefore, when $n\to\infty$,

$$\mathbb{E}\int_\tau^T |X^{k_n}(t)-X^k(t)|^4\mathrm{d}t + \mathbb{E}|\mathcal{X}^{k_n}(T)-\mathcal{X}^k(T)|^4 \leq K\|k-k_n\|^4_{L^4_{\mathcal{F}_\tau}(\Omega;\mathbb{B})} \to 0,$$

$k:=\xi,\ \eta$, which implies that for any $\tau\in[0,T]$, $\lim\limits_{n\to\infty}\mathbb{E}|f_1^{\xi,\eta}(\tau)-f_1^{\xi_n,\eta_n}(\tau)|=0$.

On the other hand, by the estimates in Lemma 3.11, and the above conclusion of the simple random variables case,

$$\mathbb{E}\Big|f_1^{\xi,\eta}(\tau)-\xi_1^\top[\mathcal{B}_1(\tau)\eta]-\int_0^T \xi_2(t)^\top\big(\mathcal{B}_2(\tau)\eta\big)(t)\mathrm{d}t\Big|=0.$$

which obviously indicates the conclusion. □

Now we give the proof of Lemma 3.8.

Proof. Given $\delta\bar\sigma(\cdot,\tau),\ \Delta\bar\sigma(\cdot,\tau)$ in (3.21), $\mathbb{X}(\cdot)$ in (3.22), like (3.41) we introduce

$$\mathcal{X}^{\bar\sigma}(T):=\delta\bar\sigma(T,\tau)+\int_\tau^T \bar b_x(T,s)\mathbb{X}(s)\mathrm{d}s+\int_\tau^T\bar\sigma_x(T,s)\mathbb{X}(s)\mathrm{d}W(s).$$

We also introduce $f_1^{\Delta\bar\sigma,\Delta\bar\sigma}(\tau)$

$$f_1^{\Delta\bar\sigma,\Delta\bar\sigma}(\tau):=\mathbb{E}_\tau\int_\tau^T\mathbb{X}(s)^\top\bar H_{xx}(s)\mathbb{X}(s)\mathrm{d}s+\mathbb{E}_\tau\big[\mathcal{X}^{\bar\sigma}(T)^\top\bar h_{xx}(T)\mathcal{X}^{\bar\sigma}(T)\big].$$

By virtue of Lemma 3.11, there exist $\mathcal{B}_1(\cdot,\cdot)$, $\mathcal{B}_2(\cdot,\cdot)$ satisfying (3.39) and

$$f_1^{\delta\bar\sigma,\delta\bar\sigma}(\tau)=\delta\bar\sigma(T,\tau)^\top\big(\mathcal{B}_1(\tau)\Delta\bar\sigma(\cdot,\tau)\big)+\int_0^T\delta\bar\sigma(s,\tau)^\top\big(\mathcal{B}_2(\tau)\Delta\bar\sigma(\cdot,\tau)\big)(s)\mathrm{d}s.$$

The integrability of $\mathbb{X}(\cdot)$ and $\delta\bar\sigma(\cdot,\tau)$ yield $\mathbb{E}_\tau\big|\mathcal{X}^{\bar\sigma}(T)-\mathbb{X}(T)\big|^4=0$, a.s., $\tau\in[0,T]$.
As a result, the conclusion (3.38) follows immediately. □

3.4. Maximum principles of optimal control problems for SVIEs

Given $\delta\bar\sigma(T,t)$, $\Delta\bar\sigma(\cdot,t)$ in (3.21), $\mathcal{B}_1(\cdot)$, $\mathcal{B}_2(\cdot)$ satisfy (3.39), (3.40), $\mathscr{P}(\cdot)$ in Lemma 3.6, $u\in U$, we define

$$\begin{aligned}
&\Delta H(t):=H(t,\bar X(t),\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot,t),u)-H(t,\bar X(t),\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot,t),\bar u(t)),\\
&\mathscr{H}(t,u):=\Delta H(t)+\frac12\delta\bar\sigma(T,t)^\top[\mathcal{B}_1(t)\Delta\bar\sigma(\cdot,t)]+\frac12\int_0^T\delta\bar\sigma(s,t)^\top\big[\mathcal{B}_2(t)\Delta\bar\sigma(\cdot,t)\big](s)\mathrm{d}s,\\
&\mathscr{H}_0(t,u):=\Delta H(t)+\frac12\delta\bar\sigma(t)^\top\mathscr{P}(t)\delta\bar\sigma(t).
\end{aligned}\tag{3.50}$$

We present the main result of this paper.

Theorem 3.12. *Let (H1)-(H2) hold and $(\bar{X}(\cdot), \bar{u}(\cdot))$ be an optimal pair. Then*

$$\min_{u\in U} \mathscr{H}(t,u) = \mathscr{H}(t,\bar{u}(t)) = 0, \quad \mathbb{P}-a.s., \quad t\in[0,T]. \quad a.e. \tag{3.51}$$

If $\delta\bar{\sigma}(\cdot,\tau)\equiv\delta\bar{\sigma}(\tau)$. Then we have

$$\min_{u\in U} \mathscr{H}_0(t,u) = \mathscr{H}_0(t,\bar{u}(t)) = 0, \quad \mathbb{P}-a.s., \quad t\in[0,T]. \quad a.e. \tag{3.52}$$

Proof. Given optimal control $\bar{u}(\cdot)$, we introduce the perturbation $u^\varepsilon(\cdot)$ as in (3.1), first-order variational equation (3.2), first-order adjoint equation (3.9), Hamiltonian function (3.10).

By the arguments from (3.7) to (3.19), and the optimality of $\bar{u}(\cdot)$, we have

$$\lim_{\varepsilon\to0}\Big\{\frac{1}{\varepsilon}\mathbb{E}\int_0^T \Delta H^\varepsilon(t)\mathrm{d}t + \mathscr{E}(\varepsilon)\Big\} \geq 0, \tag{3.53}$$

where $\Delta H^\varepsilon(\cdot)$, $\mathscr{E}(\varepsilon)$ are defined in (3.20). Now we treat the two terms on the left hand respectively.

The definition of ΔH^ε implies that, for almost $\tau\in[0,T]$,

$$\lim_{\varepsilon\to0}\frac{1}{\varepsilon}\mathbb{E}\int_0^T \Delta H^\varepsilon(t)\mathrm{d}t = \lim_{\varepsilon\to0}\frac{1}{\varepsilon}\mathbb{E}\int_\tau^{\tau+\varepsilon} \Delta H(t)\mathrm{d}t = \mathbb{E}\Delta H(\tau).$$

Here $\Delta H(\cdot)$ is defined in (3.50).

On the other hand, from Lemma 3.3 and notations in (3.22), there exists a $\{\varepsilon_n\}$ such that for almost $\tau\in[0,T]$, $\lim\limits_{\varepsilon_n\to0} 2\mathscr{E}(\varepsilon_n) = \mathbb{E}F^{\delta\bar{\sigma},\delta\bar{\sigma}}(\tau)$. Using Lemma 3.8, we have

$$\lim_{\varepsilon_n\to0}\mathscr{E}(\varepsilon_n) = \frac{1}{2}\mathbb{E}\Big\{\delta\bar{\sigma}(T,\tau)^\top\big(\mathcal{B}_1(\tau)\Delta\bar{\sigma}(\cdot,\tau)\big)+\int_0^T\delta\bar{\sigma}(s,\tau)^\top\big(\mathcal{B}_2(\tau)\Delta\bar{\sigma}(\cdot,\tau)\big)(s)\mathrm{d}s\Big\}, \tag{3.54}$$

To summarize, we have

$$\mathbb{E}\Big[\Delta H(\tau) + \frac{1}{2}\Big\{\delta\bar{\sigma}(T,\tau)^\top\big(\mathcal{B}_1(\tau)\Delta\bar{\sigma}(\cdot,\tau)\big)+\int_0^T\delta\bar{\sigma}(s,\tau)^\top\big(\mathcal{B}_2(\tau)\Delta\bar{\sigma}(\cdot,\tau)\big)(s)\mathrm{d}s\Big\}\Big] \geq 0.$$

The arbitrariness of $u\in U$ implies (3.51). The result of (3.52) can be obtained similarly. □

If we carefully look back at (3.51), it is indeed *pointwise* with respect to (t,ω). $\mathcal{B}_1$, $\mathcal{B}_2$ are called the *operator-valued* second-order adjoint process (SOAP) of Problem (SVIEs). Frankly speaking, we can not continue to derive the explicit forms of $\mathcal{B}_1$, $\mathcal{B}_2$ at this moment. However, the answer becomes positive if $\delta\bar{\sigma}(\cdot,\tau)$ degenerates into a random variable, see Remark 3.9. In this case, $\mathscr{P}(\cdot)$, which is shown explicitly in Lemma 3.6, is referred as the $\mathbb{R}^{n\times n}$-*valued* SOAP.

Remark 3.13. Let $\sigma(t,s,x,u) := \sigma_1(t,s,x) + \sigma_2(s,x,u)$, with proper σ_1, σ_2. In this case, for $\delta\bar{\sigma}(t,\tau)$ defined in (3.21), we have

$$\delta\bar{\sigma}(t,\tau) := \sigma(t,\tau,\bar{X}(\tau),u) - \sigma(t,\tau,\bar{X}(\tau),\bar{u}(\tau)) \equiv \delta\bar{\sigma}(\tau), \ u\in U. \tag{3.55}$$

Remark 3.14. We make some comments about the two cases in Theorem 3.12.

For the special case, we only need one SOAP $\mathscr{P}(\cdot)$ which is $\mathbb{R}^{n\times n}$-valued, continuous, and can be explicitly expressed in Lemma 3.6. The Itô's formula also works. Nevertheless, it is used with respect to $\mathcal{X}(\cdot,\cdot)$ in (3.33), but not the $\mathscr{X}(\cdot)$. This is totally different from the case of SDEs or stochastic evolution equations ([6, 8, 10, 13]). As a tradeoff, it is too special to cover the result in *e.g.*, Section 3.4.1 next.

For the general case, we need two SOAPs $\mathcal{B}_1(\cdot)$, $\mathcal{B}_2(\cdot)$ which, however, are operator-valued processes and implicit. The Itô's formula loses powerfulness. However, it covers the discussion in Section 3.4.1 later.

Remark 3.15. We point out some interesting aspects.

If there is no terminal cost in (1.2), according to the proof of Lemma 3.8, Lemma 3.11, it is unnecessary to introduce SOAP $\mathcal{B}_1(\cdot)$, and $\mathbb{B}$ reduces to $L^4(0,T;\mathbb{R}^n)$. However, if (1.2) only contains terminal term, then we may still need two SOAPs $\mathcal{B}_1(\cdot)$, $\mathcal{B}_2(\cdot)$ as the above.

If the diffusion satisfies $\delta\bar\sigma(\cdot,\tau)\equiv\delta\bar\sigma(\tau)$, then we only need one SOAP in the maximum condition. For the general case, we may have to rely on two SOAPs. Such necessity is obvious to see in Subsection 3.4.1.

If $\delta\bar\sigma(\cdot,\tau)\neq\delta\bar\sigma(\tau)$, and there is terminal cost $h(X(T))$ in (1.2), then according to the previous discussion, we need two SOAPs in the maximum conditions.

Remark 3.16. As to the maximum principle of SVIEs, we choose to use SOAP, the idea of which is similar to that in [6, 8, 10].

For the exploration of second-order adjoint equations (SOAE), at this moment it is quite challenging to show explicit forms along this line. We hope to present more details in future.

3.4.1. State-independent diffusion and drift terms

Suppose $b(\sigma)(t,s,x,u)$ does not depend on x. Then for $u\in U$, $\bar H_{xx}(t,x,\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot,t),u)$ is bounded. We define two operator-valued processes $\mathcal{B}_1'$, $\mathcal{B}_2'$, *i.e.* for $t\in[0,T]$, $s\in[0,T]$,

$$\mathcal{B}_1'(t)\Delta\bar\sigma(\cdot,t):=\mathbb{E}_t[\bar h_{xx}(T)]\delta\bar\sigma(T,t),\quad \big[\mathcal{B}_2'(t)\Delta\bar\sigma(\cdot,t)\big](s):=\mathbb{E}_t[\bar H_{xx}(s)]\delta\bar\sigma(s,t)I_{[t,T]}(s).$$

It is a direct calculation that $\mathcal{B}_1'$ and $\mathcal{B}_2'$ satisfy (3.39) and (3.38). By the uniqueness in (3.40) and Theorem 3.12, for any $u\in U$,

$$\Delta H(t)+\frac{1}{2}\delta\bar\sigma(T,t)^\top\mathbb{E}_t[\bar h_{xx}(T)]\delta\bar\sigma(T,t)+\frac{1}{2}\int_t^T\delta\bar\sigma(s,t)^\top\mathbb{E}_t\bar H_{xx}(s)\delta\bar\sigma(s,t)\mathrm{d}s\geq 0.\ a.s. \tag{3.56}$$

Corollary 3.17. *Suppose both $b(t,s,x,u)$ and $\sigma(t,s,x,u)$ do not rely on x, and $(\bar X(\cdot),\bar u(\cdot))$ is optimal pair. Then for almost $t\in[0,T]$, (3.56) holds true.*

Remark 3.18. If $b,\ \sigma:[0,T]\times\Omega\times U\to\mathbb{R}^n$, according to Remark 3.7, $\mathbb{E}.\bar h_{xx}(T)$ and $\mathbb{E}.\bar H_{xx}(\cdot)$ will merge into SOAP $P_2(\cdot)$. However, according to (3.56), in the SVIEs case, they are separated in general. This point explicitly explains the reasonability of two, rather than one, SOAPs in (3.51).

As to the optimal control problem of SDEs with state-independent diffusion and drift terms, we define one linear BSDE (1.7) as the second-order adjoint equation. However, for the analogue case of SVIEs, we have to use two BSDEs which are shown in (3.32). Obviously, there is no way to combine them together due to the dependence of $\delta\bar\sigma(t,\cdot)$ on t. Nevertheless, they can reduce to BSDE (1.7) with $\bar b_x=\bar\sigma_x=0$.

3.4.2. The convex control region

Suppose (H1), (H2) hold with convex U, l_u, b_u, σ_u are continuous. For $t\in[0,T]$, a.e. $\omega\in\Omega$, a.s., we define $v:=\bar{u}(t)+\varepsilon[u-\bar{u}(t)]$ with $u\in U$. The convexity of U shows that $v\in U$. We have,

$$0\leq\frac{\mathscr{H}(t,v)}{\varepsilon}\leq\langle\, H_u(t,\bar{X}(t),\bar{X}(\cdot),\bar{Y}(\cdot),\bar{Z}(\cdot,t),\bar{u}(t)+\theta\varepsilon(u-\bar{u}(t))),u-\bar{u}(t)\,\rangle+K(t)\varepsilon,$$

with some process $K(\cdot)$, $0<\theta<1$, partial derivative H_u. Let $\varepsilon\to0$, one has

$$\langle\, H_u(t,\bar{X}(t),\bar{X}(\cdot),\bar{Y}(\cdot),\bar{Z}(\cdot,t),\bar{u}(t)),u-\bar{u}(t)\,\rangle\geq0. \quad a.s. \tag{3.57}$$

Corollary 3.19. *Let (H1), (H2) hold and $(\bar{X}(\cdot),\bar{u}(\cdot))$ be optimal with convex U. Then there exists a pair of $(\bar{Y}(\cdot),\bar{Z}(\cdot,\cdot))$ satisfying (3.9) such that for almost $t\in[0,T]$, $u\in U$, (3.57) is satisfied.*

The above particular convex region case was studied in *e.g.* [19, 20]. Next we give one example to indicate some interesting role of second-order adjoint process.

Example 3.20. Let $n=m=1$, $\varphi=0$, $b=0$, $T=1$, $h(x)=x^2$, $\sigma(t,s,x,u)=D(t,s)u$, $l(s,x,u)=x^2+R(s)u^2$ with deterministic and bounded function $R(\cdot)$, $D(\cdot,\cdot)$. In this case,

$$\begin{aligned}
&H(t,x,\bar{X}(\cdot),\bar{Y}(\cdot),\bar{Z}(\cdot,t),u)=(x^2+R(t)u^2)+\Big[D(1,t)\bar{\pi}(t)+\mathbb{E}_t\int_t^1 D(s,t)\bar{Z}(s,t)\mathrm{d}s\Big]u,\\
&\bar{X}(t)=\int_0^t D(t,s)\bar{u}(s)\mathrm{d}W(s),\quad t\in[0,1],\quad \bar{X}(1)=\mathbb{E}\bar{X}(1)+\int_0^1\frac{\bar{\pi}(s)}{2}\mathrm{d}W(s),\\
&\bar{Y}(t)=2\bar{X}(t)+2R(t)\bar{u}(t),\quad \bar{Y}(t)=\mathbb{E}\bar{Y}(t)+\int_0^t\bar{Z}(t,s)\mathrm{d}W(s).
\end{aligned}$$

Moreover, if we plug $X(\cdot)$ into $J(u(\cdot))$, and define $\mathscr{D}(\cdot):=\int_\cdot^1|D(s,\cdot)|^2\mathrm{d}s+|D(1,\cdot)|^2$, then

$$J(u(\cdot))=\mathbb{E}\int_0^1\Big[\mathscr{D}(r)+R(r)\Big]u^2(r)\mathrm{d}r. \tag{3.58}$$

(i) Let $U=[-1,1]$, $R(r)=-\frac{1}{2}\mathscr{D}(r)\leq0$. Then $\bar{u}=0$ is an optimal control, $\bar{\pi}=0$, $\bar{Y}=0$, $\bar{Z}=0$. Since $R(\cdot)\leq0$, the following function is concave with respect to u,

$$H(t,\bar{X}(\cdot),\bar{X}(\cdot),\bar{Y}(\cdot),\bar{Z}(\cdot,t),u)=R(t)u^2.$$

Hence $\min_{u\in U}H(t,u)=H(t,\bar{u}(t))$ does not hold true. However, after a calculation, $\mathscr{H}(t,u)=\frac{1}{2}\mathscr{D}(t)u^2$. The convexity naturally leads to $\min_{u\in U}\mathscr{H}(t,u)=\mathscr{H}(t,\bar{u}(t))$. Thus the introducing of second-order adjoint processes turns the concave function $u\mapsto H(t,\bar{X}(\cdot),\bar{X}(\cdot),\bar{Y}(\cdot),\bar{Z}(\cdot,t),u)$ into convex function $u\mapsto\mathscr{H}(t,u)$.

(ii) Let $U=[0,\infty)$. According to (3.58), a sufficient condition of $\bar{u}=0$ being optimal is $R(\cdot)+\mathscr{D}(\cdot)\geq0$. Conversely, if $\bar{u}=0$ is optimal, the above (3.57), see also [20], is tri*via*lly true, and no further information can be obtained. Nevertheless, according to our study, $\mathscr{H}(t,u)=[R(t)+\mathscr{D}(t)]u^2$, and $R(\cdot)+\mathscr{D}(\cdot)\geq0$ becomes an optimality necessary condition.

3.4.3. The linear quadratic case.

We discuss the case when $U := \mathbb{R}^m$,

$$\begin{aligned} &b = [A_1(t,s)x + B_1(t,s)u], \quad \sigma = [A_2(t,s)x + B_2(t,s)u], \\ &h = \frac{1}{2}x^\top Gx, \quad l = \frac{1}{2}\big[x^\top Qx + 2u^\top Sx + u^\top Ru\big]. \end{aligned}$$

Here $A_1, B_1, A_2, B_2, Q, S, R, G$ are bounded and random such that

$$|f(t,s) - f(t',s)| \leq \rho(|t-t'|), \quad t, t', s \in [0,T], \quad f := A_1, B_1, A_2, B_2,$$

with some modulus function $\rho(\cdot)$. For optimal $(\bar{X}, \bar{u})$, we define

$$\begin{aligned} \mathscr{H}_1(t) := {} & S(t)\bar{X}(t) + B_1(T,t)^\top \mathbb{E}_t\big[G\bar{X}(T)\big] + \mathbb{E}_t \int_t^T B_1(s,t)^\top \bar{Y}(s)\mathrm{d}s \\ & + B_2(T,t)^\top \bar{\pi}(t) + \mathbb{E}_t \int_t^T B_2(s,t)^\top \bar{Z}(s,t)\mathrm{d}s, \end{aligned} \tag{3.59}$$

where $(\bar{Y}, \bar{Z}, \bar{\pi})$ is in (3.9) accordingly. Next we consider two special cases.

Case I: Suppose $A_1 \equiv A_2 \equiv 0$. For any $u \in \mathbb{R}^m$, $t \in [0,T]$, a.e., (3.56) becomes

$$[u - \bar{u}(t)]^\top \mathscr{H}_1(t) + \frac{1}{2}\big[u^\top R(t)u - \bar{u}(t)^\top R(t)\bar{u}(t)\big] + \frac{1}{2}[u - \bar{u}(t)]^\top \mathscr{B}(t)[u - \bar{u}(t)] \geq 0, \ \ a.s.$$

where $\mathscr{H}_1(t)$ is in (3.59), $\mathscr{B}(t) := \mathscr{B}_1(t) + \mathscr{B}_2(t)$,

$$\mathscr{B}_1(t) := B_2(T,t)^\top \big[\mathbb{E}_t G\big] B_2(T,t), \quad \mathscr{B}_2(t) := \mathbb{E}_t \int_t^T B_2(s,t)^\top Q(s) B_2(s,t)\mathrm{d}s.$$

Here $\mathscr{B}_1$, $\mathscr{B}_2$ are $\mathbb{F}$-adapted processes. By the arbitrariness of $u \in U$, one has

$$\mathscr{H}_1(t) + R(t)\bar{u}(t) = 0, \quad R(t) + \mathscr{B}(t) \geq 0, \quad t \in [0,T]. \ \ a.s. \ \ a.e.$$

Notice that these two condition were also obtained in [16] with distinct approach.

Case II: Suppose $B_2(t,\cdot) \equiv B_2(\cdot)$. Then $\delta\bar{\sigma}(\cdot) = B_2(\cdot)[u - \bar{u}(\cdot)]$. By (3.52),

$$\begin{aligned} &[u - \bar{u}(t)]^\top \mathscr{H}_1(t) + \frac{1}{2}\big[u^\top R(t)u - \bar{u}(t)^\top R(t)\bar{u}(t)\big] \\ &\qquad + \frac{1}{2}[u - \bar{u}(t)]^\top B_2(t)^\top \mathscr{P}(t) B_2(t)[u - \bar{u}(t)] \geq 0, \end{aligned}$$

where $u \in \mathbb{R}^m$, $\mathscr{H}_1(\cdot)$ is in (3.59), $\mathscr{P}(\cdot)$ is in (3.34) accordingly. By the arbitrariness of $u \in \mathbb{R}^m$, one then obtains the following maximum condition

$$\mathscr{H}_1(t) + R(t)\bar{u}(t) = 0, \quad R(t) + B_2(t)^\top \mathscr{P}(t) B_2(t) \geq 0, \quad t \in [0,T]. \ \ a.e. \tag{3.60}$$

If A_1, A_2, B_1 are further independent of t, then according to Remark 3.7, $\mathscr{P}(\cdot)$ satisfies equations (1.7) and (3.60) naturally reduce to the counterpart in SDEs case.

4. Concluding remarks

This article is devoted to general maximum principles of controlled SVIEs which is an open problem for more than a decade. Some novelties are summed up as follows.

- Inspired by the pioneering work [13] in SDEs setting, we establish the maximum principles by introducing appropriate second-order adjoint processes (SOAPs). Nevertheless, unlike [13] with *only one* SOAP, here *two* operator-valued SOAPs are actually required in the maximum conditions. Moreover, under proper conditions it is explicitly shown that just one SOAP is not sufficient, and two SOAPs are indeed necessary.
- With proper assumption on σ, the previous two operator-valued SOAPs reduce to one continuous, $\mathbb{F}$-adapted, matrix-valued SOAP with explicit expression. In the particular SDEs case, this SOAP becomes into the classical counterpart, the traditional maximum principle is fully recovered, and the developed approach is essentially different from the existing one.
- To derive the maximum principle, we develop some techniques which are new in the literature. For example, when dealing with optimal control problem of SVIEs, it is well-known that Itô's formula is not suitable and useful as in the SDEs setting. However, here by making proper transformation of involved SVIEs, we use Itô's formula from a new standpoint and obtain the explicit SOAP aforementioned.

To obtain matrix-valued SOAP, we need some extra requirement on σ, see *e.g.*, Remark 3.13. Consequently, it still remains its importance to drop such condition, and furthermore explore proper second-order adjoint equations. We hope to discuss these topics in our forthcoming papers.

Appendix A.

This section is devote to the existence of operator-valued processes in Lemma 3.11. The first two results are standard in functional analysis.

Lemma A.1. *Suppose $f:\mathbb{B}\mapsto\mathbb{R}$ is a bounded linear functional. Then there exists a unique $y(\cdot)=(y_1,y_2(\cdot))\in\mathbb{B}'$ such that*

$$
\begin{aligned}
&f(x)=\int_0^T x_2(t)^\top y_2(t)\mathrm{d}t+x_1^\top y_1, \qquad \forall x=(x_1,x_2(\cdot))\in\mathbb{B},\\
&\|f\|_{\mathcal{L}(\mathbb{B};\mathbb{R})}=\|y\|_{\mathbb{B}'}=\max\Big\{|y_1|,\Big[\int_0^T|y_2(t)|^{\frac{4}{3}}\mathrm{d}t\Big]^{\frac{3}{4}}\Big\}.
\end{aligned}
$$

Lemma A.2. *Given Banach space X with numerable dense subset X_0, suppose $X_1:=\Big\{\sum_{i=1}^n a_ix_i,\ a_i\in\mathbb{Q},\ x_i\in X_0,\ n\in\mathbb{N}\Big\}$ with $\mathbb{Q}$ the set of rational number, $f:X_1\mapsto Y$ is map to Banach space Y such that for constant $M>0$,*

$$
\|f(x)\|_Y\le M\|x\|_X,\quad \forall x\in X_1,\quad f(ax+by)=af(x)+bf(y),\quad \forall a,b\in\mathbb{Q}.
$$

Then there exists a unique bounded linear operator $F:X\mapsto Y$ satisfying

$$
F(x)=f(x),\quad x\in X_1,\quad \|F\|_{\mathcal{L}(X,Y)}\le M.
$$

Lemma A.3. *Given constant $M>0$, suppose $f:\mathbb{B}_1\times\mathbb{B}_1\mapsto\mathbb{R}$ satisfies*

$$
\begin{aligned}
&\|f(x,y)\|\le M\|x\|_{\mathbb{B}_1}\|y\|_{\mathbb{B}_1},\ \forall x,\ y\in\mathbb{B}_1,\ f(a\bar{x}+b\widetilde{x},y)=af(\bar{x},y)+bf(\widetilde{x},y),\\
&f(x,a\bar{y}+b\widetilde{y})=af(x,\bar{y})+bf(x,\widetilde{y}),\ a,b\in\mathbb{Q},\ x,\bar{x},\widetilde{x},y,\bar{y},\widetilde{y}\in\mathbb{B}_1.
\end{aligned}
\tag{A.1}
$$

Then there exist a unique pair of linear operators $\widehat{\mathcal{B}}_1 : \mathbb{B} \mapsto \mathbb{R}^n$, $\widehat{\mathcal{B}}_2 : \mathbb{B} \mapsto L^{\frac{4}{3}}(0,T;\mathbb{R}^n)$ *such that* $\|\mathcal{B}_1\|_{\mathcal{L}(\mathbb{B};\mathbb{R}^n)} \leq M$, $\|\mathcal{B}_2\|_{\mathcal{L}(\mathbb{B};L^{\frac{4}{3}}(0,T;\mathbb{R}^n))} \leq M$, *and*

$$f(x,y) = \int_0^T x_2(t)^\top \big[\widehat{\mathcal{B}}_2 y\big](t)\mathrm{d}t + x_1^\top \big[\widehat{\mathcal{B}}_1 y\big], \quad \forall x,\ y \in \mathbb{B}_1.$$

Proof. Step 1 : Fix $y \in \mathbb{B}_1$, for any $x \in \mathbb{B}_1$, we define $\varphi_y(x) = f(x,y)$. Then

$$\varphi_y(k\bar{x} + l\widetilde{x}) = k\varphi_y(\bar{x}) + l\varphi(\widetilde{x}), \quad |\varphi_y(x)| \leq M\|x\|_{\mathbb{B}}\|y\|_{\mathbb{B}}, \quad k,l \in \mathbb{Q}, \quad \bar{x},\widetilde{x} \in \mathbb{B}_1.$$

By Lemma A.2, there exists a unique $\widehat{\varphi}_y : \mathbb{B} \mapsto \mathbb{R}$ such that

$$\widehat{\varphi}_y(x) = \varphi_y(x), \quad \|\widehat{\varphi}_y\|_{\mathcal{L}(\mathbb{B};\mathbb{R})} \leq \mathcal{M}_y := M\|y\|_{\mathbb{B}}, \quad \forall x \in \mathbb{B}_1.$$

Then from Lemma A.1, there exists a unique $y^*(\cdot) = (y_1^*, y_2^*(\cdot)) \in \mathbb{B}'$ such that

$$\begin{cases} \widehat{\varphi}_y(x) = x_1^\top y_1^* + \displaystyle\int_0^T x_2(t)^\top y_2^*(t)\mathrm{d}t, \quad x = (x_1, x_2(\cdot)) \in \mathbb{B}, \\ \|\widehat{\varphi}_y\|_{\mathcal{L}(\mathbb{B};\mathbb{R})} = \max\Big\{|y_1^*|, \Big[\displaystyle\int_0^T |y_2^*(t)|^{\frac{4}{3}}\mathrm{d}t\Big]^{\frac{3}{4}}\Big\}. \end{cases}$$

Step 2 : For any $y \in \mathbb{B}_1$, let $\mathcal{B}y := \big([\mathcal{B}y]_1, [\mathcal{B}y]_2(\cdot)\big)$, $[\mathcal{B}y]_1 := y_1^*$, $[\mathcal{B}y]_2(\cdot) = y_2^*(\cdot)$. For $k,\ l \in \mathbb{Q}$, $\bar{h},\ \widetilde{h} \in \mathbb{B}$, we claim that $\mathcal{B}(k\bar{h} + l\widetilde{h}) = k\mathcal{B}\bar{h} + l\mathcal{B}\widetilde{h}$.

In fact, since $[k\bar{h} + l\widetilde{h}] \in \mathbb{B}_1$, by *Step* 1 there exists a unique linear bounded $\widehat{\varphi}_{k\bar{h}+l\widetilde{h}}(\cdot) : \mathbb{B} \mapsto \mathbb{R}$ such that $\widehat{\varphi}_{k\bar{h}+l\widetilde{h}} = \varphi_{k\bar{h}+l\widetilde{h}}$ in $\mathbb{B}_1$. Notice that $k\widehat{\varphi}_{\bar{h}} + l\widehat{\varphi}_{\widetilde{h}}$ is another linear bounded functional on $\mathbb{B}$ such that $\big[k\widehat{\varphi}_{\bar{h}} + l\widehat{\varphi}_{\widetilde{h}}\big] = \varphi_{k\bar{h}+l\widetilde{h}}$ in $\mathbb{B}_1$. By the uniqueness in Lemma A.2, $\widehat{\varphi}_{k\bar{h}+l\widetilde{h}} = \big[k\widehat{\varphi}_{\bar{h}} + l\widehat{\varphi}_{\widetilde{h}}\big]$.

From Lemma A.1, there exists $\big[k\bar{h} + l\widetilde{h}\big]^* \in \mathbb{R}^n \times L^{\frac{4}{3}}(0,T;\mathbb{R}^n)$ such that,

$$\begin{aligned} \widehat{\varphi}_{k\bar{h}+l\widetilde{h}}(x) &= \Big[\big[k\bar{h}+l\widetilde{h}\big]_1^*\Big]^\top x_1 + \int_0^T \Big[\big[k\bar{h}+l\widetilde{h}\big]_2^*(s)\Big]^\top x_2(s)\mathrm{d}s, \\ k\widehat{\varphi}_{\bar{h}}(x) + l\widehat{\varphi}_{\widetilde{h}}(x) &= \int_0^T x_2(t)^\top \big[k\bar{y}_2^*(t) + l\widetilde{y}_2^*(t)\big]\mathrm{d}t + x_1^\top \big[k\bar{y}_1^* + l\widetilde{y}_1^*\big]. \end{aligned}$$

By the arbitrariness of $x = (x_1, x_2(\cdot)) \in \mathbb{B}$, we have the following linearity of $\mathcal{B}$,

$$\big[k\bar{h} + l\widetilde{h}\big]_2^* = k\bar{y}_2^* + l\widetilde{y}_2^*, \quad \big[k\bar{h} + l\widetilde{h}\big]_1^* = k\bar{y}_1^* + l\widetilde{y}_1^*.$$

Step 3 : We extend the definition of $\mathcal{B}$ into $\mathbb{B}$. For any $y \in \mathbb{B}_1$, from *Step* 1, 2, $\|\mathcal{B}y\|_{\mathbb{B}'} \leq M\|y\|_{\mathbb{B}}$. Hence from the linearity of $\mathcal{B}$ and Lemma A.2, there exists a unique linear $\widehat{\mathcal{B}} : \mathbb{B} \mapsto \mathbb{B}'$ such that for $x \in \mathbb{B}_1$, $\widehat{\mathcal{B}}(x) = \mathcal{B}(x)$, and $\|\widehat{\mathcal{B}}\|_{\mathcal{L}(\mathbb{B},L^{\frac{4}{3}}(0,T;\mathbb{R}^n))} \leq M$. Therefore, for any $x,\ y \in \mathbb{B}_1$,

$$f(x,y) = \varphi_y(x) = \widehat{\varphi}_y(x) = x_1^\top \big(\widehat{\mathcal{B}}y\big)_1 + \int_0^T x_2(t)^\top \big(\widehat{\mathcal{B}}y\big)_2(t)\mathrm{d}t. \tag{A.2}$$

Step 4 : To see the conclusions of $\widehat{\mathcal{B}}_1$, $\widehat{\mathcal{B}}_2$, for $y \in \mathbb{B}_1$, we define $\mathcal{B}_1 : \mathbb{B}_1 \mapsto \mathbb{R}^n$, $\mathcal{B}_2 : \mathbb{B}_1 \mapsto L^{\frac{4}{3}}(0,T;\mathbb{R}^n)$ as $\mathcal{B}_1 y = \big(\mathcal{B}y\big)_1$, $(\mathcal{B}_2 y)(\cdot) = \big(\mathcal{B}y\big)_2(\cdot)$. Of course, both of them are well-defined. For $k, l \in \mathbb{Q}$, $\bar{h}, \widetilde{h} \in \mathbb{B}_1$, by the linearity of $\mathcal{B}$, we obtain

$$\mathcal{B}_i(k\bar{h} + l\widetilde{h}) = \big[\mathcal{B}(k\bar{h} + l\widetilde{h})\big]_i = k\big(\mathcal{B}\bar{h}\big)_i + l\big(\mathcal{B}\widetilde{h}\big)_i, \quad i = 1, 2.$$

On the other hand, for any $h \in \mathbb{B}_1$, $|\mathcal{B}_i h| \le \|\mathcal{B}h\| \le M\|h\|_{\mathbb{B}}$, $i = 1, 2$. Using again Lemma A.2, there exist a unique linear bounded $\widehat{\mathcal{B}}_1 : \mathbb{B} \mapsto \mathbb{R}^n$ and a unique linear bounded $\widehat{\mathcal{B}}_2 : \mathbb{B} \mapsto L^{\frac{4}{3}}(0,T;\mathbb{R}^n)$ such that $\widehat{\mathcal{B}}_1 y = \mathcal{B}_1 y$, $\widehat{\mathcal{B}}_2 y = \mathcal{B}_2 y$ with $y \in \mathbb{B}_1$. As a result, from (A.2) we obtain the conclusion. $\square$

Acknowledgements. The author highly appreciates two anonymous referees' constructive comments to improve on earlier version of this paper. He also gratefully acknowledges Professor Jiongmin Yong and Professor Xu Zhang for their valuable suggestions and helpful discussions.

References

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

[2] K. Arrow, Optimal capital policy, the cost of capital and myopic decision rules. *Ann. Inst. Stat. Math.,* **16** (1964) 21–30.

[3] A. Bensoussan, Lectures on stochastic control, Lecture notes in mathematics, Vol. 972 of *Nonlinear filtering and stochastic control, proceedings, Cortona* (1981).

[4] J. Bismut, An introductory approach to duality in optimal stochastic control. *SIAM Rev.* **20** (1978) 62–78.

[5] V. Boltyanski, R. Gamkrelidze and L. Pontryagin, The theory of optimal processes. I: The maximum principle. *Izv. Akad. Nauk SSSr Ser. Mat.* **24** (1960) 3–42 (in Russian); English transl. in *Am. Math. Soc. Transl.* **18** (1961) 341–382.

[6] K. Du and Q. Meng, A maximum principle for optimal control of stochastic evolution equations. *SIAM J. Control Optim.* **51** (2013) 4343–4362.

[7] A. Friedman, Optimal control for hereditary processes. *Arch. Rat. Mech. Anal.* **15** (1964) 396–416.

[8] M. Fuhrman, Y. Hu and G. Tessitore, Stochastic maximum principle for optimal control of SPDEs. *Appl. Math. Optim.* **68** (2013) 181–217.

[9] P. Lin and J. Yong, Controlled singular Volterra integral equations and Pontryagin maximum principle. Preprint arXiv:1712.05911v1.

[10] Q. Lü and X. Zhang, General Pontryagin-type stochastic maximum principle and backward stochastic evolution equation in infinite dimensions, Springer Briefs in Mathematics (2014).

[11] H. Kushner, Necessary conditions for continuous parameter stochastic optimization problems. *SIAM J. Control,* **10** (1972) 550–565.

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

[13] S. Peng, A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.* **28** (1990) 966–979.

[14] 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.

[15] V. Vinokurov, Optimal control of processes described by integral equations, I, II, III. *Izv. Vysš. Učebn. Zaved. Matematika* **7** (1969) 21–33; **8** (1969) 16–23; **9** (1969) 16–25; (in Russian) English transl. in *SIAM J. Control* **7** (1967) 324–336, 337–345, 346–355.

[16] T. Wang, Linear quadratic control problems of stochastic Voltera integral equations. *ESAIM: Control Optim. Cal. Var.* **24** (2018) 1849–1879.

[17] T. Wang and J. Yong, Comparison theorems for backward stochastic Volterra integral equations. *Stochastic Process Appl.,* **125** (2015) 1756–1798.

[18] T. Wang and H. Zhang, Optimal control problems of forward-backward stochastic Volterra integral equations with closed control regions. *SIAM J. Control Optim.* **55** (2017) 2574–2602.

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

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

[21] J. Yong and X. Zhou, Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer-Verlag, New York, Berlin (2000).