SIAM J. CONTROL OPTIM.
Vol. 61, No. 6, pp. 3608–3634

SPIKE VARIATIONS FOR STOCHASTIC VOLTERRA INTEGRAL EQUATIONS*

TIANXIAO WANG† AND JIONGMIN YONG‡

Abstract. The spike variation technique plays a crucial role in deriving Pontryagin's type maximum principle of optimal controls for ordinary differential equations (ODEs), partial differential equations (PDEs), stochastic differential equations (SDEs), and (deterministic forward) Volterra integral equations (FVIEs), when the control domains are not assumed to be convex. It is natural to expect that such a technique could be extended to the case of (forward) stochastic Volterra integral equations (FSVIEs). However, by mimicking the case of SDEs, one encounters an essential difficulty of handling an involved quadratic term. To overcome this difficulty, we introduce an auxiliary process for which one can use Itô's formula, and develop new technologies inspired by stochastic linear-quadratic optimal control problems. Then the suitable representation of the above-mentioned quadratic form is obtained, and the second-order adjoint equations are derived. Consequently, the maximum principle of Pontryagin type is established. Some relevant extensions are investigated as well.

Key words. backward stochastic Volterra integral equations, single variation, second-order adjoint equations, nonconvex control domain

MSC codes. 45D05, 60H20, 93E20

DOI. 10.1137/22M1522097

1. Introduction.

We begin with a main motivation of our study. Let $T>0$ and $n,m\in\mathbb{N}$ be given, let $(\Omega,\mathcal{F},\mathbb{P})$ be a complete probability space, on which a one-dimensional standard Brownian process $W(\cdot)$ is defined, with the completed augmented natural filtration denoted by $\mathbb{F}=\{\mathcal{F}_t\}_{0\leqslant t\leqslant T}$. Consider the following controlled (forward) stochastic Volterra integral equation (FSVIE):

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

with the cost functional

$$J(u(\cdot))=\mathbb{E}\left[h(X(T))+\int_0^T g(s,X(s),u(s))ds\right]. \tag{1.2}$$

Here $u(\cdot)$ is a *control process* valued in $U\subseteq\mathbb{R}^m$, $X(\cdot)$ is the corresponding *state process* valued in $\mathbb{R}^n$, $\varphi:[0,T]\times\Omega\to\mathbb{R}^n$ is called a *free term,* $(b,\sigma):[0,T]^2\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n\times\mathbb{R}^n$ is called the *generator,* and $h:\Omega\times\mathbb{R}^n\to\mathbb{R}$, $g:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ are called the *terminal cost* and *running cost rate,* respectively. Under certain conditions, for any $u(\cdot)\in\mathscr{U}_{ad}$ (to be defined later), (1.1) admits a unique solution $X(\cdot)$ such that $J(u(\cdot))$ is well-defined. The optimal control problem can be stated as follows:

*Received by the editors September 13, 2022; accepted for publication (in revised form) September 10, 2023; published electronically December 7, 2023.
https://doi.org/10.1137/22M1522097
Funding: This work was supported by the National Natural Science Foundation of China (11971332, 11931011), the Science Development Project of Sichuan University (2020SCUNL201), and NSF grant DMS-2305475
†School of Mathematics, Sichuan University, Chengdu, 610065, China (wtxiao2014@scu.edu.cn).
‡Department of Mathematics, University of Central Florida, Orlando, FL 32816 USA (jiongmin.yong@ucf.edu).

Problem (C). Find a $\bar{u}(\cdot) \in \mathscr{U}_{ad}$ such that

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot)\in\mathscr{U}_{ad}} J(u(\cdot)). \tag{1.3}$$

Here $\bar{u}(\cdot)$ is called an *open-loop optimal control,* the corresponding $\bar{X}(\cdot)$ and $(\bar{X}(\cdot), \bar{u}(\cdot))$ are called the (*open-loop*) *optimal state process* and an (*open-loop*) *optimal pair,* respectively.

Pontryagin's maximum principle for (deterministic) controlled ordinary differential equations (ODEs) was formulated and proved by Boltyanskii, Gamkrelidze, andPontryagin in 1956 [6] (see also [25]). Since then, the research in this direction has attracted many authors. Shortly after, people extended the results to stochastic differential equations (SDEs). In 1965, Kushner first proved a maximum principle for SDEs where the diffusion is independent of the state and control [15]. In 1972, Kushner further considered the case that the diffusion of the state equation contains the state [16], in which the adjoint process was not successfully characterized. In 1973, Bismut [4] introduced the duality between an initial value problem of linear SDE (called FSDE) and a terminal value problem of a linear SDE (now called BSDE, a name coined by Pardoux and Peng in 1990 [23]). With this, he proved the maximum principle for FSDEs with the diffusion contains state and control, provided the control domain is convex; the adjoint process was successfully characterized by the adapted solution (again, named by Pardoux and Peng in 1990) of corresponding linear BSDE. There were a number of following-up works (e.g., Bensoussan [3]). The maximum principle for general controlled SDEs is of the following form:

$$X(t) = x + \int_0^t b(s, X(s), u(s))ds + \int_0^t \sigma(s, X(s), u(s))dW(s), \quad t \in [0, T], \tag{1.4}$$

without assuming the convexity of the control domain was proved by Peng in 1990 [24], with the development of spike variation technique for FSDEs. We refer to the monograph of Yong and Zhou [38] for a self-contained presentation, and Hu [13], Yong [37] for the extension to the forward-backward case. By the way, spike variation techniques for partial differential equations (PDEs) and abstract evolution equations in Hilbert/Banach spaces were also fully developed during the 1970–1990s. See, e.g., Li–Yong [18], Casas and Yong [8], and Hu and Yong [12], Fattorini [9] for systematic presentations.

The integral equation was first introduced by Abel in 1825. The follow-up contributors include Liouville, Fredholm, Hilbert, Wiener, Bellman, etc. We refer to Bôcher [5] and Yosida [39] for two early monographs in this topic. The (forward) Volterra integral equation (FVIEs) was introduced by Italian mathematician Volterra. In contrast with ODEs, FVIEs bring us new theories (e.g., Volterra operator theory in Gohberg and Krein [11]) and new phenomena (e.g., the hereditary property in Bellman and-Cooke [2], Volterra [28]). In the real world, there are many models that cannot be simply described by ODEs but by (deterministic) FVIEs. Therefore, up until now, it is still an active research topic, see, e.g., the recent monograph of Brunner [7]. For the optimal control theory of FVIEs, to the best of our knowledge, the earliest work was due to Friedman [10] in 1964. Later in 1967, Vinokurov [27] investigated the optimal control for FVIEs with various type constraints and closed control domains (described by level sets of smooth functions). There are quite a few literature concerning optimal control of FVIEs in the past several decades (e.g., Kamien and Muller [14], Medhin [21]). Recently, Lin and Yong [20] established a maximum principle for singular FVIEs by spike variation technique.

In contrast with the existing literatures for optimal control of (deterministic) FVIEs, there were much less literature on optimal control of FSVIEs, to the best of our knowledge. The simple reason is that the theory of adjoint equation for FSVIEs, i.e., the backward SVIEs (BSVIEs) was not available until the appearance of the work on Type-I BSVIEs by Lin [19] in 2002. Even then, the result was not connected with the maximum principle of optimal control for FSVIEs. In 2006, Yong introduced Type-II BSVIEs and established maximum principle for FSVIEs the first time [35], provided the control domain is convex (see also Yong [36]). There are some further extensions appeared lately. Here are some of those results. Shi, Wang, and Yong [26] studied the case of forward-backward SVIEs. Agram and Øksendal [1] presented some investigations by means of Malliavin calculus, when the control domain U is open. Wang and Zhang [34] considered the case with closed set U where the necessary conditions were obtained from those directions that approximate convex perturbations at the optimal control are admissible. Here spike variations are not necessary.

Recently, Wang [31] studied the general case without the convexity of the control domain and attempted to develop the spike variation techniques for FSVIEs. In doing so, the following form of *quadratic functional*

$$\mathscr{E}(\varepsilon)=\mathbb{E}\left[X_1^\varepsilon(T)^\top h_{xx}X_1^\varepsilon(T)+\int_0^T X_1^\varepsilon(t)^\top H_{xx}(t)X_1^\varepsilon(t)dt\right] \tag{1.5}$$

appeared in the variation of the cost functional, where h_{xx} and $H_{xx}(\cdot)$ are $\mathbb{S}^n$ (the set of all real $(n\times n)$ symmetric matrices) valued functions only depending on the optimal pair $(\bar X(\cdot),\bar u(\cdot))$, and $X_1^\varepsilon(\cdot)$ is the solution of the first-order variational equation, depending on the spike variation of the control. In the case that the control domain is convex, convex variation of the control is allowed and by doing that the above quadratic form does not appear. Thus, it is not necessary for [1, 26, 34, 35, 36] to introduce techniques of handling the above. On the other hand, recall that similar quadratic form as $\mathscr{E}(\varepsilon)$ appears in the FSDE case, which can be handled by a duality between the second-order adjoint BSDE and the variational SDE satisfied by $X_1^\varepsilon(\cdot)X_1^\varepsilon(\cdot)^\top$ (see Peng [24]). It is unlikely that for FSVIEs one can obtain a good-looking linear FSVIE for $X_1^\varepsilon(\cdot)X_1^\varepsilon(\cdot)^\top$. Hence, the idea of Peng [24] cannot be directly borrowed here. In Wang [31], the author introduced two abstract and implicit operator-valued stochastic processes, which are regarded as a replacement of the second-order adjoint processes, in the statement of the maximum principle. However, there was no second-order adjoint equation in [31].

In the current paper, we are going to develop a spike variation technique for FSVIEs. The main ideas can be described as follows. Some recent studies (e.g., Li, Sun, and Yong [17], Wang [29]) show that Lyapunov type equations, obtained via decoupling forward-backward system, play crucial rules in obtaining Riccati equations in linear-quadratic (LQ) optimal controls. Here we find a slightly different manner to derive the Lyapunov type equation (with solution $P(\cdot)$) by applying the Itô's formula to $X(\cdot)^\top P(\cdot)X(\cdot)$ with $X(\cdot)$ satisfying an SDE. It turns out that this idea can be adopted to our current SVIE's problem. Together with the introduction of a suitable auxiliary process, we are able to represent the quadratic form involved $X_1^\varepsilon(\cdot)$ in a desired fashion. Consequently, the second-order adjoint equation will be derived which leads to the maximum principle. Moreover, we will go a little further. Recall that in 2007, Mou and Yong [22] established a variational formula for an SDE type state equation and the possibly vector-valued cost/payoff functionals. Those results are applicable to multiobjective problems and multiperson differential games of SDEs.

Inspired by this, in the current paper, for FSVIEs, we will do the same thing, so that it will lead to necessary conditions for the Nash equilibria of multiperson dynamic games for FSVIEs. By letting the number of players go to infinity, and under some structure conditions, it is expected to obtain the necessary conditions for the equilibria of the mean-field games governed by FSVIEs, which includes the standard SDE case. We hope to report more details in our forthcoming publications.

The rest of this paper is organized as follows. In section 2, we introduce some preliminary notations and assumptions. In section 3, we present the variations of the state and the cost functional under the spike variation of the control process. In section 4, we give a representation of the involved quadratic functional by introducing a system of BSVIEs, called the *second-order adjoint equations.* In section 5, we prove the well-posedness of the aforementioned adjoint equations. We present the Pontryagin's type maximum principle and some extensions to multiperson dynamic games for FSVIEs in section 6. Finally, some concluding remarks are collected in section 7.

2. Preliminaries.

Let $\mathbb{R}^n$ and $\mathbb{R}^{n\times d}$ be the usual n-dimensional space of real numbers and the set of all $(n\times d)$ real matrices, respectively. Also, let $\mathbb{S}^n$ be the set of all $(n\times n)$ real symmetric matrices. Next, for any Euclidean space $\mathbb{H}$ which could be $\mathbb{R}^n,\mathbb{R}^{n\times d}$, $\mathbb{S}^n$, etc., we define

$$L^p_{\mathcal{F}_t}(\Omega;\mathbb{H})\triangleq\Big\{\xi:\Omega\to\mathbb{H}\;\big|\;\xi\text{ is }\mathcal{F}_t\text{-measurable, }\|\xi\|_p\equiv\big(\mathbb{E}|\xi|^p\big)^{\frac{1}{p}}<\infty\Big\},\quad p\in[1,\infty).$$

In an obvious way, we can define $L^\infty_{\mathcal{F}_t}(\Omega;\mathbb{H})$. For each $p\in[1,\infty]$, $L^p_{\mathcal{F}_t}(\Omega;\mathbb{H})$ is a Banach space under $\|\cdot\|_p$. When the range space $\mathbb{H}$ is clear from the context and is not necessarily to be emphasized, we will omit $\mathbb{H}$. In particular, $L^p_{\mathcal{F}_T}(\Omega)\triangleq L^p_{\mathcal{F}_T}(\Omega;\mathbb{H})$.

We introduce spaces of stochastic processes. To avoid repetition, all processes $(t,\omega)\mapsto\varphi(t,\omega)$ are assumed to be at least $\mathscr{B}[0,T]\otimes\mathcal{F}_T$-measurable without further mentioning, where $\mathscr{B}[0,T]$ is the Borel σ-field of $[0,T]$. For $p,q\in[1,\infty)$, $\tau\in[0,T)$,

$$L^p(\Omega;L^q(\tau,T;\mathbb{H}))\triangleq\left\{\varphi:[\tau,T]\times\Omega\to\mathbb{H}\;\Big|\;\mathbb{E}\left(\int_\tau^T|\varphi(t)|^q dt\right)^{\frac{p}{q}}<\infty\right\},$$

$$L^p(\Omega;L^\infty(\tau,T;\mathbb{H}))\triangleq\left\{\varphi:[\tau,T]\times\Omega\to\mathbb{H}\;\Big|\;\mathbb{E}\left(\operatorname*{esssup}_{t\in[\tau,T]}|\varphi(t)|^p\right)<\infty\right\},$$

$$L^p(\Omega;C([\tau,T];\mathbb{H}))\triangleq\Bigg\{\varphi:[\tau,T]\times\Omega\to\mathbb{H}\;\Big|\;t\mapsto\varphi(t,\omega)\text{ is continuous,}$$
$$\mathbb{E}\left(\sup_{t\in[\tau,T]}|\varphi(t)|^p\right)<\infty\Bigg\},$$

$$L^q(\tau,T;L^p(\Omega;\mathbb{H}))\triangleq\left\{\varphi:[\tau,T]\times\Omega\to\mathbb{H}\;\Big|\;\int_\tau^T\Big(\mathbb{E}|\varphi(t)|^p\Big)^{\frac{q}{p}}dt<\infty\right\},$$

$$L^\infty(\tau,T;L^p(\Omega;\mathbb{H}))\triangleq\left\{\varphi:[\tau,T]\times\Omega\to\mathbb{H}\;\Big|\;\operatorname*{esssup}_{t\in[\tau,T]}\Big(\mathbb{E}|\varphi(t)|^p\Big)^{\frac{1}{p}}<\infty\right\},$$

$$C([\tau,T];L^p(\Omega;\mathbb{H}))\triangleq\Big\{\varphi:[\tau,T]\times\Omega\to\mathbb{H}\;\Big|\;t\mapsto\varphi(t,\cdot)\text{ is continuous},\\ \sup_{t\in[0,T]}\Big(\mathbb{E}|\varphi(t)|^p\Big)^{\frac{1}{p}}<\infty\Big\}.$$

The spaces $L^\infty(\Omega;L^\infty(\tau,T;\mathbb{H}))$, $L^\infty(\Omega;C([\tau,T];\mathbb{H}))$ can be defined obviously. For all $p\in[1,\infty)$, we denote

$$L^p(\tau,T;\mathbb{H})\triangleq L^p(\tau,T;L^p(\Omega;\mathbb{H}))=L^p(\Omega;L^p(\tau,T;\mathbb{H})).$$

All of the $\mathbb{F}$-progressive version of the above spaces can be denoted by putting $\mathbb{F}$ as a subscript; for example, $C_{\mathbb{F}}([\tau,T];L^p(\Omega;\mathbb{H}))$, and so on. Finally, in the above, when the range space $\mathbb{H}$ is clear from the context, we will omit $\mathbb{H}$; for example, $L^p_{\mathbb{F}}(\Omega;C([\tau,T]))$.

Next, the upper and lower triangle domains are defined by the following:

$$\begin{aligned}&\Delta^*[\tau,T]\triangleq\big\{(r,s)\in[\tau,T]^2\;\big|\;\tau\leqslant r<s\leqslant T\big\},\\&\Delta_*[\tau,T]\triangleq\big\{(r,s)\in[\tau,T]^2\;\big|\;\tau\leqslant s<r\leqslant T\big\},\end{aligned}\qquad\tau\in[0,T).$$

We introduce the following spaces:

$$L^2_{\mathbb{F}}(\Omega;L^2(\Delta^*[\tau,T]))\triangleq\Big\{\zeta:\Delta^*[\tau,T]\times\Omega\to\mathbb{H}\;\big|\;\zeta(r,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(r,T;\mathbb{H})),\\ \text{a.e. } r\in[\tau,T],\;\mathbb{E}\int_\tau^T\int_r^T|\zeta(r,s)|^2dsdr<\infty\Big\},$$

$$L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\tau,T]))\triangleq\Big\{\zeta:\Delta_*[\tau,T]\times\Omega\to\mathbb{H}\;\big|\;\zeta(r,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(\tau,r;\mathbb{H})),\\ \text{a.e. } r\in[\tau,T],\;\mathbb{E}\int_\tau^T\int_\tau^r|\zeta(r,s)|^2dsdr<\infty\Big\},$$

$$L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\tau,T]\times[\tau,T]))\triangleq\Big\{\zeta:\Delta_*[\tau,T]\times[\tau,T]\times\Omega\to\mathbb{H}\;\big|\;(\theta,r)\in\Delta_*[\tau,T],\\ \zeta(\theta,r,\cdot)\in L^2_{\mathbb{F}}(\tau,T;\mathbb{H}),\quad\mathbb{E}\int_\tau^T\int_\tau^\theta\int_\tau^T|\zeta(\theta,r,s)|^2dsdrd\theta<\infty\Big\},$$

$$L^2_{\mathbb{F}}(\Omega;L^2([\tau,T]\times[\tau,T]))\triangleq\Big\{\zeta:\Delta^*[\tau,T]\times\Omega\to\mathbb{H}\;\big|\;\zeta(r,\cdot)\in L^2_{\mathbb{F}}(\tau,T;\mathbb{H}),\\ \text{a.e. } r\in[\tau,T],\;\mathbb{E}\int_\tau^T\int_\tau^T|\zeta(r,s)|^2dsdr<\infty\Big\}.$$

For simplicity of presentation, let K be a generic constant which could be different from line to line. For FSVIE (1.1), we introduce the following assumptions.

(H1) Let $T>0$, $\varnothing\neq U\subset\mathbb{R}^m$ be measurable. Let $b,\sigma:\Delta_*[0,T]\times\mathbb{R}^n\times U\times\Omega\to\mathbb{R}^n$ be measurable such that $s\mapsto\big(b(t,s,x,u),\sigma(t,s,x,u)\big)$ is $\mathbb{F}$-progressively measurable on $[0,t]$, $x\mapsto(b(t,s,x,u),\sigma(t,s,x,u))$ is twice continuously differentiable with uniformly bounded first and second order derivatives, and for constant $L>0$,

$$|b(t,s,0,u)|+|\sigma(t,s,0,u)|\leqslant L,\qquad(t,s,u)\in\Delta_*[0,T]\times U.$$

Further, for $f \stackrel{\triangle}{=} b, \sigma, b_x, \sigma_x$, $(t,u) \mapsto f(t,s,x,u)$ is continuous uniformly in all the arguments, and for $f \stackrel{\triangle}{=} b, \sigma$, $(x,u) \mapsto f_{xx}(t,s,x,u)$ is continuous uniformly in all arguments.

Let $\mathscr{U}_{ad} \stackrel{\triangle}{=} \big\{u : [0,T] \times \Omega \to U \;\big|\; u(\cdot) \text{ is } \mathbb{F}\text{-progressively measurable}\big\}$ which is the set of all admissible controls. Under (H1), for $\varphi(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega;\mathbb{R}^n))$ with some $p \geqslant 1$ and any $u(\cdot) \in \mathscr{U}_{ad}$, it is standard to see (1.1) admits a unique solution $X(\cdot) \equiv X(\cdot\,; u(\cdot)) \in C_{\mathbb{F}}([0,T]; L^p(\Omega;\mathbb{R}^n))$. For the functions h and g in (1.2), we make the following assumption.

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

$$|g_x(t,0,u)| \leqslant L, \quad |h_{xx}(x)| + |g_{xx}(t,x,u)| \leqslant L, \qquad (t,x,u) \in [0,T] \times \mathbb{R}^n \times U$$

for some constant $L > 0$ and $(t,x,u) \mapsto g_{xx}(t,x,u)$ being uniformly continuous.

Under (H1)–(H2), problem (C) is well-formulated. Further, if h and g take values in $\mathbb{R}^\ell$, then our study covers multiobjective problems as well as multiperson nonzero sum dynamic games governed by FSVIEs (if $\ell \geqslant 2$). Finally, we point out that (H1)–(H2) can be relaxed somehow, for example, we may allow b, σ, and g to have some growth in u. We prefer not to pursue these generalities.

3. Variations of the state and the cost functional. In this section, we present the variations of the state process and the cost functional under the spike variation of the control process.

Let $(\bar{X}(\cdot), \bar{u}(\cdot))$ be any fixed state-control pair, which could be optimal, in particular. By choosing any $u \in U$, $\tau \in [0,T)$, and sufficiently small $\varepsilon > 0$ such that $\tau + \varepsilon \leqslant T$, we define $u^\varepsilon(\cdot) \stackrel{\triangle}{=} u\mathbf{1}_{[\tau,\tau+\varepsilon]}(\cdot) + \bar{u}(\cdot)\mathbf{1}_{[0,T]\setminus[\tau,\tau+\varepsilon]}(\cdot)$. This is called a *spike variation* of $\bar{u}(\cdot)$ in the direction of u on $[\tau, \tau+\varepsilon]$. Let $X^\varepsilon(\cdot)$ be the state process corresponding to $u^\varepsilon(\cdot)$. We introduce the following abbreviations: for $f \stackrel{\triangle}{=} b, \sigma$,

$$\begin{aligned}
f_x(t,s) &\stackrel{\triangle}{=} f_x(t,s,\bar{X}(s),\bar{u}(s)),\\
f_{xx}^i(t,s) &\stackrel{\triangle}{=} f_{xx}^i(t,s,\bar{X}(s),\bar{u}(s)), \quad 1 \leqslant i \leqslant n,\\
g_x(s) &\stackrel{\triangle}{=} g_x(s,\bar{X}(s),\bar{u}(s)),\\
g_{xx}(s) &\stackrel{\triangle}{=} g_{xx}(s,\bar{X}(s),\bar{u}(s)),\\
h_x &\stackrel{\triangle}{=} h_x(\bar{X}(T)),\\
h_{xx} &\stackrel{\triangle}{=} h_{xx}(\bar{X}(T)),\\
\delta f(t,s) &\stackrel{\triangle}{=} f(t,s,\bar{X}(s),u) - f(t,s,\bar{X}(s),\bar{u}(s)),\\
\delta\sigma_x(t,s) &\stackrel{\triangle}{=} \sigma_x(t,s,\bar{X}(s),u) - \sigma_x(t,s,\bar{X}(s),\bar{u}(s)).
\end{aligned}$$

With the above notation, we introduce the following variational equations:

$$\text{(3.1)} \quad \begin{cases}
X_1^\varepsilon(t) = \displaystyle\int_0^t b_x(t,s)X_1^\varepsilon(s)ds + \int_0^t \big[\sigma_x(t,s)X_1^\varepsilon(s) + \delta\sigma(t,s)\mathbf{1}_{[\tau,\tau+\varepsilon]}(s)\big]dW(s),\\
X_2^\varepsilon(t) = \displaystyle\int_0^t \Big(b_x(t,s)X_2^\varepsilon(s) + \frac{1}{2}b_{xx}(t,s)X_1^\varepsilon(s)^2 + \delta b(t,s)\mathbf{1}_{[\tau,\tau+\varepsilon]}(s)\Big)ds\\
\qquad\qquad + \displaystyle\int_0^t \Big(\sigma_x(t,s)X_2^\varepsilon(s) + \frac{1}{2}\sigma_{xx}(t,s)X_1^\varepsilon(s)^2\\
\qquad\qquad + \delta\sigma_x(t,s)X_1^\varepsilon(s)\mathbf{1}_{[\tau,\tau+\varepsilon]}(s)\Big)dW(s), \quad t \in [0,T],
\end{cases}$$

where

$$b_{xx}(t,s)X_1^\varepsilon(s)^2 \triangleq \begin{pmatrix} X_1^\varepsilon(s)^\top b_{xx}^1(t,s)X_1^\varepsilon(s) \\ \vdots \\ X_1^\varepsilon(s)^\top b_{xx}^n(t,s)X_1^\varepsilon(s) \end{pmatrix},$$

and $\sigma_{xx}(t,s)X_1^\varepsilon(s)^2$ is similar. Under (H1), the above (3.1) has unique strong solutions $X_1^\varepsilon(\cdot)$ and $X_2^\varepsilon(\cdot)$. We also introduce the following first-order adjoint equations:

$$\begin{cases} \eta(t) = h_x^\top - \int_t^T \zeta(s)dW(s), \qquad t \in [0,T], \\ Y(t) = g_x(t)^\top + b_x(T,t)^\top h_x^\top + \sigma_x(T,t)^\top \zeta(t) \\ \qquad\qquad + \int_t^T \Big(b_x(s,t)^\top Y(s) + \sigma_x(s,t)^\top Z(s,t) \Big) ds \\ \qquad\qquad - \int_t^T Z(t,s)dW(s), \quad t \in [0,T]. \end{cases} \tag{3.2}$$

The first equation of the above is the simplest BSDE. The second is a Type-II BSVIE which has a unique adapted M-solution [36]. Both $(\eta(\cdot),\zeta(\cdot))$ and $(Y(\cdot),Z(\cdot\,,\cdot))$ of (3.2) only depend on $(\bar X(\cdot),\bar u(\cdot))$. Finally, let us define the following Hamiltonian:

$$\begin{aligned} H(s,x,u,\eta(s),\zeta(s),Y(\cdot),Z(\cdot\,,s)) &\triangleq \langle\, \eta(s), b(T,s,x,u)\,\rangle + \langle\, \zeta(s), \sigma(T,s,x,u)\,\rangle \\ + g(s,x,u) + \mathbb{E}_s \int_s^T \Big(\langle\, Y(t), b(t,s,x,u)\,\rangle &+ \langle\, Z(t,s), \sigma(t,s,x,u)\,\rangle \Big) dt. \end{aligned} \tag{3.3}$$

Hereafter, $\mathbb{E}_s \triangleq \mathbb{E}[\,\cdot\,|\mathcal{F}_s]$ is the conditional expectation operator. Then the main result of this section can be stated as follows.

THEOREM 3.1. *Let* (H1)–(H2) *hold. Then for any* $k \geqslant 1$,

$$\begin{cases} \sup_{s\in[0,T]} \mathbb{E}|X^\varepsilon(s) - \bar X(s)|^k = O(\varepsilon^{\frac{k}{2}}), \\ \sup_{s\in[0,T]} \mathbb{E}|X_1^\varepsilon(s)|^k = O(\varepsilon^{\frac{k}{2}}), \\ \sup_{s\in[0,T]} \mathbb{E}|X^\varepsilon(s) - \bar X(s) - X_1^\varepsilon(s)|^k = O(\varepsilon^k), \\ \sup_{s\in[0,T]} \mathbb{E}|X_2^\varepsilon(s)|^k = O(\varepsilon^k), \\ \sup_{s\in[0,T]} \mathbb{E}|X^\varepsilon(s) - \bar X(s) - X_1^\varepsilon(s) - X_2^\varepsilon(s)|^2 = o(\varepsilon^2). \end{cases} \tag{3.4}$$

Moreover,

$$J(u^\varepsilon(\cdot)) - J(\bar u(\cdot)) = \mathbb{E}\int_\tau^{\tau+\varepsilon} \Delta H(s)ds + \frac{1}{2}\mathscr{E}(\varepsilon) + o(\varepsilon), \tag{3.5}$$

where

$$\begin{aligned} \Delta H(s) \triangleq\; & H(s,\bar X(s),u,\eta(s),\zeta(s),Y(\cdot),Z(\cdot\,,s)) \\ & - H(s,\bar X(s),\bar u(s),\eta(s),\zeta(s),Y(\cdot),Z(\cdot\,,s)), \end{aligned}$$

$$\begin{aligned}\mathscr{E}(\varepsilon) &\overset{\triangle}{=} \mathbb{E}\bigg[X_1^\varepsilon(T)^\top h_{xx}(\bar{X}(T))X_1^\varepsilon(T) \\ &\qquad + \int_0^T X_1^\varepsilon(s)^\top H_{xx}(s,\bar{X}(s),\bar{u}(s),\eta(s),\zeta(s),Y(\cdot),Z(\cdot\,,s))X_1^\varepsilon(s)ds\bigg] \\ &\equiv \mathbb{E}\left[X_1^\varepsilon(T)^\top h_{xx}X_1^\varepsilon(T) + \int_0^T X_1^\varepsilon(s)^\top H_{xx}(s)X_1^\varepsilon(s)ds\right].\end{aligned} \tag{3.6}$$

The above is essentially shown in [31] with the main idea of [24] (see also [38]). A detailed proof can be found in [33]. We point out that in the last estimate of (3.4), 2 cannot be replaced by "any $k \geqslant 1$".

4. Representation of the quadratic form. Recalling (3.1) (suppressing ε in $X_1^\varepsilon(\cdot)$), we have $X_1(t)=0$, $t\in[0,\tau]$, and

$$X_1(t) = \int_\tau^t b_x(t,s)X_1(s)ds + \int_\tau^t \big[\sigma_x(t,s)X_1(s) + \delta\sigma^\varepsilon(t,s)\big]dW(s), \quad t\in[\tau,T],$$

where $\delta\sigma^\varepsilon(t,s) \overset{\triangle}{=} \delta\sigma(t,s)\mathbf{1}_{[\tau,\tau+\varepsilon]}(s)$. Thus $X_1(t) = \Phi[\delta\sigma^\varepsilon(\cdot\,,\cdot)](t)$, $t\in[\tau,T]$, for some operator Φ. Let $X_1(T) = \Phi[\delta\sigma^\varepsilon(\cdot\,,\cdot)](T) \equiv \widehat{\Phi}[\delta\sigma^\varepsilon(\cdot\,,\cdot)]$. Consequently, by misusing the notation,

$$\mathscr{E}(\varepsilon) = \langle\, \big(\widehat{\Phi}^* h_{xx}\widehat{\Phi} + \Phi^* H_{xx}\Phi\big)\delta\sigma^\varepsilon, \delta\sigma^\varepsilon\,\rangle\,. \tag{4.1}$$

Here the dependence of $\mathscr{E}(\varepsilon)$ on $\delta\sigma^\varepsilon(\cdot\,,\cdot)$ is indirect since the operator Φ is complicated. The purpose of this section is to obtain a more direct form, whose procedure leads to the second-order adjoint equation. To express the main idea, we look at a simple problem on SDEs.

Consider the following linear SDE:

$$\begin{cases} dX(s) = \big[A(s)X(s) + b(s)\big]ds + \big[C(s)X(s) + \sigma(s)\big]dW(s), & s\in[t,T], \\ X(t) = x, \end{cases}$$

with suitable processes $A(\cdot), C(\cdot)$, $b(\cdot), \sigma(\cdot)$. Also, we consider the following quadratic functional:

$$J = \mathbb{E}\left[X(T)^\top G X(T) + \int_t^T \Big(X(s)^\top Q(s)X(s) + 2q(s)^\top X(s)\Big)ds\right],$$

with suitable $G, Q(\cdot)$, $q(\cdot)$. In the LQ problem, A, C, Q could rely on feedback strategy (e.g., [17, 29]). Let $(P(\cdot),\Lambda(\cdot))$ be the adapted solution to the following BSDE:

$$\begin{cases} dP(t) = -\Gamma(t)dt + \Lambda(t)dW(t), & t\in[0,T], \\ P(T) = G, \end{cases} \tag{4.2}$$

with $\Gamma(\cdot)$ undetermined. Then by Itô's formula (suppressing s), we have

$$\begin{aligned} d(X^\top PX) = \Big[&X^\top\Big(-\Gamma + PA + A^\top P + C^\top PC + \Lambda C + C^\top\Lambda\Big)X^\top \\ &+X^\top(Pb + C^\top P\sigma + \Lambda\sigma) + \Big(b^\top P + \sigma^\top PC + D^\top\Lambda\Big)X \\ &+\sigma^\top P\sigma\Big]ds + \{\cdots\}dW(s).\end{aligned}$$

Consequently, the term $X(T)^\top GX(T)$ in J has been absorbed since

$$J=\mathbb{E}\int_t^T \Big[X(s)^\top\Big(-\Gamma+PA+A^\top P+C^\top PC+\Lambda C+C^\top\Lambda+Q\Big)X(s) \\ +2\Big(b^\top P+\sigma^\top PC+\sigma^\top\Lambda+q^\top\Big)X+\sigma^\top P\sigma\Big]ds+x^\top P(t)x.$$

To further eliminate the term $X(s)^\top(\cdots)X(s)$ under the integral, we may choose

$$\Gamma=PA+A^\top P+C^\top PC+\Lambda C+C^\top\Lambda+Q, \tag{4.3}$$

so that a new representation of J can be given. The ideas of obtaining Γ and J will play crucial roles below. We point out that (4.2) (with Γ in (4.3)) corresponds to the Lyapunov type equation in Markovian stochastic LQ problems (e.g., [17, 29]) and second-order adjoint equations in maximum principles (e.g., [24]).

Note that the above could not be applied directly to $\mathscr{E}(\varepsilon)$ as $t\mapsto X_1(t)$ does not satisfy an SDE, and therefore, the Itô's formula does not work. To overcome this difficulty, we introduce an auxiliary process $\mathcal{X}_1(\cdot\,,\cdot)$ as follows:

$$\mathcal{X}_1(t,r)=\int_\tau^r b_x(t,s)X_1(s)ds+\int_\tau^r\big[\sigma_x(t,s)X_1(s)+\delta\sigma^\varepsilon(t,s)\big]dW(s), \tag{4.4}$$

with $\tau\leqslant r\leqslant t\leqslant T$. A similar form also appeared in [30]. Note that $r\mapsto\mathcal{X}_1(t,r)$ satisfies an SDE on $[\tau,t]$, and thus the Itô's formula can be used. Moreover, for any $t\in[\tau,T]$, by the second condition in (3.4),

$$\sup_{t\in[\tau,T]}\mathbb{E}\left[\sup_{r\in[\tau,t]}|\mathcal{X}_1(t,r)|^p\right]=O(\varepsilon^{\frac{p}{2}}). \tag{4.5}$$

Further, it is clear that $\mathcal{X}_1(r,r)=X_1(r)$ with $r\in[\tau,T]$. Thus,

$$\mathscr{E}(\varepsilon)=\mathbb{E}\left[\mathcal{X}_1(T,T)^\top h_{xx}\mathcal{X}_1(T,T)+\int_0^T\mathcal{X}_1(s,s)^\top H_{xx}(s)\mathcal{X}_1(s,s)ds\right]. \tag{4.6}$$

We now treat the terms in $\mathscr{E}(\varepsilon)$. To this end, we first carry out a general calculation which will be used several times below.

Let $\Theta:\Delta_*[\tau,T]\times\Omega\to\mathbb{R}^{n\times n}$ be a process such that for each $t\in[\tau,T]$, $\Theta(t,\cdot)\in L^2_{\mathbb{F}}(\tau,t;\,\mathbb{R}^{n\times n})$. For each $(t,s)\in\Delta_*[\tau,T]$, applying the martingale representation theorem to the random variable $\Theta(t,s)$, one can find a unique $\Lambda(t,s,\cdot)$ such that

$$\Pi(t,s,r)\equiv\mathbb{E}_r[\Theta(t,s)]=\Theta(t,s)-\int_r^s\Lambda(t,s,\theta)dW(\theta),\qquad \tau\leqslant r\leqslant s\leqslant t\leqslant T. \tag{4.7}$$

Then, applying the Itô's formula to the map $r\mapsto\mathcal{X}_1(t,r)^\top\Pi(t,s,r)\mathcal{X}_1(s,r)$ yields

$$\begin{aligned} &d\big[\mathcal{X}_1(t,r)^\top\Pi(t,s,r)\mathcal{X}_1(s,r)\big]\\ &\quad=\Big[X_1(r)^\top\Big(b_x(t,r)^\top\Pi(t,s,r)+\sigma_x(t,r)^\top\Lambda(t,s,r)\Big)\mathcal{X}_1(s,r)\\ &\qquad+\mathcal{X}_1(t,r)^\top\Big(\Pi(t,s,r)b_x(s,r)+\Lambda(t,s,r)\sigma_x(s,r)\Big)X_1(r)\\ &\qquad+X_1(r)^\top\sigma_x(t,r)^\top\Pi(t,s,r)\sigma_x(s,r)X_1(r)+\delta\sigma^\varepsilon(t,r)^\top\Pi(t,s,r)\delta\sigma^\varepsilon(s,r)\\ &\qquad+\delta\sigma^\varepsilon(t,r)^\top\Big(\Lambda(t,s,r)\mathcal{X}_1(s,r)+\Pi(t,s,r)\sigma_x(s,r)X_1(r)\Big)\\ &\qquad+\Big(\mathcal{X}_1(r,s)^\top\Lambda(t,s,r)+X_1(r)^\top\sigma_x(t,r)^\top\Pi(t,s,r)\Big)\delta\sigma^\varepsilon(s,r)\Big]dr+\Gamma(t,s,r)dW(r), \end{aligned}$$

where

$$\Gamma(t,s,r) \overset{\triangle}{=} \Big(X_1(r)^\top \sigma_x(t,r)^\top + \delta\sigma^\varepsilon(t,r)^\top\Big)\Pi(t,s,r)\mathcal{X}_1(s,r) + \mathcal{X}_1(t,r)^\top\Big[\Lambda(t,s,r)\mathcal{X}_1(s,r) + \Pi(t,s,r)\Big(\sigma_x(s,r)X_1(r) + \delta\sigma^\varepsilon(s,r)\Big)\Big].$$

By the integrability of X_1, $\mathcal{X}_1$, and (Π,Λ) in the later specific cases, as well as (H1),

$$\begin{aligned}\mathbb{E}\int_\tau^r \Big[\delta\sigma^\varepsilon(t,\theta)^\top\Big(\Lambda(t,s,\theta)\mathcal{X}_1(s,\theta) + \Pi(t,s,\theta)\sigma_x(s,\theta)X_1(\theta)\Big)\\ +\Big(\mathcal{X}_1(\theta,s)^\top\Lambda(t,s,\theta) + X_1(\theta)^\top\sigma_x(t,\theta)^\top\Pi(t,s,\theta)\Big)\delta\sigma^\varepsilon(s,\theta)\Big]d\theta = o(\varepsilon).\end{aligned} \tag{4.8}$$

Thus, (noting $\tau \leqslant r \leqslant s \leqslant t \leqslant T$, $\mathcal{X}_1(t,r)$, and $\mathcal{X}_1(s,r)$ being $\mathcal{F}_r$-measurable)

$$\begin{aligned}&\mathbb{E}\big[\mathcal{X}_1(t,r)^\top\Theta(t,s)\mathcal{X}_1(s,r)\big] = \mathbb{E}\big[\mathcal{X}_1(t,r)^\top\Pi(t,s,r)\mathcal{X}_1(s,r)\big]\\ &= o(\varepsilon) + \mathbb{E}\int_\tau^r \Big[X_1(\theta)^\top\Big(b_x(t,\theta)^\top\Theta(t,s) + \sigma_x(t,\theta)^\top\Lambda(t,s,\theta)\Big)\mathcal{X}_1(s,\theta)\\ &\qquad + \mathcal{X}_1(t,\theta)^\top\Big(\Theta(t,s)b_x(s,\theta) + \Lambda(t,s,\theta)\sigma_x(s,\theta)\Big)X_1(\theta)\\ &\qquad + X_1(\theta)^\top\sigma_x(t,\theta)^\top\Theta(t,s)\sigma_x(s,\theta)X_1(\theta) + \delta\sigma^\varepsilon(t,\theta)^\top\Theta(t,s)\delta\sigma^\varepsilon(s,\theta)\Big]d\theta.\end{aligned} \tag{4.9}$$

In the above, to get rid of the Itô's integral of $\Gamma(t,s,\cdot)$ under the expectation, some standard localization arguments via stopping times have been used. We point out that the arguments of Itô's formula on $\mathcal{X}_1$ also appeared in [30]. Now, we look at the term of $\mathscr{E}(\varepsilon)$ (see (3.6)).

Step 1. *Treatment of the first term in* $\mathscr{E}(\varepsilon)$.

Take $t = s = T$, $\Theta(T,T) = h_{xx}$ in (4.7). We denote by $(\Pi(T,T,r),\Lambda(T,T,r)) \equiv (P_1(r),Q_1(r))$ with $r\in[\tau,T]$. Then (4.7) reads

$$P_1(r) = h_{xx} - \int_r^T Q_1(\theta)dW(\theta), \qquad r\in[\tau,T]. \tag{4.10}$$

By the uniqueness, $P_1(\cdot)$ and $Q_1(\cdot)$ take values in $\mathbb{S}^n$. Making use of (3.4) and (H1), one has

$$\begin{aligned}&\mathbb{E}\int_\tau^{\tau+\varepsilon}|X_1(s)||P_1(s)||\delta\sigma(T,s)|ds\\ &\leqslant \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|X_1(s)|^4ds\Big]^{\frac14}\Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|\delta\sigma(T,s)|^4ds\Big]^{\frac14}\Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|P_1(s)|^2ds\Big]^{\frac12} = o(\varepsilon),\\ &\mathbb{E}\int_\tau^{\tau+\varepsilon}|\mathcal{X}_1(T,s)||Q_1(s)||\delta\sigma(T,s)|ds\\ &\leqslant \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|\mathcal{X}_1(T,s)|^4ds\Big]^{\frac14}\Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|\delta\sigma(T,s)|^4ds\Big]^{\frac14}\Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}|Q_1(s)|^2ds\Big]^{\frac12} = o(\varepsilon).\end{aligned}$$

Therefore, the corresponding (4.8) follows. On the other hand, in the current case, we have

$$\begin{aligned}\Gamma(T,T,r) &= \mathcal{X}_1(T,r)^\top Q_1(r)\mathcal{X}_1(T,r) + \mathcal{X}_1(T,r)^\top P_1(r)\Big(\sigma_x(T,r)X_1(r) + \delta\sigma^\varepsilon(T,r)\Big)\\ &\quad + \Big(X_1(r)^\top\sigma_x(T,r)^\top + \delta\sigma^\varepsilon(T,r)^\top\Big)P_1(r)\mathcal{X}_1(T,r).\end{aligned}$$

Again, by a localization argument, we can get rid of the Itô's integral of $\Gamma(T,T,\cdot)$ under expectation to obtain the following, which is a particular case of (4.9),

$$
\begin{aligned}
&\mathbb{E}\big[X_1(T)^\top h_{xx}X_1(T)\big]=\mathbb{E}\big[\mathcal{X}_1(T,T)^\top P_1(T)\mathcal{X}_1(T,T)\big]\\
&\quad=\mathbb{E}\int_\tau^T\Big[X_1(r)^\top F_1(r)\mathcal{X}_1(T,r)+\mathcal{X}_1(T,r)^\top F_1(r)^\top X_1(r)\\
&\qquad\qquad+X_1(r)^\top G_1(r)X_1(r)+\delta\sigma^\varepsilon(T,r)^\top P_1(r)\delta\sigma^\varepsilon(T,r)\Big]dr+o(\varepsilon).
\end{aligned}\tag{4.11}
$$

Here, both $F_1(r)$ and $G_1(r)$, defined below, are $\mathcal{F}_r$-measurable and independent of $u^\varepsilon(\cdot)$,

$$
\begin{cases}
F_1(r)\overset{\triangle}{=}b_x(T,r)^\top P_1(r)+\sigma_x(T,r)^\top Q_1(r),\\
G_1(r)\overset{\triangle}{=}\sigma_x(T,r)^\top P_1(r)\sigma_x(T,r).
\end{cases}\tag{4.12}
$$

Consequently, at the end of this step, one has

$$
\begin{aligned}
\mathscr{E}(\varepsilon)=o(\varepsilon)+\mathbb{E}\int_\tau^T&\Big[X_1(r)^\top F_1(r)\mathcal{X}_1(T,r)+\mathcal{X}_1(T,r)^\top F_1(r)^\top X_1(r)\\
&+X_1(r)^\top\Big(H_{xx}(r)+G_1(r)\Big)X_1(r)+\delta\sigma^\varepsilon(T,r)^\top P_1(r)\delta\sigma^\varepsilon(T,r)\Big]dr.
\end{aligned}\tag{4.13}
$$

With the above procedures, the term $\mathbb{E}\big[X_1(T)^\top h_{xx}X_1(T)\big]$ is absorbed. However, the new terms of the form $X_1(r)^\top F_1(r)\mathcal{X}_1(T,r)$ and its transpose appear under integral. This will be handled by the next step.

Step 2. *Treatment of the term $X_1(r)^\top F_1(r)\mathcal{X}_1(T,r)$ and its transpose.*

Take $t=T$ in (4.7) and let $\Theta(T,s)=\Theta_2(T,r)$ be undetermined (which is $\mathcal{F}_s$-measurable). By the martingale representation theorem one has

$$
\Theta_2(T,s)=\mathbb{E}_r[\Theta_2(T,s)]+\int_r^s\Lambda_2(T,s,\theta)dW(\theta),\qquad \tau\leqslant r\leqslant s\leqslant T.\tag{4.14}
$$

In this case, (4.9) reads

$$
\begin{aligned}
&\mathbb{E}\big[\mathcal{X}_1(T,r)^\top\Theta_2(T,r)X_1(r)\big]\\
&=\mathbb{E}\int_\tau^r\Big[X_1(\theta)^\top F_2(r,\theta)\mathcal{X}_1(r,\theta)+\mathcal{X}_1(T,\theta)^\top\widetilde{F}_2(r,\theta)^\top X_1(\theta)\\
&\qquad+\frac{1}{2}X_1(\theta)^\top G_2(r,\theta)X_1(\theta)+\delta\sigma^\varepsilon(T,\theta)^\top\Theta_2(T,r)\delta\sigma^\varepsilon(r,\theta)\Big]d\theta+o(\varepsilon),
\end{aligned}\tag{4.15}
$$

where

$$
\begin{cases}
F_2(r,\theta)\overset{\triangle}{=}b_x(T,\theta)^\top\Theta_2(T,r)+\sigma_x(T,\theta)^\top\Lambda_2(T,r,\theta),\\
\widetilde{F}_2(r,\theta)\overset{\triangle}{=}b_x(r,\theta)^\top\Theta_2(T,r)^\top+\sigma_x(r,\theta)^\top\Lambda_2(T,r,\theta)^\top,\\
G_2(r,\theta)\overset{\triangle}{=}\sigma_x(T,\theta)^\top\Theta_2(T,r)\sigma_x(r,\theta)+\sigma_x(r,\theta)^\top\Theta_2(T,r)^\top\sigma_x(T,\theta).
\end{cases}\tag{4.16}
$$

Thus,

$$
\begin{aligned}
&\mathbb{E}\int_\tau^T \big[\mathcal{X}_1(T,r)^\top\Theta_2(T,r)X_1(r)+X_1(r)^\top\Theta_2(T,r)^\top\mathcal{X}_1(T,r)\big]dr\\
&=\mathbb{E}\int_\tau^T\int_r^T \Big[X_1(r)^\top F_2(\theta,r)\mathcal{X}_1(\theta,r)+\mathcal{X}_1(\theta,r)F_2(\theta,r)^\top X_1(r)\\
&\qquad\qquad +\mathcal{X}_1(T,r)^\top\widetilde{F}_2(\theta,r)^\top X_1(r)+X_1(r)^\top\widetilde{F}_2(\theta,r)\mathcal{X}_1(T,r)\\
&\qquad\qquad +X_1(r)^\top G_2(\theta,r)X_1(r)+\delta\sigma^\varepsilon(T,r)^\top\Theta_2(T,\theta)\delta\sigma^\varepsilon(\theta,r)\\
&\qquad\qquad +\delta\sigma^\varepsilon(\theta,r)^\top\Theta_2(T,\theta)^\top\delta\sigma^\varepsilon(T,r)\Big]d\theta dr+o(\varepsilon).
\end{aligned}
$$

Consequently, (4.13) becomes

$$
\begin{aligned}
\mathscr{E}(\varepsilon)&=o(\varepsilon)+\mathbb{E}\int_\tau^T \Big[X_1(r)^\top\Big(F_1(r)-\Theta_2(T,r)^\top+\int_r^T\widetilde{F}_2(\theta,r)d\theta\Big)\mathcal{X}_1(T,r)\\
&+\mathcal{X}_1(T,r)^\top\Big(F_1(r)^\top-\Theta_2(T,r)+\int_r^T\widetilde{F}_2(\theta,r)^\top d\theta\Big)X_1(r)\\
&+\int_r^T\big[X_1(r)^\top F_2(\theta,r)\mathcal{X}_1(\theta,r)+\mathcal{X}_1(\theta,r)^\top F_2(\theta,r)^\top X_1(r)\big]d\theta\\
&+X_1(r)^\top\Big(H_{xx}(r)+G_1(r)+\int_r^T G_2(\theta,r)d\theta\Big)X_1(r)+\delta\sigma^\varepsilon(T,r)^\top P_1(r)\delta\sigma^\varepsilon(T,r)\\
&+\int_r^T\big[\delta\sigma^\varepsilon(T,r)^\top\Theta_2(T,\theta)\delta\sigma^\varepsilon(\theta,r)+\delta\sigma^\varepsilon(\theta,r)^\top\Theta_2(T,\theta)^\top\delta\sigma^\varepsilon(T,r)\big]d\theta\Big]dr.
\end{aligned}
$$

Thus, to eliminate the term of the form $X_1(r)^\top(\cdots)\mathcal{X}_1(T,r)$ and its transpose, we let

$$
\Theta_2(T,r)=F_1(r)^\top+\mathbb{E}_r\int_r^T\widetilde{F}_2(\theta,r)^\top d\theta.
$$

By denoting $P_2(r)\overset{\triangle}{=}\Theta_2(T,r)^\top$, $Q_2(\theta,r)\overset{\triangle}{=}\Lambda_2(T,\theta,r)^\top$, with $\tau\leqslant r\leqslant\theta\leqslant T$, we have

$$
\begin{aligned}
P_2(r)&=b_x(T,r)^\top P_1(r)+\sigma_x(T,r)^\top Q_1(r)\\
&\qquad+\int_r^T\Big(b_x(\theta,r)^\top P_2(\theta)+\sigma_x(\theta,r)^\top Q_2(\theta,r)\Big)d\theta-\int_r^T\widetilde{Q}_2(r,\theta)dW(\theta)
\end{aligned}
$$

for some $\widetilde{Q}_2(r,\theta)$ defined for $\tau\leqslant r\leqslant\theta\leqslant T$, with a suitable measurability and integrability. At the same time, (4.14) reads

$$
P_2(s)=\mathbb{E}_r[P_2(s)]+\int_r^s Q_2(s,\theta)dW(\theta),\qquad \tau\leqslant r\leqslant s\leqslant T,
$$

with $Q_2(s,\theta)$ defined on $\Delta_*[t,T]$. We now extend $Q_2(s,\theta)$ to $[\tau,T]^2$ by $Q_2(s,\theta)=\widetilde{Q}_2(s,\theta)$ with $\tau\leqslant s\leqslant\theta\leqslant T$. Then $(P_2(\cdot),Q_2(\cdot\,,\cdot))$ is the unique adapted M-solution to Type-II BSVIE:

$$
\begin{aligned}
P_2(r)&=b_x(T,r)^\top P_1(r)+\sigma_x(T,r)^\top Q_1(r)+\int_r^T\Big(b_x(\theta,r)^\top P_2(\theta)\\
&\qquad+\sigma_x(\theta,r)^\top Q_2(\theta,r)\Big)d\theta-\int_r^T Q_2(r,\theta)dW(\theta),\quad r\in[\tau,T],
\end{aligned}
\tag{4.17}
$$

which is independent of the spike variation $u^\varepsilon(\cdot)$ of $\bar u(\cdot)$. Having the above, we obtain

(4.18)
$$\begin{aligned}\mathscr{E}(\varepsilon)=o(\varepsilon)+\mathbb{E}\int_\tau^T\Bigg[&X_1(r)^\top\left(H_{xx}(r)+G_1(r)+\int_r^T G_2(\theta,r)d\theta\right)X_1(r)\\&+\int_r^T\Big(X_1(r)^\top F_2(\theta,r)\mathcal{X}_1(\theta,r)+\mathcal{X}_1(\theta,r)^\top F_2(\theta,r)^\top X_1(r)\Big)d\theta\Bigg]dr\\&+\mathbb{E}\int_\tau^T\Bigg[\delta\sigma^\varepsilon(T,r)^\top P_1(r)\delta\sigma^\varepsilon(T,r)\\&+\int_r^T\Big(\delta\sigma^\varepsilon(T,r)^\top P_2(\theta)^\top\delta\sigma^\varepsilon(\theta,r)+\delta\sigma^\varepsilon(\theta,r)^\top P_2(\theta)\delta\sigma^\varepsilon(T,r)\Big)d\theta\Bigg]dr.\end{aligned}$$

At this moment, we can express F_2, $\widetilde F_2$, G_2 in (4.16) by (P_2,Q_2) accordingly. In $\mathscr{E}(\varepsilon)$ (see (4.18)), the terms of the forms $X_1(r)^\top(\cdots)X_1(r)$ and $X_1(r)^\top(\cdots)\mathcal{X}_1(\theta,r)$, together with its transpose, need to be handled. Next, we want to get rid of those terms simultaneously.

Step 3. *Treatment of the terms $X_1(r)^\top(\cdots)X_1(r)$, $\mathcal{X}_1(\theta,r)^\top(\cdots)X_1(r)$, and its transpose.*

To treat $X_1(r)^\top(\cdots)X_1(r)$, let $t=s$ in (4.7) with $\Theta(s,s)=\Theta_3(s)$ valued in $\mathbb{S}^n$, undetermined. Again, we have $\Lambda(s,s,\cdot)=\Lambda_3(s,s,\cdot)$. Since $\Theta_3(s)$ is symmetric, so is $\Lambda_3(s,s,r)$. Now, (4.9) becomes

$$\begin{aligned}&\mathbb{E}\big[X_1(r)^\top\Theta_3(r)X_1(r)\big]\\&\quad=\mathbb{E}\int_\tau^r\Big[X_1(\theta)^\top F_3(r,\theta)\mathcal{X}_1(r,\theta)+\mathcal{X}_1(r,\theta)^\top F_3(r,\theta)^\top X_1(\theta)\\&\qquad\qquad+X_1(\theta)^\top G_3(r,\theta)X_1(\theta)+\delta\sigma^\varepsilon(r,\theta)^\top\Theta_3(r)\delta\sigma^\varepsilon(r,\theta)\Big]d\theta+o(\varepsilon),\end{aligned}$$

where

$$\begin{cases}F_3(r,\theta)\overset{\triangle}{=}b_x(r,\theta)^\top\Theta_3(r)+\sigma_x(r,\theta)^\top\Lambda_3(r,r,\theta),\\ G_3(r,\theta)\overset{\triangle}{=}\sigma_x(r,\theta)^\top\Theta_3(r)\sigma_x(r,\theta).\end{cases}\tag{4.19}$$

Then, by integrating it over $[\tau,T]$, one has

$$\begin{aligned}&\mathbb{E}\int_\tau^T\big[X_1(r)^\top\Theta_3(r)X_1(r)\big]dr\\&\quad=\mathbb{E}\int_\tau^T\int_r^T\Big[X_1(r)^\top F_3(\theta,r)\mathcal{X}_1(\theta,r)+\mathcal{X}_1(\theta,r)^\top F_3(\theta,r)^\top X_1(r)\\&\qquad\qquad+X_1(r)^\top G_3(\theta,r)X_1(r)+\delta\sigma^\varepsilon(\theta,r)^\top\Theta_3(\theta)\delta\sigma^\varepsilon(\theta,r)\Big]d\theta dr+o(\varepsilon).\end{aligned}\tag{4.20}$$

Next, for the term $\mathcal{X}_1(\theta,r)^\top(\cdots)X_1(r)$, we let $\Theta(t,s)=\Theta_4(t,s)$, $\Lambda(t,s,\cdot)=\Lambda_4(t,s,\cdot)$ be undetermined. Then, (4.9) reads

$$\begin{aligned}&\mathbb{E}\big[\mathcal{X}_1(\theta,r)^\top\Theta_4(\theta,r)X_1(r)\big]\\&\quad=\mathbb{E}\int_\tau^r\Big[X_1(\theta')^\top F_4(\theta,r,\theta')\mathcal{X}_1(r,\theta')+\mathcal{X}_1(\theta,\theta')^\top\widetilde F_4(\theta,r,\theta')^\top X_1(\theta')\\&\qquad\qquad+\frac12X_1(\theta')^\top G_4(\theta,r,\theta')X_1(\theta')+\delta\sigma^\varepsilon(\theta,\theta')^\top\Theta_4(\theta,r)\delta\sigma^\varepsilon(r,\theta')\Big]d\theta'+o(\varepsilon),\end{aligned}$$

where (note that $\tau \leqslant \theta' \leqslant r \leqslant \theta \leqslant T$)

$$
\left\{\begin{array}{l}
F_4(\theta,r,\theta') \overset{\triangle}{=} b_x(\theta,\theta')^\top \Theta_4(\theta,r) + \sigma_x(\theta,\theta')^\top \Lambda_4(\theta,r,\theta'),\\
\widetilde{F}_4(\theta,r,\theta') \overset{\triangle}{=} b_x(r,\theta')^\top \Theta_4(\theta,r)^\top + \sigma_x(r,\theta')^\top \Lambda_4(\theta,r,\theta')^\top,\\
G_4(\theta,r,\theta') \overset{\triangle}{=} \sigma_x(\theta,\theta')^\top \Theta_4(\theta,r)\sigma_x(r,\theta') + \sigma_x(r,\theta')^\top \Theta_4(\theta,r)^\top \sigma_x(\theta,\theta').
\end{array}\right. \tag{4.21}
$$

Therefore, we have

$$
\begin{aligned}
&\mathbb{E}\int_\tau^T \int_r^T \mathcal{X}_1(\theta,r)^\top \Theta_4(\theta,r) X_1(r) d\theta dr\\
&= \mathbb{E}\int_\tau^T \int_{\theta'}^T \int_\tau^{\theta'} \Big[X_1(r)^\top F_4(\theta,\theta',r)\mathcal{X}_1(\theta',r) + \mathcal{X}_1(\theta,r)^\top \widetilde{F}_4(\theta,\theta',r)^\top X_1(r)\\
&\quad + \frac{1}{2} X_1(r)^\top G_4(\theta,\theta',r) X_1(r) + \delta\sigma^\varepsilon(\theta,r)^\top \Theta_4(\theta,\theta')\delta\sigma^\varepsilon(\theta',r)\Big] dr d\theta d\theta' + o(\varepsilon)\\
&= \mathbb{E}\int_\tau^T \int_\tau^{\theta'} \int_{\theta'}^T \Big[X_1(r)^\top F_4(\theta,\theta',r)\mathcal{X}_1(\theta',r) + \mathcal{X}_1(\theta,r)^\top \widetilde{F}_4(\theta,\theta',r)^\top X_1(r)\\
&\quad + \frac{1}{2} X_1(r)^\top G_4(\theta,\theta',r) X_1(r) + \delta\sigma^\varepsilon(\theta,r)^\top \Theta_4(\theta,\theta')\delta\sigma^\varepsilon(\theta',r)\Big] d\theta dr d\theta' + o(\varepsilon)\\
&= \mathbb{E}\int_\tau^T \Bigg[\int_r^T \int_\theta^T X_1(r)^\top F_4(\theta',\theta,r)\mathcal{X}_1(\theta,r) d\theta' d\theta\\
&\quad + \int_r^T \int_r^\theta \mathcal{X}_1(\theta,r)^\top \widetilde{F}_4(\theta,\theta',r)^\top X_1(r) d\theta' d\theta\\
&\quad + \int_r^T \int_\theta^T \Big(\frac{1}{2} X_1(r)^\top G_4(\theta',\theta,r) X_1(r)\\
&\quad + \delta\sigma^\varepsilon(\theta',r)^\top \Theta_4(\theta',\theta)\delta\sigma^\varepsilon(\theta,r)\Big) d\theta' d\theta\Bigg] dr + o(\varepsilon).
\end{aligned}
$$

Hence, combining (4.18) and (4.20) with the above, we have

$$
\begin{aligned}
&\mathscr{E}(\varepsilon)\\
&= \mathbb{E}\int_\tau^T \Bigg\{X_1(r)^\top \Bigg[H_{xx}(r) + G_1(r) + \int_r^T \Bigg(G_2(\theta,r) + G_3(\theta,r) + \int_\theta^T G_4(\theta',\theta,r)d\theta'\Bigg)d\theta\\
&\quad - \Theta_3(r)\Bigg] X_1(r) + \int_r^T \Bigg[X_1(r)^\top \Bigg(F_2(\theta,r) + F_3(\theta,r) - \Theta_4(\theta,r)^\top + \int_\theta^T F_4(\theta',\theta,r)d\theta'\\
&\quad + \int_r^\theta \widetilde{F}_4(\theta,\theta',r)d\theta'\Bigg)\mathcal{X}_1(\theta,r) + \mathcal{X}_1(\theta,r)^\top \Bigg(F_2(\theta,r) + F_3(\theta,r) - \Theta_4(\theta,r)^\top\\
&\quad + \int_\theta^T F_4(\theta',\theta,r)d\theta' + \int_r^\theta \widetilde{F}_4(\theta,\theta',r)d\theta'\Bigg)^\top X_1(r)\Bigg] d\theta + \delta\sigma^\varepsilon(T,r)^\top P_1(r)\delta\sigma^\varepsilon(T,r)\\
&\quad + \int_r^T \Bigg[\delta\sigma^\varepsilon(T,r)^\top P_2(\theta)^\top \delta\sigma^\varepsilon(\theta,r) + \delta\sigma^\varepsilon(\theta,r)^\top P_2(\theta)\delta\sigma^\varepsilon(T,r)\\
&\quad + \delta\sigma^\varepsilon(\theta,r)^\top \Theta_3(\theta)\delta\sigma^\varepsilon(\theta,r) + \int_\theta^T \Big(\delta\sigma^\varepsilon(\theta',r)^\top \Theta_4(\theta',\theta)\delta\sigma^\varepsilon(\theta,r)\\
&\quad + \delta\sigma^\varepsilon(\theta,r)^\top \Theta_4(\theta',\theta)^\top \delta\sigma^\varepsilon(\theta',r)\Big)d\theta'\Bigg] d\theta\Bigg\} dr + o(\varepsilon).
\end{aligned}
$$

Consequently, in order to eliminate the term $X_1(r)^\top(\cdots)X_1(r)$, we should let

$$\Theta_3(r) = H_{xx}(r) + G_1(r) + \int_r^T \left(G_2(\theta,r) + G_3(\theta,r) + \int_\theta^T G_4(\theta',\theta,r)d\theta' \right) d\theta \\ - \int_r^T \widetilde{Q}_3(r,\theta)dW(\theta), \quad r \in [\tau,T]$$

for some $\widetilde{Q}_3(r,\theta)$ uniquely defined for almost $\tau \leqslant r \leqslant \theta \leqslant T$. Similarly, to eliminate $X_1(r)^\top(\cdots)\mathcal{X}_1(\theta,r)$ and its transpose, we should let

$$\Theta_4(\theta,r)^\top = F_2(\theta,r) + F_3(\theta,r) + \int_\theta^T F_4(\theta',\theta,r)d\theta' + \int_r^\theta \widetilde{F}_4(\theta,\theta',r)d\theta' \\ - \int_r^T \widetilde{Q}_4(\theta,r,\theta')dW(\theta'), \quad (\theta,r) \in \Delta_*[\tau,T]$$

for some $\widetilde{Q}_4(\theta,r,\theta')$ defined for $\tau \leqslant r \leqslant \theta' \leqslant T$ and $\tau \leqslant r \leqslant \theta \leqslant T$. Let

$$P_3(\theta) \stackrel{\triangle}{=} \Theta_3(\theta), \qquad Q_3(\theta,r) \stackrel{\triangle}{=} \begin{cases} \Lambda_3(\theta,\theta,r), & \tau \leqslant r \leqslant \theta \leqslant T, \\ \widetilde{Q}_3(\theta,r), & \tau \leqslant \theta \leqslant r \leqslant T, \end{cases}$$

$$P_4(\theta,\theta') \stackrel{\triangle}{=} \Theta_4(\theta,\theta')^\top, \quad Q_4(\theta,r,\theta') \stackrel{\triangle}{=} \begin{cases} \Lambda_4(\theta,r,\theta')^\top, & \tau \leqslant \theta' \leqslant r \leqslant \theta \leqslant T, \\ \widetilde{Q}_4(\theta,r,\theta')^\top, & \tau \leqslant r \leqslant \theta' \leqslant T. \end{cases}$$

Then recall the expressions of G_i, F_j, $i = 1,2,3,4$, $j = 2,3,4$, we obtain the following system for (P_1,P_2,P_3,P_4):

$$\begin{cases} P_1(t) = h_{xx} - \int_t^T Q_1(s)dW(s), \qquad \tau \leqslant t \leqslant T, \\ P_2(t) = b_x(T,t)^\top P_1(t) + \sigma_x(T,t)^\top Q_1(t) + \int_t^T \big(b_x(s,t)^\top P_2(s) \\ \qquad +\sigma_x(s,t)^\top Q_2(s,t)\big)ds - \int_t^T Q_2(t,s)dW(s), \qquad \tau \leqslant t \leqslant T, \\ P_3(t) = H_{xx}(t) + \sigma_x(T,t)^\top P_1(t)\sigma_x(T,t) \\ \qquad + \int_t^T \Big[\sigma_x(T,t)^\top P_2(s)^\top \sigma_x(s,t) + \sigma_x(s,t)^\top P_2(s)\sigma_x(T,t) \\ \qquad + \int_s^T \Big(\sigma_x(\theta,t)^\top P_4(\theta,s)^\top \sigma_x(s,t) + \sigma_x(s,t)^\top P_4(\theta,s)\sigma_x(\theta,t)\Big)d\theta\Big] ds \\ \qquad + \int_t^T \sigma_x(s,t)^\top P_3(s)\sigma_x(s,t)ds - \int_t^T Q_3(t,s)dW(s), \quad \tau \leqslant t \leqslant T, \\ P_4(r,t) = b_x(T,t)^\top P_2(r)^\top + \sigma_x(T,t)^\top Q_2(r,t)^\top + b_x(r,t)^\top P_3(r) \\ \qquad +\sigma_x(r,t)^\top Q_3(r,t) + \int_r^T \Big(b_x(s,t)^\top P_4(s,r)^\top + \sigma_x(s,t)^\top Q_4(s,r,t)^\top\Big) ds \\ \qquad + \int_t^r \Big(b_x(s,t)^\top P_4(r,s) + \sigma_x(s,t)^\top Q_4(r,s,t)\Big) ds \\ \qquad - \int_t^T Q_4(r,t,s)dW(s), \qquad \tau \leqslant t \leqslant r \leqslant T. \end{cases} \tag{4.22}$$

By saying (P_i, Q_i) to be the adapted M-solutions of the corresponding BSVIEs, the following constraints must hold:

$$
\begin{cases}
P_2(t) = \mathbb{E}_\theta[P_2(t)] + \int_\theta^t Q_2(t,s)dW(s), & \tau \leqslant \theta \leqslant t \leqslant T, \\
P_3(t) = \mathbb{E}_\theta[P_3(t)] + \int_\theta^t Q_3(t,s)dW(s), & \tau \leqslant \theta \leqslant t \leqslant T, \\
P_4(r,t) = \mathbb{E}_\theta[P_4(r,t)] + \int_\theta^t Q_4(r,t,s)dW(s), & \tau \leqslant \theta \leqslant t \leqslant r \leqslant T.
\end{cases} \tag{4.23}
$$

The above is called the *second-order adjoint equation.* With the above system, we eventually end up with the following representation theorem.

THEOREM 4.1. *Let* (H1)–(H2) *hold, and let* $(\bar{X}(\cdot), \bar{u}(\cdot))$ *be a given state-control pair. Let* (4.22) *admit an adapted M-solution* (P_i, Q_i) *(*$1 \leqslant i \leqslant 4$*) on* $[\tau, T]$ *such that*

$$
\begin{aligned}
&(P_1, Q_1) \in L^2_{\mathbb{F}}(\Omega; C([\tau,T];\mathbb{S}^n)) \times L^2_{\mathbb{F}}(\tau,T;\mathbb{S}^n), \\
&(P_2, Q_2) \in L^2_{\mathbb{F}}(\tau,T;\mathbb{R}^{n\times n}) \times L^2_{\mathbb{F}}(\Omega; L^2([\tau,T]\times[\tau,T];\mathbb{R}^{n\times n})), \\
&(P_3, Q_3) \in L^2_{\mathbb{F}}(\tau,T;\mathbb{S}^n) \times L^2_{\mathbb{F}}(\Omega; L^2([\tau,T]\times[\tau,T];\mathbb{S}^{n\times n})), \\
&(P_4, Q_4) \in L^2_{\mathbb{F}}(\Omega; L^2(\Delta_*[\tau,T];\mathbb{R}^{n\times n})) \times L^2_{\mathbb{F}}(\Omega; L^2(\Delta_*[\tau,T]\times[\tau,T];\mathbb{R}^{n\times n})).
\end{aligned}
$$

Then $\mathscr{E}(\varepsilon)$ *admits the following representation:*

$$
\begin{aligned}
\mathscr{E}(\varepsilon) = \mathbb{E}\int_\tau^T \Bigg\{ & \delta\sigma^\varepsilon(T,r)^\top P_1(r)\delta\sigma^\varepsilon(T,r) + \int_r^T \Bigg[\delta\sigma^\varepsilon(T,r)^\top P_2(\theta)^\top \delta\sigma^\varepsilon(\theta,r) \\
& + \delta\sigma^\varepsilon(\theta,r)^\top P_2(\theta)\delta\sigma^\varepsilon(T,r) + \delta\sigma^\varepsilon(\theta,r)^\top P_3(\theta)\delta\sigma^\varepsilon(\theta,r) \\
& + \int_\theta^T \Big(\delta\sigma^\varepsilon(\theta',r)^\top P_4(\theta',\theta)^\top \delta\sigma^\varepsilon(\theta,r) \\
& + \delta\sigma^\varepsilon(\theta,r)^\top P_4(\theta',\theta)\delta\sigma^\varepsilon(\theta',r)\Big)d\theta'\Bigg]d\theta\Bigg\}dr + o(\varepsilon).
\end{aligned}
$$

This conclusion naturally leads to the maximum principle in section 6.

5. Well-posedness of the second order adjoint equations.

In this section, we establish the well-posedness of system (4.22). To begin with, let us make some observations. The existence and uniqueness of (P_1, Q_1) is easy to see, while the case of (P_2, Q_2) follows from the BSVIEs theory in [36]. For the equation of (P_3, Q_3), it is a classical linear BSVIE with given (P_1, P_2, P_4). However, as to that of (P_4, Q_4), its well-posedness cannot be given by the current theories. Therefore, we need to carefully treat it.

Note that by applying conditional expectation operator $\mathbb{E}_t$ on both sides of the fourth equation in (4.22), we end up with the following form system (noting the third equality in (4.23)):

$$
\begin{cases}
P(r,t) = F(r,t) + \int_r^T \big[b_x(s,t)^\top \mathbb{E}_t P(s,r)^\top + \sigma_x(s,t)^\top Q(s,r,t)^\top\big]ds \\
\qquad\qquad + \int_t^r \big[b_x(s,t)^\top \mathbb{E}_t P(r,s) + \sigma_x(s,t)^\top Q(r,s,t)\big]ds, \quad r \geqslant t, \\
\mathbb{E}_\theta P(r,t) = P(r,t) - \int_\theta^t Q(r,t,s)dW(s), \quad \tau \leqslant \theta \leqslant t \leqslant r \leqslant T,
\end{cases} \tag{5.1}
$$

with $P(r,t)=P_4(r,t)$, $Q(r,t,s)=Q_4(r,t,s)$, and

$$F(r,t)\overset{\triangle}{=}\mathbb{E}_t\Big[b_x(T,t)^\top P_2(r)^\top+\sigma_x(T,t)^\top Q_2(r,t)^\top+b_x(r,t)^\top P_3(r)+\sigma_x(r,t)^\top Q_3(r,t)\Big]. \tag{5.2}$$

The following gives the wellposedness for (5.1).

LEMMA 5.1. *Let $b_x(\cdot,\cdot),\sigma_x(\cdot,\cdot)$ be bounded, and let $F(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\tau,T]))$. Then (5.1) admits a unique solution $P(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\tau,T]))$. In addition, for any $\rho\in[\tau,T]$, the following estimate holds:*

$$\mathbb{E}\int_\rho^T\int_\rho^r|P(r,t)|^2dtdr\leqslant K\mathbb{E}\int_\rho^T\int_\rho^r|F(r,t)|^2dtdr, \tag{5.3}$$

where K only depends on $\|b_x\|_\infty$, $\|\sigma_x\|_\infty$, and T, but not on ρ.

Proof. Let $\rho\in[\tau,T]$ be fixed. For any $p(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\rho,T]))$, we look at

$$\begin{cases} P(r,t)=F(r,t)+\displaystyle\int_r^T\big[b_x(s,t)^\top\mathbb{E}_tp(s,r)^\top+\sigma_x(s,t)^\top q(s,r,t)^\top\big]ds\\ \qquad\qquad+\displaystyle\int_t^r\big[b_x(s,t)^\top\mathbb{E}_tp(r,s)+\sigma_x(s,t)^\top q(r,s,t)\big]ds,\quad r\geqslant t,\\ \mathbb{E}_\theta p(r,t)=p(r,t)-\displaystyle\int_\theta^t q(r,t,s)dW(s),\qquad \rho\leqslant\theta\leqslant t\leqslant r\leqslant T.\end{cases} \tag{5.4}$$

By the first equality in (5.4), we have (for any $\beta>0$)

$$\begin{aligned}&\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}|P(r,t)|^2dtdr\leqslant K\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}|F(r,t)|^2dtdr\\ &\quad+K\big\|b_x(\cdot,\cdot)\big\|_\infty^2\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r\big|p(r,s)\big|ds\right)^2+\left(\int_r^T\big|p(s,r)\big|ds\right)^2\right\}dtdr\\ &\quad+K\big\|\sigma_x(\cdot,\cdot)\big\|_\infty^2\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r\big|q(r,s,t)\big|ds\right)^2+\left(\int_r^T\big|q(s,r,t)\big|ds\right)^2\right\}dtdr.\end{aligned} \tag{5.5}$$

For the second term on the right-hand side of (5.5), by Fubini's theorem, we have

$$\begin{aligned}&\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r\big|p(r,s)\big|ds\right)^2+\left(\int_r^T\big|p(s,r)\big|ds\right)^2\right\}dtdr\\ &\quad\leqslant\frac{1}{\beta}\mathbb{E}\int_\rho^T\int_\rho^r\int_\rho^s\left(e^{\beta(r+s)}-e^{\beta(t+s)}\right)|p(r,s)|^2dtdsdr\\ &\qquad+\frac{1}{\beta}\mathbb{E}\int_\rho^T\int_r^T\int_\rho^r\left(e^{\beta(t+s)}-e^{\beta(r+s+t-T)}\right)|p(s,r)|^2dtdsdr\\ &\quad\leqslant\frac{2T}{\beta}\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+s)}|p(r,s)|^2dsdr.\end{aligned}$$

For the third term on the right-hand side of (5.5), similarly, by Fubini's theorem,

$$\begin{aligned}
&\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r |q(r,s,t)|ds\right)^2+\left(\int_r^T |q(s,r,t)|ds\right)^2\right\}dtdr\\
&\quad\leqslant \mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r e^{-\beta s}ds\right)\left(\int_t^r e^{\beta s}|q(r,s,t)|^2ds\right)\right.\\
&\qquad\qquad\left.+\left(\int_r^T e^{-\beta s}ds\right)\left(\int_r^T e^{\beta s}|q(s,r,t)|^2ds\right)\right\}dtdr\\
&\quad\leqslant \frac{2}{\beta}\mathbb{E}\int_\rho^T\int_\rho^r\int_\rho^s e^{\beta(r+s)}|q(r,s,t)|^2dtdsdr\\
&\quad\leqslant \frac{2}{\beta}\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+s)}|p(r,s)|^2dsdr.
\end{aligned}$$

To sum up the above arguments, we have $P(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\rho,T]))$, and one can define a map Ξ from $L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\rho,T]))$ to itself as $\Xi(p)=P$.

Suppose $\bar p,\widetilde p$ are two elements in $L^2_{\mathbb{F}}(\Omega;L^2(\Delta_*[\rho,T]))$, and $(\bar q,\widetilde q)$ is defined similarly as in the second equality in (5.4). By the previous arguments, we have $\bar P=\Xi(\bar p)$, $\widetilde P=\Xi(\widetilde p)$, and define $\bar Q$, $\widetilde Q$ accordingly. Also let

$$\widehat p \overset{\triangle}{=} p-\widetilde p,\quad \widehat q \overset{\triangle}{=} q-\widetilde q,\quad \widehat P \overset{\triangle}{=} P-\widetilde P,\quad \widehat Q \overset{\triangle}{=} Q-\widetilde Q.$$

Then we have

$$\text{(5.6)}\qquad \begin{cases}
\widehat P(r,t)=\displaystyle\int_r^T\big[b_x(s,t)^\top\mathbb{E}_t\widehat p(s,r)^\top+\sigma_x(s,t)^\top\widehat q(s,r,t)^\top\big]ds\\
\qquad\qquad +\displaystyle\int_t^r\big[b_x(s,t)^\top\mathbb{E}_t\widehat p(r,s)+\sigma_x(s,t)^\top\widehat q(r,s,t)\big]ds,\quad r\geqslant t,\\
\mathbb{E}_\theta\widehat p(r,t)=\widehat p(r,t)-\displaystyle\int_\theta^t\widehat q(r,t,s)dW(s),\quad \rho\leqslant\theta\leqslant t\leqslant r\leqslant T.
\end{cases}$$

By the first equality in (5.6), for any $\rho\in[\tau,T]$,

(5.7)

$$\begin{aligned}
&\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}|\widehat P(r,t)|^2dtdr\\
&\quad\leqslant K\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r |\widehat p(r,s)|ds\right)^2+\left(\int_r^T |\widehat p(s,r)|ds\right)^2\right\}dtdr\\
&\qquad+K\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}\left\{\left(\int_t^r |\widehat q(r,s,t)|ds\right)^2+\left(\int_r^T |\widehat q(s,r,t)|ds\right)^2\right\}dtdr.
\end{aligned}$$

Similar as the estimation of (5.5), we have for any $\rho\in[\tau,T]$,

$$\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}|\widehat P(r,t)|^2dtdr\leqslant\frac{K}{\beta}\mathbb{E}\int_\rho^T\int_\rho^r e^{\beta(r+t)}|\widehat p(r,t)|^2dtdr.$$

Therefore, by choosing $\beta>0$ large enough, we obtain the existence and uniqueness of P on $[\rho,T]$, as well as the above conclusion (5.3). ∎

Now we return to system (4.22), the second-order adjoint system in our scenario.

THEOREM 5.2. *Suppose* (H1)–(H2) *hold. Then* (4.22) *admits a unique adapted M-solution* (P_i, Q_i), $1 \leqslant i \leqslant 4$, *such that*

$$(P_3, P_4) \in L^2_{\mathbb{F}}(\tau, T; \mathbb{R}^{n\times n}) \times L^2_{\mathbb{F}}(\Omega; L^2(\Delta_*[\tau, T]; \mathbb{R}^{n\times n})).$$

Proof. By standard BSVIEs theory [36], one has the following:

$$(P_1, Q_1) \in L^2_{\mathbb{F}}(\Omega; C([\tau, T]; \mathbb{S}^n)) \times L^2_{\mathbb{F}}(\tau, T; \mathbb{S}^n),$$
$$(P_2, Q_2) \in L^2_{\mathbb{F}}(\tau, T; \mathbb{R}^{n\times n}) \times L^2(\tau, T; L^2_{\mathbb{F}}(\tau, T; \mathbb{R}^{n\times n})).$$

We now prove the remaining well-posedness of (P_3, P_4). Recall (4.6), we see that $H_{xx}(\cdot) \in L^2_{\mathbb{F}}(\tau, T; \mathbb{R}^{n\times n})$. By applying conditional expectation operator $\mathbb{E}_t$ on both sides of the equations of (P_3, P_4) in (4.22), we have

$$\begin{cases} P_3(t)=\mathbf{F}_3(t)+\mathbb{E}_t\int_t^T\int_s^T\Big(\sigma_x(\theta,t)^\top P_4(\theta,s)^\top\sigma_x(s,t)+\sigma_x(s,t)^\top P_4(\theta,s)\sigma_x(\theta,t)\Big)d\theta ds\\ \qquad\qquad +\mathbb{E}_t\int_t^T\sigma_x(s,t)^\top P_3(s)\sigma_x(s,t)ds, \quad \tau\leqslant t\leqslant T,\\ P_4(r,t)=\mathbf{F}_4(r,t)+b_x(r,t)^\top\mathbb{E}_tP_3(r)+\sigma_x(r,t)^\top Q_3(r,t)\\ \qquad\qquad +\mathbb{E}_t\int_r^T\Big(b_x(s,t)^\top P_4(s,r)^\top+\sigma_x(s,t)^\top Q_4(s,r,t)^\top\Big)ds\\ \qquad\qquad +\mathbb{E}_t\int_t^r\Big(b_x(s,t)^\top P_4(r,s)+\sigma_x(s,t)^\top Q_4(r,s,t)\Big)ds, \quad \tau\leqslant t\leqslant r\leqslant T,\end{cases}$$

where $\mathbf{F}_3(\cdot)$, $\mathbf{F}_4(\cdot\,,\cdot)$ are given by the following, depending on (P_1, Q_1), (P_2, Q_2),

$$\begin{aligned}\mathbf{F}_3(t) &\stackrel{\triangle}{=} H_{xx}(t)+\sigma_x(T,t)^\top P_1(t)\sigma_x(T,t)\\ &\quad +\mathbb{E}_t\int_t^T\big[\sigma_x(s,t)^\top P_2(s)\sigma_x(T,t)+\sigma_x(T,t)^\top P_2(s)^\top\sigma_x(s,t)\big]ds,\\ \mathbf{F}_4(r,t) &\stackrel{\triangle}{=} b_x(T,t)^\top\mathbb{E}_tP_2(r)^\top+\sigma_x(T,t)^\top Q_2(r,t)^\top.\end{aligned}$$

Thus,

$$\mathbb{E}\left(\int_\tau^T|\mathbf{F}_3(t)|^2dt+\int_\tau^T\int_\tau^r|\mathbf{F}_4(r,t)|^2dtdr\right)<\infty.$$

For a given $p_3(\cdot) \in L^2_{\mathbb{F}}(\tau, T) \equiv L^2_{\mathbb{F}}(\tau, T; \mathbb{R}^{n\times n})$, let the associated $q_3(\cdot\,,\cdot)$ be determined by martingale representation theorem:

$$p_3(t)=\mathbb{E}_rp_3(t)+\int_r^t q_3(t,s)dW(s), \qquad \tau\leqslant r\leqslant t\leqslant T.$$

Then

$$\mathbb{E}|p_3(t)|^2=\mathbb{E}|\mathbb{E}_rp_3(t)|^2+\mathbb{E}\int_r^t|q_3(t,s)|^2ds\geqslant\mathbb{E}\int_r^t|q_3(t,s)|^2ds. \tag{5.8}$$

Now, let us consider

$$
\tag{5.9}
\begin{cases}
P_3(t) = \mathbf{F}_3(t) + \mathbb{E}_t \int_t^T \int_s^T \Big(\sigma_x(\theta,t)^\top P_4(\theta,s)^\top \sigma_x(s,t) + \sigma_x(s,t)^\top P_4(\theta,s)\sigma_x(\theta,t)\Big) d\theta ds \\
\qquad\qquad + \mathbb{E}_t \int_t^T \sigma_x(s,t)^\top p_3(s)\sigma_x(s,t)ds, \quad \tau \leqslant t \leqslant T, \\
P_4(r,t) = \mathbf{F}_4(r,t) + b_x(r,t)^\top \mathbb{E}_t p_3(r) + \sigma_x(r,t)^\top q_3(r,t) \\
\qquad\qquad + \mathbb{E}_t \int_r^T \Big(b_x(s,t)^\top P_4(s,r)^\top + \sigma_x(s,t)^\top Q_4(s,r,t)^\top\Big) ds \\
\qquad\qquad + \mathbb{E}_t \int_t^r \Big(b_x(s,t)^\top P_4(r,s) + \sigma_x(s,t)^\top Q_4(r,s,t)\Big) ds, \quad \tau \leqslant t \leqslant r \leqslant T.
\end{cases}
$$

By Lemma 5.1, the second equation in (5.9) is solvable and (noting (5.8))

$$
\tag{5.10}
\begin{aligned}
&\mathbb{E}\int_\rho^T \int_\rho^r |P_4(r,t)|^2 dt dr \\
&\quad \leqslant K\mathbb{E}\left(\int_\rho^T \int_\rho^r |\mathbf{F}_4(r,t)|^2 dt dr + \int_\rho^T |p_3(r)|^2 dr + \int_\rho^T \int_\rho^r |q_3(r,t)|^2 dt dr\right) \\
&\quad \leqslant K\mathbb{E}\left(\int_\rho^T \int_\rho^r |\mathbf{F}_4(r,t)|^2 dt dr + \int_\rho^T |p_3(r)|^2 dr\right) < \infty \qquad \forall \rho \in [\tau, T].
\end{aligned}
$$

Therefore, for any given $\beta > 0$,

$$
\begin{aligned}
&\mathbb{E}\int_\tau^T e^{\beta t}\left|\int_t^T \int_s^T \sigma_x(s,t)^\top P_4(\theta,s)\sigma_x(\theta,t) d\theta ds\right|^2 dt \\
&\quad \leqslant K\mathbb{E}\int_\tau^T e^{\beta t}\left(\int_t^T \int_t^r |\mathbf{F}_4(r,t)|^2 dt dr + \int_t^T |p_3(r)|^2 dr\right) dt \\
&\quad \leqslant \frac{Ke^{\beta T}}{\beta}\mathbb{E}\int_\tau^T \int_\tau^r |\mathbf{F}_4(r,t)|^2 dt dr + \frac{K}{\beta}\mathbb{E}\int_\tau^T e^{\beta r}|p_3(r)|^2 dt dr < \infty.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}\int_\tau^T e^{\beta t}\left|\int_t^T \int_s^T \sigma_x(\theta,t)^\top P_4(\theta,s)^\top \sigma_x(s,t) d\theta ds\right|^2 dt \\
&\quad \leqslant \frac{Ke^{\beta T}}{\beta}\mathbb{E}\int_\tau^T \int_\tau^r |\mathbf{F}_4(r,t)|^2 dt dr + \frac{K}{\beta}\mathbb{E}\int_\tau^T e^{\beta r}|p_3(r)|^2 dt dr < \infty.
\end{aligned}
$$

Thus, for the equality of $P_3(\cdot)$ in (5.9), for any given $\beta > 0$, we have (noting (5.8))

$$
\tag{5.11}
\begin{aligned}
&\mathbb{E}\int_\tau^T e^{\beta t}|P_3(t)|^2 dt \\
&\quad \leqslant K\mathbb{E}\left[e^{\beta T}\int_\tau^T |\mathbf{F}_3(t)|^2 dt + \int_\tau^T e^{\beta t}\left(\int_t^T \int_t^\theta |P_4(\theta,s)|^2 ds d\theta + \int_t^T |p_3(s)|^2 ds\right) dt\right] \\
&\quad \leqslant K\mathbb{E}\left[e^{\beta T}\int_\tau^T |\mathbf{F}_3(t)|^2 dt + \frac{e^{\beta T}}{\beta}\int_\tau^T \int_\tau^r |\mathbf{F}_4(r,s)|^2 ds dr + \frac{1}{\beta}\int_\tau^T e^{\beta s}|p_3(s)|^2 dt ds\right].
\end{aligned}
$$

Hence, we have $P_3(\cdot) \in L^2_{\mathbb{F}}(\tau, T)$, and the map $\Xi(p_3) = P_3$ in $L^2_{\mathbb{F}}(\tau, T)$ is well-defined.

Now, suppose $\bar p_3$, $\widetilde p_3$ are two elements in $L^2_{\mathbb F}(\tau,T)$. We then have $\Xi(\bar p_3)=\bar P_3$, $\Xi(\widetilde p_3)=\widetilde P_3$, where $\bar P_3$, $\widetilde P_3$ satisfies (5.9) associated with $\bar p_3$, $\widetilde p_3$, respectively, and define

$$\widehat p_3\overset{\triangle}{=}\bar p_3-\widetilde p_3,\quad \widehat P_3\overset{\triangle}{=}\bar P_3-\widetilde P_3,\quad \widehat P_4\overset{\triangle}{=}\bar P_2-\widetilde P_4.$$

Then for any $t\in[\tau,T]$, by (5.10)–(5.11) (with the corresponding $\mathbf F_3(\cdot)=0$ and $\mathbf F_4(\cdot\,,\cdot)=0$),

$$\mathbb E\int_\tau^T e^{\beta r}|\widehat P_3(r)|^2dr\leqslant\frac{K}{\beta}\mathbb E\int_\tau^T e^{\beta t}|\widehat p_3(t)|^2dt.$$

By choosing β large, we see that the map Ξ is a contraction. Hence, (5.9) admits a unique fixed point $P_3(\cdot)$. The rest of the conclusion follows easily. ∎

Remark 5.3. In the above two results, there are two crucial points, i.e., the introduction of equivalent β-norm of $L^2_{\mathbb F}(\Omega;L^2(\Delta_*[\tau,T]))$ and $L^2_{\mathbb F}(\tau,T)$, respectively, and the subtle use of Lemma 5.1 in Theorem 5.2. We point out that the idea of using multiplier $e^{\beta t}$ (β-argument, for short) appeared in the literature (e.g., [26, 32]).

6. The Pontryagin's type maximum principles and related extensions. In this section, let us first give the main result of this paper, maximum principle of optimal controls. Then we show that the result is consistent with the maximum principle of FSDEs. At last we present some extensions to multiobjective problem and multiperson dynamic games of governed by FSVIEs.

6.1. Maximum principle for FSVIEs. To begin with, let us recall the above (3.2) and (3.3). The Pontryagin maximum principle for problem (C) can be stated as follows.

THEOREM 6.1. *Let* (H1)–(H2) *hold. Let* $(\bar X(\cdot),\bar u(\cdot))$ *be an optimal pair of problem* (C). *Then*

(6.1)
$$\begin{aligned}&H(\tau,\bar X(\tau),u,\eta(\tau),\zeta(\tau),Y(\cdot),Z(\cdot\,,\tau))-H(\tau,\bar X(\tau),\bar u(\tau),\eta(\tau),\zeta(\tau),Y(\cdot),Z(\cdot\,,\tau))\\&\quad+\frac12\Big\{\delta\sigma(T,\tau)^\top P_1(\tau)\delta\sigma(T,\tau)+\mathbb E_\tau\int_\tau^T\Big[\delta\sigma(\theta,\tau)^\top P_3(\theta)\delta\sigma(\theta,\tau)\\&\qquad+\delta\sigma(T,\tau)^\top P_2(\theta)^\top\delta\sigma(\theta,\tau)+\delta\sigma(\theta,\tau)^\top P_2(\theta)\delta\sigma(T,\tau)\\&\quad+\int_\theta^T\big[\delta\sigma(\theta',\tau)^\top P_4(\theta',\theta)^\top\delta\sigma(\theta,\tau)+\delta\sigma(\theta,\tau)^\top P_4(\theta',\theta)\delta\sigma(\theta',\tau)\big]d\theta'\Big]d\theta\Big\}\geqslant0,\\&\hspace{20em}\text{a.s., a.e. }\tau\in[0,T]\ \forall u\in U,\end{aligned}$$

where H is defined in (3.3), $(\eta(\cdot),\zeta(\cdot),\bar Y(\cdot),\bar Z(\cdot))$ *satisfies BSVIE* (3.2), $(P_i(\cdot),Q_i(\cdot))$ ($i=1,2,3,4$) *is the unique adapted M-solution of the second-order adjoint equation* (4.22).

Proof. Recall that

$$\delta\sigma^\varepsilon(t,s)\overset{\triangle}{=}\big[\sigma(t,s,\bar X(s),u)-\sigma(t,s,\bar X(s),\bar u(s))\big]\mathbf 1_{[\tau,\tau+\varepsilon]}(s). \tag{6.2}$$

Thanks to (6.2) and Lebesgue differentiability theorem, for almost $\tau \in [0,T]$, by Theorem 4.1

$$\lim_{\varepsilon\to 0}\frac{\mathscr{E}(\varepsilon)}{\varepsilon}=\mathbb{E}\bigg\{\delta\sigma(T,\tau)^\top P_1(\tau)\delta\sigma(T,\tau)+\mathbb{E}_\tau\int_\tau^T\bigg[\delta\sigma(T,\tau)^\top P_2(\theta)^\top\delta\sigma(\theta,\tau)$$
$$+\delta\sigma(\theta,\tau)^\top P_2(\theta)\delta\sigma(T,\tau)+\delta\sigma(\theta,\tau)^\top P_3(\theta)\delta\sigma(\theta,\tau)d\theta$$
$$+\int_\theta^T\big[\delta\sigma(\theta',\tau)^\top P_4(\theta',\theta)^\top\delta\sigma(\theta,\tau)+\delta\sigma(\theta,\tau)^\top P_4(\theta',\theta)\delta\sigma(\theta',\tau)\big]d\theta'\bigg]d\theta\bigg\}.$$

According to Theorems 3.1 and 5.2, by the optimality of $\bar u(\cdot)$ and the arbitrariness of $u\in U$, we obtain the above maximum condition (6.1) immediately. ∎

Remark 6.2. When the control domain is convex, our conclusion reduces to that in Yong [36]. We refer to Wang [31, Subsection 3.4.2] for more details along this line.

6.2. The case of FSDEs. In this subsection, we show that the above maximum principle reduces to the FSDEs case presented in [24, 38], when the following holds:

$$b(t,s,x,u)=b(s,x,u),\qquad \sigma(t,s,x,u)=\sigma(s,x,u),\qquad \varphi(t)=x. \tag{6.3}$$

Let us look at the second-order adjoint system (4.22) under (6.3). The first equation is unchanged. The second, third, and fourth equations become, respectively,

$$\begin{cases} P_2(t)=b_x(t)^\top\mathcal{P}_2(t)+\sigma_x(t)^\top\mathcal{Q}_2(t)-\displaystyle\int_t^T Q_2(t,s)dW(s),\\ P_3(t)=H_{xx}(t)+\sigma_x(t)^\top\mathcal{P}_3(t)\sigma_x(t)-\displaystyle\int_t^T Q_3(t,s)dW(s),\\ P_4(r,t)=b_x(t)^\top\mathcal{P}_4(r,t)+\sigma_x(t)^\top\mathcal{Q}_4(r,t)-\displaystyle\int_t^T Q_4(r,t,s)dW(s),\end{cases}$$

where

$$\begin{cases} \mathcal{P}_2(t)=P_1(t)+\displaystyle\int_t^T P_2(s)ds,\qquad \mathcal{Q}_2(t)=Q_1(t)+\displaystyle\int_t^T Q_2(s,t)ds,\\ \mathcal{P}_3(t)=P_1(t)+\displaystyle\int_t^T\Big[P_2(s)^\top+P_2(s)+P_3(s)+\int_s^T\Big(P_4(\theta,s)^\top+P_4(\theta,s)\Big)d\theta\Big]ds,\\ \mathcal{P}_4(r,t)=P_2(r)^\top+P_3(r)+\displaystyle\int_r^T P_4(s,r)^\top ds+\int_t^r P_4(r,s)ds,\\ \mathcal{Q}_4(r,t)=Q_2(r,t)^\top+Q_3(r,t)+\displaystyle\int_r^T Q_4(s,r,t)^\top ds+\int_t^r Q_4(r,s,t)ds.\end{cases}$$

The Hamiltonian takes the form

$$\begin{aligned}&H(s,x,u,\eta(s),\zeta(s),Y(\cdot),Z(\cdot\,,s))\\ &=\langle\,\mathcal{Y}(s),b(s,x,u)\,\rangle+\langle\,\mathcal{Z}(s),\sigma(s,x.u)\,\rangle+g(s,x,u)\equiv\mathcal{H}(s,x,u,\mathcal{Y}(s),\mathcal{Z}(s)),\end{aligned}$$

where

$$\mathcal{Y}(s)=\eta(s)+\mathbb{E}_s\int_s^T Y(t)dt,\qquad \mathcal{Z}(s)=\zeta(s)+\mathbb{E}_s\int_s^T Z(t,s)dt.$$

The maximum condition becomes

$$
\begin{aligned}
0 \leqslant\, & \mathcal{H}(\tau,\bar X(\tau),u,\mathcal{Y}(\tau),\mathcal{Z}(\tau)) - \mathcal{H}(\tau,\bar X(\tau),\bar u(\tau),\mathcal{Y}(\tau),\mathcal{Z}(\tau)) \\
& + \frac{1}{2}\delta\sigma(\tau)^\top \mathbb{E}_\tau \mathcal{P}_3(\tau)\delta\sigma(\tau).
\end{aligned} \tag{6.4}
$$

Next, we check that $(\mathcal{Y},\mathcal{Z})$ satisfies the following first-order adjoint equation:

$$
\begin{cases} d\mathcal{Y}(t) = -\big[g_x(t)^\top + b_x(t)^\top \mathcal{Y}(t) + \sigma_x(t)^\top \mathcal{Z}(t)\big]dt + \mathcal{Z}(t)dW(t), \quad t\in[0,T], \\ \mathcal{Y}(T) = h_x(\bar X(T))^\top. \end{cases} \tag{6.5}
$$

Note that in the current case, we have

$$
\begin{aligned}
Y(t) = g_x(t)^\top + b_x(t)^\top \left(h_x^\top + \int_t^T Y(s)ds \right) + \sigma_x(t)^\top \left(\zeta(t) + \int_t^T Z(s,t)ds \right) \\
- \int_t^T Z(t,s)dW(s).
\end{aligned}
$$

Thus, applying $\mathbb{E}_t$ on both sides, we get

$$
Y(t) = g_x(t)^\top + b_x(t)^\top \mathcal{Y}(t) + \sigma_x(t)^\top \mathcal{Z}(t). \tag{6.6}
$$

By Fubini's theorem,

$$
\int_t^T Y(s)ds = \mathbb{E}_t \int_t^T Y(s)ds + \int_t^T \int_r^T Z(s,r)ds dW(r).
$$

Hence

$$
h_x^\top + \int_t^T Y(s)ds = \mathcal{Y}(t) + \int_t^T \mathcal{Z}(r)dW(r).
$$

Plugging (6.6) into the above, we have BSDE (6.5), the first-order adjoint equation.

Finally, it suffices to prove that $\mathscr{M}(\cdot) \overset{\triangle}{=} \mathbb{E}.\mathcal{P}_3(\cdot)$ satisfies the second-order adjoint equation, i.e., for $r\in[0,T]$,

$$
\begin{aligned}
\mathscr{M}(r) =\, & h_{xx}(\bar X(T)) + \int_r^T \Big[b_x(t)^\top \mathscr{M}(t) + \sigma_x(t)^\top \mathscr{N}(t) + \mathscr{M}(t)b_x(t) \\
& + \mathscr{N}(t)\sigma_x(t) + H_{xx}(t) + \sigma_x(t)^\top \mathscr{M}(t)\sigma_x(t)\Big]dt - \int_r^T \mathscr{N}(t)dW(t),
\end{aligned} \tag{6.7}
$$

where

$$
\mathscr{N}(t) \overset{\triangle}{=} \mathcal{Q}_2(t) + \int_t^T \mathcal{Q}_4(s,t)ds. \tag{6.8}
$$

To this end, we observe the following: For the equation of P_2, one has

$$
\begin{aligned}
& \int_r^T \big[P_2(s) + P_2(s)^\top\big]ds \\
& \quad = \int_r^T \Big\{ b_x(s)^\top \mathbb{E}_s \mathcal{P}_2(s) + \sigma_x(s)^\top \mathcal{Q}_2(s) + \mathbb{E}_s \mathcal{P}_2(s)^\top b_x(s) + \mathcal{Q}_2(s)^\top \sigma_x(s) \Big\} ds.
\end{aligned}
$$

For the equation of P_3, one has

$$\int_r^T P_3(t)dt = \int_r^T H_{xx}(t)dt + \int_r^T \sigma_x(t)^\top \mathbb{E}_t \mathcal{P}_3(t)\sigma_x(t)dt.$$

For the equation of P_4, we have

$$\int_r^T \int_t^T P_4(s,t)dsdt = \int_r^T \int_t^T \Big[b_x(t)^\top \mathbb{E}_s \mathcal{P}_4(s,t) + \sigma_x(t)^\top \mathcal{Q}_4(s,t)\Big]dsdt.$$

Therefore,

$$\begin{aligned}\mathcal{P}_3(r) = h_{xx}(\bar{X}(T)) + \int_r^T \Big[& b_x(t)^\top \mathscr{M}(t) + \sigma_x(t)^\top \mathscr{N}(t) + \mathscr{M}(t)b_x(t) \\ & + \mathscr{N}(t)\sigma_x(t) + H_{xx}(t) + \sigma_x(t)^\top \mathscr{M}(t)\sigma_x(t)\Big]dt,\end{aligned} \tag{6.9}$$

where $\mathscr{N}(\cdot)$ is defined in (6.8).

By the definitions of Q_2, Q_3, and Fubini's theorem, we see that

$$\int_r^T P_i(t)dt = \mathbb{E}_r \int_r^T P_i(t)dt + \int_r^T \int_s^T Q_i(t,s)dtdW(s), \quad i = 2,3. \tag{6.10}$$

Using Fubini's theorem, we have

$$\int_r^T \int_t^T \int_r^t Q_4(s,t,\theta)dW(\theta)dsdt = \int_r^T \int_\theta^T \int_t^T Q_4(s,t,\theta)dsdtdW(\theta).$$

Therefore,

$$\begin{aligned}\int_r^T \int_t^T \big[P_4(s,t) + P_4(s,t)^\top\big]dsdt &= \mathbb{E}_r \int_r^T \int_t^T \big[P_4(s,t) + P_4(s,t)^\top\big]dsdt \\ &+ \int_r^T \int_\theta^T \int_t^T \big[Q_4(s,t,\theta) + Q_4(s,t,\theta)^\top\big]dsdtdW(\theta).\end{aligned} \tag{6.11}$$

Combining (6.10) and (6.11), we have

$$\mathcal{P}_3(r) = \mathscr{M}(r) + \int_r^T \mathscr{N}(\theta)dW(\theta).$$

Consequently, by (6.9) we have BSDE (6.7), the second-order adjoint equation.

To sum up, the maximum condition (6.4), the first-order adjoint equation (6.5) and the second-order adjoint equation (6.7) form the Pontryagin's maximum principle for optimal control of FSDEs.

6.3. Multiperson dynamic games for FSVIEs. For integer $N \geqslant 2$, we consider an N-person dynamic game for an FSVIE. In this case, the cost functional is defined as in (1.2), where $h:\Omega\times\mathbb{R}^n \to \mathbb{R}^N$, $g:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^N$ are vector-valued functions. For simplicity, we only discuss the noncooperative dynamic games, namely, the ℓth player in the game wants to minimize his/her own cost functional $J^\ell(u(\cdot))$ $(1 \leqslant \ell \leqslant N)$, regardless of other players' cost functional.

Let $u(\cdot) = (u_1(\cdot), u_2(\cdot), \ldots, u_N(\cdot))$ with $u_\ell(\cdot) \in \mathcal{U}_{ad}^\ell$. Here $\mathcal{U}_{ad}^\ell$ is defined associated with $U_\ell \subset \mathbb{R}^{m_\ell}$, $1 \leqslant \ell \leqslant N$. For notational simplicity, let

$$(u_\ell^c, v) \overset{\triangle}{=} (u_1, \ldots, u_{\ell-1}, v, u_{\ell+1}, \ldots, u_N), \quad 1 \leqslant \ell \leqslant N.$$

Then player ℓ selects $u_\ell(\cdot)\in\mathcal{U}^\ell_{ad}$ to minimize the functional

$$v(\cdot)\mapsto J^\ell(u^c_\ell(\cdot),v(\cdot))\equiv J^\ell(u_1(\cdot),\dots,u_{\ell-1},v(\cdot),u_{\ell+1},\dots,u_N(\cdot)).$$

Obviously, $J^\ell(u(\cdot))$ not only depends on $u_\ell(\cdot)$, but also $u_k(\cdot)$, $k\neq\ell$. Therefore, the optimal control of Player ℓ depends on the controls of the other players.

Definition 6.3. *An N-tuple $\bar u(\cdot)=(\bar u_1(\cdot),\cdots,\bar u_N(\cdot))\in\prod_{\ell=1}^N\mathcal{U}^\ell_{ad}$ is called an open-loop Nash equilibrium of the game if the following holds:*

$$J^\ell(\bar u(\cdot))\leqslant J^\ell(\bar u^c_\ell(\cdot),v(\cdot))\qquad \forall v(\cdot)\in\mathcal{U}^\ell_{ad},\ 1\leqslant\ell\leqslant N.$$

We have the following Pontryagin type maximum principle for Nash equilibria of the N-person dynamic game of FSVIEs. To state the result, let $(\eta^\ell,\zeta^\ell,Y^\ell,Z^\ell)$ be the solution of (3.2) associated with $(h^\ell_x(\bar X(T)),g^\ell_x(\cdot))$. Then we define $H^\ell(\cdot)$ as in (3.3) accordingly. Similarly, with $(h^\ell_{xx},H^\ell_{xx})$, let $(P^\ell_1,P^\ell_2,P^\ell_3,P^\ell_4)$ be the solution of (4.22).

Theorem 6.4. *Let* (H1)–(H2) *hold with g and h being $\mathbb{R}^N$-valued. Let $\bar u=(\bar u_1(\cdot),\dots,\bar u_N(\cdot))\in\prod_{\ell=1}^N\mathcal{U}^\ell_{ad}$ be a Nash equilibrium of the game. Then for any $u\in U^\ell$,*

$$\begin{aligned}&H^\ell(\tau,\bar X(\tau),\bar u^c_\ell(\tau),u,\eta^\ell(\tau),\zeta^\ell(\tau),Y^\ell(\cdot),Z^\ell(\cdot,\tau))\\&\quad-H^\ell(\tau,\bar X(\tau),\bar u(\tau),\eta^\ell(\tau),\zeta^\ell(\tau),Y^\ell(\cdot),Z^\ell(\cdot,\tau))\\&\quad+\frac12\bigg\{\delta\sigma(T,\tau)^\top P^\ell_1(\tau)\delta\sigma(T,\tau)+\mathbb{E}_\tau\int_\tau^T\bigg[\delta\sigma(T,\tau)^\top P^\ell_2(\theta)^\top\delta\sigma(\theta,\tau)\\&\quad+\delta\sigma(\theta,\tau)^\top P^\ell_2(\theta)\delta\sigma(T,\tau)+\delta\sigma(\theta,\tau)^\top P^\ell_3(\theta)\delta\sigma(\theta,\tau)d\theta\\&\quad+\int_\theta^T\big[\delta\sigma(\theta',\tau)^\top P^\ell_4(\theta',\theta)^\top\delta\sigma(\theta,\tau)+\delta\sigma(\theta,\tau)^\top P^\ell_4(\theta',\theta)\delta\sigma(\theta',\tau)\big]d\theta'\bigg]d\theta\bigg\}\geqslant0.\end{aligned}$$

Let us continue to look at a special case: the two-person zero-sum dynamical games. In this case, $N=2$, and $J^1(\bar u(\cdot))+J^2(\bar u(\cdot))=0$. Now, if $(\bar u_1(\cdot),\bar u_2(\cdot))\in\mathcal{U}^1_{ad}\times\mathcal{U}^2_{ad}$ is a Nash equilibrium, then

$$\begin{aligned}&J^1(\bar u_1(\cdot),\bar u_2(\cdot))\leqslant J^1(u_1(\cdot),\bar u_2(\cdot))\qquad\forall u_1(\cdot)\in\mathcal{U}^1_{ad},\\&J^2(\bar u_1(\cdot),\bar u_2(\cdot))\leqslant J^2(\bar u_1(\cdot),u_2(\cdot))\qquad\forall u_2(\cdot)\in\mathcal{U}^2_{ad}.\end{aligned}$$

This also implies that (denoting $\bar J(u(\cdot))=J^1(u(\cdot))$)

$$\bar J(\bar u_1(\cdot),u_2(\cdot))\leqslant\bar J(\bar u_1(\cdot),\bar u_2(\cdot))\leqslant\bar J(u_1(\cdot),\bar u_2(\cdot)).$$

Hence $(\bar u_1(\cdot),\bar u_2(\cdot))$ is referred to as a *saddle point* of the game. In this case, to state the maximum principle, let (η,ζ,Y,Z,P_i) satisfies (3.2), (4.22) associated with (h^1,g^1). We define

$$\begin{aligned}\mathscr{H}(\tau,u_1,u_2)&\overset{\triangle}{=}H(\tau,\bar X(\tau),u_1,u_2,\eta(\tau),\zeta(\tau),Y(\cdot),Z(\cdot,\tau))\\&=\frac12\bigg\{\delta\sigma(T,\tau)^\top P_1(\tau)\delta\sigma(T,\tau)+\mathbb{E}_\tau\int_\tau^T\bigg[\delta\sigma(T,\tau)^\top P_2(\theta)^\top\delta\sigma(\theta,\tau)\\&\quad+\delta\sigma(\theta,\tau)^\top P_2(\theta)\delta\sigma(T,\tau)+\delta\sigma(\theta,\tau)^\top P_3(\theta)\delta\sigma(\theta,\tau)d\theta\\&\quad+\int_\theta^T\big[\delta\sigma(\theta',\tau)^\top P_4(\theta',\theta)^\top\delta\sigma(\theta,\tau)+\delta\sigma(\theta,\tau)^\top P_4(\theta',\theta)\delta\sigma(\theta',\tau)\big]d\theta'\bigg]d\theta\bigg\}.\end{aligned}$$

THEOREM 6.5. *Let* (H1)–(H2) *hold with* g *and* h *being* $\mathbb{R}^2$*-valued. Suppose* $(\bar u_1(\cdot), \bar u_2(\cdot)) \in \mathcal{U}_{ad}^1 \times \mathcal{U}_{ad}^2$ *is the saddle point of the game. Then*

$$\mathscr{H}(\tau, \bar u_1(\tau), u_2) \leqslant \mathscr{H}(\tau, \bar u_1(\tau), \bar u_2) \leqslant \mathscr{H}(\tau, u_1(\tau), \bar u_2) \quad \forall (u_1, u_2) \in U^1 \times U^2.$$

7. Concluding remarks. In this paper, we have developed a spike variation technique for optimal controls of FSVIEs and obtained Pontryagin's type maximum principle. One main contribution is the derivation of the second-order adjoint equation which is different from standard BSVIE. Thus its well-posedness is a part of novelty. The developed methodologies are expected to be extended/adjusted in more general framework, such as the infinite horizon case, the infinite-dimensional case, or the forward-backward SVIEs case. We hope to show more results in the future.

In this article, we have seen that for multiperson dynamic games of FSVIEs, one can obtain the corresponding maximum principle for the Nash equilibria. It is natural to ask what happens when $N \to \infty$? Under certain structure conditions, this will relate to mean-field games of FSVIEs. The relevant results will be reported in our forthcoming work.

REFERENCES

[1] N. AGRAM AND B. ØKSENDAL, *Malliavin calculus and optimal control of stochastic Volterra equations*, J. Optim. Theory Appl., 167 (2015), pp. 1070–1094.
[2] R. BELLMAN AND K. L. COOKE, *Differential-Difference Equations*, Academic Press, New York, 1963.
[3] A. BENSOUSSAN, *Lecture on Stochastc Control*, Lecture Notes in Math. 972, Springer-Verlag, Berlin, 1981.
[4] J.-M. BISMUT, *An introductory approach to duality in optimal stochastic control*, SIAM Rev., 20 (1978), pp. 62–78, https://doi.org/10.1137/1020004.
[5] M. BÔCHER, *An Introduction to the Study of Integral Equations*, Cambridge Tracts in Math. 10, Cambridge University Press, Cambridge, 1909.
[6] V. G. BOLTYANSKII, R. V. GAMKRELDZE, AND L. S. PONTRYAGIN, *On the theory of optimal processes*, Dokl. Akad. Nauk SSSR (N.S.), 110 (1956), pp. 7–10 (in Russian).
[7] H. BRUNNER, *Volterra Integral Equations, An Introduction to Theory and Applications*, Cambridge University Press, Cambridge, 2017.
[8] E. CASAS AND J. YONG, *Maximum principle for state-constrained optimal control problems governed by quasilinear elliptic equations*, Differential Integral Equations, 8 (1995), pp. 1–18.
[9] H. O. FATTORINI, *Infinite Dimensional Optimization and Control Theory*, Cambridge University Press, Cambridge, 1999.
[10] A. FRIEDMAN, *Optimal control for hereditary processes*, Arch. Ration. Mech. Anal., 15 (1964), pp. 396–416.
[11] I. C. GOHBERG AND M. G. KREIN, *Theory and Applications of Volterra Operators in Hilbert Space,* Transl. Math. Monogr. 24, reprinted ed., AMS, Providence, RI, 2004.
[12] B. HU AND J. YONG, *Pontryagin maximum principle for semilinear parabolic equations with pointwise state constraints*, SIAM J. Control Optim., 33 (1995), pp. 1857–1880, https://doi.org/10.1137/S0363012993250074.
[13] M. HU, *Stochastic global maximum principle for optimization with recursive utilities*, Probab. Uncertain. Quant. Risk, 2 (2017), 1.
[14] M. I. KAMIEN AND E. MULLER, *Optimal control with integral state equations*, Rev. Econ. Stud., 43 (1976), pp. 469–473.
[15] H. KUSHNER, *On the stochastic maximum principle: Fixed time of control*, J. Math. Anal. Appl., 11 (1965), pp. 78–92.
[16] H. KUSHNER, *Necessary conditions for continuous parameter stochastic optimization problems*, SIAM J. Control, 10 (1972), pp. 550–565, https://doi.org/10.1137/0310041.
[17] X. LI, J. SUN, AND J. YONG, *Mean-field stochastic linear quadratic optimal control problems: Closed-loop solvability*, Probab. Uncertain. Quant. Risk, 1 (2016), 2.
[18] X. LI AND J. YONG, *Optimal Control Theory for Infinite Dimensional Systems*, Birkhäuuser, Boston, 1995.

[19] J. Lin, *Adapted solution of a backward stochastic nonlinear Volterra integral equation*, Stoch. Anal. Appl., 20 (2002), pp. 165–183.
[20] P. Lin and J. Yong, *Controlled singular Volterra integral equations and Pontryagin maximum principle*, SIAM J. Control Optim., 58 (2020), pp. 136–164, https://doi.org/10.1137/19M124602X.
[21] N. G. Medhin, *Optimal processes governed by integral equations*, J. Math. Anal. Appl., 120 (1986), pp. 1–12.
[22] L. Mou and J. Yong, *A variational formula for stochastic controls and some applications*, Pure Appl. Math. Quart., 3 (2007), pp. 539–567.
[23] É. Pardoux and S. Peng, *Adapted solution of a backward stochastic dfferential equaiton*, Systems Control Lett., 14 (1990), pp. 55–61.
[24] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim., 28 (1990), pp. 966–979, https://doi.org/10.1137/0328054.
[25] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Interscience Publishers, John Wiley & Sons, New York, London, 1962.
[26] Y. Shi, T. Wang, and J. Yong, *Optimal control problems of forward-backward stochastic Volterra integral equations*, Math. Control Relat. Fields, 5 (2015), pp. 613–649.
[27] V. R. Vinokurov, *Optimal control of processes described by integral equations,* I, II, III, SIAM J. Control, 7 (1969), pp. 324–336, 337–345, 346–355.
[28] V. Volterra, *Theory of Functional and Integral and Integro-Differential Equations*, Dover Publications, New York, 1959.
[29] T. Wang, *Necessary conditions in stochastic linear quadratic problems and their applications*, J. Math. Anal. Appl., 469 (2019), pp. 280–297.
[30] T. Wang, *Linear quadratic control problems of stochastic Voltera integral equations*, ESAIM Control Optim. Calc. Var., 24 (2018), pp. 1849–1879.
[31] T. Wang, *Necessary conditions of Pontryagin's type for general controlled stochastic Voltera integral equations*, ESAIM Control Optim. Calc. Var., 26 (2020), 16.
[32] T. Wang and J. Yong, *Comparison theorems for backward stochastic Volterra integral equations*, Stochastic Process Appl., 125 (2015), pp. 1756–1798.
[33] T. Wang and J. Yong, *Spike Variations for Stochastic Volterra Integral Equations*, preprint, https://arxiv.org/abs/2205.13486, 2022.
[34] T. Wang and H. Zhang, *Optimal control problems of forward-backward stochastic Volterra integral equations with closed control regions*, SIAM J. Control Optim., 55 (2017), pp. 2574–2602, https://doi.org/10.1137/16M1059801.
[35] J. Yong, *Backward stochastic Volterra integral equations and some related problems*, Stochastic Process Appl., 116 (2006), pp. 779–795.
[36] J. Yong, *Well-posedness and regularity of backward stochastic Volterra integral equation*, Probab. Theory Relat. Fields, 142 (2008), pp. 21–77.
[37] J. Yong, *Optimality variational principle for controlled forward-backward stochastic differential equations with mixed initial-terminal conditions*, SIAM J. Control Optim., 48 (2010), pp. 4119–4156, https://doi.org/10.1137/090763287.
[38] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, Berlin, 1999.
[39] K. Yosida, *Lectures on Differential and Integral Equations*, Interscience Publishers, Wiley, New York, London, 1960.