ESAIM: COCV 29 (2023) 58
https://doi.org/10.1051/cocv/2023050

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

A GENERAL MAXIMUM PRINCIPLE FOR PROGRESSIVE OPTIMAL CONTROL OF PARTIALLY OBSERVED MEAN-FIELD STOCHASTIC SYSTEM WITH MARKOV CHAIN

Tian Chen and Yunxiao Jia*

Abstract. In this paper, we study an optimal control problem of partially observed mean-field type stochastic control system with Markov chain in progressive structure. The control variable is allowed to enter the diffusion term of the state process and the drift term of the observation process. The control domain need not be convex. In our model, the cost functional and the observation are also of mean-field type. By virtue of a special spike variation, the related stochastic maximum principle has been obtained. The stochastic maximum principle in progressive structure is essentially different from the classical case.

Mathematics Subject Classification. 93E20, 49N30, 93C41.

Received February 11, 2023. Accepted June 23, 2023.

1. Introduction

Over the past decades, Markov regime-switching models have been widely used in economy and engineering. Compared with the traditional model, Markov chain model can better describe the changes between model states. In the economic model, a continuous time finite state Markov chain is often used to describe the change of the economic level, for example, bear market and bull market, see [9, 34, 37] for more details.

Stochastic control problem plays an important role in control theory. The stochastic maximum principle (SMP), which is a necessary condition for the optimal control problem, is one of the important tools to solve the optimal control problem. Promoting the development of SMP is not only of great significance in theory but also of great practical significance in the application of finance and engineering. It was first formulated by Pontryagin in the 1950s and it converted the optimization problems into maximizing the corresponding Hamiltonian functions. Bismut [4] introduced the linear backward stochastic differential equation (BSDE) as the adjoint equation and Peng [20] obtained the general SMP by the second-order adjoint equation. Tang and Li [27] studied the optimal control problem with Poisson random measure in predictable structure and they gave the corresponding SMP. More research on the predictable structure can be seen [21, 35] etc. With the development of forward-backward stochastic differential equation (FBSDE), more scholars have studied the SMP of forward-backward stochastic control systems, see [12, 32]. Donnelly [8] studied the sufficient SMP for the stochastic control system with regime-switching and the necessary SMP was studied by Tao and Wu [28]. In order to solve the flawed estimates (see [23]), Song *et al.* [23] introduced a new method of spike variation

Keywords and phrases: Partial observation, mean-field control, stochastic maximum principle, Markov chain, spike variation.

School of Mathematics, Shandong University, Shandong 250100, Jinan, PR China.

* Corresponding author: jiayunxiao@mail.sdu.edu.cn

and thus derived the rigorous SMP. They revealed the essential difference between the stochastic control system with jump and that without jump. More details of progressive structure can be seen [7, 24].

Mean-field is a useful tool for investigating the collective behavior arising from individuals' mutual interaction, which is very popular in physics. Because of its wide application, mean-field control and game problems have attracted more and more attention, see [13, 17, 29]. There are two typically used problem frames and models for the mean-field control. The first one is the large population system which contains many cooperative agents, and it is also called mean-field game firstly studied independently by Huang *et al.* [14] and Lasry and Lions [15]. Another one is mean-field type control system, where the dynamic is given by some mean-field system, like mean-field stochastic differential equation (MF-SDE). Andersson and Djehiche [2] studied the optimal control problem of a mean-field type stochastic control system with convex domain and that with non-convex domain is studied by Buckdahn *et al.* [5]. Yong [33] studied the linear-quadratic (LQ) optimal control problem in mean-field system and obtained the related feedback form optimal control. Zhang *et al.* [35] obtained a global SMP for a Markov regime-switching mean-field model driven by Brownian motions and Poisson random measure and the SMP of mean-field stochastic control problems with the joint distribution of the controlled state and the control process was gotten by [1]. Sun *et al.* [25] studied the two-person zero-sum mean-filed LQ stochastic differential games.

As mentioned above, they all assume that the overall information is available, *i.e.*, decision makers can obtain all information of the state. However, in practice, decision makers can only get partial information in most cases. There is rich literature about partial information control system see [3, 10, 36]. Li and Tang [18] obtained the SMP for partially observed stochastic control system with non-convex domain by a purely probabilistic approach. Wang and Wu [30] and Wu [31] studied the partially observed optimal control of forward-backward stochastic systems with convex and non-convex control domain, respectively. Brandis *et al.* [19] derived the SMP of a mean-field type stochastic control system with partial information, under which dynamic is governed by a controlled Itô-Lévy process. Since the partial information problem is a very fast-growing research field, it is very hard to exhaustively and roundly account for all the developments of it, and here we refer to [6, 16, 26, 38] for interesting readers to get more details about it.

In this paper, we derive a general SMP for partially observed mean-field type stochastic control system with Markov chain in progressive structure. When it is assumed that the admissible controls are predictable, the flawed estimates may cause some problems (see [23, 24]), thus we solve this problem in progressive structure. We extend the special spike variation which was first introduced by [23] to the case of partial information. The dynamics of state and observation are governed by MF-SDE with Markov chain. The control variables are allowed to enter into all coefficients and the control domain is non-convex. Thus, sharper estimates for the first order variational equations are given (see (4.20) and (4.37)). In order to deal with this problem, three first order adjoint equations and one second order adjoint equation are introduced. Compared with [5] and [35], the main difficulty of this paper is to give an appropriate estimate for the first order variational equations of state and observation as the coefficients are progressive rather than predictable. Actually, the auxiliary process Λ defined in Lemma 4.3 becomes more complex due to the progressive coefficients. Then the general SMP for stochastic control system with Markov chain and partial information is derived. It is worth pointing out that our partially observed SMP can degenerate into many well-known results under some appropriate assumptions.

The main contribution of this paper comparing with existing literatures can be summarized as follows:

(1) A general partially observed mean-field stochastic control model with Markov chain in progressive structure is introduced. Different from the existing literature, the coefficients in this control system are progressive. Thus the terms $\mathrm{d}V_t$ and $\mathrm{d}\tilde{V}_t$ cannot be replaced with each other (see Rem. 3.2), which is different from the predictable structure. In the progressive structure, a more accurate characterization of stochastic control system with jumps is given.

(2) We introduce the progressive structure into the partially observed model. The state and observation are defined on a reference probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ (see (3.1) and (3.2)), and the cost functional is defined on probability space $(\Omega, \mathcal{F}, \mathbb{F}, P^u)$ (see (3.4)). In order to solve this problem, we transfer the cost function of original partially observed problem to the reference probability space by virtue of Girsanov's theorem (see (3.5)). Then we can solve this problem through classical method.

(3) Due to the introduction of the mean field term, we need some more accurate estimates of variational equations of state and observation (see (4.20)). In this proof process, the introduction of auxiliary process Λ is necessary. Since all coefficients are progressive, the auxiliary process and related proof process are more complex and difficult.

(4) In virtue of the dual predictable projection, we introduce three first order adjoint processes and one second adjoint process in the progressive structure to solve this problem. Then the corresponding SMP of partially observed stochastic control system have been obtained. And our result can degenerate into many well-known results under some appropriate assumptions (see Rem. 5.2).

The rest of this paper is organized as follows. In Section 2, we give some preliminaries about the stochastic integral in progressive structure. Section 3 formulates the partially observed optimal control problem. We employ the special spike variation and introduce the first and second order variation equation in Section 4. In order to get the maximum principle in the progressive structure, the related adjoint equations are given in Section 5. In Section 6, we present an LQ optimal control problem to demonstrate the effectiveness of our SMP.

2. Preliminaries

Given a fixed time horizon $T>0$. Let $(\Omega,\mathcal{F},\mathbb{F},P)$ be a filtered probability space. On this probability space, there exists a two-dimensional Brownian motion $(B_\cdot,Y_\cdot)$ and a continuous-time finite state stationary Markov chain $\alpha_\cdot$ taking value in $D=\{1,2,\dots,n\}$ with σ-field $\mathcal{D}$. We assume that the filtration $\mathbb{F}$ generated by $B_\cdot$, $Y_\cdot$ and $\alpha_\cdot$, that is

$$\mathcal{F}_t:=\sigma\{\alpha_s,0\le s\le t\}\vee\sigma\{B_s,Y_s,0\le s\le t\}\vee\mathcal{N},\qquad \forall t\in[0,T],$$

where $\mathcal{N}$ denotes the totality of P-null sets. Then $\mathbb{F}$ satisfies the usual condition.

The generator of $\alpha_\cdot$ is $Q=(q_{ij})_{n\times n}$. Note that $q_{ij}\ge 0$, for $i\ne j$ and $\sum_{j=1}^n q_{ij}=0$, so $q_{ii}\le 0$. In what follows, we further assume that $q_{ij}>0$ for $i\ne j$, so $q_{ii}<0$. Define function $f^i:D\to R$, $f^i_x=I_{\{i\}}(x)$, $i\in D$, which has the following semimartingale decomposition,

$$f^i_{\alpha_t}=f^i_{\alpha_0}+\int_0^t\sum_{j=1}^n q_{\alpha_s j}f^i_j\mathrm{d}s+M^i_t=f^i_{\alpha_0}+\int_0^t q_{\alpha_s i}\mathrm{d}s+M^i_t,$$

where $M^i_\cdot$ is an $\mathbb{F}$ martingale satisfying $E|M^i_t|^2<\infty$. Define $V^{ij}_t=\sum_{0<s\le t}f^i_{\alpha_{s-}}f^j_{\alpha_s}$, $i\ne j$, which counts the number of times that α jumps from i to j up to time t. Since $i\ne j$, we have

$$\begin{aligned}V^{ij}_t&=\sum_{0<s\le t}f^i_{\alpha_{s-}}f^j_{\alpha_s}=\sum_{0<s\le t}f^i_{\alpha_{s-}}\left(f^j_{\alpha_s}-f^j_{\alpha_{s-}}\right)=\sum_{0<s\le t}f^i_{\alpha_{s-}}\left(\Delta f^j_\alpha\right)_s\\&=\int_0^t f^i_{\alpha_{s-}}\mathrm{d}f^j_{\alpha_s}=\int_0^t f^i_{\alpha_s}q_{\alpha_s j}\mathrm{d}s+\int_0^t f^i_{\alpha_{s-}}\mathrm{d}M^j_s.\end{aligned}$$

We give the notation $\sum_{i\ne j}^n$ as abbreviation for $\sum_{j=1}^n\sum_{i=1,i\ne j}^n$. Then we define

$$V_t:=\sum_{i\ne j}^n V^{ij}_t=\int_0^t\sum_{i\ne j}^n\left(q_{\alpha_s j}f^i_{\alpha_s}\mathrm{d}s+f^i_{\alpha_{s-}}\mathrm{d}M^j_s\right)=\int_0^t r_s\mathrm{d}s+\tilde V_t,$$

where $r_s=\sum_{i\ne j}^n q_{\alpha_s j}f^i_{\alpha_s}$, $\tilde V_t=\sum_{i\ne j}^n\int_0^t f^i_{\alpha_{s-}}\mathrm{d}M^j_s$. Now, for a given Hilbert space $\mathbb{H}$, if $\xi:\Omega\to\mathbb{H}$ is an $\mathcal{F}_T$-measurable, square-integrable random variable, we write $\xi\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{H})$; if $x:[0,T]\times\Omega\to\mathbb{H}$ is an $\mathcal{F}_t$-predictable, square-integrable process, we write $x\in M^2_{\mathbb{F}}(0,T;\mathbb{H})$; if $x:[0,T]\times\Omega\to\mathbb{H}$ is an $\mathcal{F}_t$-adapted càdlàg process satisfying $E[\sup_{t\in[0,T]}|x_t|^p]<\infty$, we write $x\in S_{\mathbb{F}}^p(0,T;\mathbb{H})$; if $x:[0,T]\times\Omega\to\mathbb{H}$ is an $\mathcal{F}_t$-predictable process satisfying $E[\int_0^T|x_t|^2\mathrm{d}V_t]<\infty$, we write $x\in F_{\mathbb{F}}^2(0,T;\mathbb{H})$.

In this paper, we consider the partially observed optimal control problem in progressive structure. Given a Euclidean space N, stochastic process $X:[0,T]\times\Omega\to N$ is called progressive (predictable) if X is $\mathcal{G}/\mathcal{B}(N)$ ($\mathcal{P}/\mathcal{B}(N)$) measurable, where $\mathcal{G}(\mathcal{P})$ is the progressive (predictable) σ-field on $[0,T]\times\Omega$. In contrast to other articles about SMP with Markov chain jump, the coefficient of the control system is progressive and the integral is progressive rather than predictable. Let $\zeta(A)=E[\int_0^T I_A\mathrm{d}V_t]$ denote the measure on $\mathcal{B}([0,T])\otimes\mathbb{F}$. Obviously, ζ is not a probability. For any $\mathbb{F}\otimes\mathcal{B}([0,T])/\mathcal{B}(R)$ measurable integrable process X, we set $\mathbb{E}[X]=E[\int_0^T X\mathrm{d}\zeta]$ and denote the Radon-Nikodym derivatives with respect to the predictable σ-field $\mathcal{P}$ by $\mathbb{E}[X|\mathcal{P}]$. It is worth pointing out that $\mathbb{E}$ has similar properties to expectation but it is not an expectation. Next, we give some properties of progressive stochastic integral and dual predictable projection (see [11, 23, 24]). For a progressive process $(K_t)_{t\in[0,T]}$ satisfying $E[\int_0^T|K_t|^2\mathrm{d}V_t]<\infty$, the stochastic integral $\int_0^t K_s\mathrm{d}\tilde{V}_s$ is well-defined and it is a square integrable martingale.

Proposition 2.1. *Suppose $K.$ is a progressive process satisfying $E[\int_0^T|K_t|^2\mathrm{d}V_t]<\infty$, then*

$$\int_0^t K_s\mathrm{d}\tilde{V}_s=\int_0^t K_s\mathrm{d}V_s-\int_0^t\mathbb{E}[K_s|\mathcal{P}]r_s\mathrm{d}s,\qquad\forall t\in[0,T].$$

Remark 2.2. Under the assumption of Proposition 2.1, we have

$$E\bigg[\int_0^t K_s\mathrm{d}V_s\bigg]=E\bigg[\int_0^t\mathbb{E}[K_s|\mathcal{P}]r_s\mathrm{d}s\bigg].\tag{2.1}$$

In particular, if $K.$ is a predictable process, we have the well-known result

$$E\bigg[\int_0^t K_s\mathrm{d}V_s\bigg]=E\bigg[\int_0^t K_sr_s\mathrm{d}s\bigg].$$

Proposition 2.3. *Suppose that $K.$ and $\tilde{K}.$ are progressive processes satisfying $E[\int_0^T(|K_t|^2+|\tilde{K}_t|^2)\mathrm{d}V_t]<\infty$, we have*

$$\bigg[\int_0^\cdot K_s\mathrm{d}\tilde{V}_s,\int_0^\cdot\tilde{K}_s\mathrm{d}\tilde{V}_s\bigg]_t=\int_0^t K_s\tilde{K}_s\mathrm{d}V_s,\qquad\Delta\bigg(\int_0^\cdot K_s\mathrm{d}\tilde{V}_s\bigg)_t=K_t\Delta V_t.$$

3. Statement of the problem

Let $\{T_n\}_{n\geq1}$ be the jump time of V, then $\{T_n\}_{n\geq1}$ is a sequence of strictly increased stopping times. Let U be a nonempty subset of R. Consider the following progressive mean-field stochastic control system with Markov chain

$$\begin{cases}\mathrm{d}X_t=b(t,X_t,E[X_t],u_t)\mathrm{d}t+\sigma(t,X_t,E[X_t],u_t)\mathrm{d}B_t\\ \qquad\qquad+\gamma(t,X_{t-},E[X_{t-}],u_t)\mathrm{d}V_t+c(t,X_{t-},E[X_{t-}],u_t)\mathrm{d}\tilde{V}_t,\\ X_0=x_0,\end{cases}\tag{3.1}$$

where $x_0 \in R$, $b, \sigma, \gamma, c, : \Omega \times [0,T] \times R \times R \times U \to R$. Assume that state process $X_\cdot$ cannot be directly observed, but we can observe a related white noise process $Y_\cdot$, which satisfies the following MF-SDE,

$$Y_t = \int_0^t h(s, X_s, E[X_s], u_s)\mathrm{d}s + \int_0^t \mathrm{d}W_s^u, \tag{3.2}$$

where $h : \Omega \times [0,T] \times R \times R \times U \to R$ and $W_\cdot^u$ denotes a stochastic process depending on the control variable u. We also assume that $\alpha_\cdot$ can be directly observed. Let $\{\mathcal{F}_t^Y\}_{t\in[0,T]}$ be the natural filtration generated by $Y_\cdot$ and $\alpha_\cdot$. Then we define the admissible control set

$$\mathcal{U}_{ad} = \Big\{ u \mid u \text{ is } \mathcal{F}_t^Y\text{-progressive process, taking values in } U, \text{ such that}$$
$$\sup_{t\in[0,T]} E\big[|u_t|^p\big] < \infty \text{ and } E\bigg[\int_0^T |u_t|\mathrm{d}V_t\bigg]^p < \infty \text{ for any } p \geq 2\Big\}.$$

Remark 3.1. For any $u_\cdot \in \mathcal{U}_{ad}$ and $p \geq 2$, we have

$$E\bigg[\int_0^T |u_t|^2\mathrm{d}V_t\bigg]^{\frac{p}{2}} = E\bigg[\sum_{n=1}^{\infty} |u_{T_n}|^2 \mathbb{1}_{\{T_n\leq T\}}\bigg]^{\frac{p}{2}} \leq E\bigg[\sum_{n=1}^{\infty} |u_{T_n}| \mathbb{1}_{\{T_n\leq T\}}\bigg]^p = E\bigg[\int_0^T |u_t|\mathrm{d}V_t\bigg]^p < \infty.$$

Then we give some assumptions to ensure the well-posedness of the controlled system.

Assumption (H1)
i) b, σ, γ, c, h are $\mathcal{G} \otimes \mathcal{B}(R) \otimes \mathcal{B}(R) \otimes \mathcal{B}(U)/\mathcal{B}(R)$ measurable.
ii) b, σ, γ, c, h are twice continuously differentiable with respect to $(x, \tilde{x})$ with bounded first and second order partial derivatives, they and their partial derivatives in $(x, \tilde{x})$ are continuous in $(x, \tilde{x}, u)$. There exists a constant $\bar{C}$ such that $|\mathcal{L}(t, x, \tilde{x}, u)| \leq \bar{C}(1 + |x| + |\tilde{x}| + |u|)$ with $\mathcal{L} = b, \sigma, \gamma, c$ and $|h(t, x, \tilde{x}, u)| \leq \bar{C}$.
iii) $E[\int_0^T |b(t, \mathbf{0})|^2\mathrm{d}t + \int_0^T |\sigma(t, \mathbf{0})|^2\mathrm{d}t + (\int_0^T |\gamma(t, \mathbf{0})|\mathrm{d}V_t)^2 + \int_0^T |c(t, \mathbf{0})|^2\mathrm{d}V_t] < \infty$.
iv) $\gamma_x(t, x, \tilde{x}, u) + c_x(t, x, \tilde{x}, u) + 1 \gg 0$.

In the above, the last condition is necessary to ensure the process $\Lambda_\cdot$ in Lemma 4.3 is well-defined. Under (H1), we know that (3.1) admits a unique solution for any admissible control. Define $\mathrm{d}P^u = Z_T^u \mathrm{d}P$, where

$$Z_t^u = \exp\bigg\{\int_0^t h(s, X_s, E[X_s], u_s)\mathrm{d}Y_s - \frac{1}{2}\int_0^t |h(s, X_s, E[X_s], u_s)|^2\mathrm{d}s\bigg\}.$$

Obviously, $Z_\cdot^u$ can be characterized as follows

$$\mathrm{d}Z_t^u = Z_t^u h(t, X_t, E[X_t], u_t)\mathrm{d}Y_t, \qquad Z_0^u = 1. \tag{3.3}$$

Let (H1) hold, we know that P^u is a new probability measure by Girsanov's theorem. Under probability measure P^u, $(B_\cdot, W_\cdot^u)$ is a two-dimensional standard Brownian motion and $\tilde{V}_\cdot$ is still a compensated martingale. We introduce the following cost functional

$$J(u_\cdot) = E^u\bigg[\int_0^T l(t, X_t, E^u[X_t], u_t)\mathrm{d}t + \int_0^T f(t, X_{t-}, E^u[X_{t-}], u_t)\mathrm{d}V_t + g(X_T, E^u[X_T])\bigg], \tag{3.4}$$

where $l, f : \Omega \times [0,T] \times R \times R \times U \to R$, $g : \Omega \times R \times R \to R$. E^u denotes mathematical expectation under probability P^u. Then the partially observed optimal control problem studied in our paper can be described as follows:

Problem (OCP). Find an admissible control $\hat{u}_\cdot \in \mathcal{U}_{ad}$ such that

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

Then we introduce some assumptions:

Assumption (H2).
i) l, f are $\mathcal{G} \otimes \mathcal{B}(R) \otimes \mathcal{B}(R) \otimes \mathcal{B}(U)/\mathcal{B}(R)$ measurable, g is $\mathcal{F}_T \otimes \mathcal{B}(R) \otimes \mathcal{B}(R)/\mathcal{B}(R)$ measurable.
ii) l, f, g are twice continuous differentiable with respect to $(x, \tilde{x})$ with bounded second order partial derivatives, they and their partial derivatives in $(x, \tilde{x})$ are continuous in $(x, \tilde{x}, u)$. There exist a constant $\bar{C}$ such that $|\mathcal{L}(t, x, \tilde{x}, u)| \le \bar{C}(1 + |x|^2 + |\tilde{x}|^2 + |u|^2)$, $|\mathcal{L}_x(t, x, \tilde{x}, u)| \le \bar{C}(1 + |x| + |\tilde{x}| + |u|)$, $|\mathcal{L}_{\tilde{x}}(t, x, \tilde{x}, u)| \le \bar{C}(1 + |x| + |\tilde{x}| + |u|)$ with $\mathcal{L} = l, f$ and $|g(x, \tilde{x})| \le \bar{C}(1 + |x|^2 + |\tilde{x}|^2)$, $|g_x(x, \tilde{x})| \le \bar{C}(1 + |x| + |\tilde{x}|)$, $|g_{\tilde{x}}(x, \tilde{x})| \le \bar{C}(1 + |x| + |\tilde{x}|)$.

Obviously, the cost function (3.4) can be written as

$$J(u_\cdot) = E\left[\int_0^T Z_t^u l(t, X_t, E[Z_t^u X_t], u_t)\mathrm{d}t + Z_T^u g(X_T, E[Z_T^u X_T]) + \int_0^T Z_{t-}^u f(t, X_{t-}, E[Z_{t-}^u X_{t-}], u_t)\mathrm{d}V_t\right] \tag{3.5}$$

Remark 3.2. Let (H1)-(H2) hold in predictable structure. If γ is $\mathcal{P} \otimes \mathcal{B}(R) \otimes \mathcal{B}(R) \otimes \mathcal{B}(U)/\mathcal{B}(R)$ measurable, we have the following result for any $X_\cdot \in S_{\mathbb{F}}^2(0, T; R)$ and any $u_\cdot \in \mathcal{U}_{ad}$,

$$|\gamma(t, X_{t-}, E[X_{t-}], u_t)| \le \bar{C}(|\gamma(t, 0, 0, 0)| + |X_{t-}| + |E[X_{t-}]| + |u_t|).$$

By Proposition 2.1, we have

$$\int_0^t \gamma(s, X_{s-}, E[X_{s-}], u_s)\mathrm{d}V_t = \int_0^t \gamma(s, X_{s-}, E[X_{s-}], u_s)\mathrm{d}\tilde{V}_s + \int_0^t \gamma(s, X_s, E[X_s], u_s) r_s \mathrm{d}s.$$

Then the state system (3.1) can be rewritten as

$$\begin{cases} \mathrm{d}X_t = [b(t, X_t, E[X_t], u_t) + \gamma(s, X_s, E[X_s], u_s)r_s]\mathrm{d}t + \sigma(t, X_t, E[X_t], u_t)\mathrm{d}B_t \\ \qquad\qquad + [\gamma(t, X_{t-}, E[X_{t-}], u_t) + c(t, X_{t-}, E[X_{t-}], u_t)]\mathrm{d}\tilde{V}_t, \\ X_0 = x_0. \end{cases}$$

This shows that γ can be omitted in the state system. If f is $\mathcal{P} \otimes \mathcal{B}(R) \otimes \mathcal{B}(R) \otimes \mathcal{B}(U)/\mathcal{B}(R)$ measurable, we have

$$J(u_\cdot) = E^u\left[\int_0^T \{l(t, X_t, E^u[X_t], u_t) + f(t, X_t, E^u[X_t], u_t)r_t\}\mathrm{d}t + g(X_T, E^u[X_T])\right].$$

We also know that f can be omitted in the cost functional.

4. Variation

In this section, we focus on the SMP for the cost functional (3.5) subject to (3.1) and (3.2). Suppose $\hat{u}_\cdot \in \mathcal{U}_{ad}$ is the optimal control, $\hat{X}_\cdot$ and $\hat{Z}_\cdot$ are the corresponding trajectories of (3.1) and (3.2). For simplification, we use the abbreviation $\hat{Z} = Z^{\hat{u}}$, $\hat{E} = E^{\hat{u}}$ and $\hat{W} = W^{\hat{u}}$. Since U is not necessary convex, we employ spike variations, for any $\bar{t} \in [0, T]$,

$$u^\epsilon = \begin{cases} v, & \text{if } (t, \omega) \in \mathcal{O} :=]\!]\bar{t}, \bar{t} + \epsilon]\!] \setminus \bigcup_{n=1}^{\infty} [\![T_n]\!], \\ \hat{u}, & \text{otherwise,} \end{cases} \tag{4.1}$$

where $[\![T_n]\!] := \{(\omega, t) \in \Omega \times [0, T] | T_n(\omega) = t\}$ is the graph of T_n, and v is a bounded $\mathcal{F}_{\bar{t}}^Y$ measurable function that takes values in U. Since T_n is a stopping time, $[\![T_n]\!]$ is a progressive set. Therefore, the spike variation $u^\epsilon_\cdot$ is progressive and we can show that $u^\epsilon_\cdot \in \mathcal{U}_{ad}$. And we can know that $\bigcup_{n=1}^{\infty} [\![T_n]\!]$ is negligible under $P \times Leb$, thus it is also negligible under $P^u \times Leb$, where Leb denotes the Lebesgue measure. For more details of this variation we can see Song *et al.* [23].

Remark 4.1. (i) As we all know, Markov chain $\alpha_\cdot$ is a Feller process, then $\alpha_\cdot$ is a quasi-left continuous process. So is V. T_n is not a predictable time, so $[\![T_n]\!]$ is not predictable, which means that $u^\epsilon_\cdot$ is not predictable; that's the reason why we allow the integrand of the stochastic integral to be progressive. In fact, T_n are totally unpredictable times.
(ii) In order to deal with the stochastic optimal control problem with Markov chain and partial information in progressive structure, we introduce a special spike variation (see (4.1)). On the one hand, this form of variation requires us to dig out the graph of all jump time of Markov chain, thus the jump time of Markov chain should be observed. On the other hand, only if the admissible control set is $\mathcal{F}_t^Y$-adapted, which is generated by observation $Y_\cdot$ and Markov chain $\alpha_\cdot$, the perturbed control $u^\epsilon_\cdot$ can belong to this admissible set again. Thus we assume that the Markov chain $\alpha_\cdot$ can be directly observed. Otherwise, the perturbed control no longer belongs to the admissible control set $\mathcal{U}_{ad}$.

For simplification, we introduce the following notations,

$$\begin{cases} \varphi(t) := \varphi(t, \hat{X}_t, E[\hat{X}_t], \hat{u}_t), \eta(t) := \eta(t, \hat{X}_{t-}, E[\hat{X}_{t-}], \hat{u}_t), \\ \varphi_x(t) := \varphi_x(t, \hat{X}_t, E[\hat{X}_t], \hat{u}_t), \varphi_{xx}(t) := \varphi_{xx}(t, \hat{X}_t, E[\hat{X}_t], \hat{u}_t), \\ \eta_x(t) := \eta_x(t, \hat{X}_{t-}, E[\hat{X}_{t-}], \hat{u}_t), \eta_{xx}(t) := \eta_{xx}(t, \hat{X}_{t-}, E[\hat{X}_{t-}], \hat{u}_t), \\ \varphi_{\tilde{x}}(t) := \varphi_{\tilde{x}}(t, \hat{X}_t, E[\hat{X}_t], \hat{u}_t), \eta_{\tilde{x}}(t) := \eta_{\tilde{x}}(t, \hat{X}_{t-}, E[\hat{X}_{t-}], \hat{u}_t), \\ \delta\varphi(t, u) := \varphi(t, \hat{X}_t, E[\hat{X}_t], u) - \varphi(t, \hat{X}_t, E[\hat{X}_t], \hat{u}_t), \\ \delta\eta(t, u) := \eta(t, \hat{X}_{t-}, E[\hat{X}_{t-}], u) - \eta(t, \hat{X}_{t-}, E[\hat{X}_{t-}], \hat{u}_t), \end{cases}$$

where $\varphi = b, \sigma, h, l, g$; $\eta = \gamma, c, f$, and the mathematical expectation E in l, f, g should be replaced by E^u.
We denote by $X^\epsilon_\cdot$ and $Z^\epsilon_\cdot$ the trajectories of $u^\epsilon_\cdot$. By the estimate of MF-SDE and noticing that $(Leb \times P)([\![T_n]\!]) = 0$, we can obtain

$$\begin{aligned} & E\left[\sup_{t \in [0,T]} |X_t^\epsilon - \hat{X}_t|\right] \\ & \leq CE\left[\left(\int_0^T |\delta b(t, u_t^\epsilon)| \mathrm{d}t\right)^p + \left(\int_0^T |\delta\sigma(t, u_t^\epsilon)|^2 \mathrm{d}t\right)^{\frac{p}{2}} + \left(\int_0^T |\delta\gamma(t, u_t^\epsilon)| \mathrm{d}V_t\right)^p + \left(\int_0^T |\delta c(t, u_t^\epsilon)|^2 \mathrm{d}V_t\right)^{\frac{p}{2}}\right] \\ & \leq CE\left[\left(\int_{\bar{t}}^{\bar{t}+\epsilon} |v - \hat{u}_t| \mathrm{d}t\right)^p + \left(\int_{\bar{t}}^{\bar{t}+\epsilon} |v - \hat{u}_t|^2 \mathrm{d}t\right)^{\frac{p}{2}}\right] \\ & = O(\epsilon^p) + O(\epsilon^{\frac{p}{2}}). \end{aligned}$$

Indeed, there is no jump on $\mathcal{O}$. Thus the jump terms γ, c do not influence the order of variation. Thanks to this, we can use the method in Peng [20] to get the desired conclusion. Then we introduce the first-order variation equation

$$\begin{cases} \mathrm{d}X_{1,t} = \{b_x(t)X_{1,t} + b_{\tilde{x}}(t)E[X_{1,t}] + \delta b(t,u_t^\epsilon)\}\mathrm{d}t + \{\sigma_x(t)X_{1,t} \\ \qquad\qquad + \sigma_{\tilde{x}}(t)E[X_{1,t}] + \delta\sigma(t,u_t^\epsilon)\}\mathrm{d}B_t + \{\gamma_x(t)X_{1,t-} \\ \qquad\qquad + \gamma_{\tilde{x}}(t)E[X_{1,t-}]\}\mathrm{d}V_t + \{c_x(t)X_{1,t-} + c_{\tilde{x}}(t)E[X_{1,t-}]\}\mathrm{d}\tilde{V}_t, \\ \mathrm{d}Z_{1,t} = \{Z_{1,t}h(t) + \hat{Z}_t h_x(t)X_{1,t} + \hat{Z}_t h_{\tilde{x}}(t)E[X_{1,t}] + \hat{Z}_t\delta h(t,u_t^\epsilon)\}\mathrm{d}Y_t, \\ X_{1,0} = 0, \qquad Z_{1,0} = 0, \qquad t \in [0,T], \end{cases} \tag{4.2}$$

and the second-order variation equation

$$\begin{cases} \mathrm{d}X_{2,t} = \{b_x(t)X_{2,t} + b_{\tilde{x}}(t)E[X_{2,t}] + \frac{1}{2}b_{xx}(t)X_{1,t}^2 + \delta b_x(t,u_t^\epsilon)X_{1,t}\}\mathrm{d}t \\ \qquad\qquad + \{\sigma_x(t)X_{2,t} + \sigma_{\tilde{x}}(t)E[X_{2,t}] + \frac{1}{2}\sigma_{xx}(t)X_{1,t}^2 + \delta\sigma_x(t,u_t^\epsilon)X_{1,t}\}\mathrm{d}B_t \\ \qquad\qquad + \{\gamma_x(t)X_{2,t-} + \gamma_{\tilde{x}}(t)E[X_{2,t-}] + \frac{1}{2}\gamma_{xx}(t)X_{1,t-}^2\}\mathrm{d}V_t \\ \qquad\qquad + \{c_x(t)X_{2,t-} + c_{\tilde{x}}(t)E[X_{2,t-}] + \frac{1}{2}c_{xx}(t)X_{1,t-}^2\}\mathrm{d}\tilde{V}_t \\ \mathrm{d}Z_{2,t} = \{Z_{2,t}h(t) + Z_{1,t}h_x(t)X_{1,t} + \hat{Z}_t h_x(t)X_{2,t} + \hat{Z}_t h_{\tilde{x}}(t)E[X_{2,t}] \\ \qquad\qquad + \frac{1}{2}\hat{Z}_t h_{xx}(t)X_{1,t}^2 + Z_{1,t}\delta h(t,u_t^\epsilon) + \hat{Z}_t\delta h_x(t,u_t^\epsilon)X_{1,t}\}\mathrm{d}Y_t \\ X_{2,0} = 0, \qquad Z_{2,0} = 0, \qquad t \in [0,T]. \end{cases} \tag{4.3}$$

Under (H1)-(H2), it is easy to show that (4.2) and (4.3) admit a unique solution, respectively. We have the following basic estimate.

Lemma 4.2. *Let (H1)-(H2) hold. For any $u_\cdot \in \mathcal{U}_{ad}$ and $k \geq 1$, then we have*

$$\sup_{t\in[0,T]} E|Z_t^u|^{2k} < \infty, \qquad \text{and} \qquad \sup_{t\in[0,T]} E|X_t|^{2k} \leq C\Big(1 + \sup_{t\in[0,T]} E|u_t|^{4k}\Big).$$

Lemma 4.3. *Suppose (H1)-(H2) hold. Let $(\Phi_t)_{t\in[0,T]}$ be a progressively measurable process, for any $p \geq 1$, there exists a positive constant C_p such that*

$$E\Big[\sup_{t\in[0,T]} |\Phi_t|^p\Big] \leq C_p. \tag{4.4}$$

Then there exists a function $\tilde{\rho} : (0,\infty) \to (0,\infty)$ with $\tilde{\rho}(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$, such that

$$\int_0^T |E[\Phi_t X_{1,t}]|^2 \,\mathrm{d}t \leq \epsilon\tilde{\rho}(\epsilon), \qquad \epsilon > 0.$$

Proof. We recall the dynamic (4.2) of $X_{1,\cdot}$ and we set

$$\Lambda_t := \exp\bigg\{ \int_0^t \Big\{\frac{1}{2}\sigma_x^2(s) - b_x(s)\Big\}\mathrm{d}s - \int_0^t \sigma_x(s)\mathrm{d}B_s - \int_0^t c_x(s)\mathrm{d}\tilde{V}_s + \int_0^t \Big\{ \ln\frac{1}{1+\gamma_x(s)+c_x(s)} + c_x(s)\Big\}\mathrm{d}V_s\bigg\}.$$

Then, by applying Itô's formula to $\Lambda_t X_{1,t}$, we get

$$\begin{aligned} X_{1,t} = \Gamma_t \int_0^t & \Lambda_s \big\{ \big(b_{\tilde{x}}(s) - \sigma_x(s)\sigma_{\tilde{x}}(s)\big) E[X_{1,s}] + \big(\delta b(s, u_s^\epsilon) - \sigma_x(s)\delta\sigma(s, u_s^\epsilon)\big) \big\} \mathrm{d}s \\ & + \Gamma_t \int_0^t \Lambda_s \big\{ \sigma_{\tilde{x}}(s) E[X_{1,s}] + \delta\sigma(s, u_s^\epsilon) \big\} \mathrm{d}B_s + \Gamma_t \int_0^t \Lambda_{s-} c_{\tilde{x}}(s) E[X_{1,s-}] \mathrm{d}\tilde{V}_s \\ & + \Gamma_t \int_0^t \Lambda_{s-} \big\{ \gamma_{\tilde{x}}(s) + \big(\gamma_{\tilde{x}}(s) + c_{\tilde{x}}(s)\big)\big(\beta(s) - c_x(s)\big) \big\} E[X_{1,s-}] \mathrm{d}V_s, \end{aligned}$$

where $\Gamma_t = \Lambda_t^{-1}$ and $\beta(t) = \frac{c_x^2(t) + c_x(t)\gamma_x(t) - \gamma_x(t)}{1 + \gamma_x(t) + c_x(t)}$. Under (H1)-(H2), for any $p \geq 1$ there exists a positive constant C_p, such that

$$E\Big[\sup_{t \in [0,T]} (|\Lambda_t|^p + |\Gamma_t|^p) \Big] \leq C_p. \tag{4.5}$$

Moreover, with the help of (4.4) it also follows that

$$E\Big[\sup_{t \in [0,T]} |\Phi_t \Gamma_t|^p \Big] \leq C_p. \tag{4.6}$$

By the Itô Representation Theorem, for every $t \in [0, T]$, there exists unique $(\alpha_t(\cdot), \kappa_t(\cdot)) \in M_{\mathbb{F}}^2(0, T; R) \times F_{\mathbb{F}}^2(0, T; R)$ such that

$$\Phi_t \Gamma_t = E[\Phi_t \Gamma_t] + \int_0^t \alpha_t(s) \mathrm{d}B_s + \int_0^t \kappa_t(s) \mathrm{d}\tilde{V}_s, \qquad P\text{-a.s.}. \tag{4.7}$$

In virtue of (4.6), Burkholder-Davis-Gundy (B-D-G) inequality and Doob's inequality, we can know that there exists a constant C_p independent of t such that for any $p > 1$,

$$\begin{aligned} E\Big[\int_0^t [|\alpha_t(s)|^2 + |\kappa_t(s)|^2 r_s] \mathrm{d}s \Big]^{\frac{p}{2}} & \leq C_p E\Big[\sup_{s \in [0,t]} \Big| \int_0^s \alpha_t(r) \mathrm{d}B_r + \int_0^s \kappa_t(r) \mathrm{d}\tilde{V}_r \Big|^p \Big] \\ & \leq C_p \left(\frac{p}{p-1} \right)^p E\Big| \int_0^t \alpha_t(s) \mathrm{d}B_s + \int_0^t \kappa_t(s) \mathrm{d}\tilde{V}_s \Big|^p \leq C_p E \big| \Phi_t \Gamma_t - E\left[\Phi_t \Gamma_t\right] \big|^p \\ & \leq C_p \Big(E\big[|\Phi_t \Gamma_t|^p \big] + \big| E[\Phi_t \Gamma_t] \big|^p \Big) \leq C_p E\Big[\sup_{t \in [0,T]} |\Phi_t \Gamma_t|^p \Big] \leq C_p. \end{aligned}$$

For $p = 1$, by using Hölder's inequality, we can easily get

$$E\Big[\Big(\int_0^t |\alpha_t(s)|^2 \mathrm{d}s + \int_0^t |\kappa_t(s)|^2 r_s \mathrm{d}s \Big)^{\frac{1}{2}} \Big] \leq C.$$

Then we have, for any $p \geq 1$

$$\sup_{t \in [0,T]} E\Big[\Big(\int_0^t |\alpha_t(s)|^2 \mathrm{d}s + \int_0^t |\kappa_t(s)|^2 r_s \mathrm{d}s \Big)^{\frac{p}{2}} \Big] \leq C_p. \tag{4.8}$$

We consider

$$\Phi_t X_{1,t} := J_1(t) + J_2(t) + J_3(t) + J_4(t), \qquad t \in [0,T],$$

where

$$\begin{aligned}
J_1(t) &= \Phi_t\Gamma_t \int_0^t \Lambda_s\big\{[b_{\tilde{x}}(s) - \sigma_x(s)\sigma_{\tilde{x}}(s)]E[X_{1,s}] + [\delta b(s,u_s^\epsilon) - \sigma_x(s)\delta\sigma(s,u_s^\epsilon)]\big\}\mathrm{d}s,\\
J_2(t) &= \Phi_t\Gamma_t \int_0^t \Lambda_s\big[\sigma_{\tilde{x}}(s)E[X_{1,s}] + \delta\sigma(s,u_s^\epsilon)\big]\mathrm{d}B_s,\\
J_3(t) &= \Phi_t\Gamma_t \int_0^t \Lambda_{s-}\theta(s)E\left[X_{1,s-}\right]\mathrm{d}V_s,\;\; J_4(t) = \Phi_t\Gamma_t \int_0^t \Lambda_{s-}c_{\tilde{x}}(s)E\left[X_{1,s-}\right]\mathrm{d}\tilde{V}_s,
\end{aligned}$$

with $\theta(t) = \gamma_{\tilde{x}}(t) + (\gamma_{\tilde{x}}(t) + c_{\tilde{x}}(t))(\beta(t) - c_x(t))$. The following estimate for $E[J_1(t)]$ holds:

$$\int_0^T \big|E[J_1(t)]\big|^2\mathrm{d}t \le C\bigg(\epsilon^2 + \int_0^T\int_0^t \big|E[X_{1,s}]\big|^2\mathrm{d}s\mathrm{d}t\bigg), \qquad t \in [0,T]. \tag{4.9}$$

Indeed, by virtue of (4.5), (4.6) and Lemma 4.2, we have

$$\begin{aligned}
\big|E[J_1(t)]\big| &\le CE\bigg[\sup_{t\in[0,T]}\big|\Phi_t\Gamma_t\big|\cdot\sup_{t\in[0,T]}\big|\Lambda_t\big|\cdot\bigg(\int_0^t\big|E[X_{1,s}]\big|\mathrm{d}s + \epsilon\bigg)\bigg]\\
&\le C\bigg(E\bigg[\sup_{t\in[0,T]}\big|\Phi_t\Gamma_t\big|^2\bigg]\bigg)^{\frac{1}{2}}\bigg(E\bigg[\sup_{t\in[0,T]}\big|\Lambda_t\big|^2\bigg]\bigg)^{\frac{1}{2}}\bigg[\int_0^t\big|E[X_{1,s}]\big|\mathrm{d}s + \epsilon\bigg]\\
&\le C\epsilon + C\int_0^t\big|E[X_{1,s}]\big|\mathrm{d}s, \qquad t \in [0,T],
\end{aligned}$$

which yields (4.9). To estimate $E[J_2(t)]$, we notice that

$$\begin{aligned}
E[J_2(t)] &= E\bigg[\Phi_t\Gamma_t\int_0^t \Lambda_s\Big(\sigma_{\tilde{x}}(s)E[X_{1,s}] + \delta\sigma(s,u_s^\epsilon)\Big)\mathrm{d}B_s\bigg]\\
&= E\bigg[\int_0^t \alpha_t(s)\Lambda_s\sigma_{\tilde{x}}(s)E[X_{1,s}]\mathrm{d}s\bigg] + E\bigg[\int_0^t \alpha_t(s)\Lambda_s\delta\sigma(s,u_s^\epsilon)\mathrm{d}s\bigg]\\
&:= J_{21}(t) + J_{22}(t), \qquad t \in [0,T].
\end{aligned}$$

In view of (4.5) and (4.8), we have

$$|J_{21}(t)|^2 \le E\bigg[\sup_{t\in[0,T]}|\Lambda_t|^2\int_0^t|\alpha_t(s)|^2\mathrm{d}s\bigg]\cdot\int_0^t|E[X_{1,s}]|^2\mathrm{d}s \le C\int_0^t|E[X_{1,s}]|^2\mathrm{d}s. \tag{4.10}$$

Moreover,

$$\begin{aligned}\big|J_{22}(t)\big| &\leq CE\Big[\sup_{t\in[0,T]}\Big(1+|\hat{X}_t|+|v|+|\hat{u}_t|\Big)\cdot\int_0^t \big|\alpha_t(s)\Lambda_s\big|I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s\Big]\\ &\leq C\epsilon^{\frac{1}{2}}\bigg(E\Big[\int_0^t \big|\alpha_t(s)\Lambda_s\big|^2 I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s\Big]\bigg)^{\frac{1}{2}}\\ &\leq C\epsilon^{\frac{1}{2}}\bigg(E\Big[\sup_{t\in[0,T]}\big|\Lambda_t\big|^4\Big]\bigg)^{\frac{1}{4}}\bigg(E\Big[\Big(\int_0^t \big|\alpha_t(s)\big|^2 I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s\Big)^2\Big]\bigg)^{\frac{1}{4}}\\ &\leq C\epsilon^{\frac{1}{2}}\bigg(E\Big[\Big(\int_0^t \big|\alpha_t(s)\big|^2 I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s\Big)^2\Big]\bigg)^{\frac{1}{4}}.\end{aligned}$$

Therefore,

$$\int_0^T \big|J_{22}(t)\big|^2\mathrm{d}t \leq C\epsilon\bigg(\int_0^T E\Big[\Big(\int_0^t \big|\alpha_t(s)\big|^2 I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s\Big)^2\Big]\mathrm{d}t\bigg)^{\frac{1}{2}}.$$

Setting $\rho_1(\epsilon) := C(\int_0^T E[(\int_0^t |\alpha_t(s)|^2 I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s)^2]\mathrm{d}t)^{\frac{1}{2}}$, from (4.8), then

$$\int_0^T E\Big[\int_0^t \big|\alpha_t(s)\big|^2 I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}s\Big]^2\mathrm{d}t \leq C_T\sup_{t\in[0,T]}E\Big[\int_0^t \big|\alpha_t(s)\big|^2\mathrm{d}s\Big]^2 < \infty.$$

Thus by the dominated convergence theorem and noticing $I_{(\bar{t},\bar{t}+\epsilon]} \to 0$ in product measure $\mathrm{d}t\mathrm{d}P$, we can easily obtain that $\rho_1(\epsilon)\to 0$ as $\epsilon \downarrow 0$. Hence,

$$\int_0^T \big|J_{22}(t)\big|^2\mathrm{d}t \leq \epsilon\rho_1(\epsilon). \tag{4.11}$$

Combining (4.10) and (4.11), we have the following estimate

$$\int_0^T \big|E[J_2(t)]\big|^2\mathrm{d}t \leq C\bigg(\epsilon\rho_1(\epsilon)+\int_0^T\int_0^t \big|E[X_{1,s}]\big|^2\mathrm{d}s\mathrm{d}t\bigg), \qquad t\in[0,T]. \tag{4.12}$$

Observing the expression (4.7), we get

$$\begin{aligned}E[J_3(t)] &= E[\Phi_t\Gamma_t]E\Big[\int_0^t \Lambda_{s-}\theta(s)E\left[X_{1,s-}\right]\mathrm{d}V_s\Big]+E\Big[\int_0^t \Lambda_{s-}\theta(s)\kappa_t(s)E\left[X_{1,s-}\right]\mathrm{d}V_t\Big]\\ &:= J_{31}(t)+J_{32}(t).\end{aligned}$$

With the help of (4.5) and (4.8), we obtain

$$\begin{aligned}\big|J_{31}(t)\big|^2 &\leq CE\Big[\sup_{t\in[0,T]}\big|\Phi_t\Gamma_t\big|^2\Big]\cdot\bigg|E\left[\int_0^t \Lambda_{s-}E[X_{1,s-}]\mathrm{d}V_t\right]\bigg|^2\\ &\leq CE\Big[\sup_{t\in[0,T]}\big|\Lambda_t\big|^2\Big]\cdot\int_0^t \big|E[X_{1,s}]\big|^2\mathrm{d}s \leq C\int_0^t \big|E[X_{1,s}]\big|^2\mathrm{d}s.\end{aligned}$$

Similarly, we can obtain

$$\big|J_{32}(t)\big|^2 \leq C\int_0^t \big|E[X_{1,s}]\big|^2 \mathrm{d}s.$$

Thus we have

$$\int_0^T \big|E[J_3(t)]\big|^2 \mathrm{d}t \leq C\int_0^T\int_0^t \big|E[X_{1,s}]\big|^2 \mathrm{d}s\mathrm{d}t, \qquad t\in[0,T]. \tag{4.13}$$

Then, with the expression (4.7) we can get

$$E[J_4(t)] = E\bigg[\int_0^t \Lambda_{s-} c_{\tilde{x}}(s)\kappa_t(s)E[X_{1,s-}]\mathrm{d}V_s\bigg].$$

We also obtain

$$\begin{aligned}|E[J_4(t)]|^2 &= \bigg|E\bigg[\int_0^t \Lambda_s \mathbb{E}[c_{\tilde{x}}(s)|\mathcal{P}]\kappa_t(s)E[X_{1,s}]r_s\mathrm{d}s\bigg]\bigg|^2\\ &\leq CE\bigg[\sup_{t\in[0,T]}|\Lambda_t|^2\int_0^t|\kappa_t(s)|^2\mathrm{d}s\bigg]\int_0^t|E[X_{1,s}]|^2\mathrm{d}s\\ &\leq C\int_0^t|E[X_{1,s}]|^2\mathrm{d}s.\end{aligned}$$

Then

$$\int_0^T \big|E[J_4(t)]\big|^2 \mathrm{d}t \leq C\int_0^T\int_0^t \big|E[X_{1,s}]\big|^2 \mathrm{d}s\mathrm{d}t, \qquad t\in[0,T]. \tag{4.14}$$

In view of (4.9) and (4.12)–(4.14), we have

$$\int_0^T \big|E[\Phi_t X_{1,t}]\big|^2 \mathrm{d}t \leq C\bigg(\epsilon^2+\epsilon\rho_1(\epsilon)+\int_0^T\int_0^t \big|E[X_{1,s}]\big|^2 \mathrm{d}s\mathrm{d}t\bigg). \tag{4.15}$$

where $C = C(\Phi)$ depends on the semimartingale Φ.

Now, taking $\Phi = 1, C_1 := C(1)$ in (4.15) we get

$$\int_0^T \big|E[X_{1,t}]\big|^2 \mathrm{d}t \leq C\bigg(\epsilon^2+\epsilon\rho_1(\epsilon)+\int_0^T\int_0^t \big|E[X_{1,s}]\big|^2 \mathrm{d}s\mathrm{d}t\bigg).$$

Then by Gronwall's inequality, we have

$$\int_0^T \big|E[X_{1,t}]\big|^2 \mathrm{d}t \leq C\left(\epsilon^2+\epsilon\rho_1(\epsilon)\right). \tag{4.16}$$

Substituting (4.16) into (4.15), we have

$$\int_0^T \big|E[\Phi_t X_{1,t}]\big|^2 \mathrm{d}t \leq C\left(\epsilon^2+\epsilon\rho_1(\epsilon)\right). \tag{4.17}$$

Finally, setting $\tilde{\rho}(\epsilon) := C\left(\epsilon+\rho_1(\epsilon)\right)$ in (4.17), we obtain the result. □

Based on the above two lemmas, we have the following estimates.

Lemma 4.4. *Let (H1)-(H2) hold. Then, for any $k \geq 1$,*

$$E\left[\sup_{t\in[0,T]}\left|X_{1,t}\right|^{2k}\right] \leq C_k\epsilon^k, \qquad E\left[\sup_{t\in[0,T]}\left|Z_{1,t}\right|^{2k}\right] \leq C_k\epsilon^k, \tag{4.18}$$

$$E\left[\sup_{t\in[0,T]}\left|X_{2,t}\right|^{2k}\right] \leq C_k\epsilon^{2k}, \qquad E\left[\sup_{t\in[0,T]}\left|Z_{2,t}\right|^{2k}\right] \leq C_k\epsilon^{2k}, \tag{4.19}$$

$$\sup_{t\in[0,T]}\left|E[X_{1,t}]\right|^2 \leq \epsilon\rho(\epsilon), \qquad \sup_{t\in[0,T]}\left|E[Z_{1,t}X_{1,t}]\right| \leq \epsilon\rho(\epsilon), \quad \epsilon > 0, \tag{4.20}$$

where $\rho : (0,\infty) \to (0,\infty)$ such that $\rho(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$, ρ may vary line by line.

Proof. The proof of the first estimate of (4.18) and (4.19) is a slight extension of Lemma 4.3 of [24], so we omit it.

From (4.2), noticing the boundedness of $\gamma_x, \gamma_{\tilde{x}}$, we have

$$E[X_{1,t}] = \int_0^t \Big(E[b_x(s)X_{1,s}] + E[b_{\tilde{x}}(s)]E[X_{1,s}] + E[\delta b(s,u_s^\epsilon)]\Big)\mathrm{d}s + E\left[\int_0^t \gamma_x(s)X_{1,s-} + \gamma_{\tilde{x}}(s)E[X_{1,s-}]\mathrm{d}V_s\right].$$

Setting $\Phi_t = b_x(t)$, with the help of Lemma 4.3, we obtain

$$\left|\int_0^t E[b_x(s)X_{1,s}]\mathrm{d}s\right|^2 \leq t\int_0^t \left|E[b_x(s)X_{1,s}]\right|^2\mathrm{d}s \leq C\epsilon\tilde{\rho}(\epsilon), \qquad t\in[0,T].$$

We also have

$$\left|\int_0^t E[\delta b(s,u_s^\epsilon)]\mathrm{d}s\right|^2 \leq C\left(E\left[\int_0^t |u_s^\epsilon - \hat{u}_s|\mathrm{d}s\right]\right)^2 \leq C\epsilon^2.$$

Moreover, noticing (2.1) and setting $\Phi_t = \mathbb{E}[\gamma_x(t)|\mathcal{P}]$, by virtue of Lemma 4.3 again, we get

$$\left|E\left[\int_0^t \gamma_x(s)X_{1,s-}\mathrm{d}V_s\right]\right|^2 = \left|E\left[\int_0^t \mathbb{E}[\gamma_x(s)|\mathcal{P}]X_{1,s}r_s\mathrm{d}s\right]\right|^2 \leq Ct\int_0^t \left|E[\mathbb{E}[\gamma_x(s)|\mathcal{P}]X_{1,s}]\right|^2\mathrm{d}s \leq C\epsilon\tilde{\rho}(\epsilon),$$

and

$$\left|E\left[\int_0^t \gamma_{\tilde{x}}(s)E[X_{1,s-}]\mathrm{d}V_s\right]\right|^2 = \left|E\left[\int_0^t \mathbb{E}[\gamma_{\tilde{x}}(s)|\mathcal{P}]E[X_{1,s}]r_s\mathrm{d}s\right]\right|^2 \leq C\int_0^t \left|E[X_{1,s}]\right|^2\mathrm{d}s.$$

Then we have

$$\left|E[X_{1,t}]\right|^2 \leq C\left(\epsilon\tilde{\rho}(\epsilon) + \epsilon^2 + \int_0^t \left|E[X_{1,s}]\right|^2\mathrm{d}s\right), \qquad t\in[0,T].$$

From Gronwall's inequality, we get the first estimate in (4.20), with $\rho(\epsilon) = C(\tilde{\rho}(\epsilon)+\epsilon), \epsilon > 0$.

For the estimation of $Z_{1,\cdot}$, noticing the boundedness of $h, h_x, h_{\tilde{x}}$, we have

$$E\bigg[\bigg(\int_0^T \big|\hat{Z}_t h_x(t) X_{1,t}\big|^2 \mathrm{d}t\bigg)^k\bigg] \le C\bigg(E\bigg[\sup_{t\in[0,T]}\big|X_{1,t}\big|^{4k}\bigg]\bigg)^{\frac{1}{2}} \le C\epsilon^k.$$

Similarly,

$$E\bigg(\int_0^T \big|\hat{Z}_t h_{\tilde{x}}(t) E[X_{1,t}]\big|^2 \mathrm{d}t\bigg)^k \le C\bigg(\sup_{t\in[0,T]}\big|E[X_{1,t}]\big|^2\bigg)^k \le C\epsilon^k \rho^k(\epsilon).$$

By virtue of L^p estimate, we have

$$\begin{aligned}
&E\bigg[\sup_{t\in[0,T]}\big|Z_{1,t}\big|^{2k}\bigg]\\
&\le CE\bigg[\sup_{t\in[0,T]}\bigg|\int_0^t Z_{1,s} h(s)\mathrm{d}Y_s\bigg|^{2k}\bigg] + CE\bigg[\sup_{t\in[0,T]}\bigg|\int_0^t \hat{Z}_s h_x(s) X_{1,s}\mathrm{d}Y_s\bigg|^{2k}\bigg]\\
&\quad + CE\bigg[\sup_{t\in[0,T]}\bigg|\int_0^t \hat{Z}_s h_{\tilde{x}}(s) E[X_{1,s}]\mathrm{d}Y_s\bigg|^{2k}\bigg] + CE\bigg[\sup_{t\in[0,T]}\bigg|\int_0^t \hat{Z}_s \delta h(s,u_s^\epsilon) I_{\mathcal{O}}\mathrm{d}Y_s\bigg|^{2k}\bigg]\\
&\le CE\bigg[\int_0^T \big|\hat{Z}_t h_x(t) X_{1,t}\big|^2\mathrm{d}t\bigg]^k + CE\bigg[\int_0^T \big|\hat{Z}_t \delta h(t,u_t^\epsilon)\big|^2\mathrm{d}t\bigg]^k + CE\bigg[\int_0^T \big|\hat{Z}_t h_{\tilde{x}}(t) E[X_{1,t}]\big|^2\mathrm{d}t\bigg]^k\\
&\le C\epsilon^k.
\end{aligned}$$

By applying Itô's formula to $Z_{1,t}X_{1,t}$, we have

$$\begin{aligned}
E[Z_{1,t}X_{1,t}] &= E\bigg[\int_0^t \Big\{Z_{1,s} b_x(s) X_{1,s} + Z_{1,s} b_{\tilde{x}}(s) E[X_{1,s}] + Z_{1,s}\delta b(s,u_s^\epsilon)\Big\}\mathrm{d}s\bigg]\\
&\qquad + E\bigg[\int_0^t Z_{1,s-}\Big\{\gamma_x(s) X_{1,s-} + \gamma_{\tilde{x}}(s) E[X_{1,s-}]\Big\}\mathrm{d}V_s\bigg]\\
&= E\bigg[\int_0^t \Big\{Z_{1,s}[b_x(s) + \mathbb{E}[\gamma_x(s)|\mathcal{P}] r_s] X_{1,s} + Z_{1,s}[b_{\tilde{x}}(s) + \mathbb{E}[\gamma_{\tilde{x}}(s)] r_s] E[X_{1,s}] + Z_{1,s}\delta b(s,u_s^\epsilon)\Big\}\mathrm{d}s\bigg].
\end{aligned}$$

Similar to above method and noticing (2.1), we obtain

$$\begin{aligned}
|E[Z_{1,t}X_{1,t}]| &\le C\int_0^t |E[Z_{1,s}X_{1,s}]|\mathrm{d}s + C\bigg(E\bigg[\sup_{t\in[0,T]}|Z_{1,t}|^2\bigg]\bigg)^{\frac{1}{2}}\bigg[\sup_{t\in[0,T]}|E[X_{1,t}]|^2\bigg]^{\frac{1}{2}}\\
&\qquad + C\bigg(E\bigg[\sup_{t\in[0,T]}\big|Z_{1,t}\big|^2\bigg]\bigg)^{\frac{1}{2}}\bigg(E\bigg[\bigg|\int_0^t \delta b(s,u_s^\epsilon)\mathrm{d}s\bigg|^2\bigg]\bigg)^{\frac{1}{2}}\\
&\le C\bigg(\int_0^t \big|E[Z_{1,s}X_{1,s}]\big|\mathrm{d}s + \epsilon\rho(\epsilon) + \epsilon^{\frac{3}{2}}\bigg).
\end{aligned}$$

From Gronwall's inequality, we get the second estimate in (4.20).

For the estimation of $Z_{2,\cdot}$, noticing the boundedness of $h, h_x, h_{\tilde{x}}, h_{xx}$, with the same discussion above we have

$$\begin{aligned}
E\Big[\sup_{t\in[0,T]}|Z_{2,t}|^{2k}\Big] &\le CE\Big[\int_0^T |Z_{1,t}h_x(t)X_{1,t}|^2\mathrm{d}t\Big]^k + CE\Big[\int_0^T |Z_{1,t}\delta h(t,u_t^\epsilon)|^2\mathrm{d}t\Big]^k \\
&\quad + CE\Big[\int_0^T |\hat{Z}_t h_x(t)X_{2,t}|^2\mathrm{d}t\Big]^k + CE\Big[\int_0^T |\hat{Z}_t X_{1,t}\delta h_x(t,u_t^\epsilon)|^2\mathrm{d}t\Big]^k \\
&\quad + CE\Big[\int_0^T |\hat{Z}_t h_{\tilde{x}}(t)E[X_{2,t}]|^2\mathrm{d}t\Big]^k + CE\Big[\int_0^T |\frac{1}{2}\hat{Z}_t h_{xx}(t)X_{1,t}^2|^2\mathrm{d}t\Big]^k \\
&\le C\epsilon^{2k}.
\end{aligned}$$

□

Lemma 4.5. *Let (H1)-(H2) hold. Then we have*

$$E\Big[\sup_{t\in[0,T]}\big|X_t^\epsilon - \hat{X}_t - X_{1,t} - X_{2,t}\big|^2\Big] = \epsilon^2\rho(\epsilon), \tag{4.21}$$

$$E\Big[\sup_{t\in[0,T]}\big|Z_t^\epsilon - \hat{Z}_t - Z_{1,t} - Z_{2,t}\big|^2\Big] = \epsilon^2\rho(\epsilon). \tag{4.22}$$

Proof. Set $\Delta X_\cdot := X_{1,\cdot} + X_{2,\cdot}$ $\Delta Z_\cdot := Z_{1,\cdot} + Z_{2,\cdot}$, $\psi_\cdot^\epsilon := X_\cdot^\epsilon - \hat{X}_\cdot - \Delta X_\cdot$, we have

$$\mathrm{d}\psi_t^\epsilon = \Psi^\epsilon(t;b)\mathrm{d}t + \Psi^\epsilon(t;\sigma)\mathrm{d}B_t + \Psi^\epsilon(t;\gamma)\mathrm{d}V_t + \Psi^\epsilon(t;c)\mathrm{d}\tilde{V}_t,$$

where, for $\phi = b, \sigma$, respectively

$$\begin{aligned}
\Psi^\epsilon(t;\phi) &:= \phi(t, X_t^\epsilon, E[X_t^\epsilon], u_t^\epsilon) - \phi(t,\hat{X}_t, E[\hat{X}_t], \hat{u}_t) - \phi_x(t)\Delta X_t \\
&\quad - \phi_{\tilde{x}}(t)E[\Delta X_t] - \frac{1}{2}\phi_{xx}(t)X_{1,t}^2 - \delta\phi(t,u_t^\epsilon) - \delta\phi_x(t,u_t^\epsilon)X_{1,t},
\end{aligned}$$

for $\phi = \gamma, c$, respectively

$$\begin{aligned}
\Psi^\epsilon(t;\phi) &:= \phi(t, X_{t-}^\epsilon, E[X_{t-}^\epsilon], u_t^\epsilon) - \phi(t,\hat{X}_{t-}, E[\hat{X}_{t-}], \hat{u}_t) - \phi_x(t)\Delta X_{t-} \\
&\quad - \phi_{\tilde{x}}(t)E[\Delta X_{t-}] - \frac{1}{2}\phi_{xx}(t)X_{1,t-}^2 - \delta\phi(t,u_t^\epsilon) - \delta\phi_x(t,u_t^\epsilon)X_{1,t-}.
\end{aligned}$$

According to spike variations, we know that $\delta\phi(t,u_t^\epsilon) = \delta\phi_x(t,u_t^\epsilon) = 0$, ζ-a.s.. In order to be consistent with b, σ, we add these two terms to the above equation.

Now we first consider the estimate of $\Psi^\epsilon(t;b)$. For each $t \in [0,T]$, we have

$$b(t, X_t^\epsilon, E[X_t^\epsilon], u_t^\epsilon) - b(t,\hat{X}_t, E[\hat{X}_t], u_t^\epsilon) = \int_0^1 b_x^\theta(t)(X_t^\epsilon - \hat{X}_t) + b_{\tilde{x}}^\theta(t)E[X_t^\epsilon - \hat{X}_t]\mathrm{d}\theta,$$

for the subscript i which indicates the first and the second order derivatives of b, respectively, with respect to $i = x, \tilde{x}, xx, x\tilde{x}, \tilde{x}\tilde{x}$, and for real $\lambda \in [0,1]$,

$$b_i^\lambda(t) := b_i(t, \hat{X}_t + \lambda(X_t^\epsilon - \hat{X}_t), E[\hat{X}_t + \lambda(X_t^\epsilon - \hat{X}_t)], u_t^\epsilon).$$

Moreover

$$\begin{aligned}&b(t,X_t^\epsilon,E[X_t^\epsilon],u_t^\epsilon)-b(t,\hat{X}_t,E[\hat{X}_t],u_t^\epsilon)-b_x(t)\Delta X_t-b_{\tilde{x}}(t)E[\Delta X_t]\\&=\int_0^1\big\{b_x^\theta(t)\psi_t^\epsilon+b_{\tilde{x}}^\theta(t)E[\psi_t^\epsilon]+(b_x^\theta(t)-b_x(t))\Delta X_t+(b_{\tilde{x}}^\theta(t)-b_{\tilde{x}}(t))E[\Delta X_t]\big\}\mathrm{d}\theta.\end{aligned}$$

Similarly, we have

$$b_x^\theta(t)-b_x(t)=\theta\int_0^1\{b_{xx}^{\theta\varrho}(t)\psi_t^\epsilon+b_{x\tilde{x}}^{\theta\varrho}(t)E[\psi_t^\epsilon]+b_{xx}^{\theta\varrho}(t)\Delta X_t+b_{x\tilde{x}}^{\theta\varrho}(t)E[\Delta X_t]\}\mathrm{d}\varrho+\delta b_x(t,u_t^\epsilon),$$

and

$$b_{\tilde{x}}^\theta(t)-b_{\tilde{x}}(t)=\theta\int_0^1\big\{b_{x\tilde{x}}^{\theta\varrho}(t)\psi_t^\epsilon+b_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)E[\psi_t^\epsilon]+b_{x\tilde{x}}^{\theta\varrho}(t)\Delta X_t+b_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)E[\Delta X_t]\big\}\mathrm{d}\varrho+\delta b_{\tilde{x}}(t,u_t^\epsilon).$$

Then we have

$$\Psi^\epsilon(t;b)=\int_0^1\{b_x^\theta(t)\psi_t^\epsilon+b_{\tilde{x}}^\theta(t)E[\psi_t^\epsilon]\mathrm{d}\theta+\Psi_1^\epsilon(t;b)+\Psi_2^\epsilon(t;b)+\Psi_3^\epsilon(t;b)+\Psi_4^\epsilon(t;b),$$

where

$$\begin{cases}\Psi_1^\epsilon(t;b)=\displaystyle\int_0^1\theta\int_0^1\big\{\big(b_{xx}^{\theta\varrho}(t)\psi_t^\epsilon+b_{x\tilde{x}}^{\theta\varrho}(t)E[\psi_t^\epsilon]\big)\Delta X_t+b_{x\tilde{x}}^{\theta\varrho}(t)\psi_t^\epsilon E[\Delta X_t]+b_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)E[\psi_t^\epsilon]E[\Delta X_t]\big\}\mathrm{d}\varrho\mathrm{d}\theta,\\ \Psi_2^\epsilon(t;b)=\displaystyle\int_0^1\theta\int_0^1\big\{b_{xx}^{\theta\varrho}(t)\big[(\Delta X_t)^2-X_{1,t}^2\big]+2b_{x\tilde{x}}^{\theta\varrho}(t)\Delta X_tE[\Delta X_t]+b_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)\big(E[\Delta X_t]\big)^2\big\}\mathrm{d}\varrho\mathrm{d}\theta,\\ \Psi_3^\epsilon(t;b)=\displaystyle\int_0^1\theta\int_0^1\big(b_{xx}^{\theta\varrho}(t)-b_{xx}(t)\big)X_{1,t}^2\mathrm{d}\varrho\mathrm{d}\theta,\\ \Psi_4^\epsilon(t;b)=\delta b_x(t,u_t^\epsilon)X_{2,t}+\delta b_{\tilde{x}}(t,u_t^\epsilon)E[\Delta X_t].\end{cases}$$

From the definition of $\psi_\cdot^\epsilon$ and the Lemma 4.4, we have

$$E\bigg[\sup_{t\in[0,T]}|\psi_t^\epsilon|^2\bigg]\le C\epsilon^2,$$

together with (4.18) and (4.19), this yields

$$E\bigg[\sup_{t\in[0,T]}|\Psi_1^\epsilon(t;b)|^2\bigg]\le C\epsilon^3. \tag{4.23}$$

With the help of (4.18) and (4.19), we also have

$$E\bigg[\sup_{t\in[0,T]}|(\Delta X_t)^2-X_{1,t}^2|^2\bigg]\le C\epsilon^3.$$

Using (4.18)–(4.20), we get

$$E\Big[\sup_{t\in[0,T]}|\Delta X_t\cdot E[\Delta X_t]|^2\Big]\le \epsilon^2\rho(\epsilon),$$

Similarly, we obtain that

$$\sup_{t\in[0,T]}\Big|E[\Delta X_t]\Big|^4\le \epsilon^2\rho(\epsilon),$$

Thus

$$E\Big[\sup_{t\in[0,T]}|\Psi_2^{\epsilon}(t;b)|^2\Big]\le \epsilon^2\rho(\epsilon). \tag{4.24}$$

Using Hölder's inequality, we obtain

$$\begin{aligned}E\Big[\int_0^T|\Psi_3^{\epsilon}(t;b)|^2\mathrm{d}t\Big]&\le CE\Big[\sup_{t\in[0,T]}|X_{1,t}|^4\cdot\int_0^T\int_0^1\int_0^1|b_{xx}^{\theta\varrho}(t)-b_{xx}(t)\big)|^2\mathrm{d}\theta\mathrm{d}\varrho\mathrm{d}t\Big]\\&\le C\Big(E\Big[\sup_{t\in[0,T]}|X_{1,t}|^8\Big]\Big)^{\frac12}\Big(E\Big[\int_0^T\int_0^1\int_0^1|b_{xx}^{\theta\varrho}(t)-b_{xx}(t)\big)|^4\mathrm{d}\theta\mathrm{d}\varrho\mathrm{d}t\Big]\Big)^{\frac12}.\end{aligned}$$

Set $\bar\rho_1(\epsilon):=C(E[\int_0^T\int_0^1\int_0^1|b_{xx}^{\theta\varrho}(t)-b_{xx}(t)|^4\mathrm{d}\theta\mathrm{d}\varrho\mathrm{d}t])^{\frac12}$ and then by the dominated convergence theorem, we can prove $\bar\rho_1(\epsilon)\to0$ as $\epsilon\downarrow0$. Hence, in view of (4.18), it holds that

$$E\Big[\int_0^T|\Psi_3^{\epsilon}(t;b)|^2\mathrm{d}t\Big]\le\epsilon^2\bar\rho_1(\epsilon). \tag{4.25}$$

Finally, noticing (4.19) and (4.20), we have

$$\begin{aligned}&E\Big[\int_0^T|\delta b(t,u_t^{\epsilon})X_{2,t}|^2\mathrm{d}t\Big]\le C\epsilon E\Big[\sup_{t\in[0,T]}|X_{2,t}|^2\Big]\le C\epsilon^2,\\&E\Big[\int_0^T|\delta b_{\tilde x}(t,u_t^{\epsilon})E[X_{1,t}]|^2\mathrm{d}t\Big]\le C\epsilon\sup_{t\in[0,T]}\big|E[X_{1,t}]\big|^2\le C\epsilon^2\rho(\epsilon),\\&E\Big[\int_0^T|\delta b_{\tilde x}(t,u_t^{\epsilon})E[X_{2,t}]|^2\mathrm{d}t\Big]\le C\epsilon\sup_{t\in[0,T]}\big|E[X_{2,t}]\big|^2\le C\epsilon^3.\end{aligned}$$

Therefore,

$$E\Big[\int_0^T|\Psi_4^{\epsilon}(t;b)|^2\mathrm{d}t\Big]\le\epsilon^2\rho(\epsilon), \tag{4.26}$$

Finally, by virtue of (4.23), (4.24)–(4.26), we have

$$E\Big[\int_0^T|\Psi^{\epsilon}(t;b)|^2\mathrm{d}t\Big]\le\epsilon^2\rho(\epsilon)+C\int_0^TE\big[|\psi_t^{\epsilon}|^2\big]\mathrm{d}t, \tag{4.27}$$

Similarly, we can prove that

$$E\Big[\int_0^T |\Psi^\epsilon(t;\sigma)|^2 \mathrm{d}t\Big] \le \epsilon^2\rho(\epsilon) + C\int_0^T E\big[|\psi_t^\epsilon|^2\big]\mathrm{d}t. \tag{4.28}$$

For $\Psi^\epsilon(t;\gamma)$ and $\Psi^\epsilon(t;c)$

$$\Psi^\epsilon(t;\gamma) = \int_0^1 \big\{\gamma_x^\theta(t)\psi_{t-}^\epsilon + \gamma_{\tilde{x}}^\theta(t)E[\psi_{t-}^\epsilon]\big\}\mathrm{d}\theta + \Psi_1^\epsilon(t;\gamma) + \Psi_2^\epsilon(t;\gamma) + \Psi_3^\epsilon(t;\gamma),$$

where

$$\begin{cases} \Psi_1^\epsilon(t;\gamma) = \int_0^1 \theta\int_0^1 \big\{\big(\gamma_{xx}^{\theta\varrho}(t)\psi_{t-}^\epsilon + \gamma_{x\tilde{x}}^{\theta\varrho}(t)E[\psi_{t-}^\epsilon]\big)\Delta X_{t-} + \gamma_{x\tilde{x}}^{\theta\varrho}(t)\psi_{t-}^\epsilon E[\Delta X_{t-}] + \gamma_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)E[\psi_{t-}^\epsilon]E[\Delta X_{t-}]\big\}\mathrm{d}\varrho\mathrm{d}\theta, \\ \Psi_2^\epsilon(t;\gamma) = \int_0^1 \theta\int_0^1 \big\{\gamma_{xx}^{\theta\varrho}(t)[(\Delta X_{t-})^2 - X_{1,t-}^2] + 2\gamma_{x\tilde{x}}^{\theta\varrho}(t)\Delta X_{t-}E[\Delta X_{t-}] + \gamma_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)\big(E[\Delta X_{t-}]\big)^2\big\}\mathrm{d}\varrho\mathrm{d}\theta, \\ \Psi_3^\epsilon(t;\gamma) = \int_0^1 \theta\int_0^1 \big[\gamma_{xx}^{\theta\varrho}(t) - \gamma_{xx}(t)\big]X_{1,t-}^2\mathrm{d}\varrho\mathrm{d}\theta. \end{cases}$$

By the similar method, we also have

$$E\Big[\sup_{t\in[0,T]}|\Psi_1^\epsilon(t;\gamma)|^2\Big] \le C\epsilon^3, \qquad \text{and} \qquad E\Big[\sup_{t\in[0,T]}|\Psi_2^\epsilon(t;\gamma)|^2\Big] \le \epsilon^2\rho(\epsilon). \tag{4.29}$$

Since the coefficients are progressive rather than predictable, the proof of the estimation of $\Psi_3^\epsilon(t;\gamma)$ is different. Using Hölder's inequality, we get

$$E\Big[\Big(\int_0^T |\Psi_3^\epsilon(t;\gamma)|\mathrm{d}V_t\Big)^2\Big] \le C\Big(E\Big[\sup_{t\in[0,T]}|X_{1,t}|^8\Big]\Big)^{\frac{1}{2}}\bar{\rho}_2^{\frac{1}{2}}(\epsilon),$$

where $\bar{\rho}_2(\epsilon) = E(\int_0^T\int_0^1\int_0^1 |\gamma_{xx}^{\theta\varrho}(t) - \gamma_{xx}(t)\big)|\mathrm{d}\theta\mathrm{d}\varrho\mathrm{d}V_t)^4$. It is easy to verify that $\gamma_{xx}^{\theta\varrho}(t) - \gamma_{xx}(t) \to 0$ as $\epsilon \downarrow 0$.

$$E\Big[\Big(\int_0^T |\Psi_3^\epsilon(t;\gamma)|\mathrm{d}V_t\Big)^2\Big] \le \epsilon^2\bar{\rho}_2(\epsilon). \tag{4.30}$$

By virtue of (4.29)–(4.30), we obtain

$$E\Big[\Big(\int_0^T |\Psi^\epsilon(t;\gamma)|\mathrm{d}V_t\Big)^2\Big] \le \epsilon^2\rho(\epsilon) + C\int_0^T E\big[|\psi_t^\epsilon|^2\big]\mathrm{d}t, \tag{4.31}$$

Similarly, we also have

$$E\Big[\int_0^T |\Psi^\epsilon(t;c)|^2\mathrm{d}V_t\Big] \le \epsilon^2\rho(\epsilon) + C\int_0^T E\big[|\psi_t^\epsilon|^2\big]\mathrm{d}t, \tag{4.32}$$

From the definition of ψ_t^ϵ, by virtue of B-D-G inequality and (4.28), (4.29), (4.31) and (4.32), we have

$$\begin{aligned} E\Big[\sup_{t\in[0,T]}|\psi_t^\epsilon|^2\Big] &\le CE\Big[\int_0^T|\Psi^\epsilon(t;b)|^2\mathrm{d}t\Big]+CE\Big[\int_0^T|\Psi^\epsilon(t;\gamma)|\mathrm{d}V_t\Big]^2\\ &\quad+CE\Big[\int_0^T|\Psi^\epsilon(t;\sigma)|^2\mathrm{d}t\Big]+CE\Big[\int_0^T|\Psi^\epsilon(t;c)|^2\mathrm{d}V_t\Big]\\ &\le C\Big[\epsilon^2\rho(\epsilon)+\int_0^T E\big[|\psi_t^\epsilon|^2\big]\mathrm{d}t\Big]. \end{aligned}$$

Hence, by Gronwall's inequality, we obtain

$$E\Big[\sup_{t\in[0,T]}|\psi_t^\epsilon|^2\Big]\le C\epsilon^2\rho(\epsilon)$$

Then we can obtain the estimate (4.21). And the estimate (4.22) can be proved similarly. □

Now we give the variation equation for the cost functional. Define

$$\begin{aligned} \hat{J}=E\Big[&\int_0^T\Big\{\hat{Z}_t l_x(t)\Delta X_t+\hat{Z}_t\delta l(t,u_t^\epsilon)+\Delta Z_t l(t)+Z_{1,t}l_x(t)X_{1,t}+\hat{Z}_t l_{\tilde{x}}(t)E\big[\hat{Z}_t\Delta X_t+\Delta Z_t\hat{X}_t\big]\\ &+\frac{1}{2}\hat{Z}_t l_{xx}(t)X_{1,t}^2\Big\}\mathrm{d}t+\int_0^T\Big\{\hat{Z}_{t-}f_x(t)\Delta X_{t-}+\Delta Z_{t-}f(t)+\hat{Z}_{t-}f_{\tilde{x}}(t)E\big[\hat{Z}_{t-}\Delta X_{t-}+\Delta Z_{t-}\hat{X}_{t-}\big]\\ &+Z_{1,t-}f_x(t)X_{1,t-}+\frac{1}{2}\hat{Z}_{t-}f_{xx}(t)X_{1,t-}^2\Big\}\mathrm{d}V_t+\hat{Z}_T g_x(T)\Delta X_T+\frac{1}{2}\hat{Z}_T g_{xx}(T)X_{1,T}^2\\ &+\Delta Z_T g(T)+Z_{1,T}g_x(T)X_{1,T}+\hat{Z}_T g_{\tilde{x}}(T)E\big[\hat{Z}_T\Delta X_T+\Delta Z_T\hat{X}_T\big]\Big]. \end{aligned} \tag{4.33}$$

Lemma 4.6. *Let (H1)-(H2) hold. Then we have*

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}(J(u_\cdot^\epsilon)-J(\hat{u}_\cdot)-\hat{J})=0.$$

Proof. By (4.33), we have

$$J(u_\cdot^\epsilon)-J(\hat{u}_\cdot)-\hat{J}=E\Big[\int_0^T\Psi^\epsilon(t;l)\mathrm{d}t+\int_0^T\Psi^\epsilon(t;f)\mathrm{d}V_t+\Psi^\epsilon(T;g)\Big],$$

where

$$\begin{aligned} \Psi^\epsilon(t;l)&=Z_t^\epsilon l(t,X_t^\epsilon,E[Z_t^\epsilon X_t^\epsilon],u_t^\epsilon)-\hat{Z}_t\big\{l(t)+l_x(t)\Delta X_t+\frac{1}{2}l_{xx}(t)X_{1,t}^2\\ &\qquad+\delta l(t,u_t^\epsilon)+l_{\tilde{x}}(t)E\big[\hat{Z}_t\Delta X_t+\Delta Z_t\hat{X}_t\big]\big\}-\Delta Z_t l(t)-Z_{1,t}l_x(t)X_{1,t},\\ \Psi^\epsilon(t;f)&=Z_{t-}^\epsilon f(t,X_{t-}^\epsilon,E[Z_{t-}^\epsilon X_{t-}^\epsilon],u_t^\epsilon)-\hat{Z}_{t-}\big\{f(t)+f_x(t)\Delta X_{t-}+\frac{1}{2}f_{xx}(t)X_{1,t-}^2\\ &\qquad+f_{\tilde{x}}(t)E\big[\hat{Z}_{t-}\Delta X_{t-}+\Delta Z_{t-}\hat{X}_{t-}\big]\big\}-\Delta Z_{t-}f(t)-Z_{1,t-}f_x(t)X_{1,t-},\\ \Psi^\epsilon(T;g)&=Z_T^\epsilon g(X_T^\epsilon,E[Z_T^\epsilon X_T^\epsilon])-\hat{Z}_T\big\{g(T)+g_x(T)\Delta X_T+\frac{1}{2}g_{xx}(T)X_{1,T}^2\\ &\qquad-g_{\tilde{x}}(T)E\big[\hat{Z}_T\Delta X_T+\Delta Z_T\hat{X}_T\big]\big\}-\Delta Z_T g(T)-Z_{1,T}g_x(T)X_{1,T}. \end{aligned}$$

We first consider the estimate of $\Psi^\epsilon(t;l)$. Set $L(t,Z_t,X_t,E[Z_tX_t],u_t)=Z_tl(t,X_t,E[Z_tX_t],u_t)$. Similar to the proof of Lemma 4.5, we have

$$\Psi^\epsilon(t;l)=\int_0^1\big\{L_z^\theta(t)\phi_t^\epsilon+L_x^\theta(t)\psi_t^\epsilon+L_{\tilde{x}}^\theta(t)E[\varphi_t^\epsilon]\big\}\mathrm{d}\theta+\Psi_3^\epsilon(t;L)+\Psi_4^\epsilon(t;L)+\Psi_5^\epsilon(t;L),$$

where $\phi_\cdot^\epsilon:=Z_\cdot^\epsilon-\hat{Z}_\cdot-\Delta Z_\cdot$, $\varphi_\cdot^\epsilon:=Z_\cdot^\epsilon X_\cdot^\epsilon-\hat{Z}_\cdot\Delta X_\cdot-\Delta Z_\cdot\hat{X}_\cdot-\hat{Z}_\cdot\hat{X}_\cdot$ and

$$\begin{aligned}
\Psi_1^\epsilon(t;L)=&\int_0^1\theta\int_0^1\{[L_{z\tilde{x}}^{\theta\varrho}(t)\phi_t^\epsilon+L_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)\psi_t^\epsilon+L_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)E[\varphi_t^\epsilon]]E[\hat{Z}_t\Delta X_t+\Delta Z_t\hat{X}_t]\\
&+(L_{xz}^{\theta\varrho}(t)\phi_t^\epsilon+L_{xx}^{\theta\varrho}(t)\psi_t^\epsilon+L_{x\tilde{x}}^{\theta\varrho}(t)E[\varphi_t^\epsilon])\Delta X_t+(L_{zx}^{\theta\varrho}(t)\psi_t^\epsilon+L_{z\tilde{x}}^{\theta\varrho}(t)E[\varphi_t^\epsilon])\Delta Z_t\}\mathrm{d}\varrho\mathrm{d}\theta,\\
\Psi_2^\epsilon(t;L)=&\int_0^1\theta\int_0^1\{2L_{zx}^{\theta\varrho}(t)(\Delta X_t\Delta Z_t-X_{1,t}Z_{1,t})+L_{xx}^{\theta\varrho}(t)(\Delta X_t)^2-X_{1,t}^2)+(L_{\tilde{x}z}^{\theta\varrho}(t)+L_{z\tilde{x}}^{\theta\varrho}(t))\Delta Z_tE[\hat{Z}_t\Delta X_t]\\
&+L_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)(E[\hat{Z}_t\Delta X_t])^2+2L_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)E[\hat{Z}_t\Delta X_t]E[\Delta Z_t\hat{X}_t]+2L_{x\tilde{x}}^{\theta\varrho}(t)\Delta X_tE[\hat{Z}_t\Delta X_t]\}\mathrm{d}\varrho\mathrm{d}\theta,\\
\Psi_3^\epsilon(t;L)=&\int_0^1\theta\int_0^1(L_{xx}^{\theta\varrho}(t)-L_{xx}(t))X_{1,t}^2+2(L_{zx}^{\theta\varrho}(t)-L_{zx}(t))Z_{1,t}X_{1,t})\mathrm{d}\varrho\mathrm{d}\theta,\\
\Psi_4^\epsilon(t;L)=&\int_0^1\theta\int_0^1\{[(L_{z\tilde{x}}^{\theta\varrho}(t)+L_{\tilde{x}z}^{\theta\varrho}(t))\Delta Z_t+2L_{x\tilde{x}}^{\theta\varrho}(t)\Delta X_t]E[\Delta Z_t\hat{X}_t]+L_{\tilde{x}\tilde{x}}^{\theta\varrho}(t)(E[\Delta Z_t\hat{X}_t])^2\}\mathrm{d}\varrho\mathrm{d}\theta,\\
\Psi_5^\epsilon(t;L)=&\,\delta L_z(t,u_t^\epsilon)\Delta Z_t+\delta L_x(t,u_t^\epsilon)\Delta X_t+\delta L_{\tilde{x}}(t,u_t^\epsilon)E[\hat{Z}_t\Delta X_t+\Delta Z_t\hat{X}_t].
\end{aligned}$$

From the definition of $\psi_\cdot^\epsilon,\phi_\cdot^\epsilon$ and $\varphi_\cdot^\epsilon$, noticing Lemma 4.5,

$$E\Big[\sup_{t\in[0,T]}|\psi_t^\epsilon|^2\Big]\le\epsilon^2\rho(\epsilon),\qquad E\Big[\sup_{t\in[0,T]}|\phi_t^\epsilon|^2\Big]\le\epsilon^2\rho(\epsilon),$$

by virtue of (4.20), we have

$$E\big[|\varphi_t^\epsilon|\big]\le E\big[|\phi_t^\epsilon X_t^\epsilon|+|(Z_t+\Delta Z_t)\psi_t^\epsilon|+|\Delta Z_t\Delta X_t|\big]\le\rho(\epsilon).$$

With (4.18) and (4.19), we have

$$E\Big[\sup_{t\in[0,T]}|\Psi_1^\epsilon(t;L)|\Big]\le C\epsilon^{\frac{3}{2}}.\tag{4.34}$$

Noticing (4.18) and (4.19), we have

$$E\Big[\sup_{t\in[0,T]}|\Delta X_t\Delta Z_t-X_{1,t}Z_{1,t}|\Big]+E\Big[\sup_{t\in[0,T]}|(\Delta X_t)^2-X_{1,t}^2|\Big]\le C\epsilon^{\frac{3}{2}}.$$

Again, using (4.18)–(4.20), we have

$$\begin{aligned}
&E\Big[\sup_{t\in[0,T]}|\Delta Z_tE[\hat{Z}_t\Delta X_t]|\Big]\le\epsilon\rho(\epsilon),\qquad E\Big[\sup_{t\in[0,T]}|\Delta X_tE[\hat{Z}_t\Delta X_t]|\Big]\le\epsilon\rho(\epsilon),\\
&\sup_{t\in[0,T]}|E[\hat{Z}_t\Delta X_t]E[\Delta Z_t\hat{X}_t]|\le\epsilon\rho(\epsilon),\qquad \sup_{t\in[0,T]}|E[\hat{Z}_t\Delta X_t]|^2\le\epsilon\rho(\epsilon).
\end{aligned}$$

Hence,

$$E\bigg[\sup_{t\in[0,T]}|\Psi_2^\epsilon(t;L)|\bigg]\leq \epsilon\rho(\epsilon). \tag{4.35}$$

Using Hölder inequality, we get

$$\begin{aligned}
&E\bigg[\int_0^T\int_0^1\theta\int_0^1|(L_{xx}^{\theta\varrho}(t)-L_{xx}(t))X_{1,t}^2|\mathrm{d}\varrho\mathrm{d}\theta\mathrm{d}t\bigg]\\
&\leq C\bigg(E\bigg[\sup_{t\in[0,T]}X_{1,t}^4\bigg]\bigg)^{\frac{1}{2}}\bigg(E\bigg[\int_0^T\int_0^1\int_0^1|(L_{xx}^{\theta\varrho}(t)-L_{xx}(t))|^2\mathrm{d}\varrho\mathrm{d}\theta\mathrm{d}t\bigg]\bigg)^{\frac{1}{2}}\\
&\leq \epsilon\bar{\rho}_3(\epsilon),
\end{aligned}$$

by dominated convergence theorem, we can prove $\bar{\rho}_3(\epsilon):=C(E[\int_0^T\int_0^1\int_0^1|L_{xx}^{\theta\varrho}(t)-L_{xx}(t)|^2\ \mathrm{d}\varrho\mathrm{d}\theta\mathrm{d}t])^{\frac{1}{2}}$ satisfies $\bar{\rho}_3(\epsilon)\to 0$ as $\epsilon\downarrow 0$. Similarly, we also have

$$E\bigg[\int_0^T\int_0^1\theta\int_0^1|(L_{zx}^{\theta\varrho}(t)-L_{zx}(t))Z_{1,t}X_{1,t}|\mathrm{d}\varrho\mathrm{d}\theta\mathrm{d}t\bigg]\leq \epsilon\rho(\epsilon).$$

Then we have

$$E\bigg[\sup_{t\in[0,T]}|\Psi_3^\epsilon(t;L)|\bigg]\leq \epsilon\rho(\epsilon). \tag{4.36}$$

Applying Itô's formula, we can verify that the solution of $Z_{1,\cdot}$ has the following explicit expression

$$\begin{aligned}
Z_{1,t}=&-\hat{Z}_t\int_0^t\Big(h_x(s)X_{1,s}+h_{\tilde{x}}(s)E[X_{1,s}]+\delta h(s,u_s^\epsilon)\Big)\mathrm{d}s\\
&+\hat{Z}_t\int_0^t\Big(h_x(s)X_{1,s}+h_{\tilde{x}}(s)E[X_{1,s}]+\delta h(s,u_s^\epsilon)\Big)\mathrm{d}Y_s,
\end{aligned}$$

then we have

$$E[Z_{1,t}]=-\int_0^t\Big(E[\hat{Z}_sh_x(s)X_{1,s}]+E[\hat{Z}_sh_{\tilde{x}}(s)]E[X_{1,s}]+E[\hat{Z}_s\delta h(s,u_s^\epsilon)]\Big)\mathrm{d}s.$$

In fact, using Doob's inequality,

$$E\bigg[\sup_{t\in[0,T]}|\hat{Z}_sh_x(s)|^p\bigg]\leq C\bigg(\frac{p}{p-1}\bigg)^pE\big[|\hat{Z}_T|^p\big]\leq C_p.$$

With the help of Lemma 4.3, by setting $\Phi_t=\hat{Z}_th_x(t)$, we obtain

$$\bigg|\int_0^tE[\hat{Z}_sh_x(s)X_{1,s}]\mathrm{d}s\bigg|^2\leq t\int_0^t\big|E[\hat{Z}_sh_x(s)X_{1,s}]\big|^2\mathrm{d}s\leq T\epsilon\tilde{\rho}(\epsilon).$$

By view of (H1)-(H2) and Lemma 4.4, it holds that

$$\big|E[Z_{1,t}]\big| \leq (\epsilon\rho(\epsilon))^{\frac{1}{2}}. \tag{4.37}$$

Then we have the following estimate with a similar argument

$$E\bigg[\sup_{t\in[0,T]}|\Psi_4^\epsilon(t;L)|\bigg] \leq \epsilon\rho(\epsilon). \tag{4.38}$$

In virtue of (4.18)–(4.20), we know that,

$$E\bigg[\sup_{t\in[0,T]}|\Psi_5^\epsilon(t;L)|\bigg] \leq \epsilon\rho(\epsilon). \tag{4.39}$$

According to Lemma 4.5 and (4.34)–(4.36), (4.38) and (4.39), we have

$$E\bigg[\int_0^T|\Psi^\epsilon(t;l)|\mathrm{d}t\bigg] \leq C\epsilon\rho(\epsilon).$$

Similarly, we show that

$$E\bigg[\int_0^T|\Psi^\epsilon(t,e;f)|\mathrm{d}V_t\bigg] \leq C\epsilon\rho(\epsilon), \qquad \text{and} \qquad E\big[|\Psi^\epsilon(T;g)|\big] \leq C\epsilon\rho(\epsilon).$$

Then we get the result directly. □

Now we can rewrite (4.33) as

$$\begin{aligned}\hat{J} = \hat{E}\bigg[&\int_0^T\big\{l_x(t)\Delta X_t + \frac{1}{2}l_{xx}(t)X_{1,t}^2 + l_{\tilde{x}}(t)\hat{E}[\Delta X_t + \Delta\Gamma_t\hat{X}_t] + \delta l(t,u_t^\epsilon) + \Delta\Gamma_t l(t) + \Gamma_{1,t}l_x(t)X_{1,t}\big\}\mathrm{d}t\\ &+\int_0^T\big\{f_x(t)\Delta X_{t-} + \frac{1}{2}f_{xx}(t)X_{1,t-}^2 + \Delta\Gamma_{t-}f(t) + f_{\tilde{x}}(t)\hat{E}[\Delta X_{t-} + \Delta\Gamma_t\hat{X}_{t-}] + \Gamma_{1,t-}f_x(t)X_{1,t-}\big\}\mathrm{d}V_t\\ &+g_x(T)\Delta X_T + \frac{1}{2}g_{xx}(T)X_{1,T}^2 + g_{\tilde{x}}(T)\hat{E}[\Delta X_T + \Delta\Gamma_t\hat{X}_T] + \Delta\Gamma_T g(T) + \Gamma_{1,T}g_x(T)X_{1,T}\bigg],\end{aligned}$$

where $\Delta\Gamma_\cdot = \Gamma_{1,\cdot} + \Gamma_{2,\cdot}$ with $\Gamma_{1,\cdot} = \hat{Z}_\cdot^{-1}Z_{1,\cdot}$ and $\Gamma_{2,\cdot} = \hat{Z}_\cdot^{-1}Z_{2,\cdot}$.

5. Adjoint equations and the maximum principle

By Itô's formula and simple calculation, the solutions to (4.2) and (4.3) have the following explicit expressions:

$$\begin{aligned}Z_{1,t} &= \hat{Z}_t\int_0^t\big\{h_x(s)X_{1,s} + h_{\tilde{x}}(s)E[X_{1,s}] + \delta h(s,u_s^\epsilon)\big\}\mathrm{d}\hat{W}_s,\\ Z_{2,t} &= \hat{Z}_t\int_0^t\big\{h_x(s)X_{2,s} + h_{\tilde{x}}(s)E[X_{2,s}] + \delta h_x(s,u_s^\epsilon)X_{1,s}\\ &\qquad + \frac{1}{2}h_{xx}(s)X_{1,s}^2 + \hat{Z}_s^{-1}Z_{1,s}[\delta h(s,u_s^\epsilon) + h_x(s)X_{1,s}]\big\}\mathrm{d}\hat{W}_s.\end{aligned}$$

Then we have

$$\Gamma_{1,t} = \int_0^t \big\{h_x(s)X_{1,s} + h_{\tilde{x}}(s)E[X_{1,s}] + \delta h(s,u_s^\epsilon)\big\}\mathrm{d}\hat{W}_s,$$
$$\Gamma_{2,t} = \int_0^t \big\{h_x(s)X_{2,s} + h_{\tilde{x}}(s)E[X_{2,s}] + \delta h_x(s,u_s^\epsilon)X_{1,s} + \frac{1}{2}h_{xx}(s)X_{1,s}^2 + \Gamma_{1,s}[\delta h(s,u_s^\epsilon) + h_x(s)X_{1,s}]\big\}\mathrm{d}\hat{W}_s.$$

Now we introduce the first order and second order adjoint equations.

$$\begin{cases} \mathrm{d}z_t = -\big\{l(t) + \mathbb{E}[f(t)|\mathcal{P}]r_t + \hat{X}_t\hat{E}[\mathbb{E}[f_{\tilde{x}}(t)|\mathcal{P}]r_t] + \hat{X}_t\hat{E}[l_{\tilde{x}}(t)]\big\}\mathrm{d}t \\ \qquad\quad + z_{1,t}\mathrm{d}B_t + z_{2,t}\mathrm{d}\hat{W}_t + \varsigma_t\mathrm{d}\tilde{V}_t, \\ z_T = g(T) + \hat{X}_T\hat{E}[g_{\tilde{x}}(T)], \qquad t\in[0,T]. \end{cases} \tag{5.1}$$

$$\begin{cases} \mathrm{d}\xi_t = -\big\{\xi_t b_x(t) + \zeta_t\sigma_x(t) + z_{2,t}h_x(t) + l_x(t) + \mathbb{E}[\gamma_x(t)|\mathcal{P}]r_t\xi_t + \mathbb{E}[\gamma_x(t) \\ \qquad\quad + c_x(t)|\mathcal{P}]r_t\vartheta_t + \mathbb{E}[f_x(t)|\mathcal{P}]r_t\big\}\mathrm{d}t + \zeta_t\mathrm{d}B_t + \tilde{\zeta}_t\mathrm{d}\hat{W}_t + \vartheta_t\mathrm{d}\tilde{V}_t, \\ \xi_T = g_x(T), \qquad t\in[0,T]. \end{cases} \tag{5.2}$$

$$\begin{cases} \mathrm{d}p_t = -\big\{p_t b_x(t) + q_t\sigma_x(t) + z_{2,t}h_x(t) + l_x(t) + \hat{E}[l_{\tilde{x}}(t)] + \mathbb{E}[\gamma_x(t)|\mathcal{P}]p_t r_t \\ \qquad\quad + \mathbb{E}[\gamma_x(t) + c_x(t)|\mathcal{P}]k_t r_t + \mathbb{E}[f_x(t)|\mathcal{P}]r_t + \hat{E}[\mathbb{E}[f_{\tilde{x}}(t)|\mathcal{P}]r_t] \\ \qquad\quad + \hat{Z}_t^{-1}\hat{E}\big[p_t b_{\tilde{x}}(t) + q_t\sigma_{\tilde{x}}(t) + z_{2,t}h_{\tilde{x}}(t) + \mathbb{E}[\gamma_{\tilde{x}}(t) + c_{\tilde{x}}(t)|\mathcal{P}]k_t r_t \\ \qquad\quad + \mathbb{E}[\gamma_{\tilde{x}}(t)|\mathcal{P}]p_t r_t\big]\big\}\mathrm{d}t + q_t\mathrm{d}B_t + \tilde{q}_t\mathrm{d}\hat{W}_t + k_t\mathrm{d}\tilde{V}_t, \\ p_T = g_x(T) + \hat{E}[g_{\tilde{x}}(T)], \qquad t\in[0,T]. \end{cases} \tag{5.3}$$

$$\begin{cases} \mathrm{d}P_t = -\big\{2b_x(t)P_t + P_t\sigma_x^2(t) + 2\sigma_x(t)Q_t + b_{xx}(t)p_t + \sigma_{xx}(t)q_t + h_{xx}(t)z_{2,t} \\ \qquad\quad + 2h_x(t)\tilde{\zeta}_t + l_{xx}(t) + \big(2\mathbb{E}[\gamma_x(t)|\mathcal{P}]P_t + 2\mathbb{E}[\gamma_x(t) + c_x(t)|\mathcal{P}]K_t \\ \qquad\quad + \mathbb{E}[(\gamma_x(t) + c_x(t))^2|\mathcal{P}](P_t + K_t) + \mathbb{E}[\gamma_{xx}(t) + c_{xx}(t)|\mathcal{P}]k_t \\ \qquad\quad + p_t\mathbb{E}[\gamma_{xx}(t)|\mathcal{P}] + \mathbb{E}[f_{xx}(t)|\mathcal{P}]\big)r_t\big\}\mathrm{d}t + Q_t\mathrm{d}B_t + \tilde{Q}_t\mathrm{d}\hat{W}_t + K_t\mathrm{d}\tilde{V}_t, \\ P_T = g_{xx}(T), \qquad t\in[0,T]. \end{cases} \tag{5.4}$$

Since the partial derivatives of the state coefficients are bounded, we can easy to verify the well-posedness of the above MF-BSDEs (see [22], [27]).

Applying Itô's formula to $z_t\Delta\Gamma_t$, we have

$$\begin{aligned} &\hat{E}\big[g(T)\Delta\Gamma_T + \hat{E}[g_{\tilde{x}}(T)]\hat{X}_T\Delta\Gamma_T\big] \\ &= \hat{E}\bigg[\int_0^T \Big\{-\Delta\Gamma_t\big(l(t) + \hat{X}_t\hat{E}[l_{\tilde{x}}(t)] + \mathbb{E}[f(t)|\mathcal{P}]r_t + \hat{X}_t\hat{E}[\mathbb{E}[f_{\tilde{x}}(t)|\mathcal{P}]]r_t\big) + z_{2,t}\big(h_x(t)\Delta X_t \\ &\quad + h_{\tilde{x}}(t)E[\Delta X_t] + \delta h(t,u_t^\epsilon) + \delta h_x(t,u_t^\epsilon)X_{1,t} + \frac{1}{2}h_{xx}(t)X_{1,t}^2 + \Gamma_{1,t}\big[\delta h(t,u_t^\epsilon) + h_x(t)X_{1,t}\big]\big)\Big\}\mathrm{d}t\bigg]. \end{aligned}$$

Noticing

$$\hat{E}\Big[\int_0^T z_{2,t}\Big(\delta h_x(t,u_t^\epsilon)X_{1,t}+\Gamma_{1,t}\delta h(t,u_t^\epsilon)\Big)\mathrm{d}t\Big]=o(\epsilon).$$

By virtue of (2.1), we have

$$\begin{aligned}&\hat{E}\Big[\int_0^T\{\Delta\Gamma_t l(t)+l_{\tilde{x}}(t)\hat{E}[\Delta\Gamma_t\hat{X}_t]\}\mathrm{d}t+\int_0^T\{f_{\tilde{x}}(t)\hat{E}[\Delta\Gamma_{t-}\hat{X}_{t-}]+\Delta\Gamma_{t-}f(t)\}\mathrm{d}V_t+g_{\tilde{x}}(T)\hat{E}[\Delta\Gamma_T\hat{X}_T]+g(T)\Delta\Gamma_T\Big]\\&=\hat{E}\Big[\int_0^T z_{2,t}\Big\{h_x(t)\Delta X_t+h_{\tilde{x}}(t)E[\Delta X_t]+\delta h(t,u_t^\epsilon)+\frac{1}{2}h_{xx}(t)X_{1,t}^2+\Gamma_{1,t}h_x(t)X_{1,t}\Big\}\mathrm{d}t\Big]+o(\epsilon).\end{aligned}\tag{5.5}$$

Applying Itô's formula to $\xi_t\Gamma_{1,t}X_{1,t}$, we have

$$\begin{aligned}\hat{E}\big[\Gamma_{1,T}g_x(T)X_{1,T}\big]=\hat{E}\Big[&\int_0^T\Big\{\tilde{\zeta}_t h_x(t)X_{1,t}^2-\Gamma_{1,t}l_x(t)X_{1,t}-z_{2,t}h_x(t)\Gamma_{1,t}X_{1,t}+\big[\tilde{\zeta}_t h_{\tilde{x}}(t)X_{1,t}+\big(\xi_t b_{\tilde{x}}(t)\\&+\sigma_{\tilde{x}}\zeta_t\big)\Gamma_{1,t}\big]E[X_{1,t}]+\xi_t\Gamma_{1,t}\delta b(t,u_t^\epsilon)+\zeta_t\Gamma_{1,t}\delta\sigma(t,u_t^\epsilon)+\tilde{\zeta}_t X_{1,t}\delta h(t,u_t^\epsilon)\Big\}\mathrm{d}t\\&+\int_0^T\Big\{\big[\xi_{t-}\gamma_{\tilde{x}}(t)+\vartheta_t\gamma_{\tilde{x}}(t)+\vartheta_t c_{\tilde{x}}(t)\big)\Gamma_{1,t-}E[X_{1,t-}]+\Gamma_{1,t-}f_x(t)X_{1,t-}\Big\}\mathrm{d}V_t\Big].\end{aligned}$$

Similar to Lemma 3.3 of [5], noticing (4.20) we have

$$\begin{aligned}\hat{E}\Big[\int_0^T\{\tilde{\zeta}_t h_{\tilde{x}}(t)X_{1,t}+\big(\xi_t b_{\tilde{x}}(t)+\sigma_{\tilde{x}}\zeta_t\big)\Gamma_{1,t}\}E[X_{1,t}]\mathrm{d}t\Big]&=o(\epsilon),\\ \hat{E}\Big[\int_0^T\Big(\xi_t\Gamma_{1,t}\delta b(t,u_t^\epsilon)+\zeta_t\Gamma_{1,t}\delta\sigma(t,u_t^\epsilon)+\tilde{\zeta}_t X_{1,t}\delta h(t,u_t^\epsilon)\Big)\mathrm{d}t\Big]&=o(\epsilon),\\ \hat{E}\Big[\int_0^T\big[\big(\xi_{t-}+\vartheta_t\big)\gamma_{\tilde{x}}(t)+\vartheta_t c_{\tilde{x}}(t)\big]\Gamma_{1,t-}E[X_{1,t-}]\mathrm{d}V_t\Big]&=o(\epsilon).\end{aligned}$$

Then we have

$$\begin{aligned}&\hat{E}\Big[\int_0^T\Gamma_{1,t}l_x(t)X_{1,t}\mathrm{d}t+\int_0^T\Gamma_{1,t-}f_x(t)X_{1,t-}\mathrm{d}V_t+\Gamma_{1,T}g_x(T)X_{1,T}\Big]\\&=\hat{E}\Big[\int_0^T\Big(\tilde{\zeta}_t h_x(t)X_{1,t}^2-z_{2,t}h_x(t)\Gamma_{1,t}X_{1,t}\Big)\mathrm{d}t\Big]+o(\epsilon).\end{aligned}\tag{5.6}$$

Applying Itô's formula to $p_t\Delta X_t$, we have

$$\begin{aligned}&\hat{E}\big[\big(g_x(T)+\hat{E}[g_{\tilde{x}}(T)]\big)\Delta X_T\big]\\&=\hat{E}\Big[\int_0^T\Big\{p_t[\delta b(t,u_t^\epsilon)+\delta b_x(t,u_t^\epsilon)X_{1,t}]+q_t[\delta\sigma(t,u_t^\epsilon)+\delta\sigma_x(t,u_t^\epsilon)X_{1,t}]-\big(l_x(t)\\&\quad+z_{2,t}h_x(t)+\hat{E}[l_{\tilde{x}}(t)]+\hat{E}[z_{2,t}h_{\tilde{x}}(t)]\hat{Z}_t^{-1}+\mathbb{E}[f_x(t)|\mathcal{P}]r_t+\hat{E}\big[\mathbb{E}[f_{\tilde{x}}(t)|\mathcal{P}]\big]r_t\big)\Delta X_t\\&\quad+\frac{1}{2}[p_t b_{xx}(t)+q_t\sigma_{xx}(t)]X_{1,t}^2\Big\}\mathrm{d}t+\frac{1}{2}\int_0^T\{(p_{t-}+k_t)\gamma_{xx}(t)+k_t c_{xx}(t)\}X_{1,t-}^2\mathrm{d}V_t\Big].\end{aligned}$$

It is easy to verify that

$$\hat{E}\Big[\int_0^T\Big(p_t\delta b_x(t,u_t^\epsilon)+q_t\delta\sigma_x(t,u_t^\epsilon)\Big)X_{1,t}\mathrm{d}t\Big]=o(\epsilon),$$

then we have

$$\begin{aligned}&\hat{E}\Big[\int_0^T\big\{l_x(t)\Delta X_t+l_{\tilde{x}}(t)\hat{E}[\Delta X_t]\big\}\mathrm{d}t+\int_0^T\big\{f_x(t)\Delta X_{t-}+f_{\tilde{x}}(t)\hat{E}[\Delta X_{t-}]\big\}\mathrm{d}V_t+g_x(T)\Delta X_T+g_{\tilde{x}}(T)\hat{E}[\Delta X_T]\Big]\\&=\hat{E}\Big[\int_0^T\Big\{p_t\delta b(t,u_t^\epsilon)-z_{2,t}\big(h_x(t)\Delta X_t+h_{\tilde{x}}(t)E[\Delta X_t]\big)+q_t\delta\sigma(t,u_t^\epsilon)+\frac{1}{2}\big(p_tb_{xx}(t)+q_t\sigma_{xx}(t)\big)X_{1,t}^2\Big\}\mathrm{d}t\\&\quad+\frac{1}{2}\int_0^T\big\{\big(p_{t-}+k_t\big)\gamma_{xx}(t)+k_tc_{xx}(t)\big\}X_{1,t-}^2\mathrm{d}V_t\Big]+o(\epsilon).\end{aligned}\tag{5.7}$$

Applying Itô's formula to $P_tX_{1,t}^2$, we have

$$\begin{aligned}\hat{E}\big[g_{xx}(T)X_{1,T}^2\big]=\hat{E}\Big[&\int_0^T\Big\{2\big[(Q_t+P_t\sigma_x(t))X_{1,t}+P_t\sigma_{\tilde{x}}(t)E[X_{1,t}]\big]\delta\sigma(t,u_t^\epsilon)+P_t[\delta\sigma(t,u_t^\epsilon)]^2+2\big(P_tb_{\tilde{x}}+Q_t\sigma_{\tilde{x}}\\&+P_t\sigma_x(t)\sigma_{\tilde{x}}(t)\big)X_{1,t}E[X_{1,t}]-\big(p_tb_{xx}(t)+q_t\sigma_{xx}(t)+l_{xx}(t)+z_{2,t}h_{xx}(t)+2\tilde{\zeta}_th_x(t)\big)X_{1,t}^2\\&+P_t\sigma_{\tilde{x}}^2(t)\big(E[X_{1,t}]\big)^2-\big((k_t+p_t)\mathbb{E}[\gamma_{xx}(t)|\mathcal{P}]+k_t\mathbb{E}[c_{xx}(t)|\mathcal{P}]+\mathbb{E}[f_{xx}(t)|\mathcal{P}]\big)X_{1,t}^2r_t\\&+2P_tX_{1,t}\delta b(t,u_t^\epsilon)\Big\}\mathrm{d}t+\int_0^T\Big\{2[P_{t-}\varpi(t)+(c_{\tilde{x}}(t)+\varpi(t))K_t]X_{1,t-}E[X_{1,t-}]\\&+(P_{t-}+K_t)[\gamma_{\tilde{x}}(t)+c_{\tilde{x}}(t)]^2\big(\hat{E}[X_{1,t-}]\big)^2\Big\}\mathrm{d}V_t\Big],\end{aligned}$$

where $\varpi(t)=\gamma_{\tilde{x}}(t)[1+\gamma_x(t)+c_x(t)]+c_{\tilde{x}}(t)[c_x(t)+\gamma_x(t)]$. Similar to Lemma 3.3 of [5], noticing (4.20) we have

$$\begin{aligned}&\hat{E}\Big[\int_0^T\Big\{\big[P_tb_{\tilde{x}}(t)+(Q_t+P_t\sigma_x(t))\sigma_{\tilde{x}}(t)\big]X_{1,t}E[X_{1,t}]+P_t\sigma_{\tilde{x}}^2(t)\big(E[X_{1,t}]\big)^2\Big\}\mathrm{d}t\Big]=o(\epsilon),\\&\hat{E}\Big[\int_0^T\Big\{P_tX_{1,t}\delta b(t,u_t^\epsilon)+\big[(Q_t+P_t\sigma_x(t))X_{1,t}+P_t\sigma_{\tilde{x}}(t)E[X_{1,t}]\big]\delta\sigma(t,u_t^\epsilon)\Big\}\mathrm{d}t\Big]=o(\epsilon),\\&\hat{E}\Big[\int_0^T\big\{\big(P_{t-}+K_t\big)\varpi(t)+K_tc_{\tilde{x}}(t)\big\}X_{1,t-}E[X_{1,t-}]\mathrm{d}V_t\Big]=o(\epsilon),\\&\hat{E}\Big[\int_0^T\big\{(P_{t-}+K_t)(\gamma_{\tilde{x}}(t)+c_{\tilde{x}}(t))^2\big(E[X_{1,t-}]\big)^2\big\}\mathrm{d}V_t\Big]=o(\epsilon).\end{aligned}$$

Thus we have

$$\begin{aligned}&\hat{E}\Big[\int_0^Tl_{xx}(t)X_{1,t}^2\mathrm{d}t+\int_0^Tf_{xx}(t)X_{1,t-}^2\mathrm{d}V_t+g_{xx}(T)X_{1,T}^2\Big]\\&=\hat{E}\Big[\int_0^T\Big\{P_t[\delta\sigma(t,u_t^\epsilon)]^2I_{\mathcal{O}}-\big[p_tb_{xx}(t)+q_t\sigma_{xx}(t)+z_{2,t}h_{xx}(t)+2\tilde{\zeta}_th_x(t)\big]X_{1,t}^2\Big\}\mathrm{d}t\\&\quad-\int_0^T\Big\{p_{t-}\gamma_{xx}(t)X_{1,t-}^2+k_t\big(\gamma_{xx}(t)+c_{xx}(t)\big)X_{1,t-}^2\Big\}\mathrm{d}V_t\Big]+o(\epsilon).\end{aligned}\tag{5.8}$$

From (5.5)–(5.8), we have

$$\hat{J} = \hat{E}\Big[\int_0^T \Big\{p_t\delta b(t,u_t^\epsilon) + q_t\delta\sigma(t,u_t^\epsilon) + z_{2,t}\delta h(t,u_t^\epsilon) + \delta l(t,u_t^\epsilon) + \frac{1}{2}P_t[\delta\sigma(t,u_t^\epsilon)]^2\Big\}\mathrm{d}t\Big] + o(\epsilon). \tag{5.9}$$

Finally, let us define the Hamiltonian associated with random variable $X \in L^1(\Omega,\mathbb{F},P)$,

$$H(t,X,u,p,q,z_2) := pb(t,X,E[X],u) + q\sigma(t,X,E[X],u) + l(t,X,E^u[X],u) + z_2h(t,X,E[X],u).$$

Then we have the main result of this paper.

Theorem 5.1. *Let (H1)-(H2) hold. Assume that $\hat{u}_\cdot$ is the optimal control, $\hat{X}_\cdot$ is the corresponding trajectory, $(p_\cdot,q_\cdot,\tilde{q}_\cdot,k_\cdot)$ satisfies (5.3), $(z_\cdot,z_{1,\cdot},z_{2,\cdot},\varsigma_\cdot)$ satisfies (5.1) and $(P_\cdot,Q_\cdot,\tilde{Q}_\cdot,K_\cdot)$ satisfies (5.4). Then we have a.e. a.s. for any $u \in U$,*

$$\hat{E}\Big[H(t,\hat{X}_t,u,p_t,q_t,z_{2,t}) - H(t,\hat{X}_t,\hat{u}_t,p_t,q_t,z_{2,t}) + \frac{1}{2}P_t\big(\sigma(t,\hat{X}_t,E[\hat{X}_t],u) - \sigma(t,\hat{X}_t,E[\hat{X}_t],\hat{u}_t)\big)^2\Big|\mathcal{F}_t^Y\Big] \geq 0. \tag{5.10}$$

Proof. Notice that $\bigcup_{n=1}^\infty [\![T_n]\!]$ is negligible under $\hat{P}\times Leb$, thus by (5.9) we have

$$\hat{J} = \hat{E}\Big[\int_0^T \Big\{p_t\delta b(t,v) + q_t\delta\sigma(t,v) + z_{2,t}\delta h(t,v) + \delta l(t,v) + \frac{1}{2}P_t[\delta\sigma(t,v)]^2\Big\}I_{(\bar{t},\bar{t}+\epsilon]}\mathrm{d}t\Big] + o(\epsilon).$$

Then both sides are divided by ϵ and letting $\epsilon \to 0$, we have for a.e. $\bar{t}$

$$\hat{E}\Big[H(\bar{t},\hat{X}_{\bar{t}},v,p_{\bar{t}},q_{\bar{t}},z_{2,\bar{t}}) - H(\bar{t},\hat{X}_{\bar{t}},\hat{u}_{\bar{t}},p_{\bar{t}},q_{\bar{t}},z_{2,\bar{t}}) + \frac{1}{2}P_{\bar{t}}\big(\delta\sigma(\bar{t},v)\big)^2\Big] \geq 0.$$

We also have

$$\hat{E}\Big[\hat{E}\big[H(\bar{t},\hat{X}_{\bar{t}},v,p_{\bar{t}},q_{\bar{t}},z_{2,\bar{t}}) - H(\bar{t},\hat{X}_{\bar{t}},\hat{u}_{\bar{t}},p_{\bar{t}},q_{\bar{t}},z_{2,\bar{t}}) + \frac{1}{2}P_{\bar{t}}\big(\delta\sigma(\bar{t},v)\big)^2\big|\mathcal{F}_{\bar{t}}^Y\big]\Big] \geq 0.$$

Then for any $A \in \mathcal{F}_{\bar{t}}^Y$ and $u \in U$, let $v = uI_A + \hat{u}I_{A^c}$, we have

$$\hat{E}\Big[I_A\hat{E}\big[H(\bar{t},\hat{X}_{\bar{t}},u,p_{\bar{t}},q_{\bar{t}},z_{2,\bar{t}}) - H(\bar{t},\hat{X}_{\bar{t}},\hat{u}_{\bar{t}},p_{\bar{t}},q_{\bar{t}},z_{2,\bar{t}}) + \frac{1}{2}P_{\bar{t}}\big(\delta\sigma(\bar{t},u)\big)^2\big|\mathcal{F}_{\bar{t}}^Y\big]\Big] \geq 0,$$

which means a.e. a.s.

$$\hat{E}\Big[H(t,\hat{X}_t,u,p_t,q_t,z_{2,t}) - H(t,\hat{X}_t,\hat{u}_t,p_t,q_t,z_{2,t}) + \frac{1}{2}P_t\big(\sigma(t,\hat{X}_t,E[\hat{X}_t],u) - \sigma(t,\hat{X}_t,E[\hat{X}_t],\hat{u}_t)\big)^2\Big|\mathcal{F}_t^Y\Big] \geq 0.$$

□

Remark 5.2. Under some proper assumptions, our conclusion can degenerate to the preceding results.

i) If the stochastic control system is completely observed and all coefficients do not depend on the expected value of the state process. Suppose $\gamma = f = 0$, it is easy to know that the adjoint equations (5.1) and (5.3) are not necessary. And the adjoint equations (5.2) and (5.4) degenerate into (5.1) and (5.2) in Song and Wu [24], respectively. Then the SMP also degenerates into the main result (5.7) in Theorem 5.1 of [24].

ii) If all coefficients of the stochastic control system do not depend on the expected value of the state process. Suppose $\gamma = c = f = 0$, we can know that the adjoint equation (5.3) is not necessary. And the adjoint equations

(5.1), (5.2) and (5.4) degenerate into (2.5) in Li and Tang [18]. Then the SMP also degenerates into the main result (2.6) in Theorem 2.1 of [18].

6. Application to an LQ problem

In this section, we present an LQ example to illustrate that how to find the optimal control by our results. Let us consider the following partially observed stochastic control system with jumps,

$$\begin{cases} \mathrm{d}X_t = \big\{A_tX_t + \bar{A}_tE[X_t] + B_tu_t\big\}\mathrm{d}t + \big\{C_tX_t + \bar{C}_tE[X_t] + D_tu_t\big\}\mathrm{d}B_t \\ \qquad\qquad + \big\{F_tX_t + \bar{F}_tE[X_t]\big\}\mathrm{d}V_t + \big\{H_tX_t + \bar{H}_tE[X_t]\big\}\mathrm{d}\tilde{V}_t \\ X_0 = x, \end{cases} \tag{6.1}$$

where $A., \bar{A}., B., C., \bar{C}., D., F., \bar{F}., H., \bar{H}.$ are all deterministic uniformly bounded functions. And the observation process $Y.$ is given by the following SDE,

$$\begin{cases} \mathrm{d}Y_t = h_t\mathrm{d}s + \mathrm{d}W_t, \\ Y_0 = 0, \end{cases}$$

where $h.$ is a deterministic uniformly bounded function. An admissible control process $u.$ is an $\mathcal{F}_t^Y$-progressive process with values in $U \subset R$ such that $\sup_{t\in[0,T]} E[|u_t|^p] < \infty$ and $E[\int_0^T |u_t|\mathrm{d}V_t]^p < \infty$ for any $p \geq 2$, then the admissible control set can be denoted by $\mathcal{U}_{ad}$. In practice, the range of control variable is limited for some reasons such as $U = (-\infty, -1] \bigcup [1, +\infty)$, which means that in the "on" direction, the minimal value is 1, while in the "off" direction, the maximal value is -1. Now we introduce a new probability measure P^u by $\mathrm{d}P^u = Z_T\mathrm{d}P$ with $Z_t = \exp\{\int_0^t h_s\mathrm{d}Y_s - \frac{1}{2}\int_0^t |h_s|^2\mathrm{d}s\}$, which is the unique solution to SDE: $\mathrm{d}Z_t = Z_th_t\mathrm{d}Y_t$ with $Z_0 = 1$. Then the optimal control problem is to find an optimal admissible control $\hat{u}.$ to minimize

$$J(u.) = \frac{1}{2}E^u\bigg[\int_0^T R_tu_t^2\mathrm{d}t + Q(E^u[X_T])^2\bigg], \tag{6.2}$$

where $R.$ is a deterministic uniformly bounded function such that $R. \gg 0$ and Q is a non-negative constant. It is obvious that the cost functional (6.2) can be rewritten as

$$J(u.) = \frac{1}{2}E\bigg[\int_0^T Z_tR_tu_t^2\mathrm{d}t + Z_TQ(E[Z_TX_T])^2\bigg],$$

Noticing that the diffusion term of state equation (6.1) depends on control variable u and U is not convex, Theorem 5.1 is still workable in this case. Then we can obtain the maximum condition by Theorem 5.1 as follows,

$$\hat{E}\bigg[\frac{1}{2}R_t\hat{u}_t^2 + B_tp_t\hat{u}_t + D_tq_t\hat{u}_t\big|\mathcal{F}_t^Y\bigg] \leq \hat{E}\bigg[\frac{1}{2}R_tu^2 + B_tp_tu + D_tq_tu\big|\mathcal{F}_t^Y\bigg], \qquad \forall u \in U.$$

Thus we have

$$\hat{u}_t = \begin{cases} \nu_t, & \nu_t \in (-\infty, -1]\bigcup[1, +\infty), \\ 1, & 0 \leq \nu_t < 1, \\ -1, & -1 < \nu_t < 0, \end{cases} \tag{6.3}$$

where $\nu_t = -\frac{B_t \hat{E}[p_t|\mathcal{F}_t^Y] + D_t \hat{E}[q_t|\mathcal{F}_t^Y]}{R_t}$. Here $(p_\cdot, q_\cdot, \tilde{q}_\cdot, k_\cdot)$ is the solution to the following adjoint equation,

$$\begin{cases} \mathrm{d}p_t = -\big\{A_t p_t + C_t q_t + F_t r_t p_t + (F_t + H_t) r_t k_t + Z_t^{-1} \hat{E}\big[\bar{A}_t p_t + \bar{C}_t q_t \\ \qquad\qquad + (\bar{F}_t + \bar{H}_t) r_t k_t + \bar{F}_t r_t p_t\big]\big\}\mathrm{d}t + q_t \mathrm{d}B_t + \tilde{q}_t \mathrm{d}W_t + k_t \mathrm{d}\tilde{V}_t, \\ p_T = Q\hat{E}[\hat{X}_T]. \end{cases} \tag{6.4}$$

Next we would like to show that an admissible control $\hat{u}_\cdot$ defined by (6.3) is indeed an optimal control.

Proposition 6.1. *Let $\hat{u}_\cdot$ be a given control, $\hat{X}_\cdot$ is the related state and $(p_\cdot, q_\cdot, \tilde{q}_\cdot, k_\cdot)$ is the corresponding solution to the adjoint equation (6.4). Then $\hat{u}_\cdot$ defined by (6.3) is an optimal control.*

Proof. Suppose $X_\cdot$ is the trajectory of system (6.1) corresponding to any admissible control $u_\cdot \in \mathcal{U}_{ad}$. Since $Q \geq 0$, we obtain $Q(E[X_T])^2 - Q(\hat{E}[\hat{X}_T])^2 \geq 2Q\hat{E}[\hat{X}_T](E[X_T] - \hat{E}[\hat{X}_T])$. Applying Itô's formula to $p_t(X_t - \hat{X}_t)$ on the interval $[0, T]$, we have

$$Q\hat{E}[\hat{X}_T]\big(E^u[X_T] - \hat{E}[\hat{X}_T]\big) = \hat{E}\bigg[\int_0^T \big\{B_t p_t (u_t - \hat{u}_t) + D_t q_t (u_t - \hat{u}_t)\big\}\mathrm{d}t\bigg].$$

Then we obtain

$$\begin{aligned} J(u_\cdot) - J(\hat{u}_\cdot) &= \frac{1}{2}E^u\bigg[\int_0^T R_t u_t^2 \mathrm{d}t\bigg] - \frac{1}{2}\hat{E}\bigg[\int_0^T R_t \hat{u}_t^2 \mathrm{d}t\bigg] + \frac{1}{2}Q\big(E^u[X_T]\big)^2 - \frac{1}{2}Q\big(\hat{E}[\hat{X}_T]\big)^2 \\ &\geq \frac{1}{2}E\bigg[\int_0^T \big\{R_t Z_t u_t^2 - R_t Z_t \hat{u}_t^2\big\}\mathrm{d}t\bigg] + Q\hat{E}[\hat{X}_T]\big(E^u[X_T] - \hat{E}[\hat{X}_T]\big) \\ &= \hat{E}\bigg[\int_0^T \Big\{\frac{1}{2}R_t(u_t^2 - \hat{u}_t^2) + B_t p_t (u_t - \hat{u}_t) + D_t q_t (u_t - \hat{u}_t)\Big\}\mathrm{d}t\bigg] \\ &= \hat{E}\bigg[\int_0^T \hat{E}\Big[\frac{1}{2}R_t(u_t^2 - \hat{u}_t^2) + B_t p_t (u_t - \hat{u}_t) + D_t q_t (u_t - \hat{u}_t)\Big|\mathcal{F}_t^Y\Big]\mathrm{d}t\bigg] \\ &\geq 0, \qquad \forall u_\cdot \in \mathcal{U}_{ad}. \end{aligned}$$

Therefore, $\hat{u}_\cdot$ is the optimal control and $\hat{X}_\cdot$ is the optimal trajectory. □

7. Conclusion

In this paper, we derive the general SMP for partially observed mean-field type stochastic control system with Markov chain in progressive structure. The dynamic of state and observation are governed by MF-SDEs. The L^p estimate of the stochastic system with random jump in predictable structure is essentially different from that in progressive structure. Since the usual estimates in predictable structure are flawed, which may cause some problems. We introduce a special spike variation to overcome this flaw. Because of the introduction of the mean-field term, we need sharper estimates for the first order variational equations and the product term. Three first order adjoint equations and one second order adjoint equation are introduced to deal with the state process and observation. The form of our partially observed SMP with Markov chain is similar to the form of the partially observed SMP in [18] without jumps. The reason is that both SMP hold a.e. a.s.. Since the measure of all jumps' graphs of Markov chain is a negligible set under $P^u \times Leb$, jumps do not influence our result. Our partially observed SMP characterizes the behaviour of the optimal control on the area that V is continuous, it has no information about the optimal control on the jump's time. Our future research is to find a way to characterize the optimal control on the jumps with the general control domain.

References

[1] B. Acciaio, J. Backhoff-Veraguas and R. Carmona, Extended mean field control problems: stochastic maximum principle and transport perspective. *SIAM J. Control Optim.* **57** (2019) 3666–3693.

[2] D. Andersson and B. Djehiche, A maximum principle for SDEs of mean-field type. *Appl. Math. Optim.* **63** (2011) 341–356.

[3] A. Bensoussan, Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions. *Stochastics* **9** (1983) 169–222.

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

[5] R. Buckdahn, B. Djehiche and J. Li, A general stochastic maximum principle for SDEs of mean-field type. *Appl. Math. Optim.* **64** (2011) 197–216.

[6] T. Chen and Z. Wu, A general maximum principle for partially observed mean-field stochastic system with random jumps in progressive structure. *Math. Control Relat. Fields* **13** (2023) 664–694.

[7] T. Chen, Y. Song and Z. Wu, The maximum principle for stochastic control problem with Markov chain in progressive structure. *Syst. Control Lett.* **166** (2022) 105303.

[8] C. Donnelly, Sufficient stochastic maximum principle in a regime-switching diffusion model. *Appl. Math. Optim.* **64** (2011) 155–169.

[9] R.J. Elliott, T.K. Siu and A. Badescu, On pricing and hedging options in regime-switching models with feedback effect. *J. Econ. Dyn. Control* **35** (2011) 694–713.

[10] U.G. Haussmann, The maximum principle for optimal control of diffusions with partial information. *SIAM J. Control Optim.* **25** (1987) 341–361.

[11] S. He, J. Wang and J. Yan, Semimartingale Theory and Stochastic Calculus. Routledge (2019).

[12] M. Hu, Stochastic global maximum principle for optimization with recursive utilities. *Probab. Uncertain. Quantit. Risk* **2** (2017) 1–20.

[13] J. Huang, X. Li and J. Yong, A linear-quadratic optimal control problem for mean-field stochastic differential equations in infinite horizon. *Math. Control Relat. Fields* **5** (2015) 97–139.

[14] M. Huang, R.P. Malhamé and P.E. Caines, Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006) 221–252.

[15] J.M. Lasry and P.L. Lions, Mean field games. *Jap. J. Math.* **2** (2007) 229–260.

[16] R. Li and F. Fu, The maximum principle for partially observed optimal control problems of mean-field FBSDEs. *Int. J. Control* **92** (2019) 2463–2472.

[17] X. Li, J. Sun and J. Yong, Mean-field stochastic linear quadratic optimal control problems: closed-loop solvability. *Probab. Uncertain. Quantit. Risk* **1** (2016) 1–24.

[18] X. Li and S. Tang, General necessary conditions for partially observed optimal stochastic controls. *J. Appl. Probab.* **32** (1995) 1118–1137.

[19] T. Meyer-Brandis, B. Øksendal and X.Y. Zhou, A mean-field stochastic maximum principle via Malliavin calculus. *Stochastics* **84** (2012) 643–666.

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

[21] Y. Shen, Q. Meng and P. Shi, Maximum principle for mean-field jump-diffusion stochastic delay differential equations and its application to finance. *Automatica* **50** (2014) 1565–1579.

[22] Y. Shen and T.K. Siu, The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem. *Nonlinear Anal.-Theory Methods Appl.* **86** (2013) 58–73.

[23] Y. Song, S. Tang and Z. Wu, The maximum principle for progressive optimal stochastic control problems with random jumps. *SIAM J. Control Optim.* **58** (2020) 2171–2187.

[24] Y. Song and Z. Wu, A general maximum principle for progressive optimal stochastic control problems with Markov regime-switching. *ESAIM-Control Optim. Calc. Var.* **28** (2022) 61.

[25] J. Sun, H. Wang and Z. Wu, Mean-field linear-quadratic stochastic differential games. *J. Differ. Equ.* **296** (2021) 299–334.

[26] Z. Sun and O. Menoukeu-Pamen, The maximum principles for partially observed risk-sensitive optimal controls of Markov regime-switching jump-diffusion system. *Stoch. Anal. Appl.* **36** (2018) 782–811.

[27] S. Tang and X. Li, Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim.* **32** (1994) 1447–1475.

[28] R. Tao and Z. Wu, Maximum principle for optimal control problems of forward-backward regime-switching system and applications. *Syst. Control Lett.* **61** (2012) 911–917.

[29] B.C. Wang, H. Zhang and J.F. Zhang, Mean field linear-quadratic control: Uniform stabilization and social optimality. *Automatica* **121** (2020) 109088.

[30] G. Wang and Z. Wu, The maximum principles for stochastic recursive optimal control problems under partial information. *IEEE Trans. Autom. Control* **54** (2009) 1230–1242.

[31] Z. Wu, A maximum principle for partially observed optimal control of forward-backward stochastic control systems. *Sci. China-Inf. Sci.* **53** (2010) 2205–2214.

[32] Z. Wu, A general maximum principle for optimal control of forward–backward stochastic systems. *Automatica* **49** (2013) 1473–1480.

[33] J. Yong, Linear-quadratic optimal control problems for mean-field stochastic differential equations. *SIAM J. Control Optim.* **51** (2013) 2809–2838.

[34] Q. Zhang, Stock trading: an optimal selling rule. *SIAM J. Control Optim.* **40** (2001) 64–87.
[35] X. Zhang, Z. Sun and J. Xiong, A general stochastic maximum principle for a Markov regime switching jump-diffusion model of mean-field type. *SIAM J. Control Optim.* **56** (2018) 2563–2592.
[36] X.Y. Zhou, On the necessary conditions of optimal controls for stochastic partial differential equations. *SIAM J. Control Optim.* **31** (1993) 1462–1478.
[37] X.Y. Zhou and G. Yin, Markowitz's mean-variance portfolio selection with regime switching: a continuous-time model. *SIAM J. Control Optim.* **42** (2003) 1466–1482.
[38] Q. Zhou, Y. Ren and W. Wu, On optimal mean-field control problem of mean-field forward-backward stochastic system with jumps under partial information. *J. Syst. Sci. Complex.* **30** (2017) 828–856.