ESAIM: COCV 28 (2022) 61
https://doi.org/10.1051/cocv/2022054

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

A GENERAL MAXIMUM PRINCIPLE FOR PROGRESSIVE OPTIMAL STOCHASTIC CONTROL PROBLEMS WITH MARKOV REGIME-SWITCHING*

Yuanzhuo Song and Zhen Wu**

Abstract. In this paper, we give a general maximum principle for optimal controls of stochastic systems driven by Markov chains. The control is allowed to enter both diffusion and jump terms and the control domain is not necessarily convex. We apply a new spike variation and the stochastic integral of progressive processes to obtain the main result.

Mathematics Subject Classification. 60H10, 93E20.

Received February 20, 2021. Accepted August 18, 2022.

1. Introduction

Markov regime-switching models have been widely used in finance and stochastic optimal controls problems in the past few years. It modulates the system with a continuous-time finite-state Markov chain with each state representing a regime of the system or level of an economic indicator, which depends on the market mode that switches among finite number states. For example, in the stock market, the up-trend volatility of a stock tends to be smaller than its down-trend volatility, therefore, it is reasonable to describe the market trends by a two-state Markov chain. For details, the readers can refer to [13].

The maximum principle, a necessary condition for optimal control, is one of the central results in stochastic control problems. There is a very extensive literature on the stochastic maximum principles for various types of optimal control problems. Peng [7] proved the general maximum principle for the forward stochastic control system without jump by using a second-order variation equation to overcome the difficulty appearing along with the nonconvex control domain and control entering the diffusion term. Donnelly [3] studied the sufficient maximum principle in a regime-switching model. Zhang *et al.* [14] develops a sufficient stochastic maximum principle for a forward system driven by Markov regime switching and Poisson random measure. Lv and Wu [6] obtained the sufficient stochastic maximum principle of forward-backward Markov regime switching jump diffusion system. In Tao and Wu [11], a stochastic maximum principle of forward-backward system was obtained. Tang and Li [10] proved the maximum principle for the forward control system driven by Poisson random measure where the control variable is allowed into both diffusion and jump coefficients. Song *et al.* [9] fixed the deficiencies of [10] by introducing a new form of variation and allowing the control to be progressive instead

*This work was supported by the Natural Science Foundation of China (11831010, 61961160732), Shandong Provincial Natural Science Foundation (ZR2019ZD42) and the Taishan Scholars Climbing Program of Shandong (No. TSPD20210302).

Keywords and phrases: The maximum principle, Markov chain, regime-switching, spike variation.

School of Mathematics and Zhongtai Securities Institute for Financial Study, Shandong University, Jinan, China.

** Corresponding author: wuzhen@sdu.edu.cn

of predictable. As we know, Markov chains are pure jump processes that are quasi-left continuous. Under this condition, the counting process related to a Markov chain (the process V in this paper) has many similar properties to Poisson processes. When we consider stochastic integrals of continuous martingales, there is no difference between integrand being predictable or progressive. For example, we usually assume the integrand to be progressive when we consider the stochastic integral of Brownian Motion. However, when it comes to the stochastic integrals of martingales with jumps, the difference is huge. Therefore, there is a fundamental difference between "predictable" assumptions and "progressive" assumptions of controls when the system is driven by Markov chains. Zhang *et al.* [15] proved the maximum principle for the system driven by Markov chain and Poisson random measure with mean-field terms. They assumed that all admissible controls are predictable and cited some main estimates of [10]. Similar to [10], the flawed estimates may cause some problems.

In this paper, we will prove the maximum principle for the system driven by Markov chains with all admissible controls being progressive. Compared with [15], our model which overcomes the main difficulty caused by the jump term is more rigorous and subtle in mathematics. The rest of this paper is organized as follows. In Section 2, we give some preliminaries for Markov chains and introduce the stochastic integral of progressive processes with respect to $\tilde{V}$. In Sections 3–5, we give a rigorous proof of the maximum principle. To make the approach in [7] effective, we introduce a new spike variation in Section 4. Since the new variation is not necessarily predictable, our admissible control set is a set of progressive processes. That is the reason why we need the stochastic integral of progressive processes. In Section 5, we give our main result. It has a similar form to the result in [7], but indeed they are fundamentally different because the adjoint equations are different. Our result is also different from that in [15]. In the Appendix, we prove the existence and uniqueness of the solutions of SDEs and give an L^p estimate of the solutions.

2. Preliminaries

Suppose we are given a complete probability space with filtration $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, P)$. Let $\mathscr{P}$ be the predictable σ-field and $\mathscr{G}$ be the progressive σ-field.

2.1. Continuous-Time Markov Chains

We suppose that E is a finite set with n elements. Set $E = \{1, 2, \dots, n\}$ without loss of generality. The topology on E is the discrete topology. For readers' convenience, we give the following definition which is from Section 6.5 of [5].

Definition 2.1. A continuous-time Markov chain is a Markov process taking values in a finite set E with càdlàg sample paths.

Let α_t be a continuous-time Markov chain, $\mathscr{L}$ be its generator. Since the state space E is finite, it is clear that the domain of $\mathscr{L}$ is R^E. For each $i \in E$, let $S(i)$ be the first jump of α whose initial value is i, then $S(i)$ is exponentially distributed with parameter $q(i)$. In this paper, we assume that $q(i) > 0$ for each i, *i.e.*, there is no absorbing point. Now for each $i \in E$, we define a function $f^i : E \to R$, $f^i(j) = I_{\{i\}}(j)$. Clearly $f^i \in D(\mathscr{L})$ and $f^i(\alpha_t)$ has the following semimartingale decomposition,

$$\begin{aligned} f^i(\alpha_t) &= f^i(\alpha_0) + \int_0^t \sum_{j=1}^n L(\alpha_s, j) f^i(j) \mathrm{d}s + M_t^i \\ &= f^i(\alpha_0) + \int_0^t L(\alpha_s, i) \mathrm{d}s + M_t^i, \end{aligned}$$

where L is the transition intensity matrix of α, M^i is a martingale. Since for each $i \in E, q(i) > 0$, we have $L(i,i) < 0$ for each i.

If $i \neq j$, for each (ω, s), only one of $f^i(\alpha_{s-}), f^j(\alpha_{s-})$ can not be 0. Let V_t^{ij} be the counting process which counts the number of jumps from i to j up to time t, *i.e.*,

$$\begin{aligned} V_t^{ij} &= \sum_{s\le t} f^i(\alpha_{s-}) f^j(\alpha_s) = \sum_{s\le t} f^i(\alpha_{s-}) \left(f^j(\alpha_s) - f^j(\alpha_{s-})\right) = \int_0^t f^i(\alpha_{s-}) \mathrm{d} f^j(\alpha_s) \\ &= \int_0^t f^i(\alpha_s) L(\alpha_s, j) \mathrm{d}s + \int_0^t f^i(\alpha_{s-}) \mathrm{d}M_s^j. \end{aligned}$$

Define

$$V_t := \sum_{j=1}^n \sum_{i\neq j} V_t^{ij} = \int_0^t \sum_{j=1}^n \sum_{i\neq j} f^i(\alpha_s) L(\alpha_s, j) \mathrm{d}s + \sum_{j=1}^n \sum_{i\neq j} \int_0^t f^i(\alpha_{s-}) \mathrm{d}M_s^j = \int_0^t r_s \mathrm{d}s + M_t,$$

where $r_s = \sum_{j=1}^n \sum_{i\neq j} f^i(\alpha_s) L(\alpha_s, j)$, $M_t = \sum_{j=1}^n \sum_{i\neq j} \int_0^t f^i(\alpha_{s-}) \mathrm{d}M_s^j$. If $\alpha_t(\omega) = i_0$ for some $i_0 \in E$, it is easy to verify that $r_t(\omega) = q(i_0)$, *i.e.*, $r_t = q(\alpha_t)$. Therefore r_t can only have finite values. Hence r_t is bounded from above and below.

Since V is a counting process, we have

$$[V, V]_t = \sum_{s\le t} |(\Delta V)_s|^2 = \sum_{s\le t} (\Delta V)_s = V_t. \tag{2.1}$$

Since all f^i are bounded, M_t is a martingale. Then the compensator (*i.e.*, dual predictable projection) of V is $\int_0^t r_s \mathrm{d}s$. In addition,

$$\langle V, V\rangle_t = ([V, V]^p)_t = (V^p)_t = \int_0^t r_s \mathrm{d}s,$$

where "p" stands for compensator. Observing the compensator of V is continuous, we know that V is quasi-left continuous. From now, we use the new notation $\tilde{V}_t = V_t - \int_0^t r_s \mathrm{d}s$ to represent M_t. By (2.1), it is obvious that

$$[\tilde{V}, \tilde{V}]_t = [V, V]_t = V_t; \tag{2.2}$$

$$\langle \tilde{V}, \tilde{V}\rangle_t = \int_0^t r_s \mathrm{d}s. \tag{2.3}$$

2.2. Stochastic integral with progressive integrand

Fix $T > 0$, the measure υ generated by $(V_t)_{t\le T}$ is a measure on $(\Omega \times [0, T], \mathscr{F} \otimes \mathscr{B}([0, T]))$ defined in the following way,

$$\upsilon(A) := E \int_0^T I_A(\omega, t) \mathrm{d}V_t.$$

Then

$$\upsilon(\Omega \times [0, T]) := E\left[\int_0^T \mathrm{d}V_t\right] = E\left[V_T\right] < \infty,$$

which means that υ is a finite measure. For any υ-integrable process X, set $\mathbb{E}[X] := \int X\mathrm{d}\upsilon$, and $\mathbb{E}[X|\mathscr{P}]$ the R-N derivative with respect to the σ-field $\mathscr{P}$. $\mathbb{E}$ is not an expectation (for υ is not a probability measure), though it has similar properties to expectation.

Suppose that $(H_t)_{t\leq T}$ is a progressive process satisfing

$$E\left[\int_0^T |H_t|^2 \mathrm{d}V_t\right] < \infty. \tag{2.4}$$

Then the stochastic integral $\int_0^t H_s \mathrm{d}\tilde{V}_s$ is well-defined; it is a martingale. For further details, the readers can refer to ([4], no. 2, Chap. 9). Next we give some important properties of the stochastic integral without proof. Noticing that the compensator of $\int_0^t H_s \mathrm{d}V_s$ is $\int_0^t \mathbb{E}[H|\mathscr{P}]r_s \mathrm{d}s$ (the Remark after [4], Thm. 5.25), we have

Proposition 2.2. *Suppose H is a progressive process satisfing (2.4), then for each $t \in [0,T]$ we have*

$$\int_0^t H_s \mathrm{d}\tilde{V}_s = \int_0^t H_s \mathrm{d}V_s - \int_0^t \mathbb{E}[H|\mathscr{P}]r_s \mathrm{d}s. \tag{2.5}$$

Remark 2.3. By the above proposition, we can get

$$E\left[\int_0^t H_s \mathrm{d}V_s\right] = E\left[\int_0^t \mathbb{E}[H|\mathscr{P}]r_s \mathrm{d}s\right].$$

Specifically when H is predictable we have

$$E\left[\int_0^t H_s \mathrm{d}V_s\right] = E\left[\int_0^t H_s r_s \mathrm{d}s\right].$$

Since every Feller process is quasi-left continuous ([1], Thm. 5.40, Chap. 9), α_t is quasi-left continuous. Noticing that V_t has the same jump time to α_t, V_t is also quasi-left continuous. Therefore by Corollary 9.9 of [12] we have the following property:

Proposition 2.4. *Suppose H, K are progressive processes satisfing (2.4), then for each $t \in [0,T]$ we have*

$$\left[\int_0^\cdot H_s \mathrm{d}\tilde{V}_s, \int_0^\cdot K_s \mathrm{d}\tilde{V}_s\right]_t = \int_0^t H_s K_s \mathrm{d}V_s, \tag{2.6}$$

and

$$\Delta\left(\int_0^\cdot H_s \mathrm{d}\tilde{V}_s\right)_t = H_t \Delta V_t. \tag{2.7}$$

For processes with jumps, we give the following Itô's formula referring to Theorems 32, 33 of [8] or Theorem 9.35 of [4].

Lemma 2.5. *Let $X^1, X^2, \ldots, X^d$ be semimartingales, and F be a C^2 function on R^d. Set $X = (X^1, X^2, \ldots, X^d)$, then*

$$F(X_t)-F(X_0)=\sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X_{s-})dX_s^i + \frac{1}{2}\sum_{i=1,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X_{s-})d[X^i, X^j]_s + \sum_{s\leq t}\eta_s(F), \tag{2.8}$$

where

$$\eta_s(F) = F(X_s) - F(X_{s-}) - \sum_{i=1}^{d} \frac{\partial F}{\partial x_i}(X_{s-})\Delta X_s^i - \frac{1}{2}\sum_{i=1,j=1}^{d} \frac{\partial^2 F}{\partial x_i \partial x_j}(X_{s-})\Delta X_s^i \Delta X_s^j,$$

and

$$\Delta X_s^i = X_s^i - X_{s-}^i.$$

2.3. Example

Indeed, the difference between progressive controls and predictable controls may be significant. Let us consider the following stochastic control problem driven by Markov chain.

$$X_t = x_0 + \int_0^t r_s u_s \mathrm{d}s + \int_0^t u_s \mathrm{d}\tilde{V}_s,$$

with the cost functional

$$J(u) = E\left[\int_0^T r_t X_t^2 \mathrm{d}t\right],$$

and two admissible control sets

$$\mathscr{U}_1 = \left\{ u \mid u \text{ is predictable and } E\left[\int_0^T u_t^2 \mathrm{d}t\right] < \infty \right\},$$

and

$$\mathscr{U}_2 = \left\{ u \mid u \text{ is progressive and } E\left[\int_0^T u_t^2 \mathrm{d}t + \int_0^T u_t^2 \mathrm{d}V_s\right] < \infty \right\}.$$

First let us find an optimal control in $\mathscr{U}_1$ that minimizes J. Define

$$p_t = 1 - e^{-\int_t^T r_s \mathrm{d}s}.$$

For any $u \in \mathscr{U}_1$, applying Ito's lemma we have

$$\begin{aligned}
0 = p_T X_T^2 &= p_0 x_0^2 + \int_0^T 2p_t X_{t-} \mathrm{d}X_t + \int_0^T X_t^2 \mathrm{d}p_t + \int_0^T p_t \mathrm{d}[X,X]_t \\
&= p_0 x_0^2 + \int_0^T 2r_t p_t X_t u_t + r_t p_t X_t^2 - r_t X_t^2 \mathrm{d}t + \int_0^T p_t u_t^2 \mathrm{d}V_s \\
&+ \int_0^T 2p_t u_t X_{t-} \mathrm{d}\tilde{V}_s.
\end{aligned}$$

Since u is predictable, we have

$$E\left[\int_0^T r_tX_t^2\mathrm{d}t\right] = p_0x_0^2 + E\left[\int_0^T 2r_tp_tX_tu_t + r_tp_tX_t^2 + r_tp_tu_t^2\mathrm{d}t\right]$$
$$= p_0x_0^2 + E\left[\int_0^T r_tp_t\left(u_t+X_t\right)^2\mathrm{d}t\right]$$

Let $\bar{X}$ be the solution of the following equation,

$$X_t = x_0 - \int_0^t r_sX_s\mathrm{d}s - \int_0^t X_{s-}\mathrm{d}\tilde{V}_s.$$

Then $u_t^1 = -\bar{X}_{t-}$ is an optimal control in $\mathscr{U}_1$.

If $u \in \mathscr{U}_2$, we have

$$E\left[\int_0^T r_tX_t^2\mathrm{d}t\right] = p_0x_0^2 + E\left[\int_0^T 2r_tp_tX_tu_t + r_tp_tX_t^2\mathrm{d}t\right] + E\left[\int_0^T p_tu_t^2\mathrm{d}V_t\right]$$
$$= p_0x_0^2 + E\left[\int_0^T r_tp_t\left(u_t+X_t\right)^2\mathrm{d}t\right] + E\left[\int_0^T p_tu_t^2\mathrm{d}V_t\right]$$
$$- E\left[\int_0^T r_tp_tu_t^2\mathrm{d}t\right].$$

Let $\{T_n\}_{n=1}^\infty$ be the jump times of V_t, define

$$\hat{u} = \begin{cases} 0, \text{ if } (s,\omega) \in O := \bigcup_{n=1}^\infty [\![T_n]\!], \\ -X_{t-}, \text{ otherwise.} \end{cases}$$

It is obvious that $\hat{u}$ is progressive but not predictable since V is quasi-left continuous. By (2.5), the original equation turns to

$$\begin{aligned} X_t &= x_0 - \int_0^t X_s\mathrm{d}s + \int_0^t \hat{u}_s\mathrm{d}V_s + \int_0^t r_sX_s\mathbb{E}\left[I_{O^c}|\mathscr{P}\right]\mathrm{d}s \\ &= x_0 + \int_0^t b_sX_s\mathrm{d}s \end{aligned} \tag{2.9}$$

where $b = r\mathbb{E}\left[I_{O^c}|\mathscr{P}\right] - 1$ is a bounded process. So there is a unique continuous solution of (2.9), denoted by $\hat{X}$. If $x_0 \neq 0$, then $\hat{X}_t \neq 0$ which means that $p_t\hat{u}_t^2 \neq 0$. So

$$E\left[\int_0^T \hat{X}_t^2\mathrm{d}t\right] = p_0x_0^2 - E\left[\int_0^T r_tp_t\hat{u}_t^2\mathrm{d}t\right] < p_0x_0^2 = J(u^1) = \min_{u\in\mathscr{U}_1} J(u).$$

Therefore we find a control in $\mathscr{U}_2$ which is more optimal than the optimal control in $\mathscr{U}_1$. This example shows that there are not enough controls taken into consideration when the admissible controls are predictable. On the other hand, we can not ensure that $\hat{u}$ is an optimal control in $\mathscr{U}_2$ and even can not ensure the existence of optimal controls in $\mathscr{U}_2$.

3. Statement of the problem

On $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{0\le t\le T}, P)$, we are given an $\{\mathscr{F}_t\}_{0\le t\le T}$ Markov chain $(\alpha_t)_{0\le t\le T}$ and an $\{\mathscr{F}_t\}_{0\le t\le T}$ Brownian Motion $(B_t)_{0\le t\le T}$; α_t and B_t are independent. We also assume that $\{\mathscr{F}_t\}_{0\le t\le T}$ is the completion of the filtration generated by α_t and B_t. Let U be a nonempty subset of R^l. We define the admissible control set

$$U_{ad}=\Bigg\{u : [0,T]\times\Omega \to U \;|u \text{ is progressive; for any } p\ge 2,\ \sup_{0\le t\le T} E\left[|u_t|^p\right]<\infty \text{ and } E\left(\int_0^T |u_t|^2 \mathrm{d}V_t\right)^{\frac{p}{2}}<\infty\Bigg\}.$$

We consider the following progressive system with jumps:

$$X_t = x_0 + \int_0^t b(s, X_s, u_s)\mathrm{d}s + \int_0^t \sigma(s, X_s, u_s)\mathrm{d}B_s + \int_0^t c(s, X_{s-}, u_s)\mathrm{d}\tilde{V}_s \tag{3.1}$$

along with the cost functional:

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

where $b : \Omega\times[0,T]\times R^n\times R^l \to R^n$, $\sigma : \Omega\times[0,T]\times R^n\times R^l \to R^n$, $c : \Omega\times[0,T]\times R^n\times R^l \to R^n$, $f : \Omega\times[0,T]\times R^n\times R^l \to R$, $g : \Omega\times R^n \to R$. The control problem is to find an element $u\in U_{ad}$ such that

$$J(u) = \inf_{v\in U_{ad}} J(v).$$

We aim at finding necessary conditions for an optimal control in U_{ad}. We need the following assumptions.
Assumption H:

- b, σ, c are $\mathscr{G}\otimes\mathscr{B}(R^n)\otimes\mathscr{B}(R^l)/\mathscr{B}(R^n)$ measurable; f is $\mathscr{G}\otimes\mathscr{B}(R^n)\otimes\mathscr{B}(R^l)/\mathscr{B}(R)$ measurable; g is $\mathscr{F}_T\otimes\mathscr{B}(R^n)/\mathscr{B}(R)$ measurable.
- b, σ, c are twice continuously differentiable w.r.t x with bounded first and second order derivatives. In addition, there is a constant C such that

$$|(b, \sigma, c)(t, 0, u)| \le C(1 + |u|).$$

- f, g are twice continuously differentiable w.r.t x with bounded second order derivatives. In addition, there is a constant C such that

$$|f_x(t, x, u)| \le C(1 + |x| + |u|), |f(t, x, u)| \le C(1 + |x|^2 + |u|^2);$$
$$|g_x(x)| \le C(1 + |x|), |g(x)| \le C(1 + |x|^2).$$

Under these assumptions, we can show that there exists a unique solution of (3.1) for any admissible control in Appendix.

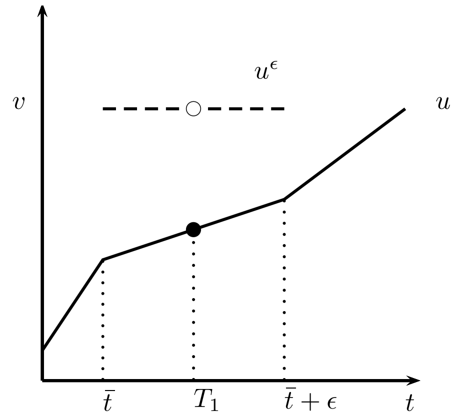

FIGURE 1. Variation.

4. VARIATION

Since U is not necessarily convex, we employ spike variations. Suppose $u \in U_{ad}$ is the optimal control. Let $\{T_n\}_{n\geq 1}$ be the jump times of V_t. For any $\bar{t} \in [0,T]$, the spike variation of u is defined by

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

where $[\![T_n]\!] := \{(\omega,t) \in \Omega \times [0,T] \mid T_n(\omega) = t\}$ is the graph of T_n, v is a bounded $\mathscr{F}_{\bar{t}}$ 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^ϵ is progressive; then it is easy to show that u^ϵ is in U_{ad}.

The method of variation is showed in Figure 1. Fix ω; we consider one path of u^ϵ and u. The difference between the new method and the traditional method is that if there are jumps in $(t, t+\epsilon]$, for example, as Figure 1 shows that if $T_1(\omega)$ is in $(t, t+\epsilon]$, then the value of u^ϵ at $T_1(\omega)$ is equal to u rather than v.

Remark 4.1. As we know, T_n is not a predictable time, so $[\![T_n]\!]$ is not predictable which means that u^ϵ is not predictable. That is the reason why we need the integrand of the stochastic integral to be progressive. In fact, T_n are totally inaccessible times.

Let X (resp. X^ϵ) be the trajectory of u (resp. u^ϵ). Since $(Leb \times P)([\![T_n]\!]) = 0$, by the estimate of SDE, we obtain

$$\begin{aligned} E\left[\sup_{0\leq t\leq T} |X_t^\epsilon - X_t|^p\right] \leq CE\Bigg[& \left(\int_0^T |b(t,X_t,u_t^\epsilon) - b(t,X_t,u_t)|\,\mathrm{d}t\right)^p \\ & + \left(\int_0^T |\sigma(t,X_t,u_t^\epsilon) - \sigma(t,X_t,u_t)|^2\,\mathrm{d}B_t\right)^{\frac{p}{2}} + \left(\int_0^T |c(t,X_{t-},u_t^\epsilon) - c(t,X_{t-},u_t)|^2\,\mathrm{d}V_t\right)^{\frac{p}{2}}\Bigg] \\ \leq & CE\left[\left(\int_t^{t+\epsilon} |b(t,X_t,u_t) - b(t,X_t,v)|\,\mathrm{d}t\right)^p + \left(\int_t^{t+\epsilon} |\sigma(t,X_t,u_t) - \sigma(t,X_t,v)|^2\,\mathrm{d}t\right)^{\frac{p}{2}}\right] \\ & + E\left[\left(\int_t^{t+\epsilon} I_{\mathcal{O}}\,|c(t,X_{t-},u_t) - c(t,X_{t-},v)|^2\,\mathrm{d}V_t\right)^{\frac{p}{2}}\right]. \end{aligned} \tag{4.2}$$

Since there is no jump on $\mathcal{O}$, we have

$$E\left[\sup_{0\le t\le T}|X_t^\epsilon - X_t|^p\right] = O(\epsilon^p) + O(\epsilon^{\frac{p}{2}}),$$

which means the jump term does not influence the order of variation. In fact, if we do not subtract the jump term in variation, $E\left[\left(\int_t^{t+\epsilon}|c(t,X_{t-},u_t)-c(t,X_{t-},v)|^2\,\mathrm{d}V_t\right)^{\frac{p}{2}}\right]$ is always of order $O(\epsilon)$ no matter how large p is. By this new variation, we can use the method in [7] to get the desired conclusion. Now we introduce the variation equations. The first equation is

$$\hat{X}_t = \int_0^t \left(b_x(s,X_s,u_s)\hat{X}_s + \delta b\right)\mathrm{d}s + \int_0^t \left(\sigma_x(s,X_s,u_s)\hat{X}_s + \delta\sigma\right)\mathrm{d}B_s + \int_0^t c_x(s,X_{s-},u_s)\hat{X}_{s-}\mathrm{d}\tilde{V}_s \tag{4.3}$$

and the second one is

$$\begin{aligned}\hat{Y}_t = &\int_0^t b_x(s,X_s,u_s)\hat{Y}_s + \frac{1}{2}b_{xx}(s,X_s,u_s)(\hat{X}_s,\hat{X}_s)\mathrm{d}s\\ &+\int_0^t \sigma_x(s,X_s,u_s)\hat{Y}_s + \frac{1}{2}\sigma_{xx}(s,X_s,u_s)(\hat{X}_s,\hat{X}_s) + \delta\sigma_x\hat{X}_s\mathrm{d}B_s\\ &+\int_0^t c_x(s,X_{s-},u_s)\hat{Y}_{s-} + \frac{1}{2}c_{xx}(s,X_{s-},u_s)(\hat{X}_{s-},\hat{X}_{s-})\mathrm{d}\tilde{V}_s,\end{aligned} \tag{4.4}$$

where $\delta\phi = \phi(s,X_s,u_s^\epsilon) - \phi(s,X_s,u_s)$; $\delta\phi_x = \phi_x(s,X_s,u_s^\epsilon) - \phi_x(s,X_s,u_s)$ for $\phi = b,\sigma$. We treat $\phi_{xx}(s,X_s,u_s)$ as a symmetric linear function from $R^n\times R^n$ to R^n.

It is easy to show that (4.3) and (4.4) have unique solutions by Theorem A.2 in the Appendix. Now we give some basic estimates w.r.t. $\hat{X}$ and $\hat{Y}$.

Lemma 4.2. *For $p\ge 2$, we have the following estimates,*

$$\begin{cases} E\left[\sup_{0\le t\le T}|\hat{X}_t|^p\right] \le C\epsilon^{\frac{p}{2}},\\ E\left[\sup_{0\le t\le T}|\hat{Y}_t|^p\right] \le C\epsilon^p.\end{cases}$$

Proof. By Theorem A.3 in the Appendix, for $\hat{X}$ we have

$$\begin{aligned} E\left[\sup_{0\le t\le T}|\hat{X}_t|^p\right] &\le CE\left[\left(\int_0^T|\delta b|\mathrm{d}t\right)^p\right] + CE\left[\left(\int_0^T|\delta\sigma|^2\mathrm{d}t\right)^{\frac{p}{2}}\right]\\ &\le CE\left[\left(\int_0^T|u_t^\epsilon - u_t|\mathrm{d}t\right)^p\right] + CE\left[\left(\int_0^T|u_t^\epsilon - u_t|^2\mathrm{d}t\right)^{\frac{p}{2}}\right] = O(\epsilon^p) + O(\epsilon^{\frac{p}{2}}).\end{aligned}$$

For $\hat{Y}$, by the boundedness of $b_{xx}, \sigma_{xx}, c_{xx}$, Lemma A.1 and Theorem A.3 in the Appendix, we have

$$\begin{aligned}
E\left[\sup_{0\le t\le T}|\hat{Y}_t|^p\right] \le & CE\left[\left(\int_0^T |\frac{1}{2}b_{xx}(s,X_s,u_s)(\hat{X}_s,\hat{X}_s)|\mathrm{d}t\right)^p\right] \\
& + CE\left[\left(\int_0^T |\frac{1}{2}\sigma_{xx}(s,X_s,u_s)(\hat{X}_s,\hat{X}_s)+\delta\sigma_x\hat{X}_s|^2\mathrm{d}t\right)^{\frac{p}{2}}\right] \\
& + CE\left[\left(\int_0^T |\frac{1}{2}c_{xx}(s,X_{s-},u_s)(\hat{X}_{s-},\hat{X}_{s-})|^2\mathrm{d}V_t\right)^{\frac{p}{2}}\right] \\
\le & CE\left[\sup_{0\le t\le T}|\hat{X}_t|^{2p}\right] + CE\left[\sup_{0\le t\le T}|\hat{X}_t|^p\left(\int_0^T|\delta\sigma_x|^2\mathrm{d}t\right)^{\frac{p}{2}}\right] \\
& + CE\left[\left(\int_0^T |\hat{X}_{s-}^2|^2\mathrm{d}V_t\right)^{\frac{p}{2}}\right] \\
= & O(\epsilon^p).
\end{aligned}$$

□

Lemma 4.3.

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}E\left[\sup_{0\le t\le T}|X_t^\epsilon - X_t - \hat{X}_t - \hat{Y}_t|^2\right] = 0. \tag{4.5}$$

Proof. First we find the equation that $X_t+\hat{X}_t+\hat{Y}_t$ satisfies.

$$\begin{aligned}
X_t+\hat{X}_t+\hat{Y}_t = & x_0 + \int_0^t \Big(b(s,X_s,u_s) + b_x(s,X_s,u_s)\hat{X}_s + b_x(s,X_s,u_s)\hat{Y}_s \\
& + \delta b + \frac{1}{2}b_{xx}(s,X_s,u_s)(\hat{X}_s,\hat{X}_s)\Big)\mathrm{d}s + \int_0^t \Big(\sigma(s,X_s,u_s) + \sigma_x(s,X_s,u_s)\hat{X}_s \\
& + \sigma_x(s,X_s,u_s)\hat{Y}_s + \delta\sigma + \delta\sigma_x\hat{X}_s + \frac{1}{2}\sigma_{xx}(s,X_s,u_s)(\hat{X}_s,\hat{X}_s)\Big)\mathrm{d}B_s \\
& + \int_0^t \Big(c(s,X_{s-},u_s) + c_x(s,X_{s-},u_s)\hat{X}_{s-} + c_x(s,X_{s-},u_s)\hat{Y}_{s-} \\
& + \frac{1}{2}c_{xx}(s,X_{s-},u_s)(\hat{X}_{s-},\hat{X}_{s-})\Big)\mathrm{d}\tilde{V}_s.
\end{aligned}$$

Since we have for $\phi = b, \sigma, c$,

$$\begin{aligned}
\phi(s,X_s+\hat{X}_s+\hat{Y}_s,u_s^\epsilon) - \phi(s,X_s,u_s) &= \phi(s,X_s+\hat{X}_s+\hat{Y}_s,u_s^\epsilon) - \phi(s,X_s,u_s^\epsilon) + \delta\phi \\
&= \delta\phi + \phi_x(s,X_s,u_s^\epsilon)(\hat{X}_s+\hat{Y}_s) + A_\phi(\hat{X}_s+\hat{Y}_s,\hat{X}_s+\hat{Y}_s),
\end{aligned}$$

where

$$A_\phi = \int_0^1\int_0^1 \alpha\phi_{xx}(X_s+\alpha\beta(\hat{X}_s+\hat{Y}_s),u_s^\epsilon)\mathrm{d}\alpha\mathrm{d}\beta$$

is a symmetric linear function from $R^n \times R^n \to R^n$. Then we get

$$\begin{aligned} X_t + \hat{X}_t + \hat{Y}_t = x_0 &+ \int_0^t \Big(b(s, X_s + \hat{X}_s + \hat{Y}_s, u_s^\epsilon) + \Lambda \Big) \mathrm{d}s \\ &+ \int_0^t \Big(\sigma(s, X_s + \hat{X}_s + \hat{Y}_s, u_s^\epsilon) + G \Big) \mathrm{d}B_s \\ &+ \int_0^t \Big(c(s, X_{s-} + \hat{X}_{s-} + \hat{Y}_{s-}, u_s^\epsilon) + F \Big) \mathrm{d}\tilde{V}_s, \end{aligned}$$

where

$$\begin{aligned} \Lambda &= \frac{1}{2} b_{xx}(s, X_s, u_s)(\hat{X}_s, \hat{X}_s) - (b_x(s, X_s, u_s^\epsilon) - b_x(s, X_s, u_s))(\hat{X}_s + \hat{Y}_s) - A_b(\hat{X}_s + \hat{Y}_s, \hat{X}_s + \hat{Y}_s), \\ G &= \frac{1}{2} \sigma_{xx}(s, X_s, u_s)(\hat{X}_s, \hat{X}_s) - (\sigma_x(s, X_s, u_s^\epsilon) - \sigma_x(s, X_s, u_s))\hat{Y}_s - A_\sigma(\hat{X}_s + \hat{Y}_s, \hat{X}_s + \hat{Y}_s); \end{aligned}$$

$$F = \frac{1}{2} c_{xx}(s, X_{s-}, u_s)(\hat{X}_{s-}, \hat{X}_{s-}) - A_c(\hat{X}_{s-} + \hat{Y}_{s-}, \hat{X}_{s-} + \hat{Y}_{s-}).$$

By Lemma 4.2, we have

$$E\left[\left(\int_0^T \Lambda \mathrm{d}s \right)^2 + \int_0^T G^2 \mathrm{d}s + \int_0^T F^2 \mathrm{d}V_t \right] = o(\epsilon^2).$$

Then by Theorem A.3, we obtain

$$\begin{aligned} E\left[\sup_{0 \le t \le T} |X_t^\epsilon - X_t - \hat{X}_t - \hat{Y}_t|^2 \right] \le & CE\left[\left(\int_0^T \Lambda \mathrm{d}s \right)^2 + \int_0^T G^2 \mathrm{d}s \right] \\ &+ CE\left[\int_0^T F^2 \mathrm{d}V_t \right], \end{aligned}$$

which shows the result. □

Now we get the variation equation for cost functional. We have

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

Define

$$\begin{aligned} \hat{J} = & E\left[\int_0^T f_x(t, X_t, u_t)(\hat{X}_t + \hat{Y}_t) + \frac{1}{2} f_{xx}(t, X_t, u_t)(\hat{X}_t, \hat{X}_t) + \delta f \mathrm{d}t \right] \\ &+ E\left[g_x(X_T)(\hat{X}_T + \hat{Y}_T) + \frac{1}{2} g_{xx}(X_T)(\hat{X}_T, \hat{X}_T) \right], \end{aligned} \tag{4.6}$$

where $\delta f = f(s, X_s, u^\epsilon_s) - \phi(s, X_s, u_s)$. We treat f_{xx}, g_{xx} as a symmetric linear function from $R^n \times R^n$ to R here. Indeed, f_{xx}, g_{xx} can be also treated as a matrix. In this case we have

$$f_{xx}(t, X_t, u_t)(\hat{X}_t, \hat{X}_t) = \hat{X}_t^\top f_{xx}(t, X_t, u_t)\hat{X}_t;$$
$$g_{xx}(X_T)(\hat{X}_T, \hat{X}_T) = \hat{X}_T^\top g_{xx}(X_T)\hat{X}_T.$$

Then we have the following lemma.

Lemma 4.4.

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

Proof.

$$\begin{aligned} J(u) + \hat{J} =& E\left[\int_0^T f(t, X_t, u_t) + f_x(t, X_t, u_t)(\hat{X}_t + \hat{Y}_t) + \frac{1}{2} f_{xx}(t, X_t, u_t)(\hat{X}_t, \hat{X}_t) + \delta f \mathrm{d}t\right] \\ &+ E\left[g(X_T) + g_x(X_T)(\hat{X}_T + \hat{Y}_T) + \frac{1}{2} g_{xx}(X_T)(\hat{X}_T, \hat{X}_T)\right] \\ =& E\left[\int_0^T f(t, X_t + \hat{X}_t + \hat{Y}_t, u^\epsilon_t) + H\mathrm{d}t\right] + E\left[g(X_T + \hat{X}_T + \hat{Y}_T) + I\right], \end{aligned}$$

where

$$\begin{aligned} H &= \frac{1}{2} f_{xx}(s, X_s, u_s)(\hat{X}_s, \hat{X}_s) - \delta f_x(\hat{X}_s + \hat{Y}_s) - A_f(\hat{X}_s + \hat{Y}_s, \hat{X}_s + \hat{Y}_s); \\ I &= -\left(\int_0^1 \int_0^1 \alpha g(X_T + \alpha\beta(\hat{X}_T + \hat{Y}_T))\mathrm{d}\alpha \mathrm{d}\beta\right)(\hat{X}_T + \hat{Y}_T, \hat{X}_T + \hat{Y}_T) + \frac{1}{2} g_{xx}(X_T)(\hat{X}_T, \hat{X}_T). \end{aligned}$$

Then

$$\begin{aligned} |J(u^\epsilon) - J(u) - \hat{J}|^2 \leq& CE\left[\int_0^T |f(t, X_t + \hat{X}_t + \hat{Y}_t, u^\epsilon_t) - f(t, X^\epsilon_t, u^\epsilon_t)|^2 \mathrm{d}t + \left(\int_0^T H\mathrm{d}t\right)^2\right] \\ &+ E\left[\left|g(X_T + \hat{X}_T + \hat{Y}_T) - g(X^\epsilon_T)\right|^2 + I^2\right] \\ \leq& CE\left[\sup_{0\leq t\leq T} |X^\epsilon_t - X_t - \hat{X}_t - \hat{Y}_t|^2\right] + E\left[\left(\int_0^T H\mathrm{d}t\right)^2 + I^2\right] \\ =& o(\epsilon^2). \end{aligned}$$

By the same method we can show that $E\left[\left(\int_0^T H\mathrm{d}t\right)^2 + I^2\right] = o(\epsilon^2)$, which proves the result. ☐

5. Adjoint equations and the maximum principle

We introduce the first order and second order adjoint equations.

The first order equation is

$$p_t = g_x(X_T) + \int_t^T \left(b_x p_s + \sigma_x q_s + f_x^\top + r_s \mathbb{E}[c_x|\mathscr{P}]k_s\right) \mathrm{d}s - \int_t^T q_s \mathrm{d}B_s - \int_t^T k_s \tilde{V}_s. \tag{5.1}$$

And the second order equation is

$$\begin{aligned} P_t = g_{xx}(X_T) + \int_t^T \Big(& b_x P_s + P_s b_x + \sigma_x Q_s + Q_s \sigma_x + f_{xx} + b_{xx}(p_s) + \sigma_{xx}(q_s) + \sigma_x P_s \sigma_x \\ & + r_s \mathbb{E}[c_x|\mathscr{P}]K_s + r_s K_s \mathbb{E}[c_x|\mathscr{P}] + r_s \mathbb{E}[c_x|\mathscr{P}](P_s + K_s)\mathbb{E}[c_x|\mathscr{P}] + r_s \mathbb{E}[c_{xx}|\mathscr{P}](k_s)\Big) \mathrm{d}s \\ & - \int_t^T Q_s \mathrm{d}B_s - \int_t^T K_s \mathrm{d}\tilde{V}_s, \end{aligned} \tag{5.2}$$

where $\phi_x = \phi_x(t, X_t, u_t), \phi_{xx} = \phi_{xx}(t, X_t, u_t)$ if $\phi = b, \sigma, f$; $\phi_x = \phi_x(t, X_{t-}, u_t), \phi_{xx} = \phi_{xx}(t, X_{t-}, u_t)$ if $\phi = c$. Here we treat $f_{xx}, g_{xx}(X_T)$ as a matrix; we treat $b_{xx}, \sigma_{xx}, c_{xx}$ as a linear function from R^n to $R^{n\times n}$.

For the existence and uniqueness of the two BSDEs above, we refer to [2]. Since ϕ_x, ϕ_{xx} are bounded, there exists a unique solution of (5.1) $(p, q, k) \in S^2[0, T] \times M^2[0, T] \times F^2[0, T]$ and a unique solution of (5.2) $(P, Q, K) \in S^2[0, T] \times M^2[0, T] \times F^2[0, T]$, where

$$S^2[0, T] := \left\{ Y \mid Y \text{ has càdlàg paths, adapted and } E\left[\sup_{0\le t\le T} |Y_t|^2\right] < \infty. \right\}$$

with norm $\|Y\|^2 = E\left[\sup_{0\le t\le T} |Y_t|^2\right]$,

$$M^2[0, T] = \left\{ Z \mid Z \text{ is predictable and } E\left[\int_0^T |Z_s|^2 \mathrm{d}s\right] < \infty. \right\}$$

with norm $\|Z\|^2 = E\left[\int_0^T |Z_s|^2 \mathrm{d}s\right]$, and

$$F^2[0, T] = \left\{ K \mid K \text{ is predictable and } E\left[\int_0^T |K_t|^2 r_t \mathrm{d}t\right] < \infty. \right\}$$

with norm $\|K\|^2 = E\left[\int_0^T |K_t|^2 r_t \mathrm{d}t\right]$.

Applying Itô's formula (2.8) for $\langle p_t, \hat{X}_t\rangle, \langle p_t, \hat{Y}_t\rangle$; and $\langle P_t \hat{X}_t, \hat{X}_t\rangle$, using (2.6), (2.7), we get

$$\begin{aligned} E\left[\left\langle p_T, \hat{X}_T\right\rangle\right] &= E\int_0^T \left\langle p_{t-}, d\hat{X}_t\right\rangle + E\int_0^T \left\langle \hat{X}_{t-}, \mathrm{d}p_t\right\rangle + E\left[\mathrm{Tr}([p, \hat{X}]_T)\right] \\ &= E\int_0^T \left(\delta b^\top p_t + \delta\sigma^\top q_t - f_x^\top \hat{X}_t\right) \mathrm{d}t, \end{aligned} \tag{5.3}$$

and

$$
\begin{aligned}
E\left[\left\langle p_T, \hat{Y}_T \right\rangle\right] =& E\int_0^T \left\langle p_{t-}, \mathrm{d}\hat{Y}_t \right\rangle + E\int_0^T \left\langle \hat{Y}_{t-}, \mathrm{d}p_t \right\rangle + E\left[\mathrm{Tr}([p,\hat{Y}]_T)\right] \\
=& E\int_0^T \left(\frac{1}{2}\left\langle p_t, b_{xx}(\hat{X}_t,\hat{X}_t)\right\rangle + \frac{1}{2}\left\langle q_t, \sigma_{xx}(\hat{X}_t,\hat{X}_t)\right\rangle - f_x^\top \hat{Y}_t + \left\langle q_t, \delta\sigma_x \hat{X}_t\right\rangle \right. \\
&\left. + \frac{1}{2} r_t \left\langle k_t, \mathbb{E}[c_{xx}|\mathscr{P}](\hat{X}_t,\hat{X}_t)\right\rangle \right)\mathrm{d}t,
\end{aligned}
\tag{5.4}
$$

and

$$
\begin{aligned}
E\left[\left\langle P_T\hat{X}_T, \hat{X}_T \right\rangle\right] =& E\left[\int_0^T \left\langle \hat{X}_{t-}\hat{X}_{t-}^\top, dP_t\right\rangle_{R^{n\times n}} + \int_0^T 2\left\langle P_{t-}\hat{X}_{t-}, d\hat{X}_t\right\rangle_{R^n}\right] \\
& + E\left[\int_0^T \left\langle P_{t-}, d[\hat{X},\hat{X}]_t\right\rangle_{R^{n\times n}} + \int_0^T 2\left\langle \hat{X}_{t-}, dL_t\right\rangle_{R^n} + \sum_{t\le T} \Delta\hat{X}_t^\top \Delta P_t \Delta \hat{X}_t\right] \\
=& E\int_0^T \delta\sigma^\top P_t \delta\sigma - \left(f_{xx}(\hat{X}_t,\hat{X}_t) + \left\langle p_t, b_{xx}(\hat{X}_t,\hat{X}_t)\right\rangle + \left\langle q_t, \sigma_{xx}(\hat{X}_t,\hat{X}_t)\right\rangle \right. \\
& \left. + r_t\left\langle k_t, \mathbb{E}[c_{xx}|\mathscr{P}](\hat{X}_t,\hat{X}_t)\right\rangle\right)\mathrm{d}t \\
& + E\int_0^T \left(2\delta b^\top P_t\hat{X}_t + 2\delta\sigma^\top Q_t\hat{X}_t + 2\delta\sigma^\top P_t\sigma_x\hat{X}_t\right)\mathrm{d}t,
\end{aligned}
\tag{5.5}
$$

where L is a $n\times 1$ matrix such that $L_t^i = \sum_{j=1}^n [\hat{X}^i, P^{ij}]_t$. In (5.5), we use the fact

$$
\begin{aligned}
\sum_{t\le T} \Delta\hat{X}_t^\top \Delta P_t \Delta\hat{X}_t &= \sum_{t\le T}\left(\hat{X}_{t-}^\top c_x^\top \Delta V_t\right)(K_t\Delta V_t)\left(c_x\hat{X}_{t-}\Delta V_t\right) \\
&= \sum_{t\le T} \hat{X}_{t-}^\top c_x^\top K_t c_x \hat{X}_{t-}\Delta V_t \\
&= \int_0^T \hat{X}_{t-}^\top c_x^\top K_t c_x \hat{X}_{t-}\mathrm{d}V_t.
\end{aligned}
$$

The second equality follows from the fact that $\Delta V_t = 1$ or 0.

From (5.3)(5.4)(5.5), we can get the form of $g_x(X_T)(X_T+Y_T)$ since $g_{xx}(X_T)(X_T,X_T) = \left\langle P_T\hat{X}_T, \hat{X}_T\right\rangle$. Then we have

$$
\hat{J} = E\left[\int_0^T \left(\delta b^\top p_t + \delta\sigma^\top q_t + \delta f + \frac{1}{2}\delta\sigma^\top P_t\delta\sigma\right)\mathrm{d}t\right] + o(\epsilon),
\tag{5.6}
$$

where $o(\epsilon)$ represents $E\left[\int_0^T \left(\left\langle q_t, \delta\sigma_x\hat{X}_t\right\rangle + \delta\sigma^\top P_t\sigma_x\hat{X}_t + \delta b^\top P_t\hat{X}_t + \delta\sigma^\top Q_t\hat{X}_t\right)\mathrm{d}t\right]$.

We set $H(t,x,u,p,q) := b^\top(t,x,u)p + \sigma^\top(t,x,u)q + f(t,x,u)$. Then we give our main result.

Theorem 5.1. *Suppose Assumption H is satisfied. Let u be the optimal control, and X is the trajectory of u. If (p, q) satisfies (5.1); P satisfies (5.2), then for any $w \in U$, we have the following inequality a.e. a.s.,*

$$H(t, X_t, w, p_t, q_t) - H(t, X_t, u_t, p_t, q_t) + \frac{1}{2}\left(\sigma^\top(t, X_t, w) - \sigma^\top(t, X_t, u_t)\right) P_t\left(\sigma(t, X_t, w) - \sigma(t, X_t, u_t)\right) \geq 0. \tag{5.7}$$

Proof. Noticing that $\bigcup_{n=1}^{\infty} [\![T_n]\!]$ is negligible under $P \times Leb$, by (5.6) we can obtain

$$\begin{aligned}\hat{J} =& E\Bigg[\int_0^T I_{(\bar{t},\bar{t}+\epsilon]}\Bigg(\left(b^\top(t, X_t, v) - b^\top(t, X_t, u)\right) p_t + \left(\sigma^\top(t, X_t, v) - \sigma^\top(t, X_t, u)\right) q_t \\ &+ f(t, X_t, v) - f(t, X_t, u) + \frac{1}{2}\left(\sigma^\top(t, X_t, v) - \sigma^\top(t, X_t, u_t)\right) P_t\left(\sigma(t, X_t, v) - \sigma(t, X_t, u)\right)\Bigg)\mathrm{d}t\Bigg] \\ &+ o(\epsilon).\end{aligned}$$

If we divide both sides by ϵ and let ϵ tend to 0, we obtain for a.e. $\bar{t}$

$$E\left[H(\bar{t}, X_{\bar{t}}, v, p_{\bar{t}}, q_{\bar{t}}) - H(\bar{t}, X_{\bar{t}}, u, p_{\bar{t}}, q_{\bar{t}}) + \frac{1}{2}\left(\sigma^\top(\bar{t}, X_{\bar{t}}, v) - \sigma^\top(\bar{t}, X_{\bar{t}}, u)\right) P_{\bar{t}}(\sigma(\bar{t}, X_{\bar{t}}, v) - \sigma(\bar{t}, X_{\bar{t}}, u))\right] \geq 0.$$

Then for any $A \in \mathscr{F}_{\bar{t}}$ and $w \in U$, letting $v = wI_A + uI_{A^c}$, we have

$$\begin{aligned}E\Bigg[I_A\Bigg(&H(\bar{t}, X_{\bar{t}}, w, p_{\bar{t}}, q_{\bar{t}}) - H(\bar{t}, X_{\bar{t}}, u, p_{\bar{t}}, q_{\bar{t}}) \\ &+ \frac{1}{2}\left(\sigma^\top(\bar{t}, X_{\bar{t}}, u) - \sigma^\top(\bar{t}, X_{\bar{t}}, w)\right) P_{\bar{t}}(\sigma(\bar{t}, X_{\bar{t}}, w) - \sigma(\bar{t}, X_{\bar{t}}, u))\Bigg)\Bigg] \geq 0,\end{aligned}$$

which means a.e. a.s.

$$H(\bar{t}, X_{\bar{t}}, w, p_{\bar{t}}, q_{\bar{t}}) - H(\bar{t}, X_{\bar{t}}, u, p_{\bar{t}}, q_{\bar{t}}) + \frac{1}{2}(\sigma^\top(\bar{t}, X_{\bar{t}}, w) - \sigma^\top(\bar{t}, X_{\bar{t}}, u))P_{\bar{t}}(\sigma(\bar{t}, X_{\bar{t}}, w) - \sigma(\bar{t}, X_{\bar{t}}, u)) \geq 0.$$

□

Remark 5.2. Compared with the results in [15], (5.7) is independent of the jump term c. In other words, (5.7) only describes the optimal control in the area where V does not jumps. The behavior of the optimal control on the jump times $\bigcup_{n=1}^{\infty} [\![T_n]\!]$ remains unsolved. In our progressive model, we need to give characterizations of the optimal control not only on the continuous part $\left(\bigcup_{n=1}^{\infty} [\![T_n]\!]\right)^c$ but also on the jump part $\bigcup_{n=1}^{\infty} [\![T_n]\!]$, while in the predictable model of [15] only "continuous part" need to be obtained. This because the performance of predictable processes on jump part is similar to that on continuous part. Therefore, our model gives a more comprehensive and detailed description of the optimal control.

6. Conclusion

In this paper, we give the Maximum Principle of systems driven by $\tilde{V}$, a martingale depends on a Markov chain. To apply the approach in [7], the key issue is the order of the last term of (4.2). The same issue appears in [9]. To fix this problem, we introduce the spike variation (4.1) to make the last term of (4.2) become 0.

However, another problem follows. The new control u^ϵ is not predictable since $\tilde{V}$ is quasi-left continuous. We deal with this problem by introducing the stochastic integral with respect to progressive processes. Compared with our previous work [9], we apply a similar method in this paper. This is because Poisson processes and V are both counting processes and quasi-left continuous. The quasi-left continuity ensures that the stochastic integral with respect to progressive processes makes sense. They are counting processes to make sure their respective quadratic variation processes are themselves, which is an important property. However, the systems driven by Markov chains are more widely used than the systems driven by Poisson random measures. Also because of their similarity, the flawed estimate in [10] may appear in the estimate of systems driven by $\tilde{V}$. Specifically, if we use the predictable quadratic variation (2.3) instead of quadratic variation (2.2) when applying the BDG inequality, the issue in (4.2) will not exist. Then the approach in [7] seems to be right.

The Maximum Principle in this paper is obtained in a clear and concise framework and lays a solid foundation for further related theoretical and application research. Our future work is to give a characterization of the optimal control on the jump times of V.

Appendix A. The existence and uniqueness of the solutions

We are given the following SDE

$$X_t = x_0 + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_s + \int_0^t c(s, X_{s-})\mathrm{d}\tilde{V}_s, \tag{A.1}$$

where $x_0 \in R^n$, $b, \sigma, c : \Omega \times [0, T] \times R^n \to R^n$, n is the dimension of X. We introduce a Banach space

$$S^p[0, T] := \left\{ X \mid X \text{ has càdlàg paths and adapted and } E\left[\sup_{0\le t\le T} |X_t|^p\right] < \infty. \right\}$$

with norm $\|X\|_p^p = E\left[\sup_{0\le t\le T} |X_t|^p\right]$.

Lemma A.1. *Suppose that X_t is an adapted process with càdlàg paths, then for any $p > 0$ we have*

$$E\left(\int_0^T |X_{s-}|\mathrm{d}V_s\right)^p \le Ce^{CT}TE\left[\sup_{0\le s\le T} |X_s|^p\right],$$

where C is a constant only depending on p.

Proof. We can suppose $E\left[\sup_{0\le s\le T} |X_s|^p\right] < \infty$, otherwise the conclusion is obvious. Set $A_t = \int_0^t |X_{s-}|\mathrm{d}V_s$. Since V_t is a counting process, A_t is a pure jump process. Notice that the jump time of A_t is also a jump time of V_t and the jump size of V_t is always equal to 1, so we have

$$\begin{aligned}
A_t^p &= \sum_{s\le t} A_s^p - A_{s-}^p = \sum_{s\le t} \left(A_s^p - A_{s-}^p\right) I_{\{\Delta V_s \ne 0\}} \\
&= \sum_{s\le t} \left((A_{s-} + |X_{s-}|)^p - A_{s-}^p\right) \Delta V_s \\
&= \int_0^t (A_{s-} + |X_{s-}|)^p - A_{s-}^p \mathrm{d}V_s \\
&\le C\int_0^t A_{s-}^p + |X_{s-}|^p \mathrm{d}V_s.
\end{aligned}$$

For any $k \geq 1$, since $A_{\cdot -}$, $X_{\cdot -}$ and $I_{[0,T_k]}$ are predictable, we have

$$E\left[A^p_{t\wedge T_k}\right] \leq E\left[\int_0^t \left(A^p_{s\wedge T_k} + |X_{s\wedge T_k}|^p\right) r_s \mathrm{d}s\right] \leq C\int_0^t E\left[A^p_{s\wedge T_k}\right]\mathrm{d}s + CTE\left[\sup_{0\leq t\leq T}|X_t|^p\right].$$

Since $E\left[A^p_{s\wedge T_k}\right] \leq kE\left[\sup_{0\leq t\leq T}|X_t|^p\right]$, by Gronwall's inequality we have

$$E\left[A^p_{t\wedge T_k}\right] \leq CTE\left[\sup_{0\leq t\leq T}|X_t|^p\right]e^{Ct} \leq Ce^{CT}TE\left[\sup_{0\leq t\leq T}|X_t|^p\right].$$

Let k tend to infinity, by Fatou's lemma we obtain

$$E\left[A^p_t\right] \leq \varliminf_{k\to\infty} E\left[A^p_{t\wedge T_k}\right] \leq Ce^{CT}TE\left[\sup_{0\leq t\leq T}|X_t|^p\right].$$

□

We have the following assumptions:
Assumption H1:

(i) b, σ, c are $\mathscr{G}\otimes\mathscr{B}(R^n)/\mathscr{B}(R^n)$ measurable.
(ii) b, σ, c are uniform lipschitz continuous with respect to x.
(iii) $E\left[\left(\int_0^T |b(\omega,t,0)|\mathrm{d}t\right)^p\right] < \infty, E\left[\left(\int_0^T |\sigma(\omega,t,0)|^2\mathrm{d}t\right)^{\frac{p}{2}}\right] < \infty,$
$E\left[\left(\int_0^T |c(\omega,t,0)|^2\mathrm{d}V_t\right)^{\frac{p}{2}}\right] < \infty.$

Theorem A.2. *Under Assumption H1, (A.1) has a unique solution in* $S^p[0,T]$.

Proof. Firstly, we show that there is a unique solution in small time duration. We construct a map from $S^p[0,T]$ to $S^p[0,T]$

$$\mathscr{T}(X)_t = x_0 + \int_0^t b(s,X_s)\mathrm{d}s + \int_0^t \sigma(s,X_s)\mathrm{d}B_s + \int_0^t c(s,X_{s-})\mathrm{d}\tilde{V}_s.$$

It is easy to show that the image of $\mathscr{T}$ is in $S^p[0,T]$ by (ii) and (iii) in Assumption H1 and Lemma A.1. For any $X, Y \in S^p[0,T]$, by Lemma A.1

$$\begin{aligned}\|\mathscr{T}(X)-\mathscr{T}(Y)\|_p^p \leq & CE\left[\left(\int_0^T |b(t,X_t)-b(t,Y_t)|\mathrm{d}t\right)^p\right] + CE\left[\left(\int_0^T |\sigma(t,X_t)-\sigma(t,Y_t)|^2\mathrm{d}t\right)^{\frac{p}{2}}\right] \\ & + CE\left[\left(\int_0^T |c(t,X_{t-})-c(t,Y_{t-})|^2\mathrm{d}V_t\right)^{\frac{p}{2}}\right] \\ \leq & C\left(T^p + T^{\frac{p}{2}} + Te^{CT}\right)\|X-Y\|_p^p,\end{aligned} \tag{A.2}$$

where C is a constant only depend on p. Choose T small enough that

$$C\left(T^p + T^{\frac{p}{2}} + e^{CT}T\right) < 1,$$

then $\mathscr{T}$ is a contraction.

For arbitrary T, we can split T into finite small pieces, so that we get a unique solution on each piece and connect them. □

Now we give the L^p estimate of the solution.

Theorem A.3. *Suppose that we are given two SDEs:*

$$X_t^i = x_0^i + \int_0^t b^i(s, X_s^i)\mathrm{d}s + \int_0^t \sigma^i(s, X_s^i)\mathrm{d}B_s + \int_0^t c^i(s, X_{s-}^i)\tilde{V}_s, \tag{A.3}$$

where $i = 1, 2$. Suppose the two equations satisfy Assumption H1, and X^1 (resp. X^2) is the solution of the first (resp. second) equation, then we have the following estimate,

$$\begin{aligned} E\left[\sup_{0\le t\le T}|X_t^1 - X_t^2|^p\right] \le & C|x_0^1 - x_0^2|^p + CE\left(\int_0^T \left|b^1(t, X_t^2) - b^2(t, X_t^2)\right| \mathrm{d}t\right)^p \\ & + CE\left(\int_0^T \left|\sigma^1(t, X_t^2) - \sigma^2(t, X_t^2)\right|^2 \mathrm{d}t\right)^{\frac{p}{2}} + CE\left(\int_0^T |c^1(t, X_{t-}^2) - c^2(t, X_{t-}^2)|^2 \mathrm{d}V_t\right)^{\frac{p}{2}}. \end{aligned} \tag{A.4}$$

Proof. We first suppose that T is sufficiently small. $\mathscr{T}^i$ is the contraction mapping with respect to the i'th equation in Theorem A.2. For simplicity, we set $L(T) := C\left(T + T^2 + e^{CT}T\right)$, which is the coefficient in (A.2) of $\mathscr{T}^1$; then by the argument in Theorem A.2 we have

$$\begin{aligned} \|X^1 - X^2\|_p^p &= \|\mathscr{T}^1(X^1) - \mathscr{T}^2(X^2)\|_p^p \\ &\le 2^{p-1}\left(\|\mathscr{T}^1(X^1) - \mathscr{T}^1(X^2)\|_p^p + \|\mathscr{T}^1(X^2) - \mathscr{T}^2(X^2)\|_p^p\right) \\ &\le 2^{p-1}L(T)\|X^1 - X^2\|_p^p + \|\mathscr{T}^1(X^2) - \mathscr{T}^2(X^2)\|_p^p. \end{aligned}$$

Choosing T small enough such that $2^{p-1}L(T) < 1$, we have

$$\|X^1 - X^2\|_p^p \le \frac{1}{1 - 2^{p-1}L(T)}\|\mathscr{T}^1(X^2) - \mathscr{T}^2(X^2)\|_p^p,$$

which is the desired estimate. □

References

[1] E. Çınlar, vol. 261 of *Probability and stochastics.* Springer, New York (2011).

[2] S.N. Cohen and R.J. Elliott, Solutions of backward stochastic differential equations on Markov chains. *Commun. Stoch. Anal.* **2** (2008) 251–262.

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

[4] S. He, J. Wang and J. Yan, Semimartingale Theory and Stochastic Calculus. Routledge, Abingdon, UK (2018).

[5] J.-F. Le Gall, Vol. 274 of *Brownian motion, Martingales, and Stochastic Calculus.* Springer, New York (2016).

[6] S. Lv and Z. Wu, Stochastic maximum principle for forward-backward regime switching jump diffusion systems and applications to finance. *Chin. Ann. Math. Ser. B* **39** (2018) 773–790.

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

[8] P.E. Protter, Stochastic Integration and Differential Equations. Springer, Heidelberg (2005).

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

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

[11] 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.

[12] J. Yan, Introduction to Martingales and Stochastic Integrals. Shanghai Sci. and Tech. Publ. House, Shanghai (1981).

[13] Q. Zhang, Stock trading: An optimal selling rule. *SIAM J. Control. Optim.* **40** (2001) 64–87.

[14] X. Zhang, R.J. Elliott and T.K. Siu, A stochastic maximum principle for a Markov regime-switching jump-diffusion model and its application to finance. *SIAM J. Control Optim.* **50** (2012) 964–990.

[15] 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.