J Syst Sci Complex (2017) 30: 316–328

Maximum Principle for Partially-Observed Optimal Control Problems of Stochastic Delay Systems

WU Shuang · SHU Lan

DOI: 10.1007/s11424-016-5078-4
Received: 30 March 2015 / Revised: 14 October 2015

Abstract This paper is concerned with partially-observed optimal control problems for stochastic delay systems. Combining Girsanov's theorem with a standard variational technique, the authors obtain a maximum principle on the assumption that the system equation contains time delay and the control domain is convex. The related adjoint processes are characterized as solutions to anticipated backward stochastic differential equations in finite-dimensional spaces. Then, the proposed theoretical result is applied to study partially-observed linear-quadratic optimal control problem for stochastic delay system and an explicit observable control variable is given.

Keywords Anticipated backward stochastic differential equation, maximum principle, partially-observed optimal control, stochastic delay systems.

1 Introduction and Problem Formulation

Throughout this article, we denote by $\mathbb{R}^n$ the n-dimensional Euclidean space, $\mathbb{R}^{n\times d}$ the collection of $n \times d$ matrices. For a given Euclidean space, we denote by $\langle \cdot, \cdot \rangle$ (resp. $|\cdot|$) the scalar product (resp. norm). The superscript τ denotes the transpose of vectors or matrices.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space equipped with a natural filtration

$$\mathcal{F}_t = \sigma\{W(s), Y(s); 0 \leq s \leq t\},$$

where $W(\cdot)$ and $Y(\cdot)$ are two independent standard Brownian motions valued in $\mathbb{R}^d$ and $\mathbb{R}^r$, respectively. Let $\mathcal{F} := \mathcal{F}_t$, and let $T > 0$ be the finite time duration and $0 < \delta \leq T$ be the

WU Shuang
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, *China; Department of Applied Mathematics, China University of Petroleum, Qingdao* 266580, *China.*
Email: namozhuntipusa@126.com.
SHU Lan
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, *China.* Email: shul@uestc.edu.cn.
◇*This paper was recommended for publication by Editor XIE Lihua.*

constant time delay. $\mathbb{E}$ denotes the expectation on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Moreover, we denote by $L^2(r, s; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued deterministic functions $\varphi(t)$ satisfying $\int_r^s |\varphi(t)|^2 dt < +\infty$, by $L^2(\mathcal{F}_t; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued $\mathcal{F}_t$-measurable random variables ζ satisfying $\mathbb{E}|\zeta|^2 < +\infty$, by $C([-\delta, 0]; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued continuous functions, and by $L^2_{\mathcal{F}}(r, s; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued $\mathcal{F}_t$-adapted processes $\psi(\cdot)$ satisfying $\mathbb{E}\int_r^s |\psi(t)|^2 dt < +\infty$. Define

$$\mathcal{F}_t^Y := \sigma\{Y(s); 0 \leq s \leq t\}.$$

Let U be a nonempty convex subset of $\mathbb{R}^k$. A control variable $v : [0, T] \times \Omega \to U$ is called admissible, if it is $\mathcal{F}_t^Y$-adapted and satisfies $\sup_{t\in[0,T]} \mathbb{E}\,|v_t|^m < \infty$, $m = 2, 3, \cdots$. The set of the admissible control variables is denoted by U_{ad}.

For given $v(\cdot) \in U_{ad}$, consider the following stochastic control system with time delay:

$$\begin{cases} dx^v(t) = b(t, x^v(t), x^v(t-\delta), v(t))dt + \sigma(t, x^v(t), x^v(t-\delta), v(t))dW(t), \quad t \in [0, T], \\ x^v(t) = \eta(t), \quad t \in [-\delta, 0], \end{cases} \tag{1}$$

where $\eta \in C([-\delta, 0]; \mathbb{R}^n)$ is the initial path of $x(\cdot)$ and

$$b : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U \to \mathbb{R}^n, \quad \sigma : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times d}.$$

We assume that the state processes $x^v(\cdot)$ cannot be observed directly, but the controllers can observe a related noisy process $Y(\cdot)$ of the state process which is described by

$$dY(t) = h(t, x^v(t), x^v(t-\delta), v(t))dt + d\overline{W}(t), \quad Y(0) = 0, \tag{2}$$

where $h : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U \to \mathbb{R}^r$ and $\overline{W}(\cdot)$ denotes a stochastic process depending on the control variable $v(\cdot)$.

We assume that the following hypothesis holds.

(**H1**) The functions b, σ are continuously differentiable in (x, y), and their partial derivatives are uniformly bounded; they are uniformly Lipschitz in v and there exists a constant $C > 0$ such that both b and σ are bounded by $C(1 + |x| + |y| + |v|)$; h is continuously differentiable in (x, y) and continuous in v, its derivatives and h are all uniformly bounded.

For any $v(\cdot) \in U_{ad}$, (H1) implies that (1) admits a unique $\mathcal{F}_t$-adapted solution. Define $d\mathbb{P}^v = Z^v(t)d\mathbb{P}$ with

$$Z^v(t) = \exp\left\{ \int_0^t h^\tau(s, x^v(s), x^v(s-\delta), v(s))dY(s) - \frac{1}{2}\int_0^t |h(s, x^v(s), x^v(s-\delta), v(s)|^2 ds \right\}.$$

Obviously, $Z(\cdot)$ is the unique $\mathcal{F}_t^Y$-adapted solution of

$$dZ^v(t) = Z^v(t)h^\tau(s, x^v(s), x^v(s-\delta), v(s))dY(t),\ Z^v(0) = 1. \tag{3}$$

By virtue of Itô's formula, we can prove that $\sup_{t\in[0,T]} \mathbb{E}\,|Z_t^v|^m < \infty$, $m = 2, 3, \cdots$. Hence, by Girsanov's theorem and (H1), $\mathbb{P}^v$ is a new probability measure and $(W(\cdot), \overline{W}(\cdot))$ is an $\mathbb{R}^{d+r}$-valued standard Brownian motion defined on the new probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}^v)$. We introduce the following cost functional

$$J(v(\cdot)) = \mathbb{E}^v \int_0^T l(t, x^v(t), v(t))dt + \Phi(x^v(T)), \tag{4}$$

where $\mathbb{E}^v$ denotes expectation on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}^v)$ and

$$l : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}, \quad \Phi : \mathbb{R}^n \to \mathbb{R}.$$

We need the following hypothesis.

(**H2**) (i) l is continuous in v, continuously differentiable in x, and its partial derivatives are continuous in (x, v) and bounded by $C(1 + |x| + |v|)$; (ii) Φ is continuously differentiable and Φ_x is bounded by $C(1 + |x|)$.

Our partially-observed optimal control problem is to minimize the cost functional (4) over $v(\cdot) \in U_{ad}$ subject to (1) and (2), i.e., to find $u(\cdot) \in U_{ad}$ satisfying

$$J(u(\cdot)) = \inf\{J(v(\cdot));\ v(\cdot) \in U_{ad}\}. \tag{5}$$

Obviously, cost functional (4) can be rewritten as

$$J(v(\cdot)) = \mathbb{E}\bigg[\int_0^T Z^v(t) l(t, x^v(t), v(t))dt + Z^v(T)\Phi(x^v(T))\bigg]. \tag{6}$$

Then the original problem (5) is equivalent to minimize (6) over $v(\cdot) \in U_{ad}$ subject to (1) and (3). Our main target is to find the necessary condition of the partially-observed optimal control $u(\cdot)$ in the form of Pontryagin stochastic maximum principle.

A stochastic control system whose state is described by solution of stochastic differential delay equation (SDDE) is called a time-delayed system. This kind of systems emerges naturally because some phenomena have the property of past dependence, that is to say, their behavior at time t not only depends on the current situation, but also on their past history. Due to the interesting structure and wide-range applications in physics, biology, engineering, and finance, optimal control problems of stochastic delay systems have received a lot of attentions in the past decades since the initial work of Kolmanovsky and Maizenberg[1], where a linear delay system with a quadratic cost functional was considered. In [2], Øksendal and Sulem discussed a certain class of stochastic control systems with time delay and gave sufficient stochastic maximum principle. On the other hand, the dynamic programming with time delay is taken into account by Larssen[3]. Recently, Peng and Yang[4] introduced a new type of stochastic differential equations, which were called anticipated backward stochastic differential equations (ABSDEs) and provided a new method to deal with optimal control problem with time delay. ABSDEs can be regarded as a generalization of classical BSDEs, which were introduced by Bismut[5] in the linear form and generalized to the nonlinear case by Pardoux and Peng[6]. By the duality relation between SDDEs and ABSDEs, Chen and Wu[7] obtained a maximum principle for stochastic optimal control problem with time delay. Øksendal, Sulem and Zhang[8] considered optimal control problems for stochastic delay systems with jumps and established sufficient and necessary maximum principle for an optimal control, and the adjoint processes were also shown to satisfy an ABSDE. It is remarkable that all the above ones are based on the assumption that the systems can be fully observed. However, in many practical systems, the systems' states cannot be observed directly, and the controllers have to make a decision according to their observable information. Thus, it is very natural and necessary to study partially-observed optimal control problems with time delay.

Some stochastic optimal control problems for partially-observed forward or forward-backward stochastic systems have been discussed by many authors, such as Bensoussan[9], Li and Tang[10], Tang[11], Zhang, Zhao, and Sheng[12], Wang, Zhang, and Zhang[13], Wang and Wu[14], Wu[15], Shi and Wu[16], etc. Li and Tang[10] obtained the maximum principle for the partially-observed forward stochastic optimal control problem in which the control variable is allowed to enter the diffusion and the observation coefficients. Tang[11] extended this result to the case with correlated noises between the system and the observation. Wang and Wu[14] considered risk-sensitive cost functional and non-convex control domain. Wu[15] studied the optimal control problem for partially-observed forward-backward stochastic systems when the control domain is convex. He obtained the maximum principle by the convex variational method. Shi and Wu[16] also researched this problem on the assumption that the control domain is not necessarily convex while the forward diffusion coefficient does not contain the control variable. The maximum principle was obtained by means of the spike variational technique. However, the systems in the above literatures were without time delays. Inspired by the former two types of works, in this paper, we will study a partially-observed stochastic time-delayed system given by (1) and (2). By Girsanov's theorem we can reformulate our original partially-observed optimal control problem to a completely observed one. The related adjoint processes are characterized by solutions of some finite-dimensional ABSDEs. Thus, our method is different from Bensoussan[9] in which he adopted an infinite-dimensional BSDE.

The paper is organized as follows. In Section 2, we first introduce a standard convex variation to get variational equations, then we give the corresponding adjoint equation which is finite-dimensional anticipated backward stochastic differential equation (ABSDE). By virtue of ABSDE, we derive a partially-observed stochastic maximum principle for time-delayed control systems. In Section 3, we focus on partially observed linear-quadratic (LQ) optimal control problem with time delay to illustrate the applications of our theoretical results obtained in Section 2. Under some proper conditions, we can combine maximum principle with classical linear filtering theory to find the explicit optimal observable control. Finally, we end this paper with some concluding remarks.

2 Partially-Observed Maximum Principle

In this section, combining Girsanov's theorem with a standard convex variational technique, we derive the maximum principle for the aforementioned partially observed optimal control problem with time delay.

Now let $u(\cdot)$ be optimal. Then for any $0 \le \varepsilon \le 1$ and $v(\cdot) \in U_{ad}$, we take the variational control $v^\varepsilon(\cdot) = u(\cdot) + \varepsilon v(\cdot)$. Because U is convex, $v^\varepsilon(\cdot)$ is in U_{ad}. For simplicity, we denote by $x^\varepsilon(\cdot)$, $x(\cdot)$, $Z^\varepsilon(\cdot)$, $Z(\cdot)$ the state trajectories of (1) and (2) corresponding to $v^\varepsilon(\cdot)$ and $u(\cdot)$.

For simplification, we introduce the notations

$$\theta(v^\varepsilon(t)) = \theta(t, x(t), x(t-\delta), v^\varepsilon(t)), \quad \theta(u(t)) = \theta(t, x(t), x(t-\delta), u(t)),$$

where $\theta = b, \sigma, h$ as well as their partial derivatives in the optimal trajectory (x, y).

We now introduce the following variational equations:

$$
\begin{aligned}
dx_1(t) &= [b_x(u(t))x_1(t) + b_y(u(t))x_1(t-\delta) + b_v(u(t))v(t)]dt \\
&\quad +[\sigma_x(u(t))x_1(t) + \sigma_y(u(t))x_1(t-\delta) + \sigma_v(u(t))v(t)]dW(t), \\
x_1(t) &= 0, \quad t \in [-\delta, 0],
\end{aligned} \tag{7}
$$

and

$$
\begin{aligned}
dZ_1(t) &= [Z_1(t)h(u(t)) + Z(t)h_x(u(t))x_1(t) + Z(t)h_y(u(t))x_1(t-\delta) \\
&\quad +Z(t)h_v(u(t))v(t)]^\tau dY(t), \\
Z_1(0) &= 0.
\end{aligned} \tag{8}
$$

By (H1) and Theorem 2.2 in [7], it is easy to know that (7) and (8) admit unique adapted solutions $x_1(\cdot)$ and $Z_1(\cdot)$, respectively.

The following lemma is due to Chen and Wu[7].

Lemma 2.1 *Let* (H1) *hold. Then, we have*

$$
\lim_{\varepsilon\to 0} \sup_{0\le t\le T} \mathbb{E}\left|\frac{x^\varepsilon(t) - x(t)}{\varepsilon} - x_1(t)\right|^2 = 0.
$$

We also need to obtain some ε-order estimations of the difference between the perturbed observed process $Z^\varepsilon(\cdot)$ with the sum of the optimal observed process $Z(\cdot)$ and the variational observed $Z_1(\cdot)$. The following lemma play an important role when we derive the variational inequality.

Lemma 2.2 *Let* (H1) *hold. Then, we have*

$$
\lim_{\varepsilon\to 0} \sup_{0\le t\le T} \mathbb{E}\left|\frac{Z^\varepsilon(t) - Z(t)}{\varepsilon} - Z_1(t)\right|^2 = 0.
$$

Proof By the definition of $Z(\cdot)$ and $Z_1(\cdot)$, we have

$$
\begin{aligned}
Z(t) + \varepsilon Z_1(t) &= 1 + \int_0^t Z(s)h^\tau(u(s))dY(s) \\
&\quad +\varepsilon \int_0^t [Z_1(s)h(u(s)) + Z(s)h_x(u(s))x_1(s) \\
&\qquad +Z(s)h_y(u(s))x_1(s-\delta) + Z(s)h_v(u(s))v(s)]^\tau dY(s) \\
&= 1 + \varepsilon \int_0^t Z_1(s)h^\tau(u(s))dY(s) \\
&\quad + \int_0^t Z(s)[h(s, x(s)+\varepsilon x_1(s), x(s-\delta)+\varepsilon x_1(s-\delta), u(s)+\varepsilon v(s)]^\tau dY(s) \\
&\quad -\varepsilon \int_0^t Z(s)[A^\varepsilon(s)]^\tau dY(s),
\end{aligned}
$$

where

$$
\begin{aligned}
&A^\varepsilon(s)\\
&= \int_0^1 [h_x(s, x(s) + \lambda\varepsilon x_1(s), x(s-\delta) + \lambda\varepsilon x_1(s-\delta), u(s) + \lambda\varepsilon v(s)) - h_x(u(s))] d\lambda x_1(s)\\
&\quad + \int_0^1 [h_y(s, x(s) + \lambda\varepsilon x_1(s), x(s-\delta) + \lambda\varepsilon x_1(s-\delta), u(s) + \lambda\varepsilon v(s)) - h_y(u(s))] d\lambda x_1(s-\delta)\\
&\quad + \int_0^1 [h_v(s, x(s) + \lambda\varepsilon x_1(s), x(s-\delta) + \lambda\varepsilon x_1(s-\delta), u(s) + \lambda\varepsilon v(s)) - h_v(u(s))] d\lambda v(s).
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&Z^\varepsilon(t) - Z(t) - \varepsilon Z_1(t)\\
&= \int_0^t Z^\varepsilon(s)[h(s, x^\varepsilon(s), x^\varepsilon(s-\delta), v^\varepsilon(s))]^\tau dY(s) - \varepsilon \int_0^t Z_1(s) h^\tau(u(s)) dY(s)\\
&\quad - \int_0^t Z(s)[h(s, x(s) + \varepsilon x_1(s), x(s-\delta) + \varepsilon x_1(s-\delta), u(s) + \varepsilon v(s)]^\tau dY(s)\\
&\quad + \varepsilon \int_0^t Z(s)[A^\varepsilon(s)]^\tau dY(s)\\
&= \int_0^t (Z^\varepsilon(s) - Z(s) - \varepsilon Z_1(s))[h(s, x^\varepsilon(s), x^\varepsilon(s-\delta), v^\varepsilon(s))]^\tau dY(s)\\
&\quad + \int_0^t (Z(s) + \varepsilon Z_1(s))[h(s, x^\varepsilon(s), x^\varepsilon(s-\delta), v^\varepsilon(s)) - h(s, x(s)\\
&\qquad\qquad + \varepsilon x_1(s), x(s-\delta) + \varepsilon x_1(s-\delta), u(s) + \varepsilon v(s))]^\tau dY(s)\\
&\quad + \varepsilon \int_0^t Z_1(s)[h((s, x(s) + \varepsilon x_1(s), x(s-\delta) + \varepsilon x_1(s-\delta), u(s) + \varepsilon v(s))]^\tau dY(s)\\
&\quad - \varepsilon \int_0^t Z_1(s) h^\tau(u(s)) dY(s) + \varepsilon \int_0^t Z(s)[A^\varepsilon(s)]^\tau dY(s)\\
&= \int_0^t (Z^\varepsilon(s) - Z(s) - \varepsilon Z_1(s))[h(s, x^\varepsilon(s), x^\varepsilon(s-\delta), v^\varepsilon(s))]^\tau dY(s)\\
&\quad + \int_0^t (Z(s) + \varepsilon Z_1(s))[B_1^\varepsilon(s)]^\tau dY(s) + \varepsilon \int_0^t Z_1(s)[B_2^\varepsilon(s)]^\tau dY(s)\\
&\quad + \varepsilon \int_0^t Z(s)[A^\varepsilon(s)]^\tau dY(s),
\end{aligned}
$$

where

$$
\begin{aligned}
B_1^\varepsilon(s) &= h(s, x^\varepsilon(s), x^\varepsilon(s-\delta), v^\varepsilon(s)) - h(s, x(s) + \varepsilon x_1(s), x(s-\delta) + \varepsilon x_1(s-\delta), u(s) + \varepsilon v(s)),\\
B_2^\varepsilon(s) &= h((s, x(s) + \varepsilon x_1(s), x(s-\delta) + \varepsilon x_1(s-\delta), u(s) + \varepsilon v(s)) - h(u(s)).
\end{aligned}
$$

Note that

$$
\begin{aligned}
B_1^\varepsilon(s) = \int_0^1 & [h_x(s, x(s) + \varepsilon x_1(s) + \lambda(x^\varepsilon(s) - x(s) - \varepsilon x_1(s)), x(s-\delta) + \varepsilon x_1(s-\delta) \\
& + \lambda(x^\varepsilon(s-\delta) - x(s-\delta) - \varepsilon x_1(s-\delta)), v^\varepsilon(s))] d\lambda (x^\varepsilon(s) - x(s) - \varepsilon x_1(s)) \\
& + \int_0^1 [h_y(s, x(s) + \varepsilon x_1(s) + \lambda(x^\varepsilon(s) - x(s) - \varepsilon x_1(s)), x(s-\delta) + \varepsilon x_1(s-\delta) \\
& \quad + \lambda(x^\varepsilon(s-\delta) - x(s-\delta) - \varepsilon x_1(s-\delta)), v^\varepsilon(s))] d\lambda (x^\varepsilon(s-\delta) - x(s-\delta) - \varepsilon x_1(s-\delta)).
\end{aligned}
$$

Due to the fact that

$$
\sup_{0\le s\le t} \mathbb{E}|x^\varepsilon(s-\delta) - x(s-\delta) - \varepsilon x_1(s-\delta)|^2 \le \sup_{0\le s\le t} \mathbb{E}|x^\varepsilon(s) - x(s) - \varepsilon x_1(s)|^2, \quad t \in [0,T]
$$

and Lemma 2.1, we know that

$$
\mathbb{E}\int_0^t |(Z(s) + \varepsilon Z_1(s))B_1^\varepsilon(s)|^2 ds \le C_\varepsilon \varepsilon^2, \tag{9}
$$

hereafter C_ε denotes some nonnegative constant such that $C_\varepsilon \to 0$ as $\varepsilon \to 0$.

Moreover, it is easy to see that

$$
\sup_{0\le t\le T} \mathbb{E}\bigg(\varepsilon \int_0^t Z(s)[A^\varepsilon(s)]^\tau dY(s)\bigg)^2 \le C_\varepsilon \varepsilon^2 \tag{10}
$$

and

$$
\sup_{0\le t\le T} \mathbb{E}\bigg(\varepsilon \int_0^t Z_1(s)[B_2^\varepsilon(s)]^\tau dY(s)\bigg)^2 \le C_\varepsilon \varepsilon^2. \tag{11}
$$

By (9), (10) and (11), we have

$$
\begin{aligned}
& \mathbb{E}|Z^\varepsilon(t) - Z(t) - \varepsilon Z_1(t)|^2 \\
\le\ & C\bigg[\int_0^t \mathbb{E}|Z^\varepsilon(s) - Z(s) - \varepsilon Z_1(s)|^2 ds + \mathbb{E}\int_0^t |(Z(s) + \varepsilon Z_1(s))B_1^\varepsilon(s)|^2 ds \\
& + \sup_{0\le t\le T} \mathbb{E}\bigg(\varepsilon \int_0^t Z(s)[A^\varepsilon(s)]^\tau dY(s)\bigg)^2 + \sup_{0\le t\le T} \mathbb{E}\bigg(\varepsilon \int_0^t Z_1(s)[B_2^\varepsilon(s)]^\tau dY(s)\bigg)^2\bigg] \\
\le\ & C\int_0^t \mathbb{E}|Z^\varepsilon(s) - Z(s) - \varepsilon Z_1(s)|^2 ds + C_\varepsilon \varepsilon^2.
\end{aligned}
$$

By the Gronwall's inequality, we obtain the desired result. ∎

Then we have the following variational inequality.

Lemma 2.3 *Let* (H1) and (H2) *hold. Then, we have*

$$
\begin{aligned}
& \mathbb{E}\int_0^T [Z_1(t) l(t, x(t), u(t)) + Z(t) l_x^\tau(t, x(t), u(t)) x_1(t) + Z(t) l_v^\tau(t, x(t), u(t)) v(t)] dt \\
& + \mathbb{E}[Z_1(T)\Phi(x(T))] + \mathbb{E}[Z(T)\Phi_x(x(T)) x_1(T)] \ge 0.
\end{aligned} \tag{12}
$$

Proof Using the Taylor expansion, Lemmas 2.1 and 2.2, we have

$$\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathbb{E}[Z^\varepsilon(T)\Phi(x^\varepsilon(T))-Z(T)\Phi(x(T))]=\mathbb{E}[Z_1(T)\Phi(x(T))+Z(T)\Phi_x(x(T))x_1(T)]$$

and

$$\begin{aligned}&\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathbb{E}\int_0^T[Z^\varepsilon(t)l(t,x^\varepsilon(t),u^\varepsilon(t))-Z(t)l(t,x(t),u(t))]dt\\ =&\ \mathbb{E}\int_0^T[Z_1(t)l(t,x(t),u(t))+Z(t)l_x^\tau(t,x(t),u(t))x_1(t)+Z(t)l_v^\tau(t,x(t),u(t))v(t)]dt.\end{aligned}$$

Then, by the fact that $\varepsilon^{-1}[J(u(\cdot)+\varepsilon v(\cdot))-J(u(\cdot))]\geq 0$, we draw the desired conclusion.

We now focus on a necessary condition of the optimal control $u(\cdot)$. For this purpose, we define the Hamiltonian function

$$H(t,x,y,v,p,q,\overline{z})=l(t,x,v)+\langle p,b(t,x,y,v)\rangle+\langle q,\sigma(t,x,y,v)\rangle+\langle\overline{z},h(t,x,y,v)\rangle, \tag{13}$$

where $H:[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times U\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times\mathbb{R}^r\to\mathbb{R}$.

To derive the maximum principle, we introduce the following adjoint equations:

$$\begin{aligned}&-dy(t)=l(u(t))dt-z(t)dW(t)-\overline{z}d\overline{W}(t),\quad t\in[0,T],\\ &y(T)=\Phi(x(T)),\ \overline{z}(t)=0,\quad t\in(T,T+\delta],\end{aligned}\tag{14}$$

$$\begin{aligned}-dp(t)=&\big\{b_x^\tau(u(t))p(t)+\mathbb{E}^u[b_y^\tau(u(t+\delta)p(t+\delta)|\mathcal{F}_t]+\sigma_x^\tau(u(t))q(t)\\ &+\mathbb{E}^u[\sigma_y^\tau(u(t+\delta))q(t+\delta)|\mathcal{F}_t]+h_x^\tau(u(t))\overline{z}(t)+\mathbb{E}^u[h_y^\tau(u(t+\delta))\overline{z}(t+\delta)|\mathcal{F}_t]\\ &+l_x(t,x(t),u(t))\big\}dt-q(t)dW(t)-\overline{q}(t)d\overline{W}(t),\quad t\in[0,T],\\ p(T)=&\ \Phi_x(x(T)),\quad p(t)=0,\quad t\in(T,T+\delta],\quad q(t)=0,\quad t\in[T,T+\delta].\end{aligned}\tag{15}$$

Note that $d\widetilde{Z}(t)=[h_x(u(t))x_1(t)+h_y(u(t))x_1(t-\delta)+h_v(u(t))v(t)]^\tau d\overline{W}(t),\widetilde{Z}(0)=0$, where $\widetilde{Z}(t)=Z^{-1}(t)Z_1(t)$. Moreover, we have

$$\begin{aligned}&\mathbb{E}^u\int_0^T\{\langle\overline{p}(t),-\overline{b}_y(u(t))x_1(t-\delta)\rangle+\langle\mathbb{E}^u[\overline{b}_y^\tau(u(t+\delta))\overline{p}(t+\delta)|\mathcal{F}_t],x_1(t)\rangle\}dt\\ =&\ \mathbb{E}^u\int_0^T\langle\overline{p}(t),-\overline{b}_y(u(t))x_1(t-\delta)\rangle dt+\mathbb{E}^u\int_\delta^{T+\delta}\langle\overline{b}_y^\tau(u(t))\overline{p}(t),x_1(t-\delta)\rangle dt\\ =&\ \mathbb{E}^u\int_0^\delta\langle\overline{p}(t),-\overline{b}_y(u(t))x_1(t-\delta)\rangle dt+\mathbb{E}^u\int_T^{T+\delta}\langle\overline{b}_y^\tau(u(t))\overline{p}(t),x_1(t-\delta)\rangle dt=0,\end{aligned}$$

where $\overline{p}=p,q,\overline{z}$ and $\overline{b}=b,\sigma,h$, correspondingly. Then, applying Itô's formula to $\langle y(t),\widetilde{Z}(t)\rangle+\langle p(t),x_1(t)\rangle$ and comparing it with the variational inequality (12), we can get

$$\mathbb{E}^u\int_0^T\langle b_v^\tau(u(t))p(t)+\sigma_v^\tau(u(t))q(t)+h_v^\tau(u(t))\overline{z}(t)+l_v(t,x(t),u(t)),v(t)\rangle dt\geq 0. \tag{16}$$

By the definition (13), (16) can be rewritten as

$$\mathbb{E}^u\int_0^T\langle H_v(t,x(t),x(t-\delta),u(t),p(t),q(t),\overline{z}(t)),v(t)\rangle dt\geq 0. \tag{17}$$

Using the method similar to that of References [7, 16], we can derive the main result of this paper. ▌

Theorem 2.4 *Assume that* (H1) *and* (H2) *hold. Let* $u(\cdot)$ *be optimal. Then, the maximum principle*

$$\mathbb{E}^u[\langle H_v(t,x(t),x(t-\delta),u(t),p(t),q(t),\overline{z}(t)),v-u(t)\rangle|\mathcal{F}_t^Y]\geq 0,\quad \forall v\in U,\text{a.e., a.s.},$$

holds, where the Hamiltonian function H *is defined by* (13).

Remark 2.5 In our partially-observed optimal control problem with time delay, we assume that the diffusion coefficient contains the control variable and the control domain is convex. To our best knowledge, the general maximum principle for stochastic delay systems is still an open problem even if the system is completely observed. On the other hand, it is worthwhile to note that our problem should be distinguished from the optimal control problems with partial information, where a sub-filtration is given to represent the information available to the controller instead of an observation process.

3 Application: Partially-Observed LQ Problem

In this section, we give an example of partially-observed linear-quadratic optimal control problem with time delay. Though there is no general filtering results for ABSDEs, we will try to give an explicit observable optimal control by means of our theoretical result in Section 2 and classical filtering theory. Let us consider the following stochastic control system with time delay ($d=r=1$):

$$\begin{cases} dx^v(t)=[A(t)x^v(t)+A_\delta(t)x^v(t-\delta)+B(t)v(t)]dt+C(t)dW(t),\quad t\in[0,T],\\ x^v(t)=\eta(t),\quad t\in[-\delta,0]\end{cases}\tag{18}$$

and the observation

$$dY(t)=D(t)dt+d\overline{W}(t),\quad Y(0)=0.\tag{19}$$

The cost functional is described as

$$J(v(\cdot))=\frac{1}{2}\mathbb{E}^v\int_0^T[\langle R(t)x^v(t),x^v(t)\rangle+\langle N(t)v(t),v(t)\rangle]dt+\langle Mx^v(T),x^v(T)\rangle,\tag{20}$$

where $A(\cdot),A_\delta(\cdot)$ are deterministic $n\times n$ bounded matrix-valued functions, $B(\cdot)$ is a deterministic $n\times k$ bounded matrix-valued function, $C(\cdot)$ is a deterministic $n\times d$ bounded matrix-valued function, $D(\cdot)$ is a deterministic $r\times 1$ bounded matrix-valued function, $R(\cdot)$ is a deterministic $n\times n$ non-negative symmetric bounded matrix-valued function, $N(\cdot)$ is a deterministic $k\times k$ positive symmetric bounded matrix-valued function and $N(\cdot)^{-1}$ is also bounded, and M is deterministic $n\times n$ non-negative symmetric matrix. The Hamiltonian function is given by

$$\begin{aligned}H(t,x,y,v,p,q,\overline{z})=&\frac{1}{2}[\langle R(t)x,x\rangle+\langle N(t)v,v\rangle]+\langle p,A(t)x+A_\delta(t)y+B(t)v\rangle\\&+\langle q,C(t)\rangle+\langle\overline{z},D(t)\rangle.\end{aligned}\tag{21}$$

If $u(\cdot)$ is optimal, then it follows from Theorem 2.4 and (21) that

$$u(t) = -N^{-1}(t)B^{\tau}(t)\mathbb{E}^{u}[p(t)|\mathcal{F}_t^Y], \quad 0 \le t \le T, \tag{22}$$

where $(p(\cdot), q(\cdot), x^u(\cdot))$ is the solution of the following FBSDE

$$\begin{cases} dx^u(t) = \{A(t)x^u(t) + A_\delta(t)x^u(t-\delta) - B(t)N^{-1}(t)B^{\tau}(t)\mathbb{E}^u[p(t)|\mathcal{F}_t^Y]\}dt \\ \qquad\qquad + C(t)dW(t), \quad t \in [0,T], \\ -dp(t) = \big\{A^{\tau}(t)p(t) + \mathbb{E}^u[A_\delta^{\tau}(t+\delta)p(t+\delta)|\mathcal{F}_t] + R(t)x^u(t)\big\}dt \\ \qquad\qquad - q(t)dW(t) - \overline{q}(t)d\overline{W}(t), \quad t \in [0,T], \\ p(T) = Mx^u(T), \quad p(t) = 0, \quad t \in (T, T+\delta], \\ q(t) = 0, \quad \overline{q}(t) = 0, \quad t \in [T, T+\delta]. \end{cases} \tag{23}$$

Note that the forward equation of (23) contains the conditional expectation of $p(t)$ with respect to $\mathcal{F}_t^Y$, then it is distinguished from the general FBSDE appearing in [17] and the existence and uniqueness of solution to (23) is not so clear. However, if we fix the trajectory $x^u(\cdot)$, then $((p(\cdot), q(\cdot), \overline{q}(\cdot))$ satisfies a certain ABSDE which is well-defined. Note that $\mathbb{E}^u[p(t+\delta)|\mathcal{F}_t^Y] = \mathbb{E}^u\{\mathbb{E}^u[p(t+\delta)|\mathcal{F}_{t+\delta}^Y]|\mathcal{F}_t^Y\} = \mathbb{E}^u[p(t+\delta)|\mathcal{F}_t^Y]$. Then from Theorems 8.1 and 8.4 in [18], we have

$$\begin{cases} d\widehat{x}^u(t) = [A(t)\widehat{x}^u(t) + A_\delta(t)\widehat{x}^u(t-\delta) - B(t)N^{-1}(t)B^{\tau}(t)\widehat{p}(t)]dt, \quad t \in [0,T], \\ -d\widehat{p}(t) = \big\{A^{\tau}(t)\widehat{p}(t) + \mathbb{E}^u[A_\delta^{\tau}(t+\delta)\widehat{p}(t+\delta)|\mathcal{F}_t^Y] + R(t)\widehat{x}^u(t)\big\}dt - \widehat{\overline{q}}(t)d\overline{W}(t), \quad t \in [0,T], \\ \widehat{p}(T) = M\widehat{x}^u(T), \quad \widehat{p}(t) = 0, \quad t \in (T, T+\delta], \quad \widehat{\overline{q}}(t) = 0, \quad t \in [T, T+\delta], \end{cases} \tag{24}$$

where $\widehat{\phi}(t) = \mathbb{E}^u[\phi(t)|\mathcal{F}_t^Y]$ is the filtering estimate of the state $\phi(t)$ depending on the observable filtration $\mathcal{F}_t^Y$, $\phi = x, p, \overline{q}$. From Theorem 2.1 in [17], the general FBSDE (24) has a unique solution $(\widehat{x}^u(\cdot), \widehat{p}(\cdot), \widehat{\overline{q}}(\cdot))$, and this implies that (23) admits a solution $(x^u(\cdot), p(\cdot), q(t), \overline{q}(\cdot))$.

Our next job is to prove the admissible control (22) which is determined by (23) is really optimal. Note that $\mathbb{E}^v$ and $\mathbb{E}^u$ are equivalent. Then for any admissible control $v(\cdot)$, we have

$$\begin{aligned} J(v(\cdot)) - J(u(\cdot)) = \frac{1}{2}\mathbb{E}^u \int_0^T [&\langle R(t)(x^v(t) - x(t)), x^v(t) - x(t)\rangle + \langle N(t)(v(t) - u(t)), v(t) - u(t)\rangle \\ &+2\langle R(t)x(t), x^v(t) - x(t)\rangle + 2\langle N(t)u(t), v(t) - u(t)\rangle]dt \\ &+\frac{1}{2}\mathbb{E}^u[\langle M(x^v(T) - x(T)), x^v(T) - x(T)\rangle + 2\langle Mx(T), x^v(T) - x(T)\rangle]. \end{aligned}$$

By the initial and terminal conditions, it follows that

$$\begin{aligned} &\mathbb{E}^u \int_0^T \{\langle A_\delta(t)(x^v(t-\delta) - x(t-\delta)), p(t)\rangle - \langle x^v(t) - x(t), \mathbb{E}^u[A_\delta^{\tau}(t+\delta)p(t+\delta)|\mathcal{F}_t]\rangle\}dt \\ &= \mathbb{E}^u \int_0^T \langle A_\delta(t)(x^v(t-\delta) - x(t-\delta)), p(t)\rangle dt - \mathbb{E}^u \int_\delta^{T+\delta} \langle x^v(t-\delta) - x(t-\delta), A_\delta^{\tau}(t)p(t)\rangle dt \\ &= 0. \end{aligned}$$

Applying Itô's formula to $\langle x^v(t-\delta) - x(t-\delta), p(t)\rangle$, we get

$$\mathbb{E}^u\ \langle x^v(T) - x(T), p(T)\rangle = -\mathbb{E}^u \int_0^T [\langle R(t)(x^v(t) - x(t)), x(t)\rangle + \langle B(t)(v(t) - u(t)), p(t)\rangle]dt.$$

As $R(t), M$ are nonnegative and $N(t)$ is positive, we can derive

$$\begin{aligned} J(v(\cdot)) - J(u(\cdot)) &\geq \mathbb{E}^u \int_0^T \langle N(t)u(t) + B^\tau(t)p(t), v(t) - u(t)\rangle dt \\ &= \mathbb{E}^u \int_0^T \langle N(t)u(t) + B^\tau(t)\mathbb{E}^u[p(t)|\mathcal{F}_t^Y], v(t) - u(t)\rangle dt, \end{aligned}$$

so it is clear that $u(t) = -N(t)^{-1}B^\tau(t)\mathbb{E}^u[p(t)|\mathcal{F}_t^Y]$ is optimal.

The remaining task is to compute the filtering estimate $\widehat{p}(t)$. From (24), it is obvious that $\widehat{\widehat{q}}(\cdot) = 0$, then (24) can be rewritten as

$$\begin{cases} d\widehat{x}(t) = [A(t)\widehat{x}(t) + A_\delta(t)\widehat{x}(t-\delta) - B(t)N^{-1}(t)B^\tau(t)\widehat{p}(t)]dt, \quad t \in [0, T], \\ -d\widehat{p}(t) = [A^\tau(t)\widehat{p}(t) + A_\delta^\tau(t+\delta)\widehat{p}(t+\delta) + R(t)\widehat{x}(t)]dt, \quad t \in [0, T], \\ \widehat{p}(T) = M\widehat{x}(T), \quad \widehat{p}(t) = 0, \quad t \in (T, T+\delta], \end{cases} \tag{25}$$

where we omit the subscript u for simplification. Though (25) is a deterministic FBSDE, we don't obtain its explicit solution by means of common FBSDE technique because of the delayed and advanced time durations. However, we can relate (25) to a deterministic linear quadratic optimal control problem with time delay as follows:

$$\begin{cases} dx^v(t) = [A(t)x^v(t) + A_\delta(t)x^v(t-\delta) + B(t)v(t)]dt, \quad t \in [0, T], \\ x^v(t) = \eta(t), \quad t \in [-\delta, 0], \end{cases} \tag{26}$$

and the cost functional

$$J(v(\cdot)) = \frac{1}{2}\int_0^T \langle R(t)x^v(t), x^v(t)\rangle + \langle N(t)v(t), v(t)\rangle dt + \langle Mx^v(T), x^v(T)\rangle, \tag{27}$$

where admissible control $v(\cdot)$ is deterministic satisfying $\int_0^T |v(t)|^2 dt < +\infty$. It follows that $u(t) = -N^{-1}(t)B^\tau(t)\widehat{p}(t)$ is optimal subject to (26) and (27), where $\widehat{p}(t)$ is determined by (25). Thanks to Theorem 4.2 in [19], we know that

$$u(t) = -N^{-1}(t)B^\tau(t)\left(E_0(t)\widehat{x}(t) + \int_{t-\delta}^t E_1(t, \theta - t)\widehat{x}(\theta)d\theta\right), \quad t \in [0, T] \tag{28}$$

is optimal, where $E_0(t), E_1(t, \theta), E_2(t, \theta, \zeta),\ t \in [0, T],\ \theta, \zeta \in [-\delta, 0]$ satisfy the following sets of equations

$$\begin{cases} \dot{E}_0(t) + E_0(t)A(t) + A^\tau(t)E_0(t) + E_1(t, 0) + E_1^\tau(t, 0) + R(t) \\ - E_0^\tau(t)B(t)N^{-1}(t)B^\tau(t)E_0(t) = 0, \\ \dfrac{\partial}{\partial t}E_1(t, \theta) - \dfrac{\partial}{\partial \theta}E_1(t, \theta) + A^\tau(t)E_1(t, \theta) + E_2(t, 0, \zeta) \\ - E_0^\tau(t)B(t)N^{-1}(t)B^\tau(t)E_1(t, \theta) = 0, \\ \dfrac{\partial}{\partial t}E_2(t, \theta, \zeta) - \dfrac{\partial}{\partial \theta}E_2(t, \theta, \zeta) - \dfrac{\partial}{\partial \zeta}E_2(t, \theta, \zeta) \\ - E_1^\tau(t, \theta)B(t)N^{-1}(t)B^\tau(t)E_1(t, \theta) = 0, \\ E_2(t, \theta, \zeta) = E_2(t, \zeta, \theta), \quad t \in [0, T], \quad \theta, \zeta \in [-\delta, 0], \end{cases} \tag{29}$$

with boundary conditions

$$\begin{cases} E_0(T) = M, \quad E_0(t)A_\delta(t) = E_1(t, -\delta), \quad E_1(T, \theta) = E_2(T, \theta, \zeta) = 0, \\ E_1^\tau(t, \theta)A_\delta(t) = E_2(t, \theta, -\delta), \quad t \in [0, T], \quad \theta, \zeta \in [-\delta, 0]. \end{cases}$$

Moreover, using the parallel rule, we can prove the uniqueness of optimal control, and this gives the relation between $\widehat{p}(t)$ and $\widehat{x}(t)$ that $\widehat{p}(t) = E_0(t)\widehat{x}(t) + \int_{t-\delta}^{t} E_1(t, \theta - t)\widehat{x}(\theta)d\theta$.

Now we summarize all the results obtained so far and give the following proposition.

Proposition 3.1 *For our partially-observed linear-quadratic optimal control problem* (18)–(20)*, an observable optimal control* $u(\cdot)$ *is given by* (22)*, where* $\widehat{p}(t) = \mathbb{E}^u[p(t)|\mathcal{F}_t^Y]$ *is the solution of FBSDE* (25)*. Furthermore, the feedback regulator of the filtering estimate for optimal trajectory is given by* (28)*, where* $E_0(t), E_1(t, \theta), E_2(t, \theta, \zeta),\ t \in [0, T],\ \theta, \zeta \in [-\delta, 0]$ *satisfy the equation sets* (29)*.*

Remark 3.2 Since the filtering $(\widehat{x}^u(\cdot), \widehat{p}(\cdot), \widehat{\overline{q}}(\cdot))$ solves a general FBSDE (24), it is different from the existing filtering literature. Obviously, FBSDE (24) can be regarded as a generalization of forward-backward stochastic differential filtering equations (FBSDFEs) appearing in [20]. Just like [20], we call (24) a general forward-backward stochastic differential filtering equation.

4 Conclusion

In this paper, we have discussed one kind of partially-observed optimal control problem with time delay. More specially, we use Girsanov's theorem to transform our optimal control problem to completely observable case and apply the conventional approach to get the partially-observed maximum principle. The related adjoint processes are characterized by solutions of some finite-dimensional ABSDEs. Because the maximum principle depends strongly on adjoint processes, it is necessary to investigate the filtering estimate for adjoint processes which satisfy ABSDEs when we try to get an observable optimal control. However, there is no general filtering results for ABSDEs as yet, even if $\delta = 0$. Only in some special cases we can get the filtering estimate for this kind of stochastic systems by combining ABSDEs theory with the traditional filtering theory. To show the application of our maximum principle, we give an example for partially-observed linear-quadratic optimal control problem with time delay and find an explicit observable control variable satisfying the necessary condition of optimality. Furthermore, based on the fact that the observation process does not depend on the control variable, we also give the feedback regulator of the filtering estimate for optimal trajectory.

References

[1] Kolmanovsky V B and Maizenberg T L, Optimal control of stochastic systems with aftereffect, *Automation and Remote Control*, 1973, **1**: 47–61.

[2] Øksendal B and Sulem A, A maximum principle for optimal control of stochastic systems with delay, with applications to finance, *Optimal Control and Partial Differential Equations*, Amsterdam: ISO Press, 2000, 64–79.

[3] Larssen B, Dynamic programming in stochastic control of systems with delay, *Stochastic and Stochastics Reports*, 2002, **74**(3): 651–673.

[4] Peng S and Yang Z, Anticipated backward stochastic differential equations, *Ann. Proba.*, 2009, **37**(3): 877–902.

[5] Bismut J M, An introductory approach to duality in optimal stochastic control, *SIAM Rev.*, 1978, **20**(1): 62–78.

[6] Pardoux E and Peng S, Adapted solution of backward stochastic dierential equation, *Systems Control Lett.*, 1990, **14**(1): 55–61.

[7] Chen L and Wu Z, Maximum principle for the stochastic optimal control problem with delay and application, *Automatica*, 2010, **46**(6): 1074–1080.

[8] Øksendal B, Sulem A, and Zhang T, Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations, *Advances in Applied Probability*, 2011, **43**(2): 572–596.

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

[10] Li X and Tang S, General necessary conditions for partially observed optimal stochastic controls, *J. Appl. Probab.*, 1995, **32**(4): 1118–1137.

[11] Tang S, The maximum principle for partially-observed optimal control of stochastic differential equations, *SIAM J. Control Optim.*, 1998, **36**(5): 1596–1617.

[12] Zhang W, Zhao Y, and Sheng L, Some remarks on stability of stochastic singular systems with state-dependent noise, *Automatica*, 2014, **51**: 273–277.

[13] Wang G, Zhang C, and Zhang W, Stochastic maximum principle for mean-field type optimal control under partial information, *IEEE Trans. Autom. Control*, 2014, **59**(2): 522–528.

[14] Wang G and Wu Z, General maximum principles for partially observed risk-sensitive optimal control problems and applications to finance, *J. Optim. Theory. Appl.*, 2009, **141**(3): 677–700.

[15] Wu Z, A maximum principle for partially observed optimal control of forward-backward stochastic control systems, *Sci. China Ser. F.*, 2010, **53**(11): 2205–2214.

[16] Shi J and Wu Z, Maximum principle for partially-observed optimal control of fully-coupled forward-backward stochastic systems, *J. Optim. Theory. Appl.*, 2010, **145**(3): 543–578.

[17] Chen L and Wu Z, A type of generalized forward-backward stochastic differential equations and applications, *Chin. Ann. Math. Ser. B*, 2011, **32**(2): 279–292.

[18] Liptser R S and Shiryayev A N, *Statistics of Random Process*, Springer-Verlag, New York, 1977.

[19] Huang J, Li X, and Shi J, Froward-backward linear quadratic stochastic optimal control problem with delay, *Systems Control Lett.*, 2012, **61**(5): 623–630.

[20] Huang, J, Wang G, and Xiong J, A maximum principle for partial information backward stochastic control problems with applications, *SIAM J. Control Optim.*, 2009, **48**(4): 2106–2117.