Zhang *Advances in Difference Equations* (2017) 2017:98
DOI 10.1186/s13662-017-1155-9

Advances in Difference Equations
a SpringerOpen Journal

RESEARCH Open Access

Maximum principle for near-optimality of stochastic delay control problem

Feng Zhang*

*Correspondence:
zhangfeng1104@sdufe.edu.cn
School of Mathematics and Quantitative Economics, Shandong University of Finance and Economics, Jinan, 250014, China

Abstract

This paper is concerned with near-optimality for stochastic control problems of linear delay systems with convex control domain and controlled diffusion. Necessary and sufficient conditions for a control to be near-optimal are established by Pontryagin's maximum principle together with Ekeland's variational principle.

MSC: 93Exx; 49Kxx; 60H30

Keywords: maximum principle; near-optimal control; stochastic differential delay equation; Ekeland's variational principle

1 Introduction

Many real-world systems are characteristic of dependence on the past, *i.e.* their present states not only depend on the current situation, but also on the previous history. This is called time delay. Indeed, phenomena with time delays are common in the fields of both natural and social sciences, such as physics, engineering, biology, economics and finance; see for example [1–3].

Stochastic optimal control problems with time-delay systems have received a lot of research attention recently. However, this kind of control problem remains practically intractable due to its infinite-dimensional nature. Fortunately, when the distributed (average) and pointwise time delays are involved in the state process, optimal control problems are found to be solvable under certain conditions. For the applications of the dynamic programming principle to this field, see [4, 5]. For Pontryagin's maximum principle applied to it, see [6–8]. Along this line, by a duality between linear stochastic differential delay equations (SDDEs) and anticipated backward stochastic differential equations (ABSDEs) established in [9], the maximum principle for stochastic delay optimal control problems was studied by [10–13].

Let us mention that it is inadequate to only focus on exact optimality. As is well known, optimal controls may not exist in many situations, and insisting on exact optimality is not only unrealistic but also unnecessary for many real systems. Let us give an example to show that optimal control may not exist even in deterministic optimal delay control problems. The system evolves by $X_t = \int_0^t u_{s-\delta}\,ds$ for $0 \le t \le 1$, where $\delta = 1/4$ and $u_\cdot$ is chosen from the admissible control set $\mathcal{U}$, which is the collection of measurable functions $u : [0,1] \to \{-1,1\}$. We assume that $u_t = 1$ for $-\delta \le t < -\delta/2$ and $u_t = -1$ for $-\delta/2 \le t < 0$ for any $u_\cdot \in \mathcal{U}$. The objective is to minimize $J(u_\cdot) = \int_\delta^1 (X_t)^2\,dt$ over $\mathcal{U}$. Let us show that $\inf_{u_\cdot \in \mathcal{U}} J(u_\cdot) = 0$.

Firstly, $X_\delta = 0$. Then define a sequence of admissible controls $\{u_t^n\}$, $0 \le t \le 3\delta$ by $u_t^n = (-1)^k$, $k/(4n) \le t \le (k+1)/(4n)$, $0 \le k \le 3n-1$. Then the corresponding trajectory $X_\cdot^n$ satisfies $|X_t^n| \le 1/(4n)$ for $\delta \le t \le 1$. Thus, $J(u_\cdot^n) \le 3/(64n^2)$ and so $\inf_{u_\cdot \in \mathcal{U}} J(u_\cdot) = 0$. However, there does not exist $u_\cdot^* \in \mathcal{U}$ satisfying $J(u_\cdot^*) = 0$; otherwise, we have $X_t^* = 0$ for $\delta \le t \le 1$, which implies $u_t^* = 0$ for $0 \le t \le 3\delta$, contradicting the definition of the admissible control.

As stated in [14], near-optimality has as many attractive features as exact optimality in view of both theory and applications. First, near-optimal controls may exist under mild assumptions. Second, by studying near-optimality it is possible to greatly simplify the optimization process with only a small loss in the objective of the decision makers, and a near-optimal solution can satisfactorily serve the ultimate purpose of the decision makers in most practical situations. Third, many more near-optimal controls are available than optimal ones, so it is possible to select among them appropriate ones that are easier for analysis and implementation.

Near-optimality for deterministic control problems was studied in [15–17]. Near-optimality for one kind of stochastic control problem with controlled diffusion and non-convex control domain was studied in [14], for which necessary and sufficient conditions of near-optimality were established. Following [14], various kinds of near-optimal stochastic control problems have been investigated; see for example [18–23] for forward control systems, and [24–29] for forward-backward systems.

In view of the importance and wide applicability of time-delay systems and near-optimality, this paper is the first attempt to study near-optimization for one kind of stochastic delay control problem. In the control problem, the control domain is convex, the control variable can enter the diffusion term of the control system, and both the state and the control variables involve delays. For simplicity and clarity, we only consider linear systems. Necessary as well as sufficient conditions for a control to be near-optimal are established. By using the maximum principle and Ekeland's variational principle, we first establish a necessary condition for near-optimality, which reveals the 'minimum' qualification for an admissible control to be ε-optimal. Then we prove a sufficient verification theorem for near-optimality, which can help to verify whether a candidate control is indeed near-optimal and thus can help to find near-optimal controls. Finally, the theoretical results are applied to some illustrative examples.

The main features of this paper are as follows. This is the first attempt to study near-optimal controls of stochastic delay control problems with the maximum principle method and by means of ABSDEs. We establish necessary and sufficient conditions for any near-optimal controls and give some examples. Since exact optimal control could be regarded as a particular case of ε-optimal control when $\varepsilon = 0$, this paper is a generalization of [10] in the linear system case. We give two sufficient conditions for near-optimality, which cannot contain each other in general. The functions l and Φ in the cost functional can be quadratic functions of x, which generalizes the corresponding assumptions in [14, 18, 21] and some other papers. In most existing literature, the error bound in the necessary condition for an admissible control to be ε-optimal is ε^γ with $\gamma \in [0, \frac{1}{3})$ or $\gamma \in [0, \frac{1}{3}]$, while it is improved in this paper to ε^γ with $\gamma \in [0, \frac{1}{2}]$. In two illustrative examples, we give some near-optimal controls in the explicit form.

The rest of this paper is organized as follows. In Section 2, we give the formulation of the problem and present some preliminaries. We establish the necessary conditions for near-optimal controls in Section 3 and the sufficient conditions in Section 4. The theoretical results are applied to two examples in Section 5 and a conclusion is given in Section 6.

2 Formulation of the problem and preliminaries

For $n \geq 1$, we use $\mathbb{R}^n$ to denote the n-dimensional Euclidean space with the usual norm $|\cdot|$ and inner product $\langle\cdot,\cdot\rangle$. Denote by A^T the transpose of a matrix A. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and $\mathbb{E}$ the expectation with respect to $\mathbb{P}$. By $\{\mathcal{F}_t, t\geq 0\}$ we denote the completed natural filtration of a standard Brownian motion $\{W_t, t \geq 0\}$, which is assumed to be scalar-valued for simplicity. For $a < b$, denote by $M^2(a,b;\mathbb{R}^n)$ the set of n-dimensional adapted processes $\{\phi_t, a\leq t\leq b\}$ satisfying $\mathbb{E}\int_a^b |\phi_t|^2\,dt < \infty$, and by $S^2(a,b;\mathbb{R}^n)$ the set of n-dimensional continuous adapted processes $\{\psi_t, a \leq t \leq b\}$ satisfying $\mathbb{E}[\sup_{a\leq t\leq b}|\psi_t|^2] < \infty$. We use C, C', C'' to represent positive constants, which can be different from line to line.

Assume that δ_1 and δ_2 are positive constants, and $\xi_\cdot : [-\delta_1, 0] \to \mathbb{R}^n$ is a continuous function. Given a bounded convex set $U \subset \mathbb{R}^k$ and a measurable function $\eta_\cdot : [-\delta_2, 0] \to U$, we define the admissible control set $\mathcal{U}$ as the collection of U-valued adapted processes $\{v_t, -\delta_2 \leq t \leq T\}$ satisfying $v_t = \eta_t$ for $-\delta_2 \leq t \leq 0$. For $v_\cdot \in \mathcal{U}$, the controlled system evolves by

$$\begin{cases} dX_t^v = b(t,X_t^v,X_{t-\delta_1}^v,v_t,v_{t-\delta_2})\,dt + \sigma(t,X_t^v,X_{t-\delta_1}^v,v_t,v_{t-\delta_2})\,dW_t, & 0\leq t\leq T, \\ X_t = \xi_t, \quad -\delta_1 \leq t \leq 0, \end{cases} \tag{1}$$

with

$$b(t,x,x_\delta,v,v_\delta) = A_1(t)x + B_1(t)x_\delta + C_1(t)v + D_1(t)v_\delta + E_1(t),$$
$$\sigma(t,x,x_\delta,v,v_\delta) = A_2(t)x + B_2(t)x_\delta + C_2(t)v + D_2(t)v_\delta + E_2(t),$$

where the coefficients $A_i(\cdot)$, $B_i(\cdot)$, $C_i(\cdot)$, $D_i(\cdot)$, $i = 1,2$, are bounded adapted processes with appropriate dimensions, and $E_1(\cdot)$, $E_2(\cdot) \in M^2(0,T;\mathbb{R}^n)$. The solution $X_\cdot^v$ of SDDE (1) is called the response of the control $v_\cdot$, and $(X_\cdot^v, v_\cdot)$ is called an admissible pair. The cost functional is given by

$$J(v_\cdot) = \mathbb{E}\left[\int_0^T l\big(t,X_t^v,X_{t-\delta_1}^v,v_t,v_{t-\delta_2}\big)\,dt + \Phi\big(X_T^v\big)\right], \quad v_\cdot \in \mathcal{U}, \tag{2}$$

where $l(\omega,t,x,x_\delta,v,v_\delta) : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times U\times U \to \mathbb{R}$ is an adapted function and $\Phi(\omega,x) : \Omega\times\mathbb{R}^n \to \mathbb{R}$ is a measurable function. The objective of our control problem is to find an admissible control $u_\cdot^* \in \mathcal{U}$ which satisfies

$$J\big(u_\cdot^*\big) = V \triangleq \inf_{v_\cdot\in\mathcal{U}} J(v_\cdot). \tag{3}$$

The following assumption will be in force throughout this paper.

(H1) The functions l and Φ are continuously differentiable in (x,x_δ,v,v_δ), and there exist a positive constant C and a continuous function $h(v,v_\delta) : U\times U \to \mathbb{R}$ such that the partial derivatives of l and Φ are bounded by $C(1 + |x| + |x_\delta| + h(v,v_\delta))$. Besides, $\Phi(0)$ is $\mathcal{F}_T$-measurable, and $\mathbb{E}|\Phi(0)| + \mathbb{E}\int_0^T |l(t,0,0,0,0)|\,dt < \infty$.

For later use, let us assume that $B_1(t)$ and $B_2(t)$ are well defined and bounded for $T < t \leq T + \delta_1$, $D_1(t)$ and $D_2(t)$ are well defined and bounded for $T < t \leq T + \delta_2$, $l_{x_\delta}(t, x, x_\delta, v, v_\delta) = 0$ for $T < t \leq T + \delta_1$, and $l_{v_\delta}(t, x, x_\delta, v, v_\delta) = 0$ for $T < t \leq T + \delta_2$.

By Theorem 2.1 in [11], SDDE (1) admits a unique solution $X^v_\cdot \in S^2(0, T; \mathbb{R}^n)$. Moreover, there exists $C > 0$ which is independent of $v_\cdot \in \mathcal{U}$ such that

$$\mathbb{E}\Big[\sup_{0\leq t\leq T} \big|X^v_t\big|^2\Big] \leq C, \quad \forall v_\cdot \in \mathcal{U}. \tag{4}$$

Then from (H1) it follows that J is well defined on $\mathcal{U}$ and there exists $C > 0$ which is independent of $v_\cdot \in \mathcal{U}$ such that $|J(v_\cdot)| \leq C$.

For the study of near-optimality, let us give the related definitions; see [14].

Definition 1 For $\varepsilon > 0$, $v^\varepsilon_\cdot \in \mathcal{U}$ is called ε-optimal if $|J(v^\varepsilon_\cdot) - V| \leq \varepsilon$. A family of admissible controls $\{v^\varepsilon_\cdot\}$ parameterized by $\varepsilon > 0$ is called near-optimal if $|J(v^\varepsilon_\cdot) - V| \leq r(\varepsilon)$ holds for sufficiently small ε, where $r(\varepsilon) \to 0$ as $\varepsilon \to 0$. If the error bound $r(\varepsilon)$ satisfies $r(\varepsilon) = c\varepsilon^\gamma$ for some $\gamma > 0$ independent of c, then $v^\varepsilon_\cdot$ is called near-optimal with order ε^γ.

Denote $\Theta^v_t = (t, X^v_t, X^v_{t-\delta_1}, v_t, v_{t-\delta_2})$. Let us introduce the following adjoint equation:

$$\begin{cases} dY^v_t = -\{\mathbb{E}^{\mathcal{F}_t}[B_1(t+\delta_1)^T Y^v_{t+\delta_1} + B_2(t+\delta_1)^T Z^v_{t+\delta_1} + l_{x_\delta}(\Theta^v_{t+\delta_1})] \\ \qquad\qquad + A_1(t)^T Y^v_t + A_2(t)^T Z^v_t + l_x(\Theta^v_t)\}\,dt + Z^v_t\,dW_t, \quad 0 \leq t \leq T, \\ Y^v_T = \Phi_x(X^v_T), \\ Y^v_t = 0, \qquad Z^v_t = 0, \quad T < t \leq T + \delta_1, \end{cases} \tag{5}$$

whose solution is defined to be a pair of processes $(Y^v_\cdot, Z^v_\cdot) \in M^2(0, T; \mathbb{R}^n) \times M^2(0, T; \mathbb{R}^n)$ satisfying (5). Let us assume w.o.l.g. that Y^v_t and Z^v_t vanish for $T < t \leq T + \max\{\delta_1, \delta_2\}$ for all $v_\cdot \in \mathcal{U}$.

Proposition 2 *Assume* (H1). *Then the adjoint equation* (5) *admits a unique solution* $(Y^v_\cdot, Z^v_\cdot)$ *for any* $v_\cdot \in \mathcal{U}$. *Moreover, there exists* $C > 0$ *which is independent of* $v_\cdot \in \mathcal{U}$ *such that*

$$\mathbb{E}\Big[\sup_{0\leq t\leq T} \big|Y^v_t\big|^2 + \int_0^T \big|Z^v_t\big|^2\,dt\Big] \leq C, \quad \forall v_\cdot \in \mathcal{U}. \tag{6}$$

Proof Set

$$\begin{aligned} g(t, y, z, \zeta_s, \kappa_r) &= A_1(t)^T y + A_2(t)^T z + l_x\big(\Theta^v_t\big) \\ &\quad + \mathbb{E}^{\mathcal{F}_t}\big[B_1(t+\delta_1)^T \zeta_s + B_2(t+\delta_1)^T \kappa_r + l_{x_\delta}\big(\Theta^v_{t+\delta_1}\big)\big]. \end{aligned}$$

First, g is Lipschitz continuous in $(y, z, \zeta_s, \kappa_r)$, so the assumption (H1) in [9] is satisfied. Next, we have

$$\mathbb{E}\int_0^T \big|g(t, 0, 0, 0, 0)\big|^2\,dt \leq 2\mathbb{E}\int_0^T \big|l_x\big(\Theta^v_t\big)\big|^2\,dt + 2\mathbb{E}\int_0^T \big|\mathbb{E}^{\mathcal{F}_t}\big[l_{x_\delta}\big(\Theta^v_{t+\delta_1}\big)\big]\big|^2\,dt.$$

Using Jensen's inequality, Fubini's theorem and a change of variables lead to

$$\mathbb{E}\int_0^T \left|\mathbb{E}^{\mathcal{F}_t}\left[l_{x_\delta}\left(\Theta_{t+\delta_1}^v\right)\right]\right|^2 dt \le \mathbb{E}\int_{\delta_1}^{T+\delta_1} \left|l_{x_\delta}\left(\Theta_t^v\right)\right|^2 dt \le \mathbb{E}\int_0^{T+\delta_1} \left|l_{x_\delta}\left(\Theta_t^v\right)\right|^2 dt.$$

Since it is assumed that $l_{x_\delta}(t,x,x_\delta,v,v_\delta) = 0$ for $T < t \le T + \delta_1$, we have

$$\mathbb{E}\int_0^{T+\delta_1} \left|l_{x_\delta}\left(\Theta_t^v\right)\right|^2 dt = \mathbb{E}\int_0^T \left|l_{x_\delta}\left(\Theta_t^v\right)\right|^2 dt.$$

Thus,

$$\mathbb{E}\int_0^T \left|g(t,0,0,0,0)\right|^2 dt \le 2\mathbb{E}\int_0^T \left(\left|l_x\left(\Theta_t^v\right)\right|^2 + \left|l_{x_\delta}\left(\Theta_t^v\right)\right|^2\right) dt.$$

Recall that U is a bounded set. Then, in view of (H1), we can use (4) to show that there exists $C > 0$, which is independent of $v_\cdot$, such that

$$\mathbb{E}\int_0^T \left|g(t,0,0,0,0)\right|^2 dt \le C.$$

Besides, $\mathbb{E}|\Phi_x(X_T^v)|^2 \le C'\mathbb{E}(1 + |X_T^v|^2) \le C$. Consequently, by Theorem 4.2 in [9] we conclude that (5) admits a unique solution. Finally, the estimate (6) can easily be obtained by Proposition 4.4 in [9]. □

Let us define a metric d on $\mathcal{U}$ by

$$d(u_\cdot,v_\cdot) = \sqrt{\mathbb{E}\int_0^T |u_t - v_t|^2 dt}.$$

Then it is well known that $(\mathcal{U},d)$ is a complete metric space.

Next result gives the continuity of $X_\cdot^v$ in $v_\cdot \in \mathcal{U}$.

Proposition 3 *Assume* (H1). *Then there exists* $C > 0$ *satisfying*

$$\mathbb{E}\Big[\sup_{0\le t\le T}\left|X_t^u - X_t^v\right|^2\Big] \le Cd(u_\cdot,v_\cdot)^2, \quad \forall u_\cdot,v_\cdot \in \mathcal{U}.$$

Proof Using the estimate (3) in [11], we get

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le t\le T}\left|X_t^u - X_t^v\right|^2\Big] &\le \mathbb{E}\int_0^T \left|b\left(\Theta_t^u\right) - b\left(t,X_t^u,X_{t-\delta_1}^u,v_t,v_{t-\delta_2}\right)\right|^2 dt \\ &\quad + \mathbb{E}\int_0^T \left|\sigma\left(\Theta_t^u\right) - \sigma\left(t,X_t^u,X_{t-\delta_1}^u,v_t,v_{t-\delta_2}\right)\right|^2 dt.\end{aligned}$$

Then it follows that

$$\mathbb{E}\Big[\sup_{0\le t\le T}\left|X_t^u - X_t^v\right|^2\Big] \le C\mathbb{E}\int_0^T |u_t - v_t|^2 dt + C\mathbb{E}\int_0^T |u_{t-\delta_2} - v_{t-\delta_2}|^2 dt,$$

where by the definition of admissible controls, we can use a change of variables to get

$$\mathbb{E}\int_0^T |u_{t-\delta_2} - v_{t-\delta_2}|^2\,dt = \mathbb{E}\int_{-\delta_2}^{T-\delta_2} |u_t - v_t|^2\,dt \le \mathbb{E}\int_0^T |u_t - v_t|^2\,dt.$$

Thus, the proof is complete. □

Let us assume, moreover,

(H2) $(\Phi_x, l_x, l_{x_\delta}, l_v, l_{v_\delta})$ are Lipschitz in $(x, x_\delta, v, v_\delta)$.

The following result shows that $(Y^v_\cdot, Z^v_\cdot)$ is continuous in $v_\cdot \in \mathcal{U}$.

Proposition 4 *Assume* (H1) *and* (H2). *Then there exists* $C > 0$ *such that*

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|Y^u_t - Y^v_t\right|^2 + \int_0^T \left|Z^u_t - Z^v_t\right|^2 dt\right] \le Cd(u_\cdot, v_\cdot)^2, \quad \forall u_\cdot, v_\cdot \in \mathcal{U}.$$

Proof Set $\bar{Y}_t = Y^u_t - Y^v_t$, $\bar{Z}_t = Z^u_t - Z^v_t$. Let us prove by dividing $[0,T]$ backwardly. Firstly, for $t \in I_1 = [T-\delta_1, T]$, $(\bar{Y}_t, \bar{Z}_t)$ solves a linear BSDE

$$\begin{cases} -d\bar{Y}_t = \{A_1(t)^T\bar{Y}_t + A_2(t)^T\bar{Z}_t + l_x(\Theta^u_t) - l_x(\Theta^v_t)\}\,dt - \bar{Z}_t\,dW_t, \\ \bar{Y}_T = \Phi_x(X^u_T) - \Phi_x(X^v_T). \end{cases}$$

The basic prior estimate of BSDEs gives

$$\mathbb{E}\left[\sup_{t\in I_1}|\bar{Y}_t|^2 + \int_{t\in I_1}|\bar{Z}_t|^2\,dt\right] \le C\mathbb{E}\left[\left|\Phi_x\left(X^u_T\right) - \Phi_x\left(X^v_T\right)\right|^2 + \int_{t\in I_1}\left|l_x\left(\Theta^u_t\right) - l_x\left(\Theta^v_t\right)\right|^2 dt\right].$$

Then, in view of (H2), using Proposition 3 and a change of variables lead to

$$\mathbb{E}\left[\sup_{t\in I_1}|\bar{Y}_t|^2 + \int_{t\in I_1}|\bar{Z}_t|^2\,dt\right] \le Cd(u_\cdot, v_\cdot)^2. \tag{7}$$

Next, on $I_2 = [T-2\delta_1, T-\delta_1]$, $(\bar{Y}_\cdot, \bar{Z}_\cdot)$ solves a BSDE with terminal value $\bar{Y}_{T-\delta_1}$ and generator function $f(t,y,z) = A_1(t)^T y + A_2(t)^T z + \Delta(t)$, where $\Delta(t) = l_x(\Theta^u_t) - l_x(\Theta^v_t) + \mathbb{E}^{\mathcal{F}_t}[B_1(t+\delta_1)^T\bar{Y}_{t+\delta_1} + B_2(t+\delta_1)^T\bar{Z}_{t+\delta_1} + l_{x_\delta}(\Theta^u_{t+\delta_1}) - l_{x_\delta}(\Theta^v_{t+\delta_1})]$. On the one hand, by (7), $\mathbb{E}|\bar{Y}_{T-\delta_1}|^2 \le Cd(u_\cdot, v_\cdot)^2$. On the other hand, by Jensen's inequality and a change of variables we get

$$\begin{aligned}\mathbb{E}\int_{t\in I_2}\left|\Delta(t)\right|^2 dt &\le C\mathbb{E}\int_{t\in I_2}\left|l_x\left(\Theta^u_t\right) - l_x\left(\Theta^v_t\right)\right|^2 dt \\ &\quad + C\mathbb{E}\int_{t\in I_1}\left(|\bar{Y}_t|^2 + |\bar{Z}_t|^2 + \left|l_{x_\delta}\left(\Theta^u_t\right) - l_{x_\delta}\left(\Theta^v_t\right)\right|^2\right) dt.\end{aligned}$$

Then, by (H2), (7) and Proposition 3, we can use a change of variables again to get

$$\mathbb{E}\int_{t\in I_2}\left|\Delta(t)\right|^2 dt \le Cd(u_\cdot, v_\cdot)^2.$$

So,

$$\mathbb{E}\left[\sup_{t\in I_2}|\bar{Y}_t|^2 + \int_{t\in I_2}|\bar{Z}_t|^2\,dt\right] \le C'\mathbb{E}\left[|\bar{Y}_{T-\delta_1}|^2 + \int_{t\in I_2}\left|\Delta(t)\right|^2 dt\right] \le Cd(u_\cdot, v_\cdot)^2.$$

Thus, we derive

$$\mathbb{E}\left[\sup_{T-2\delta_1\le t\le T}|\bar{Y}_t|^2+\int_{T-2\delta_1}^{T}|\bar{Z}_t|^2\,dt\right]\le Cd(u.,v.)^2.$$

In the same way, we can get the result after finite steps. □

Next we prove that J is a continuous functional of $v.\in\mathcal{U}$.

Proposition 5 *Assume* (H1). *Then there exists* $C>0$ *such that* $|J(u.)-J(v.)|\le Cd(u.,v.)$ *holds for all* $u.,v.\in\mathcal{U}$.

Proof Set $\bar{X}_t=X_t^u-X_t^v$, $\bar{v}_t=u_t-v_t$. We have

$$\Phi(X_T^u)-\Phi(X_T^v)=\int_0^1\langle\Phi_x(X_T^v+\lambda\bar{X}_T),\bar{X}_T\rangle\,d\lambda,$$

$$l(\Theta_t^u)-l(\Theta_t^v)=\int_0^1\{\langle l_x(\Lambda_t),\bar{X}_t\rangle+\langle l_{x_\delta}(\Lambda_t),\bar{X}_{t-\delta_1}\rangle+\langle l_v(\Lambda_t),\bar{v}_t\rangle+\langle l_{v_\delta}(\Lambda_t),\bar{v}_{t-\delta_2}\rangle\}\,d\lambda,$$

with $\Lambda_t=(t,X_t^v+\lambda\bar{X}_t,X_{t-\delta_1}^v+\lambda\bar{X}_{t-\delta_1},v_t+\lambda\bar{v}_t,v_{t-\delta_1}+\lambda\bar{v}_{t-\delta_1})$. By (H1), (4) and Proposition 3, we can use the Cauchy-Schwartz inequality to get

$$\mathbb{E}\left|\Phi(X_T^u)-\Phi(X_T^v)\right|\le Cd(u.,v.).$$

With a similar method, together with a change of variables, we have

$$\mathbb{E}\int_0^T\left|l(\Theta_t^u)-l(\Theta_t^v)\right|dt\le Cd(u.,v.).$$

Thus the proof is complete. □

The following Ekeland's variational principle will play a key role in what follows, for which one can see [15].

Lemma 6 *Let* (S,d) *be a complete metric space and* $F:S\to\mathbb{R}$ *a lower-semicontinuous and bounded from below function. Assume that* $v^\varepsilon\in S$ *satisfies* $F(v^\varepsilon)\le\inf_{v\in S}F(v)+\varepsilon$ *for some* $\varepsilon\ge 0$. *Then, for any* $\lambda>0$, *there exists* $v^\lambda\in S$ *such that* $F(v^\lambda)\le F(v^\varepsilon)$, $d(v^\lambda,v^\varepsilon)\le\lambda$, *and* $F(v^\lambda)\le F(v)+\frac{\varepsilon}{\lambda}d(v,v^\lambda)$ *for all* $v\in S$.

3 Necessary condition for near-optimality

This section is devoted to establishing necessary conditions for near-optimal controls of the stochastic control problem (1)-(3).

Recall from the previous section that $J(v.)$ is a continuous and bounded from below functional on the complete metric space $(\mathcal{U},d)$. Now let $u^\varepsilon_\cdot\in\mathcal{U}$ be an ε-optimal control of problem (1)-(3) with $\varepsilon>0$, that is, $J(u^\varepsilon_\cdot)\le\inf_{v.\in\mathcal{U}}J(v.)+\varepsilon$. Then applying Lemma 6 with $\lambda=\sqrt{\varepsilon}$ leads to the existence of $\tilde{u}^\varepsilon_\cdot\in\mathcal{U}$ such that

$$J(\tilde{u}^\varepsilon_\cdot)\le J(u^\varepsilon_\cdot),\tag{8}$$

$$d(\tilde{u}^{\varepsilon}_{\cdot}, u^{\varepsilon}_{\cdot}) \leq \sqrt{\varepsilon}, \tag{9}$$

$$J(\tilde{u}^{\varepsilon}_{\cdot}) \leq J(v_{\cdot}) + \sqrt{\varepsilon} d(v_{\cdot}, \tilde{u}^{\varepsilon}_{\cdot}), \quad \forall v_{\cdot} \in \mathcal{U}. \tag{10}$$

In what follows, we first study $\tilde{u}^{\varepsilon}_{\cdot}$, and then turn to $u^{\varepsilon}_{\cdot}$. Let $u_{\cdot} \in M^2(-\delta_2, T)$ satisfy $\tilde{u}^{\varepsilon}_{\cdot} + u_{\cdot} \in \mathcal{U}$. Then it is easy to see that $u_t = 0$ for $-\delta_2 \leq t < 0$, and the convexity of $\mathcal{U}$ shows that $u^{\theta}_{\cdot} \triangleq \tilde{u}^{\varepsilon}_{\cdot} + \theta u_{\cdot} \in \mathcal{U}$ for any $\theta \in [0,1]$. Since U is bounded, there exists $C > 0$, independent of ε and θ, such that $d(u^{\theta}_{\cdot}, \tilde{u}^{\varepsilon}_{\cdot}) \leq C\theta$. So (10) leads to

$$J(u^{\theta}_{\cdot}) - J(\tilde{u}^{\varepsilon}_{\cdot}) \geq -C\sqrt{\varepsilon}\theta. \tag{11}$$

Let $X^{\varepsilon}_{\cdot}$, $\tilde{X}^{\varepsilon}_{\cdot}$, $X^{\theta}_{\cdot}$ be, respectively, the trajectories associated with $u^{\varepsilon}_{\cdot}$, $\tilde{u}^{\varepsilon}_{\cdot}$, $u^{\theta}_{\cdot}$. Let $(Y^{\varepsilon}_{\cdot}, Z^{\varepsilon}_{\cdot})$ and $(\tilde{Y}^{\varepsilon}_{\cdot}, \tilde{Z}^{\varepsilon}_{\cdot})$ be, respectively, the solutions of the adjoint equation (5) associated with $(u^{\varepsilon}_{\cdot}, X^{\varepsilon}_{\cdot})$ and $(\tilde{u}^{\varepsilon}_{\cdot}, \tilde{X}^{\varepsilon}_{\cdot})$. Set $\Theta^{\varepsilon}_t = (t, X^{\varepsilon}_t, X^{\varepsilon}_{t-\delta_1}, u^{\varepsilon}_t, u^{\varepsilon}_{t-\delta_2})$, $\tilde{\Theta}^{\varepsilon}_t = (t, \tilde{X}^{\varepsilon}_t, \tilde{X}^{\varepsilon}_{t-\delta_1}, \tilde{u}^{\varepsilon}_t, \tilde{u}^{\varepsilon}_{t-\delta_2})$, $\Theta^{\theta}_t = (t, X^{\theta}_t, X^{\theta}_{t-\delta_1}, u^{\theta}_t, u^{\theta}_{t-\delta_2})$. Let us introduce the following variational equation:

$$\begin{cases} dX^1_t = [A_1(t)X^1_t + B_1(t)X^1_{t-\delta_1} + C_1(t)u_t + D_1(t)u_{t-\delta_2}]\,dt \\ \qquad\quad + [A_2(t)X^1_t + B_2(t)X^1_{t-\delta_1} + C_2(t)u_t + D_2(t)u_{t-\delta_2}]\,dW_t, \quad 0 \leq t \leq T, \\ X^1_t = 0, \quad -\delta_1 \leq t \leq 0. \end{cases} \tag{12}$$

It is easy to check that (12) admits a unique solution $X^1_{\cdot} \in S^2(0, T; \mathbb{R}^n)$.

The following result is a necessary condition for $\tilde{u}^{\varepsilon}_{\cdot}$.

Proposition 7 *Assume* (H1)-(H2). *Then there exists* $C > 0$, *independent of* ε, *such that*

$$\begin{aligned} \mathbb{E}\int_0^T \big\langle \mathbb{E}^{\mathcal{F}_t}\big[D_1(t+\delta_2)^T \tilde{Y}^{\varepsilon}_{t+\delta_2} + D_2(t+\delta_2)^T \tilde{Z}^{\varepsilon}_{t+\delta_2} + l_{v_\delta}\big(\tilde{\Theta}^{\varepsilon}_{t+\delta_2}\big)\big] \\ + C_1(t)^T \tilde{Y}^{\varepsilon}_t + C_2(t)^T \tilde{Z}^{\varepsilon}_t + l_v\big(\tilde{\Theta}^{\varepsilon}_t\big), v - \tilde{u}^{\varepsilon}_t\big\rangle\,dt \geq -C\sqrt{\varepsilon}, \quad \forall v \in U. \end{aligned} \tag{13}$$

Proof Following the proof of Lemma 3.3 in [11], we have

$$\begin{aligned} \lim_{\theta \downarrow 0} \theta^{-1}\big[J(u^{\theta}_{\cdot}) - J(\tilde{u}^{\varepsilon}_{\cdot})\big] = \mathbb{E}\bigg\{\big\langle \Phi_x(\tilde{X}^{\varepsilon}_T), X^1_T\big\rangle + \int_0^T \big[\big\langle l_x\big(\tilde{\Theta}^{\varepsilon}_t\big), X^1_t\big\rangle \\ + \big\langle l_{x_\delta}\big(\tilde{\Theta}^{\varepsilon}_t\big), X^1_{t-\delta_1}\big\rangle + \big\langle l_v\big(\tilde{\Theta}^{\varepsilon}_t\big), u_t\big\rangle + \big\langle l_{v_\delta}\big(\tilde{\Theta}^{\varepsilon}_t\big), u_{t-\delta_2}\big\rangle\big]\,dt\bigg\}. \end{aligned}$$

Using a change of variables leads to

$$\mathbb{E}\int_0^T \big\langle l_{x_\delta}\big(\tilde{\Theta}^{\varepsilon}_t\big), X^1_{t-\delta_1}\big\rangle\,dt = \mathbb{E}\int_{-\delta_1}^{T-\delta_1} \big\langle \mathbb{E}^{\mathcal{F}_t}\big[l_{x_\delta}\big(\tilde{\Theta}^{\varepsilon}_{t+\delta_1}\big)\big], X^1_t\big\rangle\,dt.$$

Then, since $X^1_t = 0$ for $-\delta_1 \leq t < 0$ and $l_{x_\delta}(t, x, x_\delta, v, v_\delta) = 0$ for $T < t \leq T + \delta_1$, we have

$$\mathbb{E}\int_0^T \big\langle l_{x_\delta}\big(\tilde{\Theta}^{\varepsilon}_t\big), X^1_{t-\delta_1}\big\rangle\,dt = \mathbb{E}\int_0^T \big\langle \mathbb{E}^{\mathcal{F}_t}\big[l_{x_\delta}\big(\tilde{\Theta}^{\varepsilon}_{t+\delta_1}\big)\big], X^1_t\big\rangle\,dt.$$

Similarly,

$$\mathbb{E}\int_0^T \big\langle l_{v_\delta}\big(\tilde{\Theta}^{\varepsilon}_t\big), u_{t-\delta_2}\big\rangle\,dt = \mathbb{E}\int_0^T \big\langle \mathbb{E}^{\mathcal{F}_t}\big[l_{v_\delta}\big(\tilde{\Theta}^{\varepsilon}_{t+\delta_2}\big)\big], u_t\big\rangle\,dt.$$

Consequently, from (11) it follows that

$$\begin{aligned}
&\mathbb{E}\langle \Phi_x(\tilde{X}_T^\varepsilon), X_T^1\rangle + \mathbb{E}\int_0^T \langle l_x(\tilde{\Theta}_t^\varepsilon) + \mathbb{E}^{\mathcal{F}_t}[l_{x_\delta}(\tilde{\Theta}_{t+\delta_1}^\varepsilon)], X_t^1\rangle\,dt \\
&\quad + \mathbb{E}\int_0^T \langle l_v(\tilde{\Theta}_t^\varepsilon) + \mathbb{E}^{\mathcal{F}_t}[l_{v_\delta}(\tilde{\Theta}_{t+\delta_2}^\varepsilon)], u_t\rangle\,dt \ge -C\sqrt{\varepsilon}.
\end{aligned} \tag{14}$$

On the other hand, applying Itô's formula to $\langle X_t^1, \tilde{Y}_t^\varepsilon\rangle$ gives

$$\begin{aligned}
&\mathbb{E}\langle \Phi_x(\tilde{X}_T^\varepsilon), X_T^1\rangle + \mathbb{E}\int_0^T \langle l_x(\tilde{\Theta}_t^\varepsilon) + \mathbb{E}^{\mathcal{F}_t}[l_{x_\delta}(\tilde{\Theta}_{t+\delta_1}^\varepsilon)], X_t^1\rangle\,dt \\
&\quad = \mathbb{E}\int_0^T \big(\langle B_1(t)^T\tilde{Y}_t^\varepsilon, X_{t-\delta_1}^1\rangle - \langle \mathbb{E}^{\mathcal{F}_t}[B_1(t+\delta_1)^T\tilde{Y}_{t+\delta_1}^\varepsilon], X_t^1\rangle\big)\,dt \\
&\qquad + \mathbb{E}\int_0^T \big(\langle B_2(t)^T\tilde{Z}_t^\varepsilon, X_{t-\delta_1}^1\rangle - \langle \mathbb{E}^{\mathcal{F}_t}[B_2(t+\delta_1)^T\tilde{Z}_{t+\delta_1}^\varepsilon], X_t^1\rangle\big)\,dt \\
&\qquad + \mathbb{E}\int_0^T \langle C_1(t)^T\tilde{Y}_t^\varepsilon + C_2(t)^T\tilde{Z}_t^\varepsilon, u_t\rangle\,dt + \mathbb{E}\int_0^T \langle D_1(t)^T\tilde{Y}_t^\varepsilon + D_2^T(t)\tilde{Z}_t^\varepsilon, u_{t-\delta_2}\rangle\,dt.
\end{aligned}$$

Then we can use a change of variables to get

$$\begin{aligned}
&\mathbb{E}\langle \Phi_x(\tilde{X}_T^\varepsilon), X_T^1\rangle + \mathbb{E}\int_0^T \langle l_x(\tilde{\Theta}_t^\varepsilon) + \mathbb{E}^{\mathcal{F}_t}[l_{x_\delta}(\tilde{\Theta}_{t+\delta_1}^\varepsilon)], X_t^1\rangle\,dt \\
&\quad = \mathbb{E}\int_0^T \langle C_1(t)^T\tilde{Y}_t^\varepsilon + C_2(t)^T\tilde{Z}_t^\varepsilon + \mathbb{E}^{\mathcal{F}_t}[D_1(t+\delta_2)^T\tilde{Y}_{t+\delta_2}^\varepsilon + D_2(t+\delta_2)^T\tilde{Z}_{t+\delta_2}^\varepsilon], u_t\rangle\,dt.
\end{aligned}$$

Combining this equality and (14) gives

$$\begin{aligned}
&\mathbb{E}\int_0^T \langle \mathbb{E}^{\mathcal{F}_t}[D_1(t+\delta_2)^T\tilde{Y}_{t+\delta_2}^\varepsilon + D_2(t+\delta_2)^T\tilde{Z}_{t+\delta_2}^\varepsilon + l_{v_\delta}(\tilde{\Theta}_{t+\delta_2}^\varepsilon)] \\
&\quad + C_1(t)^T\tilde{Y}_t^\varepsilon + C_2(t)^T\tilde{Z}_t^\varepsilon + l_v(\tilde{\Theta}_t^\varepsilon), u_t\rangle\,dt \ge -C\sqrt{\varepsilon}.
\end{aligned}$$

Recall that $u_\cdot$ is any process in $M^2(-\delta_2, T)$ satisfying $\tilde{u}_\cdot^\varepsilon + u_\cdot \in \mathcal{U}$. For any $v \in U$, let us define $v_t = v$ when $0 < t \le T$ and $v_t = \eta_t$ when $-\delta_2 \le t \le 0$. Replacing u_t in the previous inequality with $v_t - \tilde{u}_t^\varepsilon$ leads to the conclusion. □

Let us define

$$\begin{aligned}
&H(t,x,x_\delta,y,z,v,v_\delta) = \langle b(t,x,x_\delta,v,v_\delta), y\rangle + \langle \sigma(t,x,x_\delta,v,v_\delta), z\rangle + l(t,x,x_\delta,v,v_\delta), \\
&\mathcal{H}^\varepsilon(t,v) = H(\Sigma_t^\varepsilon, v, u_{t-\delta_2}^\varepsilon) + \mathbb{E}^{\mathcal{F}_t}[H(\Sigma_{t+\delta_2}^\varepsilon, u_{t+\delta_2}^\varepsilon, v)],
\end{aligned}$$

where $\Sigma_t^\varepsilon = (t, X_t^\varepsilon, X_{t-\delta_1}^\varepsilon, Y_t^\varepsilon, Z_t^\varepsilon)$. By the dominated convergence theorem, $\mathcal{H}^\varepsilon(t,v)$ is differentiable in v and

$$\mathcal{H}_v^\varepsilon(t,v) = H_v(\Sigma_t^\varepsilon, v, u_{t-\delta_2}^\varepsilon) + \mathbb{E}^{\mathcal{F}_t}[H_{v_\delta}(\Sigma_{t+\delta_2}^\varepsilon, u_{t+\delta_2}^\varepsilon, v)].$$

We are now in a position to establish the necessary condition for near-optimal controls of the stochastic control problem (1)-(3).

Theorem 8 *Assume* (H1)-(H2). *There exists* $C > 0$ *such that for any* $\gamma \in [0, \frac{1}{2}]$, *any* $\varepsilon > 0$ *and any* ε*-optimal control pair* $(X^\varepsilon_\cdot, u^\varepsilon_\cdot)$ *of the stochastic control problem* (1)-(3), *we have*

$$\mathbb{E}\int_0^T \langle \mathcal{H}^\varepsilon_v(t, u^\varepsilon_t), v - u^\varepsilon_t \rangle \, dt \geq -C\varepsilon^\gamma, \quad \forall v \in U.$$

Proof The inequality is just

$$\begin{aligned}\mathbb{E}\int_0^T \langle &\mathbb{E}^{\mathcal{F}_t}\big[D_1(t+\delta_2)^T Y^\varepsilon_{t+\delta_2} + D_2(t+\delta_2)^T Z^\varepsilon_{t+\delta_2} + l_{v_\delta}\big(\Theta^\varepsilon_{t+\delta_2}\big)\big] \\ &+ C_1(t)^T Y^\varepsilon_t + C_2(t)^T Z^\varepsilon_t + l_v\big(\Theta^\varepsilon_t\big), v - u^\varepsilon_t \rangle \, dt \geq -C\varepsilon^\gamma, \quad \forall v \in U. \end{aligned} \tag{15}$$

In view of (13), we only need to show that the difference between the terms on the left-hand sides of (13) and (15) is not more than $C\varepsilon^\gamma$ for some constant C that is independent of ε and γ. Note that $\varepsilon < 1$, $\gamma \leq \frac{1}{2}$.

Firstly consider

$$\Gamma_1 = \mathbb{E}\int_0^T \big\{ \langle C_1(t)^T \tilde{Y}^\varepsilon_t, v - \tilde{u}^\varepsilon_t \rangle - \langle C_1(t)^T Y^\varepsilon_t, v - u^\varepsilon_t \rangle \big\} \, dt.$$

Note that $\Gamma_1 = \Gamma_{11} + \Gamma_{12}$ with

$$\Gamma_{11} = \mathbb{E}\int_0^T \langle C_1(t)^T \big(\tilde{Y}^\varepsilon_t - Y^\varepsilon_t\big), v - \tilde{u}^\varepsilon_t \rangle \, dt, \qquad \Gamma_{12} = \mathbb{E}\int_0^T \langle C_1(t)^T Y^\varepsilon_t, u^\varepsilon_t - \tilde{u}^\varepsilon_t \rangle \, dt.$$

Since U is bounded, there exists $C > 0$, which is independent of ε, such that $\Gamma_{11} \leq C\mathbb{E}\int_0^T |\tilde{Y}^\varepsilon_t - Y^\varepsilon_t| \, dt$. Then, by Proposition 4, applying the Cauchy-Schwartz inequality we get $\Gamma_{11} \leq Cd(u^\varepsilon_\cdot, \tilde{u}^\varepsilon_\cdot)$, and furthermore $\Gamma_{11} \leq C\sqrt{\varepsilon}$ due to (9). On the other hand, using the Cauchy-Schwartz inequality again, in view of (6) and (9), we get $\Gamma_{12} \leq C\sqrt{\varepsilon}$. Thus, $\Gamma_1 \leq C\sqrt{\varepsilon} \leq C\varepsilon^\gamma$. Similarly, we can prove

$$\mathbb{E}\int_0^T \big\{ \langle C_2(t)^T \tilde{Z}^\varepsilon_t, v - \tilde{u}^\varepsilon_t \rangle - \langle C_2(t)^T Z^\varepsilon_t, v - u^\varepsilon_t \rangle \big\} \, dt \leq C\sqrt{\varepsilon}.$$

Next, let us consider

$$\Gamma_2 = \mathbb{E}\int_0^T \big\{ \langle l_v\big(\tilde{\Theta}^\varepsilon_t\big), v - \tilde{u}^\varepsilon_t \rangle - \langle l_v\big(\Theta^\varepsilon_t\big), v - u^\varepsilon_t \rangle \big\} \, dt.$$

We have $\Gamma_2 = \Gamma_{21} + \Gamma_{22}$, where

$$\Gamma_{21} = \mathbb{E}\int_0^T \langle l_v\big(\tilde{\Theta}^\varepsilon_t\big) - l_v\big(\Theta^\varepsilon_t\big), v - \tilde{u}^\varepsilon_t \rangle \, dt, \qquad \Gamma_{22} = \mathbb{E}\int_0^T \langle l_v\big(\Theta^\varepsilon_t\big), u^\varepsilon_t - \tilde{u}^\varepsilon_t \rangle \, dt.$$

Note that

$$\Gamma_{21} \leq C\mathbb{E}\int_0^T \big| l_v\big(\tilde{\Theta}^\varepsilon_t\big) - l_v\big(\Theta^\varepsilon_t\big) \big| \, dt.$$

By (H2), we can use a change of variables and the Cauchy-Schwartz inequality to get

$$\Gamma_{21} \le C\sqrt{\mathbb{E}\int_0^T \left|\tilde{X}_t^\varepsilon - X_t^\varepsilon\right|^2 dt} + Cd\left(u_\cdot^\varepsilon, \tilde{u}_\cdot^\varepsilon\right).$$

Then by Proposition 3 and (9) we get $\Gamma_{21} \le C\sqrt{\varepsilon}$. Besides, (H1) gives

$$\Gamma_{22} \le C\mathbb{E}\int_0^T \left(1 + \left|X_t^\varepsilon\right| + \left|X_{t-\delta_1}^\varepsilon\right|\right)\left|u_t^\varepsilon - \tilde{u}_t^\varepsilon\right| dt,$$

so, by (4) and (9), we can use the Cauchy-Schwartz inequality again to get $\Gamma_{22} \le C\sqrt{\varepsilon}$. Thus, $\Gamma_2 \le C\sqrt{\varepsilon} \le C\varepsilon^\gamma$. Finally let us consider

$$\begin{aligned}\Gamma_3 = \mathbb{E}\int_0^T &\Big\{\Big\langle \mathbb{E}^{\mathcal{F}_t}\big[D_1(t+\delta_2)^T\tilde{Y}_{t+\delta_2}^\varepsilon + D_2(t+\delta_2)^T\tilde{Z}_{t+\delta_2}^\varepsilon + l_{v_\delta}\big(\tilde{\Theta}_{t+\delta_2}^\varepsilon\big)\big], v - \tilde{u}_t^\varepsilon\Big\rangle \\ &- \Big\langle \mathbb{E}^{\mathcal{F}_t}\big[D_1(t+\delta_2)^T Y_{t+\delta_2}^\varepsilon + D_2(t+\delta_2)^T Z_{t+\delta_2}^\varepsilon + l_{v_\delta}\big(\Theta_{t+\delta_2}^\varepsilon\big)\big], v - u_t^\varepsilon\Big\rangle\Big\} dt.\end{aligned}$$

In fact, by using Fubini's theorem, a change of variables and recalling our assumptions we get

$$\begin{aligned}\Gamma_3 &= \mathbb{E}\int_{\delta_2}^T \Big\{\Big\langle D_1(t)^T\tilde{Y}_t^\varepsilon + D_2(t)^T\tilde{Z}_t^\varepsilon + l_{v_\delta}\big(\tilde{\Theta}_t^\varepsilon\big), v - \tilde{u}_{t-\delta_2}^\varepsilon\Big\rangle \\ &\quad - \Big\langle D_1(t)^T Y_t^\varepsilon + D_2(t)^T Z_t^\varepsilon + l_{v_\delta}\big(\Theta_t^\varepsilon\big), v - u_{t-\delta_2}^\varepsilon\Big\rangle\Big\} dt \\ &= \mathbb{E}\int_{\delta_2}^T \Big\{\Big\langle D_1(t)^T\tilde{Y}_t^\varepsilon, v - \tilde{u}_{t-\delta_2}^\varepsilon\Big\rangle - \Big\langle D_1(t)^T Y_t^\varepsilon, v - u_{t-\delta_2}^\varepsilon\Big\rangle\Big\} dt \\ &\quad + \mathbb{E}\int_{\delta_2}^T \Big\{\Big\langle D_2(t)^T\tilde{Z}_t^\varepsilon, v - \tilde{u}_{t-\delta_2}^\varepsilon\Big\rangle - \Big\langle D_2(t)^T Z_t^\varepsilon, v - u_{t-\delta_2}^\varepsilon\Big\rangle\Big\} dt \\ &\quad + \mathbb{E}\int_{\delta_2}^T \Big\{\Big\langle l_{v_\delta}\big(\tilde{\Theta}_{t+\delta_2}^\varepsilon\big), v - \tilde{u}_{t-\delta_2}^\varepsilon\Big\rangle - \Big\langle l_{v_\delta}\big(\Theta_t^\varepsilon\big), v - u_{t-\delta_2}^\varepsilon\Big\rangle\Big\} dt.\end{aligned}$$

Then similar to Γ_1 and Γ_2 we have $\Gamma_3 \le C\varepsilon^\gamma$. Thus, (15) can be obtained, and the proof is complete. □

4 Sufficient conditions for near-optimality

In this section, we study under what conditions an admissible control turns out to be near-optimal. For this purpose, let us assume, moreover,

(H3) l and Φ are convex in $(x, x_\delta, v, v_\delta)$.

(H4) l is Lipschitz in (v, v_δ).

Theorem 9 *Let* $(X_\cdot^\varepsilon, u_\cdot^\varepsilon)$ *be an admissible pair and* $(Y_\cdot^\varepsilon, Z_\cdot^\varepsilon)$ *the corresponding solution of the adjoint equation* (5).

(i) *Assume* (H1)-(H3). *If* $u_\cdot^\varepsilon$ *satisfies*

$$\mathbb{E}\int_0^T \big\langle \mathcal{H}_v^\varepsilon\big(t, u_t^\varepsilon\big), v_t - u_t^\varepsilon\big\rangle dt \ge -\varepsilon, \quad \forall v_\cdot \in \mathcal{U}, \tag{16}$$

then $J(u_\cdot^\varepsilon) \le V + \varepsilon$.

(ii) *Assume* (H1)-(H4). *If* $u^\varepsilon_\cdot$ *satisfies*

$$\inf_{v_\cdot \in \mathcal{U}} \mathbb{E} \int_0^T \left[\mathcal{H}^\varepsilon(t, v_t) - \mathcal{H}^\varepsilon\left(t, u^\varepsilon_t\right)\right] dt \geq -\varepsilon^2, \tag{17}$$

then there exists $C' > 0$, *which is independent of* ε, *such that* $J(u^\varepsilon_\cdot) \leq V + C'\varepsilon$.

Proof For any $v_\cdot \in \mathcal{U}$, set $\hat{v}_t = v_t - u^\varepsilon_t$ and $\hat{X}_t = X^v_t - X^\varepsilon_t$. Applying Itô's formula to $\langle \hat{X}_t, Y^\varepsilon_t\rangle$ yields

$$\begin{aligned}
&\mathbb{E}\langle \Phi_x\left(X^\varepsilon_T\right), \hat{X}_T\rangle + \mathbb{E}\int_0^T \langle l_x\left(\Theta^\varepsilon_t\right) + \mathbb{E}^{\mathcal{F}_t}\left[l_{x_\delta}\left(\Theta^\varepsilon_{t+\delta_1}\right)\right], \hat{X}_t\rangle\, dt \\
&\quad = \mathbb{E}\int_0^T \left(\langle B_1(t)^T Y^\varepsilon_t, \hat{X}_{t-\delta_1}\rangle - \langle \mathbb{E}^{\mathcal{F}_t}\left[B_1(t+\delta_1)^T Y^\varepsilon_{t+\delta_1}\right], \hat{X}_t\rangle\right) dt \\
&\qquad + \mathbb{E}\int_0^T \left(\langle B_2(t)^T Z^\varepsilon_t, \hat{X}_{t-\delta_1}\rangle - \langle \mathbb{E}^{\mathcal{F}_t}\left[B_2(t+\delta_1)^T Z^\varepsilon_{t+\delta_1}\right], \hat{X}_t\rangle\right) dt \\
&\qquad + \mathbb{E}\int_0^T \langle C_1(t)^T Y^\varepsilon_t + C_2(t)^T Z^\varepsilon_t, \hat{v}_t\rangle\, dt + \mathbb{E}\int_0^T \langle D_1(t)^T Y^\varepsilon_t + D_2^T(t) Z^\varepsilon_t, \hat{v}_{t-\delta_2}\rangle\, dt.
\end{aligned}$$

Then by a change of variables we get

$$\begin{aligned}
&\mathbb{E}\langle \Phi_x\left(X^\varepsilon_T\right), \hat{X}_T\rangle + \mathbb{E}\int_0^T \left(\langle l_x\left(\Theta^\varepsilon_t\right), \hat{X}_t\rangle + \langle l_{x_\delta}\left(\Theta^\varepsilon_t\right), \hat{X}_{t-\delta_1}\rangle\right) dt \\
&\quad = \mathbb{E}\int_0^T \langle C_1(t)^T Y^\varepsilon_t + C_2(t)^T Z^\varepsilon_t + \mathbb{E}^{\mathcal{F}_t}\left[D_1(t+\delta_2)^T Y^\varepsilon_{t+\delta_2} + D_2(t+\delta_2)^T Z^\varepsilon_{t+\delta_2}\right], \hat{v}_t\rangle\, dt.
\end{aligned}$$

On the other hand, thanks to (H3), we can use a change of variables again to get

$$\begin{aligned}
J(v_\cdot) - J\left(u^\varepsilon_\cdot\right) \geq{}& \mathbb{E}\langle \Phi_x\left(X^\varepsilon_T\right), \hat{X}_T\rangle + \mathbb{E}\int_0^T \left(\langle l_x\left(\Theta^\varepsilon_t\right), \hat{X}_t\rangle + \langle l_{x_\delta}\left(\Theta^\varepsilon_t\right), \hat{X}_{t-\delta_1}\rangle\right) dt \\
&+ \mathbb{E}\int_0^T \langle l_v\left(\Theta^\varepsilon_t\right) + \mathbb{E}^{\mathcal{F}_t}\left[l_{v_\delta}\left(\Theta^\varepsilon_{t+\delta_2}\right)\right], \hat{v}_t\rangle\, dt.
\end{aligned}$$

Combining them gives

$$J(v_\cdot) - J\left(u^\varepsilon_\cdot\right) \geq \mathbb{E}\int_0^T \langle \mathcal{H}^\varepsilon_v\left(t, u^\varepsilon_t\right), \hat{v}_t\rangle\, dt. \tag{18}$$

So, if (16) holds, then $J(u^\varepsilon_\cdot) \leq J(v_\cdot) + \varepsilon$. Thus, the conclusion (i) follows from the arbitrariness of $v_\cdot \in \mathcal{U}$.

Next we prove (ii). To this end, define $\tilde{d}$ by

$$\tilde{d}\left(u_\cdot, u'_\cdot\right) = \mathbb{E}\int_0^T \nu^\varepsilon_t \left|u_t - u'_t\right| dt$$

with $\nu^\varepsilon_t = 1 + |Y^\varepsilon_t| + |Z^\varepsilon_t| + \mathbb{E}^{\mathcal{F}_t}[|Y^\varepsilon_{t+\delta_2}| + |Z^\varepsilon_{t+\delta_2}|]$. Then $(\mathcal{U}, \tilde{d})$ is a complete metric space. Define a new functional $f(u_\cdot) : \mathcal{U} \to \mathbb{R}$ by

$$f(u_\cdot) = \mathbb{E}\int_0^T \mathcal{H}^\varepsilon(t, u_t)\, dt.$$

Then, by (H4), there exists $L > 0$ such that $|f(u_\cdot) - f(u'_\cdot)| \leq L\tilde{d}(u_\cdot, u'_\cdot)$, which shows that f is continuous on $(\mathcal{U}, \tilde{d})$. Besides, the assumption (17) shows that

$$f(u^\varepsilon_\cdot) \leq \inf_{u_\cdot \in \mathcal{U}} f(u_\cdot) + \varepsilon^2.$$

Consequently, applying Lemma 6 for $\lambda = \varepsilon$ leads to the existence of $\tilde{u}^\varepsilon_\cdot \in \mathcal{U}$ which satisfies

$$\tilde{d}(\tilde{u}^\varepsilon_\cdot, u^\varepsilon_\cdot) \leq \varepsilon, \tag{19}$$

$$F(\tilde{u}^\varepsilon_\cdot) \leq \inf_{u_\cdot \in \mathcal{U}} F(u_\cdot), \tag{20}$$

where

$$F(u_\cdot) \triangleq f(u_\cdot) + \varepsilon\tilde{d}(\tilde{u}^\varepsilon_\cdot, u_\cdot) = \mathbb{E}\int_0^T \left[\mathcal{H}^\varepsilon(t, u_t) + \varepsilon v^\varepsilon_t \left|\tilde{u}^\varepsilon_t - u_t\right|\right] dt.$$

Note that (20) implies a pointwise maximum principle, that is, for a.e. $t \in [0, T]$, a.s., $\mathcal{H}^\varepsilon(t, v) + \varepsilon v^\varepsilon_t |\tilde{u}^\varepsilon_t - v|$ attains its minimum over U at $\tilde{u}^\varepsilon_t$. By Propositions 2.3.2 and 2.3.3 in [30], this yields

$$0 \in \partial_v \mathcal{H}^\varepsilon(t, \tilde{u}^\varepsilon_t) + \left[-\varepsilon v^\varepsilon_t, \varepsilon v^\varepsilon_t\right],$$

where $\partial\varphi(x)$ denotes Clarke's generalized gradient of φ at x. Since $\mathcal{H}^\varepsilon(t, v)$ is differentiable in v, the previous inclusion implies the existence of $\beta^\varepsilon_t \in [-\varepsilon v^\varepsilon_t, \varepsilon v^\varepsilon_t]$ such that $\mathcal{H}^\varepsilon_v(t, \tilde{u}^\varepsilon_t) = -\beta^\varepsilon_t$. Thus,

$$\left|\mathcal{H}^\varepsilon_v(t, \tilde{u}^\varepsilon_t)\right| \leq \varepsilon v^\varepsilon_t. \tag{21}$$

Then, by (H4) and the equality

$$\begin{aligned}\mathcal{H}^\varepsilon_v(t, u^\varepsilon_t) = {} & \mathcal{H}^\varepsilon_v(t, \tilde{u}^\varepsilon_t) + \left[H_v(\Sigma^\varepsilon_t, u^\varepsilon_t, u^\varepsilon_{t-\delta_2}) - H_v(\Sigma^\varepsilon_t, \tilde{u}^\varepsilon_t, u^\varepsilon_{t-\delta_2})\right] \\ & + \mathbb{E}^{\mathcal{F}_t}\left[H_{v_\delta}(\Sigma^\varepsilon_{t+\delta_2}, u^\varepsilon_{t+\delta_2}, u^\varepsilon_t) - H_{v_\delta}(\Sigma^\varepsilon_{t+\delta_2}, u^\varepsilon_{t+\delta_2}, \tilde{u}^\varepsilon_t)\right],\end{aligned}$$

there exists $C > 0$, independent of ε, such that

$$\left|\mathcal{H}^\varepsilon_v(t, u^\varepsilon_t)\right| \leq \varepsilon v^\varepsilon_t + C v^\varepsilon_t \left|\tilde{u}^\varepsilon_t - u^\varepsilon_t\right|.$$

This, together with (6) and (19), leads to the existence of $C'' > 0$, independent of ε, such that

$$\mathbb{E}\int_0^T \left|\mathcal{H}^\varepsilon_v(t, u^\varepsilon_t)\right| dt \leq C''\varepsilon.$$

Consequently, considering the boundedness of U, we can use (18) to derive

$$J(v_\cdot) - J(u^\varepsilon_\cdot) \geq -\mathbb{E}\int_0^T \left|\mathcal{H}^\varepsilon_v(t, u^\varepsilon_t)\right| |\hat{v}_t| \, dt \geq -C'\varepsilon.$$

Since $v_\cdot \in \mathcal{U}$ is arbitrarily chosen, this completes the proof. □

Remark 10 Theorem 9(i) shows that, under (H1)-(H3), an admissible control $u^\varepsilon_\cdot$ of problem (1)-(3) is ε-optimal if it satisfies (16). By Theorem 9(ii), we know that, under (H1)-(H4), if an admissible control $u^\varepsilon_\cdot$ of problem (1)-(3) satisfies

$$\inf_{v_\cdot \in \mathcal{U}} \mathbb{E} \int_0^T \left[\mathcal{H}^\varepsilon(t, v_t) - \mathcal{H}^\varepsilon\left(t, u_t^\varepsilon\right)\right] dt \geq -\left(\varepsilon/C'\right)^2,$$

then it is indeed ε-optimal. Note that the conclusions in Theorem 9(i) and (ii) cannot contain each other in general.

5 Applications

In this section, the theoretical results are applied to two examples.

Example 1 Take $U = [0,1]$. Assume that $X_\cdot$ satisfies

$$dX_t^v = v_{t-\delta}\, dW_t, \quad 0 \leq t \leq T; \qquad X_t^v = \xi_t, \quad -\delta \leq t \leq 0.$$

The objective is to minimize

$$J(v_\cdot) = \mathbb{E}\left[\int_0^T v_t\, dt + \frac{1}{2}\left(X_T^v\right)^2\right].$$

In this case, the adjoint equation is described by

$$dY_t^v = Z_t^v\, dW_t, \quad 0 \leq t \leq T; \qquad Y_T^v = X_T^v; \qquad Y_t^v = Z_t^v = 0, \quad T < t \leq T + \delta.$$

Comparing the adjoint equation with the system equation, by the uniqueness of the solutions, we get $(Y_t^v, Z_t^v) = (X_t^v, v_{t-\delta})$ for $0 \leq t \leq T$.

Note that $H(t, x, x_\delta, y, z, v, v_\delta) = v_\delta z + v$ and

$$\mathcal{H}^\varepsilon(t, v) = \left(1 + \mathbb{E}^{\mathcal{F}_t}\left[Z_{t+\delta}^\varepsilon\right]\right)v + \left(Z_t^\varepsilon u_{t-\delta}^\varepsilon + \mathbb{E}^{\mathcal{F}_t}\left[u_{t+\delta}^\varepsilon\right]\right).$$

Thus, $\mathcal{H}_v^\varepsilon(t, u_t^\varepsilon) = 1 + \mathbb{E}^{\mathcal{F}_t}[Z_{t+\delta}^\varepsilon]$. Besides, since $Z_t^\varepsilon = u_{t-\delta}^\varepsilon$ for $0 \leq t \leq T$ and $Z_t^\varepsilon = 0$ for $T < t \leq T + \delta$, we have $1 + \mathbb{E}^{\mathcal{F}_t}[Z_{t+\delta}^\varepsilon] = f^\varepsilon(t)$, with $f^\varepsilon(t) = 1 + u_t^\varepsilon$ for $0 \leq t \leq T - \delta$ and $f^\varepsilon(t) = 1$ for $T - \delta < t \leq T$. Thus,

$$\inf_{v_\cdot \in \mathcal{U}} \mathbb{E} \int_0^T \mathcal{H}_v^\varepsilon\left(t, u_t^\varepsilon\right)\left(v_t - u_t^\varepsilon\right) dt = -\mathbb{E} \int_0^T f^\varepsilon(t) u_t^\varepsilon\, dt.$$

By Theorem 9(i), an admissible control $u^\varepsilon_\cdot$ is ε-optimal if

$$\mathbb{E} \int_0^T f^\varepsilon(t) u_t^\varepsilon\, dt = \mathbb{E}\left[\int_0^{T-\delta} \left(1 + u_t^\varepsilon\right) u_t^\varepsilon\, dt + \int_{T-\delta}^T u_t^\varepsilon\, dt\right] \leq \varepsilon$$

and thus if

$$\mathbb{E} \int_0^T u_t^\varepsilon\, dt \leq \varepsilon/2.$$

Finally, let us give some examples of ε-optimal controls with sufficiently small ε:

$$u_t^\varepsilon = \frac{\varepsilon}{2T}, \frac{\varepsilon t}{T^2}, \frac{\min\{W_t, \varepsilon\}}{3T}.$$

Example 2 We consider a cash management problem. Denote by $X_\cdot$ the cash flow of an agent, and $v_\cdot$ the control strategy which is the rate of cash disturbance (cash inflow or cash outflow). Since there exist necessary and unavoidable time delays in practice, we assume that the dynamics of the cash flow is described by

$$\begin{cases} dX_t^v = [B_1(t)X_{t-\delta_1}^v + D_1(t)v_{t-\delta_2}]\,dt + [B_2(t)X_{t-\delta_1}^v + D_2(t)v_{t-\delta_2}]\,dW_t, & 0 \le t \le T, \\ X_t^v = \xi_t, \quad -\delta_1 \le t \le 0, \end{cases}$$

where the time-varying coefficients are bounded adapted processes. Our objective is to minimize the following functional:

$$J(v_\cdot) = \mathbb{E}\left[\int_0^T \frac{1}{2}N(t)\big(v_t - \alpha(t)\big)^2\,dt - QX_T^v\right],$$

where $N(\cdot)$ and $\alpha(\cdot)$ are bounded adapted process, and Q is a bounded $\mathcal{F}_T$-measurable random variable. $N(\cdot)$ and Q are weight coefficients, and $\alpha(\cdot)$ is interpreted as a dynamic benchmark. For clarity, we assume that $U = [c,d]$ with suitable constants c and d, $c \ge 0$, $N(t) > 0$ and $Q > 0$. In this case, the objective contains two parts: one is to maximize an expected terminal reward, and the other to minimize a square criterion on the control strategy $v_\cdot$, which is to prevent it from large deviation. Let us assume w.o.l.g. that $\alpha(t) \in U$ for all $t \in [0,T]$, and $v_t = c$ for all admissible control $v_\cdot$ and $t \in (T-\delta_2, T]$.

It is easy to check that the assumptions (H1)-(H4) hold true for this example. The adjoint equation takes the following form:

$$\begin{cases} -dY_t = \mathbb{E}^{\mathcal{F}_t}[B_1(t+\delta_1)Y_{t+\delta_1} + B_2(t+\delta_1)Z_{t+\delta_1}]\,dt - Z_t\,dW_t, & 0 \le t \le T, \\ Y_T = -Q, \\ Y_t = 0, \qquad Z_t = 0, \quad T < t \le T+\delta_1. \end{cases}$$

Note that the solution is independent of the control. Similar to [11], if the coefficients Q, $B_1(\cdot)$, $B_2(\cdot)$ are Malliavin differentiable, then this ABSDE can be solved interval by interval in Malliavin's sense to get its unique solution $(Y_\cdot, Z_\cdot)$.

The Hamiltonian H takes the following form:

$$H(t,x,x_\delta,y,z,v,v_\delta) = N(t)\big(v-\alpha(t)\big)^2/2 + \big[D_1(t)y + D_2(t)z\big]v_\delta + \big[B_1(t)y + B_2(t)z\big]x_\delta.$$

Set $\lambda(t) = \mathbb{E}^{\mathcal{F}_t}[D_1(t+\delta_2)Y_{t+\delta_2} + D_2(t+\delta_2)Z_{t+\delta_2}]$ and $\mathbb{H}(t,v) = N(t)(v-\alpha(t))^2/2 + \lambda(t)v$. Then by the definition of $\mathcal{H}^\varepsilon(t,v)$ we have

$$\mathcal{H}_v^\varepsilon(t,v) = N(t)\big(v-\alpha(t)\big) + \lambda(t), \qquad \mathcal{H}^\varepsilon(t,v) - \mathcal{H}^\varepsilon(t,u) = \mathbb{H}(t,v) - \mathbb{H}(t,u).$$

Set $P_t = (\alpha(t)N(t) - \lambda(t))/N(t)$ and $\gamma_t^\varepsilon = \inf_{v_\cdot \in \mathcal{U}} \mathbb{E}\int_0^T [\mathcal{H}^\varepsilon(t,v_t) - \mathcal{H}^\varepsilon(t,u_t^\varepsilon)]\,dt$. Then

$$\gamma_t^\varepsilon = \inf_{v_\cdot \in \mathcal{U}} \mathbb{E}\int_0^T \big[\mathbb{H}^\varepsilon(t,v_t) - \mathbb{H}^\varepsilon\big(t,u_t^\varepsilon\big)\big]\,dt.$$

By Remark 10, an admissible control $u^\varepsilon_\cdot$ is ε-optimal if it satisfies

$$\gamma^\varepsilon_t \geq -\left(\varepsilon/C'\right)^2.$$

Particularly, if $P_t \in U$ for all $t \in [0,T]$, then it is easy to check that

$$\gamma^\varepsilon_t = -\mathbb{E}\int_0^T \frac{1}{2}N(t)\left(u^\varepsilon_t - P_t\right)^2 dt.$$

Consequently, an adapted process $u^\varepsilon_\cdot$ is ε-optimal if it takes values in U and satisfies

$$\mathbb{E}\int_0^T N(t)\left(u^\varepsilon_t - P_t\right)^2 dt \leq 2\left(\varepsilon/C'\right)^2. \tag{22}$$

By (22), in order to find an ε-optimal control, we need to compute C'. To this end, we follow the proof of Theorem 9(ii). Let (17) hold. Recall that $U = [c,d]$ with $c \geq 0$, and $\alpha(t) \in U$. Since

$$\mathcal{H}^\varepsilon(t,v) - \mathcal{H}^\varepsilon(t,u) = \left[N(t)\left(u + v - 2\alpha(t)\right)/2 + \lambda(t)\right](v-u),$$

we have

$$\left|\mathcal{H}^\varepsilon(t,v) - \mathcal{H}^\varepsilon(t,u)\right| \leq \nu_t|v-u|,$$

where

$$\nu_t = 1 + dN(t) + \mathbb{E}^{\mathcal{F}_t}\left[\left|D_1(t+\delta_2)Y_{t+\delta_2}\right| + \left|D_2(t+\delta_2)Z_{t+\delta_2}\right|\right].$$

So $f(u.) \triangleq \mathbb{E}\int_0^T \mathcal{H}^\varepsilon(t,u_t)\,dt$ satisfies

$$\left|f(u.) - f(v.)\right| \leq \tilde{d}(u.,v.) \triangleq \int_0^T \nu_t|u_t - v_t|\,dt.$$

On the one hand, by (21) we have $|\mathcal{H}^\varepsilon_v(t,\tilde{u}^\varepsilon_t)| \leq \varepsilon\nu_t$. On the other hand, $\mathcal{H}^\varepsilon_v(t,u^\varepsilon_t) = \mathcal{H}^\varepsilon_v(t,\tilde{u}^\varepsilon_t) + N(t)(u^\varepsilon_t - \tilde{u}^\varepsilon_t)$. Thus,

$$\left|\mathcal{H}^\varepsilon_v\left(t,u^\varepsilon_t\right)\right| \leq \varepsilon\nu_t + N(t)\left|u^\varepsilon_t - \tilde{u}^\varepsilon_t\right| \leq \varepsilon\nu_t + \nu_t\left|\tilde{u}^\varepsilon_t - u^\varepsilon_t\right|/d,$$

and so

$$\left|\mathcal{H}^\varepsilon_v\left(t,u^\varepsilon_t\right)\left(v_t - u^\varepsilon_t\right)\right| \leq d\varepsilon\nu_t + \nu_t\left|\tilde{u}^\varepsilon_t - u^\varepsilon_t\right|.$$

Therefore,

$$J(v.) - J\left(u^\varepsilon_\cdot\right) \geq -\mathbb{E}\int_0^T \left|\mathcal{H}^\varepsilon_v\left(t,u^\varepsilon_t\right)\left(v_t - u^\varepsilon_t\right)\right| dt \geq -d\mathbb{E}\int_0^T \nu_t\,dt\varepsilon - \tilde{d}\left(\tilde{u}^\varepsilon_\cdot, u^\varepsilon_\cdot\right).$$

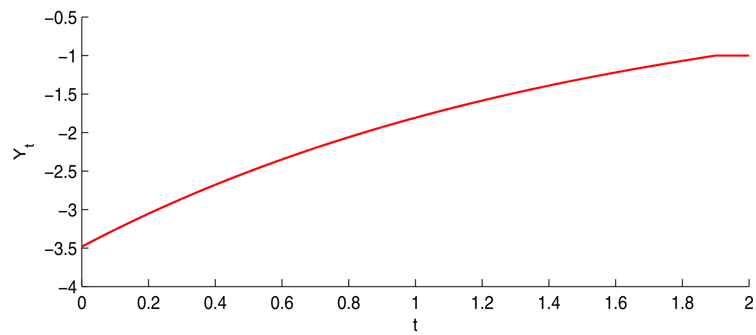

Figure 1 The function Y_t.

Next, in view of (19), we have

$$J(\nu_\cdot) - J(u^\varepsilon_\cdot) \geq -\left(1 + d\mathbb{E}\int_0^T \nu_t\,dt\right)\varepsilon,$$

and thus

$$J(u^\varepsilon_\cdot) \leq V + \left(1 + d\mathbb{E}\int_0^T \nu_t\,dt\right)\varepsilon,$$

due to the arbitrariness of $\nu_\cdot \in \mathcal{U}$. So C' could be any constant satisfying

$$C' \geq 1 + d\mathbb{E}\int_0^T \nu_t\,dt.$$

Finally, let us give a numerical simulation. Assume that the coefficients are all deterministic and time-invariant. Take $c = 0$, $d = 3$, $T = 2$, $\delta_1 = \delta_2 = 0.1$, $B_1(t) = B_2(t) = 0.7$, $Q = 1$, $D_1(t) = D_2(t) = 0.5$, $N(t) = 1$, $\alpha(t) = 0$. In this case, it is easy to check that $Z_t = 0$ for $0 \leq t \leq 2.1$, and Y_t solves the following ODE:

$$Y_t' = -0.7Y_{t+0.1}, \quad 0 \leq t \leq 2; \qquad Y_2 = -1; \qquad Y_t = 0, \quad 2 < t \leq 2.1,$$

which can be solved explicitly by subdividing $[0, 2]$ backwardly to get

$$Y_t = -1, \quad 1.9 \leq t \leq 2,$$
$$Y_t = -1 - 0.7(1.9 - t), \quad 1.8 \leq t \leq 1.9,$$
$$\vdots$$

The graph of Y_t is shown in Figure 1. Then it is easy to check that $1 + d\int_0^T \nu_t\,dt < 30$, so we can take $C' = 100$. Since $P_t = -0.5Y_{t+0.1} \in U$, we can conclude that an adapted process $u^\varepsilon_\cdot$ is ε-optimal if it takes values in U and satisfies

$$\mathbb{E}\int_0^2 \left(u^\varepsilon_t + 0.5Y_{t+0.1}\right)^2 dt \leq 2(\varepsilon/100)^2.$$

Let us give an example of ε-optimal control for sufficiently small ε:

$$u_t^\varepsilon = \begin{cases} -0.5Y_{t+0.1} + \varepsilon/100, & 0 \le t \le 1.9; \\ 0, & 1.9 < t \le 2. \end{cases}$$

6 Conclusion

We study near-optimal controls for one kind of stochastic delay control problem with convex control domain. By the stochastic maximum principle and Ekeland's variational principle, we establish necessary conditions for a control to be near-optimal. Sufficient conditions are also given, which show when an admissible control is indeed near-optimal. Two illustrative examples are given, for which some near-optimal controls in the explicit form are obtained. Future work includes the nonconvex control domain case and linear quadratic problems in terms of the Riccati equations.

Competing interests

The author declares that there is no competing interest regarding the publication of this paper.

Acknowledgements

The author would like to thank the editor and the reviewers for their valuable comments and suggestions that helped to improve this paper. This work is supported by the National Natural Science Foundation of China (11371228, 11471192), Natural Science Foundation of Shandong Province (ZR2015JL003, ZR2015JL021), Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions, Special Funds of Taishan Scholar Project (tsqn20161041) and Project of Shandong Province Higher Educational Science and Technology Program (J15LI02).

Publisher's Note

Received: 18 November 2016 Accepted: 21 March 2017 Published online: 31 March 2017

References

1. Mohammed, SE: Stochastic Functional Differential Equations. Pitman, London (1984)
2. Mao, XR: Stochastic Differential Equations and Their Applications. Horwood, New York (1997)
3. Arriojas, M, Hu, Y, Mohammed, SE, Pap, G: A delayed Black and Scholes formula. Stoch. Anal. Appl. **25**(2), 471-492 (2007)
4. Elsanousi, I, Oksendal, B, Sulem, A: Some solvable stochastic control problems with delay. Stochastics **71**(1-2), 69-89 (2000)
5. Larrsen, B: Dynamic programming in stochastic control of systems with delay. Stochastics **74**(3), 651-673 (2002)
6. Oksendal, B, Sulem, A, Zhang, TS: Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations. Adv. Appl. Probab. **43**, 572-596 (2011)
7. Agram, N, Haadem, S, Oksendal, B, Proske, F: A maximum principle for infinite horizon delay equations. SIAM J. Math. Anal. **45**(4), 2499-2522 (2013)
8. Shen, Y, Meng, QX, Shi, P: Maximum principle for mean-field jump-diffusion stochastic delay differential equations and its application to finance. Automatica **50**(6), 1565-1579 (2014)
9. Peng, SG, Yang, Z: Anticipated backward stochastic differential equations. Ann. Probab. **37**(3), 877-902 (2009)
10. Chen, L, Wu, Z: Maximum principle for the stochastic optimal control problem with delay and application. Automatica **46**(6), 1074-1080 (2010)
11. Yu, ZY: The stochastic maximum principle for optimal control problems of delay systems involving continuous and impulse controls. Automatica **48**(10), 2420-2432 (2012)
12. Huang, JH, Shi, JT: Maximum principle for optimal control of fully coupled forward-backward stochastic differential delayed equations. ESAIM Control Optim. Calc. Var. **18**(4), 1073-1096 (2012)
13. Du, H, Huang, JH, Qin, YL: A stochastic maximum principle for delayed mean-field stochastic differential equations and its applications. IEEE Trans. Autom. Control **58**(12), 3212-3217 (2013)
14. Zhou, XY: Stochastic near-optimal controls: necessary and sufficient conditions for near-optimality. SIAM J. Control Optim. **36**(3), 929-947 (1998)
15. Ekeland, I: On the variational principle. J. Math. Anal. Appl. **47**(2), 324-353 (1974)
16. Zhou, XY: Deterministic near-optimal controls. Part I: necessary and sufficient conditions for near optimality. J. Optim. Theory Appl. **85**(2), 473-488 (1995)
17. Zhou, XY: Deterministic near-optimal controls. Part II: dynamic programming and viscosity solution approach. Math. Oper. Res. **21**(3), 655-674 (1996)
18. Chighoub, F, Mezerdi, B: Near optimality conditions in stochastic control of jump diffusion processes. Syst. Control Lett. **60**(11), 907-916 (2011)

19. Hafayed, M, Abbas, S, Veverka, P: On necessary and sufficient conditions for near-optimal singular stochastic controls. Optim. Lett. **7**(5), 949-966 (2013)
20. Hafayed, M, Abbas, S: Stochastic near-optimal singular controls for jump diffusions: necessary and sufficient conditions. J. Dyn. Control Syst. **19**(4), 503-517 (2013)
21. Hafayed, M, Abbas, S: On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality. J. Optim. Theory Appl. **160**(3), 778-808 (2014)
22. Hafayed, M, Abba, A, Abbas, S: On mean-field stochastic maximum principle for near-optimal controls for Poisson jump diffusion with applications. Int. J. Dyn. Control **2**(3), 262-284 (2014)
23. Meng, QX, Shen, Y: A revisit to stochastic near-optimal controls: the critical case. Syst. Control Lett. **82**, 79-85 (2015)
24. Bahlali, K, Khelfallah, N, Mezerdi, B: Necessary and sufficient conditions for near-optimality in stochastic control of FBSDEs. Syst. Control Lett. **58**, 857-864 (2009)
25. Huang, JH, Li, X, Wang, GC: Near-optimal control problems for linear forward-backward stochastic systems. Automatica **46**(2), 397-404 (2010)
26. Hafayed, M, Veverka, P, Abbas, S: On maximum principle of near-optimality for diffusions with jumps, with application to consumption-investment problem. Differ. Equ. Dyn. Syst. **20**(2), 111-125 (2012)
27. Hafayed, M, Veverka, P, Abbas, S: On near-optimal necessary and sufficient conditions for forward-backward stochastic systems with jumps, with applications to finance. Appl. Math. **59**(4), 407-440 (2014)
28. Zhang, LQ, Huang, JH, Li, X: Necessary condition for near optimal control of linear forward-backward stochastic differential equations. Int. J. Control **88**(8), 1594-1608 (2015)
29. Hafayed, M, Abba, A, Boukaf, S: On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications. Int. J. Model. Identif. Control **25**(1), 1-16 (2016)
30. Clarke, F: Optimization and Nonsmooth Analysis. SIAM, Philadelphia (1990)