SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

November 2016, Vol. 59 112206:1–112206:14
doi: 10.1007/s11432-015-0777-0

The near-optimal maximum principle of impulse control for stochastic recursive system

Jianhui HUANG[1] & Detao ZHANG[2*]

[1]*Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China;*
[2]*School of Economics, Shandong University, Jinan 250100, China*

Received November 16, 2015; accepted January 15, 2016; published online October 14, 2016

Abstract Here, we discuss the near-optimality for a class of stochastic impulse control problems. The state process in our problem is given by forward-backward stochastic differential equations (FBSDEs) with two control components involved: the regular and impulse control. More specially, the impulse control is defined on a sequence of prescribed stopping times. A recursive cost functional is introduced and the maximum principle for its near-optimality (both necessary and sufficient conditions) is derived with the help of Ekeland's principle and variational analysis. For illustration, one concrete example is studied with our maximum principle and the corresponding near-optimal control is characterized.

Keywords Ekeland's principle, FBSDE, impulse control, maximum principle, near optimality.

Citation Huang J H, Zhang D T. The near-optimal maximum principle of impulse control for stochastic recursive system. Sci China Inf Sci, 2016, 59(11): 112206, doi: 10.1007/s11432-015-0777-0

1 Introduction

The problems of stochastic impulse control have been extensively studied in the last few decades. Their applications can be found in various fields including engineering, operational research, mathematical finance and economics, etc. A typical example can be addressed when we study the portfolio selection with transaction costs (e.g., [1]): the regular control component is introduced to characterize the consumption process while the impulse control is to characterize the transaction costs which may occur at given stopping times. In addition, there also exist considerable decision-making and management problems which can be framed into a dynamic game between a sequence of short-time impulses and a long-run patient variable (see [2, 3]). Consequently, the stochastic impulse control problems arise naturally in this framework.

It is remarkable that in many cases, stochastic impulse control problems are discussed with the employment of the dynamic programming principle (DPP). Based on it, the value function is shown to satisfy some HJB quasi-variational inequalities. On the other hand, an alternative method is to study the associated maximum principle by which the necessary condition satisfied by the optimality can be figured out with the (convex or spike) perturbation. Concerning this research line, some relevant literature includes [4] for forward stochastic singular control, [5–7] for forward-backward stochastic impulse control.

* Corresponding author (email: zhangdetao@sdu.edu.cn)

Unlike the above mentioned work, this paper aims to study the "near-optimality" of stochastic impulse control for FBSDE system. To our best knowledge, this problem has never been touched although it deserves some research attention due to the following motivations.

The first motivation is as follows. In general, the existence of "exact" optimal control for given stochastic control problem cannot always be guaranteed under rather mild conditions. Furthermore, even it exists, the characterization of exact-optimality is not feasible to get in considerable situations. For example, when the datum of our control problem has no sufficient regularity (for instance, the coefficients in state dynamics or cost functional are not smooth enough), an approximation procedure should be applied and the near optimality should thus be addressed. Actually, for a given control problem, its near-optimality always exists which usually meets our practical requirements. In this sense, the near-optimal controls are more available and feasible than exact optimal ones. Some related literature which is more close to our current work is briefly stated below: in [8], the necessary and sufficient conditions for forward stochastic differential system are systematically discussed with the idea of Ekeland's variational principle. Based on it, Ref. [9] studies the near-optimal control where the state is driven by a linear forward-backward stochastic differential equation, while in [10], the near-optimal control for general nonlinear forward-backward state system is investigated. Recently, a critical case to stochastic system is studied in [11] which provides an interesting insightful viewpoint to revisit the near-optimal control problem.

The second motivation follows from the special structure of our stochastic impulse control problem for FBSDE system. Roughly speaking, there exist two control variables (regular and impulse) and two state components (forward and backward state) in our control problem. This structure feature makes our datum set (e.g., coefficients or stopping times) of control problem more liable to be irregular hence it is more reasonable to investigate the corresponding near-optimality. For example, the impulse control problem for forward-backward system arises when the decision making policies involve the pre-scripted random time horizons and the utility functional is of stochastic recursive type.

As the response to above motivations, in this paper we consider the maximum principle of near-optimal control for stochastic recursive systems involving impulse controls. We aim to derive the necessary and sufficient conditions of this kind of control problems. The rest of this paper is structured as follows. Section 2 presents the formulation of our problem. In Section 3, the necessary condition to near-optimality is derived while Section 4 gives its sufficient condition. For illustration, one example is introduced and studied in Section 5 based on our theoretical results.

2 Preliminaries and problem formulation

Let $[0,T]$ be a finite time horizon. $(\Omega,\mathcal{F},\mathbb{P})$ is a complete probability space on which $\{B_t, 0\leqslant t\leqslant T\}$ is d-dimensional Brownian motion. Let $\{\mathcal{F}_t\}_{0\leqslant t\leqslant T}$ be the natural filtration generated by $\{B_t\}$ augmented by $\mathbb{P}$-null sets. Throughout this paper, we make use of the following notations:

$\mathcal{L}^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$: the set of all $\mathbb{R}^n$-valued, $\mathcal{F}_t$-adapted and square-integrable processes;

ϕ_x : the partial derivative of ϕ with respect to x;

$|\cdot|$: the norm of an Euclidean space;

M^τ : the transpose of a given matrix or vector M;

$\mathcal{X}_{\mathcal{S}}$: the indicator function of a set $\mathcal{S}$;

$X+Y$: the set $\{x+y : x\in X, y\in Y\}$ for any X and Y.

Let $\{\tau_i\}$ be a given sequence of increasing $\mathcal{F}_t$-stopping times such that $\tau_i\uparrow+\infty$. We denote by $\mathcal{I}$ the class of processes $\eta(\cdot)=\sum_{i\geqslant 1}\eta_i\mathcal{X}_{[\tau_i,T]}(\cdot)$, where $\eta_i\in K$ is an $\mathcal{F}_{\tau_i}$-measurable bounded random variable, K is a nonempty convex subset of $\mathbb{R}$. It is worth noting that, the assumption $\tau_i\uparrow+\infty$ implies there exist at most finite impulses occurring on $[0,T]$. Let $U\subseteq\mathbb{R}^k$ be a nonempty convex closed set and denote by $\mathcal{U}$ the class of processes $v:[0,T]\times\Omega\rightarrow U$ such that v_t is an $\mathcal{F}_t$-adapted process and $\mathbb{E}\int_0^T|v_t|^4\mathrm{d}t<+\infty$. Let $\mathcal{K}$ be the class of impulse processes $\eta\in\mathcal{I}$ such that $\mathbb{E}\sum_{i\geqslant 1}|\eta_i|^4<\infty$. We call $\mathcal{A}\triangleq\mathcal{U}\times\mathcal{K}$ the admissible control set.

Now we consider the following stochastic recursive control system:

$$\begin{cases} \mathrm{d}X_t = b(t, X_t, v_t)\mathrm{d}t + \sigma(t, X_t, v_t)\mathrm{d}B_t + C_t\mathrm{d}\eta_t, \\ -\mathrm{d}Y_t = f(t, X_t, Y_t, Z_t, v_t)\mathrm{d}t - Z_t\mathrm{d}B_t + D_t\mathrm{d}\eta_t, \\ X_0 = x_0, \quad Y_T = g(X_T), \quad t \in [0, T], \end{cases} \tag{1}$$

where $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times d}$, $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times U \to \mathbb{R}^m$ are measurable mappings, and $C : [0, T] \to \mathbb{R}^{n\times n}$, $D : [0, T] \to \mathbb{R}^{m\times n}$ are continuous functions. The cost functional to be minimized is given by

$$\mathcal{J}(v, \eta) = \mathbb{E}\bigg\{\phi(X_T) + \gamma(Y_0) + \int_0^T h(t, X_t, Y_t, Z_t, v_t)\mathrm{d}t + \sum_{i\geqslant 1} l(\tau_i, \eta_i)\bigg\}, \tag{2}$$

where $\phi : \mathbb{R}^n \to \mathbb{R}$, $\gamma : \mathbb{R}^m \to \mathbb{R}$, $h : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times U \to \mathbb{R}$ and $l : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ are measurable mappings.

Next we want to illustrate the motivation and practical meaning of the above cost functionals. In short selling models, endowed with a positive initial wealth x_0, a small investor tries to maximize an expected functional which includes three parts. The first part is the utility from the reward $\mathbb{E}[-\frac{1}{2}(X_T - a)^2]$, the second part is a recursive utility functional (which is introduced by a backward SDE) with a generator $f(t, x, y, z)$, and the third part is the utility derived from the piecewise consumption process $\eta(\cdot)$. More precisely, for any admissible control pair $(v(\cdot), \eta(\cdot))$, the utility functional can be defined by

$$J(v, \eta) = \mathbb{E}\bigg\{ -\frac{1}{2}(X_T^{v,\eta} - a)^2 + y^{v,\eta}(0) - \frac{1}{2}S\sum_{i\geqslant 1}\eta_i^2\bigg\}.$$

In what follows we assume Assumption 2.1.

Assumption 2.1. (H1) For any $0 \leqslant t \leqslant T$, b, σ, f, g, ϕ, γ, h are continuous and they are continuous differentiable w.r.t. (x, y, z, v). l is continuous, and it is continuous differentiable in η. Moreover, there exists a constant $C > 0$ such that

$$|b(t, x, v)| + |\sigma(t, x, v)| + |f(t, x, y, z, v)| + |g(x)| \leqslant C(1 + |x| + |y| + |z|);$$

(H2) The derivatives of b, σ, f, g are uniformly bounded;

(H3) For any $(t, x, y, z) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$, the partial derivatives of ϕ_x, γ_y, h_x, h_y, h_z are continuous and there exists a constant C such that

$$\begin{aligned} &(1 + |x|)^{-1}|\phi_x(x)| + (1 + |y|)^{-1}|\gamma_y(y)| \leqslant C; \\ &|h_x(t, x, y, z, v)| + |h_y(t, x, y, z, v)| + |h_z(t, x, y, z, v)| \leqslant C(1 + |x| + |y| + |z|); \\ &|l_\eta(t, \eta)| \leqslant C(1 + |\eta|), \qquad \forall\, \eta. \end{aligned}$$

The control problem under consideration is to find an admissible control which minimizes or nearly minimizes the cost functional $\mathcal{J}(v, \eta)$ over all admissible controls $\big(v(\cdot), \eta(\cdot) = \sum_{i\geqslant 1}\eta_i\mathcal{X}_{[\tau_i,T]}(\cdot)\big) \in \mathcal{A}$. The value function of our problem is thus defined as

$$V(0; x_0) \triangleq \inf_{(v,\eta)\in\mathcal{A}} \mathcal{J}(v, \eta). \tag{3}$$

The utility maximization problems with piecewise consumption processes can be regarded as an above stochastic optimal control problem with impulse control.

From the Propositions 2.1 and 2.2 in [5, 6], it follows that the forward-backward SDEs (1) admit a unique solution $(X(\cdot), Y(\cdot), Z(\cdot)) \in \mathcal{L}^2_{\mathcal{F}}(0, T; \mathbb{R}^n) \times \mathcal{L}^2_{\mathcal{F}}(0, T; \mathbb{R}^m) \times \mathcal{L}^2_{\mathcal{F}}(0, T; \mathbb{R}^{m\times d})$. For any $(v, \eta) \in \mathcal{U} \times \mathcal{K}$, the cost functional $\mathcal{J}$ in (2) is well defined.

Now we present the following two definitions, which are similar to those of [8].

Definition 2.2 (Optimal control). An admissible control $\big(u(\cdot),\xi(\cdot)=\sum_{i\geqslant 1}\xi_i\mathcal{X}_{[\tau_i,T]}(\cdot)\big)\in\mathcal{A}$ is called the optimal, if $\big(u(\cdot),\xi(\cdot)\big)$ attains the minimum of $\mathcal{J}(0,x_0;v,\eta)$.

Definition 2.3 (Near-optimal control). Both a family of admissible controls $\big\{\big(u^\varepsilon(\cdot),\xi^\varepsilon(\cdot)\big)\big\}$ parameterized by $\varepsilon>0$, and any element $\big(u^\varepsilon(\cdot),\xi^\varepsilon(\cdot)\big)$ in the family is called the near-optimal if

$$|\mathcal{J}(0,x_0;u^\varepsilon(\cdot),\xi^\varepsilon(\cdot))-V(0,x_0)|\leqslant\delta(\varepsilon)$$

holds for sufficiently small ε, where δ is a function of ε satisfying $\delta(\varepsilon)\to 0$ as $\varepsilon\to 0$. The estimate $\delta(\varepsilon)$ is called an error bound. If $\delta(\varepsilon)=c\varepsilon^\beta$ for some $\beta>0$ independent of the constant c, then $\big(u^\varepsilon(\cdot),\xi^\varepsilon(\cdot)\big)$ is called the near-optimal with order e^β.

Lemma 2.4 (Ekeland's principle [12]). Let (S,d) be a complete metric space and $F(\cdot):S\to\mathbb{R}$ be lower-semicontinuous and bounded from below. Suppose that $v^\varepsilon\in S$ satisfies

$$F(v^\varepsilon)\leqslant\inf_{v\in S}F(v)+\varepsilon.$$

Then there exists $v^\lambda\in S$ such that for any $\lambda>0$,

(1) $F(v^\lambda)\leqslant F(v^\varepsilon)$,

(2) $d(v^\varepsilon,v^\lambda)\leqslant\lambda$,

(3) $F(v^\lambda)\leqslant F(v)+\frac{\varepsilon}{\lambda}d(v,v^\lambda)$ for all $v\in S$.

3 Necessary condition for near-optimality

3.1 Some prior estimates

This subsection gives some prior estimates which play an important role in deriving our maximum principle. For any $(v(\cdot),\eta(\cdot))$ and $(v'(\cdot),\eta'(\cdot))$, let us introduce a metric on the admissible control set $\mathcal{A}$ as follows:

$$d\big((v,\eta),(v',\eta')\big)\triangleq\left(\mathbb{E}\int_0^T|v(t)-v'(t)|^2\mathrm{d}t+\mathbb{E}\sum_{\tau_i\leqslant T}|\eta_i-\eta_i'|^2\right)^{\frac{1}{2}}.\tag{4}$$

Lemma 3.1. Under Assumption 2.1, for any given $\big(v(\cdot),\eta(\cdot)=\sum_{i\geqslant 1}\eta_i\mathcal{X}_{[\tau_i,T]}(\cdot)\big)\in\mathcal{A}$, $\forall\,p\geqslant 2$, and $\mathbb{E}[\sum_{i\geqslant 1}|\eta_i|^p]\leqslant\infty$, Eq. (1) admits one unique solution $(X_t,Y_t,Z_t)\in\mathcal{L}^p_{\mathcal{F}}(0,T;\mathbb{R}^n)\times\mathcal{L}^p_{\mathcal{F}}(0,T;\mathbb{R}^m)\times\mathcal{L}^p_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d})$. Moreover, it holds that

$$\mathbb{E}\sup\nolimits_{0\leqslant t\leqslant T}|X_t|^p\leqslant C,\quad \sup\nolimits_{0\leqslant t\leqslant T}\mathbb{E}|Y_t|^p\leqslant C,\quad \mathbb{E}\sup\nolimits_{0\leqslant t\leqslant T}|Y_t|^p\leqslant C,\quad \mathbb{E}\int_0^T|Z_t|^p\mathrm{d}t\leqslant C,\tag{5}$$

where C is a constant.

Proof. Similar proof with Propositions 2.1 and 2.2 in [5, 6].

To solve the problem (NOC), we introduce the following FBSDE (adjoint equation):

$$\begin{cases}\mathrm{d}P_t=[f_y^\tau(t,X_t,Y_t,Z_t,v_t)P_t-h_y^\tau(t,X_t,Y_t,Z_t,v_t)]\mathrm{d}t\\ \qquad\quad +[f_z^\tau(t,X_t,Y_t,Z_t,v_t)P_t-h_z^\tau(t,X_t,Y_t,Z_t,v_t)]\mathrm{d}B_t,\\ -\mathrm{d}Q_t=[-f_x(t,X_t,Y_t,Z_t,v_t)P_t+b_x^\tau(t,X_t,v_t)Q(t)+\sigma_x^\tau(t,X_t,v_t)K_t\\ \qquad\quad +h_x(t,X_t,Y_t,Z_t,v_t)]\mathrm{d}t-K_t\mathrm{d}B_t,\\ P_0=-\gamma_y(Y_0),\quad Q_T=-g_x(X_T)P_T+\phi_x(X_T),\quad t\in[0,T],\end{cases}\tag{6}$$

where $v(\cdot)$ is the admissible control and (X_t,Y_t,Z_t,v_t) is the corresponding trajectory. Define the following Hamiltonian function:

$$H(t,x,y,z,v,p,q,k)=\langle b(t,x,v),q\rangle-\langle f(t,x,y,z,v),p\rangle+\langle\sigma(t,x,v),k\rangle+h(t,x,y,z,v),\tag{7}$$

with $H:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}$. Then Eq. (6) can be rewritten as

$$\begin{cases}\mathrm{d}P_t=-H_y(t,X_t,Y_t,Z_t,v_t,P_t,Q_t,K_t)\mathrm{d}t+H_z(t,X_t,Y_t,Z_t,v_t,P_t,Q_t,K_t)\mathrm{d}B_t,\\ -\mathrm{d}Q_t=H_x(t,X_t,Y_t,Z_t,v_t,P_t,Q_t,K_t)\mathrm{d}t-K_t\mathrm{d}B_t,\\ P_0=-\gamma_y(Y_0),\quad Q_T=-g_x(X_T)P_T+\phi_x(X_T),\quad t\in[0,T].\end{cases}$$

Under Assumption 2.1, Eq. (6) admits a unique solution $(P_t,Q_t,K_t)\in\mathcal{L}^2_{\mathcal{F}}(0,T;\mathbb{R}^m)\times\mathcal{L}^2_{\mathcal{F}}(0,T;\mathbb{R}^n)\times\mathcal{L}^2_{\mathcal{F}}(0,T;\mathbb{R}^{n\times d})$.

Lemma 3.2. Let Assumption 2.1 hold true, for $p\geqslant 2$, and $\mathbb{E}[\sum_{i\geqslant 1}|\eta_i|^p]\leqslant\infty$, then there exists a constant $C>0$ such that

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}|P_t|^p\leqslant C,\quad \sup_{0\leqslant t\leqslant T}\mathbb{E}|Q_t|^p\leqslant C,\quad \mathbb{E}\sup_{0\leqslant t\leqslant T}|Q_t|^p\leqslant C,\quad \mathbb{E}\int_0^T|K_t|^p\mathrm{d}t\leqslant C. \tag{8}$$

Proof. Similar to Lemma 3.1, and Propositions 2.1 and 2.2 in [5,6].

The following lemma is concerned with the continuity of state process under metric d.

Lemma 3.3. Let Assumption 2.1 hold true, then there exists a constant C such that for any $(v(\cdot),\eta(\cdot))$ and $(v'(\cdot),\eta'(\cdot))$ along with the corresponding state processes (X,Y,Z), (X',Y',Z'),

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}|X_t-X'_t|^2\leqslant Cd^2\big((v,\eta),(v',\eta')\big); \tag{9}$$

$$\sup_{0\leqslant t\leqslant T}\mathbb{E}|Y_t-Y'_t|^2+\mathbb{E}\int_0^T|Z_t-Z'_t|^2\mathrm{d}t\leqslant Cd^2\big((v,\eta),(v',\eta')\big). \tag{10}$$

Proof. After direct computation, we have

$$\begin{aligned}\mathbb{E}\sup\nolimits_{0\leqslant t\leqslant T}|X_t-X'_t|^2&\leqslant C\bigg\{\mathbb{E}\int_0^T\big[|X_t-X'_t|^2+|v_t-v'_t|^2\big]\mathrm{d}t+\mathbb{E}\sum_{\tau_i\leqslant T}|\eta_i-\eta'_i|^2\bigg\}\\&\leqslant C\bigg[\int_0^T\mathbb{E}\sup_{0\leqslant t\leqslant r}|X_t-X'_t|^2\mathrm{d}r+d^2\big((v,\eta),(v',\eta')\big)\bigg].\end{aligned}$$

Hence Eq. (9) follows from the Gronwall's inequality. We now proceed to prove the second estimate. Set $\overline{Y}_t:=Y_t-Y'_t$, $\overline{Z}_t:=Z_t-Z'_t$, $\Theta_t:=(X_t,Y_t,Z_t)$ and $\Theta'_t:=(X'_t,Y'_t,Z'_t)$. Taking square in both sides of

$$\overline{Y}_t+\int_t^T\overline{Z}_s\mathrm{d}B_s=g(X_T)-g(X'_T)+\int_t^T\big[f(s,\Theta_s,v_s)-f(s,\Theta'_s,v'_s)\big]\mathrm{d}s+\sum_{t\leqslant\tau_i\leqslant T}C_{\tau_i}(\eta_i-\eta'_i),$$

and using the fact that $\mathbb{E}[\overline{Y}_t^\tau\int_t^T\overline{Z}_s\mathrm{d}B_s]=0$, we have

$$\begin{aligned}\mathbb{E}|\overline{Y}_t|^2+\mathbb{E}\int_t^T|\overline{Z}_s|^2\mathrm{d}s&\leqslant C_1\mathbb{E}|g(X_T)-g(X'_T)|^2+C_1\mathbb{E}\bigg\{\int_t^T|f(s,\Theta_s,v_s)-f(s,\Theta'_s,v'_s)|\mathrm{d}s\bigg\}^2\\&\quad+C_1\mathbb{E}\bigg\{\sum_{t\leqslant\tau_i\leqslant T}C_{\tau_i}(\eta_i-\eta'_i)\bigg\}^2\\&\leqslant C_2\mathbb{E}|X_T-X'_T|^2+C_2\mathbb{E}\bigg\{\int_t^T[|X_s-X'_s|+|Y_s-Y'_s|+|Z_s-Z'_s|\\&\quad+|v_s-v'_s|]\mathrm{d}s\bigg\}^2+C_2\mathbb{E}\sum_{t\leqslant\tau_i\leqslant T}|\eta_i-\eta'_i|^2\\&\leqslant C_3\mathbb{E}|X_T-X'_T|^2+C_3T\mathbb{E}\int_t^T[|X_s-X'_s|^2+|Y_s-Y'_s|^2+|v_s-v'_s|^2]\mathrm{d}s\\&\quad+C_3(T-t)\mathbb{E}\int_t^T|Z_s-Z'_s|^2\mathrm{d}s+C_3\mathbb{E}\sum_{t\leqslant\tau_i\leqslant T}|\eta_i-\eta'_i|^2.\end{aligned}$$

For $t \in [T-\delta, T]$ where $\delta = \frac{1}{2C_3}$, using the first estimate of (9), we obtain

$$\mathbb{E}|\overline{Y}_t|^2 + \tfrac{1}{2}\mathbb{E}\int_t^T |\overline{Z}_s|^2 \mathrm{d}s \leqslant C_4 \mathbb{E}\int_t^T |\overline{Y}_s|^2 \mathrm{d}s + C_4 d^2\big((v,\eta),(v',\eta')\big).$$

By Gronwall's inequality, we obtain

$$\mathbb{E}|\overline{Y}_t|^2 + \mathbb{E}\int_t^T |\overline{Z}_s|^2 \mathrm{d}s \leqslant C_5 d^2\big((v,\eta),(v',\eta')\big), \quad t \in [T-\delta, T].$$

Similarly we can get

$$\mathbb{E}|\overline{Y}_t|^2 + \mathbb{E}\int_t^{T-\delta} |\overline{Z}_s|^2 \mathrm{d}s \leqslant C_5 d^2\big((v,\eta),(v',\eta')\big), \quad t \in [T-2\delta, T-\delta].$$

After finite iterations, we get (10).

The following lemma gives the continuous-dependence of the solutions to adjoint equations. It plays an important role in proving the necessary condition. Before giving the lemma, we need to introduce the following assumptions.

Assumption 3.4. For any $t \in [0,T]$, (x,y,z,v) and $(x',y',z',v') \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times U$, there exists a constant $C > 0$ such that

$$\begin{aligned}
&|h_x(x) - h_x(x')| \leqslant C|x-x'|; \quad |\gamma_y(y) - \gamma_y(y')| \leqslant C|y-y'|;\\
&|b_\alpha(t,x,v) - b_\alpha(t,x',v')| + |\sigma_\alpha(t,x,v) - \sigma_\alpha(t,x',v')| \leqslant C(|x-x'| + |v-v'|);\\
&|f_\beta(t,x,y,z,v) - f_\beta(t,x',y',z',v')| + |l_\beta(t,x,y,z,v) - l_\beta(t,x',y',z',v')|\\
&\quad \leqslant C(|x-x'| + |y-y'| + |z-z'| + |v-v'|),
\end{aligned}$$

where $\alpha = x,\ v$ and $\beta = x,\ y,\ z,\ v$.

Lemma 3.5. Let Assumptions 2.1 and 3.4 hold, then there exists a constant C, such that for any $(v(\cdot),\eta(\cdot))$ and $(v'(\cdot),\eta'(\cdot))$ along with the corresponding adjoint processes $(P(\cdot),Q(\cdot),K(\cdot))$, $(P'(\cdot),Q'(\cdot),K'(\cdot))$, we have

$$\mathbb{E}\sup_{0\leqslant t\leqslant T} |P_t - P'_t|^2 \leqslant C d^2\big((v,\eta),(v',\eta')\big); \tag{11}$$

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}|Q_t - Q'_t|^2 + \mathbb{E}\int_0^T |K_t - K'_t|^2 \mathrm{d}t \leqslant C d^2\big((v,\eta),(v',\eta')\big). \tag{12}$$

3.2 Necessary condition

Now we can state the necessary condition for the near-optimality for problem NOC. To simplify notations, throughout the paper we denote $\Theta_t := (X_t, Y_t, Z_t)$, $\Theta_t^\varepsilon := (X_t^\varepsilon, Y_t^\varepsilon, Z_t^\varepsilon)$, $\widetilde{\Theta}_t := (\widetilde{X}_t, \widetilde{Y}_t, \widetilde{Z}_t)$, and $\widetilde{\Theta}_t^\varepsilon := (\widetilde{X}_t^\varepsilon, \widetilde{Y}_t^\varepsilon, \widetilde{Z}_t^\varepsilon)$.

Theorem 3.6. If Assumptions 2.1 and 3.4 hold true, then there exists a constant C such that for any $\varepsilon > 0$ and near-optimal pair $(X^\varepsilon(\cdot), Y^\varepsilon(\cdot), Z^\varepsilon(\cdot), v^\varepsilon(\cdot), \eta^\varepsilon(\cdot))$, we have

$$\begin{aligned}
&\mathbb{E}\int_0^T H_v\big(t, \Theta_t^\varepsilon, v_t^\varepsilon, P_t^\varepsilon, Q_t^\varepsilon, K_t^\varepsilon\big)(v_t - v_t^\varepsilon)\mathrm{d}t \geqslant -C\varepsilon^{\frac{1}{3}},\\
&\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i, \eta_i^\varepsilon) + C_{\tau_i} Q^\varepsilon(\tau_i) - D_{\tau_i} P^\varepsilon(\tau_i)\big)(\eta_i - \eta_i^\varepsilon) \geqslant -C\varepsilon^{\frac{1}{3}},
\end{aligned} \tag{13}$$

for any $(v_t, \eta_t) \in \mathcal{U} \times \mathcal{K}$.

Proof. To ease the symbol burden, let us first introduce some notations. By virtue of Assumption 2.1, $\mathcal{J}(0, x_0; v, \eta)$ is lower-semicontinuous under metric (4). Combining Ekeland's principle from Lemma 2.4 with $\lambda = \varepsilon^{\frac{2}{3}}$, there exists $(\widetilde{X}^\varepsilon(\cdot), \widetilde{v}^\varepsilon(\cdot), \widetilde{\eta}^\varepsilon(\cdot))$ such that

$$d\big((v^\varepsilon(\cdot), \eta^\varepsilon(\cdot)), (\widetilde{v}^\varepsilon(\cdot), \widetilde{\eta}^\varepsilon(\cdot))\big) \leqslant \varepsilon^{\frac{2}{3}}, \tag{14}$$

and

$$\widetilde{\mathcal{J}}(\widetilde{v}^{\varepsilon}(\cdot),\widetilde{\eta}^{\varepsilon}(\cdot)) \leqslant \widetilde{\mathcal{J}}(v(\cdot),\eta(\cdot)) \quad \text{for all } (v(\cdot),\eta(\cdot)) \in \mathcal{A}, \tag{15}$$

where the cost functional

$$\widetilde{\mathcal{J}}(v(\cdot),\eta(\cdot)) \triangleq \mathcal{J}(v(\cdot),\eta(\cdot)) + \varepsilon^{\frac{1}{3}} d\big(\big(v(\cdot),\eta(\cdot)\big),\big(\widetilde{v}^{\varepsilon}(\cdot),\widetilde{\eta}^{\varepsilon}(\cdot)\big)\big). \tag{16}$$

This implies that $(\widetilde{X}^{\varepsilon}(\cdot),\widetilde{v}^{\varepsilon}(\cdot),\widetilde{\eta}^{\varepsilon}(\cdot))$ is optimal for control system (1) with the new cost functional (16). For any $\rho > 0$, define the perturbed control variation

$$v^{\varepsilon,\rho}(t) = \widetilde{v}^{\varepsilon}(t) + \rho(v(t) - \widetilde{v}^{\varepsilon}(t)), \quad \forall\, v \in \mathcal{U}, \tag{17}$$

$$\eta^{\varepsilon,\rho}(t) = \widetilde{\eta}^{\varepsilon}(t) + \rho(\eta(t) - \widetilde{\eta}^{\varepsilon}(t)), \quad \forall\, \eta \in \mathcal{K}. \tag{18}$$

We can get the following fact:

(1) $\widetilde{\mathcal{J}}(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon}) \leqslant \widetilde{\mathcal{J}}(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho})$,

(2) $d\big(\big(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho}\big),\big(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon}\big)\big) \leqslant C\rho$, and

$$\begin{aligned}\mathcal{J}(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho}) - \mathcal{J}(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon}) &= \widetilde{\mathcal{J}}(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho}) - \varepsilon^{\frac{1}{3}} d\big(\big(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho}\big),\big(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon}\big)\big) - \widetilde{\mathcal{J}}(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon})\\ &\geqslant -\varepsilon^{\frac{1}{3}} d\big(\big(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho}\big),\big(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon}\big)\big) \geqslant -C\varepsilon^{\frac{1}{3}}\rho.\end{aligned} \tag{19}$$

Dividing both sides of the above inequality by ρ and sending it to zero, we have

$$\lim_{\rho\to 0} \rho^{-1}(\mathcal{J}(v^{\varepsilon,\rho},\eta^{\varepsilon,\rho}) - \mathcal{J}(\widetilde{v}^{\varepsilon},\widetilde{\eta}^{\varepsilon})) \geqslant -C\varepsilon^{\frac{1}{3}}. \tag{20}$$

Following similar arguments with [5, 13], the left hand side of the above inequality leads to

$$\begin{aligned}&\mathbb{E}\int_0^T H_v\big(t,\widetilde{\Theta}_t^{\varepsilon},\widetilde{v}_t^{\varepsilon},\widetilde{P}_t^{\varepsilon},\widetilde{Q}_t^{\varepsilon},\widetilde{K}_t^{\varepsilon}\big)(v_t - \widetilde{v}_t^{\varepsilon})\mathrm{d}t \geqslant -C_1\varepsilon^{\frac{1}{3}},\\ &\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\widetilde{\eta}_i^{\varepsilon}) + C_{\tau_i}\widetilde{Q}^{\varepsilon}(\tau_i) - D_{\tau_i}\widetilde{P}^{\varepsilon}(\tau_i)\big)(\eta_i - \widetilde{\eta}_i^{\varepsilon}) \geqslant -C_1\varepsilon^{\frac{1}{3}}.\end{aligned} \tag{21}$$

Now to prove the desired results (13), we need to estimate

$$\begin{aligned}&\mathbb{E}\int_0^T H_v\big(t,\widetilde{\Theta}_t^{\varepsilon},\widetilde{v}_t^{\varepsilon},\widetilde{P}_t^{\varepsilon},\widetilde{Q}_t^{\varepsilon},\widetilde{K}_t^{\varepsilon}\big)(v_t - \widetilde{v}_t^{\varepsilon})\mathrm{d}t - \mathbb{E}\int_0^T H_v\big(t,\Theta_t^{\varepsilon},v_t^{\varepsilon},P_t^{\varepsilon},Q_t^{\varepsilon},K_t^{\varepsilon}\big)(v_t - v_t^{\varepsilon})\mathrm{d}t,\\ &\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\widetilde{\eta}_i^{\varepsilon}) + C_{\tau_i}\widetilde{Q}^{\varepsilon}(\tau_i) - D_{\tau_i}\widetilde{P}^{\varepsilon}(\tau_i)\big)(\eta_i - \widetilde{\eta}_i^{\varepsilon})\\ &\quad - \mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\eta_i^{\varepsilon}) + C_{\tau_i}Q^{\varepsilon}(\tau_i) - D_{\tau_i}P^{\varepsilon}(\tau_i)\big)(\eta_i - \eta_i^{\varepsilon}).\end{aligned} \tag{22}$$

For sake of simplicity, we set

$$\begin{aligned}I_1(t) &:= H_v\big(t,\widetilde{\Theta}_t^{\varepsilon},\widetilde{v}_t^{\varepsilon},\widetilde{P}_t^{\varepsilon},\widetilde{Q}_t^{\varepsilon},\widetilde{K}_t^{\varepsilon}\big)\widetilde{v}_t^{\varepsilon} - H_v\big(t,\Theta_t^{\varepsilon},v_t^{\varepsilon},P_t^{\varepsilon},Q_t^{\varepsilon},K_t^{\varepsilon}\big)v_t^{\varepsilon};\\ I_2(t) &:= H_v\big(t,\widetilde{\Theta}_t^{\varepsilon},\widetilde{v}_t^{\varepsilon},\widetilde{P}_t^{\varepsilon},\widetilde{Q}_t^{\varepsilon},\widetilde{K}_t^{\varepsilon}\big)v - H_v\big(t,\Theta_t^{\varepsilon},v_t^{\varepsilon},P_t^{\varepsilon},Q_t^{\varepsilon},K_t^{\varepsilon}\big)v;\\ I_3(t) &:= \big(l_\eta(\tau_i,\widetilde{\eta}_i^{\varepsilon}) + C_{\tau_i}\widetilde{Q}_{\tau_i}^{\varepsilon} - D_{\tau_i}\widetilde{P}_{\tau_i}^{\varepsilon}\big)\widetilde{\eta}_i^{\varepsilon} - \big(l_\eta(\tau_i,\eta_i^{\varepsilon}) + C_{\tau_i}Q_{\tau_i}^{\varepsilon} - D_{\tau_i}P_{\tau_i}^{\varepsilon}\big)\eta_i^{\varepsilon};\\ I_4(t) &:= \big(l_\eta(\tau_i,\widetilde{\eta}_i^{\varepsilon}) + C_{\tau_i}\widetilde{Q}_{\tau_i}^{\varepsilon} - D_{\tau_i}\widetilde{P}_{\tau_i}^{\varepsilon}\big)\widetilde{\eta}_i - \big(l_\eta(\tau_i,\eta_i^{\varepsilon}) + C_{\tau_i}Q_{\tau_i}^{\varepsilon} - D_{\tau_i}P_{\tau_i}^{\varepsilon}\big)\eta_i.\end{aligned} \tag{23}$$

Then we have

$$\begin{aligned}&\mathbb{E}\int_0^T H_v\big(t,\widetilde{\Theta}_t^{\varepsilon},\widetilde{v}_t^{\varepsilon},\widetilde{P}_t^{\varepsilon},\widetilde{Q}_t^{\varepsilon},\widetilde{K}_t^{\varepsilon}\big)(v_t - \widetilde{v}_t^{\varepsilon})\mathrm{d}t - \mathbb{E}\int_0^T H_v\big(t,\Theta_t^{\varepsilon},v_t^{\varepsilon},P_t^{\varepsilon},Q_t^{\varepsilon},K_t^{\varepsilon}\big)(v_t - v_t^{\varepsilon})\mathrm{d}t\\ &\quad = \mathbb{E}\int_0^T \big(I_2(t) - I_1(t)\big)\mathrm{d}t,\end{aligned} \tag{24}$$

and

$$
\begin{aligned}
&\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\widetilde{\eta}_i^\varepsilon)+C_{\tau_i}\widetilde{Q}^\varepsilon_{\tau_i}-D_{\tau_i}\widetilde{P}^\varepsilon_{\tau_i}\big)(\eta_i-\widetilde{\eta}_i^\varepsilon)-\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\eta_i^\varepsilon)+C_{\tau_i}Q^\varepsilon_{\tau_i}-D_{\tau_i}P^\varepsilon_{\tau_i}\big)(\eta_i-\eta_i^\varepsilon)\\
&\quad=\mathbb{E}\sum_{i\geqslant 1}\big(I_4(t)-I_3(t)\big)\mathrm{d}t. \qquad (25)
\end{aligned}
$$

Using Schwarz's inequality and the boundness of H_v, we have

$$
\begin{aligned}
\mathbb{E}\int_0^T|I_1(t)|\mathrm{d}t &\leqslant \mathbb{E}\int_0^T|H_v\big(t,\widetilde{\Theta}_t^\varepsilon,\widetilde{v}_t^\varepsilon,\widetilde{P}_t^\varepsilon,\widetilde{Q}_t^\varepsilon,\widetilde{K}_t^\varepsilon\big)-H_v\big(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon\big)||\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&\quad+\mathbb{E}\int_0^T|H_v\big(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon\big)||v_t^\varepsilon-\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&\leqslant C_2\bigg(\mathbb{E}\int_0^T|v_t^\varepsilon-\widetilde{v}_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}+C_2\mathbb{E}\int_0^T|\langle b_v(t,\widetilde{X}_t^\varepsilon,\widetilde{v}_t^\varepsilon),\widetilde{Q}_t^\varepsilon\rangle-\langle b_v(t,X_t^\varepsilon,v_t^\varepsilon),Q_t^\varepsilon\rangle||\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&\quad+C_2\mathbb{E}\int_0^T|\langle \sigma_v(t,\widetilde{X}_t^\varepsilon,\widetilde{v}_t^\varepsilon),\widetilde{K}_t^\varepsilon\rangle-\langle \sigma_v(t,X_t^\varepsilon,v_t^\varepsilon),K_t^\varepsilon\rangle||\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&\quad+C_2\mathbb{E}\int_0^T|\langle f_v(t,\widetilde{\Theta}_t^\varepsilon,\widetilde{v}_t^\varepsilon),\widetilde{K}_t^\varepsilon\rangle-\langle f_v(t,\Theta_t^\varepsilon,v_t^\varepsilon),K_t^\varepsilon\rangle||\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&\quad+C_2\mathbb{E}\int_0^T|h_v(t,\widetilde{\Theta}_t^\varepsilon,\widetilde{v}_t^\varepsilon)-h_v(t,\Theta_t^\varepsilon,v_t^\varepsilon)||\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&=C_2\bigg(\mathbb{E}\int_0^T|v_t^\varepsilon-\widetilde{v}_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}+C_2\bigg(\mathbb{J}_1+\mathbb{J}_2+\mathbb{J}_3+\mathbb{J}_4\bigg). \qquad (26)
\end{aligned}
$$

Firstly, for $\mathbb{J}_1$, by the boundness of b_v, the fourth integral of $v(t)$, and Lemma 3.2, we have

$$
\begin{aligned}
\mathbb{J}_1&\leqslant \mathbb{E}\int_0^T|\langle b_v(t,\widetilde{X}_t^\varepsilon,\widetilde{v}_t^\varepsilon)-b_v(t,X_t^\varepsilon,v_t^\varepsilon),\widetilde{Q}_t^\varepsilon\rangle||\widetilde{v}_t^\varepsilon|\mathrm{d}t+\mathbb{E}\int_0^T|\langle b_v(t,X_t^\varepsilon,v_t^\varepsilon),\widetilde{Q}_t^\varepsilon-Q_t^\varepsilon\rangle||\widetilde{v}_t^\varepsilon|\mathrm{d}t\\
&\leqslant C_3\mathbb{E}\int_0^T\bigg(|\widetilde{X}_t^\varepsilon-X_t^\varepsilon|+|\widetilde{v}_t^\varepsilon-v_t^\varepsilon||\widetilde{Q}_t^\varepsilon||\widetilde{v}_t^\varepsilon|\bigg)\mathrm{d}t+C_3\bigg(\mathbb{E}\int_0^T|\widetilde{Q}_t^\varepsilon-Q_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\\
&\leqslant \widetilde{C}_4\bigg(\mathbb{E}\int_0^T|\widetilde{X}_t^\varepsilon-X_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\bigg(\mathbb{E}\int_0^T|\widetilde{Q}_t^\varepsilon\widetilde{v}_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\\
&\quad+\widetilde{C}_4\bigg(\mathbb{E}\int_0^T|\widetilde{v}_t^\varepsilon-v_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\bigg(\mathbb{E}\int_0^T|\widetilde{Q}_t^\varepsilon\widetilde{v}_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}+\widetilde{C}_4\bigg(\mathbb{E}\int_0^T|\widetilde{Q}_t^\varepsilon-Q_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\\
&\leqslant \widehat{C}_4\bigg(\mathbb{E}\int_0^T|\widetilde{X}_t^\varepsilon-X_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\bigg(\mathbb{E}\sup_{0\leqslant t\leqslant T}|\widetilde{Q}_t^\varepsilon|^4\mathrm{d}t\bigg)^{\frac14}\bigg(\mathbb{E}\int_0^T|\widetilde{v}_t^\varepsilon|^4\mathrm{d}t\bigg)^{\frac14}\\
&\quad+\widehat{C}_4\bigg(\mathbb{E}\int_0^T|\widetilde{v}_t^\varepsilon-v_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\bigg(\mathbb{E}\sup_{0\leqslant t\leqslant T}|\widetilde{Q}_t^\varepsilon|^4\mathrm{d}t\bigg)^{\frac14}\bigg(\mathbb{E}\int_0^T|\widetilde{v}_t^\varepsilon|^4\mathrm{d}t\bigg)^{\frac14}+\widehat{C}_4\bigg(\mathbb{E}\int_0^T|\widetilde{Q}_t^\varepsilon-Q_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\\
&\leqslant C_4\bigg(\mathbb{E}\int_0^T|\widetilde{X}_t^\varepsilon-X_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}+C_4\bigg(\mathbb{E}\int_0^T|\widetilde{v}_t^\varepsilon-v_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}+C_4\bigg(\mathbb{E}\int_0^T|\widetilde{Q}_t^\varepsilon-Q_t^\varepsilon|^2\mathrm{d}t\bigg)^{\frac12}\\
&\leqslant C_5 d\big((v_t,\eta_t),(\widetilde{v}_t^\varepsilon,\widetilde{\eta}_t^\varepsilon)\big)\\
&\leqslant C_5\varepsilon^{\frac23}. \qquad (27)
\end{aligned}
$$

Using similar arguments for $\mathbb{J}_1$, $\mathbb{J}_2$, $\mathbb{J}_3$, $\mathbb{J}_4$, we can prove that

$$
\mathbb{E}\int_0^T|I_1(t)|\mathrm{d}t\leqslant C_6\varepsilon^{\frac23}. \qquad (28)
$$

About $I_2(t)$, $I_3(t)$ and $I_4(t)$, using the similar method with $I_1(t)$, we can conclude that

$$\mathbb{E}\int_0^T \big(|I_1(t)|+|I_2(t)|\big)\mathrm{d}t \leqslant C_7\varepsilon^{\frac{2}{3}}, \tag{29}$$

$$\mathbb{E}\sum_{i\geqslant 1}\big(|I_3(t)|+|I_4(t)|\big) \leqslant C_7\varepsilon^{\frac{2}{3}}. \tag{30}$$

Then we have

$$\begin{aligned}&\mathbb{E}\int_0^T H_v\big(t,\widetilde{\Theta}_t^\varepsilon,\widetilde{v}_t^\varepsilon,\widetilde{P}_t^\varepsilon,\widetilde{Q}_t^\varepsilon,\widetilde{K}_t^\varepsilon\big)(v_t-\widetilde{v}_t^\varepsilon)\mathrm{d}t\\&-\mathbb{E}\int_0^T H_v\big(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon\big)(v_t-v_t^\varepsilon)\mathrm{d}t \leqslant C_7\varepsilon^{\frac{2}{3}},\end{aligned} \tag{31}$$

and

$$\begin{aligned}&\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\widetilde{\eta}_i^\varepsilon)+C_{\tau_i}\widetilde{Q}_{\tau_i}^\varepsilon-D_{\tau_i}\widetilde{P}_{\tau_i}^\varepsilon\big)(\eta_i-\widetilde{\eta}_i^\varepsilon)\\&-\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\eta_i^\varepsilon)+C_{\tau_i}Q_{\tau_i}^\varepsilon-D_{\tau_i}P_{\tau_i}^\varepsilon\big)(\eta_i-\eta_i^\varepsilon) \leqslant C_7\varepsilon^{\frac{2}{3}}.\end{aligned} \tag{32}$$

Finally, combining (21), (31), and (32), we can obtain

$$\begin{aligned}&\mathbb{E}\int_0^T H_v\big(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon\big)(v_t-v_t^\varepsilon)\mathrm{d}t\\&=\mathbb{E}\int_0^T H_v\big(t,\widetilde{\Theta}_t^\varepsilon,\widetilde{v}_t^\varepsilon,\widetilde{P}_t^\varepsilon,\widetilde{Q}_t^\varepsilon,\widetilde{K}_t^\varepsilon\big)(v_t-\widetilde{v}_t^\varepsilon)\mathrm{d}t-\mathbb{E}\int_0^T H_v\big(t,\widetilde{\Theta}_t^\varepsilon,\widetilde{v}_t^\varepsilon,\widetilde{P}_t^\varepsilon,\widetilde{Q}_t^\varepsilon,\widetilde{K}_t^\varepsilon\big)(v_t-\widetilde{v}_t^\varepsilon)\mathrm{d}t\\&\quad-\mathbb{E}\int_0^T H_v\big(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon\big)(v_t-v_t^\varepsilon)\mathrm{d}t\\&\geqslant -C\varepsilon^{\frac{1}{3}}-C_7\varepsilon^{\frac{2}{3}}\\&\geqslant -C_8\varepsilon^{\frac{1}{3}}.\end{aligned} \tag{33}$$

Similarly we can get the second estimate of (13). The proof is complete.

Remark 3.7. In case $\varepsilon=0$ in (13), Theorem 3.6 is just the necessary condition for exact optimality of FBSDEs with impulse control.

4 Sufficient condition for near-optimality

This section gives a sufficient condition for the near-optimality under convexity. Our main result is as follows.

Theorem 4.1. Suppose Assumptions 2.1 and 3.4 hold true. Let v^ε be admissible control, $(X^\varepsilon,Y^\varepsilon,Z^\varepsilon)$ and $(P^\varepsilon,Q^\varepsilon,K^\varepsilon)$ be the solutions of (1) and (6) respectively. Moreover, for $M\in\mathbb{R}^{m\times n}$, $\zeta\in L^2(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^m)$, $Y_T^{v,\eta}$ has the following form: $Y_T^{v,\eta}=MX_T^{v,\eta}+\zeta,\ \forall\,(v,\eta)\in\mathcal{U}\times\mathcal{K}$. Suppose that $H(t,\cdot,\cdot,\cdot,\cdot,P^\varepsilon,Q^\varepsilon,K^\varepsilon)$ is convex for a.e. $t\in[0,T]$, $\mathbb{P}-$a.s.. Moreover, ϕ, γ, l are convex. If for some $\varepsilon>0$,

$$\begin{aligned}&\mathbb{E}\int_0^T H_v\big(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon\big)(v_t-v_t^\varepsilon)\mathrm{d}t\geqslant -\varepsilon^{\frac{1}{3}},\\&\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\eta_i^\varepsilon)+C_{\tau_i}Q^\varepsilon(\tau_i)-D_{\tau_i}P^\varepsilon(\tau_i)\big)(\eta_i-\eta_i^\varepsilon)\geqslant -\varepsilon^{\frac{1}{3}},\end{aligned} \tag{34}$$

then

$$\mathcal{J}(v^\varepsilon,\eta^\varepsilon)\leqslant \min_{(v,\eta)\in\mathcal{U}\times\mathcal{K}}\mathcal{J}(v,\eta)+C\varepsilon^{\frac{1}{3}}. \tag{35}$$

Proof. From (2) and the definition of Hamiltonian function H, we have

$$\mathcal{J}(v^\varepsilon,\eta^\varepsilon)-\mathcal{J}(v,\eta)=I_1+I_2+I_3+I_4-I_5, \tag{36}$$

where

$$\begin{aligned}
&I_1:=\mathbb{E}[\phi(X_T^\varepsilon)-\phi(X_T)];\\
&I_2:=\mathbb{E}[\gamma(Y_0^\varepsilon)-\gamma(Y_0)];\\
&I_3:=\mathbb{E}\int_0^T H(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon)\mathrm{d}t-\mathbb{E}\int_0^T H(t,\Theta_t,v_t,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon)\mathrm{d}t;\\
&I_4:=\mathbb{E}\sum_{\tau_i\leqslant T}[l(\tau_i,\eta_i^\varepsilon)-l(\tau_i,\eta_i)];\\
&I_5:=\mathbb{E}\int_0^T\langle b(t,X_t^\varepsilon,v_t^\varepsilon)-b(t,X_t,v_t),Q_t^\varepsilon\rangle\mathrm{d}t+\mathbb{E}\int_0^T\langle -f(t,\Theta_t^\varepsilon,v_t^\varepsilon)+f(t,\Theta_t,v_t),P_t^\varepsilon\rangle\mathrm{d}t\\
&\qquad+\mathbb{E}\int_0^T\langle\sigma(t,X_t^\varepsilon,v_t^\varepsilon)-\sigma(t,X_t,v_t),K_t^\varepsilon\rangle\mathrm{d}t.
\end{aligned}\tag{37}$$

Since both H and $\mathcal{U}$ are convex,

$$\begin{aligned}
&I_1\leqslant\mathbb{E}\langle\phi_x(X_T^\varepsilon)(X_T^\varepsilon-X_T)\rangle;\\
&I_2\leqslant\mathbb{E}\langle\gamma_y(Y_0^\varepsilon),(Y_0^\varepsilon-Y_0)\rangle;\\
&I_3\leqslant\mathbb{E}\int_0^T\Big[\langle H_x(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),X_t^\varepsilon-X_t\rangle+\langle H_y(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),Y_t^\varepsilon-Y_t\rangle\Big]\mathrm{d}t\\
&\qquad+\mathbb{E}\int_0^T\Big[\langle H_z(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),Z_t^\varepsilon-Z_t\rangle\mathrm{d}t+\langle H_v(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),v_t^\varepsilon-v_t\rangle\Big]\mathrm{d}t;\\
&I_4\leqslant\mathbb{E}\sum_{\tau_i\leqslant T}l_\eta(\tau_i,\eta_i^\varepsilon)(\eta_i^\varepsilon-\eta_i).
\end{aligned}\tag{38}$$

Applying Ito's formula to $\langle Q_t^\varepsilon,X_t^\varepsilon-X_t\rangle$ and $\langle P_t^\varepsilon,Y_t^\varepsilon-Y_t\rangle$, we have

$$\begin{aligned}
\mathbb{E}\langle\phi_x(X_T^\varepsilon),X_T^\varepsilon-X_T\rangle=\;&\mathbb{E}\int_0^T\langle H_x(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),X_t-X_t^\varepsilon\rangle\mathrm{d}t\\
&+\mathbb{E}\int_0^T\langle b(t,X_t^\varepsilon,v_t^\varepsilon)-b(t,X_t,v_t),Q_t^\varepsilon\rangle\mathrm{d}t\\
&+\mathbb{E}\int_0^T\langle\sigma(t,X_t^\varepsilon,v_t^\varepsilon)-\sigma(t,X_t,v_t),K_t^\varepsilon\rangle\mathrm{d}t\\
&+\mathbb{E}[\langle M^\tau P_T^\varepsilon,X_T^\varepsilon-X_T\rangle]+\mathbb{E}\sum_{i\geqslant1}C_{\tau_i}Q^\varepsilon(\tau_i)(\eta_i^\varepsilon-\eta_i);
\end{aligned}\tag{39}$$

and

$$\begin{aligned}
\mathbb{E}\langle\gamma_y(Y_0^\varepsilon),Y_0^\varepsilon-Y_0\rangle=\;&\mathbb{E}\int_0^T\langle H_y(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),Y_t-Y_t^\varepsilon\rangle\mathrm{d}t\\
&+\mathbb{E}\int_0^T\langle H_z(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),Z_t-Z_t^\varepsilon\rangle\mathrm{d}t\\
&+\mathbb{E}\int_0^T\langle -f(t,X_t^\varepsilon,v_t^\varepsilon)+f(t,X_t,v_t),P_t^\varepsilon\rangle\mathrm{d}t\\
&+\mathbb{E}[\langle P_T^\varepsilon,M(X_T-X_T^\varepsilon)\rangle]-\mathbb{E}\sum_{i\geqslant1}D_{\tau_i}P^\varepsilon(\tau_i)(\eta_i^\varepsilon-\eta_i).
\end{aligned}\tag{40}$$

Substitute (37)–(40) into (36), and

$$\mathcal{J}(v^\varepsilon,\eta^\varepsilon)-\mathcal{J}(v,\eta)\leqslant\mathbb{E}\int_0^T\langle H_v(t,\Theta_t^\varepsilon,v_t^\varepsilon,P_t^\varepsilon,Q_t^\varepsilon,K_t^\varepsilon),v_t^\varepsilon-v_t\rangle\mathrm{d}t$$
$$+\mathbb{E}\sum_{i\geqslant1}\big(l_\eta(\tau_i,\eta_i^\varepsilon)+C_{\tau_i}Q^\varepsilon(\tau_i)-D_{\tau_i}P^\varepsilon(\tau_i)\big)(\eta_i^\varepsilon-\eta_i). \tag{41}$$

From the condition (34), we can easily get the desired results. The proof is complete.

Remark 4.2. In case $\varepsilon=0$ in (34), Theorem 4.1 is just the sufficient condition for exact optimality of FBSDEs.

5 One example

In this section, we propose one example and show how to get its near-optimal control using our maximum principle. To simplify our analysis, here we set the admissible control domain $U\times K=[0,1]\times[0,1]$ and (X,Y,Z) are all one-dimensional.

Example 5.1. The controlled forward-backward stochastic system is

$$\begin{cases}\mathrm{d}X_t=(b_1X_t+b_2v_t)\mathrm{d}t+\sigma_1X_t\mathrm{d}B_t+C_t\mathrm{d}\eta_t,\\ \mathrm{d}Y_t=-(f_1X_t+f_2Y_t+f_3Z_t+f_4v_t)\mathrm{d}t+Z_t\mathrm{d}B_t-D_t\mathrm{d}\eta_t,\\ X_0=a,\quad Y_T=hX_T,\quad t\in[0,T].\end{cases}\tag{42}$$

Here, X_t represents some stochastic value (such as the surplus process in insurance management, or stochastic production process in inventory management) process with control variable v_t, and Y_t denotes a recursive utility which has some coupling structure to the value state X_t. Moreover, the cost functional to be minimized over the admissible control set $\mathcal{A}$ is

$$\mathcal{J}^\varepsilon(v,\eta)=\frac{1}{2}\mathbb{E}\bigg[LX_T+MY_0+\int_0^T\varepsilon h(v_t)\mathrm{d}t+S\sum_{\tau_i\leqslant T}\eta_i^2\bigg],$$

where $\varepsilon>0$ is small enough, $S>0$ and h (independent of ε) is a nonlinear convex function satisfying

$$0\leqslant h'(v)\leqslant C,\quad 0\leqslant h''(v)\leqslant C,\quad \forall\, v\in\mathcal{U},$$

for constant C.

The cost functional in our system is nonlinear and we consider the near-optimal control. For further analysis, we assume $C_t=D_t=0$ and consider the case of the impulse control appearing in cost functional only. Our following result can also be generalized to the case $C_t^2+D_t^2\neq0$ with more additional computations. Now we can write down the Hamiltonian function of our example as follows:

$$H^\varepsilon(t,x,y,z,v,p,q,k)=(b_1x+b_2v)q+\sigma_1xk-(f_1x+f_2y+f_3z+f_4v)p+\frac{1}{2}\varepsilon h(v).$$

It follows that

$$H_v=b_2q-f_4p+\frac{1}{2}\varepsilon h'(v).$$

The adjoint equation becomes

$$\begin{cases}\mathrm{d}P_t=f_2P_t\mathrm{d}t+f_3P_t\mathrm{d}B_t,\\ -\mathrm{d}Q_t=(b_1Q_t+\sigma_1K_t-f_1P_t)\mathrm{d}t-K_t\mathrm{d}B_t,\\ P_0=-M,\quad Q_T=-hP_T+L,\quad t\in[0,T].\end{cases}\tag{43}$$

From the necessary condition for near-optimality, for $\varepsilon > 0$, $\forall\,(v_t,\eta_t) \in \mathcal{U}\times\mathcal{K}$, the near-optimal pair $(X^\varepsilon(\cdot), Y^\varepsilon(\cdot), Z^\varepsilon(\cdot), v^\varepsilon(\cdot), \eta^\varepsilon(\cdot))$ should satisfy

$$\mathbb{E}\int_0^T H_v\big(t, X_t^\varepsilon, Y_t^\varepsilon, Z_t^\varepsilon, v_t^\varepsilon, P_t^\varepsilon, Q_t^\varepsilon, K_t^\varepsilon\big)(v_t - v_t^\varepsilon)\mathrm{d}t \geqslant -C\varepsilon^{\frac{1}{3}},$$

and

$$\mathbb{E}\sum_{i\geqslant 1}\big(l_\eta(\tau_i,\eta_i^\varepsilon) + C_{\tau_i}Q^\varepsilon(\tau_i) - D_{\tau_i}P^\varepsilon(\tau_i)\big)(\eta_i - \eta_i^\varepsilon) \geqslant -C\varepsilon^{\frac{1}{3}}. \tag{44}$$

That is,

$$\mathbb{E}\int_0^T \left(b_2Q_t^\varepsilon - f_4P_t^\varepsilon + \frac{1}{2}\varepsilon h'(v_t^\varepsilon)\right)(v_t - v_t^\varepsilon)\mathrm{d}t \geqslant -C\varepsilon^{\frac{1}{3}}, \tag{45}$$

and

$$\mathbb{E}\sum_{i\geqslant 1} S\eta_i^\varepsilon(\eta_i - \eta_i^\varepsilon) \geqslant -C\varepsilon^{\frac{1}{3}}. \tag{46}$$

First we need to find the near-optimal control of v^ε. Using some basic techniques, we know the inequality (45) takes the following form:

$$\begin{aligned}&\mathbb{E}\int_0^T \left(b_2Q_t^\varepsilon - f_4P_t^\varepsilon + \frac{1}{2}\varepsilon h'(v_t^\varepsilon)\right)v_t\mathrm{d}t\\ &\geqslant \mathbb{E}\int_0^T \left[\frac{1}{2}\varepsilon h'(v_t^\varepsilon)v_t^\varepsilon + (b_2Q_t^\varepsilon - f_4P_t^\varepsilon)v_t^\varepsilon - \frac{C\varepsilon^{\frac{1}{3}}}{T}\right]\mathrm{d}t.\end{aligned} \tag{47}$$

Under the condition $0 \leqslant h'(v) \leqslant C$ and $v_t \in U = [0,1]$, the left hand side of the above inequality can be relaxed to

$$\mathbb{E}\int_0^T \left(b_2Q_t^\varepsilon - f_4P_t^\varepsilon + \frac{1}{2}\varepsilon h'(v_t^\varepsilon)\right)v_t\mathrm{d}t \geqslant \mathbb{E}\int_0^T (b_2Q_t^\varepsilon - f_4P_t^\varepsilon)v_t\mathrm{d}t. \tag{48}$$

Then we only need to find the proper v_t^ε such that

$$\begin{aligned}\mathbb{E}\int_0^T \left(b_2Q_t^\varepsilon - f_4P_t^\varepsilon + \frac{1}{2}\varepsilon h'(v_t^\varepsilon)\right)v_t\mathrm{d}t &\geqslant \mathbb{E}\int_0^T (b_2Q_t^\varepsilon - f_4P_t^\varepsilon)v_t\mathrm{d}t\\ &\geqslant \mathbb{E}\int_0^T \left[\frac{1}{2}\varepsilon h'(v_t^\varepsilon)v_t^\varepsilon + (b_2Q_t^\varepsilon - f_4P_t^\varepsilon)v_t^\varepsilon - \frac{C\varepsilon^{\frac{1}{3}}}{T}\right]\mathrm{d}t.\end{aligned} \tag{49}$$

As to the coefficients in our system, we give the following assumptions:

$$h > 0,\quad a > 0,\quad b_2 > 0,\quad f_1 > 0,\quad f_4 > 0,\quad L < 0,\quad M < 0,\quad S > 0. \tag{50}$$

In adjoint equation (43), we can get

$$P_t = -M\exp\left[\left(f_2 - \frac{1}{2}(f_3)^2\right)t - f_3(B_t - B_0)\right]. \tag{51}$$

By the condition $M < 0$, we can deduce that $P_t > 0$, $\forall\,t \in [0,T]$. In particular, $P_T > 0$.

Under the condition $h > 0$, $L < 0$ and $f_1 > 0$, we know for any given $\mathcal{F}_{\tau_i}$-adapted square integral process P_t, there exists a unique (Q_t, K_t) satisfying the second equation of (43). Moreover, using the comparison theorem of BSDE, we have $Q_t < 0$, $\forall\,t \in [0,T]$. This implies that $b_2Q_t^\varepsilon - f_4P_t^\varepsilon < 0$ for all $t \in [0,T]$. Then the problem can be transformed to find some proper v^ε such that

$$\mathbb{E}\int_0^T\left(b_2Q_t^\varepsilon - f_4P_t^\varepsilon + \frac{1}{2}\varepsilon h'(v_t^\varepsilon)\right)v_t\mathrm{d}t \geqslant \mathbb{E}\int_0^T(b_2Q_t^\varepsilon - f_4P_t^\varepsilon)v_t\mathrm{d}t \geqslant \mathbb{E}\int_0^T(b_2Q_t^\varepsilon - f_4P_t^\varepsilon)\mathrm{d}t$$
$$\geqslant \mathbb{E}\int_0^T\left[\frac{1}{2}\varepsilon h'(v_t^\varepsilon)v_t^\varepsilon + (b_2Q_t^\varepsilon - f_4P_t^\varepsilon)v_t^\varepsilon - \frac{C\varepsilon^{\frac{1}{3}}}{T}\right]\mathrm{d}t. \tag{52}$$

That is,

$$\mathbb{E}\int_0^T\left[\frac{1}{2}\varepsilon h'(v_t^\varepsilon)v_t^\varepsilon + (b_2Q_t^\varepsilon - f_4P_t^\varepsilon)(v_t^\varepsilon - 1) - \frac{C\varepsilon^{\frac{1}{3}}}{T}\right]\mathrm{d}t \leqslant 0. \tag{53}$$

For any given small enough $\varepsilon > 0$, we know

$$v^\varepsilon(t) \equiv 1 - \varepsilon^{\frac{1}{2}} \tag{54}$$

is a near-optimal control of this problem.

Next we turn to find the near-optimal impulse control η^ε. From Eq. (46) and the control domain of η, we know Eq. (46) leads to

$$\mathbb{E}\sum_{i\geqslant 1}S\eta_i^\varepsilon\eta_i \geqslant \mathbb{E}\sum_{i\geqslant 1}S(\eta_i^\varepsilon)^2 - C\varepsilon^{\frac{1}{3}}. \tag{55}$$

It follows that the left side of above inequality is non-negative. Then the problem turns to find the near-optimal control such that

$$\mathbb{E}\sum_{i\geqslant 1}S(\eta_i^\varepsilon)^2 \leqslant C\varepsilon^{\frac{1}{3}}.$$

For any given sufficiently small $\varepsilon > 0$, we know

$$\eta^\varepsilon \equiv \varepsilon^{\frac{1}{2}}$$

is a near-optimal control for our impulse control component.

6 Conclusion

The subject of near optimality has been studied by some researchers, such as Zhou [8], etc. In this paper, we investigate the near-optimality for a class of stochastic impulse control problems. To our best knowledge, this problem has never been touched although it deserves some research attention. Some classical techniques for optimal controls do not work in our problem, and some new techniques were introduced. In this sense, our study is different from the prior studies relating to classical exact optimality. By Ekeland's principle and some delicate estimates, this paper establishes a necessary condition and a sufficient condition of near-optimality with stochastic impulse control problems in terms of a small parameter ε. Our work is partly based on the work from [8,9,10], and hopefully this result derived in this paper may inspire some applications in finance or engineering.

Acknowledgements This work was supported by RGC Earmarked (Grant Nos. 501010, 502412) and National Natural Science Foundation of China (Grant No. 11401345).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Davis M H, Norman A. Portfolio selection with transaction costs. Math Oper Res, 1990, 15: 676–713
2 Fudenberg D, Levine K. A dual-self model of impulse control. American Econ Rev, 2006, 96: 1449–1476

3 Samuel M, David M, Laibson I, et al. Separate neural systems value immediate and delayed monetary rewards. Science, 2004, 306: 503–507

4 Cadenillas A, Haussmann U G. The stochastic maximum principle for a singular control problem. Stoch Stoch Reports, 1994, 49: 211–237

5 Wu Z, Zhang F. Stochastic maximum principle for optimal control problems of forward-backward systems involving impulse control. IEEE Trans Autom Control, 2011, 56: 1401–1406

6 Wu Z, Zhang F. Maximum principle for stochastic recursive optimal control problems involving impulse controls. Abstract Appl Anal, 2012, 2012: 709682

7 Yu Z. The stochastic maximum principle for optimal control problems of delay systems involving continuous and impulse controls. Automatica, 2012, 48: 2420–2432

8 Zhou X Y. Stochastic near-optimal controls: necessary and sufficient conditions for near-optimality. SIAM J Control Optimiz, 1998, 36: 929–947

9 Huang J, Li X, Wang G. Near-optimal control problems for linear forward-backward stochastic systems. Automatica, 2010, 46: 397–404

10 Hui E, Huang J, Wang G. Near-optimal control for stochastic recursive problems. Syst Control Lett, 2011, 60: 161–168

11 Meng Q, Shen Y. A revisit to stochastic near-optimal controls: the critical case. Syst Control Lett, 2015, 82: 79–85

12 Ekeland I. Nonconvex minimization problems. Bull Amer Math Soc, 1979, 1: 443–474

13 Peng S. Backward stochastic differetial equations and applications to optimal control. Appl Math Optim, 1993, 27: 125–144