Systems & Control Letters 58 (2009) 857–864

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

# Necessary and sufficient conditions for near-optimality in stochastic control of FBSDEs

Khaled Bahlali [a], Nabil Khelfallah [b], Brahim Mezerdi [b,*]

[a] IMATH, UFR Sciences, USTV, B.P. 20132, 83957 La Garde Cedex, France

[b] Laboratory of Applied Mathematics, University Med Khider, Po. Box 145, Biskra (07000), Algeria

ARTICLE INFO

*Article history:*
Received 26 June 2009
Accepted 19 October 2009
Available online 14 November 2009

*Keywords:*
Forward–backward stochastic differential equations
Stochastic control
Maximum principle
Adjoint equation
Ekeland's variational principal

ABSTRACT

We consider control problems for systems governed by a nonlinear forward backward stochastic differential equation (FBSDE). We establish necessary as well as sufficient conditions for near optimality, satisfied by all near optimal controls. These conditions are described by two adjoint processes, corresponding to the forward and backward components and a nearly maximum condition on the Hamiltonian. The proof of the main result is based on Ekeland's variational principle and some estimates on the state and the adjoint processes with respect to the control variable. As is well known, optimal controls may fail to exist even in simple cases. This justifies the use of near optimal controls, which exist under minimal assumptions and are sufficient in most practical cases. Moreover, since there are many nearly optimal controls, it is possible to choose suitable ones, that are convenient for implementation.



## 1. Introduction

Since the work of Pardoux & Peng [1], the theory of backward stochastic differential equations (BSDEs) has found important applications and has become a powerful tool in many fields, such as mathematical finance, optimal control, stochastic games, partial differential equations and homogenization, see [2,3]. Therefore, it is natural to investigate control problems for systems governed by this kind of stochastic equations. Stochastic control problems for systems driven by FBSDEs or BSDEs have been studied by many authors, see e.g. [4–7] and the references therein. These papers have been devoted to various forms of the stochastic maximum principle, for systems of BSDEs and FBSDEs. As is well known, optimal controls may fail to exist even in simple cases. This justifies the use of near optimal controls, which exist under minimal assumptions and are sufficient in most practical cases. Moreover, since there are many nearly optimal controls, it is possible to choose suitable ones that are convenient for analysis and implementation.

Our goal in this paper is to study near optimal, rather than optimal controls for systems driven by FBSDEs. More precisely, we establish necessary as well as sufficient conditions of near-optimality. These conditions are described in terms of two adjoint processes, corresponding to the forward and backward components and a nearly maximum condition on the hamiltonian. In a second step, we prove that under additional concavity conditions, these necessary conditions of near-optimality are also sufficient. Our main result is based on Ekeland's variational principal and some estimates on the state and the adjoint processes.

Let us recall that many works have been devoted to the problem of near-optimality. In deterministic control problems, driven by ordinary differential equations, the first result on necessary conditions for near optimality has been proved by Ekeland [8], see also [9]. Near optimal control problems for systems driven by Itô stochastic differential equations with an uncontrolled diffusion coefficient, have been investigated in [10–12]. The general case of systems driven by SDE with controlled diffusion coefficient has been treated by Zhou [13], where necessary as well as sufficient conditions are established, for all nearly optimal controls. See also [9].

The paper is organised as follows. The assumptions, notations and some basic definitions are given in Section 2. In Sections 3 and 4, we establish necessary as well as sufficient conditions of near optimality.

## 2. Statement of the problem

Let $(\Omega, \mathcal{F}, P)$ be a probability space, equipped with a filtration $(\mathcal{F}_t)$, satisfying the usual conditions, on which a $\mathbb{R}^d$-valued standard Brownian motion $W(.)$ is defined. We assume that $(\mathcal{F}_t)$ is the $P$-augmentation of the natural filtration of $W(.)$.

* Corresponding author. Tel.: +213 33 74 77 88; fax: +213 33 74 77 88.
*E-mail addresses:* bahlali@univ-tln.fr (K. Bahlali), nabilkhefallah@yahoo.fr (N. Khelfallah), bmezerdi@yahoo.fr (B. Mezerdi).

0167-6911/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2009.10.005

We consider a stochastic control problem, where the control domain need not be convex and the system is governed by a forward–backward stochastic differential equation (FBSDE in short) of the type

$$\begin{cases} \mathrm{d}x(t) = f(t, x(t), u(t))\,\mathrm{d}t + \sigma(t, x(t))\,\mathrm{d}W_t, \\ x(0) = x_0, \\ \mathrm{d}y(t) = g(t, x(t), y(t), z(t), u(t))\,\mathrm{d}t + z(t)\,\mathrm{d}W_t, \\ y(T) = h(x(T)), \end{cases}$$

where $f, \sigma, g$ and $h$ are given maps. The control variable $u = (u_t)$ is an $\mathcal{F}_t$-adapted process with values in some set $U$ of $\mathbb{R}^k$. We denote by $\mathcal{U}_{ad}$ the set of all admissible controls. The expected cost on the time interval $[0, T]$ is

$$J(u(.)) = E\varphi(y(0)),$$

and the value function is defined as follows :

$$V = \inf_{u(.)\in\mathcal{U}_{ad}[0,T]} J(u(.)).$$

Since the objective of this paper is to study near-optimal rather than optimal controls of the system, we give the precise definition of near-optimality as given in Zhou [13]. For a given $\varepsilon > 0$, $u^\varepsilon(.)$ is called $\varepsilon$-optimal if

$$|J(u^\varepsilon(.)) - V| \le r(\varepsilon),$$

holds for sufficiently small $\varepsilon$, where $r$ is a function of $\epsilon$ satisfying $r(\varepsilon) \to 0$ as $\epsilon \to 0$. The estimate $r(\varepsilon)$ is called an error bound. If $r(\varepsilon) = C\varepsilon^\delta$ for some $\delta > 0$ independent of the constant $C$, then $u^\varepsilon(.)$ is called near-optimal with order $\varepsilon^\delta$.

Throughout this paper we assume the following

$$f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n,$$
$$\sigma : [0, T] \times \mathbb{R}^n \to \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^n\right),$$
$$g : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^m\right) \times \mathbb{R}^k \to \mathbb{R}^m,$$
$$h : \mathbb{R}^n \to \mathbb{R}^m.$$
$$\varphi : \mathbb{R}^m \to \mathbb{R}.$$

$(H_1)$: $f, g, \sigma, h, \varphi$ are continuous in $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^m\right) \times \mathbb{R}^k$, and continuously differentiable with respect to $(x, y, z)$.

$(H_2)$: The derivatives of $f, g, \sigma, h, \varphi$ with respect to $x, y, z$ are bounded and there is a constant $C > 0$ such that $g$ is bounded by $C(1 + |x| + |y|)$.

$(H_3)$: There is a constant $C > 0$ and $\beta \in [0, 1]$ such that

$$\left|f_x(t, x, u) - f_x\left(t, \acute{x}, u\right)\right| + \left|\sigma_x(t, x) - \sigma_x\left(t, \acute{x}\right)\right| + \left|h_x(x) - h_x\left(\acute{x}\right)\right| \le C\left|x - \acute{x}\right|^\beta,$$

$$\left|\rho(t, x, y, z, u) - \rho\left(t, \acute{x}, \acute{y}, \acute{z}, u\right)\right| \le C\left(\left|x - \acute{x}\right|^\beta + \left|y - \acute{y}\right|^\beta + \left|z - \acute{z}\right|^\beta\right) \quad \text{for } \rho = g_x, g_y, g_z.$$

Under assumptions $(H_1)$ and $(H_2)$, there is a unique triple

$$(x(.), y(.), z(.)) \in L^2_{\mathcal{F}}\left(0, T, \mathbb{R}^n\right) \times L^2_{\mathcal{F}}\left(0, T, \mathbb{R}^m\right) \times L^2_{\mathcal{F}}\left(0, T, \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^m\right)\right),$$

which solves $(E)$, where $L^2_{\mathcal{F}}(0, T, \mathbb{R}^n)$ denotes the Hilbert space of $\mathcal{F}_t$-adapted processes $X$ such that $E\int_0^T |X(s)|^2\,\mathrm{d}s < +\infty$.

For any $u(.) \in \mathcal{U}_{ad}$ with its corresponding state trajectory $(x(.), y(.), z(.))$, we introduce the adjoint equations and the Hamiltonian function for our problem. The adjoint equations are defined by

$$\begin{cases} -\mathrm{d}\Psi(t) = \left(f_x^*\Psi(t) + g_x^*Q(t) + \sigma_x^*K(t)\right)\mathrm{d}t - K(t)\,\mathrm{d}W_t, \\ \Psi(T) = -h_x^*(x(T))\,Q(T), \\ -\mathrm{d}Q(t) = g_y^*Q(t)\,\mathrm{d}t + g_z^*Q(t)\,\mathrm{d}W_t, \\ Q(0) = -\varphi_y(y(0)), \end{cases} \tag{2.1}$$

and the Hamiltonian function

$$H : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^m\right) \times \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^m \times \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^n\right) \to \mathbb{R}^n$$

is given by

$$H(t, x, y, z, u, \Psi, Q, K) := -\langle\Psi, f(t, x, u)\rangle - \langle Q, g(t, x, y, z, u)\rangle - \langle K, \sigma(t, x)\rangle.$$

The adjoint equations can be rewritten in terms of the derivatives of the Hamiltonian as

$$\begin{cases} \mathrm{d}\Psi(t) = H_x(t, x(t), y(t), z(t), u(t), \Psi(t), Q(t), K(t))\,\mathrm{d}t \\ \quad + K(t)\,\mathrm{d}W_t, \\ \Psi(T) = -h_x^*(x(T))\,Q(T), \\ \mathrm{d}Q(t) = H_y(t, x(t), y(t), z(t), u(t), \Psi(t), Q(t), K(t))\,\mathrm{d}t \\ \quad + H_z(t, x(t), y(t), z(t), u(t), \Psi(t), Q(t), K(t))\,\mathrm{d}W_t, \\ Q(0) = -\varphi_y(y(0)). \end{cases}$$

Note that the couple $(\Psi, K)$ is the adjoint process corresponding to the forward component $x(.)$ and $Q(.)$ is the adjoint process corresponding to the backward component $(y(.), z(.))$. It is a well known fact that under assumptions $(H_1)$, $(H_2)$, the backward adjoint equation admits one and only one $\mathcal{F}_t$-adapted solution $(\Psi, K) \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^n) \times L^2_{\mathcal{F}}\left(0, T, \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^n\right)\right)$ and the forward adjoint equation admits one and only one $\mathcal{F}_t$-adapted solution $Q \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^m)$. Moreover, since $f_x, \sigma_x, g_x, g_y, g_z$ are bounded by, there exists a constant $C_1 > 0$, independent of $(x(.), y(.), z(.), u(.))$, such that the solutions of adjoint equations satisfy the following estimate:

$$E(\sup_{0\le t\le T}|\Psi(t)|^2) + E(\sup_{0\le t\le T}|Q(t)|^2) + E\int_0^T |K(s)|^2\,\mathrm{d}s \le C_1. \tag{2.2}$$

Let us recall Ekeland's variational principle, which will be used in the sequel.

**Lemma 1** (*Ekeland [8]*). *Let $(V, d)$ be a complete metric space and $f : V \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, bounded from below. If for each $\varepsilon > 0$, there exists $u^\varepsilon \in V$ such that $f(u^\varepsilon) \le \inf_{u\in V} f(u) + \varepsilon$. Then for any $\delta > 0$, there exists $u^\delta \in V$ such that*

(i) $f\left(u^\delta\right) \le f(u^\varepsilon)$,
(ii) $d\left(u^\delta, u^\varepsilon\right) \le \delta$,
(iii) $f\left(u^\delta\right) \le f(u) + \dfrac{\varepsilon}{\delta} d\left(u, u^\delta\right)$, *for all $u \in V$.*

For $u, v$ in $\mathcal{U}_{ad}$, we define

$$d(u, v) = \mathrm{d}t \otimes P\{(t, \omega) \in [0, T] \times \Omega : u(t, \omega) \ne v(t, \omega)\}, \tag{2.3}$$

where $\mathrm{d}t \otimes P$ is the product measure of the Lebesgue measure $\mathrm{d}t$ with the probability measure $P$. It is well known that $(\mathcal{U}_{ad}, d)$ is a complete metric space (see [11,13]).

## 3. Necessary conditions of near-optimality

In this section we derive necessary conditions for a control to be near-optimal. This is the main result of this paper.

**Theorem 2.** *For any $\gamma \in \left[0, \frac{1}{3}\right)$, there exist a constant $C_1 = C_1(\gamma) > 0$ such that for any $\varepsilon > 0$ and any $\varepsilon$-optimal control $u^\varepsilon$, it holds that*

$$E\int_0^T \{\Psi^\varepsilon(t)\,[f(t, x^\varepsilon(t), u) - f(t, x^\varepsilon(t), u^\varepsilon(t))] + Q^\epsilon(t)\,[g(t, \lambda^\epsilon(t), u) - g(t, \lambda^\epsilon(t), u^\varepsilon(t))]\}\,\mathrm{d}t \ge -C_1\varepsilon^\gamma, \quad \text{for all } u \in U, \tag{3.1}$$

*where $\lambda^\epsilon(.) := (x^\varepsilon(.), y^\varepsilon(.), z^\varepsilon(.))$ denotes the solution of the state equation and $(\Psi^\varepsilon(t), K^\varepsilon(t), Q^\epsilon(t))$ is the solution of the adjoint equation, corresponding to $u^\varepsilon$.*

**Corollary 3.** *Under the conditions of Theorem 2, it holds that*

$$E\int_0^T H\left(t,\lambda^{\epsilon}(t),u^{\varepsilon}(t),\Lambda^{\varepsilon}(t)\right)\mathrm{d}t \geq E\int_0^T H\left(t,\lambda^{\epsilon}(t),u(t),\Lambda^{\varepsilon}(t)\right)\mathrm{d}t - C_1\varepsilon^{\gamma}, \quad \text{for all } u\in\mathcal{U}_{ad} \tag{3.2}$$

*where* $\Lambda^{\varepsilon}(t)=(\Psi^{\varepsilon}(t),K^{\varepsilon}(t),Q^{\epsilon}(t))$ *and* $\lambda^{\epsilon}(t):=(x^{\varepsilon}(t),y^{\varepsilon}(t),z^{\varepsilon}(t))$.

To prove Theorem 2 and Corollary 3, we need the following auxiliary results on the stability of the state and adjoint processes with respect to the control variable.

**Lemma 4.** *For any* $0<\alpha<1$ *and* $0<p\leq 2$, *there is a constant* $C_1=C_1(\alpha,p)>0$, *such that for any* $u(.),\acute{u}(.)\in\mathcal{U}_{ad}$, *along with the corresponding trajectories* $\lambda(.):=(x(.),y(.),z(.))$, $\acute{\lambda}(.):=(\acute{x}(.),\acute{y}(.),\acute{z}(.))$, *it holds that*

$$E\left[\sup_{0\leq t\leq T}\left|x(t)-\acute{x}(t)\right|^p\right]\leq C_1 d\left(u(.),\acute{u}(.)\right)^{\frac{\alpha p}{2}}, \tag{3.3}$$

$$\sup_{0\leq t\leq T}E\left|y(t)-\acute{y}(t)\right|^p+E\int_0^T\left|z(t)-\acute{z}(t)\right|^p\mathrm{d}s \leq C_1 d\left(u(.),\acute{u}(.)\right)^{\frac{\alpha p}{2}}. \tag{3.4}$$

**Proof.** The proof of (3.3) can be found in [13] Lemma 2.1. We need only treat (3.4). First we assume $p=2$. Squaring both sides of

$$-\left(y(t)-\acute{y}(t)\right)-\int_t^T\left(z(s)-\acute{z}(s)\right)\mathrm{d}W_s = -\left(h(x(T))-h\left(\acute{x}(T)\right)\right) + \int_t^T\left(g(s,\lambda(s),u(s))-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right)\mathrm{d}s$$

and using the fact that

$$E\left[y(t)-\acute{y}(t)\int_t^T\left(z(s)-\acute{z}(s)\right)\mathrm{d}W_s\right]=0,$$

we get

$$E\left|y(t)-\acute{y}(t)\right|^2+E\int_t^T\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s \leq C_2E\left|h(x(T))-h\left(\acute{x}(T)\right)\right|^2 + C_2E\left\{\int_t^T\left|g(s,\lambda(s),u(s))-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|\mathrm{d}s\right\}^2. \tag{3.5}$$

Let us estimate the first term in the right hand side of (3.5)

$$E\left|h(x(T))-h\left(\acute{x}(T)\right)\right|^2 \leq CE\left|x(T)-\acute{x}(T)\right|^2\leq C_3 d\left(u(.),\acute{u}(.)\right)^{\alpha}. \tag{3.6}$$

Now let us turn to the second term of (3.5)

$$\begin{aligned}
&E\left\{\int_t^T\left|g(s,\lambda(s),u(s))-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|\mathrm{d}s\right\}^2\\
&\leq E\left\{\int_t^T\left|g(s,\lambda(s),u(s))-g\left(s,\acute{\lambda}(s),u(s)\right)\right|\mathrm{d}s + \int_t^T\left|g\left(s,\acute{\lambda}(s),u(s)\right)-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right| \times\chi_{u(s)\neq\acute{u}(s)}(s)\,\mathrm{d}s\right\}^2,\\
&\leq C_4E\left\{\int_t^T\left|g(s,\lambda(s),u(s))-g\left(s,\acute{\lambda}(s),u(s)\right)\right|\mathrm{d}s\right\}^2 + C_4E\int_t^T\left|g\left(s,\acute{\lambda}(s),u(s)\right)-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|^2 \times\chi_{u(s)\neq\acute{u}(s)}(s)\,\mathrm{d}s.
\end{aligned}$$

Taking $\acute{q}=\frac{1}{1-\alpha}>1$ and $\acute{p}=\frac{1}{\alpha}>1$ such that $\frac{1}{\acute{p}}+\frac{1}{\acute{q}}=1$ and applying Hölder's inequality, we obtain

$$\begin{aligned}
&E\left\{\int_t^T\left|g\left(s,\acute{\lambda}(s),u(s)\right)-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|^2\chi_{u(s)\neq\acute{u}(s)}(s)\,\mathrm{d}s\right\}\\
&\leq\left\{E\int_t^T\left|g\left(s,\acute{\lambda}(s),u(s)\right)-g\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|^{\frac{2}{1-\alpha}}\mathrm{d}s\right\}^{1-\alpha} \times\left\{E\int_t^T\chi_{u(s)\neq\acute{u}(s)}(s)\,\mathrm{d}s\right\}^{\alpha}\\
&\quad\times\left\{1+E\left(\sup_{s\in[t,T]}\left|\acute{x}(s)\right|^{\frac{2}{1-\alpha}}\right)+E\left(\sup_{s\in[t,T]}\left|\acute{y}(s)\right|^{\frac{2}{1-\alpha}}\right)\right\}^{1-\alpha} \times\left\{E\int_t^T\chi_{u(s)\neq\acute{u}(s)}(s)\,\mathrm{d}s\right\}^{\alpha}\\
&\leq C_5 d\left(u(.),\acute{u}(.)\right)^{\alpha}.
\end{aligned} \tag{3.7}$$

Then

$$E\left\{\int_t^T\left|g(s,\lambda(s),u(s))-g\left(s,\acute{\lambda}(s),u(s)\right)\right|\mathrm{d}s\right\}^2 \leq C_6TE\int_t^T\left|x(s)-\acute{x}(s)\right|^2\mathrm{d}s+C_6TE\int_t^T\left|y(s)-\acute{y}(s)\right|^2\mathrm{d}s + C_6(T-t)E\int_t^T\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s. \tag{3.8}$$

By (3.6)–(3.8) we obtain

$$E\left|y(t)-\acute{y}(t)\right|^2+E\int_t^T\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s \leq C_7 d\left(u(.),\acute{u}(.)\right)^{\alpha}+C_6TE\int_t^T\left|y(s)-\acute{y}(s)\right|^2\mathrm{d}s + C_6(T-t)E\int_t^T\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s.$$

For every $\delta$, such that $T-t=\delta$, we obtain by choosing $\delta=\frac{1}{2C_8}$,

$$E\left|y(t)-\acute{y}(t)\right|^2+\frac{1}{2}E\int_{T-\delta}^T\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s \leq C_9 d\left(u(.),\acute{u}(.)\right)^{\alpha}+C_9E\int_{T-\delta}^T\left|y(s)-\acute{y}(s)\right|^2\mathrm{d}s.$$

Applying Gronwall's inequality, we obtain

$$E\left|y(t)-\acute{y}(t)\right|^2+\frac{1}{2}E\int_t^T\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s \leq C_{10} d\left(u(.),\acute{u}(.)\right)^{\alpha},\quad t\in[T-\delta,T].$$

Similarly we get

$$E\left|y(t)-\acute{y}(t)\right|^2+E\int_t^{T-\delta}\left|z(s)-\acute{z}(s)\right|^2\mathrm{d}s \leq C_{10} d\left(u(.),\acute{u}(.)\right)^{\alpha},\quad t\in[T-2\delta,T-\delta].$$

After a finite number of iterations, we obtain the desired result. Now assume $0 < p < 2$. Then (3.4) follows immediately from Holder's inequality. This completes the proof of Lemma 4. ■

The following lemma gives the continuity of the solutions to the adjoint equations with respect to the control variable. It plays a key role in proving the necessary condition.

**Lemma 5.** *For any $0 < \alpha < 1$ and $1 < p < 2$ satisfying $(1+\alpha\beta)p < 2$, there is a constant $C_1 = C_1(\alpha,\beta,p) > 0$ such that for any $u(.), \acute{u}(.) \in \mathcal{U}_{ad}$, along with the corresponding trajectories $(x(.), y(.), z(.))$, $(\acute{x}(.), \acute{y}(.), \acute{z}(.))$ and the solutions $(\Psi(.), Q(.), K(.))$, $(\acute{\Psi}(.), \acute{Q}(.), \acute{K}(.))$ of the corresponding adjoint equations, it hold that*

$$E\int_0^T \left\{\left|\Psi(t)-\acute{\Psi}(t)\right|^p + \left|K(t)-\acute{K}(t)\right|^p\right\} dt \leq C_1 d\left(u(.),\acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}, \tag{3.9}$$

*and*

$$E\int_0^T \left|Q(t)-\acute{Q}(t)\right|^p dt \leq C_1 d\left(u(.),\acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.10}$$

**Proof.** We proceed to prove (3.10) then (3.9). First, note that $\bar{Q}(t) = Q(t) - \acute{Q}(t)$ satisfies the following SDE

$$\begin{cases} d\bar{Q}(t) = \left\{g_y\bar{Q}(t) + G_y(t)\right\}dt + \left\{g_z\bar{Q}(t) + G_z(t)\right\}dW_t \\ \bar{Q}(0) = -\left\{\varphi_y(y(0)) - \varphi_y(\acute{y}(0))\right\}, \end{cases} \tag{3.11}$$

in which

$$G_y(t) = \left\{g_y(t,\lambda(t),u(t)) - g_y\left(t,\acute{\lambda}(t),u(t)\right)\right\}\acute{Q}(t),$$

$$G_z(t) = \left\{g_z(t,\lambda(t),u(t)) - g_z\left(t,\acute{\lambda}(t),u(t)\right)\right\}\acute{Q}(t).$$

First we assume $p = 2$. Since we have

$$-\bar{Q}(t) = \left\{\varphi_y(y(0)) - \varphi_y(\acute{y}(0))\right\} + \int_0^t \left\{g_y\bar{Q}(s) + G_y(s)\right\}ds + \int_0^t \left\{g_z\bar{Q}(s) + G_z(s)\right\}dW_s,$$

then by squaring both sides of the above equality, taking the expectation and using the fact that $g_y, g_z$ are bounded by $C$, we get

$$E\left|\bar{Q}(t)\right|^2 \leq C_2\left\{E\left|y(0)-\acute{y}(0)\right|^2 + E\int_0^T\left|\bar{Q}(s)\right|^2 ds + E\int_0^T\left\{\left|G_y(s)\right|^2 + \left|G_z(s)\right|^2\right\}ds\right\}.$$

Then by Gronwall's lemma it holds that

$$E\left|\bar{Q}(t)\right|^2 \leq C_3\left\{E\left|y(0)-\acute{y}(0)\right|^2 + E\int_0^T\left\{\left|G_y(s)\right|^2 + \left|G_z(s)\right|^2\right\}ds\right\}.$$

We shall estimate the right hand side of above inequality. From Lemma 5, we have

$$E\left|y(0)-\acute{y}(0)\right|^2 \leq C_4 d\left(u(.),\acute{u}(.)\right)^{\alpha}.$$

Then

$$\begin{aligned} E\int_0^T\left|G_y(s)\right|^2 ds &\leq C_5E\int_0^T\left|g_y(s,\lambda(s),u(s)) - g_y\left(s,\acute{\lambda}(s),u(s)\right)\right|^2\left|\acute{Q}(s)\right|^2 dt \\ &+ C_5E\int_0^T\left|g_y\left(s,\acute{\lambda}(s),u(s)\right) - g_y\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|^2 \\ &\times\left|\acute{Q}(s)\right|^2 dt, \\ &\leq C_5E\int_0^T\left\{\chi_{u(s)\neq\acute{u}(s)}(s)\left|\acute{Q}(s)\right|^2 + \left|x(s)-\acute{x}(s)\right|^{2\beta}\left|\acute{Q}(s)\right|^2\right. \\ &\left.+\left|y(s)-\acute{y}(s)\right|^{2\beta}\left|\acute{Q}(s)\right|^2 + \left|z(s)-\acute{z}(s)\right|^{2\beta}\left|\acute{Q}(s)\right|^2\right\}ds, \end{aligned}$$

By Hölder's inequality, we obtain

$$\begin{aligned} E\int_0^T\left|G_y(s)\right|^2 ds &\leq C_6\left\{E\int_0^T\left|\acute{Q}(s)\right|^{\frac{2}{1-\alpha\beta}}ds\right\}^{1-\alpha\beta} d\left(u(.),\acute{u}(.)\right)^{\alpha\beta}, \\ &+ C_6\left\{E\int_0^T\left|\acute{Q}(s)\right|^{\frac{2}{1-\beta}}ds\right\}^{1-\beta}\left\{E\int_0^T\left|x(s)-\acute{x}(s)\right|^2 ds\right\}^{\beta} \\ &+ C_6\left\{E\int_0^T\left|\acute{Q}(s)\right|^{\frac{2}{1-\beta}}ds\right\}^{1-\beta}\left\{E\int_0^T\left|y(s)-\acute{y}(s)\right|^2 ds\right\}^{\beta} \\ &+ C_6\left\{E\int_0^T\left|\acute{Q}(s)\right|^{\frac{2}{1-\beta}}ds\right\}^{1-\beta}\left\{E\int_0^T\left|z(s)-\acute{z}(s)\right|^2 ds\right\}^{\beta}. \end{aligned}$$

From Lemma 5 and using the fact that $E\sup_{0\leq t\leq T}\left|\acute{Q}(t)\right|^p < \infty$ for any $p$, it can easily checked that

$$\begin{aligned} E\int_0^T\left|G_y(s)\right|^2 ds = E\int_0^T &\left|g_y(s,\lambda(s),u(s))\right. \\ &\left.- g_y\left(s,\acute{\lambda}(s),\acute{u}(s)\right)\right|^2\left|\acute{Q}(s)\right|^2 ds, \\ &\leq C_7 d\left(u(.),\acute{u}(.)\right)^{\alpha\beta}. \end{aligned}$$

It follows easily by the same arguments that

$$E\int_0^T\left|G_z(s)\right|^2 ds \leq C_8 d\left(u(.),\acute{u}(.)\right)^{\alpha\beta}.$$

So

$$E\left|\bar{Q}(t)\right|^2 \leq C_9 d\left(u(.),\acute{u}(.)\right)^{\alpha\beta}.$$

Now assume $0 < p < 2$. Then by Hölder's inequality we obtain the inequality (3.10)

Now, let us prove (3.9). Noting that $(\bar{\Psi}(t), \bar{K}(t)) \equiv (\Psi(t) - \acute{\Psi}(t), K(t) - \acute{K}(t))$ satisfies the following BSDE:

$$\begin{cases} d\bar{\Psi}(t) - \left\{\left(f_x^*(t,x(t),u(t))\bar{\Psi}(t) + \sigma_x^*(t,x(t))\bar{K}(t)\right)\right. \\ \quad\left. + F(t)\right\}dt + \bar{K}(t)dW_t, \\ \bar{\Psi}(T) = -\left\{h_x^*(x(T))Q(T) - h_x^*\left(\acute{x}(T)\right)\acute{Q}(T)\right\}, \end{cases}$$

where

$$\begin{aligned} F(t) &= \left\{f_x^*(t,x(t),u(t)) - f_x^*\left(t,\acute{x}(t),\acute{u}(t)\right)\right\}\acute{\Psi}(t) \\ &+ g_x^*(t,\lambda(t),u(t))Q(t) - g_x^*\left(t,\acute{\lambda}(t),\acute{u}(t)\right)\acute{Q}(t) \\ &+ \left\{\sigma_x^*(t,x(t)) - \sigma_x^*\left(t,\acute{x}(t)\right)\right\}\acute{K}(t). \end{aligned}$$

Let $\eta$ be the solution of the following linear SDE:

$$\begin{cases} d\eta(t) = \left\{ f_x(t, x(t), u(t))\,\eta(t) + \left|\bar{\Psi}(t)\right|^{p-1} \operatorname{sgn}\left(\bar{\Psi}(t)\right) \right\} dt \\ \quad + \left\{ \sigma_x(t, x(t))\,\eta(t) + \left|\bar{K}(t)\right|^{p-1} \operatorname{sgn}\left(\bar{K}(t)\right) \right\} dW_t, \\ \eta(0) = 0, \end{cases}$$

where $\operatorname{sgn}(a) \equiv \left(\operatorname{sgn}\left(a^1\right), \ldots, \operatorname{sgn}(a^n)\right)^*$ for any vector $a \equiv \left(a^1, \ldots, a^n\right)^*$. Note that the existence and uniqueness of solutions to the above equation are verified by assumptions $(H_1)$, $(H_2)$ and the fact that

$$E \int_0^T \left\{ \left| \left|\bar{\Psi}(t)\right|^{p-1} \operatorname{sgn}\left(\bar{\Psi}(t)\right) \right|^2 + \left| \left|\bar{K}(t)\right|^{p-1} \operatorname{sgn}\left(\bar{K}(t)\right) \right|^2 \right\} dt < +\infty.$$

Let $q > 2$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$E \sup_{0 \le t \le T} |\eta(t)|^q \le C_2 E \int_0^T \left\{ \left|\bar{\Psi}(t)\right|^{pq-q} + \left|\bar{K}(t)\right|^{pq-q} \right\} dt$$
$$\le C_2 E \int_0^T \left\{ \left|\bar{\Psi}(t)\right|^p + \left|\bar{K}(t)\right|^p \right\} dt. \tag{3.12}$$

Note that the right hand side term of (3.12) is bounded due to (2.2). On the other hand, applying Ito's formula to $\bar{\Psi}(t) \cdot \eta(t)$ and taking expectations, we obtain

$$\begin{aligned} &E \int_0^T \left\{ \bar{\Psi}(t) \cdot \left[ \left|\bar{\Psi}(t)\right|^{p-1} \operatorname{sgn}\left(\bar{\Psi}(t)\right) \right] + \bar{K}(t) \cdot \left[ \left|\bar{K}(t)\right|^{p-1} \operatorname{sgn}\left(\bar{K}(t)\right) \right] \right\} dt \\ &= E \left\{ \int_0^T [F(t) \cdot \eta(t)]\, dt + \left[ h_x^*(x(T))\,Q(T) - h_x^*\left(\acute{x}(T)\right) \acute{Q}(T) \right] \cdot \eta(T) \right\} \\ &\le \left\{ E \int_0^T |F(t)|^p\, dt \right\}^{\frac{1}{p}} \left\{ E \int_0^T |\eta(t)|^q\, dt \right\}^{\frac{1}{q}} \\ &\quad + \left\{ E \left| h_x^*(x(T))\,Q(T) - h_x^*\left(\acute{x}(T)\right) \acute{Q}(T) \right|^p \right\}^{\frac{1}{p}} \left\{ E\,|\eta(T)|^q \right\}^{\frac{1}{q}} \\ &\le C_3 \left\{ E \int_0^T \left[ \left|\bar{\Psi}(t)\right|^p + \left|\bar{K}(t)\right|^p \right] dt \right\}^{\frac{1}{q}} \left\{ \left[ E \int_0^T |F(t)|^p\, dt \right]^{\frac{1}{p}} \right. \\ &\quad \left. + \left[ E \left| h_x^*(x(T))\,Q(T) - h_x^*\left(\acute{x}(T)\right) \acute{Q}(T) \right|^p \right]^{\frac{1}{p}} \right\}, \end{aligned}$$

Thus

$$E \int_0^T \left\{ \left|\bar{\Psi}(t)\right|^p + \left|\bar{K}(t)\right|^p \right\} dt \le C_4 \left\{ E \int_0^T |F(t)|^p\, dt + E \left| h_x^*(x(T))\,Q(T) - h_x^*\left(\acute{x}(T)\right) \acute{Q}(T) \right|^p \right\}. \tag{3.13}$$

We proceed to estimate the second term in the right hand side of (3.13). First,

$$\begin{aligned} &E \left| h_x^*(x(T))\,Q(T) - h_x^*\left(\acute{x}(T)\right) \acute{Q}(T) \right|^p \\ &= E \left| \left\{ h_x(x(T)) - h_x\left(\acute{x}(T)\right) \right\}^* Q(T) - h_x^*\left(\acute{x}(T)\right) \bar{Q}(T) \right|^p \\ &\le C_5 E \left\{ |Q(T)|^2 \right\}^{\frac{p}{2}} \left\{ E \left| h_x(x(T)) - h_x\left(\acute{x}(T)\right) \right|^{\frac{2p}{2-p}} \right\}^{1-\frac{p}{2}} + C_5 E \left|\bar{Q}(T)\right|^p \\ &\le C_6 \left\{ E \left| x(T) - \acute{x}(T) \right|^{\frac{2\beta p}{2-p}} \right\}^{\frac{2-p}{2}} + C_5 E \left|\bar{Q}(T)\right|^p. \end{aligned}$$

Using Lemma 5 and (3.10) it is easy to see that

$$E \left| h_x^*(x(T))\,Q(T) - h_x^*\left(\acute{x}(T)\right) \acute{Q}(T) \right|^p \le C_7 d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.14}$$

Next, we proceed to estimate the first term in the right side of (3.13). First

$$\begin{aligned} &E \int_0^T \left| g_x^*(t, \lambda(t), u(t))\,Q(t) - g_x^*\left(t, \acute{\lambda}(t), \acute{u}(t)\right) \acute{Q}(t) \right|^p dt \\ &\le C_8 E \int_0^T \left| g_x^*(t, \lambda(t), u(t)) - g_x^*\left(t, \acute{\lambda}(t), \acute{u}(t)\right) \right|^p \\ &\quad \times |Q(t)|^p\, dt + C_8 E \int_0^T \left| g_x^*\left(t, \acute{\lambda}(t), \acute{u}(t)\right) \right|^p \left|\bar{Q}(t)\right|^p dt. \end{aligned} \tag{3.15}$$

We estimate now the right hand side of (3.15). Using similar arguments to estimate the term $E \int_0^T \left|G_y(s)\right|^p ds$, we get

$$E \int_0^T \left| g_x^*(t, \lambda(t), u(t)) - g_x^*\left(t, \acute{\lambda}(t), \acute{u}(t)\right) \right|^p |Q(t)|^p\, dt \le C_9 d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.16}$$

Then by (3.10) and using the fact that $g_x$ is bounded, we obtain

$$E \int_0^T \left| g_x^*\left(t, \acute{\lambda}(t), \acute{u}(t)\right) \right|^p \left|\bar{Q}(t)\right|^p dt \le C_{10} d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.17}$$

From (3.16) and (3.17) we conclude

$$E \int_0^T \left| g_x^*(t, \lambda(t), u(t))\,Q(t) - g_x^*\left(t, \acute{\lambda}(t), \acute{u}(t)\right) \acute{Q}(t) \right|^p dt \le C_{11} d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.18}$$

Using similar arguments developed above, we can easily prove that

$$E \int_0^T \left| f_x^*(t, x(t), u(t)) - f_x^*\left(t, \acute{x}(t), \acute{u}(t)\right) \right|^p \left|\acute{\Psi}(t)\right|^p dt \le C_{12} d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.19}$$

and

$$E \int_0^T \left| \sigma_x^*(t, x(t)) - \sigma_x^*\left(t, \acute{x}(t)\right) \right|^p \left|\acute{K}(t)\right|^p dt \le C_{13} d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.20}$$

It follows from (3.18)–(3.20) that

$$E \int_0^T |F(t)|^p\, dt \le C_{14} d\left(u(.), \acute{u}(.)\right)^{\frac{\alpha\beta p}{2}}. \tag{3.21}$$

The desired result (3.9) follows immediately from (3.14) and (3.21) This completes the proof of Lemma 5. ∎

**Proof of Theorem 2.** By assumptions $(H_1)$, $(H_2)$, it is easy to see that $J(u(.))$ is continuous on $\mathcal{U}_{ad}$ endowed with the metric defined by (2.3). Applying Ekeland's variational principle with $\delta = \varepsilon^{\frac{2}{3}}$, there is an admissible control $\tilde{u}^{\varepsilon}(.)$ such that

$$d\left(u^{\varepsilon}(.), \tilde{u}^{\varepsilon}(.)\right) \leq \varepsilon^{\frac{2}{3}},$$

$$\tilde{J}\left(\tilde{u}^{\varepsilon}(.)\right) \leq \tilde{J}(u(.)) \quad \text{for any } u(.) \in \mathcal{U}_{ad},$$

where

$$\tilde{J}(u(.)) = J(u(.)) + \varepsilon^{\frac{1}{3}} d\left(u(.), \tilde{u}^{\varepsilon}(.)\right).$$

This means that $\tilde{u}^{\varepsilon}(.)$ is optimal for the system with the new cost function $\tilde{J}$. Let $t_0 \in [0, T)$ and $u \in U$ be fixed. For any $\theta > 0$, define the spike variation $u^{\theta} \in \mathcal{U}_{ad}[0, T]$ of $\tilde{u}^{\varepsilon}$:

$$u^{\theta}(t) = \begin{cases} u & t \in [t_0, t_0 + \theta] \\ \tilde{u}^{\varepsilon}(t), & \text{otherwise.} \end{cases}$$

Let us denote $\tilde{\lambda}^{\epsilon}(.) := \left(\tilde{x}^{\varepsilon}(.), \tilde{y}^{\varepsilon}(.), \tilde{z}^{\varepsilon}(.)\right)$. The fact that

$$\tilde{J}\left(\tilde{u}^{\varepsilon}(.)\right) \leq \tilde{J}\left(u^{\theta}(.)\right)$$

and

$$d\left(u^{\theta}(.), \tilde{u}^{\varepsilon}(.)\right) \leq \theta,$$

imply that

$$J\left(u^{\theta}(.)\right) - J\left(\tilde{u}^{\varepsilon}(.)\right) \geq -\varepsilon^{\frac{1}{3}}\theta. \tag{3.22}$$

Arguing as in [6], Theorem 1, the left hand side of inequality (3.22) is equal to

$$E\int_{t_0}^{t_0+\theta} \left\{\tilde{\Psi}^{\varepsilon}(t)\left[f\left(t, \tilde{x}^{\varepsilon}(t), u\right) - f\left(t, \tilde{x}^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right]\right.$$
$$\left. + \tilde{Q}^{\epsilon}(t)\left[g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right]\right\} \mathrm{d}t.$$

Dividing (3.22) by $\theta$ and sending $\theta$ to 0, we conclude that

$$E\left\{\tilde{\Psi}^{\varepsilon}(t)\left[f\left(t, \tilde{x}^{\varepsilon}(t), u\right) - f\left(t, \tilde{x}^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right]\right.$$
$$\left. + \tilde{Q}^{\epsilon}(t)\left[g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right]\right\} \mathrm{d}t$$
$$\geq -\varepsilon^{\frac{1}{3}}. \tag{3.23}$$

To prove (3.1), it remains to estimate the following difference

$$E\int_0^T \tilde{Q}^{\epsilon}(t)\left[g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right] \mathrm{d}t$$
$$- E\int_0^T Q^{\epsilon}(t)\left[g\left(t, \lambda^{\epsilon}(t), u\right) - g\left(t, \lambda^{\epsilon}(t), u^{\varepsilon}(t)\right)\right] \mathrm{d}t$$
$$= E\int_0^T \left\{\tilde{Q}^{\epsilon}(t) - Q^{\epsilon}(t)\right\}\left\{g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right)\right.$$
$$\left. - g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right\} \mathrm{d}t$$
$$+ E\int_0^T Q^{\epsilon}(t)\left\{g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right) - g\left(t, \lambda^{\epsilon}(t), u\right)\right\} \mathrm{d}t$$
$$- E\int_0^T Q^{\epsilon}(t)\left\{g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right) - g\left(t, \lambda^{\epsilon}(t), u^{\varepsilon}(t)\right)\right\} \mathrm{d}t$$
$$= I_1 + I_2 + I_3.$$

For any $\gamma \in \left[0, \frac{1}{3}\right)$, let $\alpha\beta = 3\gamma \in [0, 1)$. Then, in view of Lemma 5, we have

$$I_1 \leq \left\{E\int_0^T \left|\tilde{Q}^{\epsilon}(t) - Q^{\epsilon}(t)\right|^2 \mathrm{d}t\right\}^{\frac{1}{2}}$$
$$\times \left\{E\int_0^T \left|g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right|^2 \mathrm{d}t\right\}^{\frac{1}{2}}$$
$$\leq C_2\left\{d\left(u^{\varepsilon}(.), \tilde{u}^{\varepsilon}(.)\right)^{\alpha\beta}\right\}^{\frac{1}{2}}$$
$$\times C_3\left\{E\int_0^T \left(1 + \left|\tilde{x}^{\varepsilon}(t)\right|^2 + \left|\tilde{y}^{\varepsilon}(t)\right|^2\right) \mathrm{d}s\right\}^{\frac{1}{2}}$$
$$\leq C_4\left\{\varepsilon^{\frac{2\alpha\beta}{3}}\right\}^{\frac{1}{2}} = C_4\varepsilon^{\gamma}.$$

Next, by using Lemma 5, we get

$$I_2 \leq \left\{E\int_0^T \left|Q^{\epsilon}(t)\right|^2 \mathrm{d}t\right\}^{\frac{1}{2}}$$
$$\times \left\{E\int_0^T \left|g\left(t, \tilde{\lambda}^{\epsilon}(t), u\right) - g\left(t, \lambda^{\epsilon}(t), u\right)\right|^2 \mathrm{d}t\right\}^{\frac{1}{2}}$$
$$\leq C_4\left\{E\int_0^T \left|\tilde{x}^{\varepsilon}(t) - x^{\varepsilon}(t)\right|^2 \mathrm{d}t + E\int_0^T \left|\tilde{y}^{\varepsilon}(t) - y^{\varepsilon}(t)\right|^2 \mathrm{d}t\right.$$
$$\left. + E\int_0^T \left|\tilde{z}^{\varepsilon}(t) - z^{\varepsilon}(t)\right|^2 \mathrm{d}t\right\}^{\frac{1}{2}}$$
$$\leq C_5\left\{d\left(u^{\varepsilon}(.), \tilde{u}^{\varepsilon}(.)\right)^{\alpha\beta}\right\}^{\frac{1}{2}}$$
$$\leq C_5\left\{\varepsilon^{\frac{2\alpha\beta}{3}}\right\}^{\frac{1}{2}} = C_5\varepsilon^{\gamma}.$$

Further, fix $p \in (1, 2)$ so that $(1 + \alpha\beta)p < 2$. and taking $q \in (2, +\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, it holds by using Lemma 5, that

$$I_3 = -E\int_0^T Q^{\epsilon}(t)\left\{g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right.$$
$$\left. - g\left(t, \lambda^{\epsilon}(t), u^{\varepsilon}(t)\right)\right\} \mathrm{d}t$$
$$= -E\int_0^T Q^{\epsilon}(t)\left\{g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right.$$
$$\left. - g\left(t, \tilde{\lambda}^{\epsilon}(t), u^{\varepsilon}(t)\right)\right\} \mathrm{d}t, -E\int_0^T Q^{\epsilon}(t)$$
$$\times \left\{g\left(t, \tilde{\lambda}^{\epsilon}(t), u^{\varepsilon}(t)\right) - g\left(t, \lambda^{\epsilon}(t), u^{\varepsilon}(t)\right)\right\} \mathrm{d}t$$
$$\leq \left\{E\int_0^T \left|Q^{\epsilon}(t)\right|^q \mathrm{d}t\right\}^{\frac{1}{q}} \left\{E\int_0^T \left|g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right)\right.\right.$$
$$\left.\left. - g\left(t, \tilde{\lambda}^{\epsilon}(t), u^{\varepsilon}(t)\right)\right|^p \chi_{\tilde{u}^{\varepsilon}(t) \neq u^{\varepsilon}(t)}(t) \mathrm{d}t\right\}^{\frac{1}{p}} + C_5\varepsilon^{\gamma}.$$

By Hölder's inequality, one has

$$E\int_0^T \left|g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), u^{\varepsilon}(t)\right)\right|^p$$
$$\times \chi_{\tilde{u}^{\varepsilon}(t) \neq u^{\varepsilon}(t)}(t) \mathrm{d}t$$
$$\leq \left\{E\int_0^T \left|g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), u^{\varepsilon}(t)\right)\right|^2 \mathrm{d}t\right\}^{\frac{p}{2}}$$
$$\times \left\{E\int_0^T \chi_{\tilde{u}^{\varepsilon}(t) \neq u^{\varepsilon}(t)}(t) \mathrm{d}t\right\}^{1-\frac{p}{2}},$$
$$\leq C_6 d\left(u^{\varepsilon}(.), \tilde{u}^{\varepsilon}(.)\right)^{1-\frac{p}{2}},$$
$$\leq C_6 d\left(u^{\varepsilon}(.), \tilde{u}^{\varepsilon}(.)\right)^{\frac{\alpha\beta p}{2}}.$$

Therefore

$$\left\{E\int_0^T \left|g\left(t, \tilde{\lambda}^{\epsilon}(t), \tilde{u}^{\varepsilon}(t)\right) - g\left(t, \tilde{\lambda}^{\epsilon}(t), u^{\varepsilon}(t)\right)\right|^p\right.$$

$$\times \chi_{\tilde{u}^\varepsilon(t)\neq u^\varepsilon(t)}(t)\,dt\Big\}^{\frac{1}{p}}$$
$$\leq C_6 d\left(u^\varepsilon(.),\tilde{u}^\varepsilon(.)\right)^{\frac{\alpha\beta}{2}}$$
$$\leq C_6\varepsilon^{\frac{\alpha\beta}{3}} = C_6\varepsilon^{\gamma}.$$

Thus, we have proved that

$$E\int_0^T \tilde{Q}^\epsilon(t)\left[g\left(t,\tilde{\lambda}^\epsilon(t),u\right)-g\left(t,\tilde{\lambda}^\epsilon(t),\tilde{u}^\varepsilon(t)\right)\right]dt$$
$$-E\int_0^T Q^\epsilon(t)\left[g\left(t,\lambda^\epsilon(t),u\right)-g\left(t,\lambda^\epsilon(t),u^\varepsilon(t)\right)\right]dt$$
$$\leq C_7\varepsilon^{\gamma}. \tag{3.24}$$

Similarly, we obtain

$$E\int_0^T \left\{\tilde{\Psi}^\varepsilon(t)\left[f\left(t,\tilde{x}^\varepsilon(t),u\right)-f\left(t,\tilde{x}^\varepsilon(t),\tilde{u}^\varepsilon(t)\right)\right]\right.$$
$$\left.-\ \Psi^\varepsilon(t)\left[f\left(t,x^\varepsilon(t),u\right)-f\left(t,x^\varepsilon(t),u^\varepsilon(t)\right)\right]\right\}dt$$
$$\leq C_8\varepsilon^{\gamma}. \tag{3.25}$$

Finally inequality (3.1) follows from combining (3.24) and (3.25). ■

**Proof of Corollary 3.** In the definition of the perturbed control $u^\theta(.)$, the point $u\in U$ may be replaced by any admissible control $u(.)\in\mathcal{U}_{ad}$, and the subsequent argument still goes through. So (3.1) holds for any $u(.)\in\mathcal{U}_{ad}$. ■

## 4. Sufficient conditions of near-optimality

In this section, we will prove that under additional hypothesis, the near-maximum condition on the Hamiltonian function is a sufficient condition for near-optimality. First, we restrict ourselves to the one dimensional case $n=m=d=1$ and we assume that:

$(H_4)$: $\rho$ is differentiable in $u$ for $\rho = f,\sigma,g$, and there is a constant $C>0$ such that:

$$\begin{aligned}&\left|\rho(t,x,u)-\rho\left(t,x,\acute{u}\right)\right|+\left|\rho_u(t,x,u)-\rho_u\left(t,x,\acute{u}\right)\right|\\&\quad\leq C\left|u-\acute{u}\right|,\quad \text{for } \rho=f,\sigma\\&\left|g(t,x,y,z,u)-g\left(t,x,y,z,\acute{u}\right)\right|\\&\quad+\left|g_u(t,x,y,z,u)-g_u\left(t,x,y,z,\acute{u}\right)\right|\leq C\left|u-\acute{u}\right|.\end{aligned} \tag{4.1}$$

**Theorem 6.** *Let $(\lambda^\varepsilon(.),u^\varepsilon(.))$ be near-optimal solution of $(E)$, and $\Lambda^\varepsilon(.) := (\Psi^\varepsilon(.),K^\varepsilon(.),Q^\epsilon(.))$ are the solutions of the adjoint equations, corresponding to $(\lambda^\varepsilon(.),u^\varepsilon(.))$. Assume that $H(t,.,.,.,.,\Lambda^\varepsilon(t))$ is concave for a.e. $t\in[0,T]$, $P-a.s.$, $h(.)$ is concave, $\varphi(.)$ is convex and decreasing. If for some $\varepsilon>0$,*

$$E\int_0^T H\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)dt$$
$$\geq \sup_{u(.)\in\mathcal{U}_{ad}[0,T]} E\int_0^T H\left(t,\lambda^\epsilon(t),u(t),\Lambda^\varepsilon(t)\right)dt-\varepsilon^{\gamma} \tag{4.2}$$

*then*

$$J(u^\varepsilon(.))\leq \inf_{u(.)\in\mathcal{U}_{ad}[0,T]} J(u(.))+C_1\varepsilon^{\frac{1}{2}}, \tag{4.3}$$

*where $C_1>0$ is a constant, which is independent from $\varepsilon$.*

**Proof.** Let us fix $\varepsilon>0$ and define a new metric $\tilde{d}$ on $\mathcal{U}_{ad}$ as follows

$$\tilde{d}\left(u(.),u^\varepsilon(.)\right)=E\int_0^T \nu^\varepsilon(t)\left|u(t)-u^\varepsilon(t)\right|dt, \tag{4.4}$$

where

$$\nu^\varepsilon(t)=1+\left|\Psi^\varepsilon(t)\right|+\left|Q^\epsilon(t)\right|\geq 1. \tag{4.5}$$

Define a functional $\mathcal{B}$ on $\mathcal{U}_{ad}[0,T]$ by:

$$\mathcal{B}(u(.))=E\int_0^T H\left(t,\lambda^\varepsilon(t),u(t),\Lambda^\varepsilon(t)\right)dt.$$

A simple computation shows that

$$\left|\mathcal{B}(u(.))-\mathcal{B}\left(\acute{u}(.)\right)\right|\leq CE\int_0^T \nu^\varepsilon(t)\left|u(t)-u^\varepsilon(t)\right|dt,$$

which implies that $\mathcal{B}$ is continuous on $\mathcal{U}_{ad}$ with respect to $\tilde{d}$.

By (4.2) and Ekeland's lemma, there exists $\tilde{u}^\varepsilon(.)\in\mathcal{U}_{ad}[0,T]$ such that

$$\tilde{d}\left(u^\varepsilon(.),\tilde{u}^\varepsilon(.)\right)\leq\varepsilon$$

and

$$E\int_0^T \tilde{H}\left(t,\lambda^\epsilon(t),\tilde{u}^\varepsilon(t)\right)dt$$
$$=\max_{u(.)\in\mathcal{U}_{ad}[0,T]} E\int_0^T \tilde{H}\left(t,\lambda^\epsilon(t),u(t)\right)dt \tag{4.6}$$

where

$$\tilde{H}\left(t,\lambda^\epsilon(t),u(t)\right)$$
$$=H\left(t,\lambda^\epsilon(t),u(t),\Lambda^\varepsilon(t)\right)-\varepsilon^{\frac{1}{2}}\nu^\varepsilon(t)\left|u(t)-\tilde{u}^\varepsilon(t)\right|. \tag{4.7}$$

By standard arguments, it can be proved that the integral maximum condition (4.7) implies a pointwise maximum condition, namely for a.e. $t\in[0,T]$, and $P-a.s.$,

$$\tilde{H}\left(t,\lambda^\epsilon(t),\tilde{u}^\epsilon(t)\right)=\max_{u\in U}\tilde{H}\left(t,\lambda^\epsilon(t),u\right). \tag{4.8}$$

Using assumption $(H_4)$ we can prove that

$$H_u\left(t,\lambda^\epsilon(t),\tilde{u}^\varepsilon(t),\Lambda^\varepsilon(t)\right)$$
$$\leq C_3\nu^\varepsilon(t)\left|u^\varepsilon(t)-\tilde{u}^\varepsilon(t)\right|+\varepsilon^{\frac{1}{2}}\nu^\varepsilon(t). \tag{4.9}$$

By the concavity of $H\left(t,.,.,.,.,\Psi^\varepsilon(t),Q^\epsilon(t),K^\varepsilon(t)\right)$, we have

$$\begin{aligned}&H\left(t,\lambda(t),u(t),\Lambda^\varepsilon(t)\right)-H\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\\&\quad\leq H_x\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(x(t)-x^\varepsilon(t)\right)\\&\quad\leq H_y\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(y(t)-y^\varepsilon(t)\right)\\&\qquad+H_z\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(z(t)-z^\varepsilon(t)\right)\\&\qquad+H_u\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(u(t)-u^\varepsilon(t)\right).\end{aligned} \tag{4.10}$$

By integrating both sides and noting (4.1) and (4.9) we obtain

$$E\int_0^T \left\{H\left(t,\lambda(t),u(t),\Lambda^\varepsilon(t)\right)\_H\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\right\}dt$$
$$\leq E\int_0^T H_x\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(x(t)-x^\varepsilon(t)\right)dt$$
$$+E\int_0^T H_y\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(y(t)-y^\varepsilon(t)\right)dt$$
$$\times E\int_0^T H_z\left(t,\lambda^\epsilon(t),u^\varepsilon(t),\Lambda^\varepsilon(t)\right)\left(z(t)-z^\varepsilon(t)\right)dt$$
$$+C_5\varepsilon^{\frac{1}{2}}. \tag{4.11}$$

On the other hand, by applying Ito's formula respectively to $\Psi^\varepsilon(t)\left(x(t)-x^\varepsilon(t)\right)$ and $Q^\epsilon(t)\left(y(t)-y^\varepsilon(t)\right)$, and by taking expectations, we obtain

$$E\left[\Psi^\varepsilon(T)\left(x(T)-x^\varepsilon(T)\right)\right]+E\left[Q^\epsilon(T)\left(y(T)-y^\varepsilon(T)\right)\right]$$

$$
\begin{aligned}
&-E\left[Q^{\epsilon}(0)\left(y(0)-y^{\varepsilon}(0)\right)\right]\\
&=E\int_0^T H_x\left(t,\lambda^{\epsilon}(t),u^{\varepsilon}(t),\Lambda^{\varepsilon}(t)\right)\left(x(t)-x^{\varepsilon}(t)\right)\mathrm{d}t\\
&\quad+E\int_0^T H_y\left(t,\lambda^{\epsilon}(t),u^{\varepsilon}(t),\Lambda^{\varepsilon}(t)\right)\left(y(t)-y^{\varepsilon}(t)\right)\mathrm{d}t\\
&\quad+E\int_0^T H_z\left(t,\lambda^{\epsilon}(t),u^{\varepsilon}(t),\Lambda^{\varepsilon}(t)\right)\left(z(t)-z^{\varepsilon}(t)\right)\mathrm{d}t\\
&\quad+E\int_0^T \Psi^{\varepsilon}(t)\left[f\left(t,x(t),u(t)\right)-f\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t)\right)\right]\mathrm{d}t\\
&\quad+E\int_0^T Q^{\epsilon}(t)\left[g\left(t,\lambda(t),u(t)\right)-g\left(t,\lambda^{\epsilon}(t),u^{\varepsilon}(t)\right)\right]\mathrm{d}t\\
&+E\int_0^T K^{\varepsilon}(t)\left[\sigma\left(t,x(t)\right)-\sigma\left(t,x^{\varepsilon}(t)\right)\right]\mathrm{d}t. \qquad (4.12)
\end{aligned}
$$

Combining (4.11) and (4.12) we obtain

$$
\begin{aligned}
&-E\left[Q^{\epsilon}(T)\,h_x\left(x^{\varepsilon}(T)\right)\left(x(T)-x^{\varepsilon}(T)\right)\right]+EQ^{\epsilon}(T)\left(h\left(x(T)\right)\right.\\
&\quad\left.-h\left(x^{\varepsilon}(T)\right)\right)+E\left[\varphi_y\left(y^{\varepsilon}(0)\right)\left(y(0)-y^{\varepsilon}(0)\right)\right]\\
&\geq -C\varepsilon^{\frac{1}{2}}.
\end{aligned}
$$

Using the facts that $h(.)$ is concave and $\varphi(.)$ is convex and decreasing, we obtain

$$
J\left(u^{\varepsilon}(.)\right)\leq J\left(u(.)\right)+C_1\varepsilon^{\frac{1}{2}}.
$$

Since $u(.)$ is arbitrary, the desired result follows. ■

**Remark 7.** When $\varepsilon = 0$, Theorem 2 reduces to the stochastic maximum principal developed in [6].


## Acknowledgement

Partially supported by PHC Tassili 07MDU705 and Marie Curie ITN N°. 213841-2.


## References


[1] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett. 14 (1) (1990) 55–61.
[2] E. Pardoux, BSDEs, Weak convergence and homogenization of semilinear PDEs, in: F.H. Clarke, R.J. Stern (Eds.), Nonlinear Analysis, Differential Equations and Control, Kluwer Academic Publishers, 1999, pp. 503–549.
[3] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7 (1997) 1–71.
[4] N. Dokuchaev, X.Y. Zhou, Stochastic controls with terminal contingent conditions, J. Math. Anal. Appl. 238 (1999) 143–165.
[5] S. Peng, Backward stochastic differential equation and application to optimal control, Appl. Math. Optim. 27 (1993) 125–144.
[6] W. Xu, Stochastic maximum principle for optimal control problem of forward and backward, J. Aust. Math. Soc. Ser. B 37 (1995) 172–1785.
[7] S. Peng, Z. Wu, Fully coupled forward backward stochastic differential equations and application to optimal control, SIAM J. Control Optim. 37 (3) (1999) 825–843.
[8] I. Ekeland, Non convex minimisation problems, Bull. Amer. Math. Soc., NS 1 (1979) 324–353.
[9] J. Yong, X.Y. Zhou, Stochastic Controls, Hamiltonian Systems and HJB Equations, Springer Verlag, 1999.
[10] R.J Elliott, M. Kohlmann, The variational principle and stochastic optimal control, Stochastics 3 (1980) 229–241.
[11] B. Mezerdi, Necessary conditions for optimality for a diffusion with non smooth drift, Stochastics 24 (1988) 305–326.
[12] K. Bahlali, B. Mezerdi, Y. Ouknine, The maximum principle for optimal control of diffusions with non-smooth coefficients, Stoch. Stoch. Rep. 57 (1996) 303–316.
[13] X.Y. Zhou, Stochastic near-optimal controls: Necessary and sufficient conditions for near-optimality, SIAM.J. Control Optim. 36 (3) (1998) 929–947.