On necessary and sufficient conditions for near-optimal singular stochastic controls

Mokhtar Hafayed · Syed Abbas · Petr Veverka

Received: 24 October 2011 / Accepted: 11 April 2012 / Published online: 26 April 2012
© Springer-Verlag 2012

Abstract In this paper we discuss the necessary and sufficient conditions for near-optimal singular stochastic controls for the systems driven by a nonlinear stochastic differential equations (SDEs in short). The proof of our result is based on Ekeland’s variational principle and some delicate estimates of the state and adjoint processes. It is well known that optimal singular controls may fail to exist even in simple cases. This justifies the use of near-optimal singular controls, which exist under minimal conditions and are sufficient in most practical cases. Moreover, since there are many near-optimal singular controls, it is possible to choose suitable ones, that are convenient for implementation. This result is a generalization of Zhou’s stochastic maximum principle for near-optimality to singular control problem.

Keywords Near-optimal singular stochastic control · Maximum principle · Necessary and sufficient conditions · Ekeland’s variational principle

M. Hafayed
Laboratory of Applied Mathematics, Mohamed Khider University, Po Box 145, Biskra 07000, Algeria
e-mail: hafa.mokh@yahoo.com

S. Abbas (✉)
School of Basic Sciences, Indian Institute of Technology Mandi, Mandi, HP 175001, India
e-mail: sabbas.iitk@gmail.com

P. Veverka
Faculty of Nuclear Sciences and Physical Engineering, Department of Mathematics, Czech Technical University, Trojanova 13, Prague 12000, Czech Republic
e-mail: panververka@seznam.cz
1 Introduction

In this paper, we consider the singular stochastic control problem for the systems governed by nonlinear controlled diffusion of the following type

\[
\begin{align*}
 dx_t &= f(t, x_t, u_t) \, dt + \sigma(t, x_t, u_t) \, dW_t + G_t \, d\eta_t, \\
 x_s &= y,
\end{align*}
\]  

(1.1)

where \((W_t)\) is a standard \(l\)-dimensional Brownian motion defined on the filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})\). The minimized criteria associated with the state Eq. (1.1) is defined by

\[
J(s, y, u, \eta) = \mathbb{E} \left[ h(x_T) + \int_s^T \ell(t, x_t, u_t) \, dt + \int_s^T k_t \, d\eta_t \right],
\]  

(1.2)

and the value function is defined as

\[
V(s, y) = \inf_{(u, \eta) \in U([s,T])} \{ J(s, y, u, \eta) \}.
\]  

(1.3)

This kind of stochastic control problem has been investigated extensively, both by the Bellman’s dynamic programming method [6] and by Pontryagin’s maximum principle [18]. In this paper, we are concerned to the second method. Peng [17] introduced the second-order adjoint equation and obtained the global maximum principle of optimality, in which the control is present both in the drift and in the diffusion coefficients. Studying near-optimal controls makes a good sense as studying optimal controls from both theoretical as well as applications point of view. Many more near-optimal controls are available than optimal ones. Indeed, optimal controls may not even exist in many situations, while near-optimal controls always exist. The near-optimal deterministic controls problem has been treated by many authors, including [12,15,20,21]. Recently, in an interesting paper, Zhou [22] established the second-order necessary as well as sufficient conditions for near-optimal stochastic controls for general controlled diffusion with two adjoint processes. The near-optimal control problem for systems described by Volterra integral equations has been studied in [16]. However, Chighoub et al. [8] extended Zhou’s maximum principle of near-optimality to SDEs with jumps. The similar problem for systems driven by forward-backward stochastic differential equations has been solved in [5]. For justification of establishing a theory of near-optimal controls one can see [20,22, Introduction].

Singular stochastic control problem is an important and challenging class of problems in control theory, it appears in various fields like mathematical finance, problem of optimal consumption etc. Stochastic maximum principle for singular controls was considered by many authors, see for instance [1–3,7,10,11,13,14]. The first version of maximum principle for singular stochastic control problems was obtained by Cadenillas et al. [7]. The first-order weak stochastic maximum principle has been studied in [3]. In [10], the authors derived stochastic maximum principle where the
sufficient conditions for existence of optimal singular control have been obtained in [11].

The main objective of this paper is to establish necessary as well as sufficient conditions for near-optimal singular control for SDEs. The control domain is not necessarily convex. These conditions are given in terms of second-order adjoint processes corresponding to the controlled SDEs and a nearly maximum conditions on the Hamiltonian. Moreover in a second step, we prove that under additional concavity conditions on the Hamiltonian function, these necessary conditions of near-optimality are also sufficient.

The structure of the paper is as follows. The assumptions and statement of the control problem is given in the Sect. 2. In Sects. 3 and 4, we establish the main result of this paper.

2 Assumptions and statement of the problem

We consider stochastic optimal control of the following kind. Let $T$ be a fixed strictly positive real number and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a fixed filtered probability space satisfying the usual conditions in which an $l-$dimensional Brownian motion $W = \{W_t : s \leq t \leq T\}$ with $s \in [0, T]$ and $W_s = 0$ is defined. Let $\mathbb{A}_1$ be a closed convex subset of $\mathbb{R}^m$ and $\mathbb{A}_2 := (0, \infty)^m$. Let $\mathbb{U}_1$ be the class of measurable, $\mathcal{F}_t-$adapted processes $u : [s, T] \times \Omega \to \mathbb{A}_1$ and $\mathbb{U}_2$ is the class of measurable, $\mathcal{F}_t-$adapted processes $\eta : [s, T] \times \Omega \to \mathbb{A}_2$. We make use the following notation in this paper: any element $x \in \mathbb{R}^n$ will be identified to a column vector with $i-th$ component, and the norm $|x| = |x_1| + \cdots + |x_n|$. The scalar product of any two vectors $x$ and $y$ on $\mathbb{R}^n$ is denoted by $\langle x, y \rangle$. A real number $t$ is denoted by $|t|$. The transpose of any vector or matrix $A$ is denoted by $A^*$ the transpose of any vector or matrix $A$. We denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$. For a function $f$, we denote by $f_x$ (resp. $f_{xx}$) the gradient or Jacobian (resp. the Hessian) of a scalar function $f$ with respect to the variable $x$. For a set $B$, we denote by $\chi_B$ the indicator function of $B$. We denote by $L^2_{\mathcal{F}}([s, T], \mathbb{R}^n)$ the Hilbert space of $\mathcal{F}_t-$adapted processes $x_t$ such that $\mathbb{E} \int_s^T |x_t|^2 \, dt < +\infty$.

Since the objective of this paper is to study near-optimal singular stochastic controls, we give here the precise definition of the singular part of an admissible control.

**Definition 1** An admissible control is a pair $(u, \eta)$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathcal{F}_t-$ adapted processes, such that

1. $\eta$ is of bounded variation, nondecreasing continuous on the left with right limits and $\eta_s = 0$.
2. $\mathbb{E} \left[ \sup_{t \in [s, T]} |u_t|^2 + |\eta_T|^2 \right] < \infty$.

We denote $\mathbb{U} = \mathbb{U}_1 \times \mathbb{U}_2$, the set of all admissible controls. Since $d\eta_t$ may be singular with respect to Lebesgue measure $dt$, we call $\eta$ the singular part of the control and the process $u$ its absolutely continuous part.

Throughout this paper, we assume the following:

(H1) $f : [s, T] \times \mathbb{R}^n \times \mathbb{A}_1 \to \mathbb{R}^n$, $\sigma : [s, T] \times \mathbb{R}^n \times \mathbb{A}_1 \to \mathcal{M}_{n \times 1}(\mathbb{R})$ and $\ell : [s, T] \times \mathbb{R}^n \times \mathbb{A}_1 \to \mathbb{R}$ are measurable in $(t, x, u)$ and twice continuously differentiable.
in $x$, and there exists a constant $C > 0$ such that, for $\varphi = f, \sigma, \ell$:

$$|\varphi(t, x, u) - \varphi(t, x', u)| + |\varphi_x(t, x, u) - \varphi_x(t, x', u)| \leq C|x - x'|. \quad (2.1)$$

$$|\varphi(t, x, u)| \leq C(1 + |x|). \quad (2.2)$$

(H2) $h : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable in $x$, and there exists a constant $C > 0$ such that

$$|h(x) - h(x')| + |h_x(x) - h_x(x')| \leq C|x - x'|. \quad (2.3)$$

$$|h(x)| \leq C(1 + |x|). \quad (2.4)$$

(H3) $G : [s, T] \to \mathcal{M}_{n \times m}(\mathbb{R}), k : [s, T] \to ([0, \infty))^m$, for each $t \in [s, T] : G$ is continuous and bounded, also $k$ is continuous.

Under the above assumptions, the SDE (1.1) has a unique strong solution $x_t$ which is given by

$$x_t = y + \int_s^t f(r, x_r, u_r) \, dr + \int_s^t \sigma(r, x_r, u_r) \, dW_r + \int_s^t G_r \, d\eta_r,$$

and by standard arguments it is easy to show that for any $q > 0$, it holds that

$$\mathbb{E}(\sup_{t \in [s, T]} |x_t|^q) < C(q),$$

where $C(q)$ is a constant depending only on $q$ and the functional $J$ is well defined.

For any $(u, \eta) \in U$ and the corresponding state trajectory $x$, we define the first-order adjoint process $\Psi_t$ and the second-order adjoint process $Q_t$ as the ones satisfying the following two backward SDEs respectively

\[
\begin{align*}
\frac{d\Psi_t}{dt} &= -\left[ f_x^* (t, x_t, u_t) \Psi_t + \sigma_x^* (t, x_t, u_t) K_t + \ell_x (t, x_t, u_t) \right] dt \\
\Psi_T &= h_x (x_T),
\end{align*}
\]

(2.5)

and

\[
\begin{align*}
\frac{dQ_t}{dt} &= -\left[ f_x^* (t, x_t, u_t) Q_t + Q_t f_x^* (t, x_t, u_t) + \sigma_x^* (t, x_t, u_t) Q_t \sigma_x^* (t, x_t, u_t) \\
&\quad + \sigma_x^* (t, x_t, u_t) R_t + R_t \sigma_x (t, x_t, u_t) + \Gamma_t \right] dt + R_t dW_t, \\
Q_T &= h_{xx} (x_T),
\end{align*}
\]

(2.6)

where

$$\Gamma_t = \ell_{xx} (t, x_t, u_t) + \sum_{i=1}^n \left( \Psi_t^i f_{xx}^i (t, x_t, u_t) + K_t^i \sigma_{xx}^i (t, x_t, u_t) \right).$$
As it is well known that under conditions (H1), (H2) and (H3) the first-order adjoint
equation (2.5) admits one and only one \( F \) and the second-order adjoint equation (2.6) admits one and only one \( \mathcal{F}_t \)-adapted solution pair \((\Psi_t, K) \in L^2_{\mathcal{F}}([s, T], \mathbb{R}^n) \times L^2_{\mathcal{F}}([s, T], \mathbb{R}^n)\). Also the second-order adjoint equation (2.6) admits one and only one \( \mathcal{F}_t \)-adapted solution pair \((Q, R) \in L^2_{\mathcal{F}}([s, T], \mathbb{R}^{n \times n}) \times L^2_{\mathcal{F}}([s, T], \mathbb{R}^{n \times n})\).

Moreover, since \( f, \sigma, \ell, \) and \( h \) are bounded, we have the following estimate

\[
\mathbb{E} \left[ \sup_{s \leq t \leq T} |\Psi_t|^2 + \int_s^T |K_t|^2 \, dt + \sup_{s \leq t \leq T} |Q_t|^2 + \int_s^T |R_t|^2 \right] \leq C.
\]

Define the usual Hamiltonian

\[
H(t, x, u, p, q) := -p f(t, x, u) - q \sigma(t, x, u) - \ell(t, x, u),
\]

for \((t, x, u) \in [s, T] \times \mathbb{R}^n \times \mathbb{R}^n\). Furthermore, we define the \( \mathcal{H} \) functional corresponding to a given admissible pair \((x, u)\) as follows

\[
\mathcal{H}^{(x, u)}(t, x, u) = H(t, x, u, \Psi_t, K_t - Q_t \sigma(t, x, u)) - \frac{1}{2} \sigma^*(t, x, u) Q_t \sigma(t, x, u),
\]

for \((t, x, u, p, q) \in [s, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n\), where \( \Psi_t, K_t \) and \( Q_t \) are determined by adjoint equations (2.5) and (2.6) corresponding to \((x, u)\).

Before concluding this section, let us recall Ekeland’s variational principle and the Clarke’s generalized gradient, which will be used in the sequel.

**Lemma 1** (Ekeland’s Lemma [12]) Let \((F, \rho)\) be a complete metric space and \(f : F \to \mathbb{R}\) be a lower semi-continuous function which is bounded from below. For a given \( \varepsilon > 0 \), suppose that \( u^\varepsilon \in F \) satisfying \( f(u^\varepsilon) \leq \inf(f) + \varepsilon \). Then for any \( \lambda > 0 \), there exists \( u^\lambda \in F \) such that

1. \( f(u^\lambda) \leq f(u^\varepsilon) \).
2. \( \rho(u^\lambda, u^\varepsilon) \leq \lambda \).
3. \( f(u^\lambda) \leq f(u) + \frac{\varepsilon}{\lambda} \rho(u, u^\lambda) \) for all \( u \in F \).

To apply Ekeland’s variational principle to our problem, we define a distance function \( d \) on the space of admissible controls \( U \) such that \((U, d)\) becomes a complete metric space. To achieve this goal, we define for any \((u, \eta)\) and \((v, \xi)\) in \( U \):

\[
d((u, \eta), (v, \xi)) = d_1(u, v) + d_2(\eta, \xi),
\]

where

\[
d_1(u, v) = \mathbb{P} \otimes dt \{ (w, t) \in \Omega \times [0, T] : u(w, t) \neq v(w, t) \},
\]

\[
d_2(\eta, \xi) = \mathbb{P} \otimes dt \{ (w, t) \in \Omega \times [0, T] : \eta(w, t) \neq \xi(w, t) \}.
\]
and
\[ d_2(\eta, \xi) = \left[ \mathbb{E}(\sup_{t \in [s,T]} |\eta_t - \xi_t|^2) \right]^\frac{1}{2}, \quad (2.10) \]

here \( \mathbb{P} \otimes dt \) is the product measure of \( \mathbb{P} \) with the Lebesgue measure \( dt \) on \( [s,T] \). It is easy to see that \( (\mathbb{U}_2, d_2) \) is a complete metric space. Moreover, it has been shown in Yong et al. [19, pp. 146–147] that \( (\mathbb{U}_1, d_1) \) is a complete metric space. Hence \( (\mathbb{U}, d) \) as a product of two complete metric spaces is a complete metric space under \( d \).

**Definition 2** (Clarke’s generalized gradient [9]) Let \( E \) be a convex set in \( \mathbb{R}^n \) and let \( f: E \to \mathbb{R} \) be a locally Lipschitz function. The Clarke’s generalized gradient of \( f \) at \( \hat{x} \in E \), denoted by \( \partial_x f(\hat{x}) \), is a set defined by
\[ \partial_x f(\hat{x}) = \{ \xi \in \mathbb{R}^n : \langle \xi, v \rangle \leq f(\hat{x} + t v) - f(\hat{x}), \text{for any } v \in \mathbb{R}^n \}, \]
where \( f(\hat{x}, v) = \lim_{y \to \hat{x}, t \to 0} \frac{f(y + tv) - f(y)}{t} \).

### 3 Necessary conditions for near-optimal singular control

Our purpose in this paper is to establish second-order necessary and sufficient conditions for near-optimal singular control for systems governed by nonlinear SDEs. It is worth mentioning that optimal singular controls may not even exist in many situations, while near-optimal singular controls always exist. The proof follows the general ideas as in [20–22] where similar results are obtained for other class of controls.

We give the definition of near-optimal control as given in Zhou [22, Definition 2.1 and Definition 2.2].

**Definition 3** For a given \( \varepsilon > 0 \) the admissible control \((u^\varepsilon, \eta^\varepsilon)\) is near-optimal if
\[ \left| J(s, y, u^\varepsilon, \eta^\varepsilon) - V(s, y) \right| \leq O(\varepsilon), \quad (3.1) \]
where \( O(.) \) is a function of \( \varepsilon \) satisfying \( \lim_{\varepsilon \to 0} O(\varepsilon) = 0 \). The estimator \( O(\varepsilon) \) is called an error bound. If \( O(\varepsilon) = C \varepsilon^\delta \) for some \( \delta > 0 \) independent of the constant \( C \) then \((u^\varepsilon, \eta^\varepsilon)\) is called near-optimal control with order \( \varepsilon^\delta \). If \( O(\varepsilon) = \varepsilon \) the admissible control \((u^\varepsilon, \eta^\varepsilon)\) called \( \varepsilon \)-optimal.

Our first Lemma below deals with the continuity of the state processes under distance \( d \).

**Lemma 2** If \( x^{u, \eta}_t \) and \( x^{v, \xi}_t \) be the solution of the state Eq. (1.1) associated respectively with \( u, v \). For any \( 0 < \alpha < 1 \) and \( \beta > 0 \) satisfying \( \alpha \beta < 1 \), there exists a positive constants \( C = C(T, \alpha, \beta) \) such that
\[ \mathbb{E}(\sup_{s \leq t \leq T} |x^{u, \eta}_t - x^{v, \xi}_t|^2)^{\frac{2\beta}{d_1^{\alpha\beta}}} \leq Cd_1^{\alpha\beta}(u, v). \quad (3.2) \]
Proof First, we assume that \( \beta \geq 1 \). Using Burkholder–Davis–Gundy inequality for the martingale part, for any \( r \geq s \), we get

\[
\mathbb{E} \left[ \sup_{s \leq t \leq r} \left| x^u_t, \eta_t - x^v_t, \xi_t \right|^2 \beta \right] \leq C \mathbb{E} \left( \int_s^r \left\{ \left| f \left( t, x^u_t, \eta_t, u_t \right) - f \left( t, x^v_t, \xi_t, v_t \right) \right|^2 \beta \right\} \, dt \right.
\]

\[
+ \left\{ \left| \sigma \left( t, x^u_t, \eta_t, u_t \right) - \sigma \left( t, x^v_t, \xi_t, v_t \right) \right|^2 \beta \right\} \chi_{u_t \neq v_t} (t) \, dt
\]

\[
+ C \mathbb{E} \left( \eta_T - \xi_T \right)^{2\beta},
\]

\[
\leq C \mathbb{E} \left( \int_s^r \left\{ \left| f \left( t, x^u_t, \eta_t, v_t \right) - f \left( t, x^u_t, \eta_t, v_t \right) \right|^2 \beta \right\} \right.
\]

\[
+ \left\{ \left| \sigma \left( t, x^u_t, \eta_t, v_t \right) - \sigma \left( t, x^u_t, \eta_t, v_t \right) \right|^2 \beta \right\} \chi_{u_t \neq v_t} (t) \, dt
\]

\[
+ C \mathbb{E} \left( \eta_T - \xi_T \right)^{2\beta}.
\]

Now arguing as in [22, Lemma 3.1] taking \( b = \frac{1}{\alpha \beta} > 1 \) and \( a > 1 \) such that \( \frac{1}{a} + \frac{1}{b} = 1 \), and applying Cauchy–Schwarz inequality, we get

\[
\mathbb{E} \int_s^r \left| f \left( t, x^u_t, \eta_t, u_t \right) - f \left( t, x^u_t, \eta_t, v_t \right) \right|^2 \beta \chi_{u_t \neq v_t} (t) \, dt
\]

\[
\leq \left\{ \mathbb{E} \left( \int_s^r \left| f \left( t, x^u_t, \eta_t, u_t \right) - f \left( t, x^u_t, \eta_t, v_t \right) \right|^2 \beta a \right) \, dt \right\}^{\frac{1}{a}} \times \left\{ \mathbb{E} \left( \int_s^r \chi_{u_t \neq v_t} (t) \, dt \right) \right\}^{\frac{1}{b}}
\]

using definition of \( d_1 \) and linear growth condition on \( f \) we obtain

\[
\mathbb{E} \int_s^r \left| f \left( t, x^u_t, \eta_t, u_t \right) - f \left( t, x^u_t, \eta_t, v_t \right) \right|^2 \beta \chi_{u_t \neq v_t} (t) \, dt
\]

\[
\leq C \left\{ \mathbb{E} \left( \int_s^r \left( 1 + \left| x^u_t, \eta_t \right|^{2\beta a} \right) \, dt \right) \right\}^{\frac{1}{a}} \times \left\{ d_1 (u, v)^{\alpha \beta} \right\}
\]

\[
\leq C d_1 (u, v)^{\alpha \beta}.
\]
Similarly, we can prove
\[
\mathbb{E} \int_s^r \left| \sigma \left( t, x^{u, \eta}_t, u_t \right) - \sigma \left( t, x^{u, \eta}_t, v_t \right) \right|^{2\beta} \chi_{u_t \neq v_t} (t) \, dt \leq C d_1 (u, v)^{\alpha \beta}. \tag{3.3}
\]
Therefore, by using assumption (H1), we conclude that
\[
\mathbb{E} \left( \sup_{s \leq t \leq R} \left| x^{u, \eta}_t - x^{v, \xi}_t \right|^{2\beta} \right) \leq C \left\{ \mathbb{E} \int_s^r \sup_{s \leq t \leq \theta} \left| x^{u, \eta}_t - x^{v, \xi}_t \right|^{2\beta} \, d\theta + \mathbb{E} |\eta_T - \xi_T|^{2\beta} + d_1 (u, v)^{\alpha \beta} \right\}.
\]
Hence (3.1) follows immediately from Definition 1 and Gronwall’s inequality.

Now we assume \( 0 \leq \beta < 1 \). Since \( \frac{2}{\alpha} > 1 \) then the Cauchy–Schwarz inequality yields
\[
\mathbb{E} \left( \sup_{s \leq t \leq T} \left| x^{u, \eta}_t - x^{v, \xi}_t \right|^{2\beta} \right) \leq \left[ \mathbb{E} \left( \sup_{s \leq t \leq T} \left| x^{u, \eta}_t - x^{v, \xi}_t \right|^2 \right)^\beta \right] \leq \left[ C d_1 (u, v)^\alpha \right]^\beta \leq C d_1 (u, v)^{\alpha \beta}.
\]
This completes the proof of Lemma 2.

**Lemma 3** For any \( 0 < \alpha < 1 \) and \( 1 < \beta < 2 \) satisfying \( (1 + \alpha) \beta < 2 \), there exist a positive constant \( C = C (\alpha, \beta) \) such that for any \((u, \eta), (v, \xi) \in \mathbb{U} ([s, T]) \), along with the corresponding trajectories \( x^{u, \eta}, x^{v, \xi} \) and the solutions \((\Psi, K, Q, R), (\Psi', K', Q', R') \) of the corresponding adjoint equations, it holds that
\[
\mathbb{E} \int_0^T \left( \left| \Psi_t - \Psi'_t \right|^\beta + \left| K_t - K'_t \right|^\beta \right) dt \leq C d_1^{\frac{\alpha \beta}{2}} (u, v). \tag{3.1}
\]
\[
\mathbb{E} \int_0^T \left( \left| Q_t - Q'_t \right|^\beta + \left| R_t - R'_t \right|^\beta \right) dt \leq C d_1^{\frac{\alpha \beta}{2}} (u, v). \tag{3.2}
\]

**Proof** Since the adjoint processes are independent to singular part, we use similar argument as in Zhou [22, Lemma 3.2].

Now we are able to state and prove the necessary conditions for near-optimal singular control for our problem, which is the main result in this paper.

Let \((\Psi^\varepsilon, K^\varepsilon)\) and \((Q^\varepsilon, R^\varepsilon)\) be the solution of adjoint equations (2.5) and (2.6) respectively corresponding to \((x^\varepsilon, (u^\varepsilon, \eta^\varepsilon))\).
Theorem 1 (Maximum principle for any near-optimal singular control) For any $\delta \in (0, \frac{1}{3}]$, and any near-optimal singular control $(u^\varepsilon, \eta^\varepsilon)$ there exists a positive constant $C = C(\delta)$ such that for each $\varepsilon > 0$

\[
-C \varepsilon^3 \leq \mathbb{E} \int_T^s \{ \frac{1}{2} (\sigma(t, x^\varepsilon_t, u_t) - \sigma(t, x^\varepsilon_t, u^\varepsilon_t)) \sigma(t, x^\varepsilon_t, u_t) - \sigma(t, x^\varepsilon_t, u^\varepsilon_t) \}
+ \Psi^\varepsilon_t \left( f(t, x^\varepsilon_t, u_t) - f(t, x^\varepsilon_t, u^\varepsilon_t) \right) + K^\varepsilon_t \left( \sigma(t, x^\varepsilon_t, u_t) - \sigma(t, x^\varepsilon_t, u^\varepsilon_t) \right)
\]
\[
+ (\ell(t, x^\varepsilon_t, u_t) - \ell(t, x^\varepsilon_t, u^\varepsilon_t)) \} dt,
\]

(3.6)

and

\[-C \varepsilon^3 \leq \mathbb{E} \left[ \int_0^T (k_t + G^\varepsilon_t \Psi^\varepsilon_t) d(\eta - \eta^\varepsilon) \right].
\]

(3.7)

Proof

By using Ekeland’s variational principle with $\lambda = \varepsilon^3$, there is an admissible pair $(\bar{x}^\varepsilon, (\bar{u}^\varepsilon, \bar{\eta}^\varepsilon))$ such that for any $(u, \eta) \in \mathbb{U}$:

\[d((u^\varepsilon, \eta^\varepsilon), (\bar{u}^\varepsilon, \bar{\eta}^\varepsilon)) \leq \varepsilon^3,
\]

(3.8)

and

\[J^\varepsilon(s, y, u^\varepsilon, \eta^\varepsilon) \leq J^\varepsilon(s, y, u^\varepsilon, \eta^\varepsilon) + \varepsilon^3 d((u, \eta), (\bar{u}^\varepsilon, \bar{\eta}^\varepsilon)).
\]

Notice that $(u^\varepsilon, \eta^\varepsilon)$ which is near-optimal for the initial cost $J$ defined in (1.2) is optimal for the new cost $J^\varepsilon$ given by

\[J^\varepsilon(s, y, u, \eta) = J(s, y, u, \eta) + \varepsilon^3 d((u, \eta), (\bar{u}^\varepsilon, \bar{\eta}^\varepsilon)).
\]

(3.9)

Therefore we have

\[J^\varepsilon(s, y, u, \eta) \leq J^\varepsilon(s, y, u, \eta) \quad \text{for any } (u, \eta) \in \mathbb{U}([s, T]).
\]

Next, we use the spike variation techniques for $\bar{u}^\varepsilon$ to derive the first variational inequality and we use convex perturbation for $\bar{\eta}^\varepsilon$ as follows.

First variational inequality: for any $\theta > 0$, we define the following strong perturbation $(\bar{u}^\varepsilon_t, \bar{\eta}^\varepsilon_t) \in \mathbb{U}$:

\[(\bar{u}^\varepsilon_t, \bar{\eta}^\varepsilon_t) = \begin{cases}
(u^\varepsilon_t, \eta^\varepsilon_t), & t \in [t_0, t_0 + \theta], \\
(\bar{u}^\varepsilon_t, \bar{\eta}^\varepsilon_t), & \text{otherwise}.
\end{cases}
\]

(3.10)

The fact that

\[J^\varepsilon(s, y, \bar{u}^\varepsilon, \bar{\eta}^\varepsilon) \leq J^\varepsilon(s, y, \bar{u}^\varepsilon, \bar{\eta}^\varepsilon),
\]

(3.11)
and
\[
d((\bar{x}^e, \bar{u}^e), (\bar{x}^e, \bar{u}^e)) = d((\bar{x}^e, \bar{u}^e), (\bar{x}^e, \bar{u}^e)) \leq \theta,
\]

imply that
\[
J(s, y, \bar{x}^e, \bar{u}^e) - J(s, y, \bar{x}^e, \bar{u}^e) \geq -\theta \varepsilon^{\frac{1}{3}}.
\]

Since the difference \(J(s, y, \bar{x}^e, \bar{u}^e) - J(s, y, \bar{x}^e, \bar{u}^e)\) is independent to the singular part, the near-maximum condition (3.6) follows by applying similar argument as given in Zhou [22], we get
\[
- C\varepsilon^{\frac{1}{3}} \leq \mathbb{E} \int_s^T \left\{ \frac{1}{2} \left( \sigma \left( t, \bar{x}^e_t, u_t \right) - \sigma \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right)^\ast \hat{Q}_t \left( \sigma \left( t, \bar{x}^e_t, u_t \right) - \sigma \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right) \right. \\

\left. + \hat{Y}_t \left( f \left( t, \bar{x}^e_t, u_t \right) - f \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right) + \hat{K}_t \left( \sigma \left( t, \bar{x}^e_t, u_t \right) - \sigma \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right) \right\} dt.
\]

Now we are about to derive an estimate for the term similar to the right hand side of the above inequality with all the \((\bar{x}^e_t, (\bar{u}^e_t, \bar{\eta}^e_t))\) etc. replaced by \((x^e_t, (u^e_t, \eta^e_t))\) etc. To this end, we use similar method as in Zhou [22] to obtain the following estimates:
\[
\mathbb{E} \int_s^T \left[ \hat{K}_t \left( \sigma \left( t, x^e_t, u_t \right) - \sigma \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right) \right. \\
\left. - K_t^e \left( \sigma \left( t, x^e_t, u_t \right) - \sigma \left( t, x^e_t, u_t \right) \right) \right] dt \leq C\varepsilon^{\delta},
\]

and
\[
\mathbb{E} \int_s^T \left\{ \frac{1}{2} \left( \sigma \left( t, x^e_t, u_t \right) - \sigma \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right)^\ast \hat{Q}_t \left( \sigma \left( t, \bar{x}^e_t, u_t \right) - \sigma \left( t, \bar{x}^e_t, \bar{u}^e_t \right) \right) \right. \\
\left. - \frac{1}{2} \left( \sigma \left( t, x^e_t, u_t \right) - \sigma \left( t, x^e_t, u^e_t \right) \right)^\ast \hat{Q}_t \left( \sigma \left( t, x^e_t, u_t \right) - \sigma \left( t, x^e_t, u^e_t \right) \right) \right\} dt \\
\left. + \left[ \hat{Y}_t \left( f \left( t, x^e_t, u_t \right) - f \left( t, x^e_t, u_t \right) \right) - \hat{Y}_t \left( f \left( t, x^e_t, u_t \right) - f \left( t, x^e_t, u^e_t \right) \right) \right] \right\} dt \\
\leq C\varepsilon^{\delta},
\]

where \((\hat{Y}^e, \hat{K}^e)\) and \((\hat{Q}^e, \hat{R}^e)\) are the solutions of adjoint equations (2.5) and (2.6) respectively corresponding to \((\bar{x}^e, (\bar{u}^e, \bar{\eta}^e))\). The first variational inequality (3.6) follows from combining (3.12), (3.13) and (3.14).
Corollary 1 Under the assumptions of Theorem 1, we have

\[
\mathbb{E} \int_{s}^{T} \mathcal{H}^{(x^\varepsilon, u^\varepsilon)}(t, x_t^\varepsilon, u_t^\varepsilon)dt \geq \sup_{u \in \mathbb{U}_1([s, T])} \mathbb{E} \int_{s}^{T} \mathcal{H}^{(x^\varepsilon, u^\varepsilon)}(t, x_t^\varepsilon, u_t^\varepsilon)dt - C \varepsilon^\delta. \tag{3.15}
\]

Second variational inequality: to obtain the second variational inequality, we define the following convex perturbation \((\overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon, \theta) \in \mathbb{U}_1 \times \mathbb{U}_2:

\[
(\overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon, \theta) = (\overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon + \theta (\xi - \overline{\eta}_t^\varepsilon)), \tag{3.16}
\]

where \(\xi\) is an arbitrary element of the set \(\mathbb{U}_2\). Using the optimality of \((u_t^\varepsilon, \eta_t^\varepsilon)\) to the new cost \(J^\varepsilon\), we have

\[
J^\varepsilon(s, y, \overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon) \leq J^\varepsilon(s, y, u_t^\varepsilon, \eta_t^\varepsilon, \theta). \tag{3.17}
\]

A simple computation on \(d_2(\overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon, \theta)\) gives

\[
J(s, y, \overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon, \theta) - J(s, y, \overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon) \geq -C \varepsilon^\frac{1}{3} \geq -C \varepsilon^\delta.
\]

Finally, arguing as in [4] for the left-hand side of the above inequality, we have

\[
\lim_{\theta \to 0} \frac{1}{\theta} \left[ J(s, y, \overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon, \theta) - J(s, y, \overline{u}_t^\varepsilon, \overline{\eta}_t^\varepsilon) \right] = \mathbb{E} \int_{s}^{T} (k_t + G_t^* \overline{\Psi}_t) d(\eta - \overline{\eta}_t),
\]

the near-singular maximum condition is given as

\[
-C \varepsilon^\delta \leq \mathbb{E} \left[ \int_{s}^{T} (k_t + G_t^* \overline{\Psi}_t) d(\eta - \overline{\eta}_t) \right]. \tag{3.18}
\]

Now, we proceed to derive an estimates for the term similar to the right hand side of (3.18) with all \((x^\varepsilon, (u^\varepsilon, \eta^\varepsilon))\) etc. replaced by \((x^\varepsilon, (u^\varepsilon, \eta^\varepsilon))\) etc. We first estimate the following difference:

\[
\mathbb{E} \int_{s}^{T} (k_t + G_t^* \overline{\Psi}_t) d(\eta - \overline{\eta}_t) - \mathbb{E} \int_{s}^{T} (k_t + G_t^* \overline{\Psi}_t) d(\eta - \overline{\eta}_t) = \mathbb{E} \int_{s}^{T} G_t^* (\overline{\Psi}_t - \overline{\Psi}_t) d\eta_t + \mathbb{E} \int_{s}^{T} (k_t + G_t^* \overline{\Psi}_t) d\eta_t.
\]
\[-\mathbb{E} \int_{s}^{T} (k_t + G_t^* \Psi_t^e) d\eta_t^e.\]

\[= \mathbb{E} \int_{s}^{T} G_t^* \left( \frac{\Psi_t^e}{\Psi_t} - \Psi_t^e \right) d(\eta - \eta_t^e) + \mathbb{E} \int_{s}^{T} (k_t + G_t^* \Psi_t^e) d(\eta_t^e - \eta_t^e).\]

Using the bound of $G_t, k_t$, Lemma 3, Definition 1 ($\eta_s = \eta_s^e = \eta_s^e = 0$, $\mathbb{E} |\eta_T - \eta_T^e|^2 + \mathbb{E} |\eta_T^e - \eta_T^e|^2 < \infty$) and the fact that $\mathbb{E} \left( \sup_{t \leq T} |\Psi_t^e|^2 \right) < C$, we have

\[\mathbb{E} \int_{s}^{T} (k_t + G_t^* \Psi_t^e) d(\eta - \eta_t^e) - \mathbb{E} \int_{s}^{T} (k_t + G_t^* \Psi_t^e) d(\eta_t^e - \eta_t^e) \leq C \varepsilon^2. \quad (3.19)\]

Combining (3.18) and (3.19) the proof of Theorem 1 is completed.

**Corollary 2** Under the assumptions of Theorem 1, we have

\[\mathbb{E} \int_{s}^{T} (k_t + G_t^* \Psi_t^e) d\eta_t^e \leq \inf_{\eta \in \mathbb{U}_2([s, T])} \mathbb{E} \int_{s}^{T} (k_t + G_t^* \Psi_t^e) d\eta_t + C \varepsilon^2. \quad (3.20)\]

**4 Sufficient near-optimality conditions**

In this section, we prove that under an additional assumption, the near-maximum condition on the Hamiltonian function is sufficient for near-optimality. We assume:

(H4) $\varphi$ is differentiable in $u$ for $\varphi = f, \sigma, \ell$ and there is a constant $C$ such that

\[|\varphi(t, x, u) - \varphi(t, x, u')| + |\varphi_u(t, x, u) - \varphi_u(t, x, u')| \leq C |u - u'|. \quad (4.1)\]

**Theorem 2** Assume the $H(t, \cdot, \cdot, \Psi_t^e, K_t^e)$ is concave for a.e. $t \in [s, T]$, $\mathbb{P}$-a.s., and $h$ is convex. Let $(\Psi_t^e, K_t^e)$, $(Q_t^e, R_t^e)$ be the solutions of the adjoint equations (2.5)–(2.6) associated with $(u^e, \eta^e)$. If for some $\varepsilon > 0$ and for any $(u, \eta) \in \mathbb{U}$:

\[\mathbb{E} \int_{s}^{T} \mathcal{H}(x^e, u^e) (t, x_t^e, u_t^e) dt \geq \sup_{u \in \mathbb{U}_1([s, T])} \mathbb{E} \int_{s}^{T} \mathcal{H}(x^e, u^e) (t, x_t^e, u_t) dt - \varepsilon, \quad (4.2)\]

and

\[\mathbb{E} \left[ \int_{s}^{T} k_t (\eta_t - \eta^e_t) \right] \geq -C \varepsilon^{1/2}, \quad (4.3)\]

then we have

\[J(s, y, u^e, \eta^e) \leq \inf_{(u, \eta) \in \mathbb{U}([s, T])} J(s, y, u, \eta) + C \varepsilon^{1/2}, \quad (4.4)\]

where $C$ is a positive constant independent of $\varepsilon$. 

\(\copyright\) Springer
Proof First, define the cost functional

$$J(s, y, u, \eta) = J_1(s, y, u) + J_2(s, \eta),$$

(4.5)

where

$$J_1(s, y, u) = E \left[ h(x_T) + \int_s^T \ell(t, x_t, u_t) \, dt \right],$$

and

$$J_2(s, \eta) = E \left[ \int_s^T k_t \, d\eta_t \right].$$

Let us fix $\varepsilon > 0$, define a new metric $\tilde{d}$ on $U([s, T])$ as follows: for any $(u, \eta)$ and $(v, \xi) \in U:

$$\tilde{d}((u, \eta), (v, \xi)) = \tilde{d}_1(u, v) + d_2(\eta, \xi),$$

where

$$\tilde{d}_1(u, v) = E \left[ \int_s^T \zeta_t \, |u_t - v_t| \, dt \right],$$

(4.6)

and

$$\zeta_t = 1 + |\Psi_t^\varepsilon| + |K_t^\varepsilon| + |Q_t^\varepsilon| + |R_t^\varepsilon| \geq 1.$$  

(4.7)

Obviously $\tilde{d}_1$ is a metric on $(U_1, \tilde{d}_1)$, and it is a complete metric as a weighted $L^1$ norm. Hence $(U, \tilde{d})$ as a product of two complete metric spaces is a complete metric space under $\tilde{d}$.

Define a functional $\Upsilon$ on $U_1([s, T])$ by

$$\Upsilon(u) = E \left[ \int_s^T H(t, x_t^\varepsilon, u_t, \Psi_t^\varepsilon, K_t^\varepsilon, Q_t^\varepsilon) \, dt \right]$$

$$= E \left[ \int_s^T \mathcal{H}(x_t^\varepsilon, u_t) \, dt \right].$$
A simple computation shows that
\[ |\mathcal{Y}(u) - \mathcal{Y}(v)| \leq C \mathbb{E} \left[ \int_s^T \xi_t^\varepsilon |u_t - v_t| \, dt \right], \]
which implies that \( \mathcal{Y} \) is continuous on \( \mathbb{U}_1 ([s, T]) \) with respect to \( \tilde{d}_1 \). Now by using (4.2) and Ekeland’s variational principle, there exists a \( \tilde{u} \in \mathbb{U}_1 ([s, T]) \) such that
\[ \tilde{d}_1(\tilde{u}, u) \leq \epsilon^2, \]
and
\[ \mathbb{E} \int_s^T \tilde{H}(t, x_t^\varepsilon, \tilde{u}_t^\varepsilon) \, dt = \max_{u \in \mathbb{U}_1([s, T])} \mathbb{E} \int_s^T \tilde{H}(t, x_t^\varepsilon, u_t) \, dt, \tag{4.8} \]
where
\[ \tilde{H}(t, x, u) = \mathcal{H}(x, u)(t, x, u) - \varepsilon^2 \xi_t^\varepsilon |u - \tilde{u}_t^\varepsilon|. \tag{4.9} \]
The maximum condition (4.8) implies a pointwise maximum condition namely, for a.e. \( t \in [s, T] \) and \( \mathbb{P} - a.s. \),
\[ \tilde{H}(t, x_t^\varepsilon, \tilde{u}_t^\varepsilon) = \max_{u \in \mathcal{A}_1} \tilde{H}(t, x_t^\varepsilon, u). \]
Using Proposition A1 (Appendix), then we have
\[ \partial_u \tilde{H}(t, x_t^\varepsilon, \tilde{u}_t^\varepsilon) \ni 0. \]
Since \( |u - \tilde{u}_t^\varepsilon| \) is not differentiable in \( \tilde{u}_t^\varepsilon \) (locally Lipschitz), then we use Proposition A1 (Appendix) we get
\[ \partial_u \left\{ \varepsilon^2 \xi_t^\varepsilon |u - \tilde{u}_t^\varepsilon| \right\} = \left[ -\varepsilon^2 \xi_t^\varepsilon, \varepsilon^2 \xi_t^\varepsilon \right]. \]
By using (4.9) and fact that the Clarke’s generalized gradient of the sum of two functions is contained in the sum of the Clarke’s generalized gradient of the two functions, we get
\[ \partial_u \tilde{H}(t, x_t^\varepsilon, \tilde{u}_t^\varepsilon) \subset \partial_u \mathcal{H}(x, u)(t, x_t^\varepsilon, \tilde{u}_t^\varepsilon) + \left[ -\varepsilon^2 \xi_t^\varepsilon, \varepsilon^2 \xi_t^\varepsilon \right]. \]
Applying the similar method as in [22] for the rest of the proof, we obtain, for an arbitrary \( u \)
\[ J_1(s, y, u^\varepsilon) \leq J_1(s, y, u) + C \varepsilon^2. \tag{4.10} \]
Now, by using (4.3) we get

\[ J^2(s, \eta^\varepsilon) = \mathbb{E} \left[ \int_s^T k_t d\eta_t^\varepsilon \right] \leq \mathbb{E} \left[ \int_s^T k_t d\eta_t \right] + C \varepsilon^{\frac{1}{2}}, \]

which implies that for an arbitrary \( \eta \)

\[ J^2(s, \eta^\varepsilon) \leq J^2(s, \eta) + C \varepsilon^{\frac{1}{2}}. \]  \hspace{1cm} (4.11)

Combining (4.10), (4.11) and (4.5) we arrive at

\[ J(s, y, u^\varepsilon, \eta^\varepsilon) \leq J(s, y, u, \eta) + C \varepsilon^{\frac{1}{2}}. \]

Since \((u, \eta)\) is arbitrary, the desired result follows.

**Corollary 3** *Under the assumptions of Theorem 2, a sufficient condition for an admissible pair \((x^\varepsilon, u^\varepsilon, \eta^\varepsilon)\) to be \(\varepsilon\)-optimal is*

\[ \mathbb{E} \left\{ \int_s^T \mathcal{H}^{(x^\varepsilon, u^\varepsilon)}(t, x_t^\varepsilon, u_t^\varepsilon) dt - \int_s^T k_t d\eta_t^\varepsilon \right\} \geq \sup_{(u, \eta) \in \mathcal{U}([s, T])} \mathbb{E} \left\{ \int_s^T \mathcal{H}^{(x^\varepsilon, u^\varepsilon)}(t, x_t^\varepsilon, u_t) dt - \int_s^T k_t d\eta_t \right\} - \left( \frac{\varepsilon}{C} \right)^2. \]

**Example 1** Consider the one-dimensional stochastic control problem: \(n = l = 1, G_t = 1, A_1 = [0, 1], A_2 = [0, 1], \eta_1 = 1,\)

\[
\begin{cases}
    dx_t = u_t dW_t + d\eta_t, \\
    x_0 = 0
\end{cases}
\]  \hspace{1cm} (4.12)

and the cost functional being

\[ J(s, y, u, \eta) = \mathbb{E} \left[ \frac{1}{2} x_1^2 - \int_0^1 u_t dt + \int_0^1 k_t d\eta_t \right], \]  \hspace{1cm} (4.13)

For a given admissible pair \((x^\varepsilon, (u^\varepsilon, \eta^\varepsilon))\), the corresponding second-order adjoint equation is

\[
\begin{cases}
    dQ_t^\varepsilon = R_t^\varepsilon dW_t \\
    Q_1^\varepsilon = 1
\end{cases}
\]  \hspace{1cm} (4.14)
By the uniqueness of this solution, \((Q^\varepsilon, R^\varepsilon) = (1, 0)\), then for any admissible control \((u, \eta)\) we have

\[
H^{(x^\varepsilon, u^\varepsilon)}(t, x^\varepsilon_t, u_t) = u_t - (K^\varepsilon_t - Q^\varepsilon_t u_t) u_t - \frac{1}{2} Q^\varepsilon_t u_t^2
\]

\[
= \frac{1}{2} \left[ (u_t^\varepsilon - K^\varepsilon_t + 1)^2 - (u_t - u_t^\varepsilon + K^\varepsilon_t - 1)^2 \right].
\]

Replacing \(u_t = u_t^\varepsilon\), we get

\[
H^{(x^\varepsilon, u^\varepsilon)}(t, x^\varepsilon_t, u_t^\varepsilon) = \frac{1}{2} \left[ (u_t^\varepsilon - (K^\varepsilon_t - 1)^2 - (K^\varepsilon_t - 1)^2 \right]
\]

\[
= \frac{1}{2} \left[ (u_t^\varepsilon)^2 + 2u_t^\varepsilon (1 - K^\varepsilon_t) \right].
\]

Hence a simple computation shows that if

\[
u_t^\varepsilon - (K^\varepsilon_t - 1) \in [0, 1],
\]

then (3.15) and (3.20) gives

\[
\mathbb{E} \int_0^1 \frac{1}{2} \left[ (u_t^\varepsilon)^2 + 2u_t^\varepsilon (1 - K^\varepsilon_t) \right] dt
\]

\[
\geq \sup_{u \in [0, 1]} \mathbb{E} \int_0^1 \frac{1}{2} \left[ (u_t^\varepsilon - K^\varepsilon_t + 1)^2 - (u_t - u_t^\varepsilon + K^\varepsilon_t - 1)^2 \right] dt - C\varepsilon^\delta,
\]

and

\[
\mathbb{E} \int_s^T (k_t + \Psi^\varepsilon_t) d\eta^\varepsilon_t \leq \inf_{\eta \in \mathcal{U}_2([s,T])} \mathbb{E} \int_s^T (k_t + \Psi^\varepsilon_t) d\eta_t + C\varepsilon^\delta,
\]

thus

\[
\mathbb{E} \int_0^1 (K^\varepsilon_t - 1)^2 dt \leq C\varepsilon^\delta.
\]

(4.15)

We denote \(\mathcal{B} = \left\{ (w, t) \in \Omega \times [0, 1] : k_t + \Psi^\varepsilon_t \geq -C\varepsilon^{\frac{1}{2}} \right\}\) and define

\[
d\eta_t = \begin{cases} 0 & \text{if } k_t + \Psi^\varepsilon_t \geq -C\varepsilon^{\frac{1}{2}} \\ d\eta^\varepsilon_t & \text{otherwise.} \end{cases}
\]
Hence a simple calculation shows that

\[
\mathbb{E} \int_0^1 (k_t + \Psi_t^\varepsilon) d (\eta^\varepsilon - \eta_t^\varepsilon) = \mathbb{E} \int_0^1 (k_t + \Psi_t^\varepsilon) \chi_{\mathcal{B}} d (-\eta_t^\varepsilon) \geq -C \varepsilon^{\frac{1}{2}},
\]

which implies that

\[
\mathbb{E} \int_0^1 (k_t + \Psi_t^\varepsilon) \chi_{\mathcal{B}} d \eta_t^\varepsilon \leq C \varepsilon^\delta.
\]

(4.17)

It is worth mentioning that the above condition reveals the minimum qualification for the pair \((x^\varepsilon, (u^\varepsilon, \eta^\varepsilon))\) to be \(\varepsilon\)-optimal. As an example, the admissible control \((u^\varepsilon, \eta^\varepsilon) = (1 - \varepsilon^{\frac{1}{2}}, \eta^\varepsilon)\) is candidate for \(\varepsilon\)-optimality, where \(\varepsilon > 0\) is sufficiently small. Note that the first-order adjoint equation is

\[
\begin{cases}
  d\Psi_t^\varepsilon = K_t^\varepsilon dW_t, \\
  \Psi_1^\varepsilon = x_1^\varepsilon,
\end{cases}
\]

(4.18)

with the corresponding trajectories \(x_t^\varepsilon = (1 - \varepsilon^{\frac{1}{2}})W_t + \eta_t^\varepsilon\) then the unique solution pair of the first-order adjoint equation will be \((\Psi_t^\varepsilon, K_t^\varepsilon) = ((1 - \varepsilon^{\frac{1}{2}})W_t + 1, (1 - \varepsilon^{\frac{1}{2}}))\). Hence (4.15) and (4.16) will be satisfied.

Conversely, for the sufficient part, since the Hamiltonian \(H (t, x, u, p, q) = u - qu\) is concave in \((x, u)\), we use Theorem 2 to conclude that \(u_t^\varepsilon = (1 - (\frac{\varepsilon}{2})^2)\) as a candidate to be \(\varepsilon\)-optimality for \(\varepsilon\) sufficiently small is indeed an \(\varepsilon\)-optimal control.

Acknowledgments We are thankful to the anonymous reviewers and handling editor for their expert comments and suggestions which helped us to improve the manuscript considerably. The first author was supported by Algerian PNR project grant 08/u07/857. The third author was supported by the Czech CTU grant SGS 2012-2014 and MSMT grant INGO II INFRA LG12020.

Appendix

The following result gives some basic properties of the Clarke’s generalized gradient.

**Proposition A1** If \(f : \mathbb{R}^n \rightarrow \mathbb{R}\) is locally Lipschitz at \(x \in \mathbb{R}^n\), then the following statements holds

1. \(\partial_x f (x)\) is nonempty, compact, and convex set in \(\mathbb{R}^n\).
2. \(\partial_x (-f) (x) = -\partial_x (f) (x)\).
3. \(0 \in \partial_x (f) (x)\) if \(f\) attains a local minimum or maximum at \(x\).
4. If \(f\) is Fréchet-differentiable at \(x\), then \(\partial_x f (x) = \{ f' (x) \}\).
5. If \(f, g : \mathbb{R}^n \rightarrow \mathbb{R}\) are locally Lipschitz at \(x \in \mathbb{R}^d\), then

\[
\partial_x (f + g) (x) \subseteq \partial_x f (x) + \partial_x g (x).
\]
See [9] and [19, Lemma 2.3] for the detailed proof of the above Proposition.

As an example of the Clarke’s generalized gradient, we consider the absolute value function $f : x \mapsto |x|$ which is continuously differentiable everywhere except at 0. Since $f'(x) = 1$ for $x > 0$ and $f'(x) = -1$ for $x < 0$ then the Clarke’s generalized gradient of $f$ at $x = 0$ is given by

$$
\partial_x f(0) = \text{co}\{-1, 1\} = [-1, 1].
$$

References

1. Alvarez, L.H.R.: A class of solvable singular stochastic control problems. Stochastics, Stochastics Rep. 67, 83–122 (1999)
2. Alvarez, L.H.R.: Singular stochastic control linear diffusion and optimal stopping: a class of solvable problems. SIAM J. Contr. Optim. 39, 1697–1710 (2001)
3. Bahlali, S., Chala, A.: The stochastic maximum principle in optimal control of singular diffusions with non linear coefficients. Random Oper. Stoch. Equ. (ROSE), 13, 1–10 (2005)
4. Bahlali, S., Mezerdi, B.: A general stochastic maximum principle for singular control problems. Electron. J. Probab. 10, 988–1004 (2005)
5. Bahlali, K., Khelfallah, N., Mezerdi, B.: Necessary and sufficient conditions for near optimality in stochastic control of FBSDEs. Syst. Control Lett. 58, 857–864 (2009)
6. Bellman, R.: Dynamic programming. Princeton University Press, Princeton (1957)
7. Cadenillas, A., Haussman, U.: The stochastic maximum principle for singular control problem. Stochastics Rep. 49(3–4), 211–237 (1994)
8. Chighoub, F., Mezerdi, B.: Near optimality conditions in stochastic control of jump diffusion processes. Syst. Control Lett. 60, 907–916 (2011)
9. Clarke, F.H.: Optimization and nonsmooth analysis. Wiley, New York (1983)
10. Dufour, F., Miller, B.: Maximum principle for singular stochastic control problem. Siam J. Control Optim. 45(2), 668–698 (2006)
11. Dufour, F., Miller, B.: Singular stochastic control problem. Siam J. Control Optim. 43(2), 705–730 (2004)
12. Ekeland, I.: On the variational principle. J. Math. Anal. Appl. 47, 443–474 (1974)
13. Haussmann, U.G., Suo, W.: Singular optimal control I, II. SIAM J. Contr. Optim. 33(3), 916–936, 937–959 (1995)
14. Karatzas, I., Ocone, D., Wang, H., Zervos, M.: Finite-fuel singular control with discretionary stopping. Stochastics Rep. 71, 1–50 (2000)
15. Gabasov, R., Kirillova, M., Mordukhovich, B.S.: The $\varepsilon$– maximum principle for suboptimal controls. Soviet Math Dokh 27, 95–99 (1983)
16. Pan, L.P., Teo, K.L.: Near-optimal controls of class of volterra integral systems. J. Optim. Control Theory Appl. 101(2), 355–373 (1999)
17. Peng, S.: A general stochastic maximum principle for optimal control problems. SIAM J. Contr. Optim. 28(4), 966–979 (1990)
18. Pontryagin, L.S., Bolianski, V.G., Gamkrelidze, R.V.: The Mathematical Theory of Optimal Processes. Inter-scien, New York (1962)
19. Yong, J., Zhou, X.Y.: Stochastic Controls. Hamiltonian Systems and HJB Equations. Springer-Verlag. New York (1999)
20. Zhou, X.Y.: Deterministic near-optimal controls. Part I: Necessary and sufficient conditions for near optimality. J. Optim. Theory Appl. 85, 473–488 (1995)
21. Zhou, X.Y.: Deterministic near-optimal controls. Part II: Dynamic programming and viscosity solution approach. Math. Operat. Res. 21, 655–674 (1996)
22. Zhou, X.Y.: Stochastic near-optimal controls: Necessary and sufficient conditions for near-optimality. SIAM. Control. Optim. 36, 929–947 (1998)