Systems & Control Letters 60 (2011) 907–916

Contents lists available at SciVerse ScienceDirect

Systems & Control Letters

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

Near optimality conditions in stochastic control of jump diffusion processes

Farid Chighoub, Brahim Mezerdi *

Laboratory of Applied Mathematics, University Mohamed Khider, PO. Box 145 Biskra (07000), Algeria

ARTICLE INFO

Article history:
Received 12 March 2011
Received in revised form
5 June 2011
Accepted 13 July 2011
Available online 28 September 2011

Keywords:
Diffusion with jumps
Maximum principle
Near-optimal control
Adjoint process

ABSTRACT

This paper is concerned with necessary as well as sufficient conditions for near-optimality of controlled jump diffusion processes. Necessary conditions for a control to be near-optimal are derived, using Ekeland's variational principle and some stability results on the state and adjoint processes, with respect to the control variable. In a second step, we show that the necessary conditions for near-optimality, are in fact sufficient for near-optimality provided some concavity conditions are fulfilled. Finally, as an illustration some examples are solved explicitly.

1. Introduction

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P}\right)$ be a probability space such that $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets, $\mathcal{F}_T = \mathcal{F}$ and $(\mathcal{F}_t)_{t\leq T}$ satisfies the usual conditions. We assume that, $(\mathcal{F}_t)_{t\leq T}$ is the natural filtration of a d-dimensional standard Brownian motion B and an independent Poisson random measure N on $[0,T] \times E$, where $E = \mathbb{R}^m \setminus \{0\}$ for some $m \geq 1$. We assume that the compensator of N has the form $\mu(de, dt) = \nu(de)\,dt$, for some positive, σ-finite Lévy measure ν on E, endowed with its Borel σ-field $\mathcal{B}(E)$ satisfying $\int_E 1 \wedge |e|^2 \,\nu(de) < \infty$. Define the measure $\mathbb{P}\otimes\mu$ on $(\Omega \times [0,T] \times E, \mathcal{F} \times \mathcal{B}([0,T]) \times \mathcal{B}(E))$ by

$$\mathbb{P}\otimes\mu(G) = \mathbb{E}\left[\iint_{[0,T]\times E} 1_G(\omega, t, e)\,\mu(de, dt)\right],$$
$$\text{for } G \in \mathcal{F} \times \mathcal{B}([0,T]) \times \mathcal{B}(E),$$

which is called the measure generated by μ. We write $\widetilde{N} = N - \nu$ for the compensated jump martingale random measure of N.

We denote by $\left(\mathcal{F}_t^B\right)_{t\leq T}$ (resp. $\left(\mathcal{F}_t^N\right)_{t\leq T}$) the $\mathbb{P}$-augmentation of the natural filtration of B (resp. N). Obviously, we have

$$\mathcal{F}_t = \sigma\left[\iint_{A\times(0,s]} N(de, dr)\,;\, s \leq t, A \in \mathcal{B}(E)\right] \vee\sigma[B_s; s\leq t] \vee \mathcal{N},$$

where $\mathcal{N}$ denotes the totality of ν-null sets and $\sigma_1 \vee \sigma_2$ denotes the σ-field generated by $\sigma_1 \cup \sigma_2$.

In this paper, we discuss stochastic control models driven by a stochastic differential equation with jumps of the form

$$\begin{cases} dx_t = b(t, x_t, u_t)\,dt + \sigma(t, x_t, u_t)\,dB_t \\ \quad + \displaystyle\int_E r(t, x_{t-}, u_t, e)\,\widetilde{N}(dt, de), \\ x_s = a. \end{cases} \tag{1.1}$$

The expected cost functional is given by

$$J(s, y, u) = \mathbb{E}\left[\int_s^T f(t, x_t, u_t)\,dt + g(x_T)\right], \quad \text{for } u \in \mathcal{U}. \tag{1.2}$$

It is well-known that the dynamic programming principle can be extended to stochastic control problems with jumps and the corresponding Hamilton–Jacobi–Bellman equation is a nonlinear second order parabolic integrodifferential equation. Pham [1] has studied a mixed optimal stopping and stochastic control of jump diffusion processes by using the viscosity solutions approach. Some verification theorems of various types for problems governed by these kinds of SDEs are discussed in øksendal and Sulem [2]. The stochastic maximum principle is another powerful tool for solving stochastic control problems. Some results that cover the controlled jump diffusion processes are discussed; see [3–7]. The necessary and sufficient conditions of optimality for partial information control problems are given in [3]. In [6], the sufficient maximum principle and the link with the dynamic programming principle are given. The second order stochastic maximum principle for optimal controls of nonlinear dynamics with jumps and convex state constraints was developed via spike variation method by Tang and Li [7], extending the Peng maximum principle [8]. These conditions are described in terms of two adjoint processes,

* Corresponding author. Tel.: +213 33 74 77 88; fax: +213 33 74 77 88.
E-mail addresses: chighoub_farid@yahoo.fr (F. Chighoub), bmezerdi@yahoo.fr (B. Mezerdi).

0167-6911/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2011.07.009

which are linear backward SDEs. Such equations have important applications in hedging problems; see [9]. See also Becherer [10] for an application to exponential utility maximization in finance. Existence and uniqueness of solutions of BSDEs with jumps and nonlinear coefficients have been treated by Barles et al. [11], Royer [12] and Tang and Li [7]. The link with integral-partial differential equations is studied in [11]. See Bouchard and Elie [13] for discrete time approximation of decoupled FBSDE with jumps.

In this paper, instead of optimality conditions, we are interested in near optimal controls for systems governed by jump diffusion processes. Indeed, optimal controls may not even exist in many situations. This justifies the use of near optimal controls which always exist. Moreover, since there are many more candidates for near-optimal controls, it is possible to choose suitable ones, that are convenient for analysis and implementation. For deterministic control problems, the first result on necessary conditions for near-optimality has been proved in Ekeland [14], see also Zhou [15], by using Ekeland's variational principle. These necessary conditions were derived only for some near-optimal controls.

Various necessary conditions for near-optimal control problem for systems driven by Itô SDEs with an uncontrolled diffusion coefficient, have been established in [16,17]. These results played an important role in the stochastic maximum principle when the coefficients of the state dynamics and the cost functional are nonsmooth; see [18–20].

The general case of systems driven by SDE with controlled diffusion coefficient has been treated in [21]. See also [22] for systems driven by FBSDE. Zhou [21], and Bahlali et al. [22] showed that any near-optimal control (in terms of a small parameter ε) nearly maximizes the $\mathcal{H}$-function in the integral form. Under certain concavity conditions, the near-maximum condition of the $\mathcal{H}$-function in the integral form is sufficient for near-optimality.

In the second section, we formulate the problem and give the notations and assumptions used throughout the paper. In Sections 3 and 4, we derive necessary as well as sufficient conditions for near optimality respectively, which are our main results. Finally, using these results, we solve explicitly some examples from linear problems.

2. Assumptions and problem formulation

This section sets out the notations and assumptions used in the sequel.

Notation. 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 $x \cdot y$, we denote M^T the transpose of any vector or matrix M. For a function h, we denote by h_x (resp. h_{xx}) the gradient or Jacobian (resp. the Hessian) of h with respect to the variable x.

Definition 2.1. Let T be a strictly positive real number and U a non empty subset of $\mathbb{R}^n$. An admissible control is defined as a function, $u : [0, T] \times \Omega \to U$ which is Borel measurable and $\mathcal{F}_t$-predictable, such that, the SDE (2.1) has a unique solution, and write $u \in \mathcal{U}$.

Let us consider the following stochastic control problem.

For $u \in \mathcal{U}$, suppose that the state x_t of a controlled jump diffusion in $\mathbb{R}^n$ is described by the following stochastic differential equation

$$\begin{cases} dx_t = b(t, x_t, u_t)\, dt + \sigma(t, x_t, u_t)\, dB_t \\ \qquad + \int_E r(t, x_{t-}, u_t, e)\, \widetilde{N}(dt, de), \\ x_s = a \end{cases} \tag{2.1}$$

where $(s, a) \in [0, T) \times \mathbb{R}^n$ be given, representing the initial time and initial state respectively, of the system. As before $\widetilde{N}(dt, de) = \left(\widetilde{N}_1(dt, de_1), \ldots, \widetilde{N}_l(dt, de_l)\right)^T$, where $\widetilde{N}_j(dt, de_j) = N_j(dt, de_j) - \nu_j(de_j)$, $1 \le j \le m$.

Assume that the cost functional has the form

$$J(s, y, u) = \mathbb{E}\left[\int_s^T f(t, x_t, u_t)\, dt + g(x_T)\right], \quad \text{for } u \in \mathcal{U}, \tag{2.2}$$

where $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$. Here $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times d}$, $r : [0, T] \times \mathbb{R}^n \times U \times E \to \mathbb{R}^{n\times m}$, $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$, are measurable functions.

The following assumptions will be in force throughout this paper:

(H1) For each $(t, x, u, e) \in [0, T] \times \mathbb{R}^n \times U \times E$, the maps b, σ, and r are twice continuously differentiable in x, and there exists a constant $M > 0$ such that, for $h = b, \sigma$

$$\left|h(t, x, u) - h\left(t, x', u\right)\right| + \sup_{e\in E}\left\{\left|r(t, x, u, e) - r\left(t, x', u, e\right)\right|\right\} \le M\left|x - x'\right|, \tag{2.3}$$

$$\left|h_x(t, x, u) - h_x\left(t, x', u\right)\right| + \sup_{e\in E}\left\{\left|r_x(t, x, u, e) - r_x\left(t, x', u, e\right)\right|\right\} \le M\left|x - x'\right|, \tag{2.4}$$

$$|h(t, x, u)| + \sup_{e\in E}|r(t, x, u, e)| \le M(1 + |x|). \tag{2.5}$$

(H2) For each $(t, x, u) \in [0, T] \times \mathbb{R}^n \times U$, the maps f and g are twice continuously differentiable in x with bounded derivatives, and there exists a constant $M > 0$ such that

$$\left|f(t, x, u) - f\left(t, x', u\right)\right| + \left|g(x) - g\left(x'\right)\right| \le M\left|x - x'\right|, \tag{2.6}$$

$$\left|f_x(t, x, u) - f_x\left(t, x', u\right)\right| + \left|g_x(x) - g_x\left(x'\right)\right| \le M\left|x - x'\right|, \tag{2.7}$$

$$|f(t, x, u)| + |g(x)| \le M(1 + |x|). \tag{2.8}$$

Under the above hypothesis, the SDE (2.1) has a unique strong solution, and by standard arguments it is easy to show that for any $p > 0$,

$$E\left[\sup_{0\le t\le T}|x_t|^p\right] < \infty, \tag{2.9}$$

and the functional J is well defined.

The objective of the exact optimality problems, is to minimize the functional $J(u)$ over all $u \in \mathcal{U}$, i.e. we seek for u^* such that $J(u^*) = \inf_{u\in\mathcal{U}} J(u)$. Any admissible control u^* that achieves the minimum is called an optimal control, and it implies an associated optimal state evolution x^* from (2.1), and we call (x^*, u^*) an optimal solution. Then, since our objective in this paper is to study near-optimality rather than exact optimality, we start with the definition of near-optimality; see for example [21].

Definition 2.2. An admissible pair $(x^\varepsilon, u^\varepsilon)$ parameterised by $\varepsilon > 0$, is called near-optimal if

$$\left|J(u^\varepsilon) - J\left(u^*\right)\right| \le R(\varepsilon),$$

holds for sufficiently small ε, where R is a function of ε satisfying $R(\varepsilon) \to 0$ as $\varepsilon \to 0$.

The estimate $R(\varepsilon)$ is called an error bound. If $R(\varepsilon) = c\varepsilon^\alpha$ for some $\alpha > 0$ independent of the constant c, then u^ε is called *near-optimal with order* ε^α.

For any $u \in \mathcal{U}$ and the corresponding state trajectory x, we define the first order adjoint process ψ and the second-order adjoint process Ψ as solutions of the following two BSDEs, respectively

$$\begin{cases} d\psi(t) = -\Big(b_x(t,x_t,u_t)^T\psi_t + \sigma_x(t,x_t,u_t)^T\phi_t \\ \quad + \int_E r_x(t,x_t,u_t,e)^T\gamma_t(e)\,\nu(de) \\ \quad + f_x(t,x_t,u_t)\Big)dt + \phi_t dB_t + \int_E \gamma_t(e)\widetilde{N}(dt,de), \\ \psi(T) = g_x(x_T). \end{cases} \quad (2.10)$$

$$\begin{cases} d\Psi_t = -\big(b_x(t,x_t,u_t)^T\Psi_t + \Psi_t\cdot b_x(t,x_t,u_t) \\ \quad + \sigma_x(t,x_t,u_t)^T\Psi_t\sigma_x(t,x_t,u_t) \\ \quad + \sigma_x(t,x_t,u_t)^T\Phi_t + \Phi_t\cdot\sigma_x(t,x_t,u_t) \\ \quad + \int_E r_x(t,x_t,u_t,e)^T(\Gamma_t(e)+\Psi_t)\,r_x(t,x_t,u_t,e)\,\nu(de) \\ \quad + \int_E\big(\Gamma_t(e)\cdot r_x(t,x_t,u_t,e) \\ \quad + r_x(t,x_t,u_t,e)^T\Gamma_t(e)\big)\nu(de)\big)dt \\ \quad + H_{xx}(t,x_t,u_t,\psi_t,\phi_t,\gamma_t(e))\,dt + \Phi_t dB_t \\ \quad + \int_E \Gamma_t(e)\widetilde{N}(dt,de), \\ \Psi_T = g_{xx}(x_T). \end{cases} \quad (2.11)$$

Note that under assumptions (2.3)–(2.7), the linear BSDEs (2.10) and (2.11) admit unique $\mathcal{F}_t$-adapted solutions $(\psi,\phi,\gamma)\in\mathbb{R}^n\times\mathbb{R}^{n\times d}\times\mathbb{R}^{n\times m}$ and $(\Psi,\Phi,\Gamma)\in\mathbb{R}^{n\times n}\times\left(\mathbb{R}^{n\times n}\right)^d\times\left(\mathbb{R}^{n\times n}\right)^m$, with ψ and Ψ being cadlag processes. Moreover, since the coefficients b_x,σ_x,r_x,f_x and g_x are bounded by the Lipschitz constant M, we deduce from standard arguments that, there exists a constant C, independent of (x,u), such that the solutions of (2.10) and (2.11) satisfy

$$\mathbb{E}\left[\sup_{s\le t\le T}|\psi_t|^2 + \int_s^T|\phi_t|^2dt + \int_s^T\int_E|\gamma_t(e)|^2\nu(de)\,dt\right]\le C, \quad (2.12)$$

$$\mathbb{E}\left[\sup_{s\le t\le T}|\Psi_t|^2 + \int_s^T|\Phi_t|^2dt + \int_s^T\int_E|\Gamma_t(e)|^2\nu(de)\,dt\right]\le C. \quad (2.13)$$

Define the usual Hamiltonian for $(t,x,u,p,q,\mathrm{X})\in[s,T]\times\mathbb{R}^n\times U\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\times\mathbb{R}^{n\times m}$, by

$$H(t,x,u,p,q,\mathrm{X}) = -f(t,x,u) - pb(t,x,u) - q\sigma(t,x,u) - \int_E \mathrm{X}(e)\,r(t,x,u,e)\,\nu(de). \quad (2.14)$$

Furthermore, we define the $\mathcal{H}$-function corresponding to a given admissible pair (y,v) as follows

$$\begin{aligned}\mathcal{H}^{(y,v)}(t,x,u) = {} & H(t,x,u,\psi_t,\phi_t,\gamma_t(e)) \\ & + \sigma(t,x,u)^T\Psi_t\sigma(t,y_t,v_t) \\ & - \frac{1}{2}\sigma(t,x,u)^T\Psi_t\sigma(t,x,u) \\ & + \int_E\Big(r(t,x,u,e)^T(\Psi_t+\gamma_t(e)) \\ & \times r(t,y_t,v_t,e) - \frac{1}{2}r(t,x,u,e)^T \\ & \times (\Psi_t+\gamma_t(e))\,r(t,x,u,e)\Big)\nu(de),\end{aligned}$$

for $(t,x,u)\in[s,T]\times\mathbb{R}^n\times U$, where $\psi_t,\phi_t,\gamma_t(e)$, and Ψ_t are determined by adjoint equations (2.10) and (2.11) corresponding to (y_t,v_t).

3. Necessary conditions of near optimality

This section is devoted to the presentation of necessary conditions for all near-optimal controls. The proof of the main result is based on some stability results with respect to the control variable of the state process and adjoint processes, along with Ekeland's variational principle. First, we endow the set of controls with an appropriate metric

$$d(u,v) = \mathbb{P}\otimes dt\{(w,t)\in\Omega\times[0,T],\, v(\omega,t)\ne u(\omega,t)\},$$

where $\mathbb{P}\otimes dt$ is the product measure of $\mathbb{P}$ with the Lebesgue measure dt. It is easy to show that $(\mathcal{U},d)$ is a complete metric space (see e.g. [7]), and by assumptions (2.3)–(2.7) we can prove that $J(u)$ is continuous on $\mathcal{U}$ endowed with the metric d.

The main result of this section is stated in the following theorem.

Theorem 3.1. *For any $\delta\in\left[0,\frac{1}{3}\right)$, there exists a constant $C=C(\delta,\nu(E))>0$ such that for any $\varepsilon>0$, and any ε-optimal pair $(x^\varepsilon,u^\varepsilon)$ for problems (2.1) and (2.2), it holds that*

$$\begin{aligned}-C\varepsilon^\delta \le \mathbb{E}\Bigg[\int_s^T\Big\{ & \psi_t^\varepsilon\cdot\left(b\left(t,x_t^\varepsilon,u\right)-b\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right) \\ & + \phi_t^\varepsilon\cdot\left(\sigma\left(t,x_t^\varepsilon,u\right)-\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right) \\ & + \int_E\gamma_t^\varepsilon(e)\cdot\left(r\left(t,x_t^\varepsilon,u,e\right)-r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right) \\ & \times\nu(de) + \left(f\left(t,x_t^\varepsilon,u\right)-f\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right) \\ & + \frac{1}{2}\left(\sigma\left(t,x_t^\varepsilon,u\right)-\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right)^T \\ & \times\Psi_t^\varepsilon\left(\sigma\left(t,x_t^\varepsilon,u\right)-\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right) \\ & + \frac{1}{2}\int_E\left(r\left(t,x_t^\varepsilon,u,e\right)-r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right)^T \\ & \times\left(\Psi_t^\varepsilon+\gamma_t^\varepsilon(e)\right)\left(r\left(t,x_t^\varepsilon,u,e\right)\right. \\ & \left.-\ r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right)\nu(de)\Big\}dt\Bigg], \end{aligned} \quad (3.1)$$

where $(\psi^\varepsilon,\phi^\varepsilon,\gamma^\varepsilon(\cdot))$ and $(\Psi^\varepsilon,\Phi^\varepsilon,\Gamma^\varepsilon(\cdot))$ are the solutions to (2.10) and (2.11) respectively, corresponding to $(x^\varepsilon,u^\varepsilon)$.

To prove Theorem 3.1, we need some preliminary results given in the following two lemmas.

In what follows, C represents a generic constant, which can be different from line to line.

Lemma 3.2. *Let u,u' be two admissible controls, x,x' are the solutions of the state SDE (2.1) corresponding to u,u' respectively. Then, for any $\alpha\in(0,1)$ and $p\in(0,2]$ satisfying $\alpha p<1$, there is a positive constant $C=C(\alpha,p,\nu(E))$, such that*

$$\mathbb{E}\left[\sup_{s\le t\le r}\left|x_t-x_t'\right|^p\right]\le Cd\left(u,u'\right)^{\frac{\alpha p}{2}}. \quad (3.2)$$

Proof. According to Hölder's inequality, it is sufficient to prove the above estimate for $p=2$.

First of all, using the Burkholder–Davis–Gundy inequality and Proposition A.2 in the Appendix, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{s\le t\le r}\left|x_t-x_t'\right|^2\right] & \le C\mathbb{E}\left[\int_s^r\left(\left|b(t,x_t,u_t)-b\left(t,x_t',u_t'\right)\right|^2\right.\right. \\ & + \left|\sigma(t,x_t,u_t)-\sigma\left(t,x_t',u_t'\right)\right|^2 \\ & \left.\left. + \int_E\left|r(t,x_t,u_t,e)-r\left(t,x_t',u_t',e\right)\right|^2\nu(de)\right)dt\right] \\ & \le A_1+A_2,\end{aligned}$$

where

$$\begin{aligned} A_1 &= C\mathbb{E}\left[\int_s^r \left(\left|b(t,x_t,u_t) - b(t,x_t,u'_t)\right|^2 \right.\right. \\ &\quad + \left|\sigma(t,x_t,u_t) - \sigma(t,x_t,u'_t)\right|^2 \\ &\quad + \left.\nu(E)\left(\sup_{e\in E}\left|r(t,x_t,u_t,e) - r(t,x_t,u'_t,e)\right|\right)^2\right) \\ &\quad \left.\times\ 1_{\{u_t\neq u'_t\}}(t)\,dt\right] \\ A_2 &= C\mathbb{E}\left[\int_s^r \left(\left|b(t,x_t,u'_t) - b(t,x'_t,u'_t)\right|^2 \right.\right. \\ &\quad + \left|\sigma(t,x_t,u'_t) - \sigma(t,x'_t,u'_t)\right|^2 \\ &\quad \left.\left. + \nu(E)\left(\sup_{e\in E}\left|r(t,x_t,u'_t,e) - r(t,x'_t,u'_t,e)\right|\right)^2\right) dt\right]. \end{aligned}$$

Due to the linear growth of the coefficients and from the Schwarz inequality, we get

$$A_1 \le C\mathbb{E}\left[\int_s^r (1+|x_t|)^{\frac{2}{1-\alpha}}\,dt\right]^{1-\alpha} d\left(u,u'\right)^\alpha .$$

This means that $A_1 \le C\cdot d\left(u,u'\right)^\alpha$. Since the coefficients of the SDE (2.1) are Lipschitz with respect to the state variable, we get $A_2 \le C\cdot\mathbb{E}\left[\int_s^r \left|x_t - x'_t\right|^2 dt\right]$ and

$$\mathbb{E}\left[\sup_{s\le t\le r}\left|x_t - x'_t\right|^2\right] \le C\left(\int_s^r \mathbb{E}\left[\sup_{s\le t\le\theta}\left|x_t - x'_t\right|^2\right] d\theta + d\left(u,u'\right)^\alpha\right). \tag{3.3}$$

Hence (3.2) follows from the Gronwall lemma. This completes the proof. □

The next lemma is an extension of theorem 4.1 in [21] to the controlled jump diffusion processes. Note that, in [21] the author considers the Brownian case only.

Lemma 3.3. *For any $\alpha \in (0,1)$ and $p \in (1,2)$ satisfying $(1+\alpha)p < 2$, there is a positive constant $C = C(\alpha, p, \nu(E))$ such that*

$$\mathbb{E}\left[\left|\psi_t - \psi'_t\right|^p + \int_s^T \left|\phi_t - \phi'_t\right|^p dt + \int_s^T\int_E \left|\gamma_t(e) - \gamma'_t(e)\right|^p \nu(de)\,dt\right] \le Cd\left(u,u'\right)^{\frac{\alpha p}{2}}, \tag{3.4}$$

$$\mathbb{E}\left[\left|\Psi_t - \Psi'_t\right|^p + \int_s^T \left|\Phi_t - \Phi'_t\right|^p dt + \int_s^T\int_E \left|\Gamma_t(e) - \Gamma'_t(e)\right|^p \nu(de)\,dt\right] \le Cd\left(u,u'\right)^{\frac{\alpha p}{2}}, \tag{3.5}$$

where $(\psi,\phi,\gamma(\cdot))$ and $\left(\psi',\phi',\gamma'(\cdot)\right)$ (resp. $(\Psi,\Phi,\Gamma(\cdot))$ and $\left(\Psi',\Phi',\Gamma'(\cdot)\right)$) denote the unique solutions to the first-order (resp. second-order) adjoint equation (2.10) (resp. (2.11)), corresponding to the admissible pair (x,u) and $\left(x',u'\right)$.

Proof. Denote by $(\overline{\psi}_t, \overline{\phi}_t, \overline{\gamma}_t(e)) = (\psi_t - \psi'_t, \phi_t - \phi'_t, \gamma_t(e) - \gamma'_t(e))$ the unique solution of the linear backward stochastic differential equation

$$\begin{cases} d\overline{\psi}(t) = -\Big(b_x(t,x_t,u_t)^T\,\overline{\psi}_t + \sigma_x(t,x_t,u_t)^T\,\overline{\phi}_t \\ \quad + \int_E r_x(t,x_t,u_t,e)^T\,\overline{\gamma}_t(e)\,\nu(de) \\ \quad + \overline{f}(t)\Big)dt + \overline{\phi}_t dB_t + \int_E \overline{\gamma}_t(e)\,\widetilde{N}(dt,de), \\ \overline{\psi}(T) = g_x(x_T) - g_x\left(x'_T\right), \end{cases} \tag{3.6}$$

where

$$\begin{aligned} \overline{f}(t) = &\left(b_x(t,x_t,u_t)^T - b_x\left(t,x'_t,u'_t\right)^T\right)\psi'_t + \Big(\sigma_x(t,x_t,u_t)^T \\ &- \sigma_x\left(t,x'_t,u'_t\right)^T\Big)\phi'_t \\ &+ \int_E \left(r_x(t,x_t,u_t,e)^T - r_x\left(t,x'_t,u'_t,e\right)^T\right)\gamma'_t(e)\,\nu(de) \\ &+ \left(f_x(t,x_t,u_t) - f_x\left(t,x'_t,u'_t\right)\right). \end{aligned}$$

Now, let Λ be a solution of the following linear stochastic differential equation

$$\begin{cases} d\Lambda(t) = \left(b_x(t,x_t,u_t)\,\Lambda_t + \left|\overline{\psi}_t\right|^{p-1}\operatorname{sgn}\left(\overline{\psi}_t\right)\right)dt \\ \quad + \left(\sigma_x(t,x_t,u_t)\,\Lambda_t + \left|\overline{\phi}_t\right|^{p-1}\operatorname{sgn}\left(\overline{\phi}_t\right)\right)dB_t \\ \quad \times \int_E \left(r_x(t,x_t,u_t,e)\,\Lambda_t + \left|\overline{\gamma}_t(e)\right|^{p-1}\operatorname{sgn}\left(\overline{\gamma}_t(e)\right)\right) \\ \quad \times\widetilde{N}(dt,de), \\ \Lambda(s) = 0, \end{cases} \tag{3.7}$$

where $\operatorname{sgn}(a)\equiv(\operatorname{sgn}(a_1),\ldots,\operatorname{sgn}(a_n))$ for a vector $a = (a_1,\ldots,a_n)^T$. The coefficients b_x, σ_x and r_x being bounded by the Lipschitz constant and the fact that

$$\mathbb{E}\left[\int_s^T \left(\left|\left|\overline{\psi}_t\right|^{p-1}\operatorname{sgn}\left(\overline{\psi}_t\right)\right|^2 + \left|\left|\overline{\phi}_t\right|^{p-1}\operatorname{sgn}\left(\overline{\phi}_t\right)\right|^2 + \int_E \left|\left|\overline{\gamma}_t(e)\right|^{p-1}\operatorname{sgn}\left(\overline{\gamma}_t(e)\right)\right|^2\nu(de)\right)dt\right] < \infty, \tag{3.8}$$

imply that the linear SDE (3.7) satisfies the Itô conditions. Therefore, it has a unique solution. Moreover, we conclude from standard arguments, the Burkholder–Davis–Gundy inequality and the Gronwall lemma, that

$$\begin{aligned} \mathbb{E}\left[\sup_{s\le t\le T}|\Lambda_t|^q\right] &\le C\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^{(p-1)q} + \left|\overline{\phi}_t\right|^{(p-1)q} + \int_E \left|\overline{\gamma}_t(e)\right|^{(p-1)q}\nu(de)\right)dt\right] \\ &= C\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^p + \left|\overline{\phi}_t\right|^p + \int_E \left|\overline{\gamma}_t(e)\right|^p\nu(de)\right)dt\right]. \end{aligned} \tag{3.9}$$

with $q \in (2,+\infty)$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$. In view of (2.12), it yields

$$\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^p + \left|\overline{\phi}_t\right|^p + \int_E \left|\overline{\gamma}_t(e)\right|^p\nu(de)\right)dt\right] \le C. \tag{3.10}$$

On the other hand, by applying Ito's formula to $\overline{\psi}_t\cdot\Lambda_t$ and taking expectations, we obtain

$$\mathbb{E}\left[\int_s^T \bar{f}(t)\,\Lambda_t dt + \left(g_x(x_T) - g_x\left(x'_T\right)\right)\Lambda_T\right]$$
$$= \mathbb{E}\left[\int_s^T \left(\overline{\psi}_t \left|\overline{\psi}_t\right|^{p-1} \operatorname{sgn}\left(\overline{\psi}_t\right) dt + \overline{\phi}_t \left|\overline{\phi}_t\right|^{p-1} \operatorname{sgn}\left(\overline{\phi}_t\right) + \int_E \overline{\gamma}_t(e)\left|\overline{\gamma}_t(e)\right|^{p-1} \operatorname{sgn}\left(\overline{\gamma}_t(e)\right)\nu(de)\right) dt\right]. \tag{3.11}$$

First, it follows from (3.9) that

$$\mathbb{E}\left[\int_s^T \bar{f}(t)\,\Lambda_t dt + \left(g_x(x_T) - g_x\left(x'_T\right)\right)\Lambda_T\right]$$
$$\leq \mathbb{E}\left[\int_s^T \left|\bar{f}(t)\right|^p dt\right]^{\frac{1}{p}} \mathbb{E}\left[\int_s^T |\Lambda_t|^q dt\right]^{\frac{1}{q}} + \mathbb{E}\left[\left|g_x(x_T) - g_x\left(x'_T\right)\right|^p\right]^{\frac{1}{p}} \mathbb{E}\left[|\Lambda_T|^q\right]^{\frac{1}{q}}$$
$$\leq C\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^p + \left|\overline{\phi}_t\right|^p + \int_E \left|\overline{\gamma}_t(e)\right|^p \nu(de)\right) dt\right]^{\frac{1}{q}} \times \left(\mathbb{E}\left[\int_s^T \left|\bar{f}(t)\right|^p dt\right]^{\frac{1}{p}} + \mathbb{E}\left[\left|g_x(x_T) - g_x\left(x'_T\right)\right|^p\right]^{\frac{1}{p}}\right),$$

according to (3.11), we get

$$\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^p + \left|\overline{\phi}_t\right|^p + \int_E \left|\overline{\gamma}_t(e)\right|^p \nu(de)\right) dt\right]$$
$$= \mathbb{E}\left[\int_s^T \left(\overline{\psi}_t \left|\overline{\psi}_t\right|^{p-1} \operatorname{sgn}\left(\overline{\psi}_t\right) dt + \overline{\phi}_t \left|\overline{\phi}_t\right|^{p-1} \operatorname{sgn}\left(\overline{\phi}_t\right) + \int_E \overline{\gamma}_t(e)\left|\overline{\gamma}_t(e)\right|^{p-1} \operatorname{sgn}\left(\overline{\gamma}_t(e)\right)\nu(de)\right) dt\right]$$
$$\leq C\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^p + \left|\overline{\phi}_t\right|^p + \int_E \left|\overline{\gamma}_t(e)\right|^p \nu(de)\right) dt\right]^{\frac{1}{q}} \times \left(\mathbb{E}\left[\int_s^T \left|\bar{f}(t)\right|^p dt\right]^{\frac{1}{p}} + \mathbb{E}\left[\left|g_x(x_T) - g_x\left(x'_T\right)\right|^p\right]^{\frac{1}{p}}\right).$$

Since $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\mathbb{E}\left[\int_s^T \left(\left|\overline{\psi}_t\right|^p + \left|\overline{\phi}_t\right|^p + \int_E \left|\overline{\gamma}_t(e)\right|^p \nu(de)\right) dt\right] \leq C\left(\mathbb{E}\left[\int_s^T \left|\bar{f}(t)\right|^p dt\right] + \mathbb{E}\left[\left|g_x(x_T) - g_x\left(x'_T\right)\right|^p\right]\right).$$

To derive inequality (3.4), it is sufficient to prove the following two assertions

$$\mathbb{E}\left[\int_s^T \left|\bar{f}(t)\right|^p dt\right] \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}. \tag{3.12}$$

$$\mathbb{E}\left[\left|g_x(x_T) - g_x\left(x'_T\right)\right|^p\right] \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}. \tag{3.13}$$

Let us prove the second inequality. g_x being Lipschitz with respect to the state variable combined with the fact that $\frac{\alpha p}{2} < 1 - \frac{p}{2} < 1$, and Lemma 3.2, lead to (3.13).

Next, by applying the Schwarz inequality, we can estimate

$$\mathbb{E}\left[\int_s^T \left|\int_E \left(r_x(t, x_t, u_t, e)^T - r_x\left(t, x'_t, u'_t, e\right)^T\right) \times \gamma'_t(e)\,\nu(de)\right|^p dt\right] \leq B_1 + B_2$$

where

$$B_1 = \mathbb{E}\left[\int_s^T \left|\int_E \left(r_x(t, x_t, u_t, e)^T - r_x\left(t, x_t, u'_t, e\right)^T\right) \times \gamma'_t(e)\,\nu(de)\right|^p 1_{\{u_t \neq u'_t\}}(t)\,dt\right],$$

$$B_2 = \mathbb{E}\left[\int_s^T \left(\sup_{e\in E}\left|r_x\left(t, x_t, u'_t, e\right)^T - r_x\left(t, x'_t, u'_t, e\right)^T\right|\right)^p \times \left|\int_E \gamma'_t(e)\,\nu(de)\right|^p dt\right].$$

Noting that $1 - \frac{p}{2} > \frac{\alpha p}{2}$, and $d\left(u, u'\right) \leq 1$, then by the fact that r_x is bounded by the Lipschitz constant together with (2.12), we get

$$B_1 \leq C\mathbb{E}\left[\int_s^T \int_E \left|\gamma'_t(e)\right|^2 \nu(de)\,dt\right]^{\frac{p}{2}} d\left(u, u'\right)^{1-\frac{p}{2}} \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}.$$

From the Lipschitz condition on the coefficients, and $\frac{\alpha p}{2-p} < 1$, we conclude from Lemma 3.2 and estimate (2.12) that

$$B_2 \leq C\mathbb{E}\left[\int_s^T \left|\int_E \gamma'_t(e)\,\nu(de)\right|^2 dt\right]^{\frac{p}{2}} \times \mathbb{E}\left[\int_s^T \left|x_t - x'_t\right|^{\frac{2p}{2-p}} dt\right]^{1-\frac{p}{2}}$$
$$\leq C\mathbb{E}\left[\int_s^T \int_E \left|\gamma'_t(e)\right|^2 \nu(de)\,dt\right]^{\frac{p}{2}} \left(d\left(u, u'\right)^{\frac{\alpha p}{2-p}}\right)^{1-\frac{p}{2}} \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}.$$

So that

$$\mathbb{E}\left[\int_s^T \left|\int_E \left(r_x(t, x_t, u_t, e)^T - r_x\left(t, x'_t, u'_t, e\right)^T\right) \times \gamma'_t(e)\,\nu(de)\right|^p dt\right] \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}. \tag{3.14}$$

A similar argument shows that

$$\mathbb{E}\left[\int_s^T \left|\left(b_x(t, x_t, u_t)^T - b_x\left(t, x'_t, u'_t\right)^T\right)\psi'_t\right|^p dt\right] \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}, \tag{3.15}$$

$$\mathbb{E}\left[\int_s^T \left|\left(\sigma_x(t, x_t, u_t)^{\mathrm{T}} - \sigma_x\left(t, x'_t, u'_t\right)^{\mathrm{T}}\right)\phi'_t\right|^p dt\right] \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}, \tag{3.16}$$

$$\mathbb{E}\left[\int_s^T \left|f_x(t, x_t, u_t) - f_x\left(t, x'_t, u'_t\right)\right|^p dt\right] \leq Cd\left(u, u'\right)^{\frac{\alpha p}{2}}. \tag{3.17}$$

From (3.14)–(3.17), it is easy to see that $\mathbb{E}\left[\int_s^T \left|\bar{f}(t)\right|^p dt\right] \leq Cd\left(u,u'\right)^{\frac{\alpha p}{2}}$. □

Proof of Theorem 3.1. By using Ekeland's variational principle (Lemma A.1 in the Appendix), with $\lambda = \varepsilon^{\frac{2}{3}}$, there exists an admissible pair $(\widetilde{x}^\varepsilon, \widetilde{u}^\varepsilon)$ such that

$$d(u^\varepsilon,\widetilde{u}^\varepsilon) \leq \varepsilon^{\frac{2}{3}}, \quad \text{and}$$
$$\widetilde{J}(s,y;\widetilde{u}^\varepsilon) \leq \widetilde{J}(s,y;u), \quad \text{for any } u \in U, \tag{3.18}$$

where $\widetilde{J}(s,y;u) = J(s,y;u) + \varepsilon^{\frac{1}{3}} d(u,\widetilde{u}^\varepsilon)$. This means that $(\widetilde{x}^\varepsilon, \widetilde{u}^\varepsilon)$ is an optimal pair for system (2.1) with the new cost function $\widetilde{J}$.

Next, we use the spike variation technique to derive a maximum principle for $(\widetilde{x}^\varepsilon, \widetilde{u}^\varepsilon)$. We take any Borel measurable set $I^\theta \subset [s,T]$, with $\lambda\left(I^\theta\right) = \theta$ for any $\theta > 0$, where $\lambda\left(I^\theta\right)$ denote the Lebesgue measure of the set I^θ. Let $u \in U$ be fixed and define

$$\widetilde{u}_t^{\varepsilon,\theta} = \begin{cases} \widetilde{u}_t^\varepsilon & \text{if } t \in [s,T] \setminus I^\theta, \\ u & \text{if } t \in I^\theta. \end{cases} \tag{3.19}$$

The fact that $\widetilde{J}(s,y;\widetilde{u}^\varepsilon) \leq \widetilde{J}\left(s,y;\widetilde{u}^{\varepsilon,\theta}\right)$ and $d\left(\widetilde{u}^\varepsilon,\widetilde{u}^{\varepsilon,\theta}\right) \leq \theta$, imply that $-\varepsilon^{\frac{1}{3}}\theta \leq \widetilde{J}\left(s,y;\widetilde{u}^{\varepsilon,\theta}\right) - \widetilde{J}(s,y;\widetilde{u}^\varepsilon)$. Arguing as in [7] step 3. (page 1467), we obtain

$$\begin{aligned} -\varepsilon^{\frac{1}{3}} \leq \mathbb{E}\Bigg[& \widetilde{\psi}_t^\varepsilon \cdot \left(b\left(t,\widetilde{x}_t^\varepsilon,u\right) - b\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right) \\ & + \widetilde{\phi}_t^\varepsilon \cdot \left(\sigma\left(t,\widetilde{x}_t^\varepsilon,u\right) - \sigma\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right) \\ & + \int_E \widetilde{\gamma}_t^\varepsilon(e) \cdot \left(r\left(t,\widetilde{x}_t^\varepsilon,u,e\right) - r\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)\right)\nu(de) \\ & + \left(f\left(t,\widetilde{x}_t^\varepsilon,u\right) - f\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right) \\ & + \frac{1}{2}\left(\sigma\left(t,\widetilde{x}_t^\varepsilon,u\right) - \sigma\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right)^T \widetilde{\Psi}_t^\varepsilon \left(\sigma\left(t,\widetilde{x}_t^\varepsilon,u\right)\right. \\ & \left. - \sigma\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right) + \frac{1}{2}\int_E \left(r\left(t,\widetilde{x}_t^\varepsilon,u,e\right)\right. \\ & \left. - r\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)\right)^T \left(\widetilde{\Psi}_t^\varepsilon + \widetilde{\gamma}_t^\varepsilon(e)\right)\left(r\left(t,\widetilde{x}_t^\varepsilon,u,e\right)\right. \\ & \left. - r\left(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)\right)\nu(de)\Bigg]. \end{aligned} \tag{3.20}$$

Now, we are going to derive an estimate for the term similar to the left side of (3.20) with all the $\widetilde{x}_t^\varepsilon, \widetilde{u}_t^\varepsilon$, etc. replaced by $x_t^\varepsilon, u_t^\varepsilon$, etc. To this end, we first estimate the following difference

$$\begin{aligned} \mathbb{E}\Bigg[\int_s^T \int_E & \left\{\widetilde{\gamma}_t^\varepsilon(e) \cdot (r(t,\widetilde{x}_t^\varepsilon,u,e) - r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e))\right. \\ & \left. - \gamma_t^\varepsilon(e) \cdot \left(r\left(t,x_t^\varepsilon,u,e\right) - r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right)\right\}\nu(de)\,dt\Bigg] \\ \leq I_1 + I_2 - I_3, \end{aligned}$$

where

$$I_1 = \mathbb{E}\left[\int_s^T \int_E \left(\widetilde{\gamma}_t^\varepsilon(e) - \gamma_t^\varepsilon(e)\right) \cdot (r(t,\widetilde{x}_t^\varepsilon,u,e) - r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e))\nu(de)\,dt\right],$$

$$I_2 = \mathbb{E}\left[\int_s^T \int_E \gamma_t^\varepsilon(e) \cdot (r(t,\widetilde{x}_t^\varepsilon,u,e) - r(t,x_t^\varepsilon,u,e))\nu(de)\,dt\right],$$

$$I_3 = \mathbb{E}\left[\int_s^T \int_E \gamma_t^\varepsilon(e) \cdot (r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e) - r(t,x_t^\varepsilon,u_t^\varepsilon,e))\nu(de)\,dt\right].$$

For any $\delta \in \left[0,\frac{1}{3}\right)$, let $\alpha = 3\delta \in [0,1)$ and $p \in (1,2)$ so that $(1+\alpha)p < 2$. Take $q \in (2,+\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. It holds, by using Lemma 3.3, properties (2.3) and (2.5) that

$$\begin{aligned} I_1 & \leq \mathbb{E}\left[\int_s^T \int_E \left|\widetilde{\gamma}_t^\varepsilon(e) - \gamma_t^\varepsilon(e)\right|^p \nu(de)\,dt\right]^{\frac{1}{p}} \\ & \quad \times \mathbb{E}\left[\int_s^T \left(\sup_{e\in E}\left|r(t,\widetilde{x}_t^\varepsilon,u,e) - r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e)\right|\right)^q dt\right]^{\frac{1}{q}} \\ & \quad \times \nu(E)^{\frac{1}{q}} \\ & \leq \left(Cd\left(u^\varepsilon,\widetilde{u}^\varepsilon\right)^{\frac{\alpha p}{2}}\right)^{\frac{1}{p}}\left(C\mathbb{E}\left[\int_s^T \left(1+\left|\widetilde{x}_t^\varepsilon\right|^q\right)dt\right]\right)^{\frac{1}{q}} \\ & \leq C\varepsilon^{\frac{\alpha}{3}} = C\varepsilon^\delta. \end{aligned}$$

In view of the Lipschitz condition on r, together with the estimates (2.12) and (3.2), we get from the Schwartz inequality

$$\begin{aligned} I_2 & \leq \mathbb{E}\left[\int_s^T \int_E \left|\gamma_t^\varepsilon(e)\right|^2 \nu(de)\,dt\right]^{\frac{1}{2}} \\ & \quad \times \mathbb{E}\left[\int_s^T \left(\sup_{e\in E}\left|r(t,\widetilde{x}_t^\varepsilon,u,e) - r(t,x_t^\varepsilon,u_t,e)\right|\right)^2 dt\right]^{\frac{1}{2}} \\ & \quad \nu(E)^{\frac{1}{2}} \\ & \leq C\mathbb{E}\left[\int_s^T \left|\widetilde{x}_t^\varepsilon - x_t^\varepsilon\right|^2 dt\right]^{\frac{1}{2}} \\ & \leq C\left(d\left(u^\varepsilon,\widetilde{u}^\varepsilon\right)^\alpha\right)^{\frac{1}{2}} \\ & \leq C\left(\varepsilon^{\frac{2\alpha}{3}}\right)^{\frac{1}{2}} = C\varepsilon^\delta. \end{aligned}$$

Further,

$$\begin{aligned} I_3 & = \mathbb{E}\Bigg[\int_s^T \int_E \gamma_t^\varepsilon(e) \cdot \left(r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e)\right. \\ & \quad \left. - r(t,\widetilde{x}_t^\varepsilon,u_t^\varepsilon,e)\right)1_{\{\widetilde{u}^\varepsilon \neq u^\varepsilon\}}(t)\,\nu(de)\,dt\Bigg] \\ & \quad + \mathbb{E}\Bigg[\int_s^T \int_E \gamma_t^\varepsilon(e) \cdot \left(r(t,\widetilde{x}_t^\varepsilon,u_t^\varepsilon,e)\right. \\ & \quad \left. - r(t,x_t^\varepsilon,u_t^\varepsilon,e)\right)\nu(de)\,dt\Bigg] \\ & \leq \mathbb{E}\left[\int_s^T \int_E \left|\gamma_t^\varepsilon(e)\right|^2 \nu(de)\,dt\right]^{\frac{1}{2}} \\ & \quad \times \mathbb{E}\Bigg[\int_s^T \left(\sup_{e\in E}\left|r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e) - r(t,\widetilde{x}_t^\varepsilon,u_t^\varepsilon,e)\right|\right)^2 \\ & \quad \times 1_{\{\widetilde{u}^\varepsilon \neq u^\varepsilon\}}(t)\,dt\Bigg]^{\frac{1}{2}} \nu(E)^{\frac{1}{2}} \\ & \quad + C\mathbb{E}\left[\int_s^T \int_E \left|\gamma_t^\varepsilon(e)\right|^2 \nu(de)\,dt\right]^{\frac{1}{2}} \\ & \quad \times \mathbb{E}\left[\int_s^T \left|\widetilde{x}_t^\varepsilon - x_t^\varepsilon\right|^2 dt\right]^{\frac{1}{2}}. \end{aligned}$$

By the Schwarz inequality, one has

$$\mathbb{E}\left[\int_s^T \left(\sup_{e\in E}\left|r(t,\widetilde{x}_t^\varepsilon,\widetilde{u}_t^\varepsilon,e)-r(t,\widetilde{x}_t^\varepsilon,u_t^\varepsilon,e)\right|\right)^2 \times 1_{\{\widetilde{u}^\varepsilon\neq u^\varepsilon\}}(t)\,dt\right] \le C\mathbb{E}\left[\int_s^T \left(1+\left|\widetilde{x}_t^\varepsilon\right|^4\right)dt\right]^{\frac{1}{2}} d\left(\widetilde{u}^\varepsilon,u^\varepsilon\right)^{\frac{1}{2}}.$$

Thus from the first inequality in (3.18), it yields $|I_3|\le C\varepsilon^\delta$. □

Corollary 3.4. *Under the conditions of Theorem 3.1, it holds that*

$$\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right] \ge \sup_{u\in\mathcal{U}}\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t\right)dt\right]-C\varepsilon^\delta.$$

Proof. In the spike variations technique, we can replace the point $u\in U$ by any control $u\in\mathcal{U}$, and the subsequent argument still goes through. So the inequality in the estimate (3.20) holds with $u\in U$ replaced by $u\in\mathcal{U}$. □

Remark 3.5. If we assume that $\varepsilon=0$, Theorem 3.1 reduces to the maximum principle proved in [7].

4. Sufficient conditions of near optimality

In this section, we focus on proving the sufficient near-maximum principle for a stochastic control problem in the framework described in the last section. Such a maximum principle is also studied by Zhou [21] in the continuous diffusion case. Related earlier results for the exact optimality are [3,6].

We will show that, under certain concavity conditions, the near-maximum condition of the $\mathcal{H}$-function in the integral form is sufficient for near-optimality. We assume that:

(H3) ρ, r are differentiable in u for $\rho=b,\sigma,f$, and there is a constant $C>0$, such that

$$\left|\rho(t,x,u)-\rho\left(t,x,u'\right)\right|+\left|\rho_u(t,x,u)-\rho_u\left(t,x,u'\right)\right| \le C\left|u-u'\right|, \tag{4.1}$$

and

$$\sup_{e\in E}\left\{\left|r(t,x,u,e)-r\left(t,x,u',e\right)\right| + \left|r_u(t,x,u,e)-r_u\left(t,x,u',e\right)\right|\right\}\le C\left|u-u'\right|. \tag{4.2}$$

We can now state and prove the main result of this section.

Theorem 4.1. *Let $(x^\varepsilon,u^\varepsilon)$ be an admissible pair, and $(\psi^\varepsilon,\phi^\varepsilon,\gamma^\varepsilon)$ be the solution to the corresponding BSDE (2.10). Assume that $H\left(t,\cdot,\cdot,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)$ is concave for a.e. $t\in[s,T]$, P-a.s., and $g(\cdot)$ is convex. If for some $\varepsilon>0$*

$$\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t^\varepsilon\right)dt\right] \ge \sup_{u\in\mathcal{U}}\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t\right)dt\right]-\varepsilon, \tag{4.3}$$

then u^ε is a near-optimal control with order $\varepsilon^{\frac{1}{2}}$, i.e.

$$J(u^\varepsilon)\le \inf_{u\in\mathcal{U}}J(u)+C\varepsilon^{\frac{1}{2}}, \tag{4.4}$$

where $C>0$ is a constant independent of ε.

Proof. The key step in the proof is to show that $H_u(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e))$ is very small and estimate it in terms of ε. We first fix an $\varepsilon>0$, and define a new metric $\widehat{d}$ on $U[s,T]$, by setting

$$\widehat{d}\left(u,u'\right)=\mathbb{E}\left[\int_s^T \mathcal{L}^\varepsilon(t)\left|u_t-u_t'\right|dt\right], \tag{4.5}$$

where

$$\mathcal{L}^\varepsilon(t) = 1+\left|\psi_t^\varepsilon\right|+\left|\phi_t^\varepsilon\right|+2\left|\Psi_t^\varepsilon\right|\left(1+\left|x_t^\varepsilon\right|\right) + \left|\int_E \gamma_t^\varepsilon(e)\,\nu(de)\right|\left(2+\left|x_t^\varepsilon\right|\right).$$

Obviously $\widehat{d}$ is a metric, and it is a complete metric as a weighted L^1 norm. A simple computation shows that

$$\left|\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t\right)dt\right]-\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t'\right)dt\right]\right| \le C\widehat{d}\left(u,u'\right).$$

Therefore, $\mathbb{E}\left[\int_s^T \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,\cdot\right)dt\right]$ is continuous on $\mathcal{U}$ with respect to $\widehat{d}$. It follows from (4.3) and Ekeland's principle that, there exists a $\widetilde{u}^\varepsilon\in\mathcal{U}$ such that

$$\widehat{d}\left(\widetilde{u}^\varepsilon,u^\varepsilon\right)\le\varepsilon^{\frac{1}{2}}, \tag{4.6}$$

and

$$E\left[\int_s^T \overline{\mathcal{H}}\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)dt\right]=\max_{u\in U}E\left[\int_s^T \overline{\mathcal{H}}\left(t,x_t^\varepsilon,u\right)dt\right], \tag{4.7}$$

where

$$\overline{\mathcal{H}}(t,x,u)=\mathcal{H}^{x^\varepsilon,u^\varepsilon}(t,x,u)-\varepsilon^{\frac{1}{2}}\mathcal{L}^\varepsilon(t)\left|u-\widetilde{u}_t^\varepsilon\right|. \tag{4.8}$$

The integral form maximum condition (4.7) implies a pointwise maximum condition, namely, for a.e. $t\in[s,T]$ and P-a.s., $\overline{\mathcal{H}}\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)=\max_{u\in U}\overline{\mathcal{H}}\left(t,x_t^\varepsilon,u\right)$. Recall from proposition 2.3.2 in [23] and Definition A.3 in the Appendix, that

$$0\in\partial_u\overline{\mathcal{H}}\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right). \tag{4.9}$$

By (4.8) and the fact that the generalized gradient of the sum of two functions is contained in the sum of the generalized gradients of the two functions, it follows from proposition 2.3.3 in [23] that

$$\begin{aligned}\partial_u\overline{\mathcal{H}}\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\subset{}&\partial_u\mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)+\left[-\varepsilon^{\frac{1}{2}}\mathcal{L}^\varepsilon(t),\varepsilon^{\frac{1}{2}}\mathcal{L}^\varepsilon(t)\right]\\&+\sigma_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)^T\Psi_t^\varepsilon\left(\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)-\sigma\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right)\\&+\int_E r_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)^T\left(\Psi_t^\varepsilon+\gamma_t^\varepsilon(e)\right)\left(r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right.\\&\left.-r\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)\right)\nu(de).\end{aligned}$$

Since the Hamiltonian H is differentiable in u, we deduce from the inclusion (4.9) that, there is

$$K^\varepsilon(t)\in\left[-\varepsilon^{\frac{1}{2}}\mathcal{L}^\varepsilon(t),\varepsilon^{\frac{1}{2}}\mathcal{L}^\varepsilon(t)\right],$$

such that

$$\begin{aligned}H_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)={}&-K^\varepsilon(t)-\sigma_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)^T\\&\times\Psi_t^\varepsilon\left(\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)-\sigma\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right)\\&-\int_E r_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)^T\left(\Psi_t^\varepsilon+\gamma_t^\varepsilon(e)\right)\Big(r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\\&-r\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)\Big)\nu(de).\end{aligned} \tag{4.10}$$

Therefore, by assumptions (4.1) and (4.2), we get

$$\begin{aligned}&\left|H_u\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\right|\\ &\quad\le\left|H_u\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\right.\\ &\qquad\left.-\,H_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\right|\\ &\qquad+|K^\varepsilon(t)|+\left|\sigma_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)^T\right.\\ &\qquad\left.\times\,\Psi^\varepsilon(t)\left(\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)-\sigma\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon\right)\right)\right|\\ &\qquad+\left|\int_E r_u\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)^T\left(\Psi_t^\varepsilon+\gamma_t^\varepsilon(e)\right)\left(r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right.\right.\\ &\qquad\left.\left.-\,r\left(t,x_t^\varepsilon,\widetilde{u}_t^\varepsilon,e\right)\right)\nu(de)\right|\\ &\quad\le C\mathcal{L}^\varepsilon(t)\left|u_t^\varepsilon-\widetilde{u}_t^\varepsilon\right|+\varepsilon^{\frac12}\mathcal{L}^\varepsilon(t).\end{aligned}\tag{4.11}$$

By the concavity of $H\left(t,\cdot,\cdot,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)$, we have

$$\begin{aligned}&H\left(t,x_t,u_t,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)-H\left(t,x_t^\varepsilon,u^\varepsilon(t),\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\\ &\quad\le H_x\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\left(x_t-x_t^\varepsilon\right)\\ &\qquad+H_u\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\left(u_t-u_t^\varepsilon\right),\end{aligned}$$

for any admissible pair (x,u). Integrating this inequality with respect to t and taking expectations we obtain from (4.5), (4.6) and (4.11)

$$\begin{aligned}&\mathbb{E}\left[\int_s^T\left(H\left(t,x_t,u_t,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\right.\right.\\ &\qquad\left.\left.-\,H\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\right)dt\right]\\ &\quad\le\mathbb{E}\left[\int_s^T H_x\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\left(x_t-x_t^\varepsilon\right)dt\right]\\ &\qquad+C\varepsilon^{\frac12}.\end{aligned}\tag{4.12}$$

On the other hand, by the convexity of g, it yields

$$\begin{aligned}\mathbb{E}\left[g(x_T)-g\left(x_T^\varepsilon\right)\right]&\ge\mathbb{E}\left[g_x\left(x_T^\varepsilon\right)\left(x_T-x_T^\varepsilon\right)\right]\\ &=\mathbb{E}\left[\psi_T^\varepsilon\left(x_T-x_T^\varepsilon\right)\right].\end{aligned}\tag{4.13}$$

Thus it follows by the Ito formula applied to $\psi_T^\varepsilon\left(x_T-x_T^\varepsilon\right)$, together with (4.12) and (4.13)

$$\begin{aligned}&\mathbb{E}\left[\psi_T^\varepsilon\left(x_T-x_T^\varepsilon\right)\right]=\mathbb{E}\left[\int_s^T\left(H_x\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\left(x_t-x_t^\varepsilon\right)\right.\right.\\ &\qquad+\psi_t^\varepsilon.\left(b\left(t,x_t,u_t\right)-b\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right)\\ &\qquad+\phi_t^\varepsilon.\left(\sigma\left(t,x_t,u_t\right)-\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right)\\ &\qquad\left.\left.+\int_E\gamma_t^\varepsilon(e)\cdot\left(r\left(t,x_t,u_t,e\right)-r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right)\nu(de)\right)dt\right]\\ &\quad\ge\mathbb{E}\left[\int_s^T\left(H\left(t,x_t,u_t,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\right.\right.\\ &\qquad-\,H\left(t,x_t^\varepsilon,u_t^\varepsilon,\psi_t^\varepsilon,\phi_t^\varepsilon,\gamma_t^\varepsilon(e)\right)\\ &\qquad+\psi_t^\varepsilon\cdot\left(b\left(t,x_t,u_t\right)-b\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right)\\ &\qquad+\phi_t^\varepsilon\cdot\left(\sigma\left(t,x_t,u_t\right)-\sigma\left(t,x_t^\varepsilon,u_t^\varepsilon\right)\right)\\ &\qquad+\int_E\gamma_t^\varepsilon(e)\cdot\left(r\left(t,x_t,u_t,e\right)\right.\\ &\qquad\left.\left.\left.-\,r\left(t,x_t^\varepsilon,u_t^\varepsilon,e\right)\right)\nu(de)\right)dt\right]-C\varepsilon^{\frac12}\\ &\quad=\mathbb{E}\left[\int_s^T\left(f\left(t,x_t^\varepsilon,u_t^\varepsilon\right)-f\left(t,x_t,u_t\right)\right)dt\right]-C\varepsilon^{\frac12}.\end{aligned}$$

This shows that $J(u)\ge J(u^\varepsilon)-C\varepsilon^{\frac12}$. □

Corollary 4.2. *Under the assumptions of Theorem 4.1, a sufficient condition for an admissible pair $(x^\varepsilon,u^\varepsilon)$ to be ε-optimal is*

$$\begin{aligned}&\mathbb{E}\left[\int_s^T\mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t^\varepsilon\right)dt\right]\\ &\quad\ge\sup_{u\in\mathcal{U}}\mathbb{E}\left[\int_s^T\mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t\right)dt\right]-\left(\frac{\varepsilon}{C}\right)^2.\end{aligned}\tag{4.14}$$

Remark 4.3. If we assume that $\varepsilon=0$, Theorem 4.1 reduces to the maximum principle proved in [6].

5. Examples

Now, two examples are given to illustrate applications of the general results obtained. We suppose that B_t, $t\in[0,1]$, is a standard Brownian motion and $\widetilde{N}(dt,de)$ for $(t,e)\in[0,1]\times\mathbb{R}_0$, where $\mathbb{R}_0=\mathbb{R}\setminus\{0\}$, is a compensated Poisson random measure. Recall that $\mathbb{E}\left[\widetilde{N}(dt,de)^2\right]=\nu(de)\,dt$ is a finite Borel measure satisfying $\int_{\mathbb{R}_0}e^2\nu(de)<\infty$. First of all, let us consider the following particular case of the linear quadratic model.

Example 5.1. Assume that the dynamics and the cost functional are given by

$$\begin{cases}dx_t=u_t dB_t+u_t\displaystyle\int_{\mathbb{R}_0}e\widetilde{N}(dt,de),\\ x_0=0,\end{cases}\tag{5.1}$$

$$J(u)=\mathbb{E}\left[\int_0^1-u_t dt+\frac12 x_1^2\right],\quad\text{for }u\in[0,1].\tag{5.2}$$

Let $\varepsilon>0$ and $(x^\varepsilon,u^\varepsilon)$ be an admissible pair. Then the corresponding second-order adjoint equation takes the form

$$\begin{cases}d\Psi_t^\varepsilon=\Phi_t^\varepsilon dB_t+\displaystyle\int_{\mathbb{R}_0}\Gamma_t^\varepsilon(e)\,\widetilde{N}(dt,de),\\ \Psi_1^\varepsilon=1.\end{cases}\tag{5.3}$$

Eq. (5.3) has a unique solution $\left(\Psi_t^\varepsilon,\Phi_t^\varepsilon,\Gamma_t^\varepsilon(e)\right)=(1,0,0)$. Then for any admissible control u, the $\mathcal{H}$-function reduces to

$$\begin{aligned}\mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t\right)&=u_t\left(1-\phi_t^\varepsilon-\int_{\mathbb{R}_0}e\gamma_t^\varepsilon(e)\,\nu(de)\right)\\ &\quad+\left(u_tu_t^\varepsilon-\frac12u_t^2\right)\left(1+\int_{\mathbb{R}_0}e^2\nu(de)\right).\end{aligned}\tag{5.4}$$

It is clear that the supremum is attained at u_t satisfying

$$u_t-u_t^\varepsilon=\frac{\left(1-\phi_t^\varepsilon-\int_{\mathbb{R}_0}e\gamma_t^\varepsilon(e)\,\nu(de)\right)}{\left(1+\int_{\mathbb{R}_0}e^2\nu(de)\right)},\tag{5.5}$$

if the control u_t in (5.5) is admissible, i.e.

$$u_t^\varepsilon+\frac{\left(1-\phi_t^\varepsilon-\int_{\mathbb{R}_0}e\gamma_t^\varepsilon(e)\,\nu(de)\right)}{\left(1+\int_{\mathbb{R}_0}e^2\nu(de)\right)}\in[0,1].\tag{5.6}$$

Therefore, (4.3) gets the form

$$\sup_{u\in\mathcal{U}} \mathbb{E}\left[\int_0^1 \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t\right)dt\right] - \mathbb{E}\left[\int_0^1 \mathcal{H}^{x^\varepsilon,u^\varepsilon}\left(t,x_t^\varepsilon,u_t^\varepsilon\right)dt\right] = \frac{1}{2}\mathbb{E}\left[\int_0^1 \frac{\left(1-\phi_t^\varepsilon-\int_{\mathbb{R}_0} e\gamma_t^\varepsilon(e)\,\nu(de)\right)^2}{\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)}dt\right] \le C\varepsilon^\delta. \tag{5.7}$$

For example, the controls $u_t^\varepsilon = \left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}$, are candidates for near-optimality. To see this, note that the corresponding first-order adjoint equation is

$$\begin{cases} d\psi_t^\varepsilon = \phi_t^\varepsilon dB_t + \int_{\mathbb{R}_0} \gamma_t^\varepsilon(e)\,\widetilde{N}(dt,de), \\ \psi_1^\varepsilon = x_1^\varepsilon. \end{cases} \tag{5.8}$$

It is clear that, if we choose $u_t^\varepsilon = \left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}$ with the corresponding solution of the state process (5.1)

$$x_t^\varepsilon = \left(\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}\right)\left(B_t + \int_{\mathbb{R}_0} e\widetilde{N}(dt,de)\right),$$

then, the unique solution of the first-order adjoint equation (5.8) is explicitly given by

$$\begin{cases} \psi_t^\varepsilon = \left(\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}\right)\left(B_t + \int_{\mathbb{R}_0} e\widetilde{N}(dt,de)\right), \\ \phi_t^\varepsilon = \left(\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}\right), \\ \gamma_t^\varepsilon(e) = \left(\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}\right)e. \end{cases}$$

Hence, (5.6) and (5.7) are satisfied.

Moreover, the Hamiltonian $H(t,x,u,p,q,\gamma(e)) = u - qu - u\int_{\mathbb{R}_0} e\gamma(e)\,\nu(de)$ is concave in (x,u), and the final cost is convex. This shows that $u_t^\varepsilon = \left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1} - \varepsilon^{\frac{1}{2}}$ satisfies the optimality necessary and sufficient conditions of Theorem 4.1. Then u_t^ε is nearly optimal with order $\varepsilon^{\frac{1}{2}}$.

Example 5.2. Assume that we have a family of stochastic control problems parameterised by $\varepsilon > 0$, where ε may be a parameter representing the complexity of the cost functional

$$J^\varepsilon(u) = \mathbb{E}\left[\int_0^1 \varepsilon g(u_t)\,dt + \frac{1}{2}x_1^2\right]. \tag{5.9}$$

Subject to the controlled jump diffusion state process on $\mathbb{R}$

$$\begin{cases} dx_t = u_t dt + u_t dB_t + u_t\int_{\mathbb{R}_0} e\widetilde{N}(dt,de), \\ x_0 = a, \end{cases} \tag{5.10}$$

where g (independent of ε) is a nonlinear, convex function, satisfying 4.1. Explicit solution of problem (5.10), (5.9) may be a difficult problem. The idea is to approximate the cost functional in order to neglect the nonlinearity. Then by setting $\varepsilon = 0$ in (5.9), it yields

$$J^0(u) = \frac{1}{2}\mathbb{E}\left[x_1^2\right]. \tag{5.11}$$

First, consider the optimal control problem where the state is described by Eq. (5.10) with a new cost function (5.11). In a second step, we solve problem (5.10), (5.11), and obtain an optimal solution explicitly by applying the stochastic maximum principle [6,7]. Finally, we solve the nonlinear control problem (5.10), (5.11) near optimally.

By a standard argument, problem (5.10), (5.11) can be solved directly. Indeed, by applying Ito's formula to the process $\exp\left(\frac{t-1}{1+\int_{\mathbb{R}_0} e^2\nu(de)}\right)\cdot x_t^2$, we get

$$\mathbb{E}\left[x_1^2\right] = \mathbb{E}\left[a^2\exp\left(-\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1}\right)\right] + \mathbb{E}\left[\int_0^1 e^{(t-1)\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right)^{-1}}\left(\frac{x_t}{\sqrt{1+\int_{\mathbb{R}_0} e^2\nu(de)}} + u_t\sqrt{1+\int_{\mathbb{R}_0} e^2\nu(de)}\right)^2 dt\right].$$

We conclude that $u_t^* = -\frac{x_t^*}{1+\int_{\mathbb{R}_0} e^2\nu(de)}$ is the optimal control in feedback form for problem (5.10), (5.11), where x_t^* denotes the optimal state solution to (5.10) under the control u_t^*. We denote the corresponding solution to the first and second order adjoint equations by (ψ,ϕ,γ) and (Ψ,Φ,Γ), respectively. By uniqueness of the solution of BSDEs it is easy to show that $(\Psi_t,\Phi_t,\Gamma_t(e)) = (1,0,0)$ is the only adapted solution to

$$\begin{cases} d\Psi_t = \Phi_t dB_t + \int_{\mathbb{R}_0} \Gamma_t(e)\,\widetilde{N}(dt,de), \\ \Psi_1 = 1. \end{cases} \tag{5.12}$$

Then the $\mathcal{H}$-function for problem (5.10), (5.11) is

$$\mathcal{H}^{x^*,u^*}(t,x,u) = -u\left(\psi_t + \phi_t + \int_{\mathbb{R}_0} \gamma_t(e)\,e\nu(de)\right) + \left(uu_t^* - \frac{1}{2}u^2\right)\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right). \tag{5.13}$$

Since u^* is optimal, by using the stochastic maximum principle, we conclude that, it is necessary that u^* maximizes the $\mathcal{H}$-function a.s., namely

$$-\left(\psi_t + \phi_t + \int_{\mathbb{R}_0} \gamma_t(e)\,e\nu(de)\right) = 0, \quad P\text{-}a.s,\ a.e.t. \tag{5.14}$$

However, the $\mathcal{H}_\varepsilon$-function for problem (5.10) and (5.9) gets the form

$$\mathcal{H}_\varepsilon^{x^*,u^*}(t,x,u) = -u\left(\psi_t + \phi_t + \int_{\mathbb{R}_0} \gamma_t(e)\,e\nu(de)\right) + \left(uu_t^* - \frac{1}{2}u^2\right)\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right) - \varepsilon g(u), \tag{5.15}$$

the above function is maximal at u^ε, which satisfies $u^\varepsilon = u^*(t) - \frac{\varepsilon g_u(u^\varepsilon)}{1+\int_{\mathbb{R}_0} e^2\nu(de)}$, then this gives

$$\begin{aligned} &\max_{u\in U} \mathcal{H}_\varepsilon^{x^*,u^*}(t,x,u) - \mathcal{H}_\varepsilon^{x^*,u^*}\left(t,x,u_t^*\right) \\ &= \mathcal{H}_\varepsilon^{x^*,u^*}(t,x,u^\varepsilon) - \mathcal{H}_\varepsilon^{x^*,u^*}\left(t,x,u_t^*\right), \\ &= -\frac{1}{2}\left(u^\varepsilon - u_t^*\right)^2\left(1+\int_{\mathbb{R}_0} e^2\nu(de)\right) - \varepsilon\left(g(u^\varepsilon) - g(u^*)\right), \end{aligned}$$

$$= \frac{1}{2}\varepsilon^2 \left|g_u\left(u^\varepsilon\right)\right|^2 + C\varepsilon^2 \left|g_u\left(u^\varepsilon\right)\right|^2$$
$$\leq C\varepsilon^2.$$

Moreover, the Hamiltonian for problem (5.10), (5.9) is

$$H\left(t, x, u, \psi_t, \phi_t, \gamma_t\left(e\right)\right)$$
$$= -u\left(\psi_t + \phi_t + \int_{\mathbb{R}_0} e\gamma_t\left(e\right)\nu\left(de\right)\right) - \varepsilon g\left(u\right),$$
$$= -\varepsilon g\left(u\right),$$

which is concave. By Theorem 4.1, this proves that, the control u^* is indeed a near-optimal for problem (5.10), (5.9), with an error order of ε.

Appendix

Lemma A.1 (*Ekeland's Principle [14]*). *Let (S, d) be a complete metric space and $h : S \to \mathbb{R}$ be lower-semicontinuous and bounded from below. For $\varepsilon \geq 0$, suppose $u^\varepsilon \in S$ satisfies $h\left(u^\varepsilon\right) \leq \inf_{u\in S} h\left(u\right) + \varepsilon$. Then for any $\lambda > 0$, there exists $u^\lambda \in S$ such that*

$$\rho\left(u^\lambda\right) \leq \rho\left(u^\varepsilon\right),$$
$$d\left(u^\lambda, u^\varepsilon\right) \leq \lambda,$$
$$\rho\left(u^\lambda\right) \leq \rho\left(u\right) + \frac{\varepsilon}{\lambda} d\left(u, u^\lambda\right), \quad \textit{for all } u \in S.$$

The following proposition is proved in the Appendix of [13].

Proposition A.2. *Let h be a $\mathcal{P} \times \mathcal{B}\left(E\right)$-measurable function such that*

$$\mathbb{E}\left[\int_0^T \int_E \left|h\left(s, e\right)\right|^2 \nu\left(de\right) ds\right] < \infty.$$

Then, for all $p \geq 2$, there is a positive constant K depends only on p, T, and $\nu\left(E\right)$ such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \int_E h\left(s, e\right)\widetilde{N}\left(ds, de\right)\right|^p\right]$$
$$\leq K\mathbb{E}\left[\int_0^T \int_E \left|h\left(s, e\right)\right|^p \nu\left(de\right) ds\right]. \tag{A.1}$$

Note that $\mathcal{P}$ defined as the σ-algebra of $\mathcal{F}$-predictable subsets of $\Omega \times [0, T]$.

Finally, we introduce Clarke's generalized gradient.

Definition A.3 (*Clarke [23]*). Let Q be a convex set in $\mathbb{R}^d$ and let $h : Q \to \mathbb{R}$ be a locally Lipschitz function. The generalized gradient of h at $\hat{x} \in Q$, denoted by $\partial_x h$, is a set defined by

$$\partial_x h\left(\hat{x}\right) = \left\{p \in \mathbb{R}^d / p\cdot\xi \leq \limsup_{x\to\hat{x},\theta\to 0+} \frac{h\left(x+\theta\xi\right) - h\left(x\right)}{\theta};\right.$$
$$\left.\text{for any } \xi \in \mathbb{R}^d, \text{ and } x, x+\theta\xi \in Q\right\}.$$

References

[1] H. Pham, Optimal stopping of controlled jump diffusion processes: a viscosity solution approach, J. Math. Syst. Estimat. Control 8 (1998) 1–27.
[2] B. øksendal, A. Sulem, Applied Stochastic Control of Jump Diffusions, second ed., Springer, 2007.
[3] F. Baghery, B. øksendal, A maximum principle for stochastic control with partial information, Stoch. Anal. Appl. 25 (2007) 705–717.
[4] R.K. Boel, P. Varaiya, Optimal control of jump processes, SIAM J. Control Optim. 15 (1977) 92–119.
[5] M.H.A. Davis, R. Elliott, Optimal control of a jump process, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 40 (1977) 183–202.
[6] N.C. Framstad, B. øksendal, A. Sulem, Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance, J. Optim. Theory Appl. 121 (2004) 77–98; J. Optim. Theory Appl. 124 (2005) 511–512 (erratum).
[7] S. Tang, X. Li, Necessary conditions for optimal control for stochastic systems with random jumps, SIAM J. Control Optim. 32 (1994) 1447–1475.
[8] S. Peng, A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim. 28 (1990) 966–979.
[9] A. Eyraud-Loisel, Backward stochastic differential equations with enlarged filtration: option hedging of an insider trader in a financial market with jumps, Stochastic Process. Appl. 115 (2005) 1745–1763.
[10] D. Becherer, Bounded solutions to backward SDE's with jumps for utility optimization and indifference hedging, Ann. Appl. Probab. 16 (4) (2006) 2027–2054.
[11] G. Barles, R. Buckdahn, E. Pardoux, BSDE's and integral-partial differential equations, Stoch. Stoch. Rep. 60 (1997) 57–83.
[12] M. Royer, Backward stochastic differential equations with jumps and related non linear expectations, Stochastic Process. Appl. 116 (2006) 1358–1376.
[13] B. Bouchard, R. Elie, Discrete time approximation of decoupled forward–backward SDE with jumps, Stochastic Process. Appl. 118 (1) (2008) 53–75.
[14] I. Ekeland, Non convex minimization problems, Bull. Amer. Math. Soc. N.S. 1 (1979) 443–474.
[15] X.Y. Zhou, Deterministic near-optimal controls, part I: necessary and suffcient conditions for near-optimality, J. Optim. Theory Appl. 85 (1995) 473–488.
[16] B. Mezerdi, Necessary conditions for optimality for a diffusion with a non smoth drift, Stochastics 24 (1988) 305–326.
[17] R.J. Elliott, M. Kohlmann, The variational principle and stochastic optimal control, Stochastics 3 (1980) 229–241.
[18] K. Bahlali, F. Chighoub, B. Djehiche, B. Mezerdi, Optimality necessary conditions in singular stochastic control problems with non smooth data, J. Math. Anal. Appl. 355 (2009) 479–494.
[19] K. Bahlali, B. Djehiche, B. Mezerdi, On the stochastic maximum principle in optimal control of degenerate diffusions with lipschitz coefficients, Appl. Math. Optim. 56 (2007) 364–378.
[20] K. Bahlali, B. Mezerdi, Y. Ouknine, The maximum principle in optimal control of a diffusions with nonsmooth coefficients, Stoch. Stoch. Rep. 57 (1996) 303–316.
[21] X.Y. Zhou, Stochastic near optimal controls: necessary and sufficient conditions for near optimality, SIAM J. Control Optim. 36 (3) (1998) 929–947.
[22] K. Bahlali, N. Khelfallah, B. Mezerdi, Necessary and sufficient conditions for near-optimality in stochastic control of FBSDE's, Syst. Control Lett. 58 (2009) 857–864.
[23] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.