Second-Order Necessary Conditions for Optimal Control with Recursive Utilities

Yuchao Dong¹ · Qingxin Meng²

Received: 2 June 2018 / Accepted: 28 March 2019 / Published online: 5 April 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
The necessary conditions for an optimal control of a stochastic control problem with recursive utilities are investigated. The first-order condition is the well-known Pontryagin-type maximum principle. When such a first-order necessary condition is singular in some sense, certain type of the second-order necessary condition will come in naturally. The aim of this paper is to explore such kind of conditions for our optimal control problem.

Keywords Recursive optimal control · Maximum principle · Variation equation · Adjoint processes

Mathematics Subject Classification 49J53 · 49K45 · 60H99

1 Introduction

In this paper, we consider the second-order necessary condition for optimal control problems, when the cost functional is defined via a backward differential stochastic equation (BSDE for short; see its definition in Sect. 2.2). This extends the classical models for stochastic optimal control problems. In the context of mathematical finance, such functionals are sometimes called recursive utilities. They are used to represent an agent’s aversion to ambiguity (see [1] and references therein for the explanation).
One tool for the study of optimal control problems is the Pontryagin maximum principle, which is the necessary condition for the optimal pair. In many cases, this will be a pointwise first-order condition on the Hamiltonian. We refer to [2] for an early study on the first-order necessary condition for stochastic optimal controls. Many authors contributed on this topic; see [3,4] and references cited therein. Compared to the deterministic setting, new phenomenon and difficulties appear, when the diffusion term of the stochastic control system contains the control variable and the control region is nonconvex. The corresponding first-order necessary condition for this general case was established in [5]. For the recursive stochastic optimal control problem, when the control domain is convex, the local first-order maximum principle was studied in [6,7]. But for the general setting, it remains to be an open problem proposed by Peng [10] in a long time. Wu [11] and Yong [12] established the corresponding first-order maximum principle containing unknown parameters in the formulation for the maximum principle. Recently, different from their methods, Hu [13] completely solved this problem by establishing the variation equation for the BSDEs.

However, sometimes, the first-order condition turns out to be trivial. Either the gradient and the Hessian of the corresponding Hamiltonian with respect to the control variable vanish/degenerate, or the Hamiltonian is equal to a constant in the control region. We refer such control process as a singular control, whose definition will be specified in the next section. In these cases, the first-order necessary condition cannot provide enough information for the theoretical analysis and numerical computing, and therefore, one needs to study the second-order necessary condition. Along the line of necessary conditions for singular optimal control problems, the deterministic case was considered by many authors. The reader is referred to Bell and Jacobson [14], and Krener [17]. Compared to the deterministic control systems, the second-order necessary condition for stochastic optimal controls was first investigated by Tang [18]. In [18], a pointwise second-order maximum principle for stochastic singular optimal controls in the sense of Pontryagin-type maximum principle was established, which involves second-order adjoint processes, for the case that the diffusion term $σ(t, x, u)$ is independent of the control $u$, via a generalized spike variation technique together with the vector-valued measure theory and the second-order expansions of both the system and the cost functional. Recently, this direction has drawn great attention, see [19–22]. In [19], an integral-type second-order necessary condition for stochastic optimal controls was derived under the assumption that the control region $U$ is convex, while in [20], a pointwise second-order necessary condition for stochastic optimal controls was established in the case that both drift and diffusion terms may contain the control variable $u$, and the control region $U$ is still assumed to be convex. The method was further developed in [21] to obtain a pointwise second-order necessary condition in general cases, where the control region is allowed to be nonconvex; see [21] and also [22] for details.

This paper is first to investigate the second-order maximum principle for the recursive optimal control problem. The rest of this paper is organized as follows. In Sect. 2, we introduce the formulation of the optimal control problem and give the main results of this paper. Section 3 includes a quantitative analysis for the variations in the system and the cost functional between two different control actions. Section 4 contains the proof for both the first-order and second-order necessary conditions.
2 Formulation of the Problem and the Main Results

2.1 Notations

We consider a finite time horizon $T$ and a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a $d$-dimensional standard Brownian motion $W := \{W_t : t \in [0, T]\}$. Let $\mathbb{F} := \{\mathcal{F}_t : t \in [0, T]\}$ be a filtration generated by $W$ satisfying the usual conditions of right continuity and $\mathbb{P}$-completeness. We denote by $\mathcal{P}$ the predictable $\sigma$-field on $[0, T] \times \Omega$ and $\mathcal{B}(\Lambda)$ the Borel $\sigma$-algebra of any topological space $\Lambda$. Let $\mathbb{H}$ be an Euclidean space, in which the inner product and the norm are denoted by $\langle \cdot, \cdot \rangle$ and $| \cdot |$, respectively. For a function $f : \mathbb{R}^n \to \mathbb{R}$, we use $\partial f$ to denote its gradient and $\phi_x$ its Hessian (a symmetric matrix). If $\phi : \mathbb{R}^n \to \mathbb{R}^k$, where $k \geq 2$, then $\phi_x = \left(\frac{\partial \phi_j}{\partial x_i}\right)_{i=1,2,...,k;j=1,2,...,n}$ is the corresponding $(k \times n)$-Jacobian matrix.

Finally, in this paper, for the generator $f(t, x, y, z)$ of BSDE, we shall define the second-order derivative $D^2 f$ as

$$D^2 f = \begin{pmatrix} f_{xx}, & f_{xy}, & f_{xz} \\ f_{yx}, & f_{yy}, & f_{yz} \\ f_{zx}, & f_{zy}, & f_{zz} \end{pmatrix}.$$ 

Furthermore, we denote by $A^*$ the transpose of any vector or matrix $A$, and $C$ and $K$ two generic positive constants, which may be different from line to line. In this paper, for simplicity of the notation, the inner product of two vectors $a, b$ will be denoted as $ab$.

Several spaces of random variables and stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$ will be used throughout the paper. For any $\alpha, \beta \in [1, \infty]$, we define

- $L^\beta_{\mathbb{F}}(0, T; \mathbb{H})$: the space of all $\mathbb{H}$-valued and $\mathbb{F}$-adapted processes $f(\cdot) = \{f(t, \omega)| (t, \omega) \in [0, T] \times \Omega\}$ such that $\|f(\cdot)\|_{L^\beta_{\mathbb{F}}(0, T; \mathbb{H})} := \left\{\mathbb{E}\left[\int_0^T |f(t)|^\beta dt\right]\right\}^{\frac{1}{\beta}} < \infty$;

- $S^\beta_{\mathbb{F}}(0, T; \mathbb{H})$: the space of all $\mathbb{H}$-valued, $\mathbb{F}$-adapted, càdlàg processes $f(\cdot) = \{f(t, \omega)| (t, \omega) \in [0, T] \times \Omega\}$ such that $\|f(\cdot)\|_{S^\beta_{\mathbb{F}}(0, T; \mathbb{H})} := \left\{\mathbb{E}\left[\sup_{t\in[0,T]} |f(t)|^\beta\right]\right\}^{\frac{1}{\beta}} < \infty$;

- $L^\beta_{\mathcal{F}}(\Omega; \mathbb{H})$: the space of all $\mathbb{H}$-valued, $\mathcal{F}_T$-measurable random variables $\xi$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\|\xi\|_{L^\beta_{\mathcal{F}}(\Omega; \mathbb{H})} := \left\{\mathbb{E}\left[|\xi|^\beta\right]\right\}^{\frac{1}{\beta}} < \infty$;

- $L^\beta_{\mathcal{F}_{\mathbb{F}}}(\Omega; L^\alpha(0, T; \mathbb{H}))$: the space of all $L^\alpha(0, T; \mathbb{H})$-valued, $\mathcal{F}_T$-adapted processes $f(\cdot) = \{f(t, \omega)| (t, \omega) \in [0, T] \times \Omega\}$ such that $\|f(\cdot)\|_{L^\beta_{\mathcal{F}_{\mathbb{F}}}(\Omega; L^\alpha(0, T; \mathbb{H}))} := \left\{\mathbb{E}\left[\int_0^T |f(t)|^\alpha dt\right]^{\frac{\beta}{\alpha}}\right\}^{\frac{1}{\beta}} < \infty$.

In addition, we write $M^\beta_{\mathcal{F}}[0, T] := S^\beta_{\mathbb{F}}(0, T; \mathbb{R}^n) \times S^\beta_{\mathbb{F}}(0, T; \mathbb{R}) \times L^\beta_{\mathbb{F}}(\Omega; L^2(0, T; \mathbb{R}^d))$. Clearly, $M^\beta_{\mathcal{F}}[0, T]$ is a Banach space. For any triplet of processes $\Theta(\cdot) :=$...
(x(·), y(·), z(·)) in $M^p_{F}[0, T]$, the corresponding norm is defined as

$$
\|\Theta(·)\|_{M^p_{F}[0, T]} := \left\{ \mathbb{E} \left[ \sup_{t \in [0,T]} |x(t)|^p + \sup_{t \in [0,T]} |y(t)|^p + \left( \int_0^T |z(t)|^2 \, dt \right)^{\frac{p}{2}} \right] \right\}^{\frac{1}{p}}.
$$

### 2.2 Model Setup

In this paper, we consider the controlled system satisfying the following stochastic differential equation (SDE for short) driven by Brownian motion $\{W_t, 0 \leq t \leq T\}$,

$$
x(t) = x_0 + \int_0^t b(s, x(s), u(s)) \, ds + \int_0^t \sigma(s, x(s)) \, dW_s.
$$

(1)

The associated cost functional is defined via the following BSDE:

$$
y(t) = h(x(T)) + \int_t^T f(s, x(s), y(s), z(s), u(s)) \, ds - \int_t^T z(s) \, dW_s.
$$

(2)

and given as

$$
J(u(·)) := y(0).
$$

(3)

We also call the solution $(y(·), z(·))$ of (2) the cost process associated with $u(·)$. In the above system, $b : [0, T] \times \mathbb{R}^n \times \bar{U} \to \mathbb{R}^n, \sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}, f : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \times \bar{U} \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}$ are given functions, where $U$ is the control domain that is assumed to be a nonempty subset of $\mathbb{R}^m$ and not necessarily to be convex, and $\bar{U}$ is its closure. An admissible control is defined as follows.

**Definition 2.1** A control process $u(·)$ is said to be admissible, if it is an $U$-valued predictable process and satisfies

$$
||u(·)||_{U_{ad}} \triangleq \left\{ \mathbb{E} \left[ \int_0^T |u(t)|^8 \, dt \right] \right\}^{\frac{1}{8}} < \infty.
$$

Denote by $U_{ad}$ the set of all admissible control processes.

The optimal control problem is to minimize the cost over $U_{ad}$, i.e.,

**Problem 2.1** Find an admissible control $\bar{u}(·) \in U_{ad}$ such that

$$
J(\bar{u}(·)) = \inf_{u(·) \in U_{ad}} J(u(·))
$$

(4)

subject to the state equations (1), (2) and the cost functional (3).

The process $\bar{u}(·)$ is called an optimal control. The state and cost processes associated with $\bar{u}(·)$, denoted by $(\bar{x}(·), \bar{y}(·), \bar{z}(·))$, are called the optimal state and cost processes. When $f$ is independent of $(y, z)$, it is easy to check that $y(0) =$
for the definition). It is called the first-order necessary condition for the optimal control is singular in some sense. More precisely, suppose that there is a condition. In other cases, the first-order condition is insufficient, especially when sufficient to find the optimal control. For example, there is only one control satisfying essentially the Pontryagin’s maximum principle. Sometimes, such a condition is sufficient for (see (12) for the definition). It is called the first-order necessary condition for \( \tilde{u}(\cdot) \), which is essentially the Pontryagin’s maximum principle. Sometimes, such a condition is sufficient to find the optimal control. For example, there is only one control satisfying the condition. In other cases, the first-order condition is insufficient, especially when the optimal control is singular in some sense. More precisely, suppose that there is a set \( \mathcal{U}_0 \subset \mathcal{U}_{ad} \), which is different from the singleton, such that the following holds:

\[
J(u^0(\cdot)) = J(\tilde{u}(\cdot)) + \delta J_1(\tilde{u}(\cdot), u(\cdot)) + o(\delta). \tag{5}
\]

Here \( J_1(\tilde{u}(\cdot), u(\cdot)) \) is some functional of \( u(\cdot) \) and \( \tilde{u}(\cdot) \). The above can be called the first-order Taylor expansion of \( J(\cdot) \) at \( \tilde{u}(\cdot) \), and \( J_1(\tilde{u}(\cdot), u(\cdot)) \) can be regarded as the “directional derivative” of \( J(\cdot) \) at \( \tilde{u}(\cdot) \) in the “direction” \( u(\cdot) \). Hence, the minimality of \( \tilde{u}(\cdot) \) implies

\[
J_1(\tilde{u}(\cdot), u(\cdot)) \geq 0, \quad \forall u(\cdot) \in \mathcal{U}_{ad}. \tag{6}
\]

Such a condition can be transformed into the condition on the Hamiltonian. It turns out to be a second-order necessary optimality condition for Problem 2.1 with recursive utilities. We shall calculate \( J_2 \) and transform the above condition into condition on the Hamiltonian. It turns out to be a second-order condition in some sense.
2.3 Basic Assumptions

In this subsection, we first introduce some basic assumptions on the coefficients of our control problem. Let $K_0$ be some positive constant.

**Assumption 1** The functions $b, \sigma, h, f$ are Borel measurable with respect to their respective arguments, continuous in $u$, continuously differentiable in $(x, y, z)$ for each fixed $(t, u)$, and

\[
|b_x(t, x, u)|, |\sigma_x(t, x)|, |h_x(x)|, |f_x(t, x, y, z, u)|,
\]

\[
|f_z(t, x, y, z, u)| \leq K_0,
\]

\[
|b(t, x, u)| \leq K_0(1 + |x| + |u|), |\sigma(t, x)| \leq K_0(1 + |x|), |h(x)| \leq K_0(1 + |x|),
\]

\[
|f(t, x, y, z, u)| \leq K_0(1 + |x| + |y| + |z| + |u|).
\]

(10)

Moreover, all the derivatives involved above are Borel measurable and are continuous in $x$.

**Assumption 2** The first-order derivatives involved above are continuous in $u$ on $\tilde{U}$. The functions $b, \sigma, f$ and $h$ have continuous second-order derivatives in $x, y, z$. The second-order derivatives are Borel measurable with respect to $(t, x, y, z, u)$ and are bounded by the constant $K_0$, that is,

\[
|b_{xx}(t, x, u)|, |\sigma_{xx}(t, x)|, |h_{xx}(x)| \leq K_0,
\]

\[
|f_{xy}|, |f_{yx}|, |f_{yy}|, |f_{xz}|, |f_{yz}|, |f_{zz}| \leq K_0.
\]

(11)

For each $u(\cdot) \in \mathcal{U}_a$, from Assumption 1, Propositions 2.1 and 3.1 in [23] imply that the SDE (1) and BSDE (2) have unique strong solutions, which will be denoted by $(x(\cdot; u(\cdot)), y(\cdot; u(\cdot)), z(\cdot; u(\cdot)) : M^8_{\mathbb{F}}[0, T] \triangleq \mathcal{S}^8_{\mathbb{F}}(0, T; \mathbb{R}^n) \times \mathcal{S}^8_{\mathbb{F}}(0, T; \mathbb{R}) \times L^8_{\mathbb{F}}(0, T; L^2(0, T; \mathbb{R}^d)))$, or simply $(x(\cdot), y(\cdot), z(\cdot))$, if its dependence on the admissible control $u(\cdot)$ is clear from the context.

For future purposes, we recall the standard estimates of BSDEs (see [13] and the references therein).

**Lemma 2.1** Let $(Y_i, Z_i), i = 1, 2,$ be the solutions of the following BSDEs:

\[
Y_i(t) = \xi_i + \int_t^T f_i(s, Y_i(s), Z_i(s))ds - \int_t^T Z_i(s)dW_s,
\]

where $\mathbb{E}[|\xi_i|^\beta] < \infty$, $f_i = f_i(s, \omega, y, z) : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is progressively measurable, for each fixed $(y, z)$, Lipschitz in $(y, z)$, and $\mathbb{E}[(\int_0^T |f_i(s, 0, 0)|ds)^\beta] < \infty$, for some $\beta > 1$. Then, there exists a constant $C_\beta$ depending on $\beta, T$ and the Lipschitz constant such that

\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |Y_1(t) - Y_2(t)|^\beta + \left( \int_0^T |Z_1(s) - Z_2(s)|^2 ds \right)^{\beta/2} \right] \leq C_\beta \mathbb{E} \left[ |\xi_1 - \xi_2|^\beta + \left( \int_0^T |f_1(s, Y_1(s), Z_1(s)) - f_2(s, Y_1(s), Z_1(s))|ds \right)^{\beta} \right].
\]
In particular, taking \( \xi_1 = 0 \) and \( f_1 = 0 \), we have

\[
E \left[ \sup_{t \in [0,T]} |Y_2(t)|^\beta + \left( \int_0^T |Z_2(s)|^2 \, ds \right)^{\beta/2} \right] 
\leq C_\beta E \left[ |\xi_2|^\beta + \left( \int_0^T |f_2(s, 0, 0)| \, ds \right)^\beta \right].
\]

### 2.4 The Main Results

The object of this paper is to establish a general maximum principle for Problem 2.1. When the convexity assumption is not made on the control domain \( U \), the basic idea of deriving necessary conditions is to apply the spike variation to the control process and derive a Taylor-type expansion for the state process and the cost functional with respect to the spike variation in the control process. Then, using some suitable duality relations, one can obtain a maximum principle of Pontryagin-type.

Define the Hamiltonian:

\[
H(t, x, y, z, u, p, q) := \langle p, b(t, x, u) \rangle + \langle q, \sigma(t, x) \rangle + f(t, x, y, z, u).
\]

Let \( \bar{u}(\cdot) \) be an optimal control and \((\bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot))\) the associated state and cost process. To simplify the notations, we introduce the following abbreviations:

\[
\bar{b}(t) := b(t, \bar{x}(t), \bar{u}(t)), b(t; u) \\
:= b(t, \bar{x}(t), u), \delta b(t; u) := b(t, \bar{x}(t), u) - \bar{b}(t),
\]

and define similarly for \( \bar{b}_x(t), \bar{b}_{xx}(t), \delta b_x(t; u), \bar{\sigma}(t), \bar{\sigma}_x(t), \bar{\sigma}_{xx}(t), \bar{f}(t), \bar{f}_x(t), \bar{f}_{xx}(t), \delta f(t; u) \) and so on. We introduce, respectively, the following two adjoint equations:

\[
dp(t) = -\left[ \int \bar{f}_y(t) + \bar{f}_z(t)\bar{\sigma}_x(t) + \bar{b}_x(t) \right]^* p(t) + \left[ \int \bar{f}_z(t) + \bar{\sigma}_x(t) \right]^* q(t) \\
+ \bar{f}_x(t) \right] \, dt + q(t) \, dW_t,
\]

\[
p(T) = h_x(\bar{x}_T).
\]

and

\[
dP_t = -\left[ \int \bar{f}_y(t) + 2\bar{f}_z(t)\bar{\sigma}_x(t) + 2\bar{b}_x(t) + (\bar{\sigma}_x(t))^2 \right]^* P(t) \\
+ \left[ \int \bar{f}_z(t) + 2\bar{\sigma}_x(t) \right]^* Q(t) \\
+ \bar{b}_{xx}(t) p(t) + \bar{\sigma}_{xx}(t) \left[ \int \bar{f}_z(t) p(t) + q(t) \right] + \left[ 1, p(t), \bar{\sigma}_x(t) \right] P(t) \\
+ q(t) \right] D^2 \bar{f}(t) \left[ 1, p(t), \bar{\sigma}_x(t) p(t) + q(t) \right]^* \, dt + Q(t) \, dW_t,
\]

\[
P(T) = h_{xx}(\bar{x}_T).
\]
Under Assumptions 1 and 2, from Lemma 2.1, it is easy to see that, for any admissible pair \((\tilde{u}(\cdot), \tilde{x}(\cdot), \tilde{y}(\cdot), \tilde{z}(\cdot))\), BSDEs (14) and (15) admit unique solutions \((p(\cdot), q(\cdot)) \in \mathcal{S}_F^d(0, T; \mathbb{R}^n) \times L_F^8(\Omega; L^2(0, T; \mathbb{R}^{n \times d}))\) and \((P(\cdot), Q(\cdot)) \in \mathcal{S}_F^4(0, T; \mathbb{R}^{n \times n}) \times L_F^4(\Omega; L^2(0, T; \mathbb{R}^{n \times d}))\), respectively. We call (14) and (15) the first-order and the second-order adjoint equations of the control system (1)–(2), respectively, where the unique adapted solutions \((p(\cdot), q(\cdot))\) and \((P(\cdot), Q(\cdot))\) are referred as the first-order and the second-order adjoint processes. We also use the abbreviations:

\[
\begin{align*}
\bar{H}(t) &:= H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), p(t), q(t)), \\
\bar{H}_x(t) &:= H_x(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), p(t), q(t)), \\
\bar{H}_{xx}(t) &:= H_{xx}(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), p(t), q(t)), \\
\delta H(t, v) &:= H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), v, p(t), q(t)) \\
&\quad - H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), p(t), q(t)).
\end{align*}
\]

In the following, we state the main results of our paper. The first is the first-order maximum principle. Note that this result is covered by Hu [13], but as an incidental result of the proof of the second-order condition, we will still give a simple proof in Sect. 4.1.

**Theorem 2.1** Let Assumption 1 be satisfied. Let \((\tilde{u}(\cdot); \bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot))\) be an optimal pair. Then there is a subset \(I_0 \subset [0, 1]\) which is of full measure, such that at each \(t \in I_0\) the minimum condition

\[
H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), p(t), q(t)) = \min_{v \in U} H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), v, p(t), q(t)), \text{ a.s.}
\]

holds.

The maximum principle is a powerful tool for the study of optimal stochastic control problems. However, it is not always effective. For example, if the optimal admissible pair \((\tilde{u}(\cdot); \bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot))\) is such that \(h_x(\bar{x}(T)) = 0, f_x(t, \bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot), \bar{u}(\cdot)) = 0,\) a.e. a.s.. In this case, the adjoint process \((p(\cdot), q(\cdot))\), defined by BSDE (14), is identically zero, and the maximum condition (17) is trivial, giving no information about the optimal control \(\tilde{u}(\cdot)\). Such a control \(\tilde{u}(\cdot)\) is a singular one. There are other kinds of singular controls, for which the above maximum principle is ineffective. Instead of defining a singular control in the sense of the “directional derivative” as that previously, in this paper, we discuss singular optimal stochastic controls in the following sense, which is related to the Hamiltonian.

**Definition 2.2** An admissible control \(\tilde{u}(\cdot)\) is called singular on control region \(V\), if \(V \subset U\) is nonempty and, for a.e. \(t \in [0, T]\), we have

\[
H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), \tilde{p}(t), \tilde{q}(t)) = H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), v, \tilde{p}(t), \tilde{q}(t)), \quad \forall v \in V,
\]

where \((\tilde{p}, \tilde{q})\) is the solution of BSDE (14) with \((\tilde{u}, \bar{x}, \bar{y})\) replaced with \((\bar{u}, \bar{x}, \bar{y})\).
**Remark 2.1** We shall see that a singular control in the sense of Definition 2.2 will have zero “directional derivative” along some directions. In fact, the solution of BSDE (20) defined in the next section will be \((y_1, z_1) \equiv 0\), if the perturbation \(u^\varepsilon\) is \(V\)-valued. But, as we will prove later, \(J(u^\varepsilon(\cdot)) = J(\bar{u}(\cdot)) + y_1(0) + o(\varepsilon)\), which implies that the “directional derivative” is zero.

The second result of this paper is the following second-order maximum principle, which involves the second-order adjoint processes \((P(\cdot), Q(\cdot))\) given in (15).

**Theorem 2.2** Let Assumptions 1 and 2 be satisfied. Let \((\bar{u}(\cdot); \bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot))\) be an optimal pair and be singular on the control region \(V\). Then, there exists \(I_0 \subset [0, 1]\) which is of full measure, such that at each \(t \in I_0\), \(\bar{u}(\cdot)\) satisfies, in addition to the first-order maximum condition, the following second-order maximum condition:

\[
(\delta G(t; v) + \delta b^*(t; v)P(t))\delta b(t; v) \geq 0, \quad \forall v \in V, \text{ a.s.},
\]

where we have used the following shorthand notation:

\[
G(t; u) := \bar{H}_x(t; u) + \tilde{f}_y(t; u)p(t) + \tilde{f}_z(t; u)(p(t)\bar{\sigma}_x(t) + q(t)).
\]

### 3 First- and Second-Order Taylor Expansion

In this section, we introduce the first- and the second-order variation equations for the optimal pair \((\bar{u}(\cdot); \bar{x}(t), \bar{y}(t), \bar{z}(t))\) by spike variation methods and establish the dependence of the system state on control actions.

Let \(u(\cdot) \in \mathcal{U}_{\text{ad}}, \varepsilon > 0\) and \(E_\varepsilon \subset [0, T]\) be a Borel set with Borel measure \(|E_\varepsilon| = \varepsilon T\). Define the spike variation \(u^\varepsilon\) of the optimal control \(\bar{u}\) as

\[
u^\varepsilon(t) = \bar{u}(t)I_{E_\varepsilon}(t) + u(t)I_{E_\varepsilon^c}(t).
\]

The corresponding state and cost process are denoted as \((x^\varepsilon(\cdot), y^\varepsilon(\cdot), z^\varepsilon(\cdot))\). Let \(x_i(\cdot), i = 1, 2\), be the solution for the following SDEs, which are regarded as the corresponding first- and second-order variation equations for the optimal state process \(\bar{x}(\cdot)\):

\[
dx_1(t) = \left[\tilde{b}_x(t)x_1(t) + \delta b(t; u^\varepsilon(t))\right]dt + \tilde{\sigma}_x(t)x_1(t)dW_t, \quad x_1(0) = 0, \tag{18}
\]

and

\[
dx_2(t) = \left\{\tilde{b}_x(t)x_2(t) + \delta b_x(t; u^\varepsilon(t))x_1(t) + \frac{1}{2}\tilde{b}_{xx}(t)x_1(t) \otimes x_1(t)\right\}dt
\]

\[
+ \left\{\tilde{\sigma}_x(t)x_2(t) + \frac{1}{2}\tilde{\sigma}_{xx}(t)x_1(t) \otimes x_1(t)\right\}dW_t, \tag{19}
\]

\(x_2(0) = 0\).
The following lemma is a standard result and has been partially proved in [18].

**Lemma 3.1** Assume that Assumptions 1 and 2 are satisfied. Then, we have

\[
\begin{align*}
E \left[ \sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t)|^8 \right] &= O(\varepsilon^8), \\
E \left[ \sup_{0 \leq t \leq T} |x_1(t)|^8 \right] &= O(\varepsilon^8), \\
E \left[ \sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t) - x_1(t)|^4 \right] &= O(\varepsilon^8), \\
E \left[ \sup_{0 \leq t \leq T} |x_2(t)|^4 \right] &= O(\varepsilon^8), \\
E \left[ \sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t) - x_1(t) - x_2(t)|^2 \right] &= o(\varepsilon^4).
\end{align*}
\]

**Proof** All the estimations except the last one have been proved in Lemma 3.2 of [18]. Thus, we only need to prove the last estimation. Denote by \( \tilde{x} := x^\varepsilon - \bar{x} - x_1 - x_2 \). Applying Itô formula to \( \tilde{x} \), we get that

\[
d\tilde{x}(t) = I_1(t)dt + I_2(t)dW_t,
\]

with

\[
I_1(t) = b(x^\varepsilon(t), u^\varepsilon(t)) - \bar{b}(t) - \bar{b}_x(t)(x_1(t) + x_2(t)) - \frac{1}{2} \bar{b}_{xx}(t)x_1(t) \otimes x_1(t) - \delta b(t; u^\varepsilon(t)) - \delta b_x(t; u^\varepsilon(t))x_1(t),
\]

and

\[
I_2(t) = \sigma(x^\varepsilon(t)) - \bar{\sigma}(t) - \bar{\sigma}_x(t)(x_1(t) + x_2(t)) - \frac{1}{2} \bar{\sigma}_{xx}(t)x_1(t) \otimes x_1(t).
\]

Using Taylor expansion, we have

\[
b(x^\varepsilon(t), u^\varepsilon(t)) - \bar{b}(t) - \delta b(t; u^\varepsilon(t)) = b(x^\varepsilon(t), u^\varepsilon(t)) - b(\bar{x}(t), u^\varepsilon(t))
\]

\[
= (\bar{b}_x(t) + \delta b_x(t; u^\varepsilon(t))) (x^\varepsilon(t) - \bar{x}(t)) + \frac{1}{2} (\bar{b}_{xx}(t) + \delta b_{xx}(t; u^\varepsilon(t))) (x^\varepsilon(t) - \bar{x}(t)) \otimes (x^\varepsilon(t) - \bar{x}(t)) + r(t),
\]

where \( r(t) \) is the residual term of second-order Taylor expansion satisfying

\[
|r(t)| = o(|x^\varepsilon(t) - \bar{x}(t)|^2),
\]
which implies that
\[ \mathbb{E} \left[ \int_0^T |r(t)|^2 dt \right] = o(\varepsilon^4). \]

Then, we can rewrite \( I_1 \) as
\[
I_1(t) = \tilde{b}_x(t)\vec{x}(t) + \delta b_x(t; u^\varepsilon(t)) (x^\varepsilon(t) - \vec{x}(t) - x_1(t))
+ \frac{1}{2} \tilde{b}_{xx}(t) \left\{ (x^\varepsilon(t) - \vec{x}(t)) \otimes (x^\varepsilon(t) - \vec{x}(t)) - x_1(t) \otimes x_1(t) \right\}
+ \frac{1}{2} \delta b_{xx}(t; u^\varepsilon(t)) (x^\varepsilon(t) - \vec{x}(t)) \otimes (x^\varepsilon(t) - \vec{x}(t)) + r(t).
\]

We estimate them term by term. For the second term, we have
\[
\mathbb{E} \left[ \left( \int_0^T |\delta b_x(t; u^\varepsilon(t)) (x^\varepsilon(t) - \vec{x}(t) - x_1(t))| dt \right)^2 \right]
\leq \mathbb{E} \left[ \int_0^T |\delta b_x(t; u^\varepsilon(t))|^2 dt \int_0^T |x^\varepsilon(t) - \vec{x}(t) - x_1(t)|^2 dt \right]
\leq C \varepsilon \mathbb{E} \left[ \sup_{0 \leq t \leq T} |x^\varepsilon(t) - \vec{x}(t) - x_1(t)|^2 \right] = o(\varepsilon^4).
\]

For the third term,
\[
\mathbb{E} \left[ \int_0^T \frac{1}{2} \tilde{b}_{xx}(t) \left\{ (x^\varepsilon(t) - \vec{x}(t)) \otimes (x^\varepsilon(t) - \vec{x}(t)) - x_1(t) \otimes x_1(t) \right\}^2 dt \right]
\leq C \mathbb{E} \left[ \int_0^T |x^\varepsilon(t) - \vec{x}(t) - x_1(t)|^4 + |x^\varepsilon(t) - \vec{x}(t) - x_1(t)|^2 |x_1(t)|^2 dt \right]
= o(\varepsilon^4),
\]

where we use the second and the third estimations of this lemma to get the last equality. Similarly,
\[
\mathbb{E} \left[ \left( \int_0^T \left| \frac{1}{2} \delta b_{xx}(t; u^\varepsilon(t)) (x^\varepsilon(t) - \vec{x}(t)) \otimes (x^\varepsilon(t) - \vec{x}(t)) \right| dt \right)^2 \right]
\leq C \mathbb{E} \left[ \int_0^T |\delta b_{xx}(t; u^\varepsilon(t))|^2 dt \int_0^T |(x^\varepsilon(t) - \vec{x}(t)) \otimes (x^\varepsilon(t) - \vec{x}(t))|^2 dt \right]
\leq C \varepsilon \mathbb{E} \left[ \sup_{0 \leq t \leq T} |x^\varepsilon(t) - \vec{x}(t)|^4 \right] = o(\varepsilon^4).
\]
Hence, we see that 
\[ I_1(t) = \bar{b}_x(t) \tilde{x}(t) + i_1(t) \]
with 
\[ E \left[ \left( \int_0^T |i_1(t)|^2 dt \right)^2 \right] = o(\varepsilon^4). \]
For \( I_2(t) \), we can get a similar but much simpler result that 
\[ I_2(t) = \bar{\sigma}_x(t) \tilde{x}(t) + i_2(t) \]
with 
\[ E \left[ \left( \int_0^T |i_2(t)|^2 dt \right)^2 \right] = o(\varepsilon^4). \]
Then, 
\[ d\tilde{x}(t) = \{\bar{b}_x(t) \tilde{x}(t) + i_1(t)\} dt + \{\bar{\sigma}_x(t) \tilde{x}(t) + i_2(t)\} dW_t. \]

Standard estimation on SDE (see Proposition 2.1 in [23]) yields that 
\[ E \left[ \sup_{0 \leq t \leq T} |\tilde{x}(t)|^2 \right] \leq C \left( E \left[ \left( \int_0^T |i_1(t)|^2 dt \right)^2 \right] + E \left[ \int_0^T |i_2(t)|^2 dt \right] \right) = o(\varepsilon^4). \]

Hence, we finish the proof. \( \square \)

From Lemmas 2.1 and 3.1, we immediately have

**Lemma 3.2** For \( r = 2, 4 \),
\[ E \left[ \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t)|^2 + \left( \int_0^T |z^\varepsilon(t) - \bar{z}(t)|^2 dt \right)^r \right] = O(\varepsilon^{2r}). \]

Let \((y_1(\cdot), z_1(\cdot))\) be the solution of the following BSDE:
\[
dy_1(t) = -\left\{ \tilde{f}_y(t)y_1(t) + \tilde{f}_z(t)z_1(t) + p(t)\delta b(t; u_\varepsilon(t)) + \delta f(t; u_\varepsilon(t)) \right\} dt + z_1(t) dW_t, \]
\[ y_1(T) = 0. \]

**Lemma 3.3** Assume Assumption 1 to be satisfied. Then, the following estimation holds:
\[ E \left[ \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t) - p(t)x_1(t) - y_1(t)|^4 \right] = o(\varepsilon^4). \]

**Proof** Define
\[ \tilde{y}^\varepsilon(t) := y^\varepsilon(t) - \bar{y}(t) - p(t)x_1(t) - y_1(t), \]
and
\[ \tilde{z}^\varepsilon(t) := z^\varepsilon(t) - \bar{z}(t) - p(t)\bar{\sigma}_x(t)x_1(t) - q(t)x_1(t) - z_1(t). \]
Applying Itô’s formula to \( \tilde{y}^\varepsilon(\cdot) \), we have
\[ d\tilde{y}^\varepsilon(t) = -I(t) dt + \tilde{z}^\varepsilon(t) dW_t, \]
with
\[ I(t) := f(t, x^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), u^\varepsilon(t)) - \tilde{f}(t) - \tilde{f}_x(t)x_1(t) - \tilde{f}_y(t)p(t)x_1(t) + y_1(t) - \tilde{f}_z(t)(p(t)\tilde{\sigma}_x(t)x_1(t) + q(t)x_1(t) + z_1(t)) - \delta f(t; u^\varepsilon_t). \]

Thus, we see that
\[
\begin{align*}
    f(t, x^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), u^\varepsilon(t)) - \tilde{f}(t) - \delta f(t; u^\varepsilon(t))
    &= f(t, x^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), u^\varepsilon(t)) - f(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), u^\varepsilon(t))
    &\quad+ (\tilde{f}_x(t) + \delta f_x(t; u^\varepsilon(t)))(x^\varepsilon(t) - \bar{x}(t)) + (\tilde{f}_y(t) + \delta f_y(t; u^\varepsilon(t)))(y^\varepsilon(t) - \bar{y}(t))
    &\quad+ (\tilde{f}_z(t) + \delta f_z(t; u^\varepsilon(t)))(z^\varepsilon(t) - \bar{z}(t)) + r(t),
\end{align*}
\]

where \( r(t) \) is the residual term of second-order Taylor expansion. From Assumption 2 of the boundedness of the second-order derivatives of \( f \), we have
\[
    |r(t)| \leq C(|x^\varepsilon(t) - \bar{x}(t)|^2 + |y^\varepsilon(t) - \bar{y}(t)|^2 + |z^\varepsilon(t) - \bar{z}(t)|^2).
\]

Thus, due to Lemmas 3.1 and 3.2,
\[
    \mathbb{E} \left[ \left( \int_0^T |r(t)| dt \right)^4 \right] = O(\varepsilon^8).
\]

For the term \( \delta f_z(t; u^\varepsilon(t))(z^\varepsilon(t) - \bar{z}(t)) \), from the choice of \( E^\varepsilon \) and Lemma 3.2, we see that
\[
    \begin{align*}
        \mathbb{E} \left[ \left( \int_0^T |\delta f_z(t; u^\varepsilon(t))(z^\varepsilon(t) - \bar{z}(t))| dt \right)^4 \right]
        &\leq \mathbb{E} \left[ \left( \int_0^T |\delta f_z(t; u^\varepsilon(t))|^2 dt \right)^2 \left( \int_0^T |z^\varepsilon(t) - \bar{z}(t)|^2 dt \right)^2 \right]
        &\leq C\varepsilon^2 \mathbb{E} \left[ \left( \int_0^T |z^\varepsilon(t) - \bar{z}(t)|^2 dt \right)^2 \right] = o(\varepsilon^4).
    \end{align*}
\]

One can get the similar estimations for the terms \( \delta f_x(t; u^\varepsilon(t))(x^\varepsilon(t) - \bar{x}(t)) \) and \( \delta f_y(t; u^\varepsilon(t))(y^\varepsilon(t) - \bar{y}(t)) \). Thus, combining these together, we have
\[
    f(t, x^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), u^\varepsilon(t)) - \tilde{f}(t) - \delta f(t; u^\varepsilon(t))
    = \tilde{f}_x(t)(x^\varepsilon(t) - \bar{x}(t)) + \tilde{f}_y(t)(y^\varepsilon(t) - \bar{y}(t)) + \tilde{f}_z(t)(z^\varepsilon(t) - \bar{z}(t)) + i(t),
\]

where the residual term \( i(t) \) satisfies
\[
    \mathbb{E} \left[ \left( \int_0^T |i(t)| dt \right)^4 \right] = o(\varepsilon^4).
\]
Hence,

\[ I(t) = \tilde{f}_x(t)(x^\varepsilon(t) - \bar{x}(t) - x_1(t)) + \tilde{f}_y(t)\tilde{y}^\varepsilon(t) + \tilde{f}_z(t)\tilde{z}^\varepsilon(t) + i(t). \]

The standard estimate of BSDEs yields that

\[
E \left[ \sup_{0 \leq t \leq T} |\tilde{y}^\varepsilon(t)|^4 \right] \leq CE \left[ \int_0^T |x^\varepsilon(t) - \bar{x}(t) - x_1(t)| + |i(t)| \, dt \right]^4 = o(\varepsilon^4),
\]

where we use the third estimation in Lemma 3.1 to obtain the result. \(\square\)

To derive the second-order condition in the next section, we also need to expand the value function to the second order. Let \((y_2(\cdot), z_2(\cdot))\) be the solution of the following:

\[
dy_2(t) = - \left\{ \tilde{f}_y(t)y_2(t) + \tilde{f}_z(t)z_2(t) + \delta b^\varepsilon(t; u^\varepsilon_t) P(t)x_1(t) + p(t)\delta b_x(t; u^\varepsilon_t)x_1(t) + \delta f_x(t; u^\varepsilon(t)) + \delta f_y(t; u^\varepsilon(t)) p(t) + \delta f_z(t; u^\varepsilon(t)) (p(t)\bar{\sigma}_x(t) + q(t)) \right\} \, dt + z_2(t) \, dW_t,
\]

\[ y_2(T) = 0. \tag{21} \]

We now establish the following lemma.

**Lemma 3.4** Assume that Assumptions 1 and 2 are satisfied. Let \(\tilde{u}(\cdot)\) be an optimal control singular on the control region \(V\) and \(u(\cdot)\) any \(V\)-valued admissible control. We have

\[ E \left[ \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t) - p(t)(x_1(t) + x_2(t)) - \frac{1}{2} P(t)x_1(t) \otimes x_1(t) - y_2(t)|^2 \right] = o(\varepsilon^4). \tag{22} \]

**Proof** Since \(\tilde{u}\) is singular, Definition 2.2 is equivalent to that

\[ \delta H(t; u^\varepsilon(t)) = p(t)\delta b(t; u^\varepsilon(t)) + \delta f(t; u^\varepsilon(t)) = 0, \]

for all \(t \in [0, T]\), whenever the admissible control \(u(\cdot)\) is \(V\)-valued. Then, the corresponding solution \((y_1, z_1)\) of BSDE (20) is \((y_1, z_1) \equiv 0\). Hence, from Lemma 3.3, we have

\[ E \left[ \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t) - p(t)x_1(t)|^4 \right] = o(\varepsilon^4). \tag{23} \]

\(\square\) Springer
Furthermore, from the definition of $\tilde{z}^\varepsilon$ in the proof above and Lemma 2.1, we get that
\[
\mathbb{E} \left[ \left( \int_0^T |z^\varepsilon(t) - \tilde{z}(t) - (p(t)\tilde{\sigma}_x(t) + q(t))x_1(t)|^2 \, dt \right)^2 \right] = o(\varepsilon^4). \tag{24}
\]
Applying Itô’s formula, we have
\[
d \left[ p(t)(x_1(t) + x_2(t)) \right] = \left\{ p(t)\Delta b(t; u^\varepsilon(t)) + p(t)\Delta b_x(t)x_1(t) - \tilde{f}_y(t)\left[ p(t)(x_1(t) + x_2(t)) \right] \\
- \tilde{f}_z(t)\left[ \tilde{\sigma}_x(t)p(t) + q(t) \right](x_1(t) + x_2(t)) - \tilde{f}_x(t)\left[ x_1(t) + x_2(t) \right] \\
+ \frac{1}{2}\left[ p(t)\tilde{b}_{xx}(t) + q(t)\tilde{\sigma}_{xx}(t) \right]x_1(t) \otimes x_1(t) \right\} \, dt + \left\{ q(t)\left[ x_1(t) + x_2(t) \right] \\
+ p(t)\left[ \tilde{\sigma}_x(t)(x_1(t) + x_2(t)) + \frac{1}{2}\tilde{\sigma}_{xx}(t)x_1(t) \otimes x_1(t) \right] \right\} \, dW_t, \tag{25}
\]
and
\[
d \left[ \frac{1}{2}P(t)x_1(t) \otimes x_1(t) \right] = \left\{ \Delta^b(t; u^\varepsilon(t))\left[ \frac{1}{2}P(t)(x_1(t))^2 \right] \\
- \tilde{f}_z(t)\left[ \tilde{\sigma}_x(t)P(t) + \frac{1}{2}Q(t) \right](x_1(t))^2 \\
- \frac{1}{2}\left[ \tilde{b}_{xx}(t)p(t) + \tilde{\sigma}_{xx}(t)[\tilde{f}_z(t)p(t) + q(t)] \right] + [1, p(t), \tilde{\sigma}_x(t)p(t) \\
+ q(t)]D^2 \tilde{f}(t)[1, p(t), \tilde{\sigma}_x(t)p(t) + q(t)]^T x_1(t) \otimes x_1(t) \right\} \, dt \\
+ \left\{ \left[ P(t)\tilde{\sigma}_x(t) + \frac{1}{2}Q(t) \right]x_1(t) \otimes x_1(t) \right\} \, dW_t. \tag{26}
\]
Define
\[
\hat{y}^\varepsilon(t) := y^\varepsilon(t) - \hat{y}(t) - p(t)(x_1(t) + x_2(t)) - \frac{1}{2}P(t)(x_1(t))^2 - y_2(t),
\]
and
\[
\hat{z}^\varepsilon(t) := z^\varepsilon(t) - \hat{z}(t) - \left\{ q(t)\left[ x_1(t) + x_2(t) \right] \\
+ p(t)\left[ \tilde{\sigma}_x(t)(x_1(t) + x_2(t)) + \frac{1}{2}\tilde{\sigma}_{xx}(t)x_1(t) \otimes x_1(t) \right] \right\}.
\]
 Springer
\[-\left[ P(t)\bar{\sigma}_x(t) + \frac{1}{2}Q(t) \right] x_1(t) \otimes x_1(t) - z_2(t).\]

Moreover, using Taylor expansion of \( f \), we have

\[
\begin{align*}
    f(t, x^e(t), y^e(t), z^e(t), u^e(t)) - f(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), u^e(t)) &= f_x(t; u^e(t))(x^e(t) - \bar{x}(t)) + f_y(t; u^e(t))(y^e(t) - \bar{y}(t)) \\
    &+ f_z(t; u^e(t))(z^e(t) - \bar{z}(t)) + \frac{1}{2}[x^e(t) - \bar{x}(t), y^e(t) \\
    &- \bar{y}(t), z^e(t) - \bar{z}(t)]D^2 f(t; u^e(t))[x^e(t) - \bar{x}(t), y^e(t) \\
    &- \bar{y}(t), z^e(t) - \bar{z}(t)]^T + i_1(t),
\end{align*}
\]

where \( i_1(t) \) is the residual term of Taylor expansion such that

\[
|i_1(t)| = o((x^e(t) - \bar{x}(t))^2 + (y^e(t) - \bar{y}(t))^2 + (z^e(t) - \bar{z}(t))^2).
\]

Hence, one can obtain that

\[
\mathbb{E}\left[ \left( \int_0^T |i_1(t)| dt \right)^2 \right] = o(\varepsilon^4).
\]

Also, we see that

\[
f_x(t; u^e(t))(x^e(t) - \bar{x}(t)) = \tilde{f}_x(t)(x^e(t) - \bar{x}(t)) + \delta f_x(t; u^e(t))x_1(t) + i_2(t),
\]

where \( i_2(t) = \delta f_x(t; u^e(t))(x^e(t) - \bar{x}(t) - x_1(t)) \). Then, using Hölder inequality, we get

\[
\begin{align*}
    \mathbb{E}\left[ \left( \int_0^T |i_2(t)| dt \right)^2 \right] &\leq \mathbb{E}\left[ \left( \int_0^T \delta f_x(t; u^e(t))^2 dt \right) \left( \int_0^T |x^e(t) - \bar{x}(t) - x_1(t)|^2 dt \right) \right].
\end{align*}
\]

From the estimation in Lemma 3.1, we have

\[
\mathbb{E}\left[ \left( \int_0^T |i_2(t)| dt \right)^2 \right] = o(\varepsilon^4).
\]

Similarly, we have

\[
f_y(t; u^e(t))(y^e(t) - \bar{y}(t)) = \tilde{f}_y(t)(y^e(t) - \bar{y}(t)) + \delta f_y(t; u^e(t))p(t)x_1(t) + i_3(t),
\]

and
with

\[ i_3(t) = \delta f_y(t; u^\varepsilon(t))(y^\varepsilon(t) - \bar{y}(t) - p(t)x_1(t)), \]

and

\[ i_4(t) = \delta f_z(t; u^\varepsilon(t))[z^\varepsilon(t) - \bar{z}(t) - (p(t)\sigma_x(t) + q(t))x_1(t)]. \]

Using Hölder inequality again and estimations (23) and (24), we have

\[
\mathbb{E}\left[\left(\int_0^T |i_3(t)| dt\right)^2\right] + \mathbb{E}\left[\left(\int_0^T |i_4(t)| dt\right)^2\right] = o(\varepsilon^4).
\]

For the quadratic term, we see that

\[
\begin{align*}
[x^\varepsilon(t) - \bar{x}(t), y^\varepsilon(t) - \bar{y}(t), z^\varepsilon(t) - \bar{z}(t)]D^2 f(t; u^\varepsilon(t)) & \times [x^\varepsilon(t) - \bar{x}(t), y^\varepsilon(t) - \bar{y}(t), z^\varepsilon(t) - \bar{z}(t)]^T \\
= [1, p(t), p(t)\sigma_x(t) + q(t)]D^2 \bar{f}(t)[1, p(t), p(t)\sigma_x(t)] & + q(t))^T x_1(t) \otimes x_1(t) + i_5(t),
\end{align*}
\]

with \( i_5(t) \) being just the difference of the left-hand side and the first term on the right-hand side. With a similar estimation, one can check that

\[
\mathbb{E}\left[\left(\int_0^T |i_5(t)| dt\right)^2\right] = o(\varepsilon^4).
\]

Thus, finally we rewrite (27) as

\[
f(t, x^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), u^\varepsilon(t)) - \bar{f}(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), u^\varepsilon(t)) \\
= \bar{f}_x(t)(x^\varepsilon(t) - \bar{x}(t)) + \bar{f}_y(t)(y^\varepsilon(t) - \bar{y}(t)) + \bar{f}_z(t)(z^\varepsilon(t) - \bar{z}(t)) & + \delta f_x(t; u^\varepsilon(t))x_1(t) + \delta f_y(t; u^\varepsilon(t))p(t)x_1(t) + \delta f_z(t; u^\varepsilon(t))(p(t)\sigma_x(t) \\
+ q(t))x_1(t) + \frac{1}{2}[1, p(t), p(t)\sigma_x(t) + q(t)]D^2 \bar{f}(t)[1, p(t), p(t)\sigma_x(t)] & + q(t))^T x_1(t) \otimes x_1(t) + i_6(t),
\]

with \( i_6(t) \) combining all the residual terms above together and satisfying

\[
\mathbb{E}\left[\left(\int_0^T |i_6(t)| dt\right)^2\right] = o(\varepsilon^4).
\]
Combining (25), (26) and (28), we obtain that
\[
d\hat{y}^\varepsilon(t) = -\{ \bar{f}_y(t)\hat{y}^\varepsilon(t) + \bar{f}_z(t)\hat{z}^\varepsilon(t) + \bar{f}_x(t)(x^\varepsilon(t) - \bar{x}(t) - x_1(t) - x_2(t)) + \delta f(t; u^\varepsilon(t)) + p(t)\delta b(t; u^\varepsilon(t)) + i_6(t)\}dt + \hat{z}^\varepsilon(t)dW_t.
\]

We assume that the optimal control \(\bar{u}\) is singular, which implies that
\[
\delta f(t; u^\varepsilon(t)) + p(t)\delta b(t; u^\varepsilon(t)) = 0.
\]

Thus, once again applying Lemmas 2.1 and 3.1, we have
\[
\mathbb{E}\left[ \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t) - p(t)(x_1(t) + x_2(t)) - \frac{1}{2}P(t)(x_1(t))^2 - y_2(t)|^2 \right] = o(\varepsilon^4).
\]

\(\square\)

4 Proof for the Main Results

4.1 First-Order Condition

The solution of the linear BSDE (20) can be represented via the adjoint SDE. Let \(\gamma(t)\) satisfy
\[
d\gamma(t) = \bar{f}_y(t)\gamma(t)dt + \bar{f}_z(t)\gamma(t)dW_t, \quad \gamma(0) = 1. \tag{29}
\]

Applying Itô’s formula to \(\gamma(t)y_1(t)\), we shall have
\[
y_1(0) = \mathbb{E}\left[ \int_0^T \{ \gamma(t)(p(t)\delta b(t; u(t)) + \delta f(t; u(t)))1_{E^\varepsilon}(t)\}dt \right].
\]

According to the range theorem of vector-valued measures (see Li and Yao [24]), one can choose \(E^\varepsilon\) carefully such that \(|E^\varepsilon| = \varepsilon T\) and
\[
\mathbb{E}\left[ \int_0^T \{ \gamma(t)(p(t)\delta b(t; u(t)) + \delta f(t; u(t)))1_{E^\varepsilon}(t)\}dt \right] = \varepsilon\mathbb{E}\left[ \int_0^T \{ \gamma(t)(p(t)\delta b(t; u(t)) + \delta f(t; u(t)))\}dt \right] + o(\varepsilon).
\]

We have, from Lemma 3.3,
\[
J(u^\varepsilon) = y^\varepsilon(0) = \bar{y}(0) + \varepsilon\mathbb{E}\left[ \int_0^T \{ \gamma(t)(p(t)\delta b(t; u(t)) + \delta f(t; u(t)))\}dt \right] + o(\varepsilon).
\]
Since $\tilde{y}(0)$ is optimal, we shall have

$$\limsup_{\varepsilon \to 0} \frac{y^\varepsilon(0) - \tilde{y}(0)}{\varepsilon} \geq 0,$$

which implies that

$$E\left[ \int_0^T \{ \gamma(t)(p(t)\delta b(t; u(t)) + \delta f(t; u(t))) \} \, dt \right] \geq 0, \quad (30)$$

for any $u \in U_{ad}$. Finally, deduced from above inequality, we claim that (17) holds. The proof is similar to that in Tang [18, p. 1594]. For readers’ convenience, we still give the proof. Due to the arbitrariness of $u(\cdot)$ and positiveness of $\gamma(\cdot)$, it can be shown that for any $v \in U$, there is a null set $S^v$ such that

$$\delta H(t; v) \geq 0, \quad \text{for each } (t, \omega) \in [0, T] \times \Omega / S^v. \quad (31)$$

Select a countable dense subset $\{v_i\}_{i=1}^\infty \subset U$, and set

$$S_0 = \bigcup_{i=1}^\infty S^{v_i}.$$

Then, $S_0$ is a null set, and for $(t, \omega) \in S =: [0, T] \times \Omega / S_0$,

$$\delta H(t; v_i) \geq 0, \quad \text{for } i = 1, 2, \ldots, \infty. \quad (31)$$

Set the $t$-sections $S(t)$ of $S$ as

$$S(t) = \{ \omega \in \Omega | (t, \omega) \in S \}, \quad \text{for any } t.$$

It is classical to prove that $P(S(t)) = 1$ for a.e. $t \in [0, T]$. Therefore, there is a null set $T_0 \subset [0, T]$, such that for each $t_0 \in [0, T] / T_0$, (31) holds. By the continuity of the function $\delta H(t; v)$ in $v$ as well as the density of $\{v_i\}_{i=1}^\infty$ in $U$, we have for $t \in [0, T] / T_0$,

$$\delta H(t; v) \geq 0, \quad \forall v \in U, \ \text{a.s.}$$

Thus, the proof of Theorem 2.1 is completed. \hfill \Box

### 4.2 Second-Order Condition

In this subsection, we are going to prove Theorem 2.2. Recall that, as defined in the statement of Theorem 2.2,

$$G(t; u) = \tilde{H}_x(t; u) + \tilde{f}_x(t; u)p(t) + \tilde{f}_z(t; u)(p(t)\tilde{\sigma}_x(t) + q(t)).$$
The solution of BSDE (21) can be rewritten explicitly as

$$y_2(0) = \mathbb{E} \left[ \int_0^T \gamma(t) \left\{ \delta G(t, v^e(t)) + \delta b^*(t, v^e(t)) P(t) \right\} x_1(t; v^e(\cdot)) dt \right],$$

for any $v \in V_{ad}(t_1, t_2)$. Here

$$V_{ad}(t_1, t_2) := \{ v \in U_{ad} | v(t) \in V, \text{ a.s., a.e. } t \in [t_1, t_2]; \ v(t) = \bar{u}(t), \ t \in [0, T]/[t_1, t_2] \}.$$

According to the range theorem of vector-valued measures (see Li and Yao [24] or Tang [18]), one can choose a Borel set $E^\varepsilon$ such that

$$|E^\varepsilon| = \varepsilon T,$$

$$\int_{[0,t] \cap E^\varepsilon} \delta b(s; v(s)) ds = \varepsilon \int_0^t \delta b(s; v(s)) ds + \tilde{\eta}(t),$$

$$\sup_{0 \leq t \leq T} \mathbb{E}[|\tilde{\eta}(t)|^2] = O(\varepsilon^4),$$

$$\mathbb{E} \left[ \int_{E^\varepsilon} \gamma(t) \left\{ \delta G(t, v^e(t)) + \delta b^*(t, v^e(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] = \varepsilon \mathbb{E} \left[ \int_0^T \gamma(t) \left\{ \delta G(t, v(t)) + \delta b^*(t, v(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] + o(\varepsilon^2).$$

From the second relation of (32) and standard estimation for SDE (see Proposition 2.1 in [23]), we can check that

$$E \left[ \sup_{0 \leq t \leq T} |x_1(t; v^e(\cdot)) - \varepsilon x_1(t; v(\cdot))|^2 \right] = o(\varepsilon^2).$$

Then,

$$y_2(0) = \varepsilon \mathbb{E} \left[ \int_0^T \gamma(t) \left\{ \delta G(t, v^e(t)) + \delta b^*(t, v^e(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] + \varepsilon \mathbb{E} \left[ \int_0^T \gamma(t) \left\{ \delta G(t, v(t)) + \delta b^*(t, v(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] \times (x_1(t; v^e(\cdot)) - \varepsilon x_1(t; v(\cdot))) dt.$$

From (33), one can get that the second term on the right-hand side is of $o(\varepsilon^2)$ large. Using the third relation in (32), we get that

$$y_2(0) = \varepsilon^2 \mathbb{E} \left[ \int_0^T \gamma(t) \left\{ \delta G(t, v(t)) + \delta b^*(t, v(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] + o(\varepsilon^2).$$
Combining with Lemma 3.4, we have

\[
J(v^\varepsilon(\cdot)) = y^\varepsilon(0)
\]

\[
= \bar{y}(0) + \varepsilon^2 \mathbb{E}\left[ \int_0^T \gamma(t) \left\{ \delta G(t, v(t)) + \delta b^*(t, v(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] + o(\varepsilon^2).
\]

From the optimality of \(\bar{u}\), we shall have

\[
\limsup_{\varepsilon \to 0} \frac{y^\varepsilon(0) - \bar{y}(0)}{\varepsilon^2} \geq 0,
\]

which implies that

\[
\mathbb{E}\left[ \int_{t_1}^{t_2} \gamma(t) \left\{ \delta G(t, v(t)) + \delta b^*(t, v(t)) P(t) \right\} x_1(t; v(\cdot)) dt \right] \geq 0,
\]

for any \(v \in V_{ad}(t_1, t_2)\). Note that one can solve (18) explicitly:

\[
x_1(t) = \int_{t_1}^{t} \Phi(s; t) \delta b(s; v(s)) ds,
\]

where \(\Phi(\cdot; t)\) is the first component of the solution \((\Phi(\cdot; t), \Psi(\cdot; t))\) for the following matrix-valued BSDE defined on the time interval \([t_1, t]\),

\[
d\Phi(s; t) = -\left\{ \Phi(s; t) \bar{b}_x(s) + \Psi(s; t) \bar{\sigma}_x(t) \right\} ds + \Psi(s; t) dW_s, \quad \Phi(t; t) = I.
\]

Moreover, for any \(t\), \(\Phi(s; t)\) is continuous in \(s\) almost surely. Thus, we can rewrite the left-hand side of (34) as

\[
\mathbb{E}\left[ \int_{t_1}^{t} \int_{t_1}^{t} \gamma(t) \left\{ \delta G(t, v(t)) + \delta b^*(t, v(t)) P(t) \right\} \Phi(s; t) \delta b(s; v(s)) ds dt \right].
\]

Denote by \(\{r_i\}_{i=1}^{\infty}\) the totality of rational numbers in \([0, 1]\) and by \(\{u_i\}_{i=1}^{\infty}\) a dense subset of \(V\). Since \(\mathcal{F}_t\) is countable generated for \(t \in [0, T]\), we can assume that \(\{A_{ij}\}_{j=1}^{\infty}\) generates \(\mathcal{F}_{t_i}, i = 1, 2, 3, \ldots\). Set

\[
Z^v_{ij}(t) := \bar{u}(t) \chi_{A_{ij}}(\omega) \chi_{[0, r_i]}(t) + v \chi_{A_{ij}}(\omega) \chi_{[r_i, 1]}(t),
\]

for \(t \in [0, T], v \in V, i, j = 1, 2, \ldots\). Now we claim that (see in appendix for the proof) there is a null subset \(T^k_{ij} \subset [0, T]\) such that, for \(t \in [0, T]\setminus T^k_{ij}\),
Then, using above estimation together with Hölder inequality, we have
\[
\lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E} \left[ \gamma(u)(\delta G(t, Z_{ij}^{u_k}(u)) + \delta b^*(t, Z_{ij}^{u_k}(u)) P(u)) \delta b(s, Z_{ij}^{u_k}(s)) \right] ds du
\]
\[
= \lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E} \left[ \gamma(u)(\delta G(t, Z_{ij}^{u_k}(u)) + \delta b^*(t, Z_{ij}^{u_k}(u)) P(u)) \delta b(s, Z_{ij}^{u_k}(s)) I_{\{s \leq u\}} \right] ds du
\]
\[
= \frac{(\alpha + \beta)^2}{2} \mathbb{E} \left[ \gamma(t)(\delta G(t, Z_{ij}^{u_k}(t)) + \delta b^*(t, Z_{ij}^{u_k}(t)) P(t)) \delta b(t, Z_{ij}^{u_k}(t)) \right].
\]
(35)

Set
\[
\mathcal{T}_0 := \bigcup_{1 \leq i, j, k \leq \infty} \mathcal{T}_{ij}^k.
\]

Then, \(\mathcal{T}_0\) is a null subset of \([0, T]\). For \(t \in [0, T]/\mathcal{T}_0\) and the integers \(i\) such that \(r_i < t\), let the perturbed control \(v\) be \(v(s) = u(s) \chi_{[0,1/|t-r_i|]}(s) + \tilde{Z}_{ij}^{u_k}(s) \chi_{[t-r_i, t+\alpha]}(s).\)

We have
\[
\frac{1}{r^2} \mathbb{E} \left[ \int_{t-r}^{t+r} \int_{t-r}^{t+r} \gamma(u) \left\{ \delta G(u, Z_{ij}^{u_k}(u)) + \delta b^*(t, Z_{ij}^{u_k}(u)) P(u) \right\} \right. \Phi(s; u) \delta b(s; Z_{ij}^{u_k}(s)) ds du \left. \right] \geq 0.
\]

We see that \(\Phi(\cdot; u) - I\) satisfies
\[
d(\Phi(s; u) - I) = - \left\{ (\Phi(s; u) - I) \tilde{b}_\alpha(s) + \Psi(s; u) \tilde{\sigma}_\alpha(s) + \tilde{b}_\alpha(s) \right\} ds
\]
\[
+ \Psi(s; u) dW_\alpha.
\]

From Lemma 2.1 and the boundedness of the first-order derivative of the coefficient, we have
\[
\mathbb{E}[|\Phi(s; u) - I|^2] \leq C \mathbb{E} \left[ \left( \int_s^u |\tilde{b}_\alpha(r)| dr \right)^2 \right] \leq C(u - s)^2.
\]

Then, using above estimation together with Hölder inequality, we have
\[
\left| \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E} \left[ \gamma(u) \left\{ \delta G(u, Z_{ij}^{u_k}(u)) + \delta b^*(t, Z_{ij}^{u_k}(u)) P(u) \right\} (\Phi(s; u) - I) \delta b(s; Z_{ij}^{u_k}(s)) \right] ds du \right|
\]
\[ \geq 0. \]
\[
\leq C \int_{t-r}^{t+r} \mathbb{E} \left[ |\Phi(s; u) - I|^2 \right] \frac{1}{r^2} ds du \\
\leq C \int_{t-r}^{t+r} \mathbb{E} \left[ (u - s) ds du = Cr^3. \right.
\]

Thus,
\[
\frac{1}{r^2} \mathbb{E} \left[ \int_{t-r}^{t+r} \int_{t-r}^{u} \gamma(u) \left\{ \delta G(u, Z_{ij}^u(u)) + \delta b^*(t, Z_{ij}^{u_k}(u)) P(u) \right\} \delta b(s; Z_{ij}^{u_k}(s)) ds du \right] + o(1) \geq 0.
\]

Let \( r \) tend to 0 in the above inequality. From (35), we finally get that
\[
\mathbb{E} \left[ \gamma(t)(\delta G(t; u_k) + \delta b^*(t; u_k) P(t)) \delta b(t; u_k) \chi_{A_{ij}} \right] \geq 0.
\]

Since \( \{A_{ij}\}_{j=1}^\infty \) generate \( \mathcal{F}_{ri} \), we have
\[
\mathbb{E} \left[ \gamma(t)(\delta G(t; u_k) + \delta b^*(t; u_k) P(t)) \delta b(t; u_k) \right| \mathcal{F}_{ri} \right] \geq 0, a.s.\]

Since the filtration is generated by the Brownian motion, \( \mathcal{F}_t \) is also left continuous. Then, it holds that
\[
\gamma(t)(\delta G(t; u_k) + \delta b^*(t; u_k) P(t)) \delta b(t; u_k) \geq 0, \text{ a.s.}
\]

Since \( \gamma(t) \) is positive, it is equivalent to
\[
(\delta G(t; u_k) + \delta b^*(t; u_k) P(t)) \delta b(t; u_k) \geq 0, \text{ a.s.}
\]

From the continuity of the coefficients and the density of \( \{u_k\}_{k=1}^\infty \), we have
\[
(\delta G(t; u) + \delta b^*(t; u) P(t)) \delta b(t; u) \geq 0, \forall u \in V, \text{ a.s.}
\]

holds. Therefore, we finish the proof of Theorem 2.2.

\[\square\]

5 Examples

In this section, we give some examples to illustrate the application of our maximum principle.

\[\copyright\ Springer\]
**Example 5.1** The state process of the controlled system is
\[
dx(t) = \begin{pmatrix} \frac{1}{2}a^2 & -u \\ u & \frac{1}{2}a^2 \end{pmatrix} x(t) dt + \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} x(t) dW_t, \quad 0 < t < 1,\]
\[x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\]
with the cost process
\[
dy(t) = -\{\beta y(t) + \gamma z(t)\} dt + z(t) dW_t,
\]
y(T) = \frac{1}{2} |x(T)|^2,
where the valued set \( U \) of admissible controls is:
\[U = [-1, 1].\]
and \( a, \beta, \gamma \) are deterministic. For each constant control \( u \), Eq. (36) can be solved explicitly as
\[x(t; u) = \begin{pmatrix} \cos(ut + aW_t) \\ \sin(ut + aW_t) \end{pmatrix}.
\]
One can check that any admissible control \( u(\cdot) \) is optimal in this example. For the admissible reference pair \((x(\cdot), u)\) with \( u \in U_{ad} \) being constant, the associated first-order adjoint equation \((p(\cdot; u), q(\cdot; u))\) satisfies the following BSDE:
\[
dp(t) = -\left[\begin{pmatrix} \beta - \frac{1}{2}a^2 & \gamma a + u \\ -\gamma a - u & \beta - \frac{1}{2}a^2 \end{pmatrix} p(t) + \begin{pmatrix} \gamma & a \\ -a & \gamma \end{pmatrix} q(t)\right] dt + q(t) dW_t,
p(T) = x(T).
\]
It is solved as
\[p(t; u) = \exp(\beta(T - t)) \begin{pmatrix} \cos(ut + aW_t) \\ \sin(ut + aW_t) \end{pmatrix},
q(t; u) = \exp(\beta(T - t)) \begin{pmatrix} -a \sin(ut + aW_t) \\ a \cos(ut + aW_t) \end{pmatrix}.
\]
Thus, the Hamiltonian can be calculated. It shows that \( H(t, x(t; u), y(t; u), z(t; u), v, p(t; u), q(t; u)) \) is independent of \( v \). Hence, any constant control \( u \) is singular on \( U \). Consider the second-order adjoint equation:
\[
dP(t) = -[\beta P(t) + (b_x + \gamma \sigma_x)^* P(t) + P(t)(b_x + \gamma \sigma_x) + \sigma_x^* P(t) \sigma_x + \gamma Q(t) + \sigma_x^* Q(t) + Q(t) \sigma_x] dt + Q(t) dW_t,
P(1) = I,
\]
with
\[ b_x = \left( \begin{array}{cc} -\frac{1}{2}a^2 & -u \\ u & -\frac{1}{2}a^2 \end{array} \right), \quad \sigma_x = \left( \begin{array}{cc} 0 & -a \\ a & 0 \end{array} \right). \]

Obviously, \( P(t) = \exp(\beta(T - t))I \), \( Q(t) \equiv 0 \). Then, we have
\[
\delta G(t; v) = -\exp(\beta(T - t))(v - u)^2, \\
\delta b^*(t; v)P(t)\delta b(t; v) = \exp(\beta(T - t))(v - u)^2.
\]

It implies that any constant control \( u \) satisfies our second-order maximum principle.

**Example 5.2** The control system is
\[
dx(t) = u(t)dt + (x - 1)dW_t, \quad u_t \in U := \{-1, 0, 1\}, \\
x(0) = 1,
\]
and the cost process is defined as
\[
dy(t) = -f(y(t), z(t))dt + z(t)dW_t, \\
y(1) = \pm \frac{1}{2}(x(1) - 1)^2,
\]
with \( f \) be any deterministic function. For both cost functionals, the constant control \( u \equiv 0 \) is singular on \( U \) since the corresponding first-order adjoint processes are identically zero. The second adjoint processes are \((P, Q)\) with \( Q(t) \equiv 0 \), and \( P(t) \) solves the following ODE:
\[
dP(t) = -[\bar{f}_y(t) + 2\bar{f}_z(t) + 1]P(t)dt, \quad P(1) = \pm 1.
\]

From Theorem 2.2, we see that \( u \equiv 0 \) is a candidate for optimal controls at the case \( y(1) = \frac{1}{2}(x(1) - 1)^2 \) and necessarily not an optimal control at the other case.

**6 Perspectives and Open Problems**

In most general cases, we would like to assume that the control variable enters into both the drift and the diffusion terms with nonconvex control region. Under this situation, if one still uses spike variation, the main difficulty, roughly speaking, is that the Itô stochastic integral \( \int_t^{t+\varepsilon} \sigma dW_s \) is only of order \( \sqrt{\varepsilon} \), rather than \( \varepsilon \). This leads to the result that the estimations in Lemmas 3.1–3.4 will be just of half order. For example, the estimations will become
\[
\mathbb{E}\left[ \sup_{0 \leq t \leq T} |x_{\varepsilon}(t) - \bar{x}(t) - x_1(t)|^4 \right] = O(\varepsilon^4)
\]
and
\[
\mathbb{E}\left[ \sup_{0 \leq t \leq T} |x^\varepsilon(t) - \bar{x}(t) - x_1(t) - x_2(t)|^2 \right] = o(\varepsilon^2)
\]
for proper first- and second-order Taylor expansion $x_1, x_2$ of controlled system. To get the second-order necessary condition, one could imagine that we must expand the system to fourth order. In the meantime, we also need to introduce two extra BSDEs for the dual system. The exact forms of these equations are not all known to us, and their solvability also needs to be verified under proper conditions. This will make the proof quite lengthy and tedious. Besides, the definition for the singular control will be adjusted as proposed by Zhang and Zhang [21] since the first-order maximum condition will have an additional quadratic term. To get the desired result, another major problem is how to get the pointwise condition from an integral condition like (30) and (34). As we have seen in Sect. 4.2, it is not a trivial conclusion and already needs a careful proof. Whether we can obtain a pathwise condition in more general cases is still unknown. Its proof may require more detailed estimations and other mathematical tools such as Malliavin calculus; see [21] for a similar discussion. Thus, it is a challenging problem, which we would like to list it as an open question. Some further investigations will be carried out in our future works. At last, we want to mention that if the control region is convex, using convex variation instead of spike variation will simplify the question. Under this situation, the perturbation of the diffusion term will be like $\varepsilon \int_0^T \sigma dW_s$, which is of order $\varepsilon$. Thus, one does not have to expand the system to fourth order. We believe that we can still get a local second-order condition related to the first and second derivatives of the Hamiltonian. But for a general control region, especially for the bang–bang control, the spike variation seems to be inevitable.

Another interesting question is that when the necessary condition is also sufficient. There are many papers on this subject for first-order condition; see [27,28]. Most of them required a convex condition on the coefficients. To our best knowledge, there seems to be no paper studying for the second-order condition of stochastic optimal control problems. We believe that a kind of convex condition is also enough to make the second-order condition to be sufficient. This will be also carried out in our future work.

7 Conclusions

For stochastic optimal control problems with recursive utilities, we found a necessary second-order condition for the singular optimal control. The main idea is to use spike variation, since we allow the control region to be nonconvex. We expanded the controlled system and the utility process up to second order. Having introduced two BSDEs, the second-order “directional derivatives” can be explicitly calculated. In general, the optimality of the control implies the nonnegative of the derivatives. This further leads to a pointwise condition on the Hamiltonian with a quadratic term in Theorem 2.2. At last, we gave some examples to illustrate the application of our second condition. In our paper, the diffusion term is independent of control variable.
When the control variable is included in the diffusion, we discussed the difficulties that arise along the way. This will be carried out in our future works.

Acknowledgements The first author gratefully acknowledges financial support from Région Pays de la Loire through the Grant PANORisk. The second author was supported by the Natural Science Foundation of Zhejiang Province for Distinguished Young Scholar (No. LR15A010001) and the National Natural Science Foundation of China (Nos. 11871211, 11471079). Both authors would like to thank the editors and anonymous reviewers for their valuable comments and suggestions that helped improve the quality of this paper.

Appendix: Proof for the Claim in Sect. 4.2

We give the proof for the claim that there is a null subset \( T_{ij}^k \subset [0, T] \) such that, for \( t \in [0, T]/T_{ij}^k \),

\[
\lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E} \left[ \gamma(u)(\delta G(t, Z_{ij}^{Hk}(u))) + \delta b^*(t, Z_{ij}^{Hk}(u)) P(u) \right. \\
\left. + \delta b^*(s, Z_{ij}^{Hk}(s)) I_{|s| \leq u} \right] dsdu = (\alpha + \beta)^2 2 \mathbb{E} \left[ \gamma(t)(\delta G(t, Z_{ij}^{Hk}(t))) + \delta b^*(t, Z_{ij}^{Hk}(t)) P(t) \right] \delta b(t, Z_{ij}^{Hk}(t)) \right].
\]

Note that both \( \gamma(u)(\delta G(t, Z_{ij}^{Hk}(u))) + \delta b^*(t, Z_{ij}^{Hk}(u)) P(u) \) and \( \delta b(s, Z_{ij}^{Hk}(s)) \) are square integrable processes. We shall prove a more general result that for any two processes \( f, g \in L_2^2F(0, T) \), there exists a null subset \( T \subset [0, T] \) such that for \( t \in [0, T]/T \),

\[
\lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E}[f(u)g(s)I_{|s| \leq u}]dsdu = \frac{(\alpha + \beta)^2}{2} \mathbb{E}[f(t)g(t)].
\]

This result will prove our previous claim immediately. We divide the proof into three steps.

Step 1. We prove that, for \( f \in L_2^2F(0, T) \), there exists a null subset \( T \subset [0, T] \) such that, for \( t \in [0, T]/T \),

\[
\lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \mathbb{E}[f(t) - f(s)]^2 ds = 0.
\]

Let \( \zeta(t) \) be the standard mollifier (for its properties see Appendix C.4 in [25]), i.e.,

\[
\zeta(t) = \begin{cases} 
\exp \left( - \frac{1}{1 - t^2} \right), & \text{if } |t| \leq 1; \\
0, & \text{otherwise}; 
\end{cases}
\]
and $\zeta_n(t) = n \zeta(nt)$. We define $f_n(t) := f \star \zeta_n(t)$ with $\star$ representing the convolution between functions on $\mathbb{R}$. Then, we see that $f_n$ has continuous trajectories almost surely. Moreover, if we define a function $g$ on $[0, T]$ as $g_n(t) := \mathbb{E}[|f_n(t) - f(t)|^2]$, then $\|g_n\|_{L^1} = \int_0^T g_n(t)dt \longrightarrow 0$ as $n$ goes to infinity. Define

$$T_r(f)(t) := \frac{1}{r} \int_{t-rB}^{t+rB} \mathbb{E}[|f(s) - f(s)|^2]ds,$$

and $T(f)(t) = \limsup_{r \to 0} T_r(f)(t)$. It is then sufficient to prove that the set $E_a := \{ t \in [0, T] | T(f)(t) \geq 2a \}$ is a null set for any $a > 0$. One can easily obtain that

$$T_r(f) \leq C \left( \frac{1}{r} \int_{t-rB}^{t+rB} g_n(s)ds + T_r(f_n)(t) + g_n(t) \right) \leq C(Mg_n(t) + T_r(f_n)(t) + g_n(t)),$$

for any $n$, with $Mg_n$ being the Hardy–Littlewood maximal function of $g_n$. Since $f_n$ has continuous paths, we have $\lim_{r \to 0} T_r(f_n)(t) = 0$ for any $t$. Then,

$$T(f)(t) \leq C(Mg_n(t) + g_n(t)),$$

which implies that

$$E_a \subset \left\{ Mg_n \geq \frac{a}{2C} \right\} \cup \left\{ g_n \geq \frac{a}{2C} \right\}.$$

From Hardy–Littlewood maximal inequality (see [26]),

$$\left| \left\{ Mg_n \geq \frac{a}{2C} \right\} \right| \leq \frac{2C_1C}{a} \|g_n\|_{L^1},$$

where $C_1$ is the constant from the inequality. Using Chebychev’s inequality, we also obtain

$$\left| \left\{ g_n \geq \frac{a}{2C} \right\} \right| \leq \frac{2C}{a} \|g_n\|_{L^1}.$$

Letting $n$ go to infinity, we show that $|E_a| = 0$.

**Step 2. There exists a null subset $T \subset [0, T]$ such that, for $t \in [0, T] / T$,**

$$\lim_{r \to 0^+} \frac{1}{r^2} \int_{t-rB}^{t+rB} \int_{t-rB}^{t+rB} |\mathbb{E}[f(s)g(u)] - \mathbb{E}[f(t)g(t)]| dsdu = 0.$$
\[ \leq \mathbb{E}[|f(s) - f(t)|^2]^{\frac{1}{2}} \mathbb{E}[|g(u)|^2]^{\frac{1}{2}} + \mathbb{E}[|g(u) - g(t)|^2]^{\frac{1}{2}} \mathbb{E}[|f(t)|^2]^{\frac{1}{2}}. \]

Then,
\[ \frac{1}{r^2} \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E}[|f(s) - f(t)|^2]^{\frac{1}{2}} \mathbb{E}[|g(u)|^2]^{\frac{1}{2}} ds du \]
\[ = \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|f(s) - f(t)|^2]^{\frac{1}{2}} ds \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|g(u)|^2]^{\frac{1}{2}} du. \]

Using Hölder inequality, we have
\[ \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|f(s) - f(t)|^2]^{\frac{1}{2}} ds \leq (\alpha + \beta)^{\frac{1}{2}} \left( \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|f(s) - f(t)|] ds \right)^{\frac{1}{2}}, \]
and
\[ \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|g(u)|^2]^{\frac{1}{2}} du \leq (\alpha + \beta)^{\frac{1}{2}} \left( \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|g(u)|^2] du \right)^{\frac{1}{2}}. \]

From Step 1, there is a null subset \( T' \) such that, for \( t \in [0, T] \), \( \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|f(s) - f(t)|^2]^{\frac{1}{2}} ds \longrightarrow 0 \) as \( r \) goes to 0. Since \( \mathbb{E}[|g(u)|^2] \) is Lebesgue integral on \([0, T]\), there is a null set \( T'' \subset [0, T] \) such that, for \( t \in [0, T] \),
\[ \lim_{r \to 0^+} \frac{1}{r} \int_{t-r}^{t+r} \mathbb{E}[|g(u)|^2] du = \mathbb{E}[|g(t)|^2]. \]

Define \( T_1 := T' \cup T'' \). It is a null subset, and for \( t \in [0, T] \),
\[ \lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \mathbb{E}[|f(s) - f(t)|^2]^{\frac{1}{2}} \mathbb{E}[|g(u)|^2]^{\frac{1}{2}} ds du = 0. \]

Similarly, there is another null set \( T_2 \) such that, for \( t \in [0, T] \),
\[ \lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r}^{t+r} \mathbb{E}[|g(u) - g(t)|^2]^{\frac{1}{2}} \mathbb{E}[|f(t)|^2]^{\frac{1}{2}} ds du = 0. \]

Let \( T := T_1 \cup T_2 \). This is the desired subset.

**Step 3.** Now we shall prove the final result.

Note that
\[ \frac{(\alpha + \beta)^2}{2} \mathbb{E}[f(t)g(t)] = \frac{1}{r^2} \int_{t-r}^{t+r} \int_{t-r}^{t+r} \mathbb{E}[f(t)g(t)] I_{[s \leq u]} ds du. \]
Thus,

\[
\frac{1}{r^2} \int_{t-r\beta}^{t+r\alpha} \int_{t-r\beta}^{t+r\alpha} \mathbb{E}[f(u)g(s)I_{s \leq u}] ds du - \frac{1}{r^2} \int_{t-r\beta}^{t+r\alpha} \int_{t-r\beta}^{t+r\alpha} \mathbb{E}[f(t)g(t)] ds du \leq \frac{1}{r^2} \int_{t-r\beta}^{t+r\alpha} \int_{t-r\beta}^{t+r\alpha} \mathbb{E}[f(u)g(s)] - \mathbb{E}[f(t)g(t)] I_{s \leq u} ds du.
\]

From Step 2, we know that there is a null subset \( T \) such that, for \( t \in [0, T] \), the above term tends to 0, which implies that

\[
\lim_{r \to 0^+} \frac{1}{r^2} \int_{t-r\beta}^{t+r\alpha} \int_{t-r\beta}^{t+r\alpha} \mathbb{E}[f(u)g(s)I_{s \leq u}] ds du = \frac{1}{r^2} \int_{t-r\beta}^{t+r\alpha} \int_{t-r\beta}^{t+r\alpha} \mathbb{E}[f(t)g(t)] ds du.
\]

Thus, we finish the proof. \( \square \)

References

1. Chen, Z., Epstein, L.: Ambiguity, risk, and asset returns in continuous time. Econometrica 70(4), 1403–1443 (2002)
2. Kushner, H.: Necessary conditions for continuous parameter stochastic optimization problems. SIAM J. Control Optim. 10(3), 550–565 (1972)
3. Bensoussan, A.: Lectures on Stochastic Control in Nonlinear Filtering and Stochastic Control. Lectures Notes in Mathematics. Springer, Berlin (1982)
4. Bismut, J.M.: An introductory approach to duality in optimal stochastic control. SIAM Rev. 20(1), 62–78 (1978)
5. Peng, S.: A general stochastic maximum principle for optimal control problems. SIAM J. Control Optim. 28(4), 966–979 (1990)
6. Peng, S.: Backward stochastic differential equations and applications to optimal control. Appl. Math. Optim. 27(2), 125–144 (1993)
7. Dokuchaev, N., Zhou, X.Y.: Stochastic controls with terminal contingent conditions. J. Math. Anal. Appl. 238(1), 143–165 (1999)
8. Shi, J., Wu, Z.: The maximum principle for fully coupled forward–backward stochastic control system. Acta Autom. Sin. 32(2), 161 (2006)
9. Wu, Z.: Maximum principle for optimal control problem of fully coupled forward–backward stochastic systems. Syst. Sci. Math. Sci. 3, 249–259 (1998)
10. Peng, S.: Open problems on backward stochastic differential equations. In: Chen, S., Li, X., Yong, J., Zhou, X.Y. (eds.) Control of Distributed Parameter and Stochastic Systems, pp. 265–273. Springer, New York (1999)
11. Wu, Z.: A general maximum principle for optimal control of forward–backward stochastic systems. Automatica 49(5), 1473–1480 (2013)
12. Yong, J.: Optimality variational principle for controlled forward–backward stochastic differential equations with mixed initial-terminal conditions. SIAM J. Control Optim. 48(6), 4119–4156 (2010)
13. Hu, M.: Stochastic global maximum principle for optimization with recursive utilities. Probab. Uncertain. Quant. Risk 2(1), 1–20 (2017)
14. Bell, D.J., Jacobson, D.H.: Singular Optimal Control Problems. Mathematics in Science and Engineering, vol. 117. Elsevier, London (1975)
15. Gabasov, R., Kirillova, F.: High order necessary conditions for optimality. SIAM J. Control 10(1), 127–168 (1972)
16. Kazemi-Dehkordi, M.: Necessary conditions for optimality of singular controls. J. Optim. Theory Appl. 43(4), 629–637 (1984)
17. Krener, A.J.: The high order maximal principle and its application to singular extremals. SIAM J. Control Optim. 15(2), 256–293 (1977)
18. Tang, S.: A second-order maximum principle for singular optimal stochastic controls. Discrete Contin. Dyn. Syst. Ser. B 14, 1581–1599 (2010)
19. Bonnans, J.F., Silva, F.J.: First and second order necessary conditions for stochastic optimal control problems. Appl. Math. Optim. 65(3), 403–439 (2012)
20. Zhang, H., Zhang, X.: Pointwise second-order necessary conditions for stochastic optimal controls, part I: the case of convex control constraint. SIAM J. Control Optim. 53(4), 2267–2296 (2015)
21. Zhang, H., Zhang, X.: Pointwise second-order necessary conditions for stochastic optimal controls, part II: the general case. SIAM J. Control Optim. 55(5), 2841–2875 (2017)
22. Frankowska, H., Zhang, H., Zhang, X.: First and second order necessary conditions for stochastic optimal controls. J. Differ. Equ. 262(6), 3689–3736 (2017)
23. Mou, L., Yong, J.: A variational formula for stochastic controls and some applications. Pure Appl. Math. Q. 3(2), 539–567 (2007)
24. Li, X., Yao, Y.: Maximum principle of distributed parameter systems with time lags. In: Kappel, F., Kunisch, K., Schappacher, W. (eds.) Distributed Parameter Systems, pp. 410–427. Springer, Berlin (1985)
25. Evans, L.C.: Partial Differential Equations. American Mathematical Society, Providence (2010)
26. Stein, E.M., Shakarchi, R.: Real Analysis. Princeton Lectures in Analysis III. Princeton University Press, Princeton (2005)
27. Framstad, N.C., Øksendal, B., Sulem, A.: Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. J. Optim. Theory Appl. 121(1), 77–98 (2004)
28. Andersson, D., Djehiche, B.: A maximum principle for SDEs of mean-field type. Appl. Math. Optim. 63(3), 341–356 (2011)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.