Pointwise Second Order Maximum Principle for Stochastic Recursive Singular Optimal Controls Problems

Liangquan Zhang
1. School of Science,
Beijing University of Posts and Telecommunications,
Beijing 100876, China

January 7, 2019

Abstract

This paper studies two kinds of singular optimal controls (SOCs for short) problems where the systems governed by forward-backward stochastic differential equations (FBSDEs for short), in which the control has two components: the regular control, and the singular one. Both drift and diffusion terms may involve the regular control variable. The regular control domain is postulated to be convex. Under certain assumptions, in the framework of the Malliavin calculus, we derive the pointwise second-order necessary conditions for stochastic SOC in the classical sense. This condition is described by two adjoint processes, a maximum condition on the Hamiltonian supported by an illustrative example. A new necessary condition for optimal singular control is obtained as well. Besides, as a by-product, a verification theorem for SOC is derived via viscosity solutions without involving any derivatives of the value functions. It is worth pointing out that this theorem has wider applicability than the restrictive classical verification theorems. Finally, we focus on the connection between the maximum principle and the dynamic programming principle for such SOC problems without the assumption that the value function is smooth enough.

AMS subject classifications: 93E20, 60H15, 60H30.

Key words: Dynamic programming principle (DPP for short), Forward-backward stochastic differential equations (FBSDEs for short), Malliavin calculus, Maximum principle (MP for short), Singular optimal controls, Viscosity solution, Verification theorem.

1 Introduction

Singular stochastic control problem is a fundamental topic in fields of stochastic control. This problem was first introduced by Bather and Chernoff [13] in 1967 by considering a simplified model for the control of a spaceship. It was then found that there was a

*The author acknowledges the financial support partly by the National Nature Science Foundation of China (Grant No. 11701040, 11871010, 61871058 & 61603049) and Innovation Foundation of BUPT for Youth (No. 500417024 & 505018304). E-mail: xiaoquan51011@163.com.
connection between the singular control and optimal stopping problem. This link was established through the derivative of the value function of this initial singular control problem and the value function of the corresponding optimal stopping problem. Subsequently, it was considered by Benés, Shepp, Witzenshausen (see [6]) and Karatzas and Shreve (see [61, 62, 63, 64, 65]). In contrast to classical control problems, singular control problems admit both of the continuity of the cumulative displacement of the state caused by control and the jump of one in impulsive control problems, between which it is either constant or absolutely continuous.

In the classical singular control problems, the state process is governed by an \( n \)-dimensional SDE of the following type:

\[
\begin{align*}
\{ & \text{d}X_{t,x}^{t,x,v,\xi} = b\left(s, X_{t,x}^{t,x,v,\xi}, v_s\right) \text{d}s + \sigma\left(s, X_{t,x}^{t,x,v,\xi}, v_s\right) \text{d}W_s + G_s \text{d}\xi_s, \\
X_{t,x}^{t,x,v,\xi} = x, & 0 \leq t \leq s \leq T,
\end{align*}
\]

(1)
on some filtered probability space \((\Omega, \mathcal{F}, P)\), where \(b(\cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n, \sigma(\cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{n \times d}, G(\cdot) : [0, T] \to \mathbb{R}^{n \times m}\) are given deterministic functions, \((W_s)_{s \geq 0}\) is a \( d \)-dimensional Brownian motion, \((x, t)\) are initial time and state, \(v(\cdot) : [0, T] \to \mathbb{R}^k\) is a regular control process, and \(\xi(\cdot) : [0, T] \to \mathbb{R}^m\), with nondecreasing left-continuous with right limits stands for the singular control \(1\) (SC for short). To avoid the risk of confusion, we shall introduce the other definitions of singular control in various senses. Indeed, they are just a coincidence of terminology usage.

The aim is to minimize the cost functional:

\[
J(t, x; v, \xi) = \mathbb{E}\left[ \int_t^T l\left(s, X_{s}^{t,x,v,\xi}, v_s\right) \text{d}s + \int_t^T K_s \text{d}\xi_s \right],
\]

(2)
where

\[
l(\cdot, \cdot, \cdot) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R},
K(\cdot) : [0, T] \to \mathbb{R}^m_+ \triangleq \{x \in \mathbb{R}^m : x_i \geq 0, i = 1, \ldots m\}
\]
are given deterministic functions, where \(l(\cdot)\) represents the running cost rate of the problem and \(K(\cdot)\) the cost rate of applying the singular control.

There are four approaches employed: the first one is based on the theory of partial differential equations and on variational arguments, and can be found in the works of Alvarez [1, 2], Chow, Menaldi, and Robin [29], Karatzas [62], Karatzas and Shreve [65], and Menaldi and Taksar [74]. The second one is related to probabilistic methods; see Baldursson [9], Boetius [10, 11], Boetius and Kohlmann [12], El Karoui and Karatzas [37, 38], Karatzas [61], and Karatzas and Shreve [63, 64]. Third, the DPP, has been studied in a general context, for example, by Boetius [11], Haussmann and Suo [49], Fleming and Soner [39], and Zhu [106]. At last the maximum principle for optimal singular controls (see, for example, Cadenillas and Haussmann [26], Dufour and Miller [33], Dahl and Øksendal [34] see references therein). The existence for optimal singular control can be found in Haussmann, Suo [19], Dufour, Miller [32] and Fu, Host [12] via different approaches.

\footnote{Because the measure \(\text{d}\xi_s\) may be singular with respect to the Lebesgue measure \(\text{d}s\). More details can be seen in Property \(1\) below.}
It is necessary to point out that singular control problems are largely used in diverse areas such as mathematical finance (see Baldursson and Karatzas [14], Chiarolla and Hausmann [27], Kobila [66], and Karatzas and Wang [67], Davis and Norman [31]), manufacturing systems (see, Shreve, Lehoczky, and Gaver [91]), and queuing systems (see Martins and Kushner [75]). Particularly, the application of H-J-B inequality in finance can be seen in Pagès and Possamaï [88].

Completely different from the singular control introduced above, to the best of our knowledge, there are two other types of singular optimal controls, in which the first-order necessary conditions turn out to be trivial. We list briefly as follows:

- **Singular optimal control in the classical sense (SOCCS for short)**, is the optimal control for which the gradient and the Hessian of the corresponding Hamiltonian with respect to the control variable vanish/degenerate.

- **Singular optimal control in the sense of Pontryagin-type maximum principle (SOCSPMP for short)**, is the optimal control for which the corresponding Hamiltonian is equal to a constant in the control region.

When an optimal control is singular in certain senses above (SOCCS and SOCSPMP), usually the first-order necessary condition could not carry sufficient information for the further theoretical analysis and numerical computation, and consequently it is necessary to investigate the second order necessary conditions. In the deterministic setting, reader can refer many articles in this direction (see [5, 40, 44, 45, 58, 59, 60] and references therein).

As for the second-order necessary conditions for stochastic singular optimal controls (SOCCS and SOCSPMP), there are some work should be mentioned, for instance [102, 103] (noting that singular control ξ(·) in these articles does not appear in systems or \( G_s \equiv 0 \) in (1)). Tang [93] obtained a pointwise second order maximum principle for stochastic singular optimal controls in the sense of the Pontryagin-type maximum principle whenever the control variable \( u \) does not enter into the diffusion term. Meanwhile, Tang addressed an integral-type second-order necessary condition for stochastic optimal controls with convex control constraints. Recently, Dong and Meng [36] study this issue for FBSDEs, in which the control variable does not entre into diffusion terms. Zhang and Zhang [102] also establish certain pointwise second-order necessary conditions for stochastic singular (SOCCS) optimal controls, in which both drift and diffusion terms in may depend on the control variable \( u \) with convex control region \( U \) by making use of Malliavin calculus technique. Later, adopting the same idea but with large complicated analysis, Zhang et al. [103] deepen this research for the general case when the control region is nonconvex. More information can be seen in Frankowska et al. [43, 104, 105].

The theory of **backward stochastic differential equation** (BSDE for short) can be traced back to Bismut [3, 4] who studied linear BSDE motivated by stochastic control problems. Pardoux and Peng 1990 [84] proved the well-posedness for nonlinear BSDE. Duffie and Epstein (1992) introduced the notion of recursive utilities in continuous time, which is actually a type of BSDE where the generator \( f \) is independent of \( z \). El Karoui et al.
extended the recursive utility to the case where \( f \) contains \( z \). The term \( z \) can be interpreted as an ambiguity aversion term in the market (see Chen and Epstein 2002 [30]). Particularly, the celebrated Black-Scholes formula indeed provided an effective way of representing the option price (which is the solution to a kind of linear BSDE) through the solution to the Black-Scholes equation (parabolic partial differential equation actually). Since then, BSDE has been extensively studied and used in the areas of applied probability and optimal stochastic controls, particularly in financial engineering (see [56]).

By means of BSDE, Peng (1990) [82] considered the following type of stochastic optimal control problem. Minimize a cost function

\[
J(v) = \mathbb{E} \int_0^T l(x_t, v_t) \, dt + \mathbb{E} (h_T),
\]

subject to

\[
\begin{align*}
dx_t &= g(t, x_t, v_t) \, dt + \sigma(t, x_t, v_t) \, dW_t, \\
x_0 &= x,
\end{align*}
\]

(3)

over an admissible control domain which need not be convex, and the diffusion coefficients depends on the control variable. In his paper, by spike variational method and the second order adjoint equations, Peng [82] obtained a general stochastic maximum principle for the above optimal control problem. It was just the adjoint equations in stochastic optimal control problems that motivated the famous theory of BSDE (see [84]). Later Peng [83] studied a stochastic optimal control problem where state variables are described by the system of FBSDEs:

\[
\begin{align*}
dx_t &= f(t, x_t, v_t) \, dt + \sigma(t, x_t, v_t) \, dW_t, \\
dy_t &= g(t, x_t, y_t, z_t, v_t) \, dt + z_t \, dW_t, \\
x_0 &= x, y_T &= y,
\end{align*}
\]

(4)

where \( x \) and \( y \) are given deterministic constants. The optimal control problem is to minimize the cost function

\[
J(v) = \mathbb{E} \left[ \int_0^T l(t, x_t, y_t, z_t, v_t) \, dt + h(x_T) + \gamma (y_0) \right],
\]

over an admissible control domain which is convex. Xu [99] studied the following non-fully coupled forward-backward stochastic control system

\[
\begin{align*}
dx_t &= f(t, x_t, y_t, z_t, v_t) \, dt + \sigma(t, x_t, y_t, z_t, v_t) \, dW_t, \\
dy_t &= -g(t, x_t, y_t, z_t, v_t) \, dt + z_t \, dB_t, \\
x_0 &= x, y_T &= h(x_T).
\end{align*}
\]

(5)

The optimal control problem is to minimize the cost function

\[
J(v) = \mathbb{E} \gamma (y_0),
\]

over \( U_{ad} \), but the control domain is non-convex. Wu [77] firstly gave the maximum principle for optimal control problem of fully coupled forward-backward stochastic system

\[
\begin{align*}
dx_t &= f(t, x_t, y_t, z_t, v_t) \, dt + \sigma(t, x_t, y_t, z_t, v_t) \, dB_t, \\
dy_t &= -g(t, x_t, y_t, z_t, v_t) \, dt + z_t \, dB_t, \\
x_0 &= x, y_T &= \xi,
\end{align*}
\]

(6)
where $\xi$ is a random variable and the cost function

$$J(v(t)) = \mathbb{E} \left[ \int_0^T L(t, x_t, y_t, z_t, v_t) \, dt + \Phi(x_T) + h(y_0) \right].$$

The optimal control problem is to minimize the cost function $J(v(t))$ over an admissible control domain which is convex. Ji and Zhou [54] obtained a maximum principle for fully coupled forward-backward stochastic system with terminal state constraints. Shi and Wu [92] studied the maximum principle for fully coupled forward-backward stochastic system with control domain which is convex. Ji and Zhou [54] obtained a maximum principle for the optimal control problem to minimize the cost function

$$\min \{ J(v(t)) \},$$

where

$$J(v(t)) = \mathbb{E} \left[ \int_0^T l(t, x_t, y_t, z_t, v_t) \, dt + \Phi(x_T) + \gamma(y_0) \right].$$

The control domain is non-convex but the forward diffusion does not contain the control variable. Later Hu, Ji and Xue consider the MP for fully coupled FBSDEs (see [51, 52]).

Then, the second-order variational equation for the BSDE and the maximum principle are independently established the maximum principle for the recursive stochastic optimal control problem (noting the diffusion term containing control variable with non-convex control region). Nonetheless, the maximum principle derived by these method involves two unknown parameters. Therefore, the hard questions raise as follows: What is the second-order variational equation for the BSDE? How to obtain the second-order adjoint equation since the quadratic form with respect to the variation of $z$. All of which seem to be extremely complicated.

Hu [50] overcomes the above difficulties by introducing two new adjoint equations. Then, the second-order variational equation for the BSDE and the maximum principle are obtained. The main difference of his variational equations with those in Peng [82] consists in the term $\langle p(t), \delta \sigma(t) \rangle I_{E_t}(t)$ in the variation of $z$. Due to the term $\langle p(t), \delta \sigma(t) \rangle I_{E_t}(t)$ in the variation of $z$, Hu obtained a global maximum principle which is novel and different from that in Wu [98], Yong [101] and previous work, which solves completely Peng’s open problem. Furthermore, Hu’s maximum principle is stronger than the one in Wu [98], Yong [101].

Motivated by above work, in this paper, we consider singular controls problem of the following type:

$$\begin{align*}
\text{d}X_{s}^{t,x,v,\xi} &= b \left( \frac{s}{X_{s}^{t,x,v,\xi}}, v_{s} \right) \text{d}s + \sigma \left( \frac{s}{X_{s}^{t,x,v,\xi}}, v_{s} \right) \text{d}W_{s} + G_{s} \text{d}\xi_{s}, \\
\text{d}Y_{s}^{t,x,v,\xi} &= -f \left( \frac{t}{X_{s}^{t,x,v,\xi}}, Y_{s}^{t,x,v,\xi}, Z_{s}^{t,x,v,\xi}, v_{s} \right) \text{d}s + Z_{s}^{t,x,v,\xi} \text{d}W_{s} - K_{s} \text{d}\xi_{s} \\
X_{t}^{t,x,v,\xi} &= x, \quad Y_{T}^{t,x,v,\xi} = \Phi \left( X_{T}^{t,x,v,\xi} \right), \quad 0 \leq t \leq s \leq T,
\end{align*}$$

Motivated by above work, in this paper, we consider singular controls problem of the following type:
with the similar cost functional

$$J(t, x; v, \xi) = Y_t^t, x; v, \xi |_{s=t}.$$  \hspace{1cm} (9)

We postulate that $K$ is a deterministic matrix in Eq. (8). The justification will be given in Remark 28 below. Wang \[95\] firstly introduced and studied a class of singular control problems with recursive utility, where the cost function is determined by BSDE. Under certain assumptions, the author proved that the value function is a nonnegative, convex solution of the H-J-B equation. However, FBSDEs in Wang \[95\] do not contain the regular control and the generator is not general case. Later, Ji and Xue also studied the singular control problem for FBSDEs, in which the singular control does not appear in BSDE and they derive a maximum principle for SOC. In our work, using some properties of the BSDE and analysis technique, we expand the extension of the MP for SOC to the recursive control problem in Zhang and Zhang \[102\]. To the best of our knowledge, such singular optimal controls problems of FBSDEs (8) via two kinds of singular controls have not been explored before. We shall establish some pointwise second-order necessary conditions for stochastic optimal controls of FBSDEs. Both drift and diffusion terms may contain the control variable $u$, and we assume that the control region $U$ is convex. We also consider the pointwise second-order necessary condition, which is easier to verify in practical applications.

As claimed in \[102\], quite different from the deterministic setting, there exist some essential difficulties in deriving the pointwise second-order necessary condition from an integral-type one whenever the diffusion term depends on the control variable, even for the case of convex control domain. We overcome these difficulties by means of some technique from the Malliavin calculus. For general case, namely, the control region is nonconvex can be found in \[103\].

In this paper, we are interested in studying singular optimal controls for FBSDEs (8). Compared with above literature, our paper has several new features. The novelty of the formulation and the contribution in this paper may be stated as follows:

- Our control systems in this paper are governed by FBSDEs which exactly extends the work of Zhang and Zhang \[102\] to utilities. Our work is the first time to establish the pointwise second order necessary condition for stochastic singular optimal control in the classical sense for FBSDEs, a new necessary condition for singular control is involved as well. In this sense, our paper actually considers two kinds of singular controls problems simultaneously, which is interesting to deepen this research.

- We derive a new verification theorem for optimal singular controls via viscosity solution, which responses to the question raised in Zhang \[109\]. Meanwhile, we study the relationship between the adjoint equations derived and value function, which extends the smooth case considered by Cadenillas and Haussmann \[26\] to the framework of viscosity solution for stochastic recursive systems.

The rest of this paper is organized as follows: After some preliminaries in the second section, we are devoted the third section to the MP for two kinds of singular optimal controls. A concrete example is concluded with as well. Then, in Section 4 we study
the verification theorem for singular optimal controls via viscosity solutions. Finally, we establish the relationship between the DPP and MP for viscosity solution. Some proofs of lemmas will be displayed in Appendix A.

2 Preliminaries and Notations

Throughout this paper, we denote by $\mathbb{R}^n$ the space of $n$-dimensional Euclidean space, by $\mathbb{R}^{n \times d}$ the space the matrices with order $n \times d$. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ be a complete filtered probability space on which a 1-dimensional standard Brownian motion $W(\cdot)$ is defined, with $\{\mathcal{F}_t\}_{t \geq 0}$ being its natural filtration, augmented by all the $P$-null sets. Given a subset $U$ (compact) of $\mathbb{R}^k$, we will denote $U_1 \times U_2$, separately, the class of measurable, adapted processes $(v, \xi): [0, T] \times \Omega \to U \times [0, \infty)^m$, with $\xi$ nondecreasing left-continuous with right limits and $\xi_0 = 0$, moreover, $E \left[ \sup_{0 \leq t \leq T} |u_t|^2 + |\xi_t|^2 \right] < \infty$. $\xi$ is called singular control. For each $t > 0$, we denote by $\{\mathcal{F}^\xi_t, t \leq s \leq T \}$ the natural filtration of the Brownian motion $\{W_s - W_t, t \leq s \leq T\}$, augmented by the $P$-null sets of $\mathcal{F}$. $\top$ appearing as superscript denotes the transpose of a matrix. In what follows, $C$ represents a generic constant, which can be different from line to line.

Property 1 (Singular control). Let $D ([0, T]; \mathbb{R}^m)$ be the space of all functions $\xi: [0, T] \to \mathbb{R}^m$ that are right limit with left continuous. For $\xi \in D$, we define the total variation of on $[0, T]$ by $\sum_{i=1}^n |\xi^i||_{[0,T]}$ where $|\xi^i|_{[0,T]}$ is the total variation of the $i$th component of $\xi$ on $[0, T]$ in the usual sense, that is $|\xi^i|_{[0,T]} = \int_0^T |\xi^i|$. Define $SBV ([0, T]; \mathbb{R}^m) \triangleq \left\{ \xi \in D | |\xi||_{[0,T]} < \infty \right\}$. For $\xi \in D$, we define $\Delta \xi (s) = \xi (s +) - \xi (s)$ and set $\Theta (\xi) = \left\{ s \in [0, T] | \Delta \xi (s) \neq 0 \right\}$. Then, the pure jump part of $\xi$ is defined by $\xi^J (t) = \sum_{0 \leq s \leq t} \Delta \xi (s)$, and the continuous part is $\xi^C (t) = \xi (t) - \xi^J (t)$. Note that $\xi^C (t)$ is bounded variation and differentiable almost everywhere, and we have by Lebesgue decomposition Theorem that $\xi^C (t) = \xi^{ac} (t) + \xi^{sc} (t)$, $t \in [0, T]$, where $\xi^{ac} (t)$ is called the absolutely continuous part of $\xi$, and $\xi^{sc}$ the singularly continuous part of $\xi$. Thus, we obtain that

$$\xi (t) = \xi^{ac} (t) + \xi^{sc} (t) + \xi^J (t), \ t \in [0, T], \ unique!$$

Now consider the following cases:

1) If we assume that $\xi^J (t) \equiv 0, t \in [0, T]$, then the singular control reduces to a standard control problem, since we take $\xi^{ac} (t) + \xi^{sc} (t)$ as a new control variable.

2) If we assume that $\xi^{ac} (t) + \xi^{sc} (t) \equiv 0, t \in [0, T]$, then the singular control performs a special form of a pure jump process, so-called impulse control.

We now introduce the following spaces of processes:

$$S^2 (0, T; \mathbb{R}) \triangleq \left\{ \mathbb{R}^n\text{-valued } \mathcal{F}_t\text{-adapted process } \phi (t); E \left[ \sup_{0 \leq t \leq T} |\phi_t|^2 \right] < \infty \right\},$$

$$M^2 (0, T; \mathbb{R}) \triangleq \left\{ \mathbb{R}^n\text{-valued } \mathcal{F}_t\text{-adapted process } \varphi (t); E \left[ \int_0^T |\varphi_t|^2 \, dt \right] < \infty \right\},$$
and denote $\mathcal{N}^2[0, T] = \mathcal{S}^2(0, T; \mathbb{R}^n) \times \mathcal{S}^2(0, T; \mathbb{R}) \times \mathcal{M}^2(0, T; \mathbb{R}^n)$. Clearly, $\mathcal{N}^2[0, T]$ forms a Banach space.

For any $v(\cdot) \times \xi(\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2$, we study the stochastic control systems governed by FBSDEs of the following type with two adapted control processes:

$$
\begin{aligned}
&\frac{dX^t_{s,v,\xi}}{dt} = b\left(s, X^t_{s,v,\xi}, v_s\right) ds + \sigma\left(s, X^t_{s,v,\xi}, v_s\right) dW_s + G_s d\xi_s, \\
&\frac{dY^t_{s,v,\xi}}{dt} = -f\left(s, X^t_{s,v,\xi}, Y^t_{s,v,\xi}, Z^t_{s,v,\xi}, v_s\right) ds + Z^t_{s,v,\xi} dW_s - K_s d\xi_s, \\
&X^t_{0,v,\xi} = x, \quad Y^t_{T,v,\xi} = \Phi\left(X^t_{T,v,\xi}\right), \quad 0 \leq t \leq s \leq T.
\end{aligned}
$$

**Definition 2.** For any $v(\cdot) \times \xi(\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2$, a triple of processes

$$
\left(X^t_{t,v,\xi}, Y^t_{t,v,\xi}, Z^t_{t,v,\xi}\right) : [0, T] \times \Omega \to \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d
$$

is called an adapted solution of the FBSDEs (10), if

$$
\left(X^t_{t,v,\xi}, Y^t_{t,v,\xi}, Z^t_{t,v,\xi}\right) \in \mathcal{N}^2[0, T],
$$

and it satisfies (10), $\mathbb{P}$-almost surely.

We assume that the following conditions hold.

(A1) The coefficients $b : [0, T] \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, are twice continuously differentiable with respect to $x$; $b, b_x, b_{xx}, \sigma, \sigma_x, \sigma_{xx}$ are continues in $(x, u)$; $b_x, b_{xx}, \sigma_x, \sigma_{xx}$ are bounded by $C(1 + |x| + |u|)$ for some positive constant $C$. Moreover, for any $(t, x_1, u_1), (t, x_2, u_2) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^k$,

$$
|b(t, 0, x)| + |\sigma(t, 0, u)| \leq C;
$$

$$
\left|b_{(x,u)}(t, x_1, u_1) - b_{(x,u)}(t, x_2, u_2)\right| \leq C \left(|x_1 - x_2| + |u_1 - u_2|\right),
$$

$$
\left|\sigma_{(x,u)}(t, x_1, u_1) - \sigma_{(x,u)}(t, x_2, u_2)\right| \leq C \left(|x_1 - x_2| + |u_1 - u_2|\right),
$$

(A2) The coefficients $f : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$, $\Phi : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable with respect to $(x, y, z)$. $f, Df, D^2f$ are continuous in $(x, y, z), f, Df, D^2f$ are bounded, where $Df$ is the gradient of $f$ with respect to $(x, y, z), D^2f$ is the Hessian matrix of $f$ with respect to $(x, y, z)$. There exists constant $C > 0$ and $0 < \mu < 1$, such that for any $(t, x_1, y_1, z_1, u_1), (t, x_2, y_2, z_2, u_2) \in [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^k$,

$$
|f(t, x, y, z, u)| \leq C \left(1 + |x| + |y| + |z| + |u|\right),
$$

$$
|f_x(t, x, y, z, u)| + |f_y(t, x, y, z, u)| + |f_z(t, x, y, z, u)| + |f_u(t, x, y, z, u)| \leq C \left(1 + |x| + |y| + |z| + |u|\right),
$$

$$
|f_{xx}(t, x, y, z, u)| + |f_{yy}(t, x, y, z, u)| + |f_{zz}(t, x, y, z, u)| + |f_{uu}(t, x, y, z, u)|
$$

$$
+ |f_{xy}(t, x, y, z, u)| + |f_{xz}(t, x, y, z, u)| + |f_{yz}(t, x, y, z, u)| + |f_{yu}(t, x, y, z, u)| + |f_{zu}(t, x, y, z, u)| + |f_{uz}(t, x, y, z, u)| + |f_{vu}(t, x, y, z, u)| \leq C,
$$

8
where $E$ solution. Let

Under the assumptions (Optimal Control)

Definition 3 optimal 5-tuple of optimal singular problem.

We are interested in the existence and uniqueness, since the value function (12) is defined by the solution of controlled BSDE (10), so from (A1)-(A2), for any $v(\cdot) \times \xi (\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2$, it is easy to check that FBSDEs (10) admit a unique $\mathcal{F}_t$-adapted solution denoted by the triple

$(X^{t,x;v,\xi}, Y^{t,x;v,\xi}, Z^{t,x;v,\xi}) \in \mathcal{N}^2 [0, T]$

(See Pardoux and Peng [51]).

Like Peng [51], given any control processes $v(\cdot) \times \xi (\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2$, we introduce the following cost functional:

\[ J(t, x; v(\cdot), \xi (\cdot)) = Y^{t,x;v,\xi}_s \bigg|_{s=t}, \quad (t, x) \in [0, T] \times \mathbb{R}^n. \] (11)

We are interested in the value function of the stochastic optimal control problem

\[
\begin{align*}
  u(t, x) &= J(t, x; \hat{v}(\cdot), \hat{\xi}(\cdot)) \\
  &= \text{ess inf}_{v(\cdot) \times \xi (\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2} J(t, x; v(\cdot), \xi (\cdot)), \quad (t, x) \in [0, T] \times \mathbb{R}^n. \tag{12}
\end{align*}
\]

Since the value function (12) is defined by the solution of controlled BSDE (10), so from the existence and uniqueness, $u$ is well-defined.

Any $\hat{v}(\cdot) \times \hat{\xi}(\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2$ satisfying (12) is called an optimal control pair of optimal singular problem, and the corresponding state processes, denoted by $(X^{t,x;\hat{v},\hat{\xi}}, Y^{t,x;\hat{v},\hat{\xi}}, Z^{t,x;\hat{v},\hat{\xi}})$, is called optimal state process. We also refer to $(X^{t,x;\hat{v},\hat{\xi}}, Y^{t,x;\hat{v},\hat{\xi}}, Z^{t,x;\hat{v},\hat{\xi}}, \hat{v}(\cdot), \hat{\xi}(\cdot))$ as an optimal 5-tuple of optimal singular problem.

**Definition 3** (Optimal Control). Any admissible controls $\hat{v}(\cdot) \times \hat{\xi}(\cdot) \in \mathcal{U}_1 \times \mathcal{U}_2$, are called optimal, if $\hat{v}(\cdot) \times \hat{\xi}(\cdot)$ attains the minimum of $J(u(\cdot) \times \xi (\cdot))$.

We have the following properties of BSDE:

**Lemma 4.** Under the assumptions (A1)-(A2), FBSDEs (10) admit a unique strong adapted solution. Let $(y^i, z^i), i = 1, 2$, be the solution to the following

\[
y^i(t) = \xi^i + \int_t^T f^i(s, y^i(s), z^i(s)) \, ds - \int_t^T z^i(s) \, dW_s, \tag{13}
\]

where $\mathbb{E} [\xi^i]^2 < \infty$, $f^i(s, y^i, z^i)$ satisfies the conditions (A2), and

\[
\mathbb{E} \left[ \left( \int_t^T |f^i(s, y^i(s), z^i(s))| \, ds \right)^\beta \right] < \infty.
\]
Then, for some $\beta \geq 2$, there exists a positive constant $C_{\beta}$ such that

$$
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |y^1(t) - y^2(t)|^\beta + \left( \int_0^T \left| z^1(s) - z^2(s) \right|^2 \, ds \right)^{\frac{\beta}{2}} \right] 
\leq C_{\beta} \mathbb{E} \left[ |\xi^1 - \xi^2|^\beta + \left( \int_0^T \left| f^1(s, y^1(s), z^1(s)) - f^2(s, y^2(s), z^2(s)) \right| \, ds \right)^{\beta} \right].
$$

Particularly, whenever putting $\xi^2 = 0$, $f^2 = 0$, one has

$$
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |y^1(t)|^\beta + \left( \int_0^T \left| z^1(s) \right|^2 \, ds \right)^{\frac{\beta}{2}} \right] \leq C_{\beta} \mathbb{E} \left[ |\xi^1|^\beta + \left( \int_0^T \left| f^1(s, 0, 0) \right| \, ds \right)^{\beta} \right].
$$

The proof can be seen in Briand [15].

We shall recall the following basic result on BSDE. We begin with the well-known comparison theorem (see Barles, Buckdahn, and Pardoux [16], Proposition 2.6).

**Lemma 5** (Comparison theorem). Let $(y^i, z^i), \ i = 1, 2,$ be the solution to the following

$$
y^i(t) = \xi^i + \int_t^T f^i(s, y^i_s, z^i_s) \, ds - \int_t^T z^i_s \, dW_s, \quad (14)
$$

where $\mathbb{E} \left[ |\xi^i|^2 \right] < \infty, f^i(s, y^i, z^i)$ satisfies the conditions (A2), $i = 1, 2$. Under assumption (A2), BSDE (14) admits a unique adapted solution $(y^i, z^i)$, respectively, for $i = 1, 2$. Furthermore, if

(i) $\xi^1 \geq \xi^2,$ a.s.;

(ii) $f^1(t, y, z) \geq f^2(t, y, z), \ a.e.,$ for any $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d.$

Then, we have: $y^1_t \geq y^2_t,$ a.s., for all $t \in [0, T].$

Now let us recall briefly the notion of differentiation on Wiener space (see the expository papers by Nualart 1995 [76], Nualart and Pardoux [77] and Ocone 1988 [79]).

- $C^k_b(\mathbb{R}^k, \mathbb{R}^q)$ will denote the set of functions of class $C^k$ from $\mathbb{R}^k$ into $\mathbb{R}^q$ whose partial derivatives of order less than or equal to $k$ are bounded.

- Let $\mathcal{S}$ denote the set of random variables $\xi$ of the form

$$
\xi = \varphi(W(h^1), W(h^2), \ldots, W(h^k)),
$$

where $\varphi \in C^\infty_b(\mathbb{R}^k, \mathbb{R}), h^1, h^2, \ldots h^k \in L^2([0, T]; \mathbb{R}^n)$, and $W(h^i) = \int_0^T \left< h^i_s, dW(s) \right>.$

- If $\xi \in \mathcal{S}$ is of the above form, we define its derivative as being the $n$-dimensional process

$$
\mathcal{D}_\theta \xi = \sum_{j=1}^k \frac{\partial}{\partial x_j} \varphi \left( W(h^1), W(h^2), \ldots, W(h^k) \right) h^i_\theta, \ 0 \leq \theta \leq T.
$$

For $\xi \in \mathcal{S}, p > 1$, we define the norm

$$
\|\xi\|_{1,p} = \left[ \mathbb{E} \left\{ |\xi|^p + \left( \int_0^T |\mathcal{D}_\theta \xi|^2 \, d\theta \right)^{\frac{p}{2}} \right\} \right]^{\frac{1}{p}}.
$$
It can be shown (Nualart 1995) that the operator \( \mathcal{D} \) has a closed extension to the space \( \mathbb{D}^{1,p} \), the closure of \( \mathcal{S} \) with respect to the norm \( \| \cdot \|_{1,p} \). Observe that if \( \xi \) is \( \mathcal{F}_t \)-measurable, then \( \mathcal{D}_t \xi = 0 \) for \( \theta \in (t,T] \). We denote by \( \mathcal{D}_t \xi \), the \( i \)th component of \( \mathcal{D}_t \xi \).

Let \( \mathbb{L}^{1,p} (\mathbb{R}^d) \) denote the set of \( \mathbb{R}^d \)-valued progressively measurable processes \( \{ u(t,\omega), 0 \leq t \leq T; \omega \in \Omega \} \) such that

- For a.e. \( t \in [0,T] \), \( u(t,\cdot) \in \mathbb{D}^{1,p} (\mathbb{R}^n) \);
- \((t,\omega) \rightarrow \mathcal{D}u(t,\omega) \in (L^2([0,T]))^{n \times d} \) admits a progressively measurable version;
- We have
  \[
  \|u\|_{1,p} = \mathbb{E} \left[ \left( \int_0^T |u(t)|^2 \, dt \right)^{\frac{p}{2}} + \left( \int_0^T \int_0^T |\mathcal{D}_t u(t)|^2 \, d\theta dt \right)^{\frac{p}{2}} \right] < +\infty.
  \]

Note that for each \((\theta,t,\omega)\), \( \mathcal{D}_t u(t,\omega) \) is an \( n \times d \) matrix. Hence, \( \|\mathcal{D}_t u(t)\|^2 = \sum_{i,j} |\mathcal{D}_{t,i} u_{j}(t)|^2 \).

Obviously, \( \mathcal{D}_t u(t,\omega) \) is defined uniquely up to sets of \( d\theta \otimes dt \otimes dP \) measure zero. Moreover, denote by \( \mathbb{L}^{1,p}_F (\mathbb{R}^d) \) the set of all adapted processes in \( \mathbb{L}^{1,p} (\mathbb{R}^d) \).

We define the following notations from Zhang and Zhang [102]:

\[
\mathbb{L}^{1,p}_{2,+} (\mathbb{R}^d) := \left\{ \varphi(\cdot) \in \mathbb{L}^{1,p} (\mathbb{R}^d) \mid \exists \mathcal{D}^+ \varphi(\cdot) \in L^2([0,T] \times \Omega; \mathbb{R}^n) \text{ such that} \right. \\
\int_0^T f_\varepsilon(s) := \sup_{s \leq \varepsilon \leq (s+t) \in [0,T]} \mathbb{E} |\mathcal{D}_s \varphi(t) - \mathcal{D}^+ \varphi(s)|^2 < \infty, \text{ a.e. } s \in [0,T], \\
\left. f_\varepsilon(s) \text{ is measurable on } [0,T] \text{ for any } \varepsilon > 0, \text{ and } \lim_{\varepsilon \to 0^+} \int_0^T f_\varepsilon(s) \, ds = 0 \right\};
\]

and

\[
\mathbb{L}^{1,p}_{2,-} (\mathbb{R}^d) := \left\{ \varphi(\cdot) \in \mathbb{L}^{1,p} (\mathbb{R}^d) \mid \exists \mathcal{D}^- \varphi(\cdot) \in L^2([0,T] \times \Omega; \mathbb{R}^n) \text{ such that} \right. \\
\int_0^T g_\varepsilon(s) := \sup_{s \leq \varepsilon \leq (s+t) \in [0,T]} \mathbb{E} |\mathcal{D}_s \varphi(t) - \mathcal{D}^- \varphi(s)|^2 < \infty, \text{ a.e. } s \in [0,T], \\
\left. g_\varepsilon(s) \text{ is measurable on } [0,T] \text{ for any } \varepsilon > 0, \text{ and } \lim_{\varepsilon \to 0^+} \int_0^T g_\varepsilon(s) \, ds = 0 \right\}.
\]

Denote

\[
\mathbb{L}^{1,p}_{2,\pm} (\mathbb{R}^d) = \mathbb{L}^{1,p}_{2,+} (\mathbb{R}^d) \cap \mathbb{L}^{1,p}_{2,-} (\mathbb{R}^d).
\]

For any \( \varphi(\cdot) \in \mathbb{L}^{1,p}_{2,\pm} (\mathbb{R}^d) \), denote \( \nabla \varphi(\cdot) = \mathcal{D}^+ \varphi(\cdot) + \mathcal{D}^- \varphi(\cdot) \). Whenever \( \varphi \) is adapted, it follows that \( \mathcal{D}_s \varphi(t) = 0 \) for \( t < s \). Furthermore, \( \nabla \varphi(\cdot) = \mathcal{D}^+ \varphi(\cdot) \) since \( \mathcal{D}^- \varphi(\cdot) = 0 \). Put \( \mathbb{L}^{1,p}_{2,\pm} (\mathbb{R}^d) \) as the set of all adapted processes in \( \mathbb{L}^{1,p}_{2,\pm} (\mathbb{R}^d) \).

3 Maximum Principle of Singular Optimal Controls

This section will study the optimal controls separately. Due to the special structure of control systems. We shall first consider the singular control part, deriving the necessary condition, subsequently, regular part. The initial condition will fixed to be \( (0,x) \), \( x \in \mathbb{R}^n \). At the beginning let us suppose that \( (\bar{u}(\cdot), \bar{\xi}(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2 \) is an optimal control and denote by \( (X^{0,x;\bar{u},\bar{\eta}}(\cdot), Y^{0,x;\bar{u},\bar{\eta}}(\cdot), Z^{0,x;\bar{u},\bar{\eta}}(\cdot)) \) the optimal solution of (10). Our maximum principle will be proved in two steps. The first variational inequality is derived from the fact

\[
J\left(0, x, u^\varepsilon(\cdot), \bar{\xi}(\cdot)\right) - J\left(0, x, \bar{u}(\cdot), \bar{\xi}(\cdot)\right) \geq 0
\] (15)
where \( u^\varepsilon(\cdot) \) is a convex perturbation of optimal control. The second variational inequity is attained from the inequity
\[
J(0, x, \bar{u}(\cdot), \xi^\varepsilon(\cdot)) - J(0, x, \bar{u}(\cdot), \bar{\xi}(\cdot)) \geq 0
\]
where \( \xi^\varepsilon(\cdot) \) is a convex perturbation of \( \xi \).

### 3.1 Optimal Singular Control

For \( l = b(\cdot), \sigma(\cdot), f(\cdot) \), we denote
\[
\begin{align*}
\bar{l}_x (r, \cdot) &= l_x \left( r, X^{0,x;\bar{u},\bar{\xi}}(r), Y^{0,x;\bar{u},\bar{\xi}}(r), Z^{0,x;\bar{u},\bar{\xi}}(r), \bar{u}(r) \right), \\
\bar{l}_y (r, \cdot) &= l_y \left( r, X^{0,x;\bar{u},\bar{\xi}}(r), Y^{0,x;\bar{u},\bar{\xi}}(r), Z^{0,x;\bar{u},\bar{\xi}}(r), \bar{u}(r) \right).
\end{align*}
\]

Let us introduce the following proposition.

**Proposition 6.** Let (A1)-(A2) hold, and let
\[
(X^{0,x;\bar{u},\bar{\xi}}(\cdot), Y^{0,x;\bar{u},\bar{\xi}}(\cdot), Z^{0,x;\bar{u},\bar{\xi}}(\cdot)) \in \mathcal{N}^2(0, T; \mathbb{R}^n)
\]
be an optimal solution. Then, the following FBSDES:
\[
\begin{cases}
dp(r) = -\big[\bar{b}_x (r, \cdot)^T p(r) + \bar{\sigma}_x (r, \cdot)^T k(r) - \bar{f}_x (r, \cdot)^T q(r)\big] \, dr + t(r) \, dW(r), \\
dq(r) = \bar{f}_y (r, \cdot)^T q(r) \, dr + \bar{f}_z (r, \cdot)^T q(r) \, dW(r), \\
p(T) = -\Phi_x\left(X^{t,x;u^\varepsilon}(T)\right)q(T), \quad q(0) = 1,
\end{cases}
\]
(17)

admit an adapted solution \((p(\cdot), q(\cdot), \xi(\cdot)) \in \mathcal{N}^2(0, T; \mathbb{R}^n)\).

**Theorem 7.** Let (A1)-(A2) hold. If \((X^{\bar{u},\bar{\xi}}(\cdot), Y^{\bar{u},\bar{\xi}}(\cdot), Z^{\bar{u},\bar{\xi}}(\cdot), \bar{u}(\cdot), \bar{\xi}(\cdot))\) is an optimal solution of (10), then there exists a unique pair of adapted processes \((p(\cdot), q(\cdot))\) satisfying (17) such that
\[
P\left\{q(t) I_{K(i)} - \bar{p}^T(t) G(i)(t) \geq 0, \ t \in [0, T], \ \forall i \right\} = 1,
\]
and
\[
P\left\{\sum_{i=1}^m \int_0^T I_{\{q(r) K(i) - \bar{p}^T(r) G(i)(t) > 0\}} \, d\bar{\xi}_x^{(i)} = 0 \right\} = 1.
\]

Before the proof, we need some lemmas. At the beginning, we introduce the convex perturbation
\[
(\bar{u}(t), \xi^\alpha(t)) = (\bar{u}(t), \bar{\xi}(t) + \alpha (\xi(t) - \bar{\xi}(t)))
\]
where \( \alpha \in [0, 1] \) and \( \xi(\cdot) \) is an arbitrary element of \( \mathcal{U}_2 \). We now introduce the following variational equations of (10):
\[
\begin{align*}
\text{d}x^1(t) &= \bar{b}_x(t) x^1(t) \, dt + \bar{\sigma}_x(t) x^1(t) \, dW(t) + G(t) \, d(\xi(t) - \bar{\xi}(t)), \\
\text{d}y^1(t) &= -\bar{f}_x(t) x^1(t) - \bar{f}_y(t) y^1(t) - \bar{f}_z(t) z^1(t) \, dt + \bar{z}^1(t) \, dW(t) \\
&\quad - K \, d(\xi(t) - \bar{\xi}(t)), \\
y^1(T) &= \Phi_x\left(X^{0,x;\bar{u},\bar{\xi}}(T)\right) x^1(T), \quad x^1(0) = 0.
\end{align*}
\]

From (A1)-(A2) it is easy to check that (20) has a unique strong solution. Moreover, we have
Lemma 8. Under the assumptions (A1)-(A2), we have

\[
\lim_{\alpha \to 0} \mathbb{E} \left[ \frac{X^{0,x;\bar{u},\xi^\alpha} (t) - X^{0,x;\bar{u},\bar{\xi}} (t)}{\alpha} - x^1 (t) \right]^2 = 0, \quad t \in [0,T],
\]

\[
\lim_{\alpha \to 0} \mathbb{E} \left[ \frac{Y^{0,x;\bar{u},\xi^\alpha} (t) - Y^{0,x;\bar{u},\bar{\xi}} (t)}{\alpha} - y^1 (t) \right]^2 = 0, \quad t \in [0,T],
\]

\[
\lim_{\alpha \to 0} \mathbb{E} \left[ \int_0^T \frac{Z^{0,x;\bar{u},\xi^\alpha} (t) - Z^{0,x;\bar{u},\bar{\xi}} (t)}{\alpha} - z^1 (t) \right]^2 dt = 0.
\]

The proof can be seen in the Appendix.

Proof of Theorem 7. Applying Itô’s formula to \( \langle p (\cdot), x^1 (\cdot) \rangle + q (\cdot) y^1 (\cdot) \) on \([0,T]\) yields

\[
0 \leq y^1 (0) = \mathbb{E} \left[ \int_0^T \left( q (r) K - p^\top (r) G (t) \right) d\left( \xi (t) - \bar{\xi} (t) \right) dt \right]. \tag{21}
\]

In particular, let \( \xi \in \mathcal{U}_2 \) be a process satisfying \( P \left\{ \sum_i \int_0^T G (s) d\xi_s^{(i)} < \infty \right\} \) and such that (21) and

\[
d\xi_s^{(i)} = \begin{cases} 
0 & \text{if } q (r) K_{(i)} - p^\top (r) G_{(i)} (t) > 0, \\
\xi_s^{(i)} & \text{otherwise},
\end{cases}
\]

holds where \( \xi_s^{(i)} \) denotes the \( i \)th component. Then,

\[
\mathbb{E} \left[ \sum_{i=1}^m \int_0^T \left( q (r) K_{(i)} - p^\top (r) G_{(i)} (t) \right) I_{\{q(r)K_{(i)} - p^\top(r)G_{(i)}(t) > 0\}} d\left( -\bar{\xi}^{(i)} (t) \right) \right] = \mathbb{E} \left[ \int_0^T \left( q (r) K - p^\top (r) G (t) \right) d\left( \xi (t) - \bar{\xi} (t) \right) dt \right] \geq 0
\]

thus

\[
\mathbb{E} \left[ \sum_{i=1}^m \int_0^T \left( q (r) K_{(i)} - p^\top (r) G_{(i)} (t) \right) I_{\{q(r)K_{(i)} - p^\top(r)G_{(i)}(t) > 0\}} d\left( -\bar{\xi}^{(i)} (t) \right) dt \right] = 0
\]

which prove (18). Next we show that (19) is valid. For that, let us define the events:

\[
\mathcal{A}^{(i)} \triangleq \left\{ (t, \omega) \in [0,T] \times \Omega : q (r) K_{(i)} - p^\top (r) G_{(i)} (t) < 0 \right\},
\]

where \( t \in [0,T], 1 \leq i \leq m \). Define the stochastic process \( \bar{\xi}^{(i)} : [0,T] \times \Omega \to [0,\infty) \) by

\[
\bar{\xi}^{(i)} (t) = -\bar{\xi}^{(i)} (t) + \int_0^t I_{\mathcal{A}^{(i)}} (r, \omega) dr
\]

Then one can easily check that \( \bar{\xi} = (\bar{\xi}^{(1)}, \bar{\xi}^{(2)}, \ldots, \bar{\xi}^{(m)}) \) is a measurable, adapted process which is nondecreasing left-continuous with right limits and \( \bar{\xi} (0) = 0 \), and which satisfies \( P \left\{ \sum_i \int_0^T G (s) d\bar{\xi}_s^{(i)} < \infty \right\} \). Further, we have

\[
\mathbb{E} \left[ \int_0^T \left( q (r) K - p^\top (r) G (t) \right) d\left( \bar{\xi} (t) - \bar{\xi} (t) \right) dt \right] = \mathbb{E} \left[ \sum_{i=1}^m \int_0^T \left( q (r) K_{(i)} - p^\top (r) G_{(i)} (t) \right) I_{\mathcal{A}^{(i)}} dt \right] < 0,
\]

13
which obviously contradicts to (21), unless for any \( i \), we have \( (\text{Leb} \otimes P) \{ A^{(i)} \} = 0 \). We thus complete the proof.

**Remark 9.** One can easily check that

\[
q(s) = \exp \left\{ \int_t^s \left[ \dot{f}_y(r) - \frac{1}{2} \left| \dot{f}_z(r) \right|^2 \right] \, dr + \int_0^s \dot{f}_z(r) \, dW(r) \right\},
\]

which implies that \( q(r) > 0, r \in [0, T] \), \( P \)-a.s. So \( -p(\cdot)/q(\cdot) \) makes sense. Clearly, our Theorem 7 for optimal singular control is completely different from [18]. Ours contains two variables \( (p(\cdot), q(\cdot)) \). As a matter of fact, we have

\[
P\left\{ K_{(i)} + \left( -\frac{p^T(t)}{q(t)} \right) G_{(i)}(t) \geq 0, \ t \in [0, T], \ \forall i \right\} = 1.
\]

We claim that \( -p(\cdot)/q(\cdot) \) is the partial derivative of value function, which will be studied in Section 4.3. Note that Theorem 7 is slightly different from Ji et al. [55], namely, our BSDE contains the singular control.

### 3.2 Optimal Regular Control

In this subsection, we study the optimal regular controls for systems driven by Eq. (10) under the types of Pontryagin, namely, necessary maximum principles for optimal control. To this end, we fix \( \tilde{\xi} \in U_2 \) and introduce the following convex perturbation control. Taking \( u(\cdot) \in U_1 \), we define

\[
v(\cdot) = u(\cdot) - \bar{u}(\cdot),
\]

\[
u^\varepsilon(\cdot) = \bar{u}(\cdot) + \varepsilon v(\cdot),
\]

where \( \varepsilon > 0 \) is sufficiently small. Since \( U \) is convex, \( u^\varepsilon(\cdot) \in U(0, T) \). Let \( (x^\varepsilon, y^\varepsilon, z^\varepsilon, u^\varepsilon) \) be the trajectory of the control system (10) corresponding to the control \( u^\varepsilon \).

Let us introduce the following two kinds of variational equations, mainly taken from [23]. For simplicity, we omit the superscript.

\[
\begin{cases}
\quad dx_1(t) = \left[ b_x(t) x_1(t) + b_u(t) v(t) \right] \, dt + \left[ \sigma_x(t) x_1(t) + \sigma_u(t) v(t) \right] \, dW(t), \\
\quad x_1(0) = 0, \quad t \in [0, T],
\end{cases}
\]

and

\[
\begin{cases}
\quad dx_2(t) = \left[ b_x(t) x_2(t) + x_1(t)^T b_{xx}(t) x_1(t) \\
\quad + 2v(t)^T b_{xu}(t) x_1(t) + v(t)^T b_{uu}(t) v(t) \right] \, dt, \\
\quad \left[ \sigma_x(t) x_2(t) + x_1(t)^T \sigma_{xx}(t) x_1(t) \\
\quad + 2v(t)^T \sigma_{xu}(t) x_1(t) + v(t)^T \sigma_{uu}(t) v(t) \right] \, dW(t), \\
\quad x_2(0) = 0, \quad t \in [0, T].
\end{cases}
\]

From Lemma 3.5 and Lemma 3.11 in [23], we have following result.
Lemma 10. Assume that (A1)-(A2) is in force. Then, we have, for any \( \beta \geq 2 \),
\[
\|x_1\|_\infty^\beta \leq C, \quad \|x_2\|_\infty^\beta \leq C, \quad \|\delta x\|_\infty^\beta \leq C\varepsilon^\beta, \\
\|\delta x - \varepsilon x_1\|_\infty^\beta \leq C\varepsilon^{2\beta}, \quad \left\|\delta x - \varepsilon x_1 - \frac{\varepsilon^2}{2}x_2\right\|_\infty^\beta \leq C\varepsilon^{3\beta},
\]
where \( \|x_1\|_\infty^\beta = \left[\mathbb{E}\left(\sup_{t \in [0,T]} |x_1(t)|^\beta\right)\right]^\frac{1}{\beta} \).

We shall introduce the so called variational equations for FBSDEs (10) beginning from the following two adjoint equations:
\[
\begin{aligned}
-\mathrm{d}p(t) &= \Gamma(t)\mathrm{d}t - q(t)\mathrm{d}W(t), \\
p(T) &= \Phi_x(\bar{x}(T)),
\end{aligned}
\]
and
\[
\begin{aligned}
-\mathrm{d}P(t) &= \Pi(t)\mathrm{d}t - Q(t)\mathrm{d}W(t), \\
P(T) &= \Phi_{xx}(\bar{x}(T)),
\end{aligned}
\]
where \( \Gamma(\cdot), \Pi(\cdot) \) are unknown two processes to be determined.

We observe that
\[
\Phi(x^\varepsilon(T)) - \Phi(\bar{x}(T)) = \Phi_x(\bar{x}(T))\Delta x(T) + \frac{1}{2}\Phi_{xx}(\bar{x}(T))(\Delta x(T))^2 + o(\varepsilon)
\]
\[
= \varepsilon\Phi_x(\bar{x}(T))x_1(T) + \frac{\varepsilon^2}{2}\Phi_x(\bar{x}(T))x_2(T)
\]
\[
+ \frac{\varepsilon^2}{2}\Phi_{xx}(\bar{x}(T))(x_1(T))^2 + o(\varepsilon),
\]
where \( \Delta x(T) = x^\varepsilon(T) - \bar{x}(T) \), which inspires us to use the adjoint equations to expand the following
\[
\varepsilon p(t)x_1(t) + \frac{\varepsilon^2}{2}p(t)x_2(t) + \frac{\varepsilon^2}{2}x_1^T(t)P(t)x_1(t)
\]
on \([0,T] \).

By virtue of Itô’s formula, it follows that
\[
\varepsilon p(T)x_1(T) + \frac{\varepsilon^2}{2}p(T)x_2(T) + \frac{\varepsilon^2}{2}x_1^T(T)P(T)x_1(T)
\]
\[
= \varepsilon p(t)x_1(t) + \frac{\varepsilon^2}{2}p(t)x_2(t) + \frac{\varepsilon^2}{2}x_1^T(t)P(t)x_1(t)
\]
\[
+ \int_t^T \left[\Lambda_1(s)\left(\varepsilon x_1(s) + \frac{\varepsilon^2}{2}x_2(s)\right) + \varepsilon^2\Lambda_2(s)x_1(s)\right]\mathrm{d}s
\]
\[
+ \frac{\varepsilon^2}{2}\Lambda_3(s)x_1^2(s) + \Lambda_4(s)\right]\mathrm{d}s
\]
\[
+ \int_t^T \left[\Lambda_5(s)\left(\varepsilon x_1(s) + \frac{\varepsilon^2}{2}x_2(s)\right) + \varepsilon^2\Lambda_6(s)x_1(s)\right]dW(s)
\]
\[
+ \frac{\varepsilon^2}{2}\Lambda_7(s)x_1^2(s) + \Lambda_8(s)\right]dW(s)
\]
where
\[ \Lambda_1 (t) = p (t) b_x (t) + q (t) \sigma_x (t) - \Gamma (t), \]
\[ \Lambda_2 (t) = p (t) b_{xx} (t) v (t) + q (t) \sigma_{xx} (t) v (t) + P (t) b_u (t) v (t) + P (t) \sigma_x (t) \sigma_u (t) v (t) + Q (t) \sigma_u (t) v (t), \]
\[ \Lambda_3 (t) = p (t) b_{xx} (t) + q (t) \sigma_{xx} (t) + 2 P (t) b_x (t) + \sigma_x (t) P (t) 2 \sigma_x (t) + 2 Q (t) \sigma_x (t) - \Pi (t), \]
\[ \Lambda_4 (t) = \frac{\varepsilon^2}{2} p (t) v^2 (t) b_u (t) + \varepsilon q (t) \sigma_u (t) v (t) + \frac{\varepsilon^2}{2} \sigma_u^2 (t) v^2 (t) P (t), \]
\[ \Lambda_5 (t) = p (t) \sigma_x (t) + q (t), \]
\[ \Lambda_6 (t) = 2 p (t) v (t) \sigma_{xx} (t) + P (t) \sigma_u (t) v (t), \]
\[ \Lambda_7 (t) = p (t) \sigma_{xx} (t) + 2 P (t) \sigma_x (t) + Q (t), \]
\[ \Lambda_8 (t) = \varepsilon p (t) \sigma_u (t) v (t) + \frac{\varepsilon^2}{2} p (t) v^\top (t) \sigma_u u (t) v (t). \]

**Remark 11.** Note that \( \Gamma (t) \) and \( \Pi (t) \) do not appear in the \( dW (s) \)-term.

Define
\[
\begin{cases} 
    d y^\varepsilon (t) = - f ((t, x^\varepsilon (t), y^\varepsilon (t), z^\varepsilon (t), u^\varepsilon (t)) \ dt + z^\varepsilon (t) \ dW (t), \\
    y^\varepsilon (T) = \Phi (x^\varepsilon (T)).
\end{cases}
\]

Let
\[
\dot{y}^\varepsilon (t) = y^\varepsilon (t) - \left[ p (t) \left( \varepsilon x_1 (t) + \frac{\varepsilon^2}{2} x_2 (t) \right) + \frac{\varepsilon^2}{2} x_1^\top (t) P (t) x_1 (t) \right] \quad (26)
\]
\[
\dot{z}^\varepsilon (t) = z^\varepsilon (t) - \left[ \Lambda_5 (s) \left( \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s) \right) + \varepsilon^2 \Lambda_6 (s) x_1 (s) + \frac{\varepsilon^2}{2} x_1^\top (s) \Lambda_7 (s) x_1 (s) + \Lambda_8 (s) \right]. \quad (27)
\]

Clearly, from Lemma 10, we have
\[
\Phi (x^\varepsilon (T)) = \Phi (\bar{x} (T)) + p (T) \left( \varepsilon x_1 (T) + \frac{\varepsilon^2}{2} x_2 (T) \right) + \frac{\varepsilon^2}{2} x_1^\top (T) P (T) x_1 (T) + o (\varepsilon). \]

After some tedious computations, we have
\[
\dot{y}^\varepsilon (t) = \Phi (\bar{x} (T)) + \int_t^T \left[ f (s, x^\varepsilon (s), y^\varepsilon (s), z^\varepsilon (s), u^\varepsilon (s)) + \Lambda_1 (s) (\varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s)) + \varepsilon^2 x_1^\top (s) \Lambda_2 (s) + \frac{\varepsilon^2}{2} x_1^\top (s) \Lambda_3 (s) x_1 (s) + \Lambda_4 (s) \right] \ ds - \int_t^T \dot{z}^\varepsilon (s) dW (s) + o (\varepsilon). \quad (28)
\]
Put
\[
\tilde{y}^\varepsilon (t) = \tilde{y}^\varepsilon (t) - \bar{y} (t),
\]
\[
\tilde{z}^\varepsilon (t) = \tilde{z}^\varepsilon (t) - \bar{z} (t),
\]
then we attain
\[
\tilde{y}^\varepsilon (t) = \int_t^T \left[ f (s, x^\varepsilon (s), y^\varepsilon (s), z^\varepsilon (s), u^\varepsilon (s)) - f (s, \bar{x} (s), \bar{y} (s), \bar{z} (s), \bar{u} (s)) \right. \\
+ \Lambda_1 (s) \left( \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s) \right) + \varepsilon^2 x_1^\top (s) \Lambda_2 (s) \\
+ \frac{\varepsilon^2}{2} \Lambda_3 (s) x_1^2 (s) + \Lambda_4 (s) \right] \, ds - \int_t^T \tilde{z}^\varepsilon (s) \, dW (s), \quad (29)
\]
where
\[
f (s, x^\varepsilon (s), y^\varepsilon (s), z^\varepsilon (s), u^\varepsilon (s)) - f (s, \bar{x} (s), \bar{y} (s), \bar{z} (s), \bar{u} (s)) \\
= f \left( s, \bar{x} (s), \bar{y} (s), \bar{z} (s) + \varepsilon p (t) \sigma_u (t) v (t) + \frac{\varepsilon^2}{2} p (t) \sigma_{uu} (t) v^2 (t), \bar{u} (s) \right) \\
- f (s, \bar{x} (s), \bar{y} (s), \bar{z} (s), \bar{u} (s)) + f (s, x^\varepsilon (s), y^\varepsilon (s), z^\varepsilon (s), u^\varepsilon (s)) \\
- f \left( s, \bar{x} (s), \bar{y} (s), \bar{z} (s) + \varepsilon p (t) \sigma_u (t) v (t) + \frac{\varepsilon^2}{2} p (t) \sigma_{uu} (t) v^2 (t), \bar{u} (s) \right) \\
= f \left( s, \bar{x} (s), \bar{y} (s), \bar{z} (s) + \varepsilon p (t) \sigma_u (t) v (t) + \frac{\varepsilon^2}{2} p (t) \sigma_{uu} (t) v^2 (t), \bar{u} (s) \right) \\
- f (s, \bar{x} (s), \bar{y} (s), \bar{z} (s), \bar{u} (s)) \\
+ f \left( s, \bar{x} (s) + \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \bar{y} (s) + \bar{y}^\varepsilon (t) + \Lambda_9 (s), \right. \\
\tilde{z} (s) + \tilde{z}^\varepsilon (t) + \Lambda_{10} (s), u^\varepsilon (s)) \\
\left. - f (s, \bar{x} (s), \bar{y} (s), \bar{z} (s), \bar{u} (s)) + o (\varepsilon) \right).
\]
Next we shall seek $\Gamma (\cdot), \Pi (\cdot)$, which are in fact determined by the optimal quadruple $(\bar{x} (\cdot), \bar{y} (\cdot), \bar{z} (\cdot), \bar{u} (\cdot))$, such that
\[
f \left( s, \bar{x} (s) + \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \bar{y} (s) + \Lambda_9 (s), \tilde{z} (s) + \Lambda_{10} (s), u^\varepsilon (s) \right) \\
- f (s, \bar{x} (s), \bar{y} (s), \bar{z} (s), \bar{u} (s)) + \Lambda_1 (s) \left( \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s) \right) \\
+ \frac{\varepsilon^2}{2} x_1^\top (s) \Lambda_3 (s) x_1 (s),
\]
where
\[
\Lambda_9 (s) = p (t) \left( \varepsilon x_1 (t) + \frac{\varepsilon^2}{2} x_2 (t) \right) + x_1^\top (t) P (t) x_1 (t), \\
\Lambda_{10} (s) = \left[ \Lambda_5 (s) \left( \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s) \right) + \varepsilon^2 \Lambda_6 (s) x_1 (s) \\
+ \frac{\varepsilon^2}{2} x_1^\top (s) \Lambda_7 (s) x_1 (s) + \Lambda_8 (s) \right]
\]
in which $O(\varepsilon)$ does not involve the terms $x_1(\cdot)$ and $x_2(\cdot)$. Note that in BSDE (28), there appears the term $x_1^T (s) \Lambda_3 (s) x_1 (s)$. Hence, we make use of Taylor's expansion to

$$
f \left( s, \bar{x}(s) + \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \bar{y}(s) + \Lambda_9 (s), \bar{z}(s) + \Lambda_{10} (s), u^\varepsilon (s) \right)
$$

$$
= f \left( s, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s) \right) + \left[ f_x (s), f_y (s), f_z (s), f_u (s) \right] \cdot \left[ \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \Lambda_9 (s), \Lambda_{10} (s), \varepsilon v (s) \right]
$$

$$
+ \frac{1}{2} \left[ \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \Lambda_9 (s), \Lambda_{10} (s), \varepsilon v (s) \right]^T \cdot \mathbf{H} f \left( s, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s) \right)
$$

$$
= f \left( s, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s) \right) + \left[ f_x (s), f_y (s), f_z (s), f_u (s) \right] \cdot \left[ \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \Lambda_9 (s), \Lambda_{10} (s), \varepsilon v (s) \right]^T
$$

$$
+ \frac{1}{2} \left[ \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s), \Lambda_9 (s), \Lambda_{10} (s) \right]^T \cdot \mathbf{H}_1 f \left( s, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s) \right)
$$

$$
+ \varepsilon v (s) f_{xu} \left[ \varepsilon x_1 (s) + \frac{\varepsilon^2}{2} x_2 (s) \right] + \varepsilon v (s) f_{yu} \Lambda_9 (s)
$$

$$
+ \varepsilon v (s) f_{zu} \Lambda_{10} (s) + \frac{1}{2} \varepsilon^2 v^2 (s) f_{uu},
$$

where the Hessian matrix is defined by

$$
\mathbf{H} f \left( s, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s) \right) = 
\begin{bmatrix}
\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial x \partial u} \\
\frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial y \partial u} \\
\frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} & \frac{\partial^2 f}{\partial z \partial u} \\
\frac{\partial^2 f}{\partial u \partial x} & \frac{\partial^2 f}{\partial u \partial y} & \frac{\partial^2 f}{\partial u \partial z} & \frac{\partial^2 f}{\partial u^2}
\end{bmatrix},
$$

and

$$
\mathbf{H}_1 f \left( t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t) \right) = 
\begin{bmatrix}
\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial x \partial u} \\
\frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial y \partial u} \\
\frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} & \frac{\partial^2 f}{\partial z \partial u} \\
\frac{\partial^2 f}{\partial u \partial x} & \frac{\partial^2 f}{\partial u \partial y} & \frac{\partial^2 f}{\partial u \partial z} & \frac{\partial^2 f}{\partial u^2}
\end{bmatrix}.
$$

Then, we are able to obtain

$$
\Gamma \left( t \right) = p \left( t \right) (b_x (t) + f_y (t) + \sigma_x (t) f_z (t)) + q \left( t \right) (\sigma_x (t) + f_z (t)) + f_x (t),
$$

$$
\Pi \left( t \right) = p \left( t \right) b_{xx} (t) + q \left( t \right) \sigma_{xx} (t) + 2 P \left( t \right) b_x (t) + P \left( t \right) \sigma_x^2 (t)
$$

$$
+ 2 Q \left( t \right) \sigma_x (t) + f_y (t) P \left( t \right) + \left[ p \left( t \right) \sigma_{xx} (t) + 2 P \left( t \right) \sigma_x (t) + Q \left( t \right) \right] f_z (t)
$$

$$
+ \frac{1}{2} \left[ 1, p \left( t \right), p \left( t \right) \sigma_x (t) + q \left( t \right) \right] \cdot \mathbf{H}_1 f \left( t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t) \right)
$$

$$
\cdot \left[ 1, p \left( t \right), p \left( t \right) \sigma_x (t) + q \left( t \right) \right]^T.
$$
Remark 12. The adjoint equations derived here is the same as in Hu [50].

Proposition 13. Assume that (A1)-(A2) hold. We have

\[ p(s) = -\frac{p^T(s)}{q(s)}, \quad s \in [t, T], \quad P\text{-a.s.,} \]

where \((p(\cdot), q(\cdot))\) is the solution to FBSDEs (24), (17), respectively.

The proof is just to apply the Itô’s formula to \(-p^T(s)/q(s)\), so we omit it.

We define the classical Hamiltonian function:

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

Consider

\[
\dot{\hat{y}}^\varepsilon(t) = \int_t^T \left[ f(s, \bar{x}(s), \bar{y}(s), \bar{z}(s) + \varepsilon p(s) \sigma_u(s) v(s) \\
+ \frac{\varepsilon^2}{2} p(s) v^T(s) \sigma_{uu}(s) v(s), \bar{u}(s) \right) \\
- f(s, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s)) + f_y(s) \dot{\hat{y}}^\varepsilon(s) + f_z(s) \dot{\hat{z}}^\varepsilon(s) \\
+ \varepsilon^2 [p(s) b_{xu}(s) + q(s) \sigma_{xu}(s) + P(s) b_u(s) + P(s) \sigma_x(s) \sigma_u(s) + Q(s) \sigma_u(s)] x_1(s) v(s) \\
+ \varepsilon [p(s) b_u(s) + q(s) \sigma_u(s) + f_u(s)] v(s) \\
+ \frac{\varepsilon^2}{2} \left[ q(s) \sigma_{uu}(s) + p(s) b_{uu}(s) + \varepsilon^2 f_{uu}(s) + \sigma^T_u(s) P(s) \sigma_u(s) \right] v^2(s) \\
+ \varepsilon v(s) f_{xu}(s) \left[ \varepsilon x_1(s) + \frac{\varepsilon^2}{2} x_2(s) \right] + \varepsilon v^T(s) f_{yu}(s) \Lambda_9(s) \\
+ \varepsilon v(s) f_{zu}(s) \Lambda_{10}(s) \right] ds - \int_t^T \dot{\hat{z}}^\varepsilon(s) dW(s).
\]

Namely,

\[
\dot{\hat{y}}^\varepsilon(t) = \int_t^T \left[ f_y(s) \dot{\hat{y}}^\varepsilon(s) + f_z(s) \dot{\hat{z}}^\varepsilon(s) \\
+ \varepsilon^2 x_1^T(s) [H_{xu}(t) + P(s) \sigma_x(s) \sigma_u(s) + Q(s) \sigma_u(s) + p(t) f_{yu}(s) + P(s) b_u(s)] v(s) \\
+ \varepsilon [H_u(s) + f_z(s) p(s) \sigma_u(s)] v(s) \\
+ \frac{\varepsilon^2}{2} v^T(s) [H_{uu}(s) + \sigma_u^2(s) P(s) + f_{zz}(s) P^2(s) \sigma_u^2(s)] v(s) \\
+ \varepsilon^2 v^T(s) f_{zu}(s) \left[ [p(t) \sigma_x(t) + q(t)] x_1(s) + p(t) \sigma_u(t) v(t) \right] ds \\
- \int_t^T \dot{\hat{z}}^\varepsilon(s) dW(s). \tag{30}
\]
Remark 14. Note that FBSDEs (30) are completely different from (22) in Hu [50]. Specifically, the term $A_4 x_1(s) I_{E_ε}(s)$ disappears in (22) since

$$E \left[ \left( \int_0^T |A_4(s) x_1(s) I_{E_ε}(s)| ds \right)^\beta \right] = o(\varepsilon^\beta) \text{ for } \beta \geq 2$$

in [50] by using spike variational approach. Nevertheless, the corresponding term in our paper is just $\varepsilon^2 x_1^T(s) \Lambda_2(s)$. We shall see that we need this term to define a new "Hamiltonian function" as following.

Define

$$H(t,x,y,z,u,p,q,P,Q) \triangleq H_{xu}(t,x,y,z,u,p,q) + b_u(t,x,u) P + Q \sigma_u(t,x,u) \sigma_u(t,x,u) + f_{yu}(t,x,y,z,u) p + f_{zu}(t,x,y,z,u) [p \sigma_x(t,x,u) + q],$$

where $(t,x,y,z,p,q,P,Q) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$. 

Remark 15. Observe that $H$ is slightly different from $S$ in Zhang et al. [102], that is,

$$S(t,x,u,y_1,z_1,y_2,z_2) = H_{xu}(t,x,y,z,u,p,q) + b_u(t,x,u) y_2 + \varepsilon^2 \sigma_u(t,x,u) \sigma_u(t,x,u),$$

where $(y_2,z_2) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$.

We now give the adjoint equation for BSDE (30) as follow:

$$\begin{cases} 
\text{d} \chi(t) = f_{y}(t) \chi(t) \text{d}t + f_{z}(t) \chi(t) \text{d}W(t), \\
\chi(0) = 1.
\end{cases}$$

(31)

From assumptions (A1)-(A2), it is fairly easy to check that SDE (31) admits a unique adapted strong solution $\chi(t) \in \mathcal{M}^2(0,T;\mathbb{R})$.

Set

$$\begin{align*}
y^\varepsilon(t) &= \bar{y}(t) + p^\top(t) \left( \varepsilon x_1(t) + \frac{\varepsilon^2}{2} x_2(t) \right) + \frac{\varepsilon^2}{2} x_1^\top(t) P(t) x_1(t) + \hat{y}^\varepsilon(t), \\
z^\varepsilon(t) &= \bar{z}(t) + \Lambda_5(s) \left( \varepsilon x_1(s) + \frac{\varepsilon^2}{2} x_2(s) \right) + \varepsilon^2 \Lambda_6(s) x_1(s) + \frac{\varepsilon^2}{2} x_1^\top(t) \Lambda_7(s) x_1(s) + \Lambda_8(s) + \hat{z}^\varepsilon(t).
\end{align*}$$

We are able to give the variational equations as follows:

$$\begin{align*}
y^\varepsilon_1(t) &= \varepsilon p^\top(t) x_1(t), \\
z^\varepsilon_1(t) &= \varepsilon \left[ x_1^\top(s) \Lambda_5(s) + p^\top(t) \sigma_u(t) v(t) \right].
\end{align*}$$
and
\[
y_2'(t) = \frac{\varepsilon^2}{2} \left[ p^T(t)x_2(t) + x_1^T(t)P(t)x_1(t) \right] + \hat{y}^\varepsilon(t),
\]
\[
z_2'(t) = \frac{\varepsilon^2}{2} \left[ \Lambda_5(s)x_2(s) + 2\Lambda_6(s)x_1(s) + x_1^T(s)\Lambda_7(s)x_1(s) + p^T(t)v^T(t)\sigma_{uu}(t)v(t) \right] + \hat{z}^\varepsilon(t).
\]
Obviously, we have
\[
y^\varepsilon(0) - \bar{y}(0) = \hat{y}^\varepsilon(0) \geq 0, \text{ since the definition of value function},
\]
\[
z^\varepsilon(0) - \bar{z}(0) = \hat{z}^\varepsilon(0) + \varepsilon p^T(0)\sigma_u(0)v(0) + \frac{\varepsilon^2}{2} p^T(0)v^T(0)\sigma_{uu}(0)v(0).
\]

\textbf{Lemma 16.} Under the assumptions (A1)-(A2), we have the following estimate
\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\hat{y}^\varepsilon(t)|^2 + \int_0^T |\hat{z}^\varepsilon(s)|^2 \, ds \right] = O(\varepsilon^2).
\] (32)

\textbf{Proof.} To prove (32), we consider (30) again. From assumptions (A1)-(A2), one can check that the adjoint equations (24) and (25) have a unique adapted strong solution, respectively. Furthermore, by classical approach, we are able to get the following estimates for \(\beta \geq 2\),
\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} \left( |p(t)|^\beta + |P(t)|^\beta \right) + \left( \int_0^T \left( |q(t)|^2 + |Q(t)|^2 \right) dt \right)^{\frac{\beta}{2}} \right] < +\infty.
\] (33)
Applying Lemma 14 to (30), we get the desired result. \(\square\)

We shall derive a variational inequality which is crucial to establish the necessary condition for optimal control. Before this, we introduce the following the other type of singular control using the Hamiltonian function:

\textbf{Definition 17 (Singular control in the classical sense).} We call a control \(\bar{u}(\cdot) \in \mathcal{U}(0,T)\) a singular control in the classical sense if \(\bar{u}(\cdot)\) satisfies
\[
\begin{align*}
\ i) & \ H_u(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t)) \nonumber \\
& \quad + f_x(t, \bar{x}(t), \bar{y}(t), \bar{z}(t)) \bar{p}(t) + f_y(t, \bar{x}(t), \bar{y}(t), \bar{z}(t)) \bar{q}(t) + \sigma_u^2(t, \bar{x}(t), \bar{u}(t)) \bar{P}(s) = 0, \ a.s., a.e., t \in [0,T]; \\
\ ii) & \ H_{uu}(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t)) \bar{p}(t) + f_x(t, \bar{x}(t), \bar{y}(t), \bar{z}(t)) \bar{p}(t) + f_y(t, \bar{x}(t), \bar{y}(t), \bar{z}(t)) \bar{q}(t) + \sigma_u^2(t, \bar{x}(t), \bar{u}(t)) \bar{P}(s) = 0, \ a.s., a.e., t \in [0,T]; \\
\end{align*}
\] (34)
where \((\bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot))\) denotes the state trajectories driven by \(\bar{u}(\cdot)\). Moreover, \((\bar{p}(\cdot), \bar{q}(\cdot))\) and \((\bar{P}(\cdot), \bar{Q}(\cdot))\) denote the adjoint processes given respectively by (24) and (25) with \((\bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t))\) replaced by \((\bar{x}(\cdot), \bar{y}(\cdot), \bar{z}(\cdot), \bar{u}(\cdot))\). If this \(\bar{u}(\cdot)\) is also optimal, then we call it a singular optimal control in the classical sense.
Remark 18. Hu \cite{50} first considers the forward-backward stochastic control problem whenever the diffusion term $\sigma(t,x,u)$ depends on the control variable $u$ with non-convex control domain. In order to to establish the stochastic maximum principle, he first introduces the $H$-function of the following type:

$$
H(t, x, y, z, u, p, q, P) \triangleq pb(t, x, u) + q\sigma(t, x, u)
$$

$$
+ \frac{1}{2} \left( \sigma(t, x, u) - \sigma(t, \bar{x}, \bar{u}) \right) \mathbb{P} \left( \sigma(t, x, u) - \sigma(t, \bar{x}, \bar{u}) \right)
$$

$$
+ f(t, x, y, z + p \sigma(t, x, u) - \sigma(t, \bar{x}, \bar{u}), u).
$$

Note that this Hamiltonian function is slightly different from Peng 1990 \cite{82}. The main difference of this variational equations with those in (Peng 1990) \cite{82} appears in the term $p(t) \delta \sigma(t) I_{E_0}(t)$ (the similar term $\varepsilon p(t) \sigma_u(t) v(t) + \frac{\varepsilon^2}{2} p(t) \sigma_{uu}(t) v^2(t)$ in our paper) in variational equation for BSDE and maximum principle for the definition of $p(t)$ in the variation of $z$, which is $O(\varepsilon)$ for any order expansion of $f$. So it is not helpful to use the second-order Taylor expansion for treating this term. The stochastic maximum principle (see \cite{50}) says that if $(\bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t))$ is an optimal pair, then

$$
H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \bar{P}(t)) \leq \max_{u \in U} H(t, x(t), y(t), z(t), u, p(t), q(t), \bar{P}(t))
$$

(35)

Remark 18. Hu \cite{50} first considers the forward-backward stochastic control problem whenever the diffusion term $\sigma(t,x,u)$ depends on the control variable $u$ with non-convex control domain. In order to to establish the stochastic maximum principle, he first introduces the $H$-function of the following type:

$$
H(t, x, y, z, u, p, q, P) \triangleq pb(t, x, u) + q\sigma(t, x, u)
$$

$$
+ \frac{1}{2} \left( \sigma(t, x, u) - \sigma(t, \bar{x}, \bar{u}) \right) \mathbb{P} \left( \sigma(t, x, u) - \sigma(t, \bar{x}, \bar{u}) \right)
$$

$$
+ f(t, x, y, z + p \sigma(t, x, u) - \sigma(t, \bar{x}, \bar{u}), u).
$$

Note that this Hamiltonian function is slightly different from Peng 1990 \cite{82}. The main difference of this variational equations with those in (Peng 1990) \cite{82} appears in the term $p(t) \delta \sigma(t) I_{E_0}(t)$ (the similar term $\varepsilon p(t) \sigma_u(t) v(t) + \frac{\varepsilon^2}{2} p(t) \sigma_{uu}(t) v^2(t)$ in our paper) in variational equation for BSDE and maximum principle for the definition of $p(t)$ in the variation of $z$, which is $O(\varepsilon)$ for any order expansion of $f$. So it is not helpful to use the second-order Taylor expansion for treating this term. The stochastic maximum principle (see \cite{50}) says that if $(\bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t))$ is an optimal pair, then

$$
H(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \bar{P}(t)) = \max_{u \in U} H(t, x(t), y(t), z(t), u, p(t), q(t), \bar{P}(t))
$$

(35)

Apparantly, Definition \cite{17} says that a singular control in the classical sense is the real one that fulfils trivially the first and second-order necessary conditions in classical optimization theory dealing with the maximization problem \cite{50}, namely,

$$
\begin{align*}
\mathcal{H}_u \left( t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \bar{P}(t) \right) &= 0, \text{ a.s., a.e., } t \in [0, T] ; \\
\mathcal{H}_{uu} \left( t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \bar{P}(t) \right) &= 0, \text{ a.s., a.e., } t \in [0, T].
\end{align*}
$$

(36)

It is easy to verify that (36) is equivalent to (50). Certainly, one could investigate stochastic singular optimal controls for forward-backward stochastic systems in other senses, say, in the sense of process in Skorohod space, which can be seen in Zhang \cite{119} via viscosity solution approach (Hamilton-Jacobi-Bellman inequality), or in the sense of Pontryagin-type maximum principle (Tang \cite{93}). As this complete remake of the various topics is much longer than the present paper, it will be reported elsewhere.

Lemma 19 (Variational inequality). Under the assumptions (A1)-(A2), it holds that

$$
0 \leq \mathbb{E} \left[ \int_0^T \chi(s) \left[ \frac{\varepsilon^2}{2} p(s) \sigma_{uu}(s) v^2(s) f_z(s) + \varepsilon^2 x_1(s) v(s) \mathbb{H}(s) + \varepsilon^2 v^2(s) f_{zu}(s) p(t) \sigma_u(t) + \varepsilon v(s) \left[ f_z(s) p(s) \sigma_u(s) + H_u(s) \right] + \varepsilon^2 \left[ H_{uu}(s) + \sigma_u^2(s) P(s) \right] v^2(s) \right] ds \right].
$$

(37)

where $\mathbb{H}(s) = \mathbb{H}(t, \bar{x}(s), \bar{y}(s), \bar{z}(s), \bar{u}(s), p(s), q(s), P(s), Q(s))$. 

22
Proof. Using Itô’s formula to \( \langle \chi (s), \hat{y}^\varepsilon (s) \rangle \) on \([0, T]\), we get the desired result.  

**Theorem 20.** Assume that (A1)-(A2) hold. If \( \bar{u} (\cdot) \in U (0, T) \) is a singular optimal control in the classical sense, then

\[
0 \leq E \left[ \int_0^T \chi (s) \mathbb{H} (s) x_1 (s) v (s) \, ds \right], \tag{38}
\]

for any \( v (\cdot) = u (\cdot) - \bar{u} (\cdot) \), \( u (\cdot) \in U (0, T) \).

**Proof.** By virtue of Definition 17 and Lemma 19, we have

\[
0 \leq E \left[ \int_0^T \chi (s) \varepsilon^2 x_1 (s) v (s) \mathbb{H} (s) \, ds \right].
\]

According to the definition of value function, we have,

\[
J (u^\varepsilon) - J (\bar{u}) = y^\varepsilon (0) - \bar{y} (0) = \hat{y}^\varepsilon (0) \geq 0.
\]

Letting \( \varepsilon \to 0^+ \), we get the desired result.  

**Remark 21.** Clearly, if \( f \) does not contain \((y, z)\), then (38) reduces to

\[
E \left[ \int_0^T \chi (s) \mathbb{S} (s) x_1 (s) v (s) \, ds \right] \geq 0,
\]

which is just the classical case studied in Zhang et al. [102] for classical stochastic control problems. Meanwhile, our result actually extends Peng [83] to second order case.

**Remark 22.** Recall that, for deterministic system, it is possible to derive pointwise necessary conditions for optimal controls via the first suitable integral-type necessary conditions and normally there is no obstacles to establish the pointwise first-order necessary condition for optimal controls whenever an integral type one is on the hand. Nevertheless, the classical approach to handle the pointwise condition from the integral-type can not be employed directly in the framework of the pointwise second-order condition in the general stochastic setting because of certain feature the stochastic systems owning. In order to derive the second order variational equations for BSDE in Hu [50], the author there introduces two kinds of adjoint equations and a new Hamiltonian function. The main difference of this variational equations with those in (Peng 1990) [82] lies in the term \( p (t) \delta \sigma (t) I_{E_\varepsilon} (t) \). Then, it is possible to get the maximum principle basing one variational equation. Note that the order of the difference between perturbed state, optimal state and first, second order state is \( o(\varepsilon) \).

As observed in Theorem 20, there appears a term \( \mathbb{H} (s) x_1 (s) v (s) \). In order to deal with it, we give the expression of \( x_1 (\cdot) \), mainly taken from Theorem 1.6.14 in Yong and Zhou [100]. To this end, consider the following matrix-valued stochastic differential equation

\[
\begin{cases}
    d\Psi (t) = b_x (t) \Psi (t) \, dt + \sigma_x (t) \Psi (t) \, dW (t), \\
    \Psi (0) = I, \ t \in [0, T],
\end{cases}
\tag{39}
\]
Lemma 24. Assume that $A_1$-$(A_2)$ hold. Then, $\chi(\cdot) \in L_2^\infty([0,T]:\mathbb{R}^n)$. Suppose that

$$
\chi(t) = \mathbb{E}[\phi(t)] + \int_0^t \kappa(s,t) dW(s), \ a.s., \ a.e., \ t \in [0,T].
$$

The proof can be seen in Zhang et al. [102].

Lemma 23 (Martingale representation theorem). Suppose that $\phi \in L_2^\infty(\Omega; L^2([0,T]:\mathbb{R}^n))$. Then, there exists a $\kappa(\cdot,\cdot) \in L^2([0,T] \times \Omega; \mathbb{R}^n)$ such that

$$
\phi(t) = \mathbb{E}[\phi(t)] + \int_0^t \kappa(s,t) dW(s), \ a.s., \ a.e., \ t \in [0,T].
$$
Proof} We shall prove that
\[
\mathbb{E} \left[ \int_0^T |\chi(t) \mathbb{H}(t)|^2 \, dt \right] < \infty.
\]

From (A1)-(A2), we have
\[
|\psi_x| \leq C \quad \text{and} \quad |\psi_{ux}| \leq C
\]
for \( \psi = b, \sigma, f \). Besides,
\[
|f_{yu}| \leq C, \quad |f_{zu}| \leq C.
\]
Hence,
\[
\mathbb{E} \left[ \int_0^T |\chi(t) \mathbb{H}(t)|^2 \, dt \right]
\]
\[
= \mathbb{E} \left[ \int_0^T \chi(t) \left| H_{xu}(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{z}(t), \bar{\bar{u}}(t), \bar{\bar{p}}(t), \bar{\bar{q}}(t)) \right|^2 
+ Q\sigma_u(t, \bar{x}(t), \bar{u}(t)) + P\sigma_x(t, \bar{x}(t), \bar{u}(t)) \sigma_u(t, \bar{x}(t), \bar{u}(t)) + b_u(t, \bar{x}(t), \bar{u}(t)) P(t) 
+ f_{yu}(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{z}(t), \bar{\bar{u}}(t)) p(t) 
+ f_{zu}(t, \bar{x}(t), \bar{y}(t), \bar{u}(t), \bar{z}(t), \bar{\bar{u}}(t)) [p\sigma_x(t, \bar{x}(t), \bar{u}(t)) + q(t)]^2 \, dt \right]
\]
\[
\leq C \mathbb{E} \left[ \int_0^T \chi(t) \left( |p(t)|^2 + |q(t)|^2 + |P(t)|^2 + |Q(t)|^2 \right) \, dt \right]
\]
\[
\leq C \left( \mathbb{E} \left[ \int_0^T \chi(t) \, dt \right]^2 \right)^\frac{1}{2} \cdot \left( \mathbb{E} \left[ \int_0^T \left( |p(t)|^2 + |q(t)|^2 + |P(t)|^2 + |Q(t)|^2 \right) \, dt \right]^2 \right)^\frac{1}{2}
\]
\[
< \infty.
\]

The last inequality is basing on the classical estimate [33] of SDE and BSDE theories. We thus complete the proof. \( \square \)

Therefore, by our assumption (A1)-(A2) and Lemma [23], for any \( u \in U \), there exists a \( \psi^u \) \((\cdot, \cdot) \in L^2 ([0, T] : L^2_{\text{loc}} (\Omega \times [0, T] : \mathbb{R}^n))\) such that for a.e. \( t \in [0, T] \)
\[
\chi(t) \mathbb{H}^T(t) (u - \bar{u}(t)) = \mathbb{E} \left[ \chi(t) \mathbb{H}^T(t) (u - \bar{u}(t)) \right] + \int_0^t \psi^u(s, t) \, dW(s). \quad (42)
\]
Using [42], we are able to assert the following:
Example 27. Consider the following FBSDEs with \( n = 1 \) and \( U = [-1, 1] \).

\[
\begin{align*}
    \frac{dx(t)}{dt} &= u(t) dt + u(t) dW(t), \\
    -\frac{dy(t)}{dt} &= u^2(t) dt - z(t) dW(t), \\
    x(0) &= 0, \quad y(T) = \frac{1}{2}x^2(T).
\end{align*}
\] (43)

One can easily get the solutions to (37), \( \chi(t) = 1 \). Set \( (\bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{u}(t)) = (0, 0, 0, 0) \).

The corresponding adjoint equations are (26) and (27), namely,

\[
\begin{align*}
    -dp(t) &= -q(t) dW(t), \\
    p(T) &= 0,
\end{align*}
\] (44)

Theorem 25. Suppose that (A1)-(A2) are in force. Let \( \bar{u}(\cdot) \) be a singular optimal control in the classical sense, then we have

\[
\mathbb{E} \langle \chi(r) \mathbb{H}(r) b_u(r)(u - \bar{u}(r)), u - \bar{u}(r) \rangle + \partial_t^+ \left( \chi(r) \mathbb{H}^T(r)(u - \bar{u}(r)), \sigma_u(r)(u - \bar{u}(r)) \right) \geq 0, \text{ a.e., } r \in [0, T],
\]

where

\[
\partial_t^+ \left( \chi(r) \mathbb{H}^T(r)(u - \bar{u}(r)), \sigma_u(r)(u - \bar{u}(r)) \right) = 2 \limsup_{\alpha \to 0} \frac{1}{\alpha^2} \mathbb{E} \int_t^{t+\alpha} \left( \psi^n(s, t), \Psi(r) \Psi^{-1}(s) \sigma_u(s)(u - \bar{u}(s)) \right) ds dt,
\]

where \( \psi^n(s, t) \) is obtained by (42), and \( \Psi \) is determined by (39).

The proof is just to repeat the process in Theorem 3.10, [102], so we omit it.

Note that Theorem 25 is pointwise with respect to the time variable \( t \) (but also the integral form). Now if each of \( \chi(\cdot) \mathbb{H}(\cdot) \) and \( \bar{u}(\cdot) \) are regular enough, then the function \( \psi^n(\cdot, \cdot) \) admits an explicit representation.

Suppose the following:

\[(A3) \quad \bar{u}(\cdot) \in L^{1, 2}_{2, F} (\mathbb{R}^k), \chi(\cdot) \mathbb{H}^T(\cdot) \in L^{1, 2}_{2, F} (\mathbb{R}^{k \times n}) \cap L^\infty([0, T] \times \Omega; \mathbb{R}^{k \times n}).\]

Theorem 26. Suppose that (A1)-(A3) are in force. Let \( \bar{u}(\cdot) \) be a singular optimal control in the classical sense, then we have

\[
\begin{align*}
    &\langle \chi(r) \mathbb{H}(r) b_u(r)(u - \bar{u}(r)), u - \bar{u}(r) \rangle + \langle \nabla \chi(r) \mathbb{H}(r) \sigma_u(r)(u - \bar{u}(r)), u - \bar{u}(r) \rangle \\
    &- \langle \chi(r) \mathbb{H}(r) \sigma_u(r)(u - \bar{u}(r)), \nabla \bar{u}(r) \rangle \geq 0, \text{ a.e., } r \in [0, T], \forall u \in U, \text{ P-a.s.}
\end{align*}
\]

Observe that the expression (35) is similar to (3.17) in [102]. Therefore, the proof is repeated as in Theorem 3.13 in Zhang et al. [102].

3.2.1 Example

We provide a concrete example to illustrate our theoretical result (Theorem 26) by looking at an example. If the FBSDEs considered in this paper are linear, it is possible to implement our principles directly. For convenience, we still adopt the notations introduced in Section 3.2.
and
\[
\begin{aligned}
-dP(t) &= -Q(t)\,dW(t), \\
P(T) &= 1.
\end{aligned}
\] (45)

We get immediately, the solutions to (44) and (45) are
\[
(p(t),q(t)) = (0,0), \quad (P(t),Q(t)) = (1,0),
\]
respectively. Hence, from the well-known Comparison Theorem (Lemma 5), we are able to claim that
\[
(\bar{x}(t),\bar{y}(t),\bar{z}(t),\bar{u}(t)) = (0,0,0,0)
\]
is the optimal quadruple. Indeed, one can easily check that
\[
\mathcal{H}_u(t,\bar{x}(t),\bar{y}(t),\bar{z}(t),\bar{u}(t),p(t),q(t),P(t)) \equiv 0
\]
and
\[
\mathcal{H}_{uu}(t,\bar{x}(t),\bar{y}(t),\bar{z}(t),\bar{u}(t),p(t),q(t),P(t)) \equiv 0.
\]
So \(\bar{u}(t) = 0\) is a singular control in the classical sense. Moreover, we compute
\[
\nabla \bar{u}(t) = 0, \quad \nabla \mathbb{H}(t) = 1, \quad \nabla \mathbb{H}(t) \equiv 0.
\]

Therefore, we get
\[
\begin{aligned}
&\langle \chi(r) \mathbb{H}(r) b_u(r) (u - \bar{u}(r)), u - \bar{u}(r) \rangle \\
&+ \langle \nabla \chi(r) \mathbb{H}(r) \sigma_u(r) (u - \bar{u}(r)), u - \bar{u}(r) \rangle \\
&- \langle \chi(r) \mathbb{H}(r) \sigma_u(r) (u - \bar{u}(r)), \nabla \bar{u}(r) \rangle \\
&= (u - \bar{u}(r))^2 \geq 0, \quad \forall u \in [-1,1], \quad \text{a.e. } r \in [0,T], \quad \text{P-a.s.,}
\end{aligned}
\]
which indicates that Theorem (26) always holds.

4 Singular Optimal Controls via Dynamic Programming Principle

In this section, we proceed our control problem from the viewpoint of DPP. From now on, we focus on the following
\[
\begin{aligned}
dX^{t,x,v,\xi}_s &= b\left(s,X^{t,x,v,\xi}_s,v_s\right)\,ds + \sigma\left(s,X^{t,x,v,\xi}_s,v_s\right)\,dW_s + Gd\xi_s, \\
dY^{t,x,v,\xi}_s &= -f\left(s,X^{t,x,v,\xi}_s,Y^{t,x,v,\xi}_s,Z^{t,x,v,\xi}_s,v_s\right)\,ds + Z^{t,x,v,\xi}_s\,dW_s - Kd\xi_s, \\
X^{t,x,v,\xi}_t &= x, \quad Y^{t,x,v,\xi}_T = \Phi\left(X^{t,x,v,\xi}_T,\xi\right), \quad \quad 0 \leq t \leq s \leq T.
\end{aligned}
\] (46)

Remark 28. We assume that \(G_{n\times m}\) and \(K_{1\times m}\) are deterministic matrices. On the one hand, from the derivations in Theorem 5.1 of [49], it is convenient to show the “inaction” region for singular control; On the other hand, we may regard \(Y^{t,x,v,\xi}_s + K\xi_s\) together as a solution, in this way, we are able to apply the classical Itô’s formula, avoiding the appearance of jump. We believe these assumptions can be removed properly, but at present, we consider constant only in our paper. Whilst in order to get the uniqueness of the solution to H-J-B inequality [47], we add the assumption \(K^i > k_0 > 0, 1 \leq i \leq m\). More details, see Theorem 2.2 in [109].
4.1 Verification Theorem via Viscosity Solutions

Zhang [109] has given a verification theorem for smooth solution of the following H-J-B inequality:

\[
\min \left( u_x^T (t, x) G + K, \frac{\partial}{\partial x} u (t, x) \right)
+ \min_{v \in U} L (t, x, v) u (t, x) + f (t, x, u (t, x), \nabla u (t, x) \sigma (t, x, v), v) = 0, \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^n.
\]

(47)

Lemma 29. Define

\[ D_t (u) := \{ x \in \mathbb{R}^n : u (t, x) < u (t, x + Gh) + Kh, \ h \in \mathbb{R}^m, \ h \neq 0 \} \]

Then the optimal state process \( X^{t, x, \hat{v}; \hat{\xi}} \) is continuous whenever \( (r, X^{t, x, \hat{v}; \hat{\xi}}) \in D_r (u) \). To be precise, we have

\[ P \left( \Delta X^{t, x, \hat{v}; \hat{\xi}}_r \neq 0, \ X^{t, x, \hat{v}; \hat{\xi}}_r \in D_r (u) \right) = 0, \ t \leq r \leq T. \]

The proof can be seen in Zhang [109].

Proposition 30. Suppose that \( V \) is a classical solution of the H-J-B inequality [17] such that for some \( l > 1 \), \( |V (t, x)| \leq C \left( 1 + |x|^l \right) \). Then for any \( [0, T] \times \mathbb{R}^n, \ (v, \xi) \in U : \)

\[ V (t, x) \leq J (t, x, v, \xi). \]

Furthermore, if there exists \( (\hat{v}, \hat{\xi}) \in U \) such that

\[ 1 = P \left\{ (r, X^{t, x, \hat{v}; \hat{\xi}}) \in D_r (V), \ 0 \leq r \leq T \right\}, \quad \text{(48)} \]

\[ 1 = P \left\{ \int_{[t, T]} [V^T (r, x) G + K] \ dr = 0 \right\}, \quad \text{(49)} \]

\[ 1 = P \left\{ (s, X^{t, x, \hat{v}; \hat{\xi}}) \in D_r (V), \ t \leq s \leq T : \right\} \]

\[ \hat{v}_s \in \min_{v \in U} \left\{ V_t \left( s, X^{t, x, \hat{v}; \hat{\xi}} S \right) + \mathcal{L} \left( s, X^{t, x, \hat{v}; \hat{\xi}} S, v \right) V \left( s, X^{t, x, \hat{v}; \hat{\xi}} S \right) \right\} \]

\[ + f \left( s, X^{t, x, \hat{v}; \hat{\xi}} S, V \left( t, X^{t, x, \hat{v}; \hat{\xi}} S \right) \right), \]

\[ \nabla V \left( s, X^{t, x, \hat{v}; \hat{\xi}} S \right) \sigma \left( s, X^{t, x, \hat{v}; \hat{\xi}} S, v \right) \right\} \]

\[ \text{(50)} \]

and

\[ P \left\{ V \left( s, X^{t, x, \hat{v}; \hat{\xi}} S \right) = V \left( s, X^{t, x, \hat{v}; \hat{\xi}} S \right) + K \Delta \xi_s, \ t \leq s \leq T \right\} = 1. \]

\[ \text{(51)} \]

Then

\[ V (t, x) = J (t, x; \hat{v} (\cdot), \hat{\xi} (\cdot)). \]

\[ \text{(52)} \]

In this section, we remove the unreal condition, smooth on value function by means of viscosity solution\(^2\). We first recall the definition of a viscosity solution for H-J-B variational inequality [17] from [28].

\(^2\)In the classical optimal stochastic control theory, the value function is a solution to the corresponding H-J-B equation whenever it has sufficient regularity (Fleming and Rishel [11], Krylov [57]). Nevertheless, when it is only known that the value function is continuous, then, the value function is a solution to the H-J-B equation in the viscosity sense (see Lions [28]).
Definition 31. Let \( u(t,x) \in C([0,T] \times \mathbb{R}^n) \) and \( (t,x) \in [0,T] \times \mathbb{R}^n \). For every \( \varphi \in C^{1,2}([0,T] \times \mathbb{R}^n) \)

(1) for each local maximum point \( (t_0,x_0) \) of \( u - \varphi \) in the interior of \([0,T] \times \mathbb{R}^n\), we have

\[
\min \left( \varphi_t^+ G + K, \frac{\partial \varphi}{\partial t} + \min_{\omega \in \mathcal{U}} \{ \mathcal{L} \varphi + f(t_0,x_0,\varphi,\nabla \varphi \sigma,\nu) \} \right) \geq 0
\]  

(53)

at \( (t_0,x_0) \), i.e., \( u \) is a subsolution.

(2) for each local minimum point \( (t_0,x_0) \) of \( u - \varphi \) in the interior of \([0,T] \times \mathbb{R}^n\), we have

\[
\min \left( \varphi_t^+ G + K, \frac{\partial \varphi}{\partial t} + \min_{\omega \in \mathcal{U}} \{ \mathcal{L} \varphi + f(t_0,x_0,\varphi,\nabla \varphi \sigma,\nu) \} \right) \leq 0
\]  

(54)

at \( (t_0,x_0) \), i.e., \( u \) is a supersolution.

(3) \( u(t,x) \in C([0,T] \times \mathbb{R}^n) \) is said to be a viscosity solution of (47) if it is both a viscosity sub and supersolution.

We have the other definition which will be useful to verify the viscosity solutions. Below, \( \mathbb{S}^n \) will denote the set of \( n \times n \) symmetric matrices.

Definition 32. Let \( u(t,x) \in C([0,T] \times \mathbb{R}^n) \) and \( (t,x) \in [0,T] \times \mathbb{R}^n \). We denote by \( \mathcal{P}^{2,+}u(t,x) \), the “parabolic superjet” of \( u \) at \( (t,x) \) the set of triples \( (p,q,X) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n \) which are such that

\[
u(s,y) \leq u(t,x) + p(s-t) + \langle q, x-y \rangle + \frac{1}{2} \langle X(y-x), y-x \rangle + o(|s-t| + |y-x|^2).
\]

Similarly, we denote by \( \mathcal{P}^{2,-}u(t,x) \), the “parabolic subjet” of \( u \) at \( (t,x) \) the set of triples \( (p,q,X) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n \) which are such that

\[
u(s,y) \geq u(t,x) + p(s-t) + \langle q, x-y \rangle + \frac{1}{2} \langle X(y-x), y-x \rangle + o(|s-t| + |y-x|^2).
\]

Define

\[
\mathcal{H}(t,x,q,X) = \left\{ \frac{1}{2} \text{Tr}(\sigma \sigma^* (t,x,v) X) + \langle q, b(t,x,v) \rangle + f(t,x,u(t,x),q\sigma(t,x,v)) \right\}.
\]

(55)

Definition 33. (i) It can be said \( u(t,x) \in C([0,T] \times \mathbb{R}^n) \) is a viscosity subsolution of (47) if \( u(T,x) \geq \Phi(x) \), \( x \in \mathbb{R}^n \), and at any point \( (t,x) \in [0,T] \times \mathbb{R}^n \), for any \( (p,q,X) \in \mathcal{P}^{2,+}u(t,x) \),

\[
\min \left( qG + K, p + \inf_{\omega \in \mathcal{U}} \mathcal{H}(t,x,q,X) \right) \geq 0
\]  

(56)

In other words, at any point \( (t,x) \), we have both \( qG + K \geq 0 \) and

\[
p + \mathcal{H}(t,x,q,X) \geq 0.
\]

(ii) It can be said \( u(t,x) \in C([0,T] \times \mathbb{R}^n) \) is a viscosity supersolution of (47) if \( u(T,x) \leq \Phi(x) \), \( x \in \mathbb{R}^n \), and at any point \( (t,x) \in [0,T] \times \mathbb{R}^n \), for any \( (p,q,X) \in \mathcal{P}^{2,-}u(t,x) \),

\[
\min \left( qG + K, p + \inf_{\omega \in \mathcal{U}} \mathcal{H}(t,x,q,X) \right) \leq 0
\]  

(57)
In other words, at any point where \( qG + K \geq 0 \), we have
\[
p + \inf_{v \in U} (t, x, q, X) \leq 0.
\]

(iii) It can be said \( u(t, x) \in C([0, T] \times \mathbb{R}^n) \) is a viscosity solution of (47) if it is both a viscosity sub and super solution.

**Remark 34.** Definition 31 and 33 are equivalent to each other. For more details, see Fleming and Soner [39], Lemma 4.1 (page 211).

**Lemma 35.** Let \( u \in C([0, T] \times \mathbb{R}^n) \) and \((t, x) \in [0, T] \times \mathbb{R}^n\) be given. Then:

1) \((p, q, X) \in P^{2,+}u(t, x)\) if and only if there exists a function \( \varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)\) such that \( u - \varphi \) attains a strict maximum at \((t, x)\) and
\[
(\varphi(t, x) , \varphi_t(t, x) , \varphi_x(t, x) , \varphi_{xx}(t, x)) = (u(t, x) , p, q, X).
\]

2) \((p, q, X) \in P^{2,-}u(t, x)\) if and only if there exists a function \( \varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)\) such that \( u - \varphi \) attains a strict minimum at \((t, x)\) and
\[
(\varphi(t, x) , \varphi_t(t, x) , \varphi_x(t, x) , \varphi_{xx}(t, x)) = (u(t, x) , p, q, X).
\]

More details can be seen in Lemma 5.4 and 5.5 in Yong and Zhou [100].

We have the following result:

**Proposition 36.** Assume that (A1)-(A2) are in force. Then there exists at most one viscosity solution of H-J-B inequality (47) in the class of bounded and continuous functions.

**Remark 37.** We have put somewhat strong assumptions, namely, \( b, \sigma, f \) are bounded. These conditions may be removed by modifying the idea by Ishii [53].

We need a generalized Itô’s formula. Define
\[
\mathcal{L}(t, x, v) \Psi = \frac{1}{2} \text{Tr} (\sigma \sigma^* (t, x, v) D^2 \Psi) + \langle D \Psi, b(t, x, v) \rangle,
\]
\((t, x, v) \in [0, T] \times \mathbb{R}^n \times U, \ \Psi \in C^{1,2}([0, T] \times \mathbb{R}^n)\).

For any \( \Psi \in C^{1,2}([0, T] \times \mathbb{R}^n; \mathbb{R})\), by virtue of Doléans–Dade–Meyer formula (see [49] [26]), we have
\[
\Psi(s, X_s) = \Psi(t, x) + \int_t^s \Psi_t(r, X_r) + \mathcal{L}(r, X_r, v) \Psi(r, X_r) \, dr
\]
\[
+ \int_t^s \Psi_x(r, X_r) \sigma(r, X_r, v_r) \, dW_r + \int_t^s \Psi_x(r, X_r) G d\xi_r
\]
\[
+ \sum_{t \leq r \leq s} \{ \Psi(r, X_{r+}) - \Psi(r, X_r) - \Psi_x(r, X_r) \Delta X_r \}.
\]

(58)

We begin to introduce a useful lemma.
Lemma 38. Assume that (A1)-(A2) are in force. Let \((t, x) \in [0, T) \times \mathbb{R}^n\) be fixed and let \((X^{t, x; u} (\cdot), u (\cdot))\) be an admissible pair. Define processes

\[
\begin{align*}
    z_1 (r) &= b (r, X^{t, x; u} (r), u (r)) , \\
    z_2 (r) &= \sigma (r, X^{t, x; u} (r), u (r)) \sigma^* (r, X^{t, x; u} (r), u (r)) , \\
    z_3 (r) &= f (r, X^{t, x; u} (r), Y^{t, x; u} (r), Z^{t, x; u} (r), u (r)) .
\end{align*}
\]

Then

\[
\lim_{h \to 0^+} \frac{1}{h} \int_t^{t+h} | z_i (r) - z_i (t) | \, dr = 0, \quad \text{a.e. } t \in [0, T], \quad i = 1, 2, 3. \tag{59}
\]

The proof can be found in [100].

Lemma 39. Let \(g \in C ([0, T])\). Extend \(g\) to \((-\infty, +\infty)\) with \(g (t) = g (T)\) for \(t > T\), and \(g (t) = g (0)\), for \(t < 0\). Suppose that there is a integrable function \(\rho \in L^1 ([0, T]; \mathbb{R})\) and some \(h_0 > 0\), such that

\[
\frac{g (t + h) - g (t)}{h} \leq \rho (t), \quad \text{a.e. } t \in [0, T], \quad h \leq h_0.
\]

Then

\[
g (\beta) - g (\alpha) \leq \int_\alpha^\beta \limsup_{h \to 0^+} \frac{g (t + h) - g (t)}{h} \, dr, \quad \forall 0 \leq \alpha \leq \beta \leq T.
\]

The proof can be seen in Zhang [107].

The main result in this section is the following.

Theorem 40 (Verification Theorem). Assume that (A1)-(A2) hold. Let \(v \in C ([0, T] \times \mathbb{R}^n)\), be a viscosity solution of the H-J-B equations \((47)\), satisfying the following conditions:

\[
\begin{align*}
    i) & \quad v (t + h, x) - v (t, x) \leq C (1 + |x|^m) h, \quad m \geq 0, \\
    & \quad \text{for all } x \in \mathbb{R}^n, 0 < t < t + h < T, \\
    ii) & \quad v \text{ is semiconcave, uniformly in } t, \text{i.e. there exists } C_0 \geq 0 \\
    & \quad \text{such that for every } t \in [0, T], \quad v (t, \cdot) - C_0 | \cdot |^2 \text{ is concave on } \mathbb{R}^n.
\end{align*}
\]

Then we have

\[
v (t, x) \leq J (s, y; u (\cdot), \xi (\cdot)), \tag{61}
\]

for any \((t, x) \in (0, T] \times \mathbb{R}^n\) and any \(u (\cdot) \times \xi (\cdot) \in \mathcal{U} (t, T)\). Furthermore, let \((t, x) \in (0, T] \times \mathbb{R}^n\) be fixed and let

\[
\left( \tilde{X}^{t, x; u; \xi} (\cdot), \tilde{Y}^{t, x; u; \xi} (\cdot), \tilde{Z}^{t, x; u; \xi} (\cdot), \tilde{u} (\cdot), \tilde{\xi} (\cdot) \right)
\]

be an admissible pair such that there exist a function \(\varphi \in C^{1,2} ([0, T]; \mathbb{R}^n)\) and a triple

\[
(\overline{\varphi}, \overline{\varphi}, \overline{\varphi}) \in \left( L^2_{F_t} ([t, T]; \mathbb{R}) \times L^2_{F_t} ([t, T]; \mathbb{R}^n) \times L^2_{F_t} ([t, T]; \mathbb{R}^n) \right) \tag{62}
\]

satisfying

\[
\begin{align*}
    (\tilde{p} (t), \tilde{q} (t), \tilde{\Theta} (t)) & \in \mathcal{P}^{2,+} \left( t, \tilde{X}^{t, x; u; \xi} (t) \right), \\
    \left( \frac{\partial \varphi}{\partial t} (t, \tilde{X}^{t, x; u; \xi} (t)), D_x \varphi \left( t, \tilde{X}^{t, x; u; \xi} (t) \right), D^2 \varphi \left( t, \tilde{X}^{t, x; u; \xi} (t) \right) \right) & = (\overline{\varphi} (t), \overline{\varphi} (t), \overline{\varphi} (t)), \\
    \varphi (t, x) & \geq v (t, x) \quad \forall (t_0, x_0) \neq (t, x), \quad \text{a.e. } t \in [0, T], \quad P\text{-a.s.} \\
    \tilde{p} (t) G + K & = 0, \quad \text{a.e. } t \in [0, T], \quad P\text{-a.s.}
\end{align*}
\]

\[
(63)
\]
and

\[ E \left[ \int_s^T \left[ \varphi(t) + \mathcal{H} \left( t, X^{t,x;\bar{u},\bar{\xi}}(t), \varphi(t), \bar{\varphi}(t), \mathfrak{p}(t), \mathfrak{q}(t) \right) \right] dt \right] \leq 0, \quad (64) \]

where \( \varphi(t) = \varphi \left( t, X^{t,x;\bar{u},\bar{\xi}}(t) \right) \) and \( \mathcal{H} \) is defined in (52). Then \( (X^{t,x;\bar{u},\bar{\xi}}(\cdot), \bar{u}(\cdot), \bar{\xi}(\cdot)) \) is an optimal pair.

In order to prove Theorem 40 we need the following lemma:

**Lemma 41.** Let \( v \) be a viscosity subsolution of the H-J-B equations (47), satisfying the following (60). Then we have

\[ E \left[ \frac{1}{h} \left( v \left( s + h, X^{t,x;\bar{u},\bar{\xi}}(s + h) \right) - v \left( s, X^{t,x;\bar{u},\bar{\xi}}(s) \right) \right) \right] \leq \rho(s), \quad (65) \]

where \( \rho(s) \in L^1([t,T] : \mathbb{R}) \).

The proof can be seen in the Appendix.

**Proof of Theorem 40** Firstly, (61) follows from the uniqueness of viscosity solutions of the H-J-B equations (47). It remains to show that \( (X^{t,x;\bar{u},\bar{\xi}}(\cdot), \bar{u}(\cdot), \bar{\xi}(\cdot)) \) is an optimal.

We now fix \( t_0 \in [s,T] \) such that (62) and (63) hold at \( t_0 \) and (59) holds at \( t_0 \) for

\[ \begin{align*}
&z_1(\cdot) = \mathfrak{p}(\cdot), \\
&z_2(\cdot) = \varphi(\cdot) \mathfrak{p}(\cdot), \\
&z_3(\cdot) = \mathfrak{q}(\cdot).
\end{align*} \]

We claim that the set of such points is of full measure in \([s,T]\) by Lemma 38. Now we fix \( \omega_0 \in \Omega \) such that the regular conditional probability \( P \left( \cdot | \mathcal{F}^{\omega_0}_0 \right) (\omega_0) \), given \( \mathcal{F}^{\omega_0}_0 \) is well defined. In this new probability space, the random variables \( X^{t,x;\bar{u},\bar{\xi}}(t_0), \mathfrak{p}(t_0), \mathfrak{q}(t_0), \mathfrak{G}(t_0) \) are almost surely deterministic constants and equal to

\[ X^{t,x;\bar{u},\bar{\xi}}(t_0,\omega_0), \mathfrak{p}(t_0,\omega_0), \mathfrak{q}(t_0,\omega_0), \mathfrak{G}(t_0,\omega_0), \]

respectively. We remark that in this probability space the Brownian motion \( W \) is still the a standard Brownian motion although now \( W(t_0) = W(t_0,\omega_0) \) almost surely. The space is now equipped with a new filtration \( \{ \mathcal{F}^\omega_s \}_{s \leq t \leq T} \) and the control process \( \mathfrak{p}(\cdot) \) is adapted to this new filtration. For \( P \) a.s. \( \omega_0 \) the process \( X^{t,x;\bar{u},\bar{\xi}}(\cdot) \) is a solution of (1.1) on \([t_0,T]\) in \((\Omega, \mathcal{F}, P \left( \cdot | \mathcal{F}^{\omega_0}_0 \right) (\omega_0))\) with inital condition \( X^{t,x;\bar{u},\bar{\xi}}(t_0) = X^{t,x;\bar{u},\bar{\xi}}(t_0,\omega_0) \).

Then on the probability space \((\Omega, \mathcal{F}, P \left( \cdot | \mathcal{F}^{\omega_0}_0 \right) (\omega_0))\), we are going to apply Itô’s formula to \( \varphi \) on \([t_0, t_0 + h]\) for any \( h > 0 \),

\[ \varphi \left( t_0 + h, X^{t,x;\bar{u},\bar{\xi}}(t_0 + h) \right) - \varphi \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) = \int_{t_0}^{t_0 + h} \left[ \frac{\partial \varphi}{\partial r} \left( r, X^{t,x;\bar{u},\bar{\xi}}(r) \right) + \left( D_x \varphi \left( r, X^{t,x;\bar{u},\bar{\xi}}(r) \right), \mathfrak{p}(r) \right) \right] dr + \frac{1}{2} \text{tr} \left( \mathfrak{p}(r)^* D_{xx} \varphi \left( r, X^{t,x;\bar{u},\bar{\xi}}(r) \right) \mathfrak{p}(r) \right) \right] dr \]

\[ + \int_{t_0}^{t_0 + h} \varphi_x \left( r, X^{t,x;\bar{u},\bar{\xi}}(r) \right) Gd\xi_r + \int_{t_0}^{t_0 + h} \left( D_x \varphi \left( r, X^{t,x;\bar{u},\bar{\xi}}(r) \right), \mathfrak{p}(r) \right) dW_r \]

32
From Property 1, note that

\[
\begin{align*}
\mathbb{E} \left[ v \left( s + \theta, \bar{X}^{t,x;u,\xi} (s + \theta) \right) - v \left( s, \bar{X}^{t,x;u,\xi} (s) \right) \right] \\
\geq \mathbb{E} \left[ \varphi \left( s + \theta, \bar{X}^{t,x;u,\xi} (s + \theta) \right) - \varphi \left( s, \bar{X}^{t,x;u,\xi} (s) \right) \right]
\end{align*}
\]

\[
= \mathbb{E} \left[ \int_{s}^{s+\theta} \left( \frac{\partial \varphi}{\partial t} \right) \left( r, \bar{X}^{t,x;u,\xi} (r) \right) + \left\langle D_{x} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right), \bar{G} (r) \right\rangle \\
+ \frac{1}{2} tr \left\{ \bar{\sigma} (r)^{\ast} D_{xx} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) \bar{\sigma} (r) \right\} \right] dr \\
+ \int_{t_{0}}^{t_{0}+h} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) Gd\xi_{r} + \sum_{s \leq r \leq s+\theta} \left\{ \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) - \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) \Delta \bar{X}^{t,x;u,\xi} (r) \right\}
\]

\[
\geq \mathbb{E} \left[ \int_{s}^{s+\theta} \left( \frac{\partial \varphi}{\partial t} \right) \left( r, \bar{X}^{t,x;u,\xi} (r) \right) + \left\langle D_{x} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right), \bar{G} (r) \right\rangle \\
+ \frac{1}{2} tr \left\{ \bar{\sigma} (r)^{\ast} D_{xx} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) \bar{\sigma} (r) \right\} \right] dr - \int_{s}^{s+\theta} Kd\xi_{r}. \quad (66)
\]

We now handle the last two terms.

\[
\begin{align*}
\mathbb{E} \left[ \int_{t_{0}}^{t_{0}+h} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) Gd\xi_{r} \\
+ \sum_{t_{0} \leq r \leq t_{0}+h} \left\{ \varphi \left( r, \bar{X}^{t,x;u,\xi} (r+) \right) - \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) \right\} \\
- \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) \Delta \bar{X}^{t,x;u,\xi} (r) \right] 
\end{align*}
\]

From Property 1 note that

\[
\Delta \bar{X}^{t,x;u,\xi} (r) = G\xi_{r} \text{ and } \bar{X}^{t,x;u,\xi} (r+) = \bar{X}^{t,x;u,\xi} (r) + \Delta \bar{X}^{t,x;u,\xi} (r) = \bar{X}^{t,x;u,\xi} (r) + G\xi_{r}. \text{ Thus}
\]

\[
-\mathbb{E} \left[ \int_{t_{0}}^{t_{0}+h} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) Gd\xi_{r} \right] + \mathbb{E} \left[ \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) \Delta \bar{X}^{t,x;u,\xi} (r) \right]
\]

\[
= -\mathbb{E} \left[ \int_{t_{0}}^{t_{0}+h} \varphi \left( r, \bar{X}^{t,x;u,\xi} (r) \right) Gd\xi_{r} \right]
\]

\[
= \mathbb{E} \left[ \int_{t_{0}}^{t_{0}+h} Kd\xi_{r} \right]. \quad (67)
\]
We now deal the term

\[-E \left[ \sum_{t_0 \leq r \leq t_0 + h} \left\{ \varphi \left( r, X^{t,x;\bar{u},\bar{\xi}}(r+), r, X^{t,x;\bar{u},\bar{\xi}}(r) \right) \right\} \right] \]

\[= -E \left[ \sum_{t_0 \leq r \leq t_0 + h} \left\{ \int_0^1 \varphi_x \left( r, X^{t,x;\bar{u},\bar{\xi}}(r) + \theta \Delta X^{t,x;\bar{u},\bar{\xi}}(r) \right) G \Delta \bar{\xi}_r d\theta \right\} \right] \]

\[= E \left[ K \Delta \bar{\xi}_r \right]. \]  

Combining (67) and (68), we have

\[\frac{1}{h} \limsup_{h \to 0+} E^{F_{t_0}\omega} \left[ v \left( t_0 + h, X^{t,x;\bar{u},\bar{\xi}}(t_0 + h) \right) - v \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \right] \]

\[\leq \frac{1}{h} \limsup_{h \to 0+} E^{F_{t_0}\omega} \left[ v \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \right] \]

\[= \frac{1}{2} \text{tr} \left\{ \sigma(t_0)^* D_{xx} \varphi \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \sigma(t_0) \right\} - K d\xi_{t_0} \]  

Letting \( h \to 0 \), and employing the similar delicate method as in the proof of Theorem 4.1 of Gozzi et al. [46], we have

\[\frac{1}{h} \limsup_{h \to 0+} E^{F_{t_0}\omega} \left[ v \left( t_0 + h, X^{t,x;\bar{u},\bar{\xi}}(t_0 + h) \right) - v \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \right] \]

\[\leq \frac{1}{h} \limsup_{h \to 0+} E^{F_{t_0}\omega} \left[ V \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \right] \]

\[= \frac{1}{2} \text{tr} \left\{ \sigma(t_0)^* D_{xx} \varphi \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \sigma(t_0) \right\} - K d\xi_{t_0} \]

From Lemma III that there exist

\[\rho \in L^1(t_0, T; \mathbb{R}) \text{ and } \rho_1 \in L^1(\Omega; \mathbb{R})\]

such that

\[E \left[ \frac{1}{h} \left[ v \left( t + h, X^{t,x;\bar{u},\bar{\xi}}(t + h) \right) - v \left( t, X^{t,x;\bar{u},\bar{\xi}}(t) \right) \right] \right] \]

\[\leq \rho(t), \text{ for } h \leq h_0, \text{ for some } h_0 > 0 \]  

and

\[E^{F_{t_0}\omega} \left[ \frac{1}{h} \left[ v \left( t + h, X^{t,x;\bar{u},\bar{\xi}}(t + h) \right) - v \left( t, X^{t,x;\bar{u},\bar{\xi}}(t) \right) \right] \right] \]

\[\leq \rho_1(\omega), \text{ for } h \leq h_0, \text{ for some } h_0 > 0. \]  

holds, respectively. By virtue of Fatou’s Lemma, noting (71), we obtain

\[\limsup_{h \to 0+} \frac{1}{h} E \left[ v \left( t_0 + h, X^{t,x;\bar{u},\bar{\xi}}(t_0 + h) \right) - v \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \right] \]

\[= \limsup_{h \to 0+} \frac{1}{h} E^{F_{t_0}\omega} \left\{ v \left( t_0 + h, X^{t,x;\bar{u},\bar{\xi}}(t_0 + h) \right) - v \left( t_0, X^{t,x;\bar{u},\bar{\xi}}(t_0) \right) \right\} \]
\begin{align*}
\leq & \mathbb{E}\left[ \limsup_{h \to 0^+} \frac{1}{h} \mathbb{E}_{\omega}^{X_{t_0}} \left\{ v \left( t_0 + h, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} \right) (t_0 + h) - v \left( t_0, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} \left( t_0 \right) \right) \right\} \right] \\
\leq & \mathbb{E}\left[ \Phi(t_0) + \mathbb{E} \left\{ \mathcal{D} \left\{ \mathcal{D} \left( t_0 \right) \right\} - K d\tilde{\xi}_0 \right\} \right], \tag{72}
\end{align*}
for a.e. $t_0 \in [t, T]$. Then the rest of the proof goes exactly as in [40]. We apply Lemma 39 to
\begin{equation*}
g(t) = \mathbb{E} \left[ v \left( t, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (t) \right) \right],
\end{equation*}
using (70), (63) and (72) to get
\begin{align*}
\mathbb{E} \left[ v \left( T, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (T) \right) - v (s, y) \right] \\
\leq \mathbb{E} \left[ \int_s^T \mathcal{F} (t) + \mathbb{E} \left\{ \mathcal{D} \left( t_0 \right) \right\} - K d\tilde{\xi}_0 \right] \\
\leq -\mathbb{E} \left[ \int_s^T \mathcal{F} (t) dt + K d\tilde{\xi}_0 \right].
\end{align*}
From this we claim that
\begin{align*}
v(t, x) & \geq \mathbb{E} \left[ v \left( T, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (T) \right) + \int_t^T \mathcal{F} (r) dr + \int_s^T K d\tilde{\xi}_0 \right] \\
& = \mathbb{E} \left\{ \Phi \left( X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (T) \right) + \int_t^T \mathcal{F} (r) dr + \int_s^T K d\tilde{\xi}_0 \right\},
\end{align*}
where
\begin{equation*}
\mathcal{F} (r) = f \left( r, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (r), v \left( r, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (r) \right), \sigma \left( r, X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (r) \right), \tilde{u} \right).
\end{equation*}
Thus, combining the above with the first assertion [63], we prove the \( \left( X_{t_0}^t, x, \tilde{u}, \tilde{\xi} (\cdot), \tilde{\pi} (\cdot) \right) \) is an optimal pair. The proof is thus completed.

**Remark 42.** The condition (64) is just equivalent to the following:
\begin{align*}
\bar{\pi} (s) & = \min_{u \in U} \mathcal{H} \left( ts, X_{t_0}^{t, x, \tilde{u}, \tilde{\xi}} (s), \bar{\pi} (s), \bar{\pi} (s), \bar{\pi} (s), u \right) \\
& = \mathcal{H} \left( s, X_{t_0}^{t, x, \tilde{u}, \tilde{\xi}} (s), \bar{\pi} (s), \bar{\pi} (s), \bar{\pi} (s), \bar{\pi} (s) \right), \tag{73}
a.e. \ s \in [t, T], \ P-a.s.,
\end{align*}
where \( \bar{\pi} (t) \) is defined in Theorem 40. This is easily seen by recalling the fact that \( v \) is the viscosity solution of (47).

**Remark 43.** Clearly, Theorem 40 is expressed in terms of parabolic superjet. One could naturally seek whether a similar result holds for parabolic subjet. The answer was yes for the deterministic case (in terms of the first-order parabolic subjet; see Theorem 3.9 in [100]). Unfortunately, as claimed in Yong and Zhou [100], the answer is that the statement of Theorem 40 is no longer valid whenever the parabolic superjet in (63) is replaced by the parabolic subjet.
Then, it holds that

\[ \mathbb{E} \tilde{p}(s) \leq -\mathbb{E} \left[ \mathcal{H} \left( s, \bar{X}^{t,x;\bar{u},\xi}(s), \tilde{q}(s), \Theta(t) \right) \right], \text{ a.e. } s \in [t,T]. \]

**Proof** On the one hand, let \( s \in [t,T] \) and \( \omega \in \Omega \) such that \( (\tilde{p}(s), \tilde{q}(s), \Theta(s)) \in \mathcal{P}^{2,-}\left( s, \bar{X}^{t,x;\bar{u},\xi}(s) \right) \). By Lemma 35, we have a test function \( \varphi \in C^{1,2}([0,T] \times \mathbb{R}^n) \) with \( (s,x) \in [0,T] \times \mathbb{R}^n \) and \( (p,q,\Theta) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{S} \) such that \( v - \varphi \) achieves its minimum at \( (s,\bar{X}^{t,x;\bar{u},\xi}(s)) \)

\[
\left( \frac{\partial \varphi}{\partial t} \left( s, \bar{X}^{t,x;\bar{u},\xi}(s) \right), D_x \varphi \left( s, \bar{X}^{t,x;\bar{u},\xi}(s) \right), D^2 \varphi \left( s, \bar{X}^{t,x;\bar{u},\xi}(s) \right) \right) = (\tilde{p}(s), \tilde{q}(s), \Theta(s))
\]

holds. Then for sufficiently small \( \theta > 0 \), a.e. \( s \in [t,T] \).

\[
\mathbb{E} \left[ v \left( s + \theta, \bar{X}^{t,x;\bar{u},\xi}(s + \theta) \right) - v \left( s, \bar{X}^{t,x;\bar{u},\xi}(s) \right) \right] \\
\geq \mathbb{E} \left[ \varphi \left( s + \theta, \bar{X}^{t,x;\bar{u},\xi}(s + \theta) \right) - \varphi \left( s, \bar{X}^{t,x;\bar{u},\xi}(s) \right) \right] \\
= \mathbb{E} \left[ \int_s^{s+\theta} \left[ \frac{\partial \varphi}{\partial t} \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right) + \langle D_x \varphi \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right), \bar{b}(r) \rangle \right] dr \\
+ \frac{1}{2} \int_{r_0}^{r_0+h} \left\{ \bar{q}(r) \bar{X}^{t,x;\bar{u},\xi}(r), D_x \varphi \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right), \bar{b}(r) \right\} \right] dr \\
+ \int_{r_0}^{r_0+h} \varphi_x \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right) G \xi_r + \sum_{s \leq r \leq s+\theta} \left\{ \varphi \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right) - \varphi_x \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right) \Delta \bar{X}^{t,x;\bar{u},\xi}(r) \right\} \right] \\
\geq \mathbb{E} \left[ \int_s^{s+\theta} \left[ \frac{\partial \varphi}{\partial t} \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right) + \langle D_x \varphi \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right), \bar{b}(r) \rangle \right] dr \\
+ \frac{1}{2} \int_{r_0}^{r_0+h} \left\{ \bar{q}(r) \bar{X}^{t,x;\bar{u},\xi}(r), D_x \varphi \left( r, \bar{X}^{t,x;\bar{u},\xi}(r) \right), \bar{b}(r) \right\} \right] dr - \int_s^{s+\theta} K d\xi_r \right].
\]

The last inequality comes from the derivation in Theorem 40 by means of the condition \([41]\). On the other hand, since \( \left( \bar{X}^{t,x;\bar{u},\xi}(\cdot), \bar{Y}^{t,x;\bar{u},\xi}(\cdot), \bar{Z}^{t,x;\bar{u},\xi}(\cdot) \right) \) is optimal,
by DPP of optimality, it yields
\[v(\tau, X^{t,x;u,\xi}(\tau)) = \mathbb{E}^{F}_{\omega} \left[ \Phi(X^{t,x;u,\xi}(T)) + \int_{\tau}^{T} \bar{f}(r) dr + \int_{\tau}^{T} Kd\xi_{r} \right], \quad \forall \tau \in [t, T], \text{ P-a.s.,} \]
which implies that
\[\mathbb{E} \left[ v(s + \theta, X^{t,x;u,\xi}(s + \theta)) - v(s, X^{t,x;u,\xi}(s)) \right] = -\left[ \int_{s}^{s+\theta} \bar{f}(r) dr + \int_{\tau}^{T} Kd\xi_{r} \right]. \quad (76)\]
Therefore, from (73) and (76), after some simple computation, we get
\[\mathbb{E} [\bar{p}(s)] \leq -\mathbb{E} \left[ \frac{D_{\sigma}}{dt} \left( s, X^{t,x;u,\xi}(s) \right) + \left\{ D_{x} \varphi \left( s, X^{t,x;u,\xi}(s) \right), \bar{b}(s) \right\} + \frac{1}{2} \text{tr} \left\{ \mathbf{\Sigma}(s)^{T} D_{x} \varphi \left( s, X^{t,x;u,\xi}(s) \right) \mathbf{\Sigma}(s) - \bar{f}(s) \right\}, \text{ a.e. } s \in [t, T], \]
where
\[\bar{f}(r) = f\left( r, X^{t,x;u,\xi}(r), v\left( r, X^{t,x;u,\xi}(r) \right), \bar{q}(r) \sigma\left( r, X^{t,x;u,\xi}(r), \bar{u} \right) \right).\]
We thus complete the proof. \(\square\)

### 4.2 Optimal Feedback Controls

In this subsection, we describe the method to construct optimal feedback controls by the verification Theorem 40 obtained. First, let us recall the definition of admissible feedback controls.

**Definition 45.** A measurable function \((u, \xi)\) from \([0, T] \times \mathbb{R}^{n}\) to \(U \times [0, \infty)^{m}\) is called an admissible feedback control pair if for any \((t, x) \in [0, T) \times \mathbb{R}^{n}\) there is a weak solution \(X^{t,x;u,\xi}(\cdot)\) of the following SDE:

\[
\begin{cases}
    dX^{t,x;u,\xi}(r) = b\left( r, X^{t,x;u,\xi}(r), u(r) \right) dr + \sigma\left( r, X^{t,x;u,\xi}(r), u(r) \right) dW(r) + Gd\xi_{r}, \\
    dY^{t,x;u,\xi}(r) = -f\left( r, X^{t,x;u,\xi}(r), Y^{t,x;u,\xi}(r), u(r) \right) dr + dM^{t,x;u,\xi}(r) - Kd\xi_{r}, \\
    X^{t,x;u,\xi}(t) = x, \quad Y^{t,x;u,\xi}(T) = \Phi\left( X^{t,x;u,\xi}(T) \right), \quad r \in [t, T],
\end{cases}
\]

where \(M^{t,x;u,\xi}\) is an \(\mathbb{R}\)-valued \(\mathcal{F}^{t,x;u,\xi}\)-adapted right continuous and left limit martingale vanishing in \(t = 0\) which is orthogonal to the driving Brownian motion \(W\). Here \(\mathcal{F}^{t,x;u,\xi} = \left( \mathcal{F}_{s}^{t,x;u,\xi} \right)_{s \in [t, T]}\) is the smallest filtration and generated by \(X^{t,x;u,\xi}\), which is such that \(X^{t,x;u,\xi}\) is \(\mathcal{F}^{t,x;u,\xi}\)-adapted. Obviously, \(M^{t,x;u,\xi}\) is a part of the solution of BSDE of (77). Simultaneously, we suppose that \(f\) satisfies the Lipschitz condition.

\[|f(t, x, y, u) - f(t, x', y', u')| \leq L \left( |x - x'| + |y - y'| + |u - u'| \right), \quad x, x', y, y', u, u' \in U.\]

An admissible feedback control pair \((u^{*}, \xi^{*})\) is called optimal if

\[\left( X^{*}(\cdot; t, x), Y^{*}(\cdot; t, x), u^{*}(\cdot, X^{*}(\cdot; t, x)), \xi^{*}(\cdot, X^{*}(\cdot; t, x)) \right)\]

is optimal for each \((t, x)\) is a solution of (77) corresponding to \((u^{*}, \xi^{*})\).
Theorem 46. Let \((u^*, \xi^*)\) be an admissible feedback control and \(p^*, q^*\), and \(\Theta^*\) be measurable functions satisfying
\[
(p^*(t, x), q^*(t, x), \Theta(t, x)) \in \mathcal{P}^{2,+} v(t, x)
\]
for all \((t, x) \in [0, T] \times \mathbb{R}^n\). If
\[
p^*(t, x) + \mathcal{H}(t, x, V(t, x), q^*(t, x), \Theta^*(t, x), u^*(t, x))
\]
\[
= \inf_{(p, q, u) \in \mathcal{P}^{2,+} v(t, x) \times U} [p + \mathcal{H}(t, x, V(t, x), q, \Theta, u)] - q^*(t, x) G + K \geq 0
\]
and
\[
(78)
\]
for all \((t, x) \in [0, T] \times \mathbb{R}^n\), then \((u^*, \xi^*)\) is singular optimal control pair.

**Proof** From Theorem 40 we get the desired result.

Remark 47. In Eq. (77), \(Y^{t,x;u} (\cdot)\) is obviously determined by \((X^{t,x;u}(\cdot), u(\cdot), \xi(\cdot))\). Hence, we need to investigate the conditions imposed in Theorem 40 to ensure the existence and uniqueness of \(X^{t,x;u}(\cdot)\) in law and the measurability of the multifunctions \((t, x) \rightarrow \mathcal{P}^{2,+} v(t, x)\) to obtain \((p^*(t, x), q^*(t, x), \Theta(t, x)) \in \mathcal{P}^{2,+} v(t, x)\) that minimizes (78) by virtue of the celebrated Filippov’s Lemma (see [100]).

### 4.3 The Connection between DPP and MP

In Section 3 we have obtained the first and second order adjoint equations. In this part, we shall investigate the connection between the general DPP and the MP for such singular controls problem without the assumption that the value is sufficient smooth. By associated adjoint equations and delicate estimates, it is possible to establish the set inclusions among the super- and sub-jets of the value function and the first-order and second-order adjoint processes as well as the generalized Hamiltonian function.

Theorem 48. Assume that (A1)-(A2) are in force. Suppose that \((\bar{u}, \bar{\xi})\) be a singular optimal controls, \(v(\cdot, \cdot)\) is a value function, and \((X^{t,x;\bar{u},\bar{\xi}}(\cdot), Y^{t,x;\bar{u},\bar{\xi}}(\cdot), Z^{t,x;\bar{u},\bar{\xi}}(\cdot), \bar{u}(\cdot), \bar{\xi}(\cdot))\)
is optimal trajectory. Let \((p, q) \in S^2(0, T; \mathbb{R}^n) \times \mathcal{M}^2(0, T; \mathbb{R}^n)\) and \((P, Q) \in S^2(0, T; \mathbb{R}^{n \times n}) \times \mathcal{M}^2(0, T; \mathbb{R}^{n \times n})\) be the adjoint equations (22), (22), respectively. Then, we have
\[
P \{ K_{(i)} + p(t) G_{(i)} (t) \geq 0, \ t \in [0, T], \ \forall i \} = 1,
\]
\[
\{ p(s) \} \times [P(s), \infty) \subseteq \mathcal{P}^{2,+} v \left( t, X^{t,x;\bar{u},\bar{\xi}} (s) \right),
\]
\[
\mathcal{P}^{2,-} v \left( t, X^{t,x;\bar{u},\bar{\xi}} (s) \right) \subseteq \{ p(s) \} \times [-\infty, P(s)], \ a.e. \ s \in [t, T], \ P-a.s..
\]

**Proof** From Theorem 7 and Proposition 13 we get the first part of (79). From Theorem 3.1 in Nie, Shi and Wu [78], we get the second and third results of (79).
5 Concluding remarks

In this paper, on the one hand, we have derived a second order pointwise necessary condition for singular optimal control in classical sense of FBSDEs with convex control domain by means of the variation equations and two adjoint equations, which is separately extends the work by Zhang and Zhang \[102\] to stochastic recursive case, and Hu \[50\] to pointwise case in the framework of Malliavin calculus. A new necessary condition for singular control has been obtained. Moreover, we investigate the verification theorem for optimal controls via viscosity solution and establish the connection between the adjoint equations and value function also in viscosity solution.

There are still several interesting topics will be discussed as follows:

• As an important issue, the existence of optimal singular controls has never been exploited. Haussmannand and Suo \[48\] apply the compactification method to study the classical and singular control problem of Itô’s type of stochastic differential equation, where the problem is reformulated as a martingale problem on an appropriate canonical space after the relaxed form of the classical control is introduced. Under some mild continuity assumptions on the data, they obtain the existence of optimal control by purely probabilistic arguments. Note that, in the framework of BSDE with singular control, the trajectory of \(Y\) seems to be a càdlàg process (from French, for right continuous with left hand limits). Hence, we may consider \(Y\) in some space with appropriate topologies, for instance, Skorokhod \(M_1\) topology or Meyer-Zheng topology (see \[42\]) to obtain the convergence of probability measures deduced by \(Y\) involving relaxed control. Related work from the technique of PDEs can be seen in \[20, 24\] references therein. From Wang \[95\], one may construct the optimal control via the existence of diffusion with reflections (see \[29\]). However, it is interesting to extend this result to FBSDEs.

• The matrices \(K, G\) are deterministic. It is also interesting to extend this restriction to time varying matrices, even the generator \(b, \sigma, f\) involving the singular control. Whenever the coefficients are random, the H-J-B inequality will become stochastic PDEs. No doubt, stochastic viscosity solution will be applied. For this direction, reader can refer to Buckdahn, Ma \[21, 22\] and Qiu \[89\].

• As for the general cases, i.e., the control regions are assumed to be nonconvex and both the drift and diffusion terms depend on the control variable. Indeed, such a mathematical model, from view point of application, is more reasonable and urgent in many real-life problems (for instance, queueing systems in heavy traffic and some finance models for which the controls may impact the uncertainty, etc). In near future, we shall remove the condition of convex control region, employing the idea developed by Zhang et al. \[103\]. It is worth mentioning that the analysis in \[103\] is much more complicated. Some new and useful tools, such as the multilinear function valued stochastic processes, the BSDE for these processes are introduced. It will be interesting to borrow these tools to investigate the singular optimal controls problems for FBSDEs, which will definitely promote and enrich the theories of FBSDEs.
A Proofs of Lemmas

Proof of Lemma 8 We first prove the continuity of solution depending on parameter. Let

\begin{align*}
\hat{X}^\alpha (s) &= X_{0,x;\bar{u},\xi}^\alpha (s) - X_{0,x;\bar{u},\bar{\xi}}^\alpha (s), \\
\hat{Y}^\alpha (s) &= Y_{0,x;\bar{u},\xi}^\alpha (s) - Y_{0,x;\bar{u},\bar{\xi}}^\alpha (s), \\
\hat{Z}^\alpha (s) &= Z_{0,x;\bar{u},\xi}^\alpha (s) - Z_{0,x;\bar{u},\bar{\xi}}^\alpha (s), \\
\hat{Y}_{0,x;\bar{u},\xi}^\alpha (s) &= Y_{0,x;\bar{u},\xi}^\alpha (s) + K\xi^\alpha (s), \\
\hat{Y}_{0,x;\bar{u},\bar{\xi}}^\alpha (s) &= Y_{0,x;\bar{u},\bar{\xi}}^\alpha (s) + K\bar{\xi}^\alpha (s), \quad s \in [0,T].
\end{align*}

We have to show that \( (\hat{X}^\alpha (t), \hat{Y}^\alpha (t), \hat{Z}^\alpha (t)) \) converges to 0 in \( \mathcal{N}^2 [0,T] \) as \( \alpha \to 0 \). Put

\( \hat{Y}^\alpha (s) = \hat{Y}_{0,x;\bar{u},\xi}^\alpha (s) - \hat{Y}_{0,x;\bar{u},\bar{\xi}}^\alpha (s), \quad s \in [0,T]. \)

Consider

\( \hat{Y}^\alpha (s) = \hat{Y}^\alpha (T) + \int_t^T \hat{f} (s) \, ds - \int_t^T \hat{Z}^\alpha (s) \, dW (s), \)

where

\( \hat{f} (s) = f \left( s, X_{0,x;\bar{u},\xi}^\alpha (s), \hat{Y}_{0,x;\bar{u},\xi}^\alpha (s) - K\xi^\alpha (s), Z_{0,x;\bar{u},\xi}^\alpha (s), \bar{u} (s) \right) - f \left( s, X_{0,x;\bar{u},\bar{\xi}}^\alpha (s), \hat{Y}_{0,x;\bar{u},\bar{\xi}}^\alpha (s) - K\bar{\xi}^\alpha (s), Z_{0,x;\bar{u},\bar{\xi}}^\alpha (s), \bar{u} (s) \right). \)

From Lemma 8 we have

\[ \mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| \hat{Y}^\alpha (s) \right|^2 + \int_0^T \left| \hat{Z}^\alpha (s) \right|^2 \, ds \right] \leq C \mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| \hat{X}^\alpha (s) \right|^2 + \alpha^2 \left| \xi (T) - \bar{\xi} (T) \right|^2 \right]. \tag{85} \]

From standard estimates and the Burkholder-Davis-Gundy inequality, we have

\[ \mathbb{E} \left[ \left| \hat{X}^\alpha (s) \right|^2 \right] \leq 3T \mathbb{E} \left[ \int_0^s \left| b \left( r, X_{0,x;\bar{u},\xi}^\alpha (r), \bar{u} (r) \right) - b \left( r, X_{0,x;\bar{u},\bar{\xi}}^\alpha (r), \bar{u} (r) \right) \right|^2 \, dr \right] \\
+ 3 \mathbb{E} \left[ \int_0^s \left| \sigma \left( r, X_{0,x;\bar{u},\xi}^\alpha (r), \bar{u} (r) \right) - \sigma \left( r, X_{0,x;\bar{u},\bar{\xi}}^\alpha (r), \bar{u} (r) \right) \right|^2 \, dr \right] \\
+ 3\alpha^2 \mathbb{E} \left[ \int_0^s \left| G (r) (\xi (r) - \bar{\xi} (r)) \right|^2 \, dr \right] \leq C \mathbb{E} \left[ \int_0^s \left| \hat{X}^\alpha (r) \right|^2 \, dr \right] + C\alpha^2 \mathbb{E} \left| \xi (T) - \eta (T) \right|^2, \quad t \in [0,T]. \tag{86} \]

From (88), (89) and Gronwall’s lemma, we claim that \( (\hat{X}^\alpha (t), \hat{Y}^\alpha (t), \hat{Z}^\alpha (t)) \) converges to 0 in \( \mathcal{N}^2 [0,T] \) as \( \alpha \to 0 \). Next, set

\[ \Delta X^\alpha (s) = \frac{\hat{X}^\alpha (s)}{\alpha}, \quad \Delta Y^\alpha (s) = \frac{\hat{Y}^\alpha (s)}{\alpha}, \quad \Delta Z^\alpha (s) = \frac{\hat{Z}^\alpha (s)}{\alpha}. \]
Then,
\[
\begin{align*}
\frac{d \Delta X^\alpha (t)}{dt} &= \frac{\hat{b}_r}{\alpha} dt + \frac{\hat{\sigma}_r}{\alpha} dW (r), \quad \Delta X^\alpha (t) = 0, \\
\frac{d \Delta Y^\alpha (s)}{dt} &= - \frac{l(s)}{\alpha} dt + \hat{Z}^\alpha (s) dW (s), \\
\Delta Y^\alpha (T) &= \Phi \left( X^{0,x;\bar{u},\bar{\xi}} (T) \right) - \Phi \left( X^{0,x;\bar{u},\bar{\xi}} (T) \right) + \xi (T) - \bar{\xi} (T),
\end{align*}
\]
where \( \hat{l} (r) = l (r, X^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r)) = l \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right) \), \( l = b, \sigma \). We can transform the above equation into
\[
\begin{align*}
\frac{d \Delta X^\alpha (r)}{dt} &= \Delta b (r) dt + \Delta \sigma (r) dW (r), \quad \Delta X^\alpha (t) = 0, \\
\frac{d \Delta Y^\alpha (r)}{dt} &= - \Delta f (r) dt + \Delta Z^\alpha (s) dW (s), \\
\Delta Y^\alpha (T) &= \Phi \left( X^{0,x;\bar{u},\bar{\xi}} (T) \right) - \Phi \left( X^{0,x;\bar{u},\bar{\xi}} (T) \right) + \xi (T) - \bar{\xi} (T),
\end{align*}
\]
where \( \Delta l (r, x, y, z) = A^l (r) x + B^l (r) y + C^l (r) z, l = b, \sigma, f \), separately. Set
\[
\begin{align*}
A (r) &= l \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right) - l \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right) \\
B (r) &= \left( f \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), Y^{0,x;\bar{u},\bar{\xi}} (r) - K \xi (r), Z^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right) \right) \\
C (r) &= \left( f \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), Y^{0,x;\bar{u},\bar{\xi}} (r) - K \xi (r), Z^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right) \right) \\
D (r) &= \left( f \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), Y^{0,x;\bar{u},\bar{\xi}} (r) - K \xi (r), Z^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right) \right) \\
\end{align*}
\]
From the continuity result and existence and uniqueness of FBSDEs, we have
\[
\lim_{\alpha \to 0} A^l (r) = l \left( r, X^{0,x;\bar{u},\bar{\xi}} (r), \bar{u} (r) \right), \quad \text{etc.,}
\]
from which we get the desired result. \( \square \)
Proof of Lemma 41} From (61) and (6) in Gozzi et al. [46], we have that if $(p,q,P) \in \mathcal{P}^{2,\tau} v(t,x)$, then

$$v(t+h,X^{t,x;\bar{u},\bar{\xi}}(t+h)) - v(t,X^{t,x;\bar{u},\bar{\xi}}(t)) \leq C \left( 1 + |X^{t,x;\bar{u},\bar{\xi}}(t)|^m \right) h$$

$$+ \left\langle q(t), X^{t,x;\bar{u},\bar{\xi}}(t+h) - X^{t,x;\bar{u},\bar{\xi}}(t) \right\rangle$$

$$+ C_0 \left| X^{t,x;\bar{u},\bar{\xi}}(t+h) - X^{t,x;\bar{u},\bar{\xi}}(t) \right|^2$$

$$= I_1 + I_2 + I_3$$

We shall deal with $I_1$, $I_2$, $I_3$, separately. For $I_1$, we have

$$\mathbb{E} \left( 1 + |X^{t,x;\bar{u},\bar{\xi}}(t+h)|^m \right) h \leq C \left( 1 + |x|^m \right) h,$$

by classical estimate and the assumption $\mathbb{E} \left[ |\xi_t|^2 \right] < \infty$. For $I_2$, from (7) in [46], we have, by Hölder inequality,

$$\mathbb{E} \left\langle q(t), X^{t,x;\bar{u},\bar{\xi}}(t+h) - X^{t,x;\bar{u},\bar{\xi}}(t) \right\rangle \leq C \left( \mathbb{E} \left[ (1 + |X^{t,x;\bar{u},\bar{\xi}}(t)|^{m_1})^2 \right] \right)^\frac{1}{2} \left\{ \left( \mathbb{E} \left[ \left| \int_t^{t+h} b \left( r, X^{0,x;\bar{u},\bar{\xi}}(r), \bar{u}(r) \right) \right|^2 \right] \right)^\frac{1}{2} \right\}$$

$$+ \left( \mathbb{E} \left[ \left| \int_t^{t+h} G(r) \, d\bar{\xi}(r) \right|^2 \right] \right)^\frac{1}{2},$$

since $\mathbb{E} \left[ |\xi_t|^2 \right] < \infty$ and the fact $(1 + |x|^2)^\frac{1}{2} \leq 1 + |x|$. Finally,

$$C_0 \mathbb{E} \left| X^{t,x;\bar{u},\bar{\xi}}(t+h) - X^{t,x;\bar{u},\bar{\xi}}(t) \right|^2 \leq C_0 \left\{ \left( \frac{t+h}{t} \right) ^\frac{1}{2} \left( \mathbb{E} \left[ \left| \int_t^{t+h} b \left( r, X^{0,x;\bar{u},\bar{\xi}}(r), \bar{u}(r) \right) \right|^2 \right] \right)^\frac{1}{2} \right\}$$

$$+ C_0 \left( \int_t^{t+h} \sigma \left( r, X^{0,x;\bar{u},\bar{\xi}}(r), \bar{u}(r) \right) \right)^2 \right\}$$

$$+ C_0 \left( \mathbb{E} \left[ \left| \int_t^{t+h} G(r) \, d\bar{\xi}(r) \right|^2 \right] \right)^\frac{1}{2} \right\}$$

$$\leq C \left( 1 + |x|^2 \right) (2h^2 + h),$$

By Itô isometry and classical estimate on SDE, we complete the proof. \hfill \Box

References

[1] L. Alvarez, A class of solvable singular stochastic control problems, *Stochastics Stochastics Rep.* 67 (1999) 83–122.
[2] L. Alvarez, Singular stochastic control, linear diffusions, and optimal stopping: A class of solvable problems, *SIAM J. Control Optim.* 39 (2001) 1697–1710.

[3] J. M. Bismut, “Théorie Probabiliste du Contrôle des Diffusions”, *Memoirs of the American Mathematical Society*, 176, Providence, Rhode Island, 1973.

[4] J. Bismut, An introductory approach to duality in optimal stochastic control, *SIAM Rev.*, 20 (1978), 62–78.

[5] D. J. Bell and D. H. Jacobson, Singular Optimal Control Problems, Math. Sci. Eng. 117, Academic Press, New York, 1975.

[6] V.E. Benêš, L.A. Shepp, H.S. Witzenshausen, Some solvable stochastic control problems, *Stochastics* 4 (1980) 39–83.

[7] A. Bensoussan, J. L. Lions, Applications des Inéquations Variationnelles en Contrôle Stochastique, Dunod, Paris, 1978.

[8] M. I. Borodowski, A.S. Bratus, F.L. Chernous’ko, Optimal impulse correction under random perturbation, *J. Appl. Math. Mech.* 39 (1975) 797–805.

[9] F. Baldursson, Singular stochastic control and optimal stopping, *Stochastics* 21 (1987) 1-40.

[10] F. Boetius, Bounded variation singular stochastic control and associated Dynkin game, in Mathematical Finance, Trends Math., M. Kohlmann and S. Tang, eds., Birkhäuser, Basel, (2001) 111–120.

[11] F. Boetius, Singular Stochastic Control and Its Relations to Dynkin Game and Entry-Exit Problems, Ph.D. thesis, University of Konstanz, 2001.

[12] F. Boetius and M. Kohlmann, Connections between optimal stopping and singular stochastic control, *Stochastic Process. Appl.*, 77 (1998) 253-281.

[13] J. Bather and H. Chernoff, Sequential decisions in the control of a spaceship, in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Physical Sciences, University of California Press, Berkeley, CA, (1967) 181–207.

[14] F. Baldursson and I. Karatzas, Irreversible investment and industry equilibrium, *Finance Stoch.* 1 (1997) 69–89.

[15] P. Briand, B. Delyon and Y. Hu, E. Pardoux, L. Stoica, $L_p$ solutions of backward stochastic differential equations. *Stochastic Process. Appl.* 108, (2003) 109–129.

[16] G. Barles, R. Buckdahn and E. Pardoux, Backward stochastic differential equations and integral-partial differential equations, *Stochastics Stochastics Rep.*, Volume 60, Issue 1-2.

[17] A. Bensoussan, J.L. Lions, Applications des Inéquations Variationnelles en Contrôle Stochastique, Dunod, Paris, 1978.

[18] S. Bahlali, B. Mezerdi B. Djehiche, The relaxed stochastic maximum principle in singular control of diffusions, *SIAM J. Control Optim.*, 46, 427–444, (2007).
[19] R. Buckdahn and J. Li, Stochastic differential games and viscosity solutions of Hamilton-Jacobi-Bellman-Isaacs equations, *SIAM J. Control Optim.* 47 (2008) 444-475.

[20] R. Buckdahn , B. Labed , C. Rainer, and L. Tamer, Existence of an optimal control for stochastic control systems with nonlinear cost functional, *An International Journal of Probability and Stochastic Processes* Vol. 82, (2010) - Issue 3

[21] R. Buckdahn and J. Ma, Stochastic viscosity solutions for nonlinear stochastic. Part I, *Stochastic Processes and their Applications*, 93 (2001), 181–204.

[22] R. Buckdahn and J. Ma, Pathwise stochastic taylor expansions and stochastic viscosity solutions for fully nonlinear stochastic pdes, *The Annals of Probability*, 30 (2002), 1131–1171.

[23] J. F. Bonnans and F. J. Silva, First and second order necessary conditions for stochastic optimal control problems, *Appl. Math. Optim.*, 65 (2012), 403–439.

[24] K. Bahlali, B. Gherbal and B. Mezerdi, Existence of optimal controls for systems driven by FBSDEs, *Systems Control Lett.* 60 (2011) 344-349.

[25] D. J. Clements and B. D. O. Anderson, Singular Optimal Control: The Linear-Quadratic Problem, Lecture Notes in Control and Inform. Sci. 5, Springer-Verlag, New York, 1978.

[26] A. Cadenillas, U. Haussmann, The stochastic maximum principle for a singular control problem, *Stochastics Stochastics Rep* Vol 49, (1994) 211-237.

[27] M. B. Chiarolla and U. G. Haussmann, Optimal control of inflation: Central bank problem, *SIAM J. Control Optim.* 36 (1998) 1099–1132.

[28] M. G. Crandall, H. Ishii, and P.-L. Lions, User’s guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (N.S.)* 27 (1992) 1–67.

[29] P. Chow, J.-L.Menaldi and M. Robin, Additive control of stochastic linear systems with finite horizon, *SIAM J. Control Optim.* 23, (1985) 858–899.

[30] Z. Chen and L Epstein, Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70, (2002), 1403–1443.

[31] M. H. A. Davis, A.R. Norman, Portfolio Selection with Transaction Costs, *Mathematics of Operations Research* Vol. 15, No. 4, (1990) 676-713.

[32] F. Dufour, B. Miller, Singular stochastic control problems, *SIAM J. Control Optim.* Vol. 43, No. 2, (2004) 708–730.

[33] F. Dufour, B. Miller, Maximum principle for singular stochastic control problems, *SIAM J. Control Optim.* Vol. 45, No. 2, (2006) 668–698.

[34] K. R. Dahl & B. Øksendal, Singular recursive utility, *Stochastics* VOL. 89, NOS. 6–7, (2017) 994–1014.

[35] D. Duffie, L. Epstein, Stochastic differential utility, *Econometrica* 60 (1992) 353–394.
[36] Y. Dong, Q. Meng, Second-Order Necessary Conditions for Optimal Control with Recursive Utilities, [arXiv:1802.09242](https://arxiv.org/abs/1802.09242).

[37] N. El Karoui and I. Karatzas, Probabilistic aspects of finite-fuel, reflected follower problems, *Acta Appl. Math.* 11 (1988) 223–258.

[38] N. El Karoui and I. Karatzas, A new approach to the skorohod problem, and its applications, *Stochastics Stochastics Rep.* 34 (1991) 57–82.

[39] W. Fleming and H. Soner, Controlled Markov Processes and Viscosity Solutions, Springer-Verlag, New York, 1993.

[40] H. Frankowska and D. Tonon, Pointwise second-order necessary optimality conditions for the Mayer problem with control constraints, *SIAM, J. Control Optim.*, 51 (2013), 3814–3843.

[41] W. H. Fleming, R. W. Rishel , Deterministic and Stochastic Optimal Control, Springer-Verlag, New York, 1975.

[42] G. Fu, U. Horst, Mean Field Games with Singular Controls, *SIAM J. Control Optim.* Vol. 55, No. 6, (2017) 3833-3868.

[43] H. Frankowska, H. Zhang, and X. Zhang, First and second order necessary conditions for stochastic optimal controls, *J. Differential Equations*, 262 (2017), 3689–3736.

[44] R. Gabasov and F. M. Kirillova, High order necessary conditions for optimality, *SIAM, J. Control Optim.*, 10 (1972), 127–168.

[45] B. S. Goh, Necessary conditions for singular extremals involving multiple control variables, *SIAM, J. Control Optim.*, 4 (1966), 716–731.

[46] F. Gozzi, A. Swiech, X. Zhou, Erratum: “a corrected proof of the stochastic verification theorem within the framework of viscosity solutions”, *SIAM J. Control Optim.* 48 (6) (2010) 4177–4179.

[47] S. Hamadène and J. P. Lepeltier, Reflected backward stochastic differential equations and mixed game problem, *Stochastic Process. Appl.* 85 (2000) 177–188.

[48] U. Haussmann, W. Suo, Singular optimal stochastic controls I: Existence, *SIAM J. Control Optim.* Vol. 33, No. 3, (1995) 916-936.

[49] U. Haussmann, W. Suo, Singular optimal stochastic controls II: Dynamic programming, *SIAM J. Control Optim.* Vol. 33, No. 3, (1995) 937-959.

[50] M. Hu, Stochastic global maximum principle for optimization with recursive utilities, *Probability, Uncertainty and Quantitative Risk* (2017) 2:1.

[51] M. Hu, S. Ji, X. Xue, A Global Stochastic Maximum Principle for Fully Coupled Forward-Backward Stochastic Systems, *SIAM J. Control Optim.* Vol. 56, No. 6, (2018) 4309–4335.

[52] ——, A note on the global stochastic maximum principle for fully coupled forward-backward stochastic systems, [arXiv:1812.10469](https://arxiv.org/abs/1812.10469).
[53] H. Ishii, Uniqueness of unbounded viscosity solutions of Hamilton-Jacobi equations, *Indiana U. Math. J.*, 26, (1984) 721-748.

[54] S. Ji and X.Y. Zhou, A maximum principle for stochastic optimal control with terminal state constrains, and its applications. *Comm. Inf. Syst.*, 6, (2006), 321–337.

[55] S. Ji, X. Xue, The stochastic maximum principle in singular optimal control with recursive utilities, *J. Math.Anal.Appl.* 471(2019)378–391.

[56] N. El Karoui, S. Peng, and M. C. Quenez, Backward stochastic differential equations in finance, *Math. Finance* 7 (1997) 1–71.

[57] N. Krylov, Controlled Diffusion Processes, Springer-Verlag, New York, (1980).

[58] H.-W. Knobloch, Higher Order Necessary Conditions in Optimal Control Theory, Lecture Notes in Comput. Sci. 34, Springer-Verlag, New York, 1981.

[59] A. J. Krener, The high order maximal principle and its application to singular extremals, *SIAM J. Control Optim.*, 15 (1977), 256–293.

[60] H. J. Kushner, Necessary conditions for continuous parameter stochastic optimization problems, *SIAM J. Control Optim.*, 10 (1972), 550–565.

[61] I. Karatzas, The monotone follower problem in stochastic decision theory, *Appl. Math. Optim.* 7 (1981) 175–189.

[62] I. Karatzas, A class of singular control problems, *Adv. in Appl. Probab.* 15 (1983) 225–254.

[63] I. Karatzas and S. Shreve, Connections between optimal stopping and singular stochastic control I. Monotone follower problems, *SIAM J. Control Optim.* 22 (1984) 856–877.

[64] I. Karatzas and S. Shreve, Connections between optimal stopping and singular stochastic control II. Reflected follower problems, *SIAM J. Control Optim.* 23 (1985) 433–451.

[65] I. Karatzas and S. Shreve, Equivalent models for finite-fuel stochastic control, *Stochastics* 18 (1986) 245–276.

[66] T. Kobila, A class of solvable stochastic investment problems involving singular controls, *Stochastics Stochastics Rep* 43 (1993) 29–63.

[67] I. Karatzas and H. Wang, A barrier option of American type, *Appl. Math. Optim.* 42 (2000) 259–279.

[68] M. Kohlmann, X.Y. Zhou, Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach. *SIAM J. Control Optim.* 38, (2000), 1392–1407.

[69] J. Lehoczky and S. Shreve, Absolutely Continuous and Singular Stochastic Control. *Stochastics* 17, (1986) 91-109.
[70] P. L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 1: The dynamic programming principle and applications, and Part 2: Viscosity solutions and uniqueness, *Comm. Partial Differential Equations* 8, 1101-1174. (1983) 1229-1276.

[71] J. Li and Q. Wei, Optimal control problems of fully coupled FBSDEs and viscosity solutions of Hamilton-Jacobi-Bellman equations *SIAM J. Control Optim.* 52 (3), (2014) 1622-1662.

[72] A. Lim and X.Y. Zhou, Linear-quadratic control of backward stochastic differential equations. *SIAM J. Control Optim.* 40, (2001) 450–474.

[73] J. Ma, On the principle of smooth fit for a class of singular stochastic control problems for diffusion, *SIAM J. Control Optim.* 30 (1992) 975–999.

[74] J.-L. Menaldi and M. Taksar, Optimal correction problem of a multidimensional stochastic system, *Automatica* 25 (1989) 223–232.

[75] L. F. Martins and H. J. Kushner, Routing and singular control for queueing networks in heavy traffic, *SIAM J. Control Optim.* 28 (1990) 1209–1233.

[76] D. Nualart, The Malliavin Calculus and Related Topics, Probability and Its Applications. New York and Berlin: Springer-Verlag (1995).

[77] D. Nualart and E. Pardoux, Stochastic Calculus with Anticipating Integrands, *Probab. Theory Rel. Fields*, 78, 535–581 (1988).

[78] T. Nie, J. Shi, Z. Wu, Connection between MP and DPP for Stochastic Recursive Optimal Control Problems: Viscosity Solution Framework in the General Case, *SIAM J. CONTROL OPTIM.* Vol. 55, No. 5, (2017) 3258–3294.

[79] D. Ocone, Malliavin’s Calculus and Stochastic Integral Representation of Functionals of Diffusion Processes, *Stochastics*, 12, (1988), 161–185.

[80] B. Øsendal, A. Sulem, T. Zhang, Singular Control and Optimal Stopping of SPDEs, and Backward SPDEs with Reflection, *Mathematics of Operations Research*, Volume 39, Issue 2, (2014) 464-486.

[81] S. Peng, Stochastic Hamilton-Jacobi-Bellman equations, *SIAM J. Control Optim.*, 30 (1992), 284–304.

[82] S. Peng, A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.* 28, (1990) 966–979.

[83] S. Peng, Backward stochastic differential equations and applications to optimal control. *Appl. Math. Optim.* 27, (1993) 125–144.

[84] E. Pardoux, S. Peng, Adapted solution of a backward stochastic di-erential equation, *Systems Control Lett.* 14 (1990) 55–61.

[85] S. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, *Stochastics Stochastics Rep.* 37 (1991) 61–74.
[86] S. Peng, A generalized dynamic programming principle and HJB equation, *Stochastics Stochastics Rep.* 38 (1992) 119–134.

[87] S. Peng, BSDE and stochastic optimizations, in: J. Yan, S. Peng, S. Fang, L. Wu, Topics in Stochastic Analysis, Science Press, Beijing, 1997 (Chapter 2) (in Chinese).

[88] H. Pagès, D. Possamaï, A mathematical treatment of bank monitoring incentives, *Finance and Stochastics* 18 (1), (2014) 39-73.

[89] J. Qiu, Viscosity Solutions of Stochastic Hamilton–Jacobi–Bellman Equations, *SIAM J. Control Optim.* Vol. 56, No. 5, (2018) 3708–3730.

[90] H. M. Soner, S. E. Shreve, Regularity of the value function for a two-dimensional singular stochastic control problem, *SIAM J. Control Optim.* 27 (1989) 876–907.

[91] S. E. Shreve, J. P. Lehoczky, and D. P. Gaver, Optimal consumption for general diffusions with absorbing and reflecting barriers, *SIAM J. Control Optim.* 22 (1984) 55–75.

[92] J. Shi and Z. Wu, The maximum principle for fully coupled forward-backward stochastic control system. *Acta Automat. Sinica* 32, (2006) 161–169.

[93] S. Tang, A second-order maximum principle for singular optimal stochastic controls, *Discrete Contin. Dyn. Syst. Ser. B.* 14 (2010), 1581–1599.

[94] S. A. Williams, P. L. Chow, J. L. Menaldi, Regularity of the free boundary in singular stochastic control, *J. Differential Equations* 111 (1994) 175–201.

[95] B. Wang, Singular control of stochastic linear systems with recursive utility, *Systems Control Lett.* 51, (2004) 105-122.

[96] Z. Wu, Z. Yu, Dynamic programming principle for one kind of stocastic recursive optimal control problem and Hamilton-Jacobi-Bellman equation, *SIAM J. Control Optim.* Vol. 47, No. 5, (2008) 2616–2641.

[97] Z. Wu, Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems. *Syst. Sci. Math. Sci.* 11, (1998) 249–259.

[98] Z. Wu, A general maximum principle for optimal control of forward-backward stochastic systems. *Automatica* 49, (2013) 1473–1480.

[99] W. Xu, Stochastic maximum principle for optimal control problem of forward and backward system. *J. Austral. Math. Soc.* Ser. B 37, (1995) 172–185.

[100] J. Yong, X. Y. Zhou, Stochastic Controls. Hamiltonian Systems and HJB Equations, Springer-Verlag, New York, 1999.

[101] J. Yong, Optimality variational principle for controlled forward-backward stochastic differential equations with mixed initial-terminal conditions. *SIAM J. Control Optim.* 48, (2010) 4119–4156.
[102] H. Zhang, X. Zhang, Pointwise Second-order Necessary Conditions for Stochastic Optimal Controls, Part I: The Case of Convex Control Constraint, *SIAM J. CONTROL OPTIM*. Vol. 53, No. 4, (2015) 2267–2296.

[103] ——, Pointwise Second-Order Necessary Conditions for Stochastic Optimal Controls, Part II: The General Case, *SIAM J. CONTROL OPTIM*. 55:5, (2017) 2841-2875.

[104] ——, H. Zhang and X. Zhang, Some results on pointwise second-order necessary conditions for stochastic optimal controls, Sci. China Math., 59 (2016), 227–238.

[105] ——, Second-Order Necessary Conditions for Stochastic Optimal Control Problems, *SIAM Review* 60(1): (2018) 139-178.

[106] H. Zhu, Generalized solution in singular control: The nondegenerate problem, *Appl. Math. Optim*. 25 (1992) 225–245.

[107] L. Zhang, Stochastic verification theorem of forward–backward controlled systems for viscosity solutions, Systems Control Lett. 61 (2012) 649–654.

[108] L. Zhang, Q. Zhou, Near-optimal control of stochastic recursive systems via viscosity solution, *J Optim Theory Appl*. 178 (2), (2018) 363-382.

[109] L. Zhang, Singular optimal controls of stochastic recursive systems and Hamilton–Jacobi–Bellman inequality, to appear in *J. Differential Equations*, 2018.