ARTICLE IN PRESS YJDEQ:8637

Available online at www.sciencedirect.com

ScienceDirect

J. Differential Equations ••• (••••) •••–•••

Journal of Differential Equations

www.elsevier.com/locate/jde

First and second order necessary conditions for stochastic optimal controls

Hélène Frankowska [a,1], Haisen Zhang [b,*,2], Xu Zhang [c,*,3]

[a] CNRS, IMJ-PRG, UMR 7586, Sorbonne Universités, UPMC Univ. Paris 06, Univ. Paris Diderot, case 247, 4 place Jussieu, 75252 Paris, France

[b] School of Mathematics and Statistics, Southwest University, Chongqing 400715, China

[c] School of Mathematics, Sichuan University, Chengdu 610064, China

Received 13 April 2016; revised 1 November 2016

Abstract

The main purpose of this paper is to establish the first and second order necessary optimality conditions for stochastic optimal controls using the classical variational analysis approach. The control system is governed by a stochastic differential equation, in which both drift and diffusion terms may contain the control variable and the set of controls is allowed to be nonconvex. Only one adjoint equation is introduced to derive the first order necessary condition; while only two adjoint equations are needed to state the second order necessary conditions for stochastic optimal controls.

MSC: primary 93E20; secondary 49J53, 60H07, 60H10

* Corresponding authors.

E-mail addresses: helene.frankowska@imj-prg.fr (H. Frankowska), haisenzhang@yeah.net (H. Zhang), zhang_xu@scu.edu.cn (X. Zhang).

[1] The research of this author is partially supported by the Gaspard Monge Program for Optimisation and Operational Research, Jacques Hadamard Mathematical Foundation (FMJH).

[2] The research of this author is partially supported by NSFC under grants 11401404 and 11471231, the Fundamental Research Funds for the Central Universities under grants SWU114074 and XDJK2015C142.

[3] The research of this author is partially supported by NSFC under grant 11231007, the PCSIRT under grant IRT_15R53, the Chang Jiang Scholars Program from the Chinese Education Ministry. Also, he thanks the Fondation Sciences Mathématiques de Paris for the support and hospitality.

http://dx.doi.org/10.1016/j.jde.2016.11.041

0022-0396/

Keywords: Stochastic optimal control; Malliavin calculus; Necessary conditions; Adjacent cone; Variational equation; Adjoint equation

1. Introduction

Let $T > 0$ and $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space (satisfying the usual conditions), on which a 1-dimensional standard Wiener process $W(\cdot)$ is defined such that $\mathbb{F} = \{\mathcal{F}_t\}_{0\le t\le T}$ is the natural filtration generated by $W(\cdot)$ (augmented by all the P-null sets).

Let us consider the following controlled stochastic differential equation

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \quad t \in [0, T], \\ x(0) = x_0 \in K, \end{cases} \tag{1.1}$$

with the cost functional

$$J(u(\cdot), x_0) = \mathbb{E}\Big[\int_0^T f(t, x(t), u(t))dt + g(x(T))\Big]. \tag{1.2}$$

Here $u(\cdot)$ is the control variable with values in a closed nonempty subset U of $\mathbb{R}^m$ (for some fixed $m \in \mathbb{N}$), $x(\cdot)$ is the state variable with values in $\mathbb{R}^n$ (for some given $n \in \mathbb{N}$), K is a closed nonempty subset in $\mathbb{R}^n$, and $b, \sigma : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}$ and $g : \mathbb{R}^n \times \Omega \to \mathbb{R}$ are given functions (satisfying suitable conditions to be stated later). As usual, when the context is clear, we omit the ω $(\in \Omega)$ argument in the defined functions.

Denote by $\langle\cdot,\cdot\rangle$ and $|\cdot|$ respectively the inner product and norm in $\mathbb{R}^n$ or $\mathbb{R}^m$, which can be identified from the contexts, by $\mathcal{B}(X)$ the Borel σ-field of a metric space X, and by $\mathcal{U}_{ad}$ the set of $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted stochastic processes with values in U such that $\mathbb{E}\int_0^T |u(t, \omega)|^2 dt < \infty$. Any $u(\cdot) \in \mathcal{U}_{ad}$ is called an admissible control, the corresponding state $x(\cdot; x_0)$ of (1.1) with initial datum $x_0 \in K$ is called an admissible state, and (x, u, x_0) is called an admissible triple. An admissible triple $(\bar{x}, \bar{u}, \bar{x}_0)$ is called optimal if

$$J(\bar{u}(\cdot), \bar{x}_0) = \inf_{\substack{u(\cdot)\in\mathcal{U}_{ad} \\ x_0\in K}} J(u(\cdot), x_0). \tag{1.3}$$

The purpose of this paper is to establish first and second order necessary optimality conditions for problem (1.3). We refer to [4,5,16,21] and references cited therein for some early works on this subject. Although the stochastic optimal control theory was developing almost simultaneously with the deterministic one, its results are much less fruitful than those obtained for the deterministic control systems. The main reasons are due to some essential difficulties (or new phenomena) when the diffusion term of the stochastic control system depends on the control variable and the control region lacks convexity. In contrast with the deterministic case, for stochastic optimal control problems when spike variations are used as perturbations, the cost functional needs to be expanded up to the *second order* and *two adjoint equations* have to be introduced to derive the *first order necessary optimality conditions*. A stochastic maximum principle for this general case was established in [27]. On the other hand, to derive the second order necessary optimality conditions, the cost functional needs to be expanded up to the forth order and four adjoint equations have to be introduced, see [34]. Consequently, these necessary conditions narrow the field of applications, since they require so many adjoint equations and considerably strong smoothness assumptions (with respect to the state variable x) on the coefficients of the control system and the cost functional.

Can we use just one adjoint equation (resp. two adjoint equations) to derive a first (resp. second) order necessary condition for the above general stochastic optimal control problem? To answer this question, let us first turn back to the special case of convex control constraint. When the control region is convex, the usual convex variation can be used to construct a control perturbation. Only one adjoint equation is needed to establish the first order necessary condition (see [4]) and two adjoint equations are needed to establish the second order necessary condition (see [33]) for stochastic optimal controls. The main advantage of using the convex variations instead of the spike ones, is the fact that, it avoids efficiently the difficulties brought by perturbations with respect to the measure. However, when the control region is nonconvex, the traditional convex variations cannot be used, since there may exist a control $u(\cdot)$ in the set of admissible controls $\mathcal{U}_{ad}$ such that $v := u - \bar{u}$ is not an admissible direction to construct a control perturbation (of the optimal control $\bar{u}$). Nevertheless, if the perturbation direction v is chosen so that for any $\varepsilon > 0$ one can find a v^ε converging to v (in a suitable sense) when $\varepsilon \to 0^+$ and satisfying $\bar{u} + \varepsilon v^\varepsilon \in \mathcal{U}_{ad}$, then the variational approach can be adopted to deal with some optimal control problems having nonconvex control regions (we call it the *classical variational analysis approach*). Indeed, this method has been used extensively in optimization and optimal control theory in the deterministic setting. Using this method, in [17,11], some second order integral type necessary conditions for deterministic optimal controls were established. It was shown in [10,12] that these necessary conditions imply pointwise ones.

In this paper, we shall use the classical variational analysis approach to establish the first and second order necessary optimality conditions for stochastic optimal controls in the general setting, that is, when the control region is allowed to be nonconvex and the control variable enters also into the diffusion term of the control system. Let us recall that, when the diffusion term does NOT depend on the control variable, cf. [1,23,30], the situation is more or less similar to the deterministic setting like the one in [11,22]. Compared to the existing results for the case of general control constraints obtained by the spike variations [27,34], the main advantage of the classical variational analysis approach is due to weaker smoothness requirements imposed on the coefficients of the control system and the cost functional (with respect to the state variable x) and to fewer adjoint equations needed to state these conditions. Previously the first and second order integral type necessary conditions for stochastic optimal controls with convex control constraints were derived in [6] using the convex (first order) variations of optimal control. In the difference with [6], our variational approach is also valid when the control region is nonconvex and, since the second order variations of the control region are used in this paper, the corresponding second order necessary condition is more effective than the one of [6] even in the case of convex control constraints (see Example 4.1 below).

In a sense, our work can be viewed as a refinement of known optimality conditions for stochastic control problems. To see it, let us return, for a moment, to the deterministic optimal control problem, i.e., when the functions $\sigma(\cdot) \equiv 0$, $b(\cdot)$, $f(\cdot)$, $g(\cdot)$, $x(\cdot)$ and $u(\cdot)$ in (1.1)–(1.2) are independent from the sample point ω, and also, for the sake of simplicity, let $K = \{x_0\}$ for some fixed $x_0 \in \mathbb{R}^n$. Consider an optimal pair $(\bar{x}, \bar{u})$ and the solution $\psi(\cdot)$ to the following ordinary differential equation,

$$\begin{cases} \dot{\psi}(t) = -b_x(t,\bar{x}(t),\bar{u}(t))^\top \psi(t) + f_x(t,\bar{x}(t),\bar{u}(t)), \quad t \in [0,T], \\ \psi(T) = -\nabla g(\bar{x}(T)). \end{cases} \tag{1.4}$$

Define the (deterministic) Hamiltonian

$$H(t,x,u,\psi) := \langle \psi, b(t,x,u)\rangle - f(t,x,u), \qquad \forall\,(t,x,u,\psi) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n.$$

Then the following Pontryagin maximum principle [28] holds

$$H(t,\bar{x}(t),\bar{u}(t),\psi(t)) = \max_{v\in U} H(t,\bar{x}(t),v,\psi(t)), \ \text{ a.e. } \ t\in[0,T]. \tag{1.5}$$

Clearly, when U is a finite set, condition (1.5) provides an effective way to compute "$\bar{u}(\cdot)$"; while when U is convex, condition (1.5) yields

$$\big\langle H_u(t,\bar{x}(t),\bar{u}(t),\psi(t)), v-\bar{u}(t)\big\rangle \le 0, \qquad \forall\, v\in U, \text{ a.e. } t\in[0,T]. \tag{1.6}$$

What about other types of U? Are there other necessary conditions for optimal pairs? The classical monograph [28] was followed by numerous works addressing the above issues and refinements of known results on optimal control problems in the deterministic finite dimensional setting. In this respect, we refer to [3,7,10,14,15,17,19,20,26] for high order necessary conditions when the first-order necessary conditions turn out to be trivial and to [26] for a discussion on "bang-bang" controls which are very useful in applications. A very natural question concerns the stochastic counterpart of the above results. Surprisingly, very little is known about high order conditions in the stochastic framework! Indeed, as an interesting comparison, we mention that, there exists at least five research monographs [3,7,14,19,26] devoted to deterministic high order necessary conditions but one can find only a very few published articles [1,6,23,30,33] for their stochastic analogues.

The outline of the paper is as follows. In Section 2, we collect some notations and introduce some spaces and preliminary results that will be used later. In Section 3, we derive the first order necessary conditions for stochastic optimal controls. Section 4 is devoted to establishing second order necessary conditions. Finally, in the Appendix, we give the proofs of two technical results from Sections 3 and 4.

Some of preliminary results of this paper are announced (without proofs) in [13].

2. Preliminaries

This section is of preliminary nature, in which we shall introduce some useful notations and spaces, and recall some concepts and results from the set-valued analysis and the Malliavin calculus.

2.1. Notations and spaces

In this subsection, we introduce some notations and spaces which will be used in the sequel.

Denote by $C_b^\infty(\mathbb{R}^n;\mathbb{R}^m)$ the set of C^∞-smooth functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ with bounded partial derivatives. Let $\mathbb{R}^{n\times m}$ be the space of all $n\times m$-real matrices. For any $A\in\mathbb{R}^{n\times m}$, denote by $A^\top$ its transpose and by $|A| = \sqrt{tr\{AA^\top\}}$ the norm of A. Also, write $\mathbf{S}^n := \big\{A\in\mathbb{R}^{n\times n}\,\big|\, A^\top = A\big\}$.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Let $\varphi : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^d$ ($d \in \mathbb{N}$) be a given function. For a.e. $(t, \omega) \in [0, T] \times \Omega$, we denote by $\varphi_x(t, x, u, \omega)$ and $\varphi_u(t, x, u, \omega)$ respectively the first order partial derivatives of φ with respect to x and u at (t, x, u, ω), by $\varphi_{(x,u)^2}(t, x, u, \omega)$ the Hessian of φ with respect to (x, u) at (t, x, u, ω), and by $\varphi_{xx}(t, x, u, \omega)$, $\varphi_{xu}(t, x, u, \omega)$ and $\varphi_{uu}(t, x, u, \omega)$ respectively the second order partial derivatives of φ with respect to x and u at (t, x, u, ω).

For any $\alpha, \beta \in [1, +\infty)$ and $t \in [0, T]$, we denote by $L^{\beta}_{\mathcal{F}_t}(\Omega; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued, $\mathcal{F}_t$ measurable random variables ξ such that $\mathbb{E}\,|\xi|^{\beta} < +\infty$; by $L^{\beta}([0, T] \times \Omega; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable processes φ such that $\|\varphi\|_{\beta} := \big[\mathbb{E} \int_0^T |\varphi(t, \omega)|^{\beta} dt\big]^{\frac{1}{\beta}} < +\infty$; by $L^{\beta}_{\mathbb{F}}(\Omega; L^{\alpha}(0, T; \mathbb{R}^n))$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable, $\mathbb{F}$-adapted processes φ such that $\|\varphi\|_{\alpha,\beta} := \big[\mathbb{E}\,\big(\int_0^T |\varphi(t, \omega)|^{\alpha} dt\big)^{\frac{\beta}{\alpha}}\big]^{\frac{1}{\beta}} < +\infty$; by $L^{\beta}_{\mathbb{F}}(\Omega; C([0, T]; \mathbb{R}^n))$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted continuous processes φ such that $\|\varphi\|_{\infty,\beta} := \big[\mathbb{E}\,\big(\sup_{t \in [0,T]} |\varphi(t, \omega)|^{\beta}\big)\big]^{\frac{1}{\beta}} < +\infty$; by $L^{\infty}([0, T] \times \Omega; \mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable processes φ such that $\|\varphi\|_{\infty} := \operatorname{ess\,sup}_{(t,\omega) \in [0,T] \times \Omega} |\varphi(t, \omega)| < +\infty$ and by $L^{\beta}(0, T; L^{\beta}_{\mathbb{F}}([0, T] \times \Omega; \mathbb{R}^n))$ the $\mathbb{R}^n$-valued, $\mathcal{B}([0, T]) \otimes \mathcal{B}([0, T]) \otimes \mathcal{F}$ measurable functions φ such that $\|\varphi\|_{\beta} := \big[\mathbb{E} \int_0^T \int_0^T |\varphi(s, t, \omega)|^{\beta} ds dt\big]^{\frac{1}{\beta}} < +\infty$ and for any $t \in [0, T]$, the process $\varphi(\cdot, t, \cdot)$ is $\mathbb{F}$-adapted.

Let us recall that on a given filtered probability space, any $\mathbb{F}$-progressively measurable process is $\mathcal{B}([0; T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted, and every $\mathcal{B}([0; T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted process has an $\mathbb{F}$-progressively measurable modification (see [32, Proposition 2.8]).

2.2. *Some concepts and results from the set-valued analysis*

In this subsection, we recall some concepts and results from the set-valued analysis. We refer the reader to [2] for more details.

Let X be a Banach space with norm $\|\cdot\|_X$, and denote by X^* the dual space of X. For any subset $K \subset X$, denote by ∂K, $int K$ and $cl K$ its boundary, interior and closure, respectively. K is called a cone if $\alpha x \in K$ for any $\alpha \geq 0$ and $x \in K$. Define the distance between a point $x \in X$ and K by $dist\,(x, K) := \inf\limits_{y \in K} \|y - x\|_X$. Define the metric projection of x onto K by $\Pi_K(x) := \{y \in K \mid \|y - x\|_X = dist\,(x, K)\}$.

Definition 2.1. For $x \in K$, the adjacent cone $T^b_K(x)$ to K at x is defined by

$$T^b_K(x) := \Big\{v \in X \;\Big|\; \lim_{\varepsilon \to 0^+} \frac{dist\,(x + \varepsilon v, K)}{\varepsilon} = 0\Big\}.$$

If in the above $\lim_{\varepsilon \to 0^+}$ is replaced by $\liminf_{\varepsilon \to 0^+}$, then we obtain a larger cone, the so called *contingent cone* $T^B_K(x)$ to K at x. When K is convex, the adjacent cone and the contingent cone coincide with each other, and

$$T^b_K(x) = cl\Big\{\alpha(y - x) \;\Big|\; \alpha \geq 0,\ y \in K\Big\}.$$

It is not difficult to realize that $v \in T^b_K(x)$ if and only if for any $\varepsilon > 0$ there exists a $v_\varepsilon \in X$ such that $v_\varepsilon \to v$ (in X) as $\varepsilon \to 0^+$, and $x + \varepsilon v_\varepsilon \in K$.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Definition 2.2. For any $x \in K$ and $v \in T_K^b(x)$, the second order adjacent subset to K at (x, v) is defined by

$$T_K^{b(2)}(x, v) := \left\{ h \in X \;\middle|\; \lim_{\varepsilon \to 0^+} \frac{dist\,(x + \varepsilon v + \varepsilon^2 h, K)}{\varepsilon^2} = 0 \right\}.$$

Similarly to the above, $h \in T_K^{b(2)}(x, v)$ if and only if for any $\varepsilon > 0$ there exists an $h_\varepsilon \in X$ such that $h_\varepsilon \to h$ (in X) as $\varepsilon \to 0^+$ and $x + \varepsilon v + \varepsilon^2 h_\varepsilon \in K$.

Remark 2.1. Clearly, $0 \in T_K^b(x)$ for any $x \in K$ and $\alpha v \in T_K^b(x)$ for any $\alpha > 0$ and $v \in T_K^b(x)$. Therefore, $T_K^b(x)$ is a nonempty closed cone. $T_K^b(x) = X$ for any $x \in int K$. Also, $T_K^{b(2)}(x, 0) = T_K^b(x)$. When K is convex, $y - x \in T_K^b(x)$ and $0 \in T_K^{b(2)}(x, y - x)$ for any $x \in K$ and $y \in K$. When $v \neq 0$, the set $T_K^{b(2)}(x, v)$, in general, may not be a cone and it may be an empty set (some examples can be found in [2, section 4.7]).

The dual cone of the tangent cone $T_K^b(x)$, denoted by $N_K^b(x)$, is called the normal cone of K at x, i.e.,

$$N_K^b(x) := \left\{ \xi \in X^* \;\middle|\; \langle \xi, v \rangle \le 0, \ \forall\, v \in T_K^b(x) \right\}.$$

When K is convex, $N_K^b(x)$ reduces to the normal cone $N_K(x)$ of the convex analysis, where

$$N_K(x) := \left\{ \xi \in X^* \;\middle|\; \langle \xi, y - x \rangle \le 0, \ \forall\, y \in K \right\}.$$

When X is a Hilbert space, for any $\xi \in N_K^b(x)$ the second order normal cone to K at (x, ξ) is defined by

$$N_K^{b(2)}(x, \xi) := \left\{ \zeta \in \mathbf{S}(X) \;\middle|\; \langle \xi, h \rangle + \frac{1}{2} \langle \zeta v, v \rangle \le 0, \ \forall\, v \in T_K^b(x) \cap \{\xi\}^\perp, \ \forall\, h \in T_K^{b(2)}(x, v) \right\},$$

where $\mathbf{S}(X)$ is the space of symmetric, continuous linear operators from X to X and $\{\xi\}^\perp := \{v \in X \mid \langle \xi, v \rangle = 0\}$.

In the following, we recall a classical example in which the closed set K is defined by finitely many equalities and inequalities.

Example 2.1. When $K \subset \mathbb{R}^n$ is given by inequality and equality constraints and a constraint qualification holds true, there are exact expressions for the first and second order tangent sets. More precisely, consider twice continuously differentiable functions $\varphi_1, ..., \varphi_p : \mathbb{R}^n \to \mathbb{R}$ and $\psi_1, \ldots, \psi_r : \mathbb{R}^n \to \mathbb{R}$ (for some $p, r \in \mathbb{N}$), set $\varphi = (\varphi_1, ..., \varphi_p)$ and define

$$K = \left\{ x \in \mathbb{R}^n \;\middle|\; \varphi(x) = 0, \ \ \psi_j(x) \le 0, \ \forall\, j = 1, ..., r \right\}.$$

If there are no equality, resp. inequality, constraints in the definition of K, then the terms involving φ, φ_i, resp. ψ_j, are absent in the discussion below and p, resp. r, is equal to zero.

Let $x \in K$ and denote by $I(x)$ the set of all active indices, i.e. $j \in I(x)$ if and only if $\psi_j(x) = 0$. We assume that the Mangasarian–Fromowitz constraint qualification holds true: the Jacobian $\varphi'(x)$ is surjective and there exists a $v_0 \in \mathbb{R}^n$ such that

$$\varphi'(x)v_0 = 0, \ \ \langle \nabla\psi_j(x), v_0\rangle < 0, \quad \forall\, j \in I(x).$$

In the absence of equality constraints this is equivalent to the assumption that $\{\nabla\psi_j(x) \mid j \in I(x)\}$ are positively independent or, equivalently, $0 \notin \mathrm{co}\{\nabla\psi_j(x) \mid j \in I(x)\}$. Then it is well known, see for instance [2, pp. 150–151] that

$$T_K^b(x) = \left\{ v \in \mathbb{R}^n \;\middle|\; \varphi'(x)v = 0,\ \langle \nabla\psi_j(x), v\rangle \le 0,\ \ \forall\, j \in I(x) \right\},$$

$$N_K^b(x) = \sum_{i=1}^{p} \mathbb{R}\nabla\varphi_i(x) + \sum_{j\in I(x)} \mathbb{R}_+\nabla\psi_j(x).$$

If there are no equality constraints and $I(x) = \emptyset$, then $T_K^b(x) = \mathbb{R}^n$ and therefore $N_K^b(x) = \{0\}$.

Fix any $v \in T_K^b(x)$ and consider the set $I_v(x) = \{j \in I(x) \mid \langle \nabla\psi_j(x), v\rangle = 0\}$. Then the same proof as in [2, p. 177] (given there only for the second order contingent set) implies that

$$T_K^{b(2)}(x,v) = \Big\{ h \in \mathbb{R}^n \;\Big|\; \langle \nabla\varphi_i(x), h\rangle + \frac{1}{2}\langle \varphi_i''(x)v, v\rangle = 0,\ \forall\, i = 1, \cdots, p$$
$$\text{and}\ \ \langle \nabla\psi_j(x), h\rangle + \frac{1}{2}\langle \psi_j''(x)v, v\rangle \le 0,\ \forall\, j \in I_v(x) \Big\}.$$

Thus, under our assumptions, $T_K^{b(2)}(x,v) \neq \emptyset$ for all $v \in T_K^b(x)$.

Observe that $N_K^{b(2)}(x,0)$ is equal to the set of all symmetric $(n \times n)$-matrices that are semi-negative on $T_K^b(x)$.

If $I(x) \neq \emptyset$, denote by $i_1, \ldots, i_k$ all the active indices (for some $k \le r$). In the expressions below the terms involving φ_i, resp. ψ_{i_j}, are absent when there are no equality constraints, resp. when $I(x) = \emptyset$.

Fix any $0 \neq q \in N_K^b(x)$. Then for some reals $\{\mu_i\}_{i=1}^p$, $\lambda_j \ge 0$, $j = 1, \ldots, k$

$$q = \sum_{i=1}^{p} \mu_i \nabla\varphi_i(x) + \sum_{j=1}^{k} \lambda_j \nabla\psi_{i_j}(x).$$

To express $N_K^{b(2)}(x,q)$ we could apply the same method as in [11]. In order to simplify the discussion, we assume that $\{\nabla\varphi_1(x), \cdots, \nabla\varphi_p(x)\} \bigcup \{\nabla\psi_j(x) \mid j \in I_v(x)\}$ are linearly independent for every $v \in T_K^b(x) \cap \{q\}^\perp$ different from zero.

Let $v \in T_K^b(x) \cap \{q\}^\perp$. If $I(x) \neq \emptyset$, then $0 = \langle q, v\rangle = \left\langle \sum_{j=1}^k \lambda_j \nabla\psi_{i_j}(x), v\right\rangle$, which yields $\lambda_j \langle \nabla\psi_{i_j}(x), v\rangle = 0$ for every $j = 1, \ldots, k$. Hence, $\lambda_j = 0$ whenever $i_j \notin I_v(x)$. Furthermore, if the equality constraints are absent, then $I_v(x) \neq \emptyset$ for every $v \in T_K^b(x) \cap \{q\}^\perp$. Consequently,

$$\langle q, h\rangle + \frac{1}{2}\sum_{i=1}^{p} \mu_i \langle \varphi_i''(x)v, v\rangle + \frac{1}{2}\sum_{j=1}^{k} \lambda_j \langle \psi_{i_j}''(x)v, v\rangle \le 0, \qquad \forall\, h \in T_K^{\flat(2)}(x,v).$$

Therefore, by arbitrariness of $v \in T_K^b(x) \cap \{q\}^\perp$,

$$\overline{Q} := \sum_{i=1}^{p} \mu_i \varphi_i''(x) + \sum_{j=1}^{k} \lambda_j \psi_{i_j}''(x) \in N_K^{b(2)}(x,q).$$

Observe that if a symmetric $(n \times n)$-matrix Q is so that $\langle Qv, v\rangle \le \langle \overline{Q}v, v\rangle$ for every $v \in T_K^b(x) \cap \{q\}^\perp$, denoted by $Q \le \overline{Q}$, then $Q \in N_K^{b(2)}(x,q)$.

We show next that $\overline{Q}$ is the largest second order normal in the above sense. Fix any $Q \in N_K^{b(2)}(x,q)$. Let $v \in T_K^b(x) \cap \{q\}^\perp$. If $v = 0$, then $\langle Qv, v\rangle \le \langle \overline{Q}v, v\rangle$. Assume next that $v \neq 0$. If $I_v(x) \neq \emptyset$, consider the set $\{j_1, ..., j_m\}$ of all the indices that belong to $I_v(x)$. Define the $(n \times (p+m))$-matrix A such that its s-th column is $\nabla\varphi_s(x)$ for $1 \le s \le p$ and $\nabla\psi_{j_{s-p}}(x)$ for $p+1 \le s \le p+m$ (we set $m = 0$ if $I_v(x) = \emptyset$). By the linear independence assumption, we show that for any $0 \neq v \in T_K^b(x) \cap \{q\}^\perp$ there exists $z_v \in \mathbb{R}^n$ satisfying

$$z_v^\top A = -\frac{1}{2}\Big(\langle \varphi_1''(x)v, v\rangle, ..., \langle \varphi_p''(x)v, v\rangle, \langle \psi_{j_1}''(x)v, v\rangle, ..., \langle \psi_{j_m}''(x)v, v\rangle\Big).$$

Hence $z_v \in T_K^{b(2)}(x,v)$ and $\langle q, z_v\rangle = -\frac{1}{2}\sum_{i=1}^p \mu_i \langle \varphi_i''(x)v, v\rangle - \frac{1}{2}\sum_{j=1}^k \lambda_j \langle \psi_{i_j}''(x)v, v\rangle$. Thus

$$\langle q, z_v\rangle + \frac{1}{2}\langle Qv, v\rangle \le 0 = \langle q, z_v\rangle + \frac{1}{2}\sum_{i=1}^{p} \mu_i \langle \varphi_i''(x)v, v\rangle + \frac{1}{2}\sum_{j=1}^{r} \lambda_j \langle \psi_j''(x)v, v\rangle.$$

Consequently $Q \le \overline{Q}$ in the above sense.

However, in general, closed sets do not have the above representation. We refer to [11] for a very simple example of a set K given by union of two intervals in $\mathbb{R}^2$, where the first and second order tangents can be easily computed, but, at the same time, K does not satisfy the constraint qualification assumption.

We would like to underline here that to prove the celebrated Pontryagin maximum principle in optimal control just a particular subset of tangents to the set of controlled trajectories was used. The computation of the whole tangent cone is, in general, not possible. Similarly, we do not need to know the whole set of the second order tangents to eliminate some candidates for optimality.

Let $(\Xi, \mathscr{G})$ be a measurable space, and $F : \Xi \rightsquigarrow 2^X$ be a set-valued map. For any $\xi \in \Xi$, $F(\xi)$ is called the value of F at ξ. The domain of F is the subset of all $\xi \in \Xi$ such that $F(\xi)$ is nonempty, i.e., $Dom\,(F) := \{\xi \in \Xi \mid F(\xi) \neq \emptyset\}$. F is called measurable if $F^{-1}(A) := \{\xi \in \Xi \mid F(\xi) \cap A \neq \emptyset\} \in \mathscr{G}$ for any $A \in \mathcal{B}(X)$. Clearly, the domain of a measurable set-valued map is measurable.

The following result is a special case of [2, Theorem 8.5.1].

Lemma 2.1. *Suppose $(\Xi, \mathscr{G}, \mu)$ is a complete σ-finite measure space, X is a separable Banach space, $p \ge 1$ and K is a closed nonempty subset in X. Define*

$$\mathcal{K} := \Big\{\varphi(\cdot) \in L^p(\Xi, \mathscr{G}, \mu; X) \;\Big|\; \varphi(\xi) \in K,\ \mu\text{-}a.e.\ \xi \in \Xi\Big\}.$$

Then for any $\varphi(\cdot) \in \mathcal{K}$, the set-valued map $T^b_K(\varphi(\cdot))$: $\xi \rightsquigarrow T^b_K(\varphi(\xi))$ is $\mathscr{G}$-measurable, and

$$\mathcal{T} := \left\{ \psi(\cdot) \in L^p(\Xi, \mathscr{G}, \mu; X) \;\middle|\; \psi(\xi) \in T^b_K(\varphi(\xi)),\ \mu\text{-a.e. } \xi \in \Xi \right\} \subset T^b_{\mathcal{K}}(\varphi(\cdot)).$$

The following result is a special case of [2, Corollary 8.2.13].

Lemma 2.2. *Suppose $(\Xi, \mathscr{G}, \mu)$ is a complete σ-finite measure space, X is a separable Banach space, K is a closed nonempty subset in X and $\varphi(\cdot)$ is a $\mathscr{G}$-measurable single-valued mapping. Then the projection mapping $\xi \rightsquigarrow \Pi_K(\varphi(\xi))$ is $\mathscr{G}$-measurable. Moreover, if for every ξ, $\Pi_K(\varphi(\xi)) \neq \emptyset$, then there exists a $\mathscr{G}$-measurable, X-valued selection $\psi(\cdot)$ such that $\|\psi(\xi) - \varphi(\xi)\|_X = dist\,(\varphi(\xi), K)$, μ-a.e.*

As in [18], we call a measurable set-valued map $\zeta : (\Omega, \mathcal{F}) \rightsquigarrow 2^{\mathbb{R}^m}$ a set-valued random variable, and, we call a map $\Gamma : [0, T] \times \Omega \rightsquigarrow 2^{\mathbb{R}^m}$ a measurable set-valued stochastic process if Γ is $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable. We say that Γ is $\mathbb{F}$-adapted if $\Gamma(t)$ is $\mathcal{F}_t$-measurable for any $t \in [0, T]$. Define

$$\mathscr{G} := \left\{ A \in \mathcal{B}([0, T]) \otimes \mathcal{F} \;\middle|\; A_t \in \mathcal{F}_t,\ \forall\, t \in [0, T] \right\}, \tag{2.1}$$

where $A_t := \{\omega \in \Omega \mid (t, \omega) \in A\}$ is the section of A. Obviously, $\mathscr{G}$ is a sub-σ-algebra of $\mathcal{B}([0, T]) \otimes \mathcal{F}$. As pointed in [18, p. 96], the following result holds.

Lemma 2.3. *A set-valued stochastic process $\Gamma : [0, T] \times \Omega \rightsquigarrow 2^{\mathbb{R}^m}$ is $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted if and only if Γ is $\mathscr{G}$-measurable.*

Obviously, $\mathcal{U}_{ad}$ is a nonempty closed subset of the Banach space $L^2_{\mathbb{F}}(\Omega; L^2(0, T); \mathbb{R}^m)$. Using Lemmas 2.1 and 2.3, the following result was derived in [31]. It is useful later in getting the desired pointwise first order necessary condition.

Lemma 2.4. *([31, Lemma 4.6]) Let U be closed, $\tilde{u}(\cdot) \in \mathcal{U}_{ad}$, and $F : [0, T] \times \Omega \to \mathbb{R}^m$ be a $\mathcal{B}([0, T]) \times \mathcal{F}$-measurable and $\mathbb{F}$-adapted process such that*

$$\mathbb{E} \int_0^T \langle F(t), v(t) \rangle\, dt \leq 0, \quad \forall\, v(\cdot) \in T^b_{\mathcal{U}_{ad}}(\tilde{u}(\cdot)).$$

Then,

$$\langle F(t, \omega), v \rangle \leq 0, \quad \forall\, v \in T^b_U(\tilde{u}(t, \omega)),\ a.e.\ (t, \omega) \in [0, T] \times \Omega.$$

2.3. Some concepts and results from the Malliavin calculus

In this subsection, we recall some concepts and results from the Malliavin calculus (see [25] for a detailed discussion on this topic).

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

For any $\eta \in L^2(0,T)$, write $\mathcal{W}(\eta) = \int_0^T \eta(t)dW(t)$. Define

$$\mathcal{S} := \Big\{ \zeta = \varphi(\mathcal{W}(\eta_1),\ \mathcal{W}(\eta_2),\ \cdots,\ \mathcal{W}(\eta_d)) \ \Big| \ \varphi \in C_b^\infty(\mathbb{R}^d;\mathbb{R}^n), \\ \eta_1,\eta_2,\cdots,\eta_d \in L^2(0,T), d \in \mathbb{N} \Big\}. \tag{2.2}$$

Clearly, $\mathcal{S}$ is a linear subspace of $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$. For any $\zeta \in \mathcal{S}$ (as in (2.2)), its Malliavin derivative is defined as follows:

$$\mathcal{D}_s\zeta := \sum_{i=1}^d \eta_i(s)\frac{\partial\varphi}{\partial x_i}(\mathcal{W}(\eta_1),\ \mathcal{W}(\eta_2),\ \cdots,\ \mathcal{W}(\eta_d)), \quad \text{a.e. } s \in [0,T], \ \text{a.s.}$$

Write

$$|||\zeta|||_2 := \Big[\mathbb{E}\,|\zeta|^2 + \mathbb{E}\int_0^T |\mathcal{D}_s\zeta|^2 ds\Big]^{\frac{1}{2}}.$$

Obviously, $|||\cdot|||_2$ is a norm on $\mathcal{S}$. It is shown in [25] that the operator $\mathcal{D}$ has a closed extension to the space $\mathbb{D}^{1,2}(\mathbb{R}^n)$, the completion of $\mathcal{S}$ with respect to the norm $|||\cdot|||_2$. When $\zeta \in \mathbb{D}^{1,2}(\mathbb{R}^n)$, the following Clark–Ocone representation formula holds:

$$\zeta = \mathbb{E}\,\zeta + \int_0^T \mathbb{E}\,(\mathcal{D}_s\zeta \mid \mathcal{F}_s)dW(s). \tag{2.3}$$

Furthermore, if ζ is $\mathcal{F}_t$-measurable, then $\mathcal{D}_s\zeta = 0$ for any $s \in (t,T]$.

Let $\mathbb{L}^{1,2}(\mathbb{R}^n)$ denote the space of processes $\varphi \in L^2([0,T]\times\Omega;\mathbb{R}^n)$ such that

(i) for a.e. $t \in [0,T]$, $\varphi(t,\cdot) \in \mathbb{D}^{1,2}(\mathbb{R}^n)$;
(ii) the function $\mathcal{D}_\cdot\varphi(\cdot,\cdot) : \ [0,T]\times[0,T]\times\Omega \to \mathbb{R}^n$ admits a $\mathcal{B}([0,T]\times[0,T])\otimes\mathcal{F}$-measurable version;
(iii) $|||\varphi|||_{1,2} := \big[\mathbb{E}\int_0^T |\varphi(t,\omega)|^2 dt + \mathbb{E}\int_0^T\int_0^T |\mathcal{D}_s\varphi(t,\omega)|^2 ds dt\big]^{\frac{1}{2}} < +\infty$.

Denote by $\mathbb{L}^{1,2}_{\mathbb{F}}(\mathbb{R}^n)$ the set of all $\mathbb{F}$-adapted processes in $\mathbb{L}^{1,2}(\mathbb{R}^n)$.

In addition, write

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

$$\mathbb{L}^{1,2}_{2^-}(\mathbb{R}^n) := \Big\{ \varphi \in \mathbb{L}^{1,2}(\mathbb{R}^n) \Big| \; \exists\, \mathcal{D}^-\varphi \in L^2([0,T]\times\Omega;\mathbb{R}^n) \text{ s. t. for any small } \varepsilon > 0,$$

$$g_\varepsilon(s) := \sup_{(s-\varepsilon)\vee 0<t<s} \mathbb{E}\left|\mathcal{D}_s\varphi(t,\omega) - \mathcal{D}^-\varphi(s,\omega)\right|^2 < \infty, \text{ a.e. } s\in[0,T],$$

$$g_\varepsilon(\cdot) \text{ is measurable on } [0,T], \text{ and } \lim_{\varepsilon\to 0^+}\int_0^T g_\varepsilon(s)ds = 0 \Big\}.$$

Set $\mathbb{L}^{1,2}_{2}(\mathbb{R}^n) = \mathbb{L}^{1,2}_{2^+}(\mathbb{R}^n) \cap \mathbb{L}^{1,2}_{2^-}(\mathbb{R}^n)$ and define

$$\nabla\varphi = \mathcal{D}^+\varphi + \mathcal{D}^-\varphi, \quad \forall\, \varphi \in \mathbb{L}^{1,2}_{2}(\mathbb{R}^n).$$

When φ is $\mathbb{F}$-adapted, $\mathcal{D}_s\varphi(t,\omega) = 0$ a.s. for any $t < s$. In this case, $\mathcal{D}^-\varphi = 0$ and $\nabla\varphi = \mathcal{D}^+\varphi$ a.e. $t\in[0,T]$, a.s. Denote by $\mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^n)$ the set of all $\mathbb{F}$-adapted processes in $\mathbb{L}^{1,2}_{2}(\mathbb{R}^n)$.

Roughly speaking, an element $\varphi \in \mathbb{L}^{1,2}_{2}(\mathbb{R}^n)$ is a stochastic process whose Malliavin derivative has suitable continuity on some neighborhood of $\{(t,t) \mid t\in[0,T]\}$. Examples of such processes can be found in [25]. Especially, if $(s,t)\mapsto \mathcal{D}_s\varphi(t,\omega)$ is continuous from $V_\delta := \{(s,t) \big| \; |s-t| < \delta, \;\; s,t\in[0,T]\}$ (for some $\delta>0$) to $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$, then $\varphi\in\mathbb{L}^{1,2}_{2}(\mathbb{R}^n)$ and, $\mathcal{D}^+\varphi(t,\omega) = \mathcal{D}^-\varphi(t,\omega) = \mathcal{D}_t\varphi(t,\omega)$ a.e. $t\in[0,T]$, a.s.

3. First order necessary conditions

In this section, we study the first order necessary optimality conditions for the optimal control problem (1.3). Firstly, we introduce the notion of local minimizer for the problem (1.3).

Definition 3.1. An admissible triple $(\bar{x},\bar{u},\bar{x}_0) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \times \mathcal{U}_{ad} \times K$ is called a local minimizer for the problem (1.3) if there exists a $\delta>0$ such that $J(u,x_0) \ge J(\bar{u},\bar{x}_0)$ for any admissible triple $(x,u,x_0) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \times \mathcal{U}_{ad} \times K$ satisfying $\|u-\bar{u}\|_2 < \delta$ and $|\bar{x}_0 - x_0| < \delta$.

In this section, we need the following assumptions:

(C1) *The control region U is nonempty and closed.*

(C2) *The functions b, σ, f and g satisfy the following:*

(i) *For any $(x,u)\in\mathbb{R}^n\times\mathbb{R}^m$, the stochastic processes $b(\cdot,x,u,\cdot): [0,T]\times\Omega\to\mathbb{R}^n$ and $\sigma(\cdot,x,u,\cdot): [0,T]\times\Omega\to\mathbb{R}^n$ are $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted. For a.e. $(t,\omega)\in[0,T]\times\Omega$, the functions $b(t,\cdot,\cdot,\omega): \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ and $\sigma(t,\cdot,\cdot,\omega): \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ are differentiable and*

$$(x,u)\mapsto (b_x(t,x,u,\omega), b_u(t,x,u,\omega), \sigma_x(t,x,u,\omega), \sigma_u(t,x,u,\omega))$$

is uniformly continuous in $x\in\mathbb{R}^n$ and $u\in\mathbb{R}^m$. There exist a constant $L>0$ and a nonnegative $\eta\in L^\beta_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}))$ with $\eta(T,\cdot)\in L^\beta_{\mathcal{F}_T}(\Omega;\mathbb{R})$ and $\beta\ge 1$ such that for a.e. $(t,\omega)\in[0,T]\times\Omega$ and for any $x\in\mathbb{R}^n$ and $u\in\mathbb{R}^m$,

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

$$\begin{cases} |b(t,0,u,\omega)|+|\sigma(t,0,u,\omega)|\le L(\eta(t,\omega)+|u|), \\ |b_x(t,x,u,\omega)|+|b_u(t,x,u,\omega)|\le L, \\ |\sigma_x(t,x,u,\omega)|+|\sigma_u(t,x,u,\omega)|\le L; \end{cases}$$

(ii) *For any $(x,u)\in\mathbb{R}^n\times\mathbb{R}^m$, the stochastic process $f(\cdot,x,u,\cdot):\ [0,T]\times\Omega\to\mathbb{R}$ is $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted, and the random variable $g(x,\cdot)$ is $\mathcal{F}_T$-measurable. For a.e. $(t,\omega)\in[0,T]\times\Omega$, the functions $f(t,\cdot,\cdot,\omega):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ and $g(\cdot,\omega):\ \mathbb{R}^n\to\mathbb{R}$ are differentiable, and for any $x,\ \tilde{x}\in\mathbb{R}^n$ and $u,\ \tilde{u}\in\mathbb{R}^m$,*

$$\begin{cases} |f(t,x,u,\omega)|\le L(\eta(t,\omega)^2+|x|^2+|u|^2), \\ |f_x(t,0,u,\omega)|+|f_u(t,0,u,\omega)|\le L(\eta(t,\omega)+|u|), \\ |f_x(t,x,u,\omega)-f_x(t,\tilde{x},\tilde{u},\omega)|+|f_u(t,x,u,\omega)-f_u(t,\tilde{x},\tilde{u},\omega)| \\ \quad\le L(|x-\tilde{x}|+|u-\tilde{u}|), \\ |g(x,\omega)|\le L(\eta(T,\omega)^2+|x|^2),\ \ |g_x(0,\omega)|\le L\eta(T,\omega), \\ |g_x(x,\omega)-g_x(\tilde{x},\omega)|\le L|x-\tilde{x}|. \end{cases}$$

When the condition (C2) is satisfied, the state x (of (1.1)) is uniquely defined by any given initial datum $x_0\in\mathbb{R}^n$ and admissible control $u\in\mathcal{U}_{ad}$, and the cost functional (1.2) is well-defined on $\mathcal{U}_{ad}$. In what follows, C represents a generic positive constant (depending only on T, β, $\eta(\cdot)$ and L), which may be different from one place to another.

The following known result [24] is useful in the sequel.

Lemma 3.1. *Assume (C2). Then, for any $x_0\in\mathbb{R}^n$, $\beta\ge 1$ and $u\in L^{\beta}_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$, the state equation (1.1) admits a unique solution $x\in L^{\beta}_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$, and for any $t\in[0,T]$ the following estimate holds:*

$$\mathbb{E}\Big(\sup_{s\in[0,t]}|x(s,\omega)|^{\beta}\Big)\le C\mathbb{E}\Big[|x_0|^{\beta}+\Big(\int_0^t|b(s,0,u(s),\omega)|ds\Big)^{\beta} + \Big(\int_0^t|\sigma(s,0,u(s),\omega)|^2ds\Big)^{\frac{\beta}{2}}\Big]. \tag{3.1}$$

Moreover, if $\tilde{x}$ is the solution to (1.1) corresponding to $(\tilde{x}_0,\tilde{u})\in\mathbb{R}^n\times L^{\beta}_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$, then, for any $t\in[0,T]$,

$$\mathbb{E}\Big(\sup_{s\in[0,t]}|x(s,\omega)-\tilde{x}(s,\omega)|^{\beta}\Big)\le C\mathbb{E}\Big[|x_0-\tilde{x}_0|^{\beta}+\Big(\int_0^t|u(s,\omega)-\tilde{u}(s,\omega)|^2ds\Big)^{\frac{\beta}{2}}\Big]. \tag{3.2}$$

Now, let us introduce the classical first order variational control system. Let $\bar{u},v,v_\varepsilon\in L^{\beta}_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$ ($\beta\ge 1$) and $\nu_0,\nu_0^{\varepsilon}\in\mathbb{R}^n$ satisfying $v_\varepsilon\to v$ in $L^{\beta}_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$ and $\nu_0^{\varepsilon}\to\nu_0$ in $\mathbb{R}^n$ as $\varepsilon\to 0^+$. For $u^{\varepsilon}:=\bar{u}+\varepsilon v_\varepsilon$ and $x_0^{\varepsilon}:=x_0+\varepsilon\nu_0^{\varepsilon}$, let x^{ε} be the state of (1.1) corresponding to the control u^ε and the initial datum x_0^ε, and put $\delta x^\varepsilon = x^\varepsilon - \bar{x}$. For $\varphi = b, \sigma, f$, denote

$$\varphi_x(t) = \varphi_x(t, \bar{x}(t), \bar{u}(t)), \quad \varphi_u(t) = \varphi_u(t, \bar{x}(t), \bar{u}(t)).$$

Consider the following linearized stochastic control system:

$$\begin{cases} dy_1(t) = \big(b_x(t)y_1(t) + b_u(t)v(t)\big)dt + \big(\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\big)dW(t), & t \in [0,T], \\ y_1(0) = \nu_0. \end{cases} \tag{3.3}$$

We first establish the following estimates.

Lemma 3.2. *Let (C2) hold and $\beta \geq 1$. Then, for any $\bar{u}$, v, v_ε, ν_0, ν_0^ε and δx^ε as above*

$$\|y_1\|_{\infty,\beta}^\beta \leq C\big(|\nu_0|^\beta + \|v\|_{2,\beta}^\beta\big), \quad \|\delta x^\varepsilon\|_{\infty,\beta}^\beta = O(\varepsilon^\beta).$$

Furthermore,

$$\|r_1^\varepsilon\|_{\infty,\beta}^\beta \to 0, \quad \text{as } \varepsilon \to 0^+, \tag{3.4}$$

where $r_1^\varepsilon(t,\omega) := \frac{\delta x^\varepsilon(t,\omega)}{\varepsilon} - y_1(t,\omega)$.

Proof. See Appendix A.1. □

Next, define the Hamiltonian

$$H(t,x,u,p,q,\omega) := \langle p, b(t,x,u,\omega)\rangle + \langle q, \sigma(t,x,u,\omega)\rangle - f(t,x,u,\omega), \tag{3.5}$$

where $(t,x,u,p,q,\omega) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \Omega$. We introduce the first order adjoint equation for (3.3):

$$\begin{cases} dP_1(t) = -\big(b_x(t)^\top P_1(t) + \sigma_x(t)^\top Q_1(t) - f_x(t)\big)dt + Q_1(t)dW(t), & t \in [0,T], \\ P_1(T) = -g_x(\bar{x}(T)). \end{cases} \tag{3.6}$$

By [8] and (C2), for any $\beta \geq 1$, if $\bar{u} \in L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$, the equation (3.6) admits a unique strong solution $(P_1, Q_1) \in L^\beta_{\mathbb{F}}(\Omega;\ C([0,T];\mathbb{R}^n)) \times L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^n))$.

We have the following result.

Theorem 3.1. *Let (C1)–(C2) hold. If $(\bar{x}, \bar{u}, \bar{x}_0)$ is a local minimizer for the problem (1.3), then*

$$\mathbb{E}\int_0^T \langle H_u(t), v(t)\rangle\, dt \leq 0, \quad \forall\, v \in T^b_{\mathcal{U}_{ad}}(\bar{u}), \tag{3.7}$$

and

$$P_1(0) \in N_K^b(\bar{x}_0), \tag{3.8}$$

where (P_1, Q_1) *is the solution to the first order adjoint equation* (3.6) *corresponding to* $(\bar{x}, \bar{u}, \bar{x}_0)$ *and* $H_u(t) = H_u(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t))$.

Proof. Let $v \in T^b_{\mathcal{U}_{ad}}(\bar{u})$ and $\nu_0 \in T^b_K(\bar{x}_0)$. Then, for any $\varepsilon > 0$, there exist $v_\varepsilon \in L^2_{\mathbb{F}}(\Omega; L^2(0, T; \mathbb{R}^m))$ and $\nu_0^\varepsilon \in \mathbb{R}^n$ such that $\bar{u} + \varepsilon v_\varepsilon \in \mathcal{U}_{ad}$, $\bar{x}_0 + \varepsilon \nu_0^\varepsilon \in K$ and

$$\mathbb{E}\int_0^T |v(t) - v_\varepsilon(t)|^2 dt \to 0, \quad |\nu_0^\varepsilon - \nu_0| \to 0, \ \text{ as } \varepsilon \to 0^+.$$

Expanding the cost functional $J(\cdot)$ at $\bar{u}$, we have for all small $\varepsilon > 0$,

$$\begin{aligned}
0 \le \ & \frac{J(u^\varepsilon, x_0^\varepsilon) - J(\bar{u}, \bar{x}_0)}{\varepsilon} \\
= \ & \mathbb{E}\int_0^T \Big(\int_0^1 \Big\langle f_x(t, \bar{x}(t) + \theta\delta x^\varepsilon(t), \bar{u}(t) + \varepsilon v_\varepsilon(t)), \frac{\delta x^\varepsilon(t)}{\varepsilon} \Big\rangle d\theta \\
& \qquad + \int_0^1 \langle f_u(t, \bar{x}(t), \bar{u}(t) + \theta\varepsilon v_\varepsilon(t)), v_\varepsilon(t) \rangle \, d\theta \Big) dt \\
& + \mathbb{E}\int_0^1 \Big\langle g_x(\bar{x}(T) + \theta\delta x^\varepsilon(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} \Big\rangle d\theta \\
= \ & \mathbb{E}\int_0^T \big(\langle f_x(t), y_1(t) \rangle + \langle f_u(t), v(t) \rangle \big) dt + \mathbb{E}\, \langle g_x(\bar{x}(T)), y_1(T) \rangle + \rho_1^\varepsilon,
\end{aligned} \tag{3.9}$$

where

$$\begin{aligned}
\rho_1^\varepsilon = \ & \mathbb{E}\int_0^T \Big(\int_0^1 \Big\langle f_x(t, \bar{x}(t) + \theta\delta x^\varepsilon(t), \bar{u}(t) + \varepsilon v_\varepsilon(t)) - f_x(t), \frac{\delta x^\varepsilon(t)}{\varepsilon} \Big\rangle d\theta \\
& \qquad + \int_0^1 \langle f_u(t, \bar{x}(t), \bar{u}(t) + \theta\varepsilon v_\varepsilon(t)) - f_u(t), v_\varepsilon(t) \rangle \, d\theta \\
& \qquad + \Big\langle f_x(t), \frac{\delta x^\varepsilon(t)}{\varepsilon} - y_1(t) \Big\rangle + \langle f_u(t), v_\varepsilon(t) - v(t) \rangle \Big) dt \\
& + \mathbb{E}\int_0^1 \Big\langle g_x(\bar{x}(T) + \theta\delta x^\varepsilon(T)) - g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} \Big\rangle d\theta \\
& \qquad + \mathbb{E}\Big\langle g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} - y_1(T) \Big\rangle.
\end{aligned} \tag{3.10}$$

By Lemma 3.2 (with $\beta = 2$) and (C2), it follows that

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^T\int_0^1 \Big\langle f_x(t,\bar{x}(t)+\theta\delta x^\varepsilon(t),\bar{u}(t)+\varepsilon v_\varepsilon(t)) - f_x(t), \frac{\delta x^\varepsilon(t)}{\varepsilon}\Big\rangle d\theta dt\Big| \\
&\le \Big(\mathbb{E}\int_0^T\int_0^1 \big|f_x(t,\bar{x}(t)+\theta\delta x^\varepsilon(t),\bar{u}(t)+\varepsilon v_\varepsilon(t)) - f_x(t)\big|^2 d\theta dt\Big)^{\frac{1}{2}}\Big(\mathbb{E}\int_0^T \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^2 dt\Big)^{\frac{1}{2}} \\
&\le C\Big[\mathbb{E}\int_0^T \big(|\delta x^\varepsilon(t)| + |\varepsilon v_\varepsilon(t)|\big)^2 dt\Big]^{\frac{1}{2}} \cdot \Big(\mathbb{E}\int_0^T \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^2 dt\Big)^{\frac{1}{2}} \\
&\to 0, \quad \text{as } \varepsilon \to 0^+.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^T\int_0^1 \langle f_u(t,\bar{x}(t),\bar{u}(t)+\theta\varepsilon v_\varepsilon(t)) - f_u(t), v_\varepsilon(t)\rangle\, d\theta dt\Big| \\
&\le C\Big(\mathbb{E}\int_0^T |\varepsilon v_\varepsilon(t)|^2 dt\Big)^{\frac{1}{2}} \cdot \Big(\mathbb{E}\int_0^T |v_\varepsilon(t)|^2 dt\Big)^{\frac{1}{2}} \to 0, \quad \varepsilon \to 0^+
\end{aligned}$$

and

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^1 \Big\langle g_x(\bar{x}(T)+\theta\delta x^\varepsilon(T)) - g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon}\Big\rangle d\theta\Big| \\
&\le C\Big(\mathbb{E}|\delta x^\varepsilon(T)|^2\Big)^{\frac{1}{2}} \cdot \Big(\mathbb{E}\Big|\frac{\delta x^\varepsilon(T)}{\varepsilon}\Big|^2\Big)^{\frac{1}{2}} \to 0, \quad \varepsilon \to 0^+.
\end{aligned}$$

Then, by (C2) and Lemma 3.2, we obtain that

$$\begin{aligned}
\lim_{\varepsilon\to 0^+} |\rho_1^\varepsilon| &\le \limsup_{\varepsilon\to 0^+}\Big|\mathbb{E}\int_0^T \Big\langle f_x(t), \frac{\delta x^\varepsilon(t)}{\varepsilon} - y_1(t)\Big\rangle dt\Big| \\
&\quad + \limsup_{\varepsilon\to 0^+}\Big|\mathbb{E}\int_0^T \langle f_u(t), v_\varepsilon(t) - v(t)\rangle\, dt\Big| \\
&\quad + \limsup_{\varepsilon\to 0^+}\Big|\mathbb{E}\Big\langle g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} - y_1(T)\Big\rangle\Big| = 0. \qquad (3.11)
\end{aligned}$$

Therefore, from (3.9) and (3.11), we conclude that

$$0 \leq \mathbb{E}\int_0^T \big(\langle f_x(t), y_1(t)\rangle + \langle f_u(t), v(t)\rangle\big)dt + \mathbb{E}\,\langle g_x(\bar{x}(T)), y_1(T)\rangle. \tag{3.12}$$

By the duality between (3.3) and (3.6), we have

$$\begin{aligned}
&\mathbb{E}\,\langle g_x(\bar{x}(T)), y_1(T)\rangle = -\mathbb{E}\,\langle P_1(T), y_1(T)\rangle \\
&= -\langle P_1(0), \nu_0\rangle - \mathbb{E}\int_0^T \big(\langle P_1(t), b_x(t)y_1(t)\rangle + \langle P_1(t), b_u(t)v(t)\rangle \\
&\qquad + \langle Q_1(t), \sigma_x(t)y_1(t)\rangle + \langle Q_1(t), \sigma_u(t)v(t)\rangle \\
&\qquad - \Big\langle b_x(t)^\top P_1(t), y_1(t)\Big\rangle - \Big\langle \sigma_x(t)^\top Q_1(t), y_1(t)\Big\rangle + \langle f_x(t), y_1(t)\rangle\big)dt \\
&= -\langle P_1(0), \nu_0\rangle - \mathbb{E}\int_0^T \big(\langle P_1(t), b_u(t)v(t)\rangle + \langle Q_1(t), \sigma_u(t)v(t)\rangle + \langle f_x(t), y_1(t)\rangle\big)dt.
\end{aligned} \tag{3.13}$$

Substituting (3.13) in (3.12), we obtain that

$$\begin{aligned}
0 &\leq -\langle P_1(0), \nu_0\rangle - \mathbb{E}\int_0^T \big(\langle P_1(t), b_u(t)v(t)\rangle + \langle Q_1(t), \sigma_u(t)v(t)\rangle - \langle f_u(t), v(t)\rangle\big)dt \\
&= -\langle P_1(0), \nu_0\rangle - \mathbb{E}\int_0^T \langle H_u(t), v(t)\rangle\, dt.
\end{aligned} \tag{3.14}$$

For $v(\cdot) = 0$, (3.14) implies (3.8). On the other hand, for $\nu_0 = 0$ in (3.14), we have (3.7). This completes the proof of Theorem 3.1. □

From Theorem 3.1 and Lemma 2.4, it is easy to deduce the following pointwise first order necessary condition.

Theorem 3.2. *Let (C1)–(C2) hold. If $(\bar{x}, \bar{u}, \bar{x}_0)$ is a local minimizer for the problem (1.3), then,*

$$H_u(t,\omega) \in N_U^b(\bar{u}(t,\omega)), \ \text{a.e. } t \in [0,T], \ \text{a.s. and } P_1(0) \in N_K^b(\bar{x}_0). \tag{3.15}$$

Remark 3.1. When the control set U and the initial state constraint set K are also convex, $N_U^b(\bar{u})$ and $N_K^b(\bar{x}_0)$ coincide with the normal cones of convex analysis. In this case, the condition (3.15) becomes

$$H_u(t,\omega) \in N_U(\bar{u}(t,\omega)) \quad \text{a.e. } t \in [0,T], \ \text{a.s. and } P_1(0) \in N_K(\bar{x}_0).$$

Remark 3.2. If $T_U^b(\bar{u}(t,\omega)) = \{0\}$ for a.e. $(t,\omega) \in [0,T] \times \Omega$, then $N_U^b(\bar{u}(t,\omega)) = \mathbb{R}^m$, for a.e. $(t,\omega) \in [0,T] \times \Omega$, and the first condition in (3.15) turns out to be trivial. It is the case, for instance, when the control set U is a finite union of singletons. Therefore, to have the first condition in (3.15) meaningful, U should have nontrivial tangent cones. It is not difficult to verify that for every $v \in T^b_{\mathcal{U}_{ad}}(\bar{u})$, and for a.e. $(t,\omega) \in [0,T] \times \Omega$, the vector $v(t,\omega)$ belongs to the *contingent cone* $T^B_U(\bar{u}(t,\omega))$ to U at $\bar{u}(t,\omega)$. Under some suitable assumptions on U, we have $T^B_U(\bar{u}(t,\omega)) = T^b_U(\bar{u}(t,\omega))$ a.e. in $[0,T] \times \Omega$, see [2, Chapter 4] for more details. Consequently, under some convenient structural assumptions on U, if $T^b_{\mathcal{U}_{ad}}(\bar{u}) \neq \{0\}$, then $T^b_U(\bar{u}(t,\omega)) \neq \{0\}$ on a set of positive measure.

Remark 3.3. Define

$$\begin{aligned}\mathcal{H}(t,x,u,\omega) &:= H(t,x,u,P_1(t),Q_1(t),\omega) - \frac{1}{2}\left\langle P_2(t)\sigma(t,\bar{x}(t),\bar{u}(t),\omega), \sigma(t,\bar{x}(t),\bar{u}(t),\omega)\right\rangle \\ &\quad + \frac{1}{2}\Big\langle P_2(t)\big(\sigma(t,x,u,\omega) - \sigma(t,\bar{x}(t),\bar{u}(t),\omega)\big), \sigma(t,x,u,\omega) - \sigma(t,\bar{x}(t),\bar{u}(t),\omega)\Big\rangle,\end{aligned}$$

where (P_2, Q_2) is the second order adjoint process with respect to $(\bar{x},\bar{u})$ (defined by (4.3) in Section 4). The stochastic maximum principle (e.g. [27]) says that, if $(\bar{x},\bar{u})$ is an optimal pair, then

$$\mathcal{H}(t,\bar{x}(t),\bar{u}(t),\omega) = \max_{v\in U} \mathcal{H}(t,\bar{x}(t),v,\omega), \qquad \text{a.e. } t \in [0,T], \text{ a.s.} \tag{3.16}$$

When b, σ and f are differentiable with respect to the variable u, (3.16) implies that

$$\langle H_u(t,\omega), v\rangle \leq 0, \qquad \forall\, v \in T^b_U(\bar{u}(t,\omega)), \text{ a.e. } t \in [0,T], \text{ a.s.},$$

i.e., the first condition in (3.15) holds (when U is convex, this also coincides with the corresponding result in [4]). However, to derive the maximum principle (3.16) one has to assume that b, σ, f and g are differentiable up to the second order with respect to the variable x, and the second order adjoint process (P_2, Q_2) should be introduced (even it does not appear in the condition (3.15)). Therefore, in practice, under the usual structural assumptions on U, it is more convenient to use the condition (3.15) directly.

In what follows we give a simple example to demonstrate how to use the condition (3.15) to check if a given admissible control is not optimal.

Example 3.1. Let $n = m = 2$, $T = 1$, $U = \{(u_1,u_2) \in \mathbb{R}^2 \mid u_1u_2 = 0, u_1 \in [-1,1], u_2 \in [-1,1]\}$. Clearly, this U is neither a finite set nor convex in $\mathbb{R}^2$. Consider the control system

$$\begin{cases} dx_1(t) = (x_2(t) - \frac{1}{2})dt + dW(t), & t \in [0,1], \\ dx_2(t) = u_1(t)dt + u_2(t)dW(t), & t \in [0,1], \\ x_1(0) = 0, x_2(0) = 0 \end{cases} \tag{3.17}$$

with the cost functional

$$J(u) = \frac{1}{2}\mathbb{E}|x_1(1) - W(1)|^2. \tag{3.18}$$

Define the Hamiltonian of this optimal control problem

$$H(t,(x_1,x_2),(u_1,u_2),(p_1^1,p_1^2),(q_1^1,q_1^2),\omega) = p_1^1(x_2-\frac{1}{2}) + p_1^2 u_1 + q_1^1 + q_1^2 u_2, \tag{3.19}$$

for all $(t,(x_1,x_2),(u_1,u_2),(p_1^1,p_1^2),(q_1^1,q_1^2),\omega) \in [0,1]\times\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\times\Omega$. In what follows, we show that the control $(u_1(t),u_2(t)) \equiv (0,0)$ is not a local minimizer.

Obviously, the corresponding solution to the control system (3.17) is

$$(x_1(t),x_2(t)) = (W(t)-\frac{t}{2},0), \tag{3.20}$$

and the first order adjoint equation is

$$\begin{cases} dP_1^1(t) = Q_1^1(t)dW(t), \quad t\in[0,1], \\ dP_1^2(t) = -P_1^1(t)dt + Q_1^2(t)dW(t), \quad t\in[0,1], \\ P_1^1(1) = \frac{1}{2}, \quad P_1^2(1) = 0. \end{cases} \tag{3.21}$$

It is easy to verify that the solution to (3.21) is

$$(P_1^1(t),Q_1^1(t)) = (\frac{1}{2},0),\ (P_1^2(t),Q_1^2(t)) = (\frac{1-t}{2},0), \quad \text{a.e. } (t,\omega)\in[0,1]\times\Omega. \tag{3.22}$$

Note that even though the Mangasarian–Fromowitz constraint qualification does not hold at $(0,0)$, we can easily obtain that

$$T_U^b((0,0)) = \{(v_1,v_2)\in\mathbb{R}^2 \mid v_1v_2 = 0\}.$$

By the first order condition in (3.15),

$$\langle H_u(t),v\rangle = P_1^2(t)v_1 \le 0, \quad \forall\, v=(v_1,v_2)\in T_U^b((0,0)).$$

Since $P_1^2(t) = \frac{1}{2}(1-t) > 0$ for any $t\in[0,1)$, a.s., chose $(v_1,v_2) = (1,0)$ we have

$$P_1^2(t)v_1 = \frac{1}{2}(1-t) > 0, \quad \text{a.e. } (t,\omega)\in[0,1]\times\Omega,$$

which is a contradiction. Therefore, $(u_1(t),u_2(t)) \equiv (0,0)$ is not an local minimizer.

Actually, choosing $(\bar u_1(t),\bar u_2(t)) \equiv (1,0)$, we find that the corresponding state is

$$(\bar x_1(t),\bar x_2(t)) = \Big(\frac{t^2}{2}-\frac{t}{2}+W(t),t\Big), \quad \forall\,(t,\omega)\in[0,1]\times\Omega, \tag{3.23}$$

and hence $\bar x_1(1) = W(1)$, i.e., the cost functional attains its minimum 0 and $(\bar u_1(t),\bar u_2(t)) \equiv (1,0)$ is the global minimizer. In addition, a simple calculation shows that the corresponding first order adjoint process is

$$(P_1^1(t),Q_1^1(t)) = (0,0),\ (P_1^2(t),Q_1^2(t)) = (0,0), \quad \forall\,(t,\omega)\in[0,1]\times\Omega, \tag{3.24}$$

which implies that the condition (3.15) is trivially satisfied.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Remark 3.4. The approach proposed in Theorems 3.1–3.2 can be applied to more general control problems. We refer the reader to [31] for the optimal control problems involving stochastic Volterra integral equations.

4. Second order necessary conditions

In this section, we investigate the second order necessary conditions for the local minimizers $(\bar{x}, \bar{u}, \bar{x}_0)$ of (1.3). In addition to the assumptions (C1) and (C2), we suppose that

(C3) *The functions b, σ, f and g satisfy the following:*
(i) *For a.e. $(t,\omega) \in [0,T] \times \Omega$, the functions $b(t,\cdot,\cdot,\omega) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\sigma(t,\cdot,\cdot,\omega) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are twice differentiable and*

$$(x,u) \mapsto (b_{(x,u)^2}(t,x,u,\omega), \sigma_{(x,u)^2}(t,x,u,\omega))$$

is uniformly continuous in $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$, and,

$$|b_{(x,u)^2}(t,x,u,\omega)| + |\sigma_{(x,u)^2}(t,x,u,\omega)| \le L, \qquad \forall\, (x,u) \in \mathbb{R}^n \times \mathbb{R}^m;$$

(ii) *For a.e. $(t,\omega) \in [0,T] \times \Omega$, the functions $f(t,\cdot,\cdot,\omega) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $g(\cdot,\omega) : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable, and for any $x, \tilde{x} \in \mathbb{R}^n$ and $u, \tilde{u} \in \mathbb{R}^m$,*

$$\begin{cases} |f_{(x,u)^2}(t,x,u,\omega)| \le L, \\ |f_{(x,u)^2}(t,x,u,\omega) - f_{(x,u)^2}(t,\tilde{x},\tilde{u},\omega)| \le L(|x-\tilde{x}| + |u-\tilde{u}|), \\ |g_{xx}(x,\omega)| \le L, \ |g_{xx}(x,\omega) - g_{xx}(\tilde{x},\omega)| \le L|x-\tilde{x}|. \end{cases}$$

For $\varphi = b, \sigma, f$, denote

$$\varphi_{xx}(t) = \varphi_{xx}(t,\bar{x}(t),\bar{u}(t)), \quad \varphi_{xu}(t) = \varphi_{xu}(t,\bar{x}(t),\bar{u}(t)), \quad \varphi_{uu}(t) = \varphi_{uu}(t,\bar{x}(t),\bar{u}(t)).$$

4.1. *Integral-type second order necessary conditions*

In this subsection, we consider first the integral-type second order necessary conditions for the local minimizers of (1.3).

Let $\bar{u}, v, h, h_\varepsilon \in L^{2\beta}_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ $(\beta \ge 1)$ and $\nu_0, \varpi_0, \varpi_0^\varepsilon \in \mathbb{R}^m$ be such that h_ε converges to h in $L^{2\beta}_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ and $\varpi_0^\varepsilon \to \varpi_0$ in $\mathbb{R}^m$ as $\varepsilon \to 0^+$. Set

$$u^\varepsilon := \bar{u} + \varepsilon v + \varepsilon^2 h_\varepsilon, \qquad x_0^\varepsilon := \bar{x}_0 + \varepsilon \nu_0 + \varepsilon^2 \varpi_0^\varepsilon.$$

Denote by x^ε the solution of (1.1) corresponding to the control u^ε and the initial datum x_0^ε. Put

$$\delta x^\varepsilon = x^\varepsilon - \bar{x}, \qquad \delta u^\varepsilon = \varepsilon v + \varepsilon^2 h_\varepsilon.$$

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Similarly to [17], we introduce the following second-order variational equation:

$$\begin{cases} dy_2(t) = \Big(b_x(t)y_2(t) + 2b_u(t)h(t) + y_1(t)^\top b_{xx}(t)y_1(t) + 2v(t)^\top b_{xu}(t)y_1(t) \\ \qquad\qquad + v(t)^\top b_{uu}(t)v(t)\Big)dt + \Big(\sigma_x(t)y_2(t) + 2\sigma_u(t)h(t) + y_1(t)^\top \sigma_{xx}(t)y_1(t) \\ \qquad\qquad + 2v(t)^\top \sigma_{xu}(t)y_1(t) + v(t)^\top \sigma_{uu}(t)v(t)\Big)dW(t), \qquad t \in [0,T], \\ y_2(0) = 2\varpi_0, \end{cases} \tag{4.1}$$

where y_1 is the solution to the first variational equation (3.3) (for $v(\cdot)$ and ν_0 as above). We have the following estimates.

Lemma 4.1. *Let (C2)–(C3) hold and $\beta \geq 1$. Then, for $\bar{u}, v, h, h_\varepsilon \in L^{2\beta}_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ and $\nu_0, \varpi_0, \varpi_0^\varepsilon \in \mathbb{R}^m$ as above, we have*

$$\|y_2\|^\beta_{\infty,\beta} \leq C(|\varpi_0|^\beta + |\nu_0|^{2\beta} + \|v\|^{2\beta}_{4,2\beta} + \|h\|^\beta_{2,\beta}).$$

Furthermore,

$$\|r_2^\varepsilon\|^\beta_{\infty,\beta} \to 0, \quad \varepsilon \to 0^+, \tag{4.2}$$

where,

$$r_2^\varepsilon(t,\omega) := \frac{\delta x^\varepsilon(t,\omega) - \varepsilon y_1(t,\omega)}{\varepsilon^2} - \frac{1}{2}y_2(t,\omega).$$

Proof. See Appendix A.2. □

We now introduce the following adjoint equation for (4.1):

$$\begin{cases} dP_2(t) = -\Big(b_x(t)^\top P_2(t) + P_2(t)b_x(t) + \sigma_x(t)^\top P_2(t)\sigma_x(t) + \sigma_x(t)^\top Q_2(t) \\ \qquad\qquad + Q_2(t)\sigma_x(t) + H_{xx}(t)\Big)dt + Q_2(t)dW(t), \ t \in [0,T], \\ P_2(T) = -g_{xx}(\bar{x}(T)), \end{cases} \tag{4.3}$$

where $H_{xx}(t) = H_{xx}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t))$ with $(P_1(\cdot), Q_1(\cdot))$ given by (3.6).

By [8] and (C2)–(C3), it is easy to check that, if $\bar{u} \in L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$, (4.3) admits a unique strong solution $(P_2(\cdot), Q_2(\cdot)) \in L^\beta_{\mathbb{F}}(\Omega; C([0,T]; \mathbf{S}^n)) \times L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\mathbf{S}^n))$ for any $\beta \geq 1$.

To simplify the notation, we define

$$\begin{aligned} \mathbb{S}(t,x,u,y_1,z_1,y_2,z_2,\omega) := {} & H_{xu}(t,x,u,y_1,z_1,\omega) + b_u(t,x,u,\omega)^\top y_2 \\ & + \sigma_u(t,x,u,\omega)^\top z_2 + \sigma_u(t,x,u,\omega)^\top y_2\sigma_x(t,x,u,\omega), \end{aligned} \tag{4.4}$$

where $(t,x,u,y_1,z_1,y_2,z_2,\omega) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbf{S}^n \times \mathbf{S}^n \times \Omega$, and denote

$$\mathbb{S}(t) = \mathbb{S}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t), P_2(t), Q_2(t)), \quad t \in [0,T]. \tag{4.5}$$

Let $\bar{u} \in \mathcal{U}_{ad} \cap L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$. Define

$$\Upsilon_{\bar{u}} := \Big\{ v \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m)) \;\Big|\; \langle H_u(t,\omega), v(t,\omega)\rangle = 0 \text{ a.e. } t \in [0,T], \text{ a.s.} \Big\},$$

and the set of admissible second order variations by

$$\begin{aligned} \mathcal{A}_{\bar{u}} := \Big\{ (v,h) \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m)) \times L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m)) \;\Big|& \\ h(t,\omega) \in T_U^{b(2)}(\bar{u}(t,\omega), v(t,\omega)), \text{ a.e. } t \in [0,T], \text{ a.s.} \Big\}&. \end{aligned}$$

Denote

$$\mathcal{A}^1_{\bar{u}} := \Big\{ v \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m)) \;\Big|\; \exists\, h \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m)), \text{ s.t. } (v,h) \in \mathcal{A}_{\bar{u}} \Big\}.$$

We have the following result.

Theorem 4.1. *Let (C1)–(C3) hold and $(\bar{x}, \bar{u}, \bar{x}_0)$ be a local minimizer for the problem (1.3) with $\bar{u} \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$. Then for the adjoint process P_1 defined by (3.6) (relative to $(\bar{x}, \bar{u}, \bar{x}_0)$) and for all $(v,h) \in \mathcal{A}_{\bar{u}}$ satisfying $v \in \Upsilon_{\bar{u}}$,*

$$\begin{aligned} \mathbb{E}\int_0^T \Big(2\,\langle H_u(t), h(t)\rangle + \langle H_{uu}(t)v(t), v(t)\rangle& \\ + \langle P_2(t)\sigma_u(t)v(t), \sigma_u(t)v(t)\rangle + 2\,\langle \mathbb{S}(t)y_1(t), v(t)\rangle \Big) dt \le 0,& \end{aligned} \tag{4.6}$$

and

$$P_2(0) \in N_K^{b(2)}(x, P_1(0)). \tag{4.7}$$

Proof. We borrow some ideas from [11, proof of Theorem 2].

From the definition of the second order adjacent set, we deduce that, if $(v,h) \in \mathcal{A}_{\bar{u}}$, then $v(t,\omega) \in T_U^b(\bar{u}(t,\omega))$, a.e. $(t,\omega) \in [0,T] \times \Omega$, and for any $\varepsilon > 0$, there exists an $r(\varepsilon,t,\omega) \in \mathbb{R}^m$ such that

$$\bar{u}(t,\omega) + \varepsilon v(t,\omega) + \varepsilon^2 h(t,\omega) + r(\varepsilon,t,\omega) \in U,\; r(\varepsilon,t,\omega) = o(\varepsilon^2), \text{ a.e. } (t,\omega) \in [0,T] \times \Omega.$$

Furthermore, let $\ell(t,\omega) = |h(t,\omega)| + 1$, then for a.e. $(t,\omega) \in [0,T] \times \Omega$ there exists a $\rho(t,\omega) > 0$ such that

$$\begin{aligned} &dist(\bar{u}(t,\omega) + \varepsilon v(t,\omega), U) \\ &\le |\bar{u}(t,\omega) + \varepsilon v(t,\omega) - (\bar{u}(t,\omega) + \varepsilon v(t,\omega) + \varepsilon^2 h(t,\omega) + r(\varepsilon,t,\omega))| \\ &= |\varepsilon^2 h(t,\omega) + r(\varepsilon,t,\omega)| \le \varepsilon^2 \ell(t,\omega), \quad \forall\, \varepsilon \in [0, \rho(t,\omega)]. \end{aligned} \tag{4.8}$$

Motivated by the inequality (4.8), we introduce the following subset of $\mathcal{A}_{\bar{u}}$:

$$\mathcal{A}^*_{\bar{u}} = \Big\{(v,h) \in \mathcal{A}_{\bar{u}} \;\Big|\; \exists \text{ a } \rho_0 > 0 \text{ (independent of } (t,\omega)) \text{ such that}$$
$$dist(\bar{u}(t,\omega) + \varepsilon v(t,\omega), U) \le \varepsilon^2 \ell(t,\omega),\ \forall \varepsilon \in [0,\rho_0]\Big\}.$$

We fist prove that (4.6) and (4.7) hold for any $(v,h) \in \mathcal{A}^*_{\bar{u}}$ satisfying $v \in \Upsilon_{\bar{u}}$. Fix such a $(v,h) \in \mathcal{A}^*_{\bar{u}}$ and a corresponding $\rho_0 > 0$.

Using similar arguments as those in the proof of [17, Proposition 4.2], we now prove that $v \in T^b_{\mathcal{U}_{ad}}(\bar{u})$ and $h \in T^{b(2)}_{\mathcal{U}_{ad}}(\bar{u}, v)$.

Define

$$\alpha_\varepsilon(t,\omega) = dist(\bar{u}(t,\omega) + \varepsilon v(t,\omega), U).$$

The distance function being Lipschitz continuous, α_ε is a $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted process. Furthermore, since, $v(t,\omega) \in T^b_U(\bar{u}(t,\omega))$ a.s., we have $\alpha_\varepsilon(t,\omega)/\varepsilon \to 0$ a.e. $(t,\omega) \in [0,T] \times \Omega$ as $\varepsilon \to 0^+$.

On the other hand, U being a closed set in $\mathbb{R}^m$, for a.e. $(t,\omega) \in [0,T] \times \Omega$ there exists a $u_\varepsilon(t,\omega) \in U$ such that

$$\alpha_\varepsilon(t,\omega) = |u_\varepsilon(t,\omega) - \bar{u}(t,\omega) - \varepsilon v(t,\omega)| \le \varepsilon^2 \ell(t,\omega)\ \ \forall \varepsilon \in [0,\rho_0].$$

Using Lemma 2.2, we show that u_ε admits a $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted version. (Note that the metric projection mapping $(t,\omega) \rightsquigarrow \Pi_U(\bar{u}(t,\omega) + \varepsilon v(t,\omega))$ may not be $\mathcal{B}([0;T]) \otimes \mathcal{F}$-measurable, since $([0,T] \times \Omega, \mathcal{B}([0;T]) \otimes \mathcal{F}, dt \times dP)$ is not complete. Therefore, we can only obtain a measurable selection of $(t,\omega) \rightsquigarrow \Pi_U(\bar{u}(t,\omega) + \varepsilon v(t,\omega))$ on the completion of this product measure space and then modify this selection to be a $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable process.) To simplify the notation, we still denote this version by u_ε.

For $v_\varepsilon = (u_\varepsilon - \bar{u})/\varepsilon$, we have

$$|v_\varepsilon(t,\omega) - v(t,\omega)| = \left|\frac{u_\varepsilon(t,\omega) - \bar{u}(t,\omega)}{\varepsilon} - v(t,\omega)\right| = \left|\frac{\alpha_\varepsilon(t,\omega)}{\varepsilon}\right| \le \varepsilon \ell(t,\omega).$$

Since $(v,h) \in \mathcal{A}^*_{\bar{u}}$, it follows that $v_\varepsilon \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ and, by the dominated convergence theorem, $v_\varepsilon \to v$ in $L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ as $\varepsilon \to 0^+$. By the definition of v_ε, we get $\bar{u}(t,\omega) + \varepsilon v_\varepsilon(t,\omega) = u_\varepsilon(t,\omega) \in U$, a.e. $(t,\omega) \in [0,T] \times \Omega$. This proves that $v \in T^b_{\mathcal{U}_{ad}}(\bar{u})$.

Similarly, define

$$\gamma_\varepsilon(t,\omega) = dist(\bar{u}(t,\omega) + \varepsilon v(t,\omega) + \varepsilon^2 h(t,\omega), U).$$

Then, γ_ε is $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted, and, because $h(t,\omega) \in T^{b(2)}_U(\bar{u}(t,\omega), v(t,\omega))$, a.e. $(t,\omega) \in [0,T] \times \Omega$, $\gamma_\varepsilon(t,\omega)/\varepsilon^2 \to 0$ a.e. $(t,\omega) \in [0,T] \times \Omega$ as $\varepsilon \to 0^+$.

Choose a $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted processes $w_\varepsilon(t,\omega) \in U$, such that

$$\gamma_\varepsilon(t,\omega) = |w_\varepsilon(t,\omega) - \bar{u}(t,\omega) - \varepsilon v(t,\omega) - \varepsilon^2 h(t,\omega)|,\ \text{a.e. } (t,\omega) \in [0,T] \times \Omega$$

and define

$$h_\varepsilon = \frac{w_\varepsilon - \bar{u} - \varepsilon v}{\varepsilon^2}.$$

Then,

$$\begin{aligned}
|h_\varepsilon(t,\omega) - h(t,\omega)| &= \Big| \frac{w_\varepsilon - \bar{u}(t,\omega) - \varepsilon v(t,\omega)}{\varepsilon^2} - h(t,\omega) \Big| \\
&\leq \Big| \frac{u_\varepsilon - \bar{u}(t,\omega) - \varepsilon v(t,\omega) - \varepsilon^2 h(t,\omega)}{\varepsilon^2} \Big| \leq \frac{\alpha_\varepsilon(t,\omega)}{\varepsilon^2} + |h(t,\omega)| \\
&\leq \ell(t,\omega) + |h(t,\omega)|, \quad \text{a.e. } (t,\omega) \in [0,T] \times \Omega,
\end{aligned}$$

and hence $h_\varepsilon \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$. Moreover, by the definition of h_ε,

$$\bar{u}(t,\omega) + \varepsilon v(t,\omega) + \varepsilon^2 h_\varepsilon(t,\omega) = w_\varepsilon(t,\omega) \in U, \ \text{a.e. } (t,\omega) \in [0,T] \times \Omega,$$

and

$$|h_\varepsilon(t,\omega) - h(t,\omega)| = \Big| \frac{\gamma_\varepsilon(t,\omega)}{\varepsilon^2} \Big| \to 0, \qquad \text{a.e. } (t,\omega) \in [0,T] \times \Omega.$$

By the dominated convergence theorem, $h_\varepsilon \to h$ in $L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ as $\varepsilon \to 0^+$. This proves that $h \in T^{b(2)}_{\mathcal{U}_{ad}}(\bar{u}, v)$.

Let $\nu_0 \in T^b_K(\bar{x}_0) \cap \{P_1(0)\}^\perp$ and $\varpi_0 \in T^{b(2)}_K(\bar{x}_0, \nu_0)$.

Define $u^\varepsilon = \bar{u} + \varepsilon v + \varepsilon^2 h_\varepsilon$ and let x_0^ε, δx^ε and δu^ε be defined as above. Denote $\tilde{f}^\varepsilon_{xx}(t) := \int_0^1 (1-\theta) f_{xx}(t, \bar{x}(t) + \theta \delta x^\varepsilon(t), \bar{u}(t) + \theta \delta u^\varepsilon(t)) d\theta$. Mappings $\tilde{f}^\varepsilon_{xu}(t)$, $\tilde{f}^\varepsilon_{uu}(t)$ and $\tilde{g}^\varepsilon_{xx}(T)$ are defined in a similar way.

Expanding the cost functional J at $\bar{u}$, we get

$$\begin{aligned}
&\frac{J(u^\varepsilon) - J(\bar{u})}{\varepsilon^2} \\
&= \frac{1}{\varepsilon^2} \mathbb{E} \int_0^T \Big(\big\langle f_x(t), \delta x^\varepsilon(t) \big\rangle + \big\langle f_u(t), \delta u^\varepsilon(t) \big\rangle + \Big\langle \tilde{f}^\varepsilon_{xx}(t) \delta x^\varepsilon(t), \delta x^\varepsilon(t) \Big\rangle \\
&\quad + 2 \Big\langle \tilde{f}^\varepsilon_{xu}(t) \delta x^\varepsilon(t), \delta u^\varepsilon(t) \Big\rangle + \Big\langle \tilde{f}^\varepsilon_{uu}(t) \delta u^\varepsilon(t), \delta u^\varepsilon(t) \Big\rangle \Big) dt \\
&\quad + \frac{1}{\varepsilon^2} \mathbb{E} \Big(\big\langle g_x(\bar{x}(T)), \delta x^\varepsilon(T) \big\rangle + \big\langle \tilde{g}^\varepsilon_{xx}(\bar{x}(T)) \delta x^\varepsilon(T), \delta x^\varepsilon(T) \big\rangle \Big) \\
&= \mathbb{E} \int_0^T \Big[\frac{1}{\varepsilon} \langle f_x(t), y_1(t) \rangle + \frac{1}{2} \langle f_x(t), y_2(t) \rangle + \frac{1}{\varepsilon} \langle f_u(t), v(t) \rangle + \langle f_u(t), h(t) \rangle \\
&\quad + \frac{1}{2} \Big(\langle f_{xx}(t) y_1(t), y_1(t) \rangle + 2 \langle f_{xu}(t) y_1(t), v(t) \rangle + \langle f_{uu}(t) v(t), v(t) \rangle \Big) \Big] dt \\
&\quad + \mathbb{E} \Big(\frac{1}{\varepsilon} \langle g_x(\bar{x}(T)), y_1(T) \rangle + \frac{1}{2} \langle g_x(\bar{x}(T)), y_2(T) \rangle
\end{aligned}$$

$$+\frac{1}{2}\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle\Big) + \rho_2^\varepsilon,$$

where

$$\rho_2^\varepsilon = \mathbb{E}\int_0^T \Big(\big\langle f_x(t), r_2^\varepsilon(t)\big\rangle + \langle f_u(t), h_\varepsilon(t) - h(t)\rangle\Big)dt + \mathbb{E}\big\langle g_x(\bar{x}(T)), r_2^\varepsilon(T)\big\rangle$$
$$+\mathbb{E}\int_0^T \Big[\Big(\Big\langle \tilde{f}_{xx}^\varepsilon(t)\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta x^\varepsilon(t)}{\varepsilon}\Big\rangle - \frac{1}{2}\langle f_{xx}(t)y_1(t), y_1(t)\rangle\Big)$$
$$+\Big(2\Big\langle \tilde{f}_{xu}^\varepsilon(t)\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta u^\varepsilon(t)}{\varepsilon}\Big\rangle - \langle f_{xu}(t)y_1(t), v(t)\rangle\Big)$$
$$+\Big(\Big\langle \tilde{f}_{uu}^\varepsilon(t)\frac{\delta u^\varepsilon(t)}{\varepsilon}, \frac{\delta u^\varepsilon(t)}{\varepsilon}\Big\rangle - \frac{1}{2}\langle f_{uu}(t)v(t), v(t)\rangle\Big)\Big]dt$$
$$+\mathbb{E}\Big(\Big\langle \tilde{g}_{xx}^\varepsilon(\bar{x}(T))\frac{\delta x^\varepsilon(T)}{\varepsilon}, \frac{\delta x^\varepsilon(T)}{\varepsilon}\Big\rangle - \frac{1}{2}\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle\Big).$$

In the same way as in the proof of Lemma 4.1, we find that $\lim_{\varepsilon\to 0^+}\rho_2^\varepsilon = 0$. On the other hand, by (3.13) and, recalling that $v \in \Upsilon_{\bar{u}}$, $\nu_0 \in \{P_1(0)\}^\perp$, we have

$$\frac{1}{\varepsilon}\mathbb{E}\int_0^T \Big(\langle f_x(t), y_1(t)\rangle + \langle f_u(t), v(t)\rangle\Big)dt + \frac{1}{\varepsilon}\mathbb{E}\langle g_x(\bar{x}(T)), y_1(T)\rangle$$
$$= -\frac{1}{\varepsilon}\langle P_1(0), \nu_0\rangle - \frac{1}{\varepsilon}\mathbb{E}\int_0^T \langle H_u(t), v(t)\rangle\, dt = 0.$$

Therefore,

$$\begin{aligned}0 &\le \lim_{\varepsilon\to 0^+}\frac{J(u^\varepsilon(\cdot)) - J(\bar{u}(\cdot))}{\varepsilon^2}\\
&= \mathbb{E}\int_0^T \Big[\frac{1}{2}\langle f_x(t), y_2(t)\rangle + \langle f_u(t), h(t)\rangle\\
&\quad + \frac{1}{2}\Big(\langle f_{xx}(t)y_1(t), y_1(t)\rangle + 2\langle f_{xu}(t)y_1(t), v(t)\rangle + \langle f_{uu}(t)v(t), v(t)\rangle\Big)\Big]dt\\
&\quad + \frac{1}{2}\mathbb{E}\Big(\langle g_x(\bar{x}(T)), y_2(T)\rangle + \langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle\Big).\end{aligned} \tag{4.9}$$

By Itô's formula,

$$\mathbb{E}\ \langle g_x(\bar{x}(T)), y_2(T)\rangle = -\mathbb{E}\ \langle P_1(T), y_2(T)\rangle \tag{4.10}$$

$$
\begin{aligned}
&= -2\langle P_1(0), \varpi_0\rangle - \mathbb{E}\int_0^T \Big(2\langle P_1(t), b_u(t)h(t)\rangle + \Big\langle P_1(t), y_1(t)^\top b_{xx}(t)y_1(t)\Big\rangle \\
&\quad + 2\Big\langle P_1(t), v(t)^\top b_{xu}(t)y_1(t)\Big\rangle + \Big\langle P_1(t), v(t)^\top b_{uu}(t)v(t)\Big\rangle + 2\langle Q_1(t), \sigma_u(t)h(t)\rangle \\
&\quad + \Big\langle Q_1(t), y_1(t)^\top \sigma_{xx}(t)y_1(t)\Big\rangle + 2\Big\langle Q_1(t), v(t)^\top \sigma_{xu}(t)y_1(t)\Big\rangle \\
&\quad + \Big\langle Q_1(t), v(t)^\top \sigma_{uu}(t)v(t)\Big\rangle + \langle f_x(t), y_2(t)\rangle\Big)dt,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\,\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle = -\mathbb{E}\,\langle P_2(T)y_1(T), y_1(T)\rangle \qquad (4.11)\\
&= -\langle P_2(0)\nu_0, \nu_0\rangle - \mathbb{E}\int_0^T \Big(2\langle P_2(t)y_1(t), b_u(t)v(t)\rangle + 2\langle P_2(t)\sigma_x(t)y_1(t), \sigma_u(t)v(t)\rangle \\
&\quad + \langle P_2(t)\sigma_u(t)v(t), \sigma_u(t)v(t)\rangle + 2\langle Q_2(t)\sigma_u(t)v(s), y_1(t)\rangle - \langle H_{xx}(t)y_1(t), y_1(t)\rangle\Big)dt.
\end{aligned}
$$

Substituting (4.10) and (4.11) into (4.9) yields

$$
\begin{aligned}
0 &\geq \langle P_1(0), \varpi_0\rangle + \frac{1}{2}\langle P_2(0)\nu_0, \nu_0\rangle \\
&\quad + \mathbb{E}\int_0^T \Big[\Big(\langle P_1(t), b_u(t)h(t)\rangle + \langle Q_1(t), \sigma_u(t)h(t)\rangle - \langle f_u(t), h(t)\rangle\Big) \\
&\qquad + \frac{1}{2}\Big(\Big\langle P_1(t), v(t)^\top b_{uu}(t)v(t)\Big\rangle + \Big\langle Q_1(t), v(t)^\top \sigma_{uu}(t)v(t)\Big\rangle - \langle f_{uu}(t)v(t), v(t)\rangle\Big) \\
&\qquad + \frac{1}{2}\langle P_2(t)\sigma_u(t)v(t), \sigma_u(t)v(t)\rangle + \Big(\Big\langle P_1(t), v(t)^\top b_{xu}(t)y_1(t)\Big\rangle \\
&\qquad + \Big\langle Q_1(t), v(t)^\top \sigma_{xu}(t)y_1(t)\Big\rangle - \langle f_{xu}(t)y_1(t), v(t)\rangle + \Big\langle b_u(t)^\top P_2(t)y_1(t), v(t)\Big\rangle \\
&\qquad + \Big\langle \sigma_u(t)^\top P_2(t)\sigma_x(t)y_1(t), v(t)\Big\rangle + \Big\langle \sigma_u(t)^\top Q_2(t)y_1(t), v(t)\Big\rangle\Big)\Big]dt \\
&= \langle P_1(0), \varpi_0\rangle + \frac{1}{2}\langle P_2(0)\nu_0, \nu_0\rangle + \frac{1}{2}\mathbb{E}\int_0^T \Big(2\langle H_u(t), h(t)\rangle \\
&\quad + \langle H_{uu}(t)v(t), v(t)\rangle + \langle P_2(t)\sigma_u(t)v(t), \sigma_u(t)v(t)\rangle + 2\langle \mathbb{S}(t)y_1(t), v(t)\rangle\Big)dt.
\end{aligned}
$$

Then, letting $v(\cdot) = h(\cdot) = 0$ we obtain (4.7) and letting $\nu_0 = \varpi_0 = 0$, we obtain (4.6), for any $(v, h) \in \mathcal{A}_{\bar{u}}^*$ satisfying $v \in \Upsilon_{\bar{u}}$.

To prove (4.6) for any $(v, h) \in \mathcal{A}_{\bar{u}}$ satisfying $v \in \Upsilon_{\bar{u}}$, define

$$
E_i := \{(t, \omega) \in [0, T] \times \Omega \mid dist(\bar{u}(t, \omega) + \varepsilon v(t, \omega), U) \leq \varepsilon^2 \ell(t, \omega),\ \forall\, \varepsilon \in (0, \frac{1}{i}]\}.
$$

It can be proved that E_i is $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable, the family $\{E_i\}_{i=1}^{\infty}$ is nondecreasing and $\bigcup_{i=1}^{\infty} E_i$ is of full measure in $[0,T]\times\Omega$. For any $i\in\mathbb{N}$ and $(v,h)\in\mathcal{A}_{\bar u}$ satisfying $v\in\Upsilon_{\bar u}$, define

$$v^i(t,\omega):=\begin{cases} v(t,\omega), & (t,\omega)\in E_i,\\ 0, & \text{otherwise},\end{cases}\qquad h^i(t,\omega):=\begin{cases} h(t,\omega), & (t,\omega)\in E_i,\\ 0, & \text{otherwise}.\end{cases}$$

Then, $(v^i,h^i)\in\mathcal{A}_{\bar u}^*$ and $v^i\in\Upsilon_{\bar u}$. Hence,

$$\begin{aligned}\mathbb{E}\int_0^T \Big(2\Big\langle H_u(t),h^i(t)\Big\rangle+\Big\langle H_{uu}(t)v^i(t),v^i(t)\Big\rangle\\ +\Big\langle P_2(t)\sigma_u(t)v^i(t),\sigma_u(t)v^i(t)\Big\rangle+2\Big\langle \mathbb{S}(t)y_1^i(t),v^i(t)\Big\rangle\Big)dt\le 0,\end{aligned}\tag{4.12}$$

where y_1^i is the solution to the first order variational equation (3.3) with v replaced by v^i. Since $v^i\to v$, $h^i\to h$ in $L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$ as $i\to\infty$, we have $y_1^i\to y_1$ in $L^4_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$. Passing to the limit in inequality (4.12), we finally obtain (4.6). This completes the proof of Theorem 4.1. □

In what follows, we shall give a consequence of Theorem 4.1 for the case when U is represented by finitely many mixed constraints, i.e.,

$$U=\Big\{u\in\mathbb{R}^m\,\Big|\,\varphi_i(u)=0,\ \forall\, i=1,\dots,p,\ \psi_j(u)\le 0,\ \forall\, j=1,\dots,r\Big\},$$

where $\varphi_1,\dots,\varphi_p:\mathbb{R}^m\to\mathbb{R}$ and $\psi_1,\dots,\psi_r:\mathbb{R}^m\to\mathbb{R}$ (for some $p,r\in\mathbb{N}$) are twice continuously differentiable functions and for any $u\in U$,

$$\{\nabla\varphi_1(u),\cdots,\nabla\varphi_p(u)\}\bigcup\{\nabla\psi_j(u)\,|\,j\in I(u)\}\ \text{are linearly independent.}\tag{4.13}$$

Moreover, there exist two constants $L\ge 0$ and $\rho>0$ such that for every $u\in U$,

$$\begin{aligned}&|\varphi_i''(u)|\le L,\quad i=1,\dots,p,\\ &|\psi_j''(u)|\le L,\quad j\in I(u),\\ &\rho B_{Im(\Gamma_u)}\subset\Gamma_u B_{\mathbb{R}^{p+k}},\end{aligned}\tag{4.14}$$

where $I(u)$ is the set of all active indices at u, $\Gamma_u:=(\nabla\varphi_1(u),\dots,\nabla\varphi_p(u),\nabla\psi_{i_1}(u),\dots,\psi_{i_k}(u))$ with $i_1,\dots,i_k\in I(u)$ being all active indices for some $k\le r$, and $B_{Im(\Gamma_u)}$ and $B_{\mathbb{R}^{p+k}}$ are respectively the unit balls in the image space of Γ_u and $\mathbb{R}^{p+k}$.

We observe that (4.13) implies (4.14) with a ρ depending on u. In the above we required ρ to be independent of u to obtain the following result.

Corollary 4.1. *Let U be as above, (C2)–(C3) hold and $(\bar x,\bar u,\bar x_0)$ be a local minimizer for the problem (1.3) with $\bar u\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$. Then there exist $\mu_i(\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}))$, $i=1,\dots,p$ and $\lambda_j(\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}_+))$, $j=1,\dots,r$ such that for any $v(\cdot)\in\Upsilon_{\bar u}\cap$*

$L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ *satisfying* $v(t,\omega) \in T^b_U(\bar{u}(t,\omega))$, *a.e.* $(t,\omega) \in [0,T] \times \Omega$ *and the corresponding solution* y_1 *of equation (3.3) we have*

$$\mathbb{E}\int_0^T \Big(\langle H_{uu}(t)v(t), v(t)\rangle + \langle P_2(t)\sigma_u(t)v(t), \sigma_u(t)v(t)\rangle + 2\,\langle \mathbb{S}(t)y_1(t), v(t)\rangle$$

$$- \sum_{i=1}^{p} \mu_i(t)\big\langle \varphi_i''(\bar{u}(t))v(t), v(t)\big\rangle - \sum_{j\in I_v(\bar{u}(t))} \lambda_j(t)\Big\langle \psi_j''(\bar{u}(t))v(t), v(t)\Big\rangle\Big)dt \le 0, \quad (4.15)$$

where

$$I_v(\bar{u}(t,\omega)) = \{ j \in I(\bar{u}(t,\omega)) \mid \big\langle \nabla\psi_j(\bar{u}(t,\omega)), v(t,\omega)\big\rangle = 0\}.$$

Proof. The proof of this result is similar to that of [11, Theorem 3]. Obviously, condition (4.13) implies the Mangasarian–Fromowitz constraint qualification. By Example 2.1, for any (t,ω),

$$N^b_U(\bar{u}(t,\omega)) = \sum_{i=1}^{p} \mathbb{R}\nabla\varphi_i(\bar{u}(t,\omega)) + \sum_{j\in I(\bar{u}(t,\omega))} \mathbb{R}_+\nabla\psi_j(\bar{u}(t,\omega)).$$

Then, by the first order condition (3.15), we have

$$H_u(t,\omega) \in \sum_{i=1}^{p} \mathbb{R}\nabla\varphi_i(\bar{u}(t,\omega)) + \sum_{j\in I(\bar{u}(t,\omega))} \mathbb{R}_+\nabla\psi_j(\bar{u}(t,\omega)), \text{ a.e. } (t,\omega) \in [0,T]\times\Omega.$$

Define

$$\Gamma(t,\omega) = \{(\mu_1,\dots,\mu_p,\lambda_1,\dots,\lambda_r) \in \mathbb{R}^{p+r} \mid \lambda_j \ge 0,\ j=1,\dots,r,\ \lambda_j\psi_j(\bar{u}(t,\omega)) = 0\},$$

and

$$G(t,\omega,(\mu_1,\dots,\mu_p,\lambda_1,\dots,\lambda_r)) = \sum_{i=1}^{p} \mu_i\nabla\varphi_i(\bar{u}(t,\omega)) + \sum_{j=1}^{r} \lambda_j\nabla\psi_j(\bar{u}(t,\omega)).$$

By Filippov's theorem (see [2, Theorem 8.2.10]), there exists a $\mathscr{G}^*$-measurable selection

$$\gamma^*(t,\omega) = (\mu_1^*(t,\omega),\dots,\mu_p^*(t,\omega),\lambda_1^*(t,\omega),\dots,\lambda_r^*(t,\omega)) \in \Gamma(t,\omega), \text{ a.e. } (t,\omega)\in[0,T]\times\Omega$$

such that

$$H_u(t,\omega) = \sum_{i=1}^{p} \mu_i^*(t,\omega)\nabla\varphi_i(\bar{u}(t,\omega)) + \sum_{j\in I(\bar{u}(t,\omega))} \lambda_j^*(t,\omega)\nabla\psi_j(\bar{u}(t,\omega)),$$
$$\text{a.e. } (t,\omega)\in[0,T]\times\Omega$$

where $\mathscr{G}^*$ is the completion of $\mathscr{G}$ and $\mathscr{G}$ is defined by (2.1). By assumption (4.13) the process $\gamma^*(\cdot)$ is uniquely determined (up to a set of measure zero). Since $\mathbb{R}^m$ is separable, there exists a $\mathscr{G}$-measurable modification of $\gamma^*(\cdot)$:

$$\gamma(\cdot) = (\mu_1(\cdot), ..., \mu_p(\cdot), \lambda_1(\cdot), ..., \lambda_r(\cdot)).$$

By Lemma 2.3, $\gamma(\cdot)$ is $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted and

$$\begin{aligned} H_u(t,\omega) = \sum_{i=1}^{p} \mu_i(t,\omega)\nabla\varphi_i(\bar{u}(t,\omega)) + \sum_{j\in I(\bar{u}(t,\omega))} \lambda_j(t,\omega)\nabla\psi_j(\bar{u}(t,\omega)), \\ \text{a.e. } (t,\omega) \in [0,T]\times\Omega. \end{aligned} \tag{4.16}$$

By [9, Theorem 2.1] and assumption (4.14), for a.e. $(t,\omega) \in [0,T]\times\Omega$

$$|\mu_i(t,\omega)| \le \frac{1}{\rho}|H_u(t,\omega)|,\ \forall i = 1, ..., p,\ \lambda_j(t,\omega) \le \frac{1}{\rho}|H_u(t,\omega)|,\ \forall j \in I(\bar{u}(t,\omega)). \tag{4.17}$$

On the other hand, when $j \notin I(\bar{u}(t,\omega))$, $\lambda_j(t,\omega) = 0$ and therefore also $\lambda_j(t,\omega) \le \frac{1}{\rho}|H_u(t,\omega)|$. Since $H_u(\cdot) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$, we deduce that $\mu_i(\cdot) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}))$, $i = 1, ..., p$, and, $\lambda_j(\cdot) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}_+))$, $j = 1, ..., r$.

Let $v(\cdot) \in \Upsilon_{\bar{u}} \cap L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ satisfy $v(t,\omega) \in T^b_U(\bar{u}(t,\omega))$, a.e. $(t,\omega) \in [0,T]\times\Omega$. Then

$$\langle H_u(t,\omega), v(t,\omega)\rangle = 0, \quad \text{a.e. } (t,\omega) \in [0,T]\times\Omega. \tag{4.18}$$

Combining (4.18) with (4.16), one has, for a.e. $(t,\omega) \in [0,T]\times\Omega$,

$$\sum_{j\in I(\bar{u}(t,\omega))} \lambda_j(t,\omega)\big\langle\nabla\psi_j(\bar{u}(t,\omega)), v(t,\omega)\big\rangle = 0.$$

Therefore, for a.e. $(t,\omega) \in [0,T]\times\Omega$ and for any $j \notin I_v(\bar{u}(t,\omega))$, $\lambda_j(t,\omega) = 0$. Consequently,

$$\begin{aligned} H_u(t,\omega) = \sum_{i=1}^{p} \mu_i(t,\omega)\nabla\varphi_i(\bar{u}(t,\omega)) + \sum_{j\in I_v(\bar{u}(t,\omega))} \lambda_j(t,\omega)\nabla\psi_j(\bar{u}(t,\omega)), \\ \text{a.e. } (t,\omega) \in [0,T]\times\Omega. \end{aligned} \tag{4.19}$$

On the other hand, for any $(t,\omega) \in [0,T]\times\Omega$, by Example 2.1,

$$\begin{aligned} &\emptyset \ne T^{b(2)}_U(\bar{u}(t,\omega), v(t,\omega)) \\ &= \Big\{ h \in \mathbb{R}^m \,\Big|\, \langle\nabla\varphi_i(\bar{u}(t,\omega)), h\rangle + \frac{1}{2}\langle\varphi_i''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\rangle = 0,\ \forall i = 1, \cdots, p, \\ &\quad \text{and } \langle\nabla\psi_j(\bar{u}(t,\omega)), h\rangle + \frac{1}{2}\langle\psi_j''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\rangle \le 0,\ \forall j \in I_v(\bar{u}(t,\omega)) \Big\}. \end{aligned} \tag{4.20}$$

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

It follows that, for any $h \in T_U^{b(2)}(\bar{u}(t,\omega), v(t,\omega))$ and a.e. $(t,\omega) \in [0,T] \times \Omega$,

$$\begin{aligned}\langle H_u(t,\omega), h\rangle \le & -\frac{1}{2}\sum_{i=1}^{p}\mu_i(t,\omega)\left\langle\varphi_i''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle \\ & -\frac{1}{2}\sum_{j\in I_v(\bar{u}(t,\omega))}\lambda_j(t,\omega)\left\langle\psi_j''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle,\end{aligned} \tag{4.21}$$

which implies that

$$\sup_{h\in T_U^{b(2)}(\bar{u}(t,\omega),v(t,\omega))}\langle H_u(t,\omega), h\rangle < \infty, \text{ a.e. } (t,\omega) \in [0,T]\times\Omega.$$

By (4.20), $T_U^{b(2)}(\bar{u}(t,\omega), v(t,\omega))$ is a polyhedral set, cf. [29, p. 43]. By [29, Corollary 3.53] the supremum in the above is attained.

By [2, Theorems 8.2.11 and 8.2.9] (making a completion argumentation if necessary), there exists a $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted process $\tilde{h}(\cdot)$ such that $\tilde{h}(t,\omega) \in T_U^{b(2)}(\bar{u}(t,\omega), v(t,\omega))$ a.e. in $[0,T]\times\Omega$ and

$$\left\langle H_u(t,\omega), \tilde{h}(t,\omega)\right\rangle = \sup_{h\in T_U^{b(2)}(\bar{u}(t,\omega),v(t,\omega))}\langle H_u(t,\omega), h\rangle, \text{ a.e. } (t,\omega)\in[0,T]\times\Omega.$$

Then, for a.e. $(t,\omega) \in [0,T]\times\Omega$

$$\begin{aligned}&\mu_i(t,\omega)\left\langle\nabla\varphi_i(\bar{u}(t,\omega)), \tilde{h}(t,\omega)\right\rangle = -\frac{\mu_i(t,\omega)}{2}\left\langle\varphi_i''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle, \\ &\forall\, i = 1, \dots, p,\end{aligned} \tag{4.22}$$

and,

$$\lambda_j(t,\omega)\left\langle\nabla\psi_j(\bar{u}(t,\omega)), \tilde{h}(t,\omega)\right\rangle \le -\frac{\lambda_j(t,\omega)}{2}\left\langle\psi_j''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle, \ \forall\, j \in I_v(\bar{u}(t,\omega)).$$

Applying the same argument as at the end of Example 2.1 we show, using (4.19), that

$$\begin{aligned}&\lambda_j(t,\omega)\left\langle\nabla\psi_j(\bar{u}(t,\omega)), \tilde{h}(t,\omega)\right\rangle = -\frac{\lambda_j(t,\omega)}{2}\left\langle\psi_j''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle, \\ &\forall\, j \in I_v(\bar{u}(t,\omega)).\end{aligned} \tag{4.23}$$

Combining (4.19), (4.22) with (4.23), one obtains that, for a.e. $(t,\omega)\in[0,T]\times\Omega$,

$$\begin{aligned}\left\langle H_u(t,\omega), \tilde{h}(t,\omega)\right\rangle = & -\frac{1}{2}\sum_{i=1}^{p}\mu_i(t,\omega)\left\langle\varphi_i''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle \\ & -\frac{1}{2}\sum_{j\in I_v(\bar{u}(t,\omega))}\lambda_j(t,\omega)\left\langle\psi_j''(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle.\end{aligned} \tag{4.24}$$

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Now, for any $i \in \mathbb{N}$, define

$$v^i(t,\omega) := \begin{cases} v(t,\omega), & \text{if } |\tilde{h}(t,\omega)| \le i, \\ 0, & \text{otherwise}, \end{cases} \qquad h^i(t,\omega) := \begin{cases} \tilde{h}(t,\omega), & \text{if } |\tilde{h}(t,\omega)| \le i, \\ 0, & \text{otherwise}, \end{cases}$$

we have $(v^i(\cdot), h^i(\cdot)) \in \mathcal{A}_{\bar{u}}$ and $v^i(\cdot) \in \Upsilon_{\bar{u}}$. Let y_1^i be the solution to the first order variational equation (3.3) corresponding to $v^i(\cdot)$, then by (4.24) and condition (4.6), we obtain that

$$\begin{aligned} &\mathbb{E}\int_0^T \Big(\Big\langle H_{uu}(t)v^i(t), v^i(t) \Big\rangle + \Big\langle P_2(t)\sigma_u(t)v^i(t), \sigma_u(t)v^i(t) \Big\rangle + 2\Big\langle \mathbb{S}(t)y_1^i(t), v^i(t) \Big\rangle \\ &\quad - \sum_{i=1}^{p} \mu_i(t) \Big\langle \varphi_i''(\bar{u}(t))v^i(t), v^i(t) \Big\rangle - \sum_{j \in I_v(\bar{u}(t))} \lambda_j(t) \Big\langle \psi_j''(\bar{u}(t))v^i(t), v^i(t) \Big\rangle \Big) dt \\ &\le 0. \end{aligned} \tag{4.25}$$

Passing to the limit in inequality (4.25), we finally obtain condition (4.15). This completes the proof of Corollary 4.1. □

In [6], in the special case of $K = \{x_0\}$, the authors obtained the following integral-type first and second order necessary conditions for stochastic optimal controls:

Theorem 4.2. *Let (C2)–(C3) hold. If U is closed and convex and $\bar{u}$ is an optimal control, then*

$$\mathbb{E}\int_0^T \langle H_u(t), v(t) \rangle \, dt \le 0, \qquad \forall\, v \in cl_{2,2}\big(\mathcal{R}_{\mathcal{U}_{ad}}(\bar{u}) \cap L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))\big). \tag{4.26}$$

Furthermore, for any $v(\cdot) \in cl_{4,4}\big(\mathcal{R}_{\mathcal{U}_{ad}}(\bar{u}) \cap L^\infty([0,T] \times \Omega; \mathbb{R}^m) \cap \Upsilon_{\bar{u}}\big)$ the following second order necessary condition holds:

$$\begin{aligned} &\mathbb{E}\int_0^T \Big(\langle H_{xx}(t)y_1(t), y_1(t) \rangle + 2\langle H_{xu}(t)y_1(t), v(t) \rangle \\ &\qquad + \langle H_{uu}(t)v(t), v(t) \rangle \Big) dt + \mathbb{E}\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T) \rangle \le 0, \end{aligned} \tag{4.27}$$

where,

$$\mathcal{R}_{\mathcal{U}_{ad}}(\bar{u}) := \big\{ \alpha u - \alpha \bar{u} \;\big|\; u \in \mathcal{U}_{ad}, \alpha \ge 0 \big\},$$

and $cl_{2,2}(A)$ and $cl_{4,4}(A)$ are respectively the closures of a set A under the norms $\|\cdot\|_{2,2}$ and $\|\cdot\|_{4,4}$.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Remark 4.1. There are three main differences between (4.6) and (4.27): First, the control region is allowed to be nonconvex in (4.6). Second, the solutions to two adjoint equations (3.6) and (4.3) are used in (4.6), and consequently, the second order term involving y_1 (the solution to the first order variational equation (3.3)) is absent in this condition. Third, the condition (4.6) contains the second order adjacent vector h, while in (4.27) it is equal to zero, cf. Remark 2.1. Our condition (4.6) is more effective in distinguishing optimal controls from other admissible controls than (4.27), even if the diffusion term $\sigma = 0$, see [17]. See also the examples (especially Example 4.2) that we shall give below.

Example 4.1. Let U be equal to the intersection of two closed balls in $\mathbb{R}^2$ of radii 1 and centers at respectively $(1,0)$ and $(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $T=1$, $A \in \mathbb{R}^{2\times 2}$, $F=(F^1, F^2): \mathbb{R}^2 \to \mathbb{R}_+ \times \mathbb{R}$ be a given function satisfying $F(0)=0$, $F_x(0)=0$, $F_{xx}(0)=0$, and for some $L>0$,

$$|F_x(x)| + |F_{xx}(x)| \le L, \quad \forall x \in \mathbb{R}^2.$$

Consider the stochastic control system

$$\begin{cases} dx(t) = \big[F(x(t)) + u(t)\big]dt + Au(t)dW(t), & t \in [0,1], \\ x(0) = 0, \end{cases}$$

with the cost functional

$$J(u(\cdot)) = \mathbb{E}\big[x_1(1) - \cos(x_2(1))^2\big].$$

For this optimal control problem, the Hamiltonian is defined as

$$H(t,x,u,p,q,\omega) := \langle p, F(x)+u\rangle + \langle q, Au\rangle\,,$$

where $(t,x,u,p,q,\omega) \in [0,1]\times\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\times\Omega$.

Define $\bar{u}(t) \equiv (0,0)$. Then, the corresponding state $\bar{x}(t) \equiv (0,0)$. Since $F^1(x) \ge 0$ for any $x \in \mathbb{R}^2$ and $U \subset \mathbb{R}_+ \times \mathbb{R}$, we deduce that $\mathbb{E}(x_1(1)) \ge 0$ for any solution $x=(x_1,x_2)$ of the above stochastic system. Therefore $\bar{u}$ is the global minimizer. Furthermore, the first and the second order adjoint equations are

$$\begin{cases} dP_1(t) = Q_1(t)dW(t), & t \in [0,1], \\ P_1(1) = (-1,0) \end{cases} \tag{4.28}$$

and

$$\begin{cases} dP_2(t) = Q_2(t)dW(t), & t \in [0,1], \\ P_2(1) = 0. \end{cases} \tag{4.29}$$

It is easy to see that the solutions to equations (4.28) and (4.29) are $P_1(t) \equiv (-1,0)$, $Q_1(t) \equiv 0$ and $(P_2(t), Q_2(t)) \equiv (0,0)$, respectively. Then,

$$H_u(t) = P_1(t) + A^\top Q_1(t) \equiv (-1,0),\ H_{uu}(t) + \sigma_u^\top(t)P_2(t)\sigma_u(t) \equiv 0,\ \text{and } \mathbb{S}(t) \equiv 0.$$

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

By the definition of U, $T^b_U((0,0))$ is the closed convex cone generated by $\{(0,1),(1,1)\}$. Moreover $(\frac{1}{2},0) \in T^{b(2)}_U((0,0),(0,1))$.

Then the first order necessary condition

$$\langle H_u(t,\omega), v\rangle \le 0, \quad \forall\, v \in T^b_U((0,0))$$

(which corresponds to the first condition in (3.15)) is satisfied and

$$H_{xx}(t) \equiv 0, \quad H_{xu}(t) \equiv 0, \quad H_{uu}(t) \equiv 0, \quad \text{and} \quad g_{xx}(\bar{x}(1)) \equiv 0.$$

Therefore, the second order necessary condition (4.27) is satisfied trivially in this case and does not contain any additional information with respect to the first order necessary condition (4.26).

Comparatively, our second order necessary condition (4.6) provides more information about the control $\bar{u}$. For example, let $\tilde{v}(t) \equiv (0,1)$ and $\tilde{h}(t) \equiv (\frac{1}{2},0)$. Obviously $\tilde{v} \in \Upsilon_{\bar{u}}$, $(\tilde{v},\tilde{h}) \in \mathcal{A}_{\bar{u}}$, and condition (4.6) becomes

$$2\mathbb{E}\int_0^1 \Big\langle H_u(t), \tilde{h}(t)\Big\rangle dt = -1 \le 0.$$

Noting that $(\frac{1}{2},0) \notin T^b_U((0,0))$, the last inequality is different from the first order necessary condition (3.7) and from the second order necessary condition (4.27).

Example 4.2. Let $n = m = 2$, $T = 1$, and

$$U = \{(u_1,u_2) \in \mathbb{R}^2 \mid |u_1+1|^2 + |u_2|^2 = 1\} \cup \{(u_1,u_2) \in \mathbb{R}^2 \mid |u_1-1|^2 + |u_2|^2 = 1\}.$$

Clearly, this U is neither a finite set nor convex in $\mathbb{R}^2$. One can easily check that

$$T^b_U((0,0)) = \{0\} \times \mathbb{R}, \qquad T^{b(2)}_U((0,0),(0,1)) \ni (\frac{1}{2},0).$$

Consider the control system

$$\begin{cases} dx_1(t) = (x_2(t) - \frac{1}{2})dt + dW(t), & t \in [0,1],\\ dx_2(t) = u_1(t)dt + |u_2(t)|^4 dW(t), & t \in [0,1],\\ x_1(0) = 0, x_2(0) = 0 \end{cases} \tag{4.30}$$

with the cost functional

$$J(u) = \mathbb{E}\Big[\frac{1}{2}|x_1(1) - W(1)|^2 + \int_0^1 |u_2(t)|^4 dt\Big]. \tag{4.31}$$

Obviously, the only difference between (3.17) and (4.30) is that the coefficient "$u_2(t)$" in the first system is replaced by "$|u_2(t)|^4$" in the second one and, since U is a bounded set, the assumptions (C2)–(C3) are fulfilled.

The Hamiltonian of this optimal control problem is given by

$$H(t,(x_1,x_2),(u_1,u_2),(p_1^1,p_1^2),(q_1^1,q_1^2),\omega)=p_1^1(x_2-\frac{1}{2})+p_1^2u_1+q_1^1+q_1^2|u_2|^4-|u_2|^4,$$

for all $(t,(x_1,x_2),(u_1,u_2),(p_1^1,p_1^2),(q_1^1,q_1^2),\omega)\in[0,1]\times\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\times\Omega$. In what follows, we show that the admissible control $(u_1(t),u_2(t))\equiv(0,0)$ is not locally optimal.

The corresponding solution to the control system (4.30) is still given by (3.20), and the first order adjoint equation is the same as in (3.21). Therefore $(P_1^1(t),Q_1^1(t))$ and $(P_1^2(t),Q_1^2(t))$ are as in (3.22).

For the present problem,

$$H_u(t)=(P_1^2(t),4Q_1^2(u_2(t))^3-4(u_2(t))^3)=(\frac{1-t}{2},0). \tag{4.32}$$

Hence, the first order condition in (3.15),

$$\langle H_u(t),v\rangle=P_1^2(t)v_1+4(Q_1^2-1)(u_2(t))^3v_2=0,\quad \forall\, v=(v_1,v_2)\in T_U^b((0,0))$$

is trivially satisfied, and therefore we need to check the second order condition (4.6). For this, we observe that

$$H_{uu}(t)=\begin{bmatrix}0&0\\0&0\end{bmatrix},\quad b_x(t)=\begin{bmatrix}0&1\\0&0\end{bmatrix},$$
$$b_u(t)=\begin{bmatrix}0&0\\1&0\end{bmatrix},\quad \sigma_x(t)=\sigma_u(t)=\begin{bmatrix}0&0\\0&0\end{bmatrix}. \tag{4.33}$$

We now choose a direction $v=(v_1,v_2)=(0,1)$ and $\nu_0=(0,0)$. Then, the first order variational equation (3.3) becomes

$$\begin{cases}\frac{dy_1(t)}{dt}=b_x(t)y_1(t), & t\in[0,1],\\ y_1(0)=(0,0),\end{cases} \tag{4.34}$$

and hence $y_1(t)\equiv(0,0)$. This, combined with (4.33), shows that the second condition in (4.6) is specified as

$$\mathbb{E}\int_0^1\langle H_u(t),h\rangle\,dt\le 0,\qquad \forall\, h\in T_U^{b(2)}((0,0),(0,1)). \tag{4.35}$$

We now choose $h=(\frac{1}{2},0)$ in (4.35). By (4.32), we obtain that

$$\mathbb{E}\int_0^1\langle H_u(t),h\rangle\,dt=\frac{1}{8}>0,$$

which is a contradiction. Therefore, $(u_1(t),u_2(t))\equiv(0,0)$ is not locally optimal.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

4.2. Pointwise second order necessary conditions

In this subsection, under some further assumptions, we shall deduce from the integral-type second order necessary condition (4.6) a pointwise one. First, we introduce the following notion.

Definition 4.1. We call $\tilde{u} \in \mathcal{U}_{ad}$ partially singular in the classical sense if $\tilde{u}$ satisfies

$$\begin{cases} \widetilde{H}_u(t) = 0, \quad \text{a.e. } t \in [0, T], \text{ a.s.,} \\ \big\langle \big(\widetilde{H}_{uu}(t) + \tilde{\sigma}_u(t)^\top \widetilde{P}_2(t) \tilde{\sigma}_u(t)\big) v, v \big\rangle = 0, \quad \forall\, v \in T_U^b(\tilde{u}(t)), \text{ a.e. } t \in [0, T], \text{ a.s.} \end{cases} \tag{4.36}$$

where $\tilde{x}$ is the state corresponding to $\tilde{u}$, $\widetilde{H}_u(t) = H_u(t, \tilde{x}(t), \tilde{u}(t), \widetilde{P}_1(t), \widetilde{Q}_1(t))$, and similarly for $\widetilde{H}_{uu}(t)$ and $\tilde{\sigma}_u(t)$. $(\widetilde{P}_1, \widetilde{Q}_1)$ and $(\widetilde{P}_2, \widetilde{Q}_2)$ are the adjoint processes given respectively by (3.6) and (4.3) with $(\bar{x}, \bar{u}, \bar{x}_0)$ replaced by $(\tilde{x}, \tilde{u}, \tilde{x}_0)$. When $(\bar{x}, \bar{u}, \bar{x}_0)$ is a local minimizer for the problem (1.3) and $\bar{u}$ is singular, we call $(\bar{x}, \bar{u}, \bar{x}_0)$ a singular local minimizer (for the problem (1.3)).

Remark 4.2. The definition of the singular control in (4.36) is much more general than that in [33, Definition 3.3]. More precisely, by the maximality condition (3.16), if the control $\tilde{u}$ is optimal, the first and second necessary conditions in optimization theory immediately imply that, for a.e. $(t, \omega) \in [0, T] \times \Omega$,

$$\big\langle \widetilde{H}_u(t, \omega), v \big\rangle \le 0, \quad \forall\, v \in T_U^b(\tilde{u}(t, \omega)).$$

Further, if $\big\langle \widetilde{H}_u(t, \omega), v_0 \big\rangle = 0$ for some $v_0 \in T_U^b(\tilde{u}(t, \omega))$, then for any $h \in T_U^{b(2)}(\tilde{u}(t, \omega), v_0)$,

$$\big\langle \widetilde{H}_u(t, \omega), h \big\rangle + \frac{1}{2} \Big\langle \big(\widetilde{H}_{uu}(t, \omega) + \tilde{\sigma}_u(t, \omega)^\top \widetilde{P}_2(t, \omega) \tilde{\sigma}_u(t, \omega)\big) v_0, v_0 \Big\rangle \le 0. \tag{4.37}$$

Both Definition 4.1 and [33, Definition 3.3] imply that the corresponding singular controls satisfy the above first and second order necessary condition trivially, but in Definition 4.1, $\widetilde{H}_{uu}(t) + \tilde{\sigma}_u(t)^\top \widetilde{P}_2(t) \tilde{\sigma}_u(t)$ is only assumed to be degenerated, for a.e. $[0, T] \times \Omega$, in the directions from $T_U^b(\tilde{u}(t))$. We shall see in Example 4.3 below that for partially singular controls, $\widetilde{H}_{uu}(t) + \tilde{\sigma}_u(t)^\top \widetilde{P}_2(t) \tilde{\sigma}_u(t)$ may be different from 0 on a subset of $[0, T] \times \Omega$ having positive measure.

By Theorem 4.1, it is easy to verify the following second order integral-type necessary condition for the problem (1.3).

Theorem 4.3. *Let (C1)–(C3) hold. If $(\bar{x}, \bar{u}, \bar{x}_0)$ is a singular local minimizer for the problem (1.3) and $\bar{u} \in L^4_{\mathbb{F}}(\Omega; L^4(0, T; \mathbb{R}^m))$, then*

$$\mathbb{E} \int_0^T \langle \mathbb{S}(t) y_1(t), v(t) \rangle \, dt \le 0, \qquad \forall\, v \in \mathcal{A}^1_{\bar{u}}. \tag{4.38}$$

As underlined in [33], there are some essential difficulties to deduce from the above integral type second order necessary condition a pointwise one. The main reason for it is that the spike variations have to be used to get the pointwise second order necessary condition from (4.38). Substituting the explicit expression for y_1 into (4.38), the Itô integral will appear in this condition. Thus there will be a "bad" term making impossible using the Lebesgue differentiation theorem to derive the pointwise condition (see Subsection 3.2 in [33] for more details). However, when $\mathbb{S}$ and v are regular enough, a method similar to the one proposed in [33] can be used to establish the following pointwise second-order necessary condition for stochastic singular optimal controls for the problem (1.3).

Theorem 4.4. *Let (C1)–(C3) hold. If $(\bar{x}, \bar{u}, \bar{x}_0)$ is a singular local minimizer for the problem (1.3), $\bar{u} \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ and $\mathbb{S} \in \mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^{m\times n}) \cap L^\infty([0,T]\times\Omega;\mathbb{R}^{m\times n})$, then in addition to the second order transversality condition (4.7), for any $v \in \mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^m) \cap L^\infty([0,T]\times\Omega;\mathbb{R}^m) \cap \mathcal{A}^1_{\bar{u}}$, the following pointwise second order necessary condition holds:*

$$\begin{aligned}
&\langle \mathbb{S}(\tau) b_u(\tau) v(\tau), v(\tau)\rangle + \langle \nabla\mathbb{S}(\tau)\sigma_u(\tau)v(\tau), v(\tau)\rangle \\
&\quad + \langle \mathbb{S}(\tau)\sigma_u(\tau)v(\tau), \nabla v(\tau)\rangle \le 0, \quad a.e.\ \tau\in[0,T],\ a.s.
\end{aligned} \tag{4.39}$$

Proof. The proof is similar to the one of [33, Theorem 3.13]. Let $\tau \in [0,T)$, $\theta\in(0,T-\tau)$, $E_\theta = [\tau,\tau+\theta)$ and choose $A\in\mathcal{F}_\tau$. For any $v(\cdot)\in \mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^m) \cap L^\infty([0,T]\times\Omega;\mathbb{R}^m) \cap \mathcal{A}^1_{\bar{u}}$, define

$$v^{\theta,A}(t,\omega) = \begin{cases} v(t,\omega), & (t,\omega)\in E_\theta\times A,\\ 0, & (t,\omega)\in \big([0,T]\times\Omega\big)\setminus\big(E_\theta\times A\big).\end{cases}$$

Clearly, $v^{\theta,A}(\cdot)\in\mathcal{A}^1_{\bar{u}}$. Denote by $y_1^{\theta,A}(\cdot)$ the solution to the first order variational equation (3.3) with $v(\cdot)$ replaced by $v^{\theta,A}(\cdot)$. By [32, Theorem 1.6.14, p. 47], $y_1^{\theta,A}(\cdot)$ enjoys an explicit representation:

$$\begin{aligned}
y_1^{\theta,A}(t) &= \Phi(t)\int_0^t \Phi(s)^{-1}\big(b_u(s) - \sigma_x(s)\sigma_u(s)\big)v^{\theta,A}(s)ds \\
&\quad + \Phi(t)\int_0^t \Phi(s)^{-1}\sigma_u(s)v^{\theta,A}(s)dW(s),
\end{aligned} \tag{4.40}$$

where $\Phi(\cdot)$ solves the following matrix-valued stochastic differential equation

$$\begin{cases} d\Phi(t) = b_x(t)\Phi(t)dt + \sigma_x(t)\Phi(t)dW(t), & t\in[0,T],\\ \Phi(0) = I, \end{cases} \tag{4.41}$$

and I stands for the identity matrix of dimension n.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

From Theorem 4.3, it follows that

$$
\begin{aligned}
0 &\geq \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t) y_1^{\theta,A}(t), v(t)\Big\rangle \chi_A dt \\
&= \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\big(b_u(s)-\sigma_x(s)\sigma_u(s)\big)v(s)\chi_A ds, v(t)\Big\rangle \chi_A dt \\
&\quad + \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt.
\end{aligned} \tag{4.42}
$$

By the Lebesgue differentiation theorem, it is immediate that for a.e. $\tau \in [0, T)$,

$$
\begin{aligned}
&\lim_{\theta\to 0^+}\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\big(b_u(s)-\sigma_x(s)\sigma_u(s)\big)v(s)\chi_A ds, v(t)\Big\rangle \chi_A dt \\
&= \frac{1}{2}\mathbb{E}\Big(\big\langle \mathbb{S}(\tau)\big(b_u(\tau)-\sigma_x(\tau)\sigma_u(\tau)\big)v(\tau), v(\tau)\big\rangle \chi_A\Big).
\end{aligned} \tag{4.43}
$$

On the other hand, by (4.41)

$$
\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt \tag{4.44}
$$

$$
\begin{aligned}
&= \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\Phi(\tau)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt \\
&\quad + \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\int_\tau^t b_x(s)\Phi(s)ds\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt \\
&\quad + \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\int_\tau^t \sigma_x(s)\Phi(s)dW(s)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt.
\end{aligned}
$$

By the properties of the Itô integral and the Lebesgue differentiation theorem, it can be proved that

$$
\begin{aligned}
&\lim_{\theta\to 0^+}\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta} \Big\langle \mathbb{S}(t)\int_\tau^t b_x(s)\Phi(s)ds\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt \\
&= 0, \quad \text{a.e. } \tau \in [0, T),
\end{aligned} \tag{4.45}
$$

and

$$\lim_{\theta\to 0^+}\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left\langle \mathbb{S}(t)\int_\tau^t \sigma_x(s)\Phi(s)dW(s)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s),\, v(t)\right\rangle \chi_A dt$$
$$=\frac{1}{2}\mathbb{E}\left(\langle \mathbb{S}(\tau)\sigma_x(\tau)\sigma_u(\tau)v(\tau),\, v(\tau)\rangle\, \chi_A\right),\quad \text{a.e. } \tau\in[0,T). \tag{4.46}$$

Next, the assumptions on $\mathbb{S}$ and v yield

$$\mathbb{S}(\cdot)^\top v(\cdot)\in \mathbb{L}^{1,2}_{\mathbb{F}}(\mathbb{R}^n)\cap L^\infty([0,T]\times\Omega;\mathbb{R}^n).$$

Hence, by the Clark–Ocone formula, for a.e. $t\in[0,T)$,

$$\mathbb{S}(t)^\top v(t)=\mathbb{E}\left(\mathbb{S}(t)^\top v(t)\right)+\int_0^t \mathbb{E}\left(\mathcal{D}_s\big(\mathbb{S}(t)^\top v(t)\big)\,\Big|\, \mathcal{F}_s\right)dW(s). \tag{4.47}$$

Substituting (4.47) into the first term of the right hand of (4.44), it follows that

$$\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left\langle \mathbb{S}(t)\Phi(\tau)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s),\, v(t)\right\rangle \chi_A dt \tag{4.48}$$
$$=\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\langle \int_\tau^t \Phi(\tau)\Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s),\, \mathbb{E}\left(\mathbb{S}(t)^\top v(t)\right)\right\rangle \chi_A dt$$
$$+\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\langle \int_\tau^t \Phi(\tau)\Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s),\, \int_0^t \mathbb{E}\Big(\mathcal{D}_s\big(\mathbb{S}(t)^\top v(t)\big)\Big|\mathcal{F}_s\Big)dW(s)\right\rangle \chi_A dt$$
$$=\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\int_\tau^t \mathbb{E}\left\langle \Phi(\tau)\Phi(s)^{-1}\sigma_u(s)v(s),\, \mathcal{D}_s\big(\mathbb{S}(t)^\top v(t)\big)\right\rangle \chi_A ds dt.$$

Note that

$$\mathcal{D}_s\big(\mathbb{S}(t)^\top v(t)\big)=\big(\mathcal{D}_s\mathbb{S}(t)^\top\big)v(t)+\mathbb{S}(t)^\top\mathcal{D}_s v(t).$$

Using the same argument as that in [33, Theorem 3.13], we conclude that there exists a sequence $\{\theta_\ell\}_{\ell=1}^\infty$ of positive numbers such that $\lim_{\ell\to\infty}\theta_\ell=0$, and

$$\lim_{\ell\to\infty}\frac{1}{\theta_\ell^2}\mathbb{E}\int_\tau^{\tau+\theta_\ell}\left\langle \mathbb{S}(t)\Phi(\tau)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s),\, v(t)\right\rangle \chi_A dt \tag{4.49}$$
$$=\frac{1}{2}\mathbb{E}\Big(\Big\langle \nabla\mathbb{S}(\tau)^\top v(\tau),\, \sigma_u(\tau)v(\tau)\Big\rangle\chi_A\Big)+\frac{1}{2}\mathbb{E}\Big(\Big\langle \mathbb{S}(\tau)^\top\nabla v(\tau),\, \sigma_u(\tau)v(\tau)\Big\rangle\chi_A\Big),\ \text{a.e. } \tau\in[0,T).$$

Then, by (4.44), (4.49), one concludes that

$$\lim_{\ell\to\infty}\frac{1}{\theta_\ell^2}\mathbb{E}\int_\tau^{\tau+\theta_\ell}\Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)v(s)\chi_A dW(s), v(t)\Big\rangle \chi_A dt \tag{4.50}$$
$$=\frac{1}{2}\mathbb{E}\left(\langle \mathbb{S}(\tau)\sigma_x(\tau)\sigma_u(\tau)v(\tau), v(\tau)\rangle \chi_A\right)+\frac{1}{2}\mathbb{E}\left(\Big\langle \nabla\mathbb{S}(\tau)^\top v(\tau), \sigma_u(\tau)v(\tau)\Big\rangle \chi_A\right)$$
$$+\frac{1}{2}\mathbb{E}\left(\Big\langle \mathbb{S}(\tau)^\top \nabla v(\tau), \sigma_u(\tau)v(\tau)\Big\rangle \chi_A\right), \quad \text{a.e. } \tau\in[0,T).$$

Combining (4.42), (4.43) and (4.50), one has

$$0\geq \mathbb{E}\left(\langle \mathbb{S}(\tau)b_u(\tau)v(\tau)), v(\tau)\rangle \chi_A\right)+\mathbb{E}\left(\Big\langle \nabla\mathbb{S}(\tau)^\top v(\tau), \sigma_u(\tau)v(\tau)\Big\rangle \chi_A\right)$$
$$+\mathbb{E}\left(\Big\langle \mathbb{S}(\tau)^\top \nabla v(\tau), \sigma_u(\tau)v(\tau)\Big\rangle \chi_A\right), \quad \text{a.e. } \tau\in[0,T).$$

Finally, by the arbitrariness of $A\in\mathcal{F}_\tau$, we deduce that the desired second order necessary condition (4.39) holds. This completes the proof of Theorem 4.4. □

If $\bar{u}\in\mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^m)$, U is a bounded closed convex set in $\mathbb{R}^m$, $v-\bar{u}(\cdot)\in\mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^m)\cap L^\infty([0,T]\times\Omega;\mathbb{R}^m)\cap\mathcal{A}^1_{\bar{u}}$ holds true for any $v\in U$. Then, by Theorem 4.4 and the separability of U, one has

$$\langle \mathbb{S}(\tau)b_u(\tau)(v-\bar{u}(\tau)), v-\bar{u}(\tau)\rangle+\langle \nabla\mathbb{S}(\tau)\sigma_u(\tau)(v-\bar{u}(\tau)), v-\bar{u}(\tau)\rangle$$
$$-\langle \mathbb{S}(\tau)\sigma_u(\tau)(v-\bar{u}(\tau)), \nabla\bar{u}(\tau)\rangle\leq 0, \qquad \forall\, v\in U,\ \text{a.e. } \tau\in[0,T],\ \text{a.s.,} \tag{4.51}$$

which coincides with [33, Theorem 3.13]. However, when the control set U is nonconvex, some more assumptions as follows are required to establish a pointwise condition similar to (4.51).

(C4) For any $u\in\partial U$ and $v\in T^b_U(u)$, $T^{b(2)}_U(u,v)\neq\emptyset$.

When the control set U has a C^2 boundary, the assumption (C4) holds, see [10].
From the proof of Theorem 4.4, we deduce the following result.

Corollary 4.2. *Let (C1)–(C4) hold, $(\bar{x},\bar{u},\bar{x}_0)$ be a singular local minimizer for the problem (1.3). If $\mathbb{S}\in\mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^{m\times n})$, and the optimal control $\bar{u}$ is a step function as below*

$$\bar{u}(t,\omega)=\sum_{i=1}^k\sum_{j=1}^{l_i}u_{ij}\chi_{A_{ij}}\chi_{[t_i,t_{i+1})}(t,\omega),\ \text{a.e. } (t,\omega)\in[0,T]\times\Omega, \tag{4.52}$$

where $k\in\mathbb{N}$, $0=t_1<\cdots<t_{k+1}=T$, $l_i\in\mathbb{N}$, $u_{ij}\in U$ and $A_{ij}\in\mathcal{F}_{t_i}$ for $i=1,\cdots,k$ and $j=1,\cdots,l_i$, then, in addition to the second order transversality condition (4.7), the following pointwise second order necessary condition holds:

$$\begin{aligned}&\langle \mathbb{S}(\tau,\omega)b_u(\tau,\omega)v, v\rangle + \langle \nabla\mathbb{S}(\tau,\omega)\sigma_u(\tau,\omega)v, v\rangle \le 0,\\ &\forall\, v\in T_U^b(\bar u(\tau,\omega)),\ a.e.\ \tau\in[0,T],\ a.s.\end{aligned} \tag{4.53}$$

Proof. When $\bar u(t,\omega)$ is given as in (4.52), for any fixed i and j, $\bar u(t,\omega)$ has constant value u_{ij} on $[t_i,t_{i+1})\times A_{ij}$. Then, on $[t_i,t_{i+1})\times A_{ij}$, let $v_{ij}\in T_U^b(u_{ij})$, $h_{ij}\in T_U^{b(2)}(u_{ij},v_{ij})$, $\tau\in[t_i,t_{i+1})$, $\theta\in(0,t_{i+1}-\tau)$, $E_\theta=[\tau,\tau+\theta)$ and choose $A\in\mathcal{F}_{t_i}$. Define

$$v^{\theta,A}(t,\omega)=\begin{cases} v_{ij}, & (t,\omega)\in E_\theta\times(A\cap A_{ij}),\\ 0, & \text{otherwise,}\end{cases}\qquad h^{\theta,A}(t,\omega)=\begin{cases} h_{ij}, & (t,\omega)\in E_\theta\times(A\cap A_{ij}),\\ 0, & \text{otherwise.}\end{cases}$$

It is clear that $(v^{\theta,A},h^{\theta,A})\in\mathcal{A}_{\bar u}$. Then, by similar arguments as in the proof of Theorem 4.4 and noting that the Malliavin derivative of the constant-valued process v_{ij} is equal to 0, we obtain that

$$\big\langle \mathbb{S}(\tau,\omega)b_u(\tau,\omega)v_{ij}, v_{ij}\big\rangle + \big\langle \nabla\mathbb{S}(\tau,\omega)\sigma_u(\tau,\omega)v_{ij}, v_{ij}\big\rangle \le 0, \qquad \text{a.e. } (\tau,\omega)\in[t_i,t_{i+1})\times A_{ij}.$$

By the closedness of the adjacent cone, the separability of $\mathbb{R}^m$, the arbitrariness of i, j and v_{ij} it follows that

$$\langle \mathbb{S}(\tau,\omega)b_u(\tau,\omega)v, v\rangle + \langle \nabla\mathbb{S}(\tau,\omega)\sigma_u(\tau,\omega)v, v\rangle \le 0,\ \forall\, v\in T_U^b(\bar u(\tau,\omega)),\ \text{a.e. } \tau\in[0,T],\ \text{a.s.}$$

This completes the proof of Corollary 4.2. □

Example 4.3. Let the optimal control problem be the one stated in Example 3.1. We have shown that $\bar u(t)=(\bar u_1(t),\bar u_2(t))\equiv(1,0)$ is the optimal control. In the following we will prove that this optimal control is partially singular and satisfies the second order necessary condition (4.53).

In Example 3.1 we obtained that the corresponding state $(\bar x_1(t),\bar x_2(t))$ is as in (3.23) and the first order adjoint process $(P_1(t),Q_1(t))$ is as in (3.24). In addition, it is easy to see that the second order adjoint equation is

$$\begin{cases} d\begin{bmatrix} P_2^1(t) & P_2^2(t)\\ P_2^3(t) & P_2^4(t)\end{bmatrix} = \begin{bmatrix} 0 & -P_2^1(t)\\ -P_2^1(t) & -P_2^2(t)-P_2^3(t)\end{bmatrix}dt + \begin{bmatrix} Q_2^1(t) & Q_2^2(t)\\ Q_2^3(t) & Q_2^4(t)\end{bmatrix}dW(t),\ t\in[0,1],\\ \begin{bmatrix} P_2^1(1) & P_2^2(1)\\ P_2^3(1) & P_2^4(1)\end{bmatrix} = \begin{bmatrix} -1 & 0\\ 0 & 0\end{bmatrix}\end{cases}$$

and its solution is

$$\left(\begin{bmatrix} P_2^1(t) & P_2^2(t)\\ P_2^3(t) & P_2^4(t)\end{bmatrix}, \begin{bmatrix} Q_2^1(t) & Q_2^2(t)\\ Q_2^3(t) & Q_2^4(t)\end{bmatrix}\right) = \left(\begin{bmatrix} -1 & t-1\\ t-1 & -t^2+2t-1\end{bmatrix}, \begin{bmatrix} 0 & 0\\ 0 & 0\end{bmatrix}\right).$$

A direct calculation shows that

$$T_U^b((1,0))=\{(v_1,0)\in\mathbb{R}^2 \mid v_1\le 0\}.$$

Then, we have

$$H_u(t)=0,\ H_{uu}(t)=0,\quad \forall\,(t,\omega)\in[0,1]\times\Omega,$$
$$\sigma_u(t)^\top P_2(t)\sigma_u(t)=\begin{bmatrix}0&0\\0&1\end{bmatrix}\begin{bmatrix}-1&t-1\\t-1&-t^2+2t-1\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}0&0\\0&-t^2+2t-1\end{bmatrix}$$

and therefore

$$\Big\langle\big(H_{uu}(t)+\sigma_u(t)^\top P_2(t)\sigma_u(t)\big)v,v\Big\rangle=0,\quad \forall\,v\in T_U^b(\bar u(t)),\ \text{a.e. } t\in[0,T],\ \text{a.s.}$$

This means that $\bar u(t)=(\bar u_1(t),\bar u_2(t))\equiv(1,0)$ is partially singular. Next, we prove that $\bar u(t)=(\bar u_1(t),\bar u_2(t))\equiv(1,0)$ satisfies the second order necessary condition in Corollary 4.2. It is clear that

$$\mathbb{S}(t)=\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}-1&t-1\\t-1&-t^2+2t-1\end{bmatrix}=\begin{bmatrix}t-1&-t^2+2t-1\\0&0\end{bmatrix}.$$

Then, $\nabla\mathbb{S}(t)\equiv0$, and

$$\begin{aligned}\langle\mathbb{S}(t)b_u(t)v,v\rangle+\langle\nabla\mathbb{S}(t)\sigma_u(t)v,v\rangle&=\begin{bmatrix}v_1&0\end{bmatrix}\begin{bmatrix}-t^2+2t-1&0\\0&0\end{bmatrix}\begin{bmatrix}v_1\\0\end{bmatrix}\\&=-(t-1)^2v_1^2\le0,\ \forall\,v\in T_U^b(\bar u(t)),\ \text{a.e. } t\in[0,T],\ \text{a.s.}\end{aligned}$$

Acknowledgment

The authors highly appreciate the constructive comments of two anonymous referees which led to several improvements of the original version.

Appendix A

In this section, we prove the two technical Lemmas 3.2 and 4.1. The fundamental idea comes from the classical calculus, see also the related results in [4,6] for the optimal control problems with convex control constraints, and [27,32] for the general control constraints.

A.1. Proof of Lemma 3.2

Proof. From (3.3) and Lemma 3.1 we deduce that

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,T]}|y_1(t)|^\beta\Big)&\le C\mathbb{E}\left[|\nu_0|^\beta+\Big(\int_0^T|b_u(t)v(t)|dt\Big)^\beta+\Big(\int_0^T|\sigma_u(t)v(t)|^2dt\Big)^{\frac{\beta}{2}}\right]\\&\le C\mathbb{E}\left[|\nu_0|^\beta+\Big(\int_0^T|v(t)|^2dt\Big)^{\frac{\beta}{2}}\right].\end{aligned}$$

Since $v_\varepsilon(\cdot)$ converges to $v(\cdot)$ in $L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$, and $\nu_0^\varepsilon \to \nu_0$ in $\mathbb{R}^n$ as $\varepsilon \to 0^+$, we deduce from (3.2) that

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |\delta x^\varepsilon(t)|^\beta\Big) \le C\mathbb{E}\left[\varepsilon^\beta|\nu_0^\varepsilon|^\beta + \Big(\int_0^T |\varepsilon v_\varepsilon(t)|^2 ds\Big)^{\frac{\beta}{2}}\right] = O(\varepsilon^\beta).$$

Consequently, by the Hölder inequality,

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |\delta x^\varepsilon(t)|\Big) \le \left[\mathbb{E}\Big(\sup_{t\in[0,T]} |\delta x^\varepsilon(t)|^\beta\Big)\right]^{1/\beta} = O(\varepsilon) \tag{A.1}$$

and

$$\mathbb{E}\int_0^T |v_\varepsilon(t)-v(t)|dt \le C\left[\mathbb{E}\Big(\int_0^T |v_\varepsilon(t)-v(t)|^2 dt\Big)^{\frac{\beta}{2}}\right]^{\frac{1}{\beta}} \to 0,\ \ \varepsilon \to 0^+. \tag{A.2}$$

Denote $\tilde{b}_x^\varepsilon(t) := \int_0^1 b_x(t, \bar{x}(t)+\theta\delta x^\varepsilon(t), \bar{u}(t)+\theta\varepsilon v_\varepsilon(t))d\theta$. Mappings $\tilde{b}_u^\varepsilon(t)$, $\tilde{\sigma}_x^\varepsilon(t)$ and $\tilde{\sigma}_u^\varepsilon(t)$ are defined in a similar way. Then, $\delta x^\varepsilon(\cdot)$ is the solution to the following stochastic differential equation

$$\begin{cases} d\delta x^\varepsilon(t) = \big(\tilde{b}_x^\varepsilon(t)\delta x^\varepsilon(t) + \varepsilon\tilde{b}_u^\varepsilon(t)v_\varepsilon(t)\big)dt \\ \qquad\qquad + \big(\tilde{\sigma}_x^\varepsilon(t)\delta x^\varepsilon(t) + \varepsilon\tilde{\sigma}_u^\varepsilon(t)v_\varepsilon(t)\big)dW(t), \quad t\in[0,T], \\ \delta x^\varepsilon(0) = \varepsilon\nu_0^\varepsilon, \end{cases}$$

and $r_1^\varepsilon(\cdot)$ satisfies the following stochastic differential equation

$$\begin{cases} dr_1^\varepsilon(t) = \Big[\tilde{b}_x^\varepsilon(t)r_1^\varepsilon(t) + \big(\tilde{b}_x^\varepsilon(t)-b_x(t)\big)y_1(t) + \tilde{b}_u^\varepsilon(t)\big(v_\varepsilon(t)-v(t)\big) \\ \qquad + \big(\tilde{b}_u^\varepsilon(t)-b_u(t)\big)v(t)\Big]dt + \Big[\tilde{\sigma}_x^\varepsilon(t)r_1^\varepsilon(t) + \big(\tilde{\sigma}_x^\varepsilon(t)-\sigma_x(t)\big)y_1(t) \\ \qquad + \tilde{\sigma}_u^\varepsilon(t)\big(v_\varepsilon(t)-v(t)\big) + \big(\tilde{\sigma}_u^\varepsilon(t)-\sigma_u(t)\big)v(t)\Big]dW(t), \quad t\in[0,T], \\ r_1^\varepsilon(0) = \nu_0^\varepsilon - \nu_0. \end{cases} \tag{A.3}$$

For any sequence $\{\varepsilon_j\}_{j=1}^\infty$ of positive numbers converging to 0 as $j\to\infty$, we can find a subsequence $\{j_k\}_{k=1}^\infty \subset \mathbb{N}$ such that $\sup_{t\in[0,T]}|\delta x^{\varepsilon_{j_k}}(t)| \to 0$ a.s. and $\varepsilon_{j_k}v_{\varepsilon_{j_k}}(t) \to 0$ a.s. for a.e. $t\in[0,T]$, as $k\to\infty$. The assumption (C2) yields, $\big|(\tilde{b}_x^{\varepsilon_{j_k}}(t) - b_x(t))y_1(t)\big| \to 0$ a.s. for a.e. $t\in[0,T]$, as $k\to\infty$. Hence,

$$\big|(\tilde{b}_x^{\varepsilon_j}(\cdot) - b_x(\cdot))y_1(\cdot)\big| \to 0 \ \text{ in measure, as } j\to\infty.$$

Then, using Lebesgue's dominated convergence theorem, we conclude that

$$\mathbb{E}\Big(\int_0^T |\big(\tilde{b}_x^{\varepsilon_j}(t) - b_x(t)\big)y_1(t)|^2 dt\Big)^{\frac{\beta}{2}} \to 0, \quad j\to\infty. \tag{A.4}$$

A slight modification of the above discussion shows that

$$\mathbb{E}\Big(\int_0^T |\big(\tilde b_u^{\varepsilon_j}(t)-b_u(t)\big)v(t)|^2dt\Big)^{\frac{\beta}{2}}+\mathbb{E}\Big(\int_0^T |\big(\tilde \sigma_x^{\varepsilon_j}(t)-\sigma_x(t)\big)y_1(t)|^2dt\Big)^{\frac{\beta}{2}}$$
$$+\mathbb{E}\Big(\int_0^T |\big(\tilde \sigma_u^{\varepsilon_j}(t)-\sigma_u(t)\big)v(t)|^2dt\Big)^{\frac{\beta}{2}}\to 0,\quad j\to\infty. \tag{A.5}$$

On the other hand

$$\mathbb{E}\Big(\int_0^T |\tilde b_u^{\varepsilon_j}(t)\big(v_{\varepsilon_j}(t)-v(t)\big)|^2dt\Big)^{\frac{\beta}{2}}+\mathbb{E}\Big(\int_0^T |\tilde \sigma_u^{\varepsilon_j}(t)\big(v_{\varepsilon_j}(t)-v(t)\big)|^2dt\Big)^{\frac{\beta}{2}}$$
$$\le C\mathbb{E}\Big(\int_0^T |v_{\varepsilon_j}(t)-v(t)|^2dt\Big)^{\frac{\beta}{2}}\to 0,\quad j\to\infty.$$

Therefore, by Lemma 3.1, we finally obtain that

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|r_1^{\varepsilon_j}(t)|^\beta\Big)\le C\mathbb{E}\Big[|\nu_0^{\varepsilon_j}-\nu_0|^\beta+\Big(\int_0^T\Big|\big(\tilde b_x^{\varepsilon_j}(t)-b_x(t)\big)y_1(t)+\tilde b_u^{\varepsilon_j}(t)\big(v_{\varepsilon_j}(t)-v(t)\big)$$
$$+\big(\tilde b_u^{\varepsilon_j}(t)-b_u(t)\big)v(t)\Big|dt\Big)^\beta+\Big(\int_0^T\Big|\big(\tilde \sigma_x^{\varepsilon_j}(t)-\sigma_x(t)\big)y_1(t)$$
$$+\tilde \sigma_u^{\varepsilon_j}(t)\big(v_{\varepsilon_j}(t)-v(t)\big)+\big(\tilde \sigma_u^{\varepsilon_j}(t)-\sigma_u(t)\big)v(t)\Big|^2dt\Big)^{\frac{\beta}{2}}\Big]\to 0,\quad j\to\infty.$$

The sequence $\varepsilon_j\to 0^+$ being arbitrary, the proof is complete. □

A.2. Proof of Lemma 4.1

Proof. By Lemma 3.2 (with β replaced by 2β), we obtain that

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|y_1(t)|^{2\beta}\Big)\le C\mathbb{E}\Big[|\nu_0|^{2\beta}+\Big(\int_0^T|v(t)|^2dt\Big)^\beta\Big]. \tag{A.6}$$

Then, by (4.1), Lemma 3.1 and the Hölder inequality, it follows that

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|y_2(t)|^\beta\Big)$$
$$\le C\mathbb{E}\Big[|\varpi_0|^\beta+\Big(\int_0^T|2b_u(t)h(t)+y_1(t)^\top b_{xx}(t)y_1(t)+2v(t)^\top b_{xu}(t)y_1(t)$$

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

$$
\begin{aligned}
&+ v(t)^\top b_{uu}(t)v(t)|dt\Big)^{\beta} + \Big(\int_0^T |2\sigma_u(t)h(t) + y_1(t)^\top \sigma_{xx}(t)y_1(t) \\
&+ 2v(t)^\top \sigma_{xu}(t)y_1(t) + v(t)^\top \sigma_{uu}(t)v(t)|^2 dt\Big)^{\frac{\beta}{2}}\Big] \\
\le\ & C\mathbb{E}\Big[|\varpi_0|^\beta + \Big(\int_0^T |h(t)|^2 dt\Big)^{\frac{\beta}{2}} + \sup_{t\in[0,T]} |y_1(t)|^{2\beta} \\
&+ \sup_{t\in[0,T]} |y_1(t)|^\beta \cdot \Big(\int_0^T |v(t)|^2 dt\Big)^{\frac{\beta}{2}} + \Big(\int_0^T |v(t)|^4 dt\Big)^{\frac{\beta}{2}}\Big] \\
\le\ & C\mathbb{E}\Big[|\varpi_0|^\beta + |\nu_0|^{2\beta} + \Big(\int_0^T |h(t)|^2 dt\Big)^{\frac{\beta}{2}} + \Big(\int_0^T |v(t)|^4 dt\Big)^{\frac{\beta}{2}}\Big].
\end{aligned}
$$

Denote $\tilde b^\varepsilon_{xx}(t) := \int_0^1 (1-\theta) b_{xx}(t, \bar x(t) + \theta\delta x^\varepsilon(t), \bar u(t) + \theta\delta u^\varepsilon(t))d\theta$. Mappings $\tilde b^\varepsilon_{xu}(t)$, $\tilde b^\varepsilon_{uu}(t)$, $\tilde\sigma^\varepsilon_{xx}(t)$, $\tilde\sigma^\varepsilon_{xu}(t)$ and $\tilde\sigma^\varepsilon_{uu}(t)$ are defined in a similar way. Then, δx^ε satisfies the following stochastic differential equation:

$$
\begin{cases}
d\delta x^\varepsilon(t) = \Big(b_x(t)\delta x^\varepsilon(t) + b_u(t)\delta u^\varepsilon(t) + \delta x^\varepsilon(t)^\top \tilde b^\varepsilon_{xx}(t)\delta x^\varepsilon(t) \\
\qquad + 2\delta x^\varepsilon(t)^\top \tilde b^\varepsilon_{xu}(t)\delta u^\varepsilon(t) + \delta u^\varepsilon(t)^\top \tilde b^\varepsilon_{uu}(t)\delta u^\varepsilon(t)\Big)dt \\
\qquad + \Big(\sigma_x(t)\delta x^\varepsilon(t) + \sigma_u(t)\delta u^\varepsilon(t) + \delta x^\varepsilon(t)^\top \tilde\sigma^\varepsilon_{xx}(t)\delta x^\varepsilon(t) \\
\qquad + 2\delta x^\varepsilon(t)^\top \tilde\sigma^\varepsilon_{xu}(t)\delta u^\varepsilon(t) + \delta u^\varepsilon(t)^\top \tilde\sigma^\varepsilon_{uu}(t)\delta u^\varepsilon(t)\Big)dW(t), \quad t\in[0,T], \\
\delta x^\varepsilon(0) = \varepsilon\nu_0 + \varepsilon^2\varpi^\varepsilon_0.
\end{cases}
$$

Therefore r_2^ε solves the following stochastic differential equation:

$$
\begin{cases}
dr_2^\varepsilon(t) = \Big\{b_x(t)r_2^\varepsilon(t) + b_u(t)\big(h_\varepsilon(t) - h(t)\big) \\
\qquad + \Big[\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big)^\top \tilde b^\varepsilon_{xx}(t)\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big) - \frac{1}{2}y_1(t)^\top b_{xx}(t)y_1(t)\Big] \\
\qquad + \Big[2\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big)^\top \tilde b^\varepsilon_{xu}(t)\Big(\frac{\delta u^\varepsilon(t)}{\varepsilon}\Big) - y_1(t)^\top b_{xu}(t)v(t)\Big] \\
\qquad + \Big[\Big(\frac{\delta u^\varepsilon(t)}{\varepsilon}\Big)^\top \tilde b^\varepsilon_{uu}(t)\Big(\frac{\delta u^\varepsilon(t)}{\varepsilon}\Big) - \frac{1}{2}v(t)^\top b_{uu}(t)v(t)\Big]\Big\}dt \\
\qquad + \Big\{\sigma_x(t)r_2^\varepsilon(t) + \sigma_u(t)\big(h_\varepsilon(t) - h(t)\big) \\
\qquad + \Big[\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big)^\top \tilde\sigma^\varepsilon_{xx}(t)\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big) - \frac{1}{2}y_1(t)^\top \sigma_{xx}(t)y_1(t)\Big] \\
\qquad + \Big[2\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big)^\top \tilde\sigma^\varepsilon_{xu}(t)\Big(\frac{\delta u^\varepsilon(t)}{\varepsilon}\Big) - y_1(t)^\top \sigma_{xu}(t)v(t)\Big] \\
\qquad + \Big[\Big(\frac{\delta u^\varepsilon(t)}{\varepsilon}\Big)^\top \tilde\sigma^\varepsilon_{uu}(t)\Big(\frac{\delta u^\varepsilon(t)}{\varepsilon}\Big) - \frac{1}{2}v(t)^\top \sigma_{uu}(t)v(t)\Big]\Big\}dW(t), \quad t\in[0,T], \\
r_2^\varepsilon(0) = \varpi^\varepsilon_0 - \varpi_0.
\end{cases}
\tag{A.7}
$$

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

Since $h_\varepsilon(\cdot)$ converges to $h(\cdot)$ in $L^{2\beta}_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$,

$$\mathbb{E}\Big(\int_0^T \Big|b_u(t)\big(h_\varepsilon(t)-h(t)\big)\Big|dt\Big)^\beta+\mathbb{E}\Big(\int_0^T \Big|\sigma_u(t)\big(h_\varepsilon(t)-h(t)\big)\Big|^2dt\Big)^{\frac{\beta}{2}}\to 0,\quad \varepsilon\to 0^+. \quad (A.8)$$

On the other hand, by the Hölder inequality,

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^T \Big|\big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\big)^\top \tilde b^\varepsilon_{xx}(t)\big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\big)-\frac{1}{2}y_1(t)^\top b_{xx}(t)y_1(t)\Big|dt\Big)^\beta\\
&\le C\mathbb{E}\Big(\int_0^T \Big|\big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\big)^\top \tilde b^\varepsilon_{xx}(t)\big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\big)-\frac{1}{2}y_1(t)^\top b_{xx}(t)y_1(t)\Big|^2dt\Big)^{\frac{\beta}{2}}\\
&\le C\mathbb{E}\Big[\int_0^T \Big|\big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\big)^\top \big(\tilde b^\varepsilon_{xx}(t)-\frac{1}{2}b_{xx}(t)\big)\big(\frac{\delta x^\varepsilon(t)}{\varepsilon}\big)\Big|^2dt\Big]^{\frac{\beta}{2}}\\
&\quad +C\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}-y_1(t)\Big|^\beta\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^\beta+\sup_{t\in[0,T]}|y_1(t)|^\beta\Big)\Big]\\
&\le C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^{2\beta}\Big)\Big]^{1/2}\Big[\mathbb{E}\Big(\int_0^T\Big|\tilde b^\varepsilon_{xx}(t)-\frac{1}{2}b_{xx}(t)\Big|^4\cdot\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^4dt\Big)^{\frac{\beta}{2}}\Big]^{1/2}\\
&\quad +C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}-y_1(t)\Big|^{2\beta}\Big)\Big]^{\frac{1}{2}}\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^{2\beta}+\sup_{t\in[0,T]}|y_1(t)|^{2\beta}\Big)\Big]^{\frac{1}{2}}. \quad (A.9)
\end{aligned}$$

Since h_ε converges to h in $L^{2\beta}_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ and ϖ_0^ε converges to ϖ_0 in $\mathbb{R}^m$ as $\varepsilon\to 0^+$, by Lemma 3.1,

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|\delta x^\varepsilon(t)|^{2\beta}\Big)\le C\mathbb{E}\Big[|\varepsilon v_0+\varepsilon^2\varpi_0^\varepsilon|^{2\beta}+\Big(\int_0^T|\varepsilon v(t)+\varepsilon^2h_\varepsilon(t)|^2dt\Big)^\beta\Big]=O(\varepsilon^{2\beta}).$$

As in the proof of (3.4) in Lemma 3.2, we obtain that

$$\mathbb{E}\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}-y_1(t)\Big|^{2\beta}\Big)\to 0,\quad \varepsilon\to 0^+.$$

For any sequence $\{\varepsilon_j\}_{j=1}^\infty$ of positive numbers converging to 0 as $j\to\infty$, one can show that

$$b_{xx}(\cdot,\bar x(\cdot)+\theta\delta x^{\varepsilon_j}(\cdot),\bar u(\cdot)+\theta\delta u^{\varepsilon_j}(\cdot))-b_{xx}(\cdot)\to 0,\quad \text{in measure, as } j\to\infty.$$

Since

$$\tilde{b}^{\varepsilon_j}_{xx}(t) - \frac{1}{2}b_{xx}(t) = \int_0^1 (1-\theta)\big(b_{xx}(t, \bar{x}(t) + \theta\delta x^{\varepsilon_j}(t), \bar{u}(t) + \theta\delta u^{\varepsilon_j}(t)) - b_{xx}(t)\big)d\theta,$$

from (C3), (A.9) and the Lebesgue dominated convergence theorem, we obtain that

$$\mathbb{E}\Big(\int_0^T \Big|\big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\big)^\top \tilde{b}^{\varepsilon_j}_{xx}(t)\big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\big) - \frac{1}{2}y_1(t)^\top b_{xx}(t)y_1(t)\Big|dt\Big)^\beta \to 0, \quad \text{as } j \to \infty. \tag{A.10}$$

Similarly,

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^T \Big|2\big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\big)^\top \tilde{b}^{\varepsilon_j}_{xu}(t)\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big) - y_1(t)^\top b_{xu}(t)v(t)\Big|dt\Big)^\beta \\
&\le C\mathbb{E}\Big(\int_0^T \Big|2\big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\big)^\top \tilde{b}^{\varepsilon_j}_{xu}(t)\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big) - y_1(t)^\top b_{xu}(t)v(t)\Big|^2 dt\Big)^{\beta/2} \\
&\le C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big|^{2\beta}\Big)\Big]^{\frac{1}{2}}\Big[\mathbb{E}\Big(\int_0^T |\tilde{b}^{\varepsilon_j}_{xu}(t) - \frac{1}{2}b_{xu}(t)|^4\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|^4 dt\Big)^{\frac{\beta}{2}}\Big]^{\frac{1}{2}} \\
&\quad + C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}\Big|\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j} - y_1(t)\Big|\Big)^{2\beta}\Big]^{\frac{1}{2}}\Big[\mathbb{E}\Big(\int_0^T \Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|^4 dt\Big)^{\frac{\beta}{2}}\Big]^{\frac{1}{2}} \\
&\quad + C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}|y_1(t)|^{2\beta}\Big)\Big]^{\frac{1}{2}}\Big[\mathbb{E}\Big(\int_0^T \Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j} - v(t)\Big|^4 dt\Big)^{\frac{\beta}{2}}\Big]^{\frac{1}{2}},
\end{aligned}$$

which implies that

$$\mathbb{E}\Big(\int_0^T \Big|2\big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\big)^\top \tilde{b}^{\varepsilon_j}_{xu}(t)\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big) - y_1(t)^\top b_{xu}(t)v(t)\Big|dt\Big)^\beta \to 0, \quad j \to \infty. \tag{A.11}$$

In a similar way, we have

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^T \Big|\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big)^\top \tilde{b}^{\varepsilon_j}_{uu}(t)\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big) - \frac{1}{2}v^\top b_{uu}(t)v(t)\Big|dt\Big)^\beta \\
&\le C\mathbb{E}\Big(\int_0^T \Big|\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big)^\top\big(\tilde{b}^{\varepsilon_j}_{uu}(t) - \frac{1}{2}b_{uu}(t)\big)\big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\big)\Big|^2 dt\Big)^{\frac{\beta}{2}}
\end{aligned}$$

$$+C\mathbb{E}\Big[\int_0^T \Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}-v(t)\Big|^2\cdot\Big(\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|^2+|v(t)|^2\Big)dt\Big]^{\frac{\beta}{2}}$$

$$\le C\mathbb{E}\Big(\int_0^T \Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|^4\Big|\tilde{b}_{uu}^{\varepsilon_j}(t)-\frac{1}{2}b_{uu}(t)\Big|^2dt\Big)^{\frac{\beta}{2}}$$

$$+C\mathbb{E}\Big[\int_0^T \Big|\varepsilon_j h_{\varepsilon_j}(t)\Big|^2\cdot\Big(\Big|v(t)+\varepsilon_j h_{\varepsilon_j}(t)\Big|^2+|v(t)|^2\Big)dt\Big]^{\frac{\beta}{2}}\to 0,\quad j\to\infty. \tag{A.12}$$

Applying the above method to the diffusion coefficient σ, we conclude that

$$\mathbb{E}\Big(\int_0^T \Big|\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top\tilde{\sigma}_{xx}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}y_1(t)^\top\sigma_{xx}(t)y_1(t)\Big|^2dt\Big)^{\frac{\beta}{2}}\to 0,\quad j\to\infty, \tag{A.13}$$

$$\mathbb{E}\Big(\int_0^T \Big|2\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top\tilde{\sigma}_{xu}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-y_1(t)^\top\sigma_{xu}(t)v(t)\Big|^2dt\Big)^{\frac{\beta}{2}}\to 0,\quad j\to\infty \tag{A.14}$$

and

$$\mathbb{E}\Big(\int_0^T \Big|\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top\tilde{\sigma}_{uu}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}v(t)^\top\sigma_{uu}(t)v(t)\Big|^2dt\Big)^{\frac{\beta}{2}}\to 0,\quad j\to\infty. \tag{A.15}$$

By Lemma 3.1, and using (A.7), (A.8) and (A.10)–(A.15), we obtain that

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|r_2^{\varepsilon_j}(t)|^\beta\Big]$$

$$\le C|\varpi_0^{\varepsilon_j}-\varpi_0|^\beta+C\mathbb{E}\Big(\int_0^T \Big|b_u(t)\big(h_{\varepsilon_j}(t)-h(t)\big)\Big|dt\Big)^\beta$$

$$+C\mathbb{E}\Big(\int_0^T \Big|\sigma_u(t)\big(h_{\varepsilon_j}(t)-h(t)\big)\Big|^2dt\Big)^{\frac{\beta}{2}}$$

$$+C\mathbb{E}\Big(\int_0^T \Big|\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top\tilde{b}_{xx}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}y_1(t)^\top b_{xx}(t)y_1(t)\Big|dt\Big)^\beta$$

$$+C\mathbb{E}\Big(\int_0^T \Big|2\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top\tilde{b}_{xu}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-y_1(t)^\top b_{xu}(t)v(t)\Big|dt\Big)^\beta$$

$$
\begin{aligned}
&+ C\mathbb{E}\Big(\int_0^T \Big|\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top \tilde{b}_{uu}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big) - \frac{1}{2}v(t)^\top b_{uu}(t)v(t)\Big|dt\Big)^\beta \\
&+ C\mathbb{E}\Big(\int_0^T \Big|\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top \tilde{\sigma}_{xx}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big) - \frac{1}{2}y_1(t)^\top \sigma_{xx}(t)y_1(t)\Big|^2 dt\Big)^{\frac{\beta}{2}} \\
&+ C\mathbb{E}\Big(\int_0^T \Big|2\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top \tilde{\sigma}_{xu}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big) - y_1(t)^\top \sigma_{xu}(t)v(t)\Big|^2 dt\Big)^{\frac{\beta}{2}} \\
&+ C\mathbb{E}\Big(\int_0^T \Big|\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)^\top \tilde{\sigma}_{uu}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big) - \frac{1}{2}v(t)^\top \sigma_{uu}(t)v(t)\Big|^2 dt\Big)^{\frac{\beta}{2}} \\
&\to 0, \quad j \to \infty.
\end{aligned}
$$

This proves (4.2). The sequence $\varepsilon_j \to 0^+$ being arbitrary, the proof is complete. □

References

[1] Ch.A. Agayeva, Second order necessary conditions of optimality for stochastic systems with variable delay, Teor. Ĭmovīr. Mat. Stat. 83 (2010) 1–12, translation in: Theory Probab. Math. Statist. 83 (2011) 1–12.

[2] J.-P. Aubin, H. Frankowska, Set-Valued Analysis, Birkhäuser, Berlin, 1990.

[3] D.J. Bell, D.H. Jacobson, Singular Optimal Control Problems, Mathematics in Science and Engineering, vol. 117, Academic Press, London–New York, 1975.

[4] A. Bensoussan, Lectures on stochastic control, in: Nonlinear Filtering and Stochastic Control, in: Lecture Notes in Math., vol. 972, Springer-Verlag, Berlin, 1981, pp. 1–62.

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

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

[7] D.J. Clements, B.D.O. Anderson, Singular Optimal Control: the Linear-Quadratic Problem, Lecture Notes in Control and Information Sciences, vol. 5, Springer-Verlag, Berlin–New York, 1978.

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

[9] H. Frankowska, Some inverse mapping theorems, Ann. Inst. H. Poincaré Anal. Non Linéaire 7 (1990) 183–234.

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

[11] H. Frankowska, N. Osmolovskii, Second-order necessary optimality conditions for the Mayer problem subject to a general control constraint, in: Analysis and Geometry in Control Theory and Its Applications, in: Springer INdAM Ser., vol. 11, Springer, Cham, 2015, pp. 171–207.

[12] H. Frankowska, D. Hoehener, Jacobson type necessary optimality conditions for general control systems, in: Proceedings of 54th IEEE Conference on Decision and Control, Osaka, Japan, December 15–18, 2015, pp. 1304–1309.

[13] H. Frankowska, H. Zhang, X. Zhang, Necessary optimality conditions for weak local minima in stochastic control, in: Proceedings of NOLCOS 2016, 10th IFAC Symposium on Nonlinear Control Systems, Monterey, CA, USA, August 23–25, 2016.

[14] R.F. Gabasov, F.M. Kirillova, Singular Optimal Controls, Izdat. "Nauka", Moscow, 1973.

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

[16] U.G. Haussmann, General necessary conditions for optimal control of stochastic systems, Math. Program. Stud. 6 (1976) 30–48.

[17] D. Hoehener, Variational approach to second-order optimality conditions for control problems with pure state constraints, SIAM J. Control Optim. 50 (2012) 1139–1173.

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041

[18] M. Kisielewicz, Stochastic Differential Inclusions and Applications, Springer, 2013.
[19] H.-W. Knobloch, Higher Order Necessary Conditions in Optimal Control Theory, Lecture Notes in Computer Science, vol. 34, Springer-Verlag, Berlin–New York, 1981.
[20] A.J. Krener, The high order maximal principle and its application to singular extremals, SIAM J. Control Optim. 15 (1977) 256–293.
[21] H.J. Kushner, Necessary conditions for continuous parameter stochastic optimization problems, SIAM J. Control Optim. 10 (1972) 550–565.
[22] H. Lou, Second-order necessary/sufficient conditions for optimal control problems in the absence of linear structure, Discrete Contin. Dyn. Syst. Ser. B 14 (2010) 1445–1464.
[23] N.I. Mahmudov, A.E. Bashirov, First order and second order necessary conditions of optimality for stochastic systems, in: Statistics and Control of Stochastic Process, Moscow, 1995/1996, World Sci. Publ., Rever Edge, NJ, 1997, pp. 283–295.
[24] L. Mou, J. Yong, A variational formula for stochastic controls and some applications, Pure Appl. Math. Q. 3 (2007) 539–567.
[25] D. Nualart, The Malliavin Calculus and Related Topics, second edition, Springer-Verlag, Berlin, 2006.
[26] N.P. Osmolovskii, H. Maurer, Applications to Regular and Bang-Bang Control. Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control, SIAM, Philadelphia, PA, 2012.
[27] S. Peng, A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim. 28 (1990) 966–979.
[28] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mishchenko, The Mathematical Theory of Optimal Processes, John Wiley, New York, 1962.
[29] R.T. Rockafellar, R.J.-B. Wets, Variational Analysis, Gründlehren der Mathematischen Wissensschaften, vol. 317, Springer Verlag, New York, 1998.
[30] S. Tang, A second-order maximum principle for singular optimal stochastic controls, Discrete Contin. Dyn. Syst. Ser. B 14 (2010) 1581–1599.
[31] T. Wang, H. Zhang, Optimal control problems for forward–backward stochastic Volterra integral equations with closed control regions, arXiv:1602.05661 [math.OC].
[32] J. Yong, X.Y. Zhou, Stochastic Controls: Hamiltonian Systems and HJB Equations, Springer-Verlag, New York, Berlin, 2000.
[33] 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. 53 (2015) 2267–2296.
[34] H. Zhang, X. Zhang, Pointwise second-order necessary conditions for stochastic optimal controls. Part II: the general case, arXiv:1509.07995 [math.OC].

Please cite this article in press as: H. Frankowska et al., First and second order necessary conditions for stochastic optimal controls, J. Differential Equations (2016), http://dx.doi.org/10.1016/j.jde.2016.11.041