HAL
open science

Necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints

Hélène Frankowska, Haisen Zhang, Xu Zhang

To cite this version:

Hélène Frankowska, Haisen Zhang, Xu Zhang. Necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints. Transactions of the American Mathematical Society, 2019, 372, pp.1289-1331. hal-02126122

HAL Id: hal-02126122

https://hal.science/hal-02126122v1

Submitted on 10 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints

Hélène Frankowska[*], Haisen Zhang[†] and Xu Zhang[‡]

Abstract

The main purpose of this work is to establish some first and second order necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints. The control may affect both the drift and the diffusion terms of the systems and the control regions are allowed to be nonconvex. A stochastic inward pointing condition is proposed to ensure the normality of the corresponding necessary conditions.

Key words: Stochastic optimal control, local minimizer, necessary optimality conditions, inward pointing condition.

AMS subject classifications: Primary 93E20; Secondary 49J53, 60H10.

1. Introduction

Let $d, m, n, k \in \mathbb{N}$ and $\mathbb{R}^{n\times m}$ be the space of all $n \times m$-real matrices. For any $A \in \mathbb{R}^{n\times m}$, denote by $A^\top$ and $|A| = \sqrt{tr(AA^\top)}$ respectively the transpose and norm of A. Also, denote by $\langle\cdot,\cdot\rangle$ and $|\cdot|$ respectively the usual inner product and norm in $\mathbb{R}^n$ or $\mathbb{R}^m$, which can be identified from the context, and by $\mathcal{B}(X)$ the Borel σ-field of a metric space X.

Let $T > 0$ and $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space with the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{0\le t\le T}$ (satisfying the usual conditions), on which a d-dimensional standard Wiener process $W(\cdot) \equiv \big(W^1(\cdot), \cdots, W^d(\cdot)\big)^\top$ is defined such that $\mathbb{F}$ is the natural filtration generated by $W(\cdot)$ (augmented by all the P-null sets). We consider the following controlled stochastic differential equation

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

with the Mayer-type cost functional

$$J(x(\cdot), u(\cdot)) = \mathbb{E}\ \phi(x(T)), \tag{1.2}$$

the initial-final states constraints

$$x_0 \in K_0, \quad \mathbb{E}\ \varphi^\ell(x(T)) \le 0,\ \ell = 1, \ldots, k \tag{1.3}$$

and the state constraint

$$\mathbb{E}\ g(x(t)) \le 0, \quad \text{for all } t \in [0, T]. \tag{1.4}$$

[*]CNRS, IMJ-PRG, Sorbonne Université, case 247, 4 place Jussieu, 75252 Paris, France. The research of this author benefited from the support of the "FMJH Program Gaspard Monge in optimization and operation research", and from the support to this program from EDF under grant PGMO 2016-2832H and CNRS-NSFC PRC Project under grant 271392. *E-mail:* `helene.frankowska@imj-prg.fr`.

[†]School of Mathematics and Statistics, Southwest University, Chongqing 400715, China and CNRS, IMJ-PRG, Sorbonne Université, case 247, 4 place Jussieu, 75252 Paris, France. The research of this author is partially supported by NSF of China under grants 11401404, 11471231 and 11701470, the State Scholarship Fund of China Scholarship Council under grant [2016]3035 and the NSF of CQ CSTC under grant 2015jcyjA00017 and the Advance and Basic Research Project of Chongqing under grant cstc2016jcyjA0239. *E-mail:* `haisenzhang@yeah.net`.

[‡]School of Mathematics, Sichuan University, Chengdu 610064, China. The research of this author is partially supported by NSF of China under grants 11221101 and 11231007, the NSFC-CNRS Joint Research Project under grant 11711530142, the PCSIRT under grant IRT_16R53 and the Chang Jiang Scholars Program from the Chinese Education Ministry. *E-mail:* `zhang_xu@scu.edu.cn`.

Here $u(\cdot) \in \mathcal{U}$ is a control, $\mathcal{U}$ is the set of $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted stochastic processes with values in a closed nonempty subset U of $\mathbb{R}^m$ such that $\|u\|_2 := \left[\mathbb{E}\int_0^T |u(t)|^2 dt\right]^{\frac{1}{2}} < \infty$, $x(\cdot)$ solves (1.1), $b : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^n$, $\sigma = (\sigma^1, \ldots, \sigma^d) : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^{n\times d}$, $\phi : \mathbb{R}^n \times \Omega \to \mathbb{R}$, $\varphi^\ell : \mathbb{R}^n \to \mathbb{R}$, $\ell = 1, \ldots, k$ and $g : \mathbb{R}^n \to \mathbb{R}$ are given functions (satisfying suitable conditions to be given later), and, K_0 is a nonempty subset of $\mathbb{R}^n$. As usual, when the context is clear, we omit writing ω ($\in \Omega$) explicitly.

A state-control pair $(x(\cdot), u(\cdot))$ is called admissible if $u(\cdot) \in \mathcal{U}$ and $x(\cdot)$, the solution of (1.1) corresponding to u, satisfies (1.3) and (1.4). In this case, we call $u(\cdot)$ an admissible control. Denote by $\mathscr{P}_{ad}$ the set of all admissible pairs. In this paper, we shall consider the following optimal control problem:

$$\inf_{(x(\cdot),u(\cdot))\in\mathscr{P}_{ad}} J(x(\cdot), u(\cdot)). \tag{1.5}$$

Similarly to its counterpart in deterministic optimal control problems, we define

Definition 1.1. *An admissible pair $(\bar{x}, \bar{u}) \in L^2_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n)) \times \mathcal{U}$ is called a local minimizer for the problem (1.5) if there exists a $\delta > 0$ such that $J((x(\cdot), u(\cdot))) \geq J(\bar{x}(\cdot), \bar{u}(\cdot))$ for any $(x(\cdot), u(\cdot)) \in \mathscr{P}_{ad}$ satisfying $|\bar{x}(0) - x(0)| < \delta$ and $\|u - \bar{u}\|_2 < \delta$.*

One of the main motivations to study the optimal control problem (1.5) with state constraint (1.4) comes from its financial applications. Let us consider an investment and consumption model as follows: Suppose there are $m+1$ assets whose price processes $S_i(\cdot)$, $i = 0, 1, \ldots, m$ are described by the following differential equations:

$$\begin{cases} dS_0(t) = rS_0(t)dt, \quad t \in [0,T], \\ S_0(0) = s_0, \end{cases} \tag{1.6}$$

and, for $i = 1, 2, \ldots, m$,

$$\begin{cases} dS_i(t) = b_i S_i(t)dt + \sum\limits_{j=1}^{d} \sigma^{ij} S_i(t) dW^j(t), \quad t \in [0,T], \\ S_i(0) = s_i, \end{cases} \tag{1.7}$$

where $r \in [0, \infty)$, $s_0, s_i, b_i, \sigma^{ij} \in \mathbb{R}$, $i = 1, 2, \ldots, m$, $j = 1, \ldots, d$. Denote by $x(t)$ the total wealth of an investor at time t and by $u_i(t)$ the market value of his or her wealth in the i-th asset at time t, $i = 1, \ldots, m$. Then, x satisfies the following controlled stochastic differential equation:

$$\begin{cases} dx(t) = \left[rx(t) + c(t) + \sum\limits_{i=1}^{m} (b_i - r)u_i(t)\right]dt + \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{d} \sigma^{ij} u_i(t) dW^j(t), \quad t \in [0,T], \\ x(0) = x_0, \end{cases} \tag{1.8}$$

where x_0 is the initial wealth and $c(\cdot)$ is a suitably chosen scalar stochastic processes. In order to prohibit bankrupt the investor needs to choose the portfolio $u(\cdot) \equiv \big(u_1(\cdot), \cdots, u_m(\cdot)\big)^\top$ such that the corresponding state x satisfies

$$x(t, \omega) \geq 0, \quad \text{a.e. } (t, \omega) \in [0,T] \times \Omega. \tag{1.9}$$

Many scholars have studied this type of portfolio selection problem, see [3, 8, 11] and the references cited therein. Generally, the portfolio that only prohibits bankrupt is not well adapted to an individual investor with initial wealth $x_0 > 0$, who wishes to keep the expected return to be not lower than a proper proportion of the initial wealth for all time t, i.e., for some $\theta > 0$,

$$\mathbb{E}\, x(t) \geq \theta x_0, \quad \text{for all } t \in [0,T]. \tag{1.10}$$

Obviously, (1.10) is a special case of (1.4). Note that, under this condition, the investor still has to face the danger of the bankruptcy with some probability.

The optimal control problems with state constraints have been studied extensively in the deterministic setting, see [7, 12, 14, 18, 22, 25] and the references therein. However, compared with their deterministic counterpart, the literature for the state constrained stochastic optimal control problems is not rich. The existing results on necessary optimality conditions for stochastic optimal control problems under state constraints only focus on some special cases, as for instance when the diffusion terms are invertible with respect

to the control variable and the state constraint is represented by (1.9). When the diffusion terms are invertible with respect to the control variable, the controlled stochastic differential equation can be rewritten as a backward stochastic differential equation without control, and, by the comparison theorem (see [10], p. 22), under some further assumptions, the state constraint (1.9) is equivalent to the sample point pointwise end-point constraint

$$x(T) \geq 0, \quad \text{a.s.}$$

Then, the state constrained optimal control problem can be reformulated as a stochastic optimization problem using backward stochastic differential equations. For more details, we refer the reader to [3, 19]. Unfortunately, this approach fails whenever the state constraint (1.4) is present.

The main purpose of this paper is to provide some first and second order necessary conditions for the local minimizers of problem (1.5). The techniques of variational analysis are used to handle the nonconvexity of the control regions. Using the separation theorem, we first establish a weak maximum principle for local minimizers. Then, we derive some second order necessary conditions for critical elements of local minimizers. Also, a sufficient condition, namely a stochastic inward pointing condition, is proposed to guarantee the normality of the weak maximum principle.

Some key ideas of this paper are taken from [7, 15] and they can be used to deal with much more general classes of state constrained stochastic optimal control problems. For instance, the state constraints may be defined by finitely many inequalities and (or) the function g may depend on the time variable t. Further, one may consider some other control problems, for example, the mean field type stochastic control problems, the forward-backward stochastic control problems, or, the same control system but with the pointwise state constraint (for some nonempty subset $\mathcal{K}$ of $\mathbb{R}^n$)

$$x(t,\omega) \in \mathcal{K}, \quad \text{a.e. } (t,\omega) \in [0,T] \times \Omega. \tag{1.11}$$

Nevertheless, in this paper, we do not intend to pursue the full generality. Instead, we shall discuss such problem in a relatively simple setting making the main idea much clearer and direct. Other cases will be discussed elsewhere.

The rest of this paper is organized as follows. In Section 2, we collect some notations and introduce some preliminary results that will be used later. In Section 3, we derive the first order necessary condition for stochastic optimal controls, and, in Section 4, we establish the second order necessary conditions. Finally, in Section 5, we present a stochastic inward pointing condition to guarantee the normality of the first order necessary condition.

2. Preliminaries

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

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

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 ϑ such that $\|\vartheta\|_\beta := \left[\mathbb{E}\int_0^T |\vartheta(t,\omega)|^\beta dt\right]^{\frac{1}{\beta}} < +\infty$; 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 ϑ such that $\|\vartheta\|_{\alpha,\beta} := \left[\mathbb{E}\left(\int_0^T |\vartheta(t,\omega)|^\alpha dt\right)^{\frac{\beta}{\alpha}}\right]^{\frac{1}{\beta}} < +\infty$; 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 ϑ such that $\|\vartheta\|_{\infty,\beta} := \left[\mathbb{E}\left(\sup_{t\in[0,T]}|\vartheta(t,\omega)|^\beta\right)\right]^{\frac{1}{\beta}} < +\infty$, by $C_{\mathbb{F}}([0,T];L^\beta(\Omega;\mathbb{R}^n))$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted processes ϑ such that the map $t \mapsto \vartheta(t,\cdot)$ is continuous in $L^\beta_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ and $\|\vartheta\|_{\beta,\infty} := \sup_{t\in[0,T]}\left(\mathbb{E}\,|\vartheta(t,\omega)|^\beta\right)^{\frac{1}{\beta}} < +\infty$, by $L^\beta_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}^n))$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted processes ϑ whose paths are right-continuous with left limits and $\|\vartheta\|_{\infty,\beta} < +\infty$, by $L^\beta_{\mathbb{F}}(\Omega;BV([0,T];\mathbb{R}^n))$ the space of processes $\vartheta \in L^\beta_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}^n))$ whose sample paths have bounded

variation such that $\|\vartheta\|_{BV,\beta} := \left(\mathbb{E}\,|\vartheta(t,\omega)|^{\beta}_{BV[0,T]}\right)^{\frac{1}{\beta}} < +\infty$, by $L^{\beta}_{\mathbb{F}}(\Omega; BV_0([0,T];\mathbb{R}^n))$ the space of processes $\vartheta \in L^{\beta}_{\mathbb{F}}(\Omega; BV([0,T];\mathbb{R}^n))$ with $\vartheta(0)=0$, and, by $L^{\infty}_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable, $\mathbb{F}$-adapted processes ϑ such that $\|\vartheta\|_{\infty} := \text{ess sup}_{(t,\omega)\in[0,T]\times\Omega}|\vartheta(t,\omega)| < +\infty$. When every sample path of a process ϑ has left limits, we denote by ϑ_- its left continuous modification.

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, $\mathbb{F}$-adapted process has an $\mathbb{F}$-progressively measurable modification (see [27, Proposition 2.8, p. 17]).

Next, we recall some concepts and results from the set-valued analysis. We refer the reader to [1] for more details.

In the sequel, we assume that X is a Banach space with a norm $\|\cdot\|_X$ and the dual X^*. For any subset $K\subset X$, denote by ∂K, $intK$, clK and coK its boundary, interior, closure and convex hull, respectively. K is called a cone if $\alpha x\in K$ for any $\alpha>0$ and $x\in K$. For a cone K, the convex closed cone $K^- := \{\xi\in X^* \mid \xi(x)\le 0,\ \forall\, x\in K\}$ is called the dual cone (or negative polar cone) of K. Define the distance between a point $x\in X$ and K by $dist\,(x,K) := \inf\limits_{y\in K}\|y-x\|_X$.

Definition 2.1. *For $x\in K$, the Clarke tangent cone $\mathcal{C}_K(x)$ to K at x is defined by*

$$\mathcal{C}_K(x) := \Big\{v\in X \;\Big|\; \lim_{\substack{\varepsilon\to 0^+\\ y\in K,\ y\to x}} \frac{dist\,(y+\varepsilon v, K)}{\varepsilon} = 0\Big\},$$

the adjacent cone $T^b_K(x)$ to K at x is defined by

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

It is well known that $\mathcal{C}_K(x)$ is a closed convex cone in X and $\mathcal{C}_K(x)\subset T^b_K(x)$. By the definition of $T^b_K(x)$, $v\in T^b_K(x)$ 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 $x+\varepsilon v_\varepsilon\in K$. Equivalently, $v\in T^b_K(x)$ if and only if for any $\varepsilon>0$ there exists an $\eta(\varepsilon)\in X$ with $\|\eta(\varepsilon)\|_X = o(\varepsilon)$ $(\varepsilon\to 0^+)$ such that $x+\varepsilon v+\eta(\varepsilon)\in K$. When K is convex, for any $x\in K$, $T^b_K(x)=\mathcal{C}_K(x)=cl\{\alpha(y-x)\,|\,\alpha\ge 0,\ y\in K\}$.

Example 2.1. *Let $\varphi^\ell:\mathbb{R}^n\to\mathbb{R}$, $\ell=1,...,k$ be continuously differentiable with globally Lipschitz derivatives. Consider the set K_T defined by*

$$K_T := \big\{\xi\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)\,\big|\,\mathbb{E}\ \varphi^\ell(\xi)\le 0,\ \ell=1,\dots,k\big\}. \tag{2.1}$$

Let $\xi\in K_T$ and define $I(\xi) := \big\{\ell\in\{1,\dots,k\}\,|\,\mathbb{E}\ \varphi^\ell(\xi)=0\big\}$. If $I(\xi)=\emptyset$, then $T^b_{K_T}(\xi)=L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$.

If $I(\xi)\neq\emptyset$ and there exists a set $A\in\mathcal{F}_T$ with $P(A)>0$ such that

$$Z(\omega) := \big\{z\in\mathbb{R}^n \;\big|\; \big\langle\varphi^\ell_x(\xi(\omega)),z\big\rangle<0,\ \forall\,\ell\in I(\xi)\big\}\neq\emptyset,\ \ a.s.\ \ in\ \ A, \tag{2.2}$$

then, $int\big(T^b_{K_T}(\xi)\big)\neq\emptyset$ and

$$T^b_{K_T}(\xi) = \big\{v\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)\,\big|\,\mathbb{E}\,\big\langle\varphi^\ell_x(\xi),v\big\rangle\le 0,\ \forall\,\ell\in I(\xi)\big\}.$$

Proof. Let $v\in T^b_{K_T}(\xi)$. Then, from the definition of the adjacent cone, for any $\varepsilon>0$, there exists a $v_\varepsilon\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ such that v_ε converges to v in $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ as $\varepsilon\to 0^+$ and $\xi+\varepsilon v_\varepsilon\in K_T$. Therefore, for any $\ell\in I(\xi)$,

$$0\ge\mathbb{E}\ \varphi^\ell(\xi+\varepsilon v_\varepsilon) = \mathbb{E}\ \varphi^\ell(\xi)+\varepsilon\mathbb{E}\,\big\langle\varphi^\ell_x(\xi),v_\varepsilon\big\rangle+o(\varepsilon) = \varepsilon\mathbb{E}\,\big\langle\varphi^\ell_x(\xi),v_\varepsilon\big\rangle+o(\varepsilon)$$

Dividing by ε the both sides of the above inequality and letting $\varepsilon\to 0^+$, we obtain that

$$\mathbb{E}\,\big\langle\varphi^\ell_x(\xi),v\big\rangle\le 0.$$

On the other hand, by the condition (2.2) and by similar arguments to those in [16, Proof of Lemma 3.3], there exists a $\bar v\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ such that

$$\mathbb{E}\,\big\langle\varphi^\ell_x(\xi),\bar v\big\rangle<0,\quad\forall\,\ell\in I(\xi). \tag{2.3}$$

From the definition of the adjacent cone, it is easy to verify that $\bar{v} \in int\big(T^b_{K_T}(\xi)\big)$. Let $v \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ be such that, for any $\ell \in I(\xi)$, $\mathbb{E}\left\langle \varphi^\ell_x(\xi), v\right\rangle \leq 0$. Then $\mathbb{E}\left\langle \varphi^\ell_x(\xi), v^\lambda\right\rangle < 0$ for all $\ell \in I(\xi)$, where $\lambda \in (0,1)$ and $v^\lambda := (1-\lambda)v + \lambda\bar{v}$. It implies that there exists a $\rho > 0$ such that for all sufficiently small $\varepsilon > 0$,

$$\mathbb{E}\ \varphi^\ell(\xi + \varepsilon v^\lambda) = \mathbb{E}\ \varphi^\ell(\xi) + \varepsilon\mathbb{E}\left\langle \varphi^\ell_x(\xi), v^\lambda\right\rangle + o(\varepsilon) < -\varepsilon\rho + o(\varepsilon) \leq 0, \quad \forall\, \ell \in I(\xi)$$

and

$$\mathbb{E}\ \varphi^\ell(\xi + \varepsilon v^\lambda) = \mathbb{E}\ \varphi^\ell(\xi) + \varepsilon\mathbb{E}\left\langle \varphi^\ell_x(\xi), v^\lambda\right\rangle + o(\varepsilon) < -\rho + \varepsilon\mathbb{E}\left\langle \varphi^\ell_x(\xi), v^\lambda\right\rangle + o(\varepsilon) \leq 0, \quad \forall\, \ell \notin I(\xi).$$

This proves that $v^\lambda \in T^b_{K_T}(\xi)$ for any $\lambda \in (0,1)$. Since $T^b_{K_T}(\xi)$ is closed, letting $\lambda \to 0$, we have $v \in T^b_{K_T}(\xi)$. □

Observe that the proof in the above example can be used to show that, under the condition (2.2), $\mathcal{C}_{K_T}(\xi) = T^b_{K_T}(\xi)$ when K_T is represented by (2.1).

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

$$T^{b(2)}_K(x,v) := \Big\{h \in X \ \Big|\ \lim_{\varepsilon\to 0^+} \frac{dist\,(x + \varepsilon v + \varepsilon^2 h, K)}{\varepsilon^2} = 0\Big\}.$$

Thus $h \in T^{b(2)}_K(x,v)$ 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. *It is not difficult to check that $v \in T^{b(2)}_K(x,0)$ if and only if $v \in T^b_K(x)$. Moreover, it follows from [12, Lemma 2.4] that, if $T^{b(2)}_K(x,v) \neq \emptyset$, then $\mathcal{C}_K(x) + T^{b(2)}_K(x,v) = T^{b(2)}_K(x,v)$. In particular, $\mathcal{C}_K(x) + T^b_K(x) = T^b_K(x)$.*

The dual cone of the Clarke tangent cone $\mathcal{C}_K(x)$, denoted by $N^C_K(x)$, is called the Clarke normal cone to K at x, i.e.,

$$N^C_K(x) := \Big\{\xi \in X^* \ \Big|\ \xi(v) \leq 0,\ \forall\, v \in \mathcal{C}_K(x)\Big\}.$$

When K is convex, $N^C_K(x)$ is the normal cone $N_K(x) := \Big\{\xi \in X^* \ \Big|\ \xi(y-x) \leq 0,\ \forall\, y \in K\Big\}$ of the convex analysis.

The following three elementary lemmas will be useful in the sequel.

Lemma 2.1. *([16, Lemma 2.4]) Let $K_1, \ldots, K_k$ be convex cones in X such that $\bigcap_{i=1}^k intK_i \neq \emptyset$. Then for any convex cone K_0 such that $K_0 \bigcap \big(\bigcap_{i=1}^k intK_i\big) \neq \emptyset$, we have $\big(\bigcap_{i=0}^k K_i\big)^- = \sum_{i=0}^k K_i^-$.*

Lemma 2.2. *([16, Lemma 2.5]) Let H be a Hilbert space (with an inner product $\langle\cdot,\cdot\rangle_H$), and K be a nonempty closed polyhedra in H, i.e., for some $\{a_1, \cdots, a_k\} \subset H \setminus \{0\}$ and $\{b_1, \cdots, b_k\} \subset \mathbb{R}$,*

$$K := \big\{x \in H \mid \langle a_i, x\rangle_H + b_i \leq 0,\ \forall\, i = 1, \cdots, k\ \big\}.$$

If $0 \neq \xi \in H$ satisfies $\sup_{x\in K}\langle \xi, x\rangle_H < +\infty$, then, this supremum is attained at some $\bar{x} \in \partial K$ and $\xi \in \sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i$, where

$$I(\bar{x}) := \big\{i \in \{1, \cdots, k\} \ \big|\ \langle a_i, \bar{x}\rangle_H + b_i = 0\big\}.$$

Lemma 2.3 ([6])**.** *Let X be a Banach space and $M_0, M_1, \ldots, M_k$ be nonempty convex subsets of X such that M_i is open for all $i \in \{1, \cdots, k\}$. Then*

$$M_0 \cap M_1 \cap \ldots \cap M_k = \emptyset \tag{2.4}$$

if and only if there are $x_0^*, x_1^*, \ldots, x_k^* \in X^*$, not vanishing simultaneously, such that

$$x_0^* + x_1^* + \ldots + x_k^* = 0, \quad \inf x_0^*(M_0) + \inf x_1^*(M_1) + \ldots + \inf x_k^*(M_k) \geq 0, \tag{2.5}$$

where $\inf x_j^*(M_j) := \inf\{x_j^*(x) \mid x \in M_j\}$ for $j \in \{0, 1, \cdots, k\}$.

Furthermore, if (2.5) holds true and for some $i \in \{0, \ldots, k\}$, there is a nonempty cone $C_i \subset X$ and $x_i \in X$ such that $x_i + C_i \subset M_i$, then $-x_i^* \in C_i^-$.

A very short proof of the above lemma can be found in [13].

Remark 2.2. *Lemma 2.3 implies that: if (2.5) holds true for a nontrivial family $x_0^*, x_1^*, \ldots, x_k^* \in X^*, k \geq 2$, and $M_0 \cap M_2 \cap \ldots \cap M_k \neq \emptyset$, then $x_1^* \neq 0$.*

Let $(\Xi, \mathscr{G})$ be a measurable space, Y be a complete separable metric space and $F : \Xi \rightsquigarrow Y$ be a set-valued map. The inverse F^{-1} of F is the set-valued map from Y to Ξ defined by

$$F^{-1}(y) := \big\{\xi \in \Xi \,\big|\, y \in F(\xi)\big\}, \qquad \forall\, y \in Y.$$

Recall that F is called measurable if $F^{-1}(A) := \{\xi \in \Xi \mid F(\xi) \cap A \neq \emptyset\} \in \mathscr{G}$ for any $A \in \mathcal{B}(Y)$. We shall need the following known result.

Lemma 2.4. *([1, Theorem 8.2.11]) Let $(\Xi, \mathscr{G}, \mu)$ be a complete σ-finite measure space, Y be a complete separable metric space, F be a measurable set-valued map from Ξ to Y with nonempty closed images, and, f be a Carathéodory function from $\Xi \times Y$ to $\mathbb{R}$. Then, the set-valued map*

$$M(\xi) := \big\{y \in F(\xi) \,\big|\, f(\xi, y) = \inf_{z \in F(\xi)} f(\xi, z)\big\},\ \forall\, \xi \in \Xi$$

is a measurable set-valued map (from Ξ to Y).

Combining [1, Theorem 8.5.1] with [16, Lemma 2.6], we have:

Lemma 2.5. *Suppose $(\Xi, \mathscr{G}, \mu)$ is a complete finite measure space, X is a separable Banach space, $p \geq 1$ 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 maps $\mathcal{C}_K(\varphi(\cdot))$: $\xi \rightsquigarrow \mathcal{C}_K(\varphi(\xi))$ and $T_K^b(\varphi(\cdot))$: $\xi \rightsquigarrow T_K^b(\varphi(\xi))$ are measurable, and

$$\begin{aligned}
&\big\{v(\cdot) \in L^p(\Xi, \mathscr{G}, \mu; X) \,\big|\, v(\xi) \in \mathcal{C}_K(\varphi(\xi)),\ \mu\text{–a.e. } \xi \in \Xi\big\} \\
&\subset \big\{v(\cdot) \in L^p(\Xi, \mathscr{G}, \mu; X) \,\big|\, v(\xi) \in T_K^b(\varphi(\xi)),\ \mu\text{–a.e. } \xi \in \Xi\big\} \\
&\subset T_{\mathcal{K}}^b(\varphi(\cdot)).
\end{aligned} \tag{2.6}$$

As in [20], we call a measurable set-valued map $\zeta : (\Omega, \mathcal{F}) \rightsquigarrow \mathbb{R}^m$ a set-valued random variable. We call a map $\Gamma : [0, T] \times \Omega \rightsquigarrow \mathbb{R}^m$ a measurable set-valued stochastic process if Γ is $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable, and, we say that Γ is $\mathbb{F}$-adapted if $\Gamma(t)$ is $\mathcal{F}_t$-measurable for any $t \in [0, T]$. Define

$$\mathscr{A} := \big\{A \in \mathcal{B}([0, T]) \otimes \mathcal{F} \,\big|\, A_t \in \mathcal{F}_t,\ \forall\, t \in [0, T]\big\}, \tag{2.7}$$

where $A_t := \{\omega \in \Omega \mid (t, \omega) \in A\}$ is the section of A. Obviously, $\mathscr{A}$ is a σ-subalgebra of $\mathcal{B}([0, T]) \otimes \mathcal{F}$. As pointed in [20, p. 96], the following result holds.

Lemma 2.6. *A set-valued stochastic process $\Gamma : [0, T] \times \Omega \rightsquigarrow \mathbb{R}^m$ is $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted if and only if Γ is $\mathscr{A}$-measurable.*

The following result will play a key role in the sequel.

Lemma 2.7. *Let p, $q \in (1, \infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then, for any bounded linear functional Λ on the Banach space $L^p_{\mathbb{F}}(\Omega; C([0, T]; \mathbb{R}^n))$, there exists a process $\psi \in L^q_{\mathbb{F}}(\Omega; BV_0([0, T]; \mathbb{R}^n))$ such that*

$$\Lambda(x(\cdot)) = \mathbb{E} \int_0^T \langle x(t), d\psi(t) \rangle, \quad \forall\, x(\cdot) \in L^p_{\mathbb{F}}(\Omega; C([0, T]; \mathbb{R}^n)), \tag{2.8}$$

and

$$\|\Lambda\|_{L^p_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n))^*} \leq \|\psi\|_{L^q_{\mathbb{F}}(\Omega; BV([0,T]; \mathbb{R}^n))}. \tag{2.9}$$

Proof. Clearly $L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ is a linear subspace of $L^p_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}^n))$. For $\Lambda\in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))^*$, by the Hahn-Banach theorem, one can find an extension $\widetilde{\Lambda}\in L^p_{\mathbb{F}}(\Omega;\ D([0,T];\mathbb{R}^n))^*$ such that

$$||\widetilde{\Lambda}||_{L^p_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}^n))^*}=||\Lambda||_{L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))^*} \tag{2.10}$$

and

$$\widetilde{\Lambda}(x(\cdot))=\Lambda(x(\cdot)),\qquad \forall\, x(\cdot)\in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n)). \tag{2.11}$$

From the proof of [9, Theorem 65, p. 254], we deduce that, there exist two processes $\psi^+(\cdot)$ and $\psi^-(\cdot)$ with bounded variation such that $\psi^+(\cdot)$ is optional and purely discontinuous, $\psi^-(\cdot)$ is predictable with $\psi^-(0)=0$,

$$\mathbb{E}\left|\int_{(0,T]}d|\psi^-(t)|+\int_{[0,T)}d|\psi^+(t)|\right|^2<\infty$$

and, for any $x(\cdot)\in L^p_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}^n))$,

$$\widetilde{\Lambda}(x(\cdot))=\mathbb{E}\left[\int_{(0,T]}\big\langle x_-(t),d\psi^-(t)\big\rangle+\int_{[0,T)}\big\langle x(t),d\psi^+(t)\big\rangle\right].$$

Define $\psi^*:=\psi^-+\psi^+$. By (2.11), we have

$$\Lambda(x(\cdot))=\mathbb{E}\int_0^T\langle x(t),d\psi^*(t)\rangle,\qquad \forall\, x(\cdot)\in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n)).$$

Letting $\psi=\psi^*-\psi^*(0)$, we obtain (2.8). (2.9) follows from (2.8). □

Let $p=q=2$. Obviously, for any $\psi\in L^2_{\mathbb{F}}(\Omega;BV_0([0,T];\mathbb{R}^n))$,

$$x(\cdot)\mapsto\mathbb{E}\int_0^T\langle x(t),d\psi(t)\rangle,\quad \forall\, x(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$$

defines a bounded linear functional on $L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$. Due to this and in order to simplify the notation, in this paper we identify the process $\psi\in L^2_{\mathbb{F}}(\Omega;BV_0([0,T];\mathbb{R}^n))$ with the above bounded linear functional, and for a convex cone $K\subset L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$, we say $\psi\in K^-$ if

$$\mathbb{E}\int_0^T\langle x(t),d\psi(t)\rangle\le 0,\quad \forall\, x(\cdot)\in K.$$

By Itô's formula for discontinuous semimartingales (see [24, Theorem 33]), we have the following result.

Lemma 2.8. *Let $\psi\in L^2_{\mathbb{F}}(\Omega;BV([0,T];\mathbb{R}^n))$ and $z(\cdot)=z(0)+\int_0^\cdot\eta(t)dt+\sum_{j=1}^d\int_0^\cdot\vartheta^j(t)dW^j(t)$ with $z(0)\in\mathbb{R}^n$, η, $\vartheta^j\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^n))$, $j=1,\dots,d$. Then,*

$$\mathbb{E}\ \langle z(T),\psi(T)\rangle=\mathbb{E}\ \langle z(0),\psi(0)\rangle+\mathbb{E}\int_0^T\langle z(t),d\psi(t)\rangle+\mathbb{E}\int_0^T\langle\psi(t),\eta(t)\rangle\,dt.$$

Proof. Since $\psi\in L^2_{\mathbb{F}}(\Omega;BV([0,T];\mathbb{R}^n))$, we have $\big\langle\psi_-(\cdot),\vartheta^j(\cdot)\big\rangle\in L^2(0,T;\mathbb{R})$ a.s., for any $j=1,\dots,d$. Let

$$A_n(\omega):=\Big\{t\in[0,T]\ \Big|\ \sum_{j=1}^d\int_0^t|\,\big\langle\psi_-(s,\omega),\vartheta^j(s,\omega)\big\rangle\,|^2ds>n\Big\}.$$

Define

$$\tau_n(\omega):=\begin{cases}\inf A_n(\omega), & \text{if } A_n(\omega)\neq\emptyset,\\ T, & \text{if } A_n(\omega)=\emptyset,\end{cases}\qquad \omega\in\Omega$$

and

$$z_n(\cdot)=z(0)+\int_0^\cdot\eta(t)dt+\sum_{j=1}^d\int_0^\cdot\vartheta^j(t)\chi_{[0,\tau_n]}(t)dW^j(t).$$

By [24, Theorem 33], we have

$$\begin{aligned}&\langle z_n(T),\psi(T)\rangle-\langle z(0),\psi(0)\rangle\\=&\int_0^T\langle z_n(t),d\psi(t)\rangle+\int_0^T\langle\psi_-(t),\eta(t)\rangle\,dt+\sum_{j=1}^d\int_0^T\left\langle\psi_-(t),\vartheta^j(t)\right\rangle\chi_{[0,\tau_n]}(t)dW^j(t),\ a.s.\end{aligned}$$

Obviously, by the property of the Lebesgue integral, $\int_0^T\langle\psi_-(t),\eta(t)\rangle\,dt=\int_0^T\langle\psi(t),\eta(t)\rangle\,dt$, a.s. In addition, since $\sum_{j=1}^d\mathbb{E}\int_0^T|\left\langle\psi_-(t),\vartheta^j(t)\right\rangle\chi_{[0,\tau_n]}(t)|^2dt\le n$, by the property of Itô's integral we have

$$\sum_{j=1}^d\mathbb{E}\int_0^T\left\langle\psi_-(t),\vartheta^j(t)\right\rangle\chi_{[0,\tau_n]}(t)dW^j(t)=0.$$

Therefore,

$$\mathbb{E}\,\langle z_n(T),\psi(T)\rangle-\langle z(0),\psi(0)\rangle=\mathbb{E}\int_0^T\langle z_n(t),d\psi(t)\rangle+\mathbb{E}\int_0^T\langle\psi(t),\eta(t)\rangle\,dt.\tag{2.12}$$

Obviously, by Lebesgue's dominated convergence theorem,

$$\mathbb{E}\int_0^T|\vartheta^j(t)\chi_{[0,\tau_n]}(t)-\vartheta^j(t)|^2dt\to0,\quad n\to\infty,\ j=1,\dots,d,$$

which implies that $\mathbb{E}\sup_{t\in[0,T]}|z_n(t)-z(t)|^2\to0$ as $n\to\infty$. Then, the desired conclusion follows by passing to the limit on both sides of (2.12). □

We end this section by the following simple result (which is certainly known but we could not find an exact reference).

Lemma 2.9. *$L^\beta_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))\subset C_{\mathbb{F}}([0,T];L^\beta(\Omega;\mathbb{R}^n))$ for any $\beta\in[1,\infty)$.*

Proof. Although the proof of this lemma is obvious, for the sake of completeness, we give below the details. Let $z\in L^\beta_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$. We only need to prove that $t\mapsto z(t)$ is a continuous function from $[0,T]$ to the Banach space $L^\beta_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$.

Let $t\in[0,T]$ and $t_n\in[0,T]$, $t_n\to t$ as $n\to\infty$. Since $z\in L^\beta_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$, $z(t_n,\omega)\to z(t,\omega)$ a.s. as $n\to\infty$. On the other hand, $\mathbb{E}\ |z(t_n,\omega)|^\beta\le\mathbb{E}\ \sup_{t\in[0,T]}|z(t,\omega)|^\beta<+\infty$. By Lebesgue's dominated convergence theorem, $\mathbb{E}\ |z(t_n,\omega)-z(t,\omega)|^\beta\to0$ as $n\to\infty$. □

3. First order necessary conditions

In this section, we study the first order necessary optimality conditions for the optimal control problem (1.5).

We need the following assumptions:

(A1) *The control region U is nonempty and closed in $\mathbb{R}^m$.*

(A2) *The functions b, σ, ϕ, g and φ^ℓ, $\ell=1,\dots,k$ satisfy the following:*

(i) *For any $(x,u)\in\mathbb{R}^n\times\mathbb{R}^m$, $b(\cdot,x,u,\cdot):\ [0,T]\times\Omega\to\mathbb{R}^n$ and $\sigma^j(\cdot,x,u,\cdot):\ [0,T]\times\Omega\to\mathbb{R}^n$ $(j=1,\dots,d)$ 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 $\sigma^j(t,\cdot,\cdot,\omega):\ \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ are differentiable and*

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

$$(x,u)\mapsto(\sigma^j_x(t,x,u,\omega),\sigma^j_u(t,x,u,\omega)),\ j=1,\dots,d$$

are uniformly continuous in $x\in\mathbb{R}^n$ and $u\in\mathbb{R}^m$. There exists a constant $L\ge0$ and a nonnegative $\eta\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}))$ such that for a.e. $(t,\omega)\in[0,T]\times\Omega$ and for any $x\in\mathbb{R}^n$ and $u\in\mathbb{R}^m$,

$$\begin{cases}|b(t,0,u,\omega)|+|\sigma^j(t,0,u,\omega)|\le L(\eta(t,\omega)+|u|),\\|b_x(t,x,u,\omega)|+|b_u(t,x,u,\omega)|\le L,\\|\sigma^j_x(t,x,u,\omega)|+|\sigma^j_u(t,x,u,\omega)|\le L,\ j=1,\dots,d;\end{cases}$$

(ii) *For any $x \in \mathbb{R}^n$, the random variable $\phi(x,\cdot)$ is $\mathcal{F}_T$-measurable, $\phi(\cdot,\omega):\ \mathbb{R}^n \to \mathbb{R}$ is differentiable a.s., and there exists a nonnegative $\eta_T \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R})$ such that for any $x,\ \tilde{x} \in \mathbb{R}^n$,*

$$\begin{cases} |\phi(x,\omega)| \le L(\eta_T(\omega)^2 + |x|^2), \quad |\phi_x(0,\omega)| \le L\eta_T(\omega),\ a.s., \\ |\phi_x(x,\omega) - \phi_x(\tilde{x},\omega)| \le L|x-\tilde{x}|,\ a.s. \end{cases}$$

(iii) *g and φ^ℓ, $\ell = 1,\dots,k$ are differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$, and, for any $x,\ \tilde{x} \in \mathbb{R}^n$,*

$$\begin{cases} |g(x)| + \sum_{\ell=1}^k |\varphi^\ell(x)| \le L(1+|x|^2), \\ |g_x(0)| + \sum_{\ell=1}^k |\varphi^\ell_x(0)| \le L, \\ |g_x(x) - g_x(\tilde{x})| + \sum_{\ell=1}^k |\varphi^\ell_x(x) - \varphi^\ell_x(\tilde{x})| \le L|x-\tilde{x}|. \end{cases}$$

Moreover we assume that, $g(x_0) < 0$ for every $x_0 \in K_0$.

When the conditions (i) and (ii) in (A2) are satisfied, the state x (of (1.1)) is uniquely defined by any given initial datum $x_0 \in \mathbb{R}^n$ and control $u \in \mathcal{U}$, and the cost functional (1.2) is well-defined for $u(\cdot) \in \mathcal{U}$. In what follows, C represents a generic positive constant (depending only on T, $\eta(\cdot)$, $\eta_T(\cdot)$ and L), which may differ from one place to another.

Let $(\bar{x},\bar{u})$ be a local minimizer and $\bar{x}_0 := \bar{x}(0)$. For $f = b,\ \sigma$, denote

$$f_x[t] = f_x(t,\bar{x}(t),\bar{u}(t)), \quad f_u[t] = f_u(t,\bar{x}(t),\bar{u}(t)).$$

Let Φ be a set-valued stochastic process satisfying

$$\begin{cases} \Phi \text{ is } \mathcal{B}([0,T]) \otimes \mathcal{F}\text{-measurable and } \mathbb{F}\text{-adapted;} \\ \text{for a.e. } (t,\omega) \in [0,T]\times\Omega, \Phi(t,\omega) \text{ is a nonempty closed convex cone in } \mathbb{R}^m; \\ \Phi(t,\omega) \subset T^b_U(\bar{u}(t,\omega)),\ \text{a.e. } (t,\omega) \in [0,T]\times\Omega. \end{cases} \tag{3.1}$$

Define

$$\mathcal{T}_\Phi(\bar{u}) := \Big\{ v(\cdot) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m)) \,\Big|\, v(t,\omega) \in \Phi(t,\omega),\ \text{a.e. } (t,\omega) \in [0,T]\times\Omega \Big\}.$$

Since $0 \in \mathcal{T}_\Phi(\bar{u})$, $\mathcal{T}_\Phi(\bar{u})$ is nonempty. Clearly it is a closed convex cone in $L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$. By Lemma 2.5, we may chose $\Phi(t,\omega) = \mathcal{C}_U(\bar{u}(t,\omega))$. However, in general, there may exist a $\Phi(t,\omega)$ as above such that $\mathcal{C}_U(\bar{u}(t,\omega)) \subsetneq \Phi(t,\omega) \subset T^b_U(\bar{u}(t,\omega))$.

Let $v \in \mathcal{T}_\Phi(\bar{u})$ and $\nu_0 \in T^b_{K_0}(\bar{x}_0)$. We consider the following first order linearized stochastic control system:

$$\begin{cases} dy_1(t) = \big(b_x[t]y_1(t) + b_u[t]v(t)\big)dt + \sum_{j=1}^d \big(\sigma^j_x[t]y_1(t) + \sigma^j_u[t]v(t)\big)dW^j(t), \quad t \in [0,T], \\ y_1(0) = \nu_0. \end{cases} \tag{3.2}$$

It is easy to see that, under the assumption (i) in (A2), for any $v \in \mathcal{T}_\Phi(\bar{u})$ and $\nu_0 \in T^b_{K_0}(\bar{x}_0)$, (3.2) admits a unique solution $y_1(\cdot) \in L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$.

By Lemma 2.5, $\mathcal{T}_\Phi(\bar{u}) \subset T^b_{\mathcal{U}}(\bar{u})$. For any $\varepsilon > 0$, choose $\nu^\varepsilon_0 \in \mathbb{R}^n$ and $v_\varepsilon \in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$ such that $x^\varepsilon_0 := \bar{x}_0 + \varepsilon\nu^\varepsilon_0 \in K_0$, $u^\varepsilon := \bar{u} + \varepsilon v_\varepsilon \in \mathcal{U}$ and $\nu^\varepsilon_0 \to \nu_0$ in $\mathbb{R}^n$, $v_\varepsilon \to v$ in $L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^m))$ as $\varepsilon \to 0^+$. Let x^ε be the solution of (1.1) corresponding to the control u^ε and the initial datum x^ε_0, and put

$$\delta x^\varepsilon = x^\varepsilon - \bar{x}, \qquad r^\varepsilon_1(t,\omega) := \frac{\delta x^\varepsilon(t,\omega)}{\varepsilon} - y_1(t,\omega).$$

The following results for the d-dimensional Wiener process can be proved in the same way as [15, Lemma 3.2] for the one-dimensional Wiener process.

Lemma 3.1. *Let (i) and (ii) in (A2) hold. Then, for any $\beta \ge 2$,*

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

b) $\|r_1^\varepsilon\|_{\infty,\beta} \to 0,$ *as* $\varepsilon \to 0^+$.

Let $\mathcal{T}_{K_0}(\bar{x}_0)$ be a nonempty closed convex cone contained in $T^b_{K_0}(\bar{x}_0)$. Denote

$$\mathcal{G}^{(1)} := \big\{ y_1(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\big|\, y_1 \text{ solves (3.2) with } v \in \mathcal{T}_\Phi(\bar{u}) \text{ and } \nu_0 \in \mathcal{T}_{K_0}(\bar{x}_0) \big\}, \tag{3.3}$$

$$I_0^g := \{ t \in [0,T] \,|\, \mathbb{E} g(\bar{x}(t)) = 0 \}, \tag{3.4}$$

and

$$\mathcal{I}_0^\varphi := \big\{ \ell \in \{1,\dots,k\} \,|\, \mathbb{E}\varphi^\ell(\bar{x}(T)) = 0 \big\}.$$

Consider the sets

$$\mathcal{Q}^{(1)} := \big\{ z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\big|\, \mathbb{E}\,\langle g_x(\bar{x}(t)), z(t)\rangle < 0,\ \forall\, t \in I_0^g \big\}, \tag{3.5}$$

and

$$\mathcal{E}^{(1)} := \big\{ z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\big|\, \mathbb{E}\,\big\langle \varphi_x^\ell(\bar{x}(T)), z(T)\big\rangle < 0,\ \forall\, \ell \in \mathcal{I}_0^\varphi \big\}. \tag{3.6}$$

When $I_0^g = \emptyset$ (*resp.* $\mathcal{I}_0^\varphi = \emptyset$) we set $\mathcal{Q}^{(1)} = L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$ (*resp.* $\mathcal{E}^{(1)} = L^2_{\mathbb{F}}(\Omega;\, C([0,T];\mathbb{R}^n))$).

Also, we define

$$\mathcal{L}^{(1)} := \big\{ z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\big|\, \mathbb{E}\,\langle \phi_x(\bar{x}(T)), z(T)\rangle < 0 \big\}. \tag{3.7}$$

Since $\mathcal{T}_\Phi(\bar{u})$ and $\mathcal{T}_{K_0}(\bar{x}_0)$ are nonempty convex cones, $\mathcal{G}^{(1)}$ is a nonempty convex cone in $L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$.

Notice that $\mathcal{E}^{(1)}$ and $\mathcal{L}^{(1)}$ are, possibly empty, open convex cones in $L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$. In Example 2.1 we have found that when the condition (2.2) (with ξ replaced by $\bar{x}(T)$) holds true, there exists a $z_T \in L^2_{\mathcal{F}_T}(\Omega)$ such that $\mathbb{E}\,\big\langle \varphi_x^\ell(\bar{x}(T)), z_T\big\rangle < 0$ for every $\ell \in \mathcal{I}_0^\varphi$.

By the Itô representation theorem (see [21, Theorem 4.3.3, p.51]), there exist $\vartheta^j \in L^2_{\mathbb{F}}(\Omega;\, L^2(0,T;\mathbb{R}^n))$, $j = 1,\dots,d$ such that

$$z_T = \mathbb{E}\ z_T + \sum_{j=1}^d \int_0^T \vartheta^j(t) dW^j(t), \quad a.s.$$

Then, $z(\cdot) := \mathbb{E}\ z_T + \sum_{j=1}^d \int_0^\cdot \vartheta^j(t) dW^j(t)$ belongs to $\mathcal{E}^{(1)}$. Similarly, when $\phi_x(\bar{x}(T)) \neq 0$ on a set $A \in \mathcal{F}_T$ with $P(A) > 0$, $\mathcal{L}^{(1)} \neq \emptyset$. Furthermore, when $\mathcal{Q}^{(1)}$ and $\mathcal{E}^{(1)}$ are nonempty sets,

$$cl_{\infty,2}\big(\mathcal{Q}^{(1)}\big) = \{ z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\big|\, \mathbb{E}\ \langle g_x(\bar{x}(t)), z(t)\rangle \le 0,\ \forall\, t \in I_0^g \},$$

and

$$cl_{\infty,2}\big(\mathcal{E}^{(1)}\big) = \big\{ z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\big|\, \mathbb{E}\,\big\langle \varphi_x^\ell(\bar{x}(T)), z(T)\big\rangle \le 0,\ \forall\, \ell \in \mathcal{I}_0^\varphi \big\}.$$

Lemma 3.2. $\mathcal{Q}^{(1)}$ *is an open convex cone in* $L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$.

Proof. It is sufficient to prove that $\mathcal{Q}^{(1)}$ is open when it is nonempty.

Let $z \in \mathcal{Q}^{(1)}$. Since $\bar{x} \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$, by Lemma 2.9, I_0^g is a compact subset of $[0,T]$. Hence, there exists a constant $\rho > 0$ such that $\mathbb{E}\,\langle g_x(\bar{x}(t)), z(t)\rangle < -\rho$ for every $t \in I_0^g$ and for some $\delta > 0$ and for any $\eta \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$ with $\|\eta\|_{\infty,2} \le \delta$,

$$\mathbb{E}\ \langle g_x(\bar{x}(t)), z(t) + \eta(t)\rangle < -\frac{\rho}{2}, \quad \forall\, t \in I_0^g.$$

This proves that $z \in int\mathcal{Q}^{(1)}$. □

We associate with the first order variational equation (3.2), the following first order adjoint equation

$$\begin{cases} dP_1(t) = -\Big[b_x[t]^\top \big(P_1(t) + \psi(t)\big) + \sum\limits_{j=1}^d \sigma_x^j[t]^\top Q_1^j(t)\Big] dt + \sum\limits_{j=1}^d Q_1^j(t) dW^j(t), \quad t \in [0,T], \\ P_1(T) = -\lambda_0 \phi_x(\bar{x}(T)) - \sum\limits_{\ell \in \mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) - \psi(T), \end{cases} \tag{3.8}$$

where $\lambda_0 \in \{0,1\}$, $\psi \in L^2_{\mathbb{F}}(\Omega; BV_0([0,T];\mathbb{R}^n))$ and $\lambda_\ell \geq 0$ for any $\ell \in \mathcal{I}_0^{\varphi}$ ($\{\lambda_0, \{\lambda_\ell\}\}$ and ψ will be specified later). Since $L^2_{\mathbb{F}}(\Omega; BV_0([0,T];\mathbb{R}^n)) \subset L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^n))$, under the assumption (A2), the equation (3.8) admits a unique strong solution $(P_1(\cdot), Q_1(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \times L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^{n\times d}))$.

Define the Hamiltonian

$$\mathscr{H}(t,x,u,p,q,r,\omega) := \langle p+r, b(t,x,u,\omega)\rangle + \sum_{j=1}^{d} \left\langle q^j, \sigma^j(t,x,u,\omega)\right\rangle, \tag{3.9}$$

where $(t,x,u,p,q,r,\omega) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times \mathbb{R}^n \times \Omega$, and denote

$$\mathscr{H}[t] = \mathscr{H}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t), \psi(t)), \quad t \in [0,T],$$

$\mathscr{H}_x[t]$, $\mathscr{H}_u[t]$, $\mathscr{H}_{xx}[t]$, $\mathscr{H}_{xu}[t]$ and $\mathscr{H}_{uu}[t]$ are defined in a similar way.

Let γ be the map from $L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$ to $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ given by

$$\gamma(x) = x(T), \quad \forall\, x \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)). \tag{3.10}$$

By Lemma 2.9, γ is a well-defined bounded linear operator. By Lemma 2.7, denote by

$$\gamma^* : L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n) \to L^2_{\mathbb{F}}(\Omega; BV_0([0,T];\mathbb{R}^n)) \tag{3.11}$$

the adjoint operator of γ.

We next state the first order necessary optimality condition in the integral form.

Theorem 3.1. *Let (A1)–(A2) hold and $(\bar{x},\bar{u})$ be a local minimizer for the problem (1.5). If $\mathbb{E}\,|g_x(\bar{x}(t))| \neq 0$ for any $t \in I_0^g$, then there exist $\lambda_0 \in \{0,1\}$, $\lambda_\ell \geq 0$ for any $\ell \in \mathcal{I}_0^{\varphi}$ and $\psi \in \left(\mathcal{Q}^{(1)}\right)^-$ with $\psi(0) = 0$ satisfying*

$$\lambda_0 + \sum_{\ell \in \mathcal{I}_0^{\varphi}} \lambda_\ell + \|\psi\|_{BV,2} \neq 0$$

such that the corresponding solution (P_1, Q_1) to the first order adjoint equation (3.8) verifies

$$\langle P_1(0), \nu_0\rangle + \mathbb{E}\int_0^T \langle \mathscr{H}_u[t], v(t)\rangle\, dt \leq 0, \;\; \forall\, \nu_0 \in \mathcal{T}_{K_0}(\bar{x}_0), \;\; \forall\, v(\cdot) \in \mathcal{T}_{\Phi}(\bar{u}). \tag{3.12}$$

In addition, the above holds true with $\lambda_0 = 1$ if $\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$.

Letting $\Phi(t,\omega) = \mathcal{C}_U(\bar{u}(t,\omega))$, a.e. $(t,\omega) \in [0,T] \times \Omega$ and $\mathcal{T}_{K_0}(\bar{x}_0) = \mathcal{C}_{K_0}(\bar{x}_0)$ and using the same arguments as those in [15, proof of Theorem 3.2], as a direct consequence of Theorem 3.1, we obtain the following pointwise first order necessary condition (Hence we omit its proof).

Theorem 3.2. *Let (A1)–(A2) hold. If $(\bar{x},\bar{u})$ is a local minimizer for the problem (1.5) such that $\mathbb{E}|g_x(\bar{x}(t))| \neq 0$ for any $t \in I_0^g$, then for (P_1, Q_1) as in Theorem 3.1,*

$$P_1(0) \in N^C_{K_0}(\bar{x}_0) \qquad \text{and} \qquad \mathscr{H}_u[t] \in N^C_U(\bar{u}(t)), \;\; a.e.\; t \in [0,T], \;\; a.s. \tag{3.13}$$

Proof. (of Theorem 3.1). We shall derive the desired result by considering several cases.

(i) If $I_0^g = \emptyset$, then $\mathbb{E}\; g(\bar{x}(t)) < 0$ for every $t \in [0,T]$ and $\mathcal{Q}^{(1)} = L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$.

Let $\nu_0 \in \mathcal{T}_{K_0}(\bar{x}_0)$ and $v \in \mathcal{T}_{\Phi}(\bar{u})$. Consider $\mu(\varepsilon) \in \mathbb{R}^n$ with $|\mu(\varepsilon)| = o(\varepsilon)$ and $\eta(\varepsilon) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^n))$ with $\|\eta(\varepsilon)\|_2 = o(\varepsilon)$ such that $x_0^\varepsilon := \bar{x}_0 + \varepsilon\nu_0 + \mu(\varepsilon) \in K_0$ and $u^\varepsilon := \bar{u} + \varepsilon v + \eta(\varepsilon) \in \mathcal{U}$. Let x^ε be the solution to the control system (1.1) with the initial datum x_0^ε and control u^ε and let y_1 be the solution to the first order linearized control system (3.2) with control v and initial datum ν_0. By Lemma 3.1, for any fixed $t \in [0,T]$, one can find $\rho(t) > 0$ and $\alpha(t) > 0$ such that for any $\varepsilon \in [0, \alpha(t)]$,

$$\begin{aligned} \mathbb{E}\; g(x^\varepsilon(t)) &= \mathbb{E}\; g(\bar{x}(t)) + \varepsilon\mathbb{E}\,\langle g_x(\bar{x}(t)), y_1(t)\rangle + o(\varepsilon) \\ &\leq -\rho(t) + C\varepsilon(|\nu_0| + \|v\|_2) + o(\varepsilon) < -\frac{\rho(t)}{2}. \end{aligned}$$

Then, by the continuity of x^ε and Lemma 2.9, there exists a $\delta(t) > 0$ such that

$$\mathbb{E}\ g(x^\varepsilon(s)) < -\frac{\rho(t)}{4}, \quad \forall\, s \in (t-\delta(t), t+\delta(t)).$$

In the above we have set $x^\varepsilon(s) \equiv x_0$ when $s \in (-\delta(0), 0]$ and $x^\varepsilon(s) \equiv x(T)$ when $s \in [T, T+\delta(T))$. Since $[0,T]$ is compact, there exist $\rho_0 > 0$ and $\varepsilon_0 > 0$ (both independent from t), such that for any $\varepsilon \in [0, \varepsilon_0]$, $\mathbb{E}\ g(x^\varepsilon(t)) \le -\rho_0$ for every $t \in [0,T]$. This proves that any $(x^\varepsilon(\cdot), u^\varepsilon(\cdot))$ obtained by a sufficiently small perturbation of $\bar{u}(\cdot)$ and $\bar{x}_0$ satisfies the state constraint (1.4). Setting $\psi \equiv 0$ and using [16, proof of Theorem 3.4] we deduce that, there exist $\lambda_0 \in \{0,1\}$, $\lambda_\ell \ge 0$ for $\ell \in \mathcal{I}_0^\varphi$, $\lambda_0 + \sum_{\ell \in \mathcal{I}_0^\varphi} \lambda_\ell \ne 0$ such that, for the adjoint process (P_1, Q_1) defined by (3.8) (with $\psi \equiv 0$),

$$\langle P_1(0), \nu_0 \rangle + \mathbb{E} \int_0^T \langle \mathscr{H}_u[t], v(t) \rangle\, dt \le 0, \ \ \forall\, \nu_0 \in \mathcal{T}_{K_0}(\bar{x}_0), \ \ \forall\, v(\cdot) \in \mathcal{T}_\Phi(\bar{u})$$

and that $\lambda_0 = 1$ if $\mathcal{G}^{(1)} \cap \mathcal{E}^{(1)} \ne \emptyset$.

(ii) $I_0^g \ne \emptyset$, then $-g_x(\bar{x}(\cdot)) \in \mathcal{Q}^{(1)}$ and therefore, $\mathcal{Q}^{(1)} \ne \emptyset$.

We first prove that

$$\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)} \cap \mathcal{L}^{(1)} = \emptyset. \tag{3.14}$$

Indeed, otherwise, there would exist a $\bar{y}_1 \in \mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)}$ such that

$$\mathbb{E}\, \langle \phi_x(\bar{x}(T)), \bar{y}_1(T) \rangle < 0. \tag{3.15}$$

Let $\bar{\nu}_0 \in \mathcal{T}_{K_0}(\bar{x}_0)$ and $\bar{v}(\cdot) \in \mathcal{T}_\Phi(\bar{u}(\cdot))$ be respectively the initial datum and control corresponding to $\bar{y}_1$. Consider $\mu(\varepsilon) \in \mathbb{R}^n$ with $|\mu(\varepsilon)| = o(\varepsilon)$ and $\eta(\varepsilon) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^n))$ with $\|\eta(\varepsilon)\|_2 = o(\varepsilon)$ such that $x_0^\varepsilon := \bar{x}_0 + \varepsilon \bar{\nu}_0 + \mu(\varepsilon) \in K_0$ and $u^\varepsilon := \bar{u} + \varepsilon \bar{v} + \eta(\varepsilon) \in \mathcal{U}$. Let x^ε be the solution to the control system (1.1) with the initial datum x_0^ε and control u^ε.

Since $\bar{y}_1 \in \mathcal{Q}^{(1)}$, there exists a $\rho_0 > 0$ such that $\mathbb{E}\ \langle g_x(\bar{x}(t)), \bar{y}_1(t) \rangle < -\rho_0$ for every $t \in I_0^g$. Then, by Lemma 2.9 and the compactness of I_0^g, there exists a $\delta > 0$ (independent of $t \in I_0^g$) such that, $\mathbb{E}\ \langle g_x(\bar{x}(s)), \bar{y}_1(s) \rangle < -\frac{\rho_0}{2}$ for every $s \in (t-\delta, t+\delta) \cap [0,T]$, $t \in I_0^g$. Using Lemma 3.1 again, we can find an $\varepsilon_0 > 0$ such that, for any $\varepsilon \in [0, \varepsilon_0]$,

$$\begin{aligned} \mathbb{E}\ g(x^\varepsilon(s)) &= \mathbb{E}\ g(\bar{x}(s)) + \varepsilon\, \mathbb{E}\ \langle g_x(\bar{x}(s)), \bar{y}_1(s) \rangle + o(\varepsilon) \\ &\le \varepsilon\, \mathbb{E}\ \langle g_x(\bar{x}(s)), \bar{y}_1(s) \rangle + o(\varepsilon) \\ &< -\frac{\varepsilon \rho_0}{4} < 0, \quad \forall\, s \in (t-\delta, t+\delta) \cap [0,T],\ t \in I_0^g. \end{aligned} \tag{3.16}$$

In addition, because $I_\delta^c := [0,T] \setminus \cup_{t \in I_0^g} (t-\delta, t+\delta)$ is a compact set, there exist $\rho_1 > 0$ and $\varepsilon_1 > 0$ such that, for any $\varepsilon \in [0, \varepsilon_1]$,

$$\begin{aligned} \mathbb{E}\ g(x^\varepsilon(t)) &= \mathbb{E}\ g(\bar{x}(t)) + \varepsilon \mathbb{E}\, \langle g_x(\bar{x}(t)), \bar{y}_1(t) \rangle + o(\varepsilon) \\ &< -\rho_1 + \varepsilon \mathbb{E}\, \langle g_x(\bar{x}(t)), \bar{y}_1(t) \rangle + o(\varepsilon) \\ &< -\frac{\rho_1}{2} < 0, \quad \forall\, t \in I_\delta^c. \end{aligned} \tag{3.17}$$

By (3.16) and (3.17), x^ε satisfies the state constraint (1.4) when $\varepsilon < \min\{\varepsilon_0, \varepsilon_1\}$.

Moreover, since $\bar{y}_1(T) \in \mathcal{E}^{(1)}$, $\mathbb{E}\, \langle \varphi_x^\ell(\bar{x}(T)), \bar{y}_1(T) \rangle < 0$ for avery $\ell \in \mathcal{I}_0^\varphi$. Similarly to the arguments in the proof of [16, Theorem 3.4], for every sufficient small ε, x^ε satisfies the end point constraint (1.3). This proves that, when ε is small enough, $(x^\varepsilon, u^\varepsilon) \in \mathscr{P}_{ad}$. By (3.15), there exists a $\rho_2 > 0$ such that for sufficiently small ε,

$$\begin{aligned} \mathbb{E}\ \phi(x^\varepsilon(T)) &= \mathbb{E}\ \phi(\bar{x}(T)) + \varepsilon \mathbb{E}\, \langle \phi_x(\bar{x}(T)), \bar{y}_1(T) \rangle + o(\varepsilon) \\ &< \mathbb{E}\ \phi(\bar{x}(T)) - \varepsilon \rho_2 + o(\varepsilon) < \mathbb{E}\ \phi(\bar{x}(T)), \end{aligned}$$

in contradiction with the local optimality of $(\bar{x}, \bar{u})$. This completes the proof of (3.14).

Now we consider three different subcases.

Case a: $\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} = \emptyset$. Since $\mathcal{Q}^{(1)}$ is a nonempty open convex set and $\mathcal{G}^{(1)}$ is nonempty and convex, by Lemma 2.7 and the separation theorem there exists a nonzero $\psi \in L^2_{\mathbb{F}}(\Omega; BV([0,T];\mathbb{R}^n))$ with $\psi(0) = 0$ such that

$$\sup_{z\in\mathcal{Q}^{(1)}} \mathbb{E}\int_0^T \langle z(t), d\psi(t)\rangle \le \inf_{y\in\mathcal{G}^{(1)}} \mathbb{E}\int_0^T \langle y(t), d\psi(t)\rangle .$$

Because $\mathcal{Q}^{(1)}$ and $\mathcal{G}^{(1)}$ are cones,

$$0 = \sup_{z\in\mathcal{Q}^{(1)}} \mathbb{E}\int_0^T \langle z(t), d\psi(t)\rangle = \inf_{y\in\mathcal{G}^{(1)}} \mathbb{E}\int_0^T \langle y(t), d\psi(t)\rangle .$$

Therefore, $\psi \in \big(\mathcal{Q}^{(1)}\big)^-$ and $-\psi \in \big(\mathcal{G}^{(1)}\big)^-$. By Lemma 2.8, for any $y_1 \in \mathcal{G}^{(1)}$,

$$0 \ge -\mathbb{E}\int_0^T \langle y_1(t), d\psi(t)\rangle = -\mathbb{E}\langle y_1(T), \psi(T)\rangle + \mathbb{E}\int_0^T \langle \psi(t), b_x[t]y_1(t) + b_u[t]v(t)\rangle\, dt. \tag{3.18}$$

On the other hand, by the duality between (3.2) and (3.8) we have

$$\begin{aligned}
& \mathbb{E}\langle P_1(T), y_1(T)\rangle - \langle P_1(0), \nu_0\rangle \\
= \;& \mathbb{E}\int_0^T \Big(\langle P_1(t), b_x[t]y_1(t)\rangle + \langle P_1(t), b_u[t]v(t)\rangle - \big\langle b_x[t]^\top (P_1(t) + \psi(t)), y_1(t)\big\rangle \\
& - \sum_{j=1}^d \Big\langle \sigma_x^j[t]^\top Q_1^j(t), y_1(t)\Big\rangle + \sum_{j=1}^d \Big\langle Q_1^j(t), \sigma_x^j[t]y_1(t)\Big\rangle + \sum_{j=1}^d \Big\langle Q_1^j(t), \sigma_u^j[t]v(t)\Big\rangle \Big) dt \\
= \;& \mathbb{E}\int_0^T \Big(\langle P_1(t), b_u[t]v(t)\rangle - \big\langle b_x[t]^\top \psi(t), y_1(t)\big\rangle + \sum_{j=1}^d \Big\langle Q_1^j(t), \sigma_u^j[t]v(t)\Big\rangle \Big) dt.
\end{aligned} \tag{3.19}$$

Set $\lambda_0 = 0$, $\lambda_\ell = 0$, $\ell \in \mathcal{I}_0^\varphi$ and $P_1(T) = -\psi(T)$. Then, $\lambda_0 + \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell + \|\psi\|_{BV,2} \ne 0$ and (3.12) follows from (3.18) and (3.19).

Case b: $\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} \ne \emptyset$ and $\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} \cap \mathcal{E}^{(1)} = \emptyset$.

If $\mathcal{E}^{(1)} = \emptyset$, then from the proof of case (ii) in [16, Theorem 3.4], we know that, for each $\ell \in \mathcal{I}_0^\varphi$ there exists a $\lambda_\ell \ge 0$ such that $\sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell > 0$ and $\sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) = 0$. Then, taking $\lambda_0 = 0$, $\psi \equiv 0$ and $P_1(T) = 0$, we have $\lambda_0 + \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell + \|\psi\|_{BV,2} \ne 0$ and the condition (3.12) holds trivially with $(P_1, Q_1) \equiv 0$.

If $\mathcal{E}^{(1)} \ne \emptyset$, then $\gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big) \cap \mathcal{E}_T^{(1)} = \emptyset$, where γ is the bounded linear operator defined by (3.10) and

$$\mathcal{E}_T^{(1)} := \big\{\zeta \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n) \;\big|\; \mathbb{E}\big\langle \varphi_x^\ell(\bar{x}(T)), \zeta\big\rangle < 0,\ \forall\, \ell \in \mathcal{I}_0^\varphi\big\}. \tag{3.20}$$

By the separation theorem, there exists a nonzero $\xi \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ such that

$$\sup_{\alpha\in\gamma(\mathcal{Q}^{(1)}\cap\mathcal{G}^{(1)})} \mathbb{E}\langle \xi, \alpha\rangle \le \inf_{\beta\in\mathcal{E}_T^{(1)}} \mathbb{E}\langle \xi, \beta\rangle .$$

Since both $\gamma(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)})$ and $\mathcal{E}_T^{(1)}$ are cones, $0 = \sup_{\alpha\in\gamma(\mathcal{Q}^{(1)}\cap\mathcal{G}^{(1)})} \mathbb{E}\langle \xi, \alpha\rangle = \inf_{\beta\in\mathcal{E}_T^{(1)}} \mathbb{E}\langle \xi, \beta\rangle$. Therefore, $\xi \in \big[\gamma(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)})\big]^-$ and $-\xi \in \big(\mathcal{E}_T^{(1)}\big)^-$. By Lemma 2.1, for each $\ell \in \mathcal{I}_0^\varphi$ there exists a $\lambda_\ell \ge 0$ such that $\sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell > 0$ and $-\xi = \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T))$. In addition,

$$0 \ge \mathbb{E}\langle \xi, \gamma(z)\rangle \quad \forall\, z \in \mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}, \tag{3.21}$$

implying that $\gamma^*(\xi) \in \big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big)^-$. By Lemma 2.1 there exists a $\psi \in \big(\mathcal{Q}^{(1)}\big)^-$ with $\psi(0) = 0$ such that $\gamma^*(\xi) - \psi \in \big(\mathcal{G}^{(1)}\big)^-$. Then, by Lemma 2.8, for any $y_1 \in \mathcal{G}^{(1)}$,

$$0 \;\ge\; \mathbb{E}\langle \xi, y_1(T)\rangle - \mathbb{E}\int_0^T \langle y_1(t), d\psi(t)\rangle$$

$$= \mathbb{E}\langle \xi, y_1(T)\rangle - \mathbb{E}\langle \psi(T), y_1(T)\rangle + \mathbb{E}\int_0^T \langle \psi(t), b_x[t]y_1(t) + b_u[t]v(t)\rangle\, dt. \tag{3.22}$$

Let $\lambda_0 = 0$. Since $\xi \neq 0$, $\lambda_0 + \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell + \|\psi\|_{BV,2} \neq 0$. Setting

$$P_1(T) = -\sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) - \psi(T), \tag{3.23}$$

by (3.22), (3.23) and (3.19) we obtain (3.12).

Case c: $\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$. It implies that $\gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big) \cap \mathcal{E}_T^{(1)} \neq \emptyset$, ($\mathcal{E}_T^{(1)}$ is defined by (3.20)). By (3.14), $0 \leq \mathbb{E}\langle \phi_x(\bar{x}(T)), z(T)\rangle$ for every $z \in \mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} \cap \mathcal{E}^{(1)}$. This yields

$$0 \leq \mathbb{E}\langle \phi_x(\bar{x}(T)), \zeta\rangle, \quad \forall\, \zeta \in \gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big) \cap \mathcal{E}_T^{(1)}.$$

Consequently, $-\phi_x(\bar{x}(T)) \in \big[\gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big) \cap \mathcal{E}_T^{(1)}\big]^-$.

By Lemma 2.1, $\big[\gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big) \cap \mathcal{E}_T^{(1)}\big]^- = \big[\gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big)\big]^- + \big(\mathcal{E}_T^{(1)}\big)^-$. Then, for each $\ell \in \mathcal{I}_0^\varphi$ there exists a $\lambda_\ell \geq 0$ such that $\xi := \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) \in \big(\mathcal{E}_T^{(1)}\big)^-$ and

$$-\phi_x(\bar{x}(T)) - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) \in \big[\gamma\big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big)\big]^-.$$

Therefore,

$$\gamma^*\Big(-\phi_x(\bar{x}(T)) - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T))\Big) \in \big(\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}\big)^- = \big(\mathcal{Q}^{(1)}\big)^- + \big(\mathcal{G}^{(1)}\big)^-.$$

Let $\psi \in \big(\mathcal{Q}^{(1)}\big)^-$ with $\psi(0) = 0$ be such that

$$\gamma^*\Big(-\phi_x(\bar{x}(T)) - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T))\Big) - \psi \in \big(\mathcal{G}^{(1)}\big)^-$$

and set $\lambda_0 = 1$,

$$P_1(T) = -\phi_x(\bar{x}(T)) - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) - \psi(T). \tag{3.24}$$

Then, $\lambda_0 + \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell + \|\psi\|_{BV,2} \neq 0$ and, by Lemma 2.8 once more, for any $y_1 \in \mathcal{G}^{(1)}$,

$$\begin{aligned} 0 &\geq -\mathbb{E}\langle \phi_x(\bar{x}(T))), y_1(T)\rangle - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \mathbb{E}\big\langle \varphi_x^\ell(\bar{x}(T)), y_1(T)\big\rangle - \mathbb{E}\int_0^T \langle y_1(t), d\psi(t)\rangle \\ &= -\mathbb{E}\langle \phi_x(\bar{x}(T))), y_1(T)\rangle - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \mathbb{E}\big\langle \varphi_x^\ell(\bar{x}(T)), y_1(T)\big\rangle - \mathbb{E}\langle \psi(T), y_1(T)\rangle \\ &\quad + \mathbb{E}\int_0^T \langle \psi(t), b_x[t]y_1(t) + b_u[t]v(t)\rangle\, dt. \end{aligned} \tag{3.25}$$

Combining (3.24) and (3.25) with (3.19), we obtain (3.12). This completes the proof of Theorem 3.1. □

To end this section, we give below a simple example.

Example 3.1. *Let $T = 2$, $m = n = 2$, $d = 1$, $K_0 = \{0\}(\subset \mathbb{R}^2)$ and*

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Consider the control system

$$\begin{cases} dx(t) = Bu(t)dt + \big(Cx(t) + Du(t)\big)dW(t),\ t \in [0,2], \\ x(0) = 0 \end{cases} \tag{3.26}$$

with the control set

$$U := \Big\{(u^1, u^2)^\top \in \mathbb{R}^2 \ \Big|\ 0 \le u^1 \le 2,\ 0 \le u^2 \le 2 - \sqrt{4 - (u^1 - 2)^2}\Big\},$$

the cost functional

$$\mathbb{E}\ \phi(x(2)) = \frac{1}{2}\mathbb{E}\ |x^1(2) - 1 - (W(2) - W(1))^2|^2 - 2\mathbb{E}\ x^1(2)$$

and the state constraint

$$\mathbb{E}\ g(x(t)) = \mathbb{E}\ x^1(t) - 2 \le 0, \quad \forall\, t \in [0,2].$$

Let $\bar{u}(\cdot) = (\bar{u}^1(\cdot), \bar{u}^2(\cdot))$ be defined by

$$\bar{u}^1(t,\omega) = \begin{cases} 0, & (t,\omega) \in [0,1) \times \Omega; \\ 2, & (t,\omega) \in [1,2] \times \Omega. \end{cases}, \quad \bar{u}^2(t,\omega) = \begin{cases} 2, & (t,\omega) \in [0,1) \times \Omega; \\ 0, & (t,\omega) \in [1,2] \times \Omega. \end{cases}$$

Then, the solution $\bar{x}(\cdot) = (\bar{x}^1(\cdot), \bar{x}^2(\cdot))$ of (3.26) corresponding to $\bar{u}(\cdot)$ is

$$\bar{x}^1(t) = \begin{cases} 2t, & t \in [0,1),\ a.s.; \\ 3 - t + (W(t) - W(1))^2, & t \in [1,2],\ a.s.; \end{cases} \quad \bar{x}^2(t) = \begin{cases} 0, & t \in [0,1),\ a.s.; \\ 2W(t) - 2W(1), & t \in [1,2],\ a.s. \end{cases}$$

Clearly, $(\bar{x}, \bar{u})$ is an optimal pair, $I_0^g = [1,2]$, and,

$$\mathcal{Q}^{(1)} = \Big\{y \in L^2_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n)) \ \Big|\ \mathbb{E}\ y^1(t) < 0, \quad \forall\, t \in [1,2]\Big\}.$$

It is easy to check that

$$T_U^b(\bar{u}(t,\omega)) = \mathcal{C}_U(\bar{u}(t,\omega)) = \begin{cases} \big\{(0, -\alpha)^\top \ \big|\ \alpha \ge 0\big\}, & (t,\omega) \in [0,1) \times \Omega; \\ \big\{(-\alpha, 0)^\top \ \big|\ \alpha \ge 0\big\}, & (t,\omega) \in [1,2] \times \Omega. \end{cases}$$

Then, we can choose $\Phi(t,\omega) = T_U^b(\bar{u}(t,\omega))$ for any $(t,\omega) \in [0,2] \times \Omega$.

Letting

$$v(t,\omega) = \begin{cases} (0,-1)^\top, & (t,\omega) \in [0,1) \times \Omega; \\ (0,0)^\top, & (t,\omega) \in [1,2] \times \Omega \end{cases}$$

and $y_1(\cdot) = (y_1^1(\cdot), y_1^2(\cdot))^\top$ be the solution to

$$\begin{cases} dy_1(t) = Bv(t)dt + \big(Cy_1(t) + Dv(t)\big)dW(t),\ t \in [0,2], \\ y_1(0) = 0, \end{cases} \tag{3.27}$$

we have $y_1 \in \mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)}$. Set

$$\psi(t,\omega) = \begin{cases} (0,0)^\top, & (t,\omega) \in [0,1) \times \Omega; \\ (1,0)^\top, & (t,\omega) \in [1,2] \times \Omega, \end{cases}$$

$\lambda_0 = 1$ and let (P_1, Q_1) be the solution to the backward stochastic differential equation

$$\begin{cases} dP_1(t) = -C^\top Q_1(t)dt + Q_1(t)dW(t),\ t \in [0,2], \\ P_1(2) = (1,0)^\top. \end{cases} \tag{3.28}$$

Then $\psi \in \big(\mathcal{Q}^{(1)}\big)^-$, $\lambda_0 + \|\psi\|_{BV,2} \ne 0$, $(P_1, Q_1) \equiv ((1,0)^\top, 0)$,

$$P_1(t,\omega) + \psi(t,\omega) = \begin{cases} (1,0)^\top, & (t,\omega) \in [0,1) \times \Omega; \\ (2,0)^\top, & (t,\omega) \in [1,2] \times \Omega, \end{cases}$$

and,

$$\langle \mathscr{H}_u[t], v\rangle = \big\langle B^\top(P_1(t) + \psi(t)), v\big\rangle \le 0, \quad \forall\, v \in \mathcal{C}_U(\bar{u}(t)), \forall\, t \in [0,T],\ a.s.$$

Therefore, the first order necessary condition (3.13) holds.

4. Second order necessary condition

In this section, we investigate the second order necessary condition for $(\bar{x},\bar{u})$ to be a local minimizer for the problem (1.5). Throughout this section, we assume that $\bar{u}\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$ and write

$$\mathcal{V}:=\mathcal{U}\cap L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m)).$$

In addition to (A1)–(A2), we assume that

(A3) *b, σ^j $(j=1,\dots,d)$, ϕ, g and φ^ℓ $(\ell=1,\dots,k)$ satisfy the following conditions:*

(i) *For a.e. $(t,\omega)\in[0,T]\times\Omega$, $b(t,\cdot,\cdot,\omega):\ \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ and $\sigma^j(t,\cdot,\cdot,\omega):\ \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ $(j=1,\dots,d)$ are twice differentiable and*

$$(x,u)\mapsto b_{(x,u)^2}(t,x,u,\omega),\quad (x,u)\mapsto \sigma^j_{(x,u)^2}(t,x,u,\omega),\quad j=1,\dots,d$$

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

$$|b_{(x,u)^2}(t,x,u,\omega)|+\sum_{j=1}^d|\sigma^j_{(x,u)^2}(t,x,u,\omega)|\le L,\qquad \forall\,(x,u)\in\mathbb{R}^n\times\mathbb{R}^m,\quad \text{a.s.};$$

(ii) *$\phi(\cdot,\omega):\ \mathbb{R}^n\to\mathbb{R}$ is twice differentiable a.s., and for any $x,\ \tilde{x}\in\mathbb{R}^n$,*

$$|\phi_{xx}(x,\omega)|\le L,\quad |\phi_{xx}(x,\omega)-\phi_{xx}(\tilde{x},\omega)|\le L|x-\tilde{x}|,\quad \text{a.s.}$$

(iii) *g, φ^ℓ, $\ell=1,\dots,k$ are twice differentiable functions, and for any $x,\ \tilde{x}\in\mathbb{R}^n$,*

$$\begin{cases}|g_{xx}(x)|+\sum\limits_{\ell=1}^k|\varphi^\ell_{xx}(x)|\le L,\\ |g_{xx}(x)-g_{xx}(\tilde{x})|+\sum\limits_{\ell=1}^k|\varphi^\ell_{xx}(x)-\varphi^\ell_{xx}(\tilde{x})|\le L|x-\tilde{x}|,\quad \text{a.s.}\end{cases}$$

For $f=b$, σ^j $(j=1,\dots,d)$, write

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

For any $\nu_0\in T^b_{K_0}(\bar{x}_0)$, $v\in T^b_{\mathcal{V}}(\bar{u})$, $\varpi_0\in T^{b(2)}_{K_0}(\bar{x}_0,\nu_0)$ and $h\in T^{b(2)}_{\mathcal{V}}(\bar{u},v)$ (Here and henceforth, for the definitions of $T^b_{\mathcal{V}}(\bar{u})$ and $T^{b(2)}_{\mathcal{V}}(\bar{u},v)$, $\mathcal{V}$ is viewed as a subset of $L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$), similarly to [18, 15], we introduce the following second-order variational equation:

$$\begin{cases}dy_2(t)=\big(b_x[t]y_2(t)+b_u[t]h(t)+\frac{1}{2}y_1(t)^\top b_{xx}[t]y_1(t)+v(t)^\top b_{xu}[t]y_1(t)\\ \qquad\quad +\frac{1}{2}v(t)^\top b_{uu}[t]v(t)\big)dt+\sum\limits_{j=1}^d\big(\sigma^j_x[t]y_2(t)+\sigma^j_u[t]h(t)+\frac{1}{2}y_1(t)^\top\sigma^j_{xx}[t]y_1(t)\\ \qquad\quad +v(t)^\top\sigma^j_{xu}[t]y_1(t)+\frac{1}{2}v(t)^\top\sigma^j_{uu}[t]v(t)\big)dW^j(t),\qquad t\in[0,T],\\ y_2(0)=\varpi_0,\end{cases}\tag{4.1}$$

where y_1 is the solution of the linearized equation (3.2).

By the definition of the second order adjacent vector, for any $\varepsilon>0$, there exist $\varpi^\varepsilon_0\in\mathbb{R}^n$ and $h_\varepsilon\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$ such that $x^\varepsilon_0:=\bar{x}_0+\varepsilon\nu_0+\varepsilon^2\varpi^\varepsilon_0\in K_0$, $u^\varepsilon:=\bar{u}+\varepsilon v+\varepsilon^2h_\varepsilon\in\mathcal{V}$, $\varpi^\varepsilon_0\to\varpi_0$ in $\mathbb{R}^n$ and h_ε converges to h in $L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$ as $\varepsilon\to0^+$. Denote by x^ε the solution of (1.1) corresponding to the initial datum x^ε_0 and the control u^ε. Put

$$\delta x^\varepsilon=x^\varepsilon-\bar{x},\qquad r^\varepsilon_2(t,\omega):=\frac{\delta x^\varepsilon(t,\omega)-\varepsilon y_1(t,\omega)-\varepsilon^2y_2(t,\omega)}{\varepsilon^2}.$$

The next result for d-dimensional Wiener process follows by the same arguments as those used to prove [15, Lemma 4.1].

Lemma 4.1. *Let (A2) (i) and (A3) (i) hold. Then, for any* $v, h, h_\varepsilon \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\ \mathbb{R}^m))$ *and* $\nu_0, \varpi_0, \varpi_0^\varepsilon \in \mathbb{R}^n$ *as above, we have*

$$\|y_2\|_{\infty,2} \le C(|\varpi_0| + \|v\|_4^2 + \|h\|_2).$$

Furthermore,

$$\|r_2^\varepsilon\|_{\infty,2} \to 0, \quad \varepsilon \to 0^+. \tag{4.2}$$

Define the critical set

$$\begin{aligned}\Upsilon(\bar{x},\bar{u}) := \Big\{ (y_1, v, \nu_0) \in L^4_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \times T^b_{\mathcal{V}}(\bar{u}) \times T^b_{K_0}(\bar{x}_0) \Big|\ & y_1 \text{ solves (3.2)}, \\ & y_1 \in cl_{\infty,2}(\mathcal{Q}^{(1)}) \cap cl_{\infty,2}(\mathcal{E}^{(1)}) \text{ and } \mathbb{E}\langle \phi_x(\bar{x}(T)), y_1(T)\rangle = 0 \Big\}.\end{aligned} \tag{4.3}$$

For any $(y_1, v, \nu_0) \in \Upsilon(\bar{x},\bar{u})$, let $\mathcal{W}(\bar{x}_0,\nu_0)$ and $\mathcal{M}(\bar{u}, v)$ be convex subsets of $T^{b(2)}_{K_0}(\bar{x}_0,\nu_0)$ and $T^{b(2)}_{\mathcal{V}}(\bar{u}, v)$, respectively. Denote

$$\mathcal{G}^{(2)}(y_1, v) := \Big\{ y_2 \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \Big|\ y_2 \text{ solves (4.1) for some } \varpi_0 \in \mathcal{W}(\bar{x}_0,\nu_0) \text{ and } h \in \mathcal{M}(\bar{u}, v) \Big\}. \tag{4.4}$$

Remark 4.1. *Let* $v \in T^b_{\mathcal{V}}(\bar{u})$. *When* U *is defined by finitely many equalities and inequalities constraints, under some constraint qualification and smoothness assumptions, the second order adjacent set* $T^{b(2)}_U(\bar{u}(t,\omega), v(t,\omega))$ *is a nonempty closed convex subset of* $\mathbb{R}^m$ *for a.e.* $(t,\omega) \in [0,T] \times \Omega$ *(see [15, Example 2.1]). Moreover, if there exist a nonnegative* $\eta(\cdot) \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}))$ *and an* $\varepsilon_0 > 0$ *such that*

$$dist(\bar{u}(t,\omega) + \varepsilon v(t,\omega), U) \le \varepsilon^2 \eta(t,\omega),\ a.e.\ (t,\omega) \in [0,T] \times \Omega,\ \forall\, \varepsilon \in [0,\varepsilon_0], \tag{4.5}$$

then, similarly to the proof of [15, Theorem 4.1], we deduce that every $h(\cdot) \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ *satisfying* $h(t,\omega) \in T^{b(2)}_U(\bar{u}(t,\omega), v(t,\omega))$ *for a.e.* $(t,\omega) \in [0,T] \times \Omega$ *belongs to* $T^{b(2)}_{\mathcal{V}}(\bar{u}, v)$. *In this case, we can choose* $\mathcal{M}(\bar{u}, v)$ *to be the set*

$$\Big\{ h(\cdot) \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m)) \Big|\ h(t,\omega) \in T^{b(2)}_U(\bar{u}(t,\omega), v(t,\omega)),\ a.e.\ (t,\omega) \in [0,T] \times \Omega \Big\}.$$

For a sufficient condition for (4.5) we refer the reader to [16, Example 4.1].

Define

$$I_1^g := \big\{ t \in I_0^g \ \big|\ \mathbb{E}\langle g_x(\bar{x}(t)), y_1(t)\rangle = 0 \big\}, \tag{4.6}$$

$$\mathcal{I}_1^\varphi := \big\{ \ell \in \mathcal{I}_0^\varphi \mid \mathbb{E}\big\langle \varphi_x^\ell(\bar{x}(T)), y_1(T)\big\rangle = 0 \big\}, \tag{4.7}$$

$$\begin{aligned}\tau^g(\bar{x}) := \big\{ t \in [0,T] \ \big|\ \exists\, \{s_i\}_{i=1}^\infty \subset [0,T] \text{ such that } \lim_{i\to\infty} s_i = t,\ \mathbb{E}\ g(\bar{x}(s_i)) < 0, \\ \mathbb{E}\langle g_x(\bar{x}(s_i)), y_1(s_i)\rangle > 0,\ \forall\, i = 1,2,\dots \big\},\end{aligned} \tag{4.8}$$

and

$$e(t) := \begin{cases} \displaystyle\limsup_{\substack{s\to t \\ \mathbb{E}\ g(\bar{x}(s))<0 \\ \mathbb{E}\langle g_x(\bar{x}(s)), y_1(s)\rangle > 0}} \frac{\big|\mathbb{E}\langle g_x(\bar{x}(s)), y_1(s)\rangle\big|^2}{4\big|\mathbb{E}\ g(\bar{x}(s))\big|}, & t \in \tau^g(\bar{x}), \\ 0, & \text{otherwise.} \end{cases} \tag{4.9}$$

Consider the set

$$\begin{aligned}\mathcal{Q}^{(2)}(y_1) \ := \ \Big\{ z \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \ \Big|\ & \text{For all } t \in I_1^g, \\ & \mathbb{E}\langle g_x(\bar{x}(t)), z(t)\rangle + \frac{1}{2}\mathbb{E}\langle g_{xx}(\bar{x}(t)) y_1(t), y_1(t)\rangle + e(t) < 0 \Big\}\end{aligned} \tag{4.10}$$

and

$$\begin{aligned}\mathcal{E}^{(2)}(y_1) \ := \ \Big\{ z \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \ \Big|\ & \text{For all } \ell \in \mathcal{I}_1^\varphi, \\ & \mathbb{E}\big\langle \varphi_x^\ell(\bar{x}(T)), z(T)\big\rangle + \frac{1}{2}\mathbb{E}\big\langle \varphi_{xx}^\ell(\bar{x}(T)) y_1(T), y_1(T)\big\rangle < 0 \Big\}.\end{aligned} \tag{4.11}$$

Remark 4.2. *From (4.9), it is easy to see that $e(\cdot)$ is an upper semicontinuous function on I_1^g. Therefore, $\mathcal{Q}^{(2)}(y_1)$ is a convex open set (possibly empty). When $y_1 \in \mathcal{Q}^{(1)}$, i.e., $\mathbb{E}\langle g_x(\bar{x}(t)), y_1(t)\rangle < 0$ for any $t \in I_0^g$, $\mathcal{Q}^{(2)}(y_1) = L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$, since for such y_1, $I_1^g = \emptyset$. In addition, if there exists a $\delta > 0$ such that*

$$\mathbb{E}\langle g_x(\bar{x}(s)), y_1(s)\rangle \le 0, \quad \forall\, s \in (t-\delta, t+\delta),\ t \in I_0^g,$$

then, $e(t) = 0$ for any $t \in I_1^g$. In this case, $\mathcal{Q}^{(2)}(y_1)$ enjoys a much simpler expression:

$$\mathcal{Q}^{(2)}(y_1) := \Big\{ z \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \,\Big|\, \text{For all } t \in I_1^g,\ \mathbb{E}\langle g_x(\bar{x}(t)), z(t)\rangle + \frac{1}{2}\mathbb{E}\langle g_{xx}(\bar{x}(t))y_1(t), y_1(t)\rangle < 0 \Big\}.$$

However, since the function $e(\cdot)$ depends on the process y_1, in general, $e(\cdot) \neq 0$. For more details about this issue in the deterministic cases we refer the reader to [23] and references cited therein.

Let (P_1, Q_1), ψ and λ_ℓ, $\ell \in \mathcal{I}_0^\varphi$ be defined as in the proof of Theorem 3.1 in the case when $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$ (See (3.3), (3.5) and (3.6) for the definitions of $\mathcal{G}^{(1)}$, $\mathcal{Q}^{(1)}$ and $\mathcal{E}^{(1)}$, respectively). In particular, $P_1(T)$ is given by (3.24). Corresponding to such fixed (P_1, Q_1), ψ and λ_ℓ for all $\ell \in \mathcal{I}_0^\varphi$, we introduce the following adjoint equation to (4.1):

$$\begin{cases} dP_2(t) = -\big(b_x[t]^\top P_2(t) + P_2(t)b_x[t] + \sum\limits_{j=1}^d \sigma_x^j[t]^\top P_2(t)\sigma_x^j[t] + \sum\limits_{j=1}^d \sigma_x^j[t]^\top Q_2^j(t) \\ \qquad\qquad + \sum\limits_{j=1}^d Q_2^j(t)\sigma_x^j[t] + \mathscr{H}_{xx}[t]\big)dt + \sum\limits_{j=1}^d Q_2^j(t)dW^j(t), \quad t \in [0,T], \\ P_2(T) = -\phi_{xx}(\bar{x}(T)) - \sum\limits_{\ell \in \mathcal{I}_0^\varphi} \lambda_\ell \varphi_{xx}^\ell(\bar{x}(T)), \end{cases} \tag{4.12}$$

where $\mathscr{H}$ is given by (3.9). Clearly, under the assumptions (A2)–(A3), the equation (4.12) admits a unique strong solution $(P_2(\cdot), Q_2(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C([0,T]; \mathbf{S}^n)) \times \big(L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbf{S}^n))\big)^d$, where $\mathbf{S}^n := \big\{A \in \mathbb{R}^{n\times n} \big|\, A^\top = A\big\}$.

To simplify the notation, we define

$$\begin{aligned} \mathcal{S}(t,x,u,y_1,z_1,r,y_2,z_2,\omega) := {} & \mathscr{H}_{xu}(t,x,u,y_1,z_1,r,\omega) + b_u(t,x,u,\omega)^\top y_2 \\ & + \sum_{j=1}^d \sigma_u^j(t,x,u,\omega)^\top z_2^j + \sum_{j=1}^d \sigma_u^j(t,x,u,\omega)^\top y_2 \sigma_x^j(t,x,u,\omega), \end{aligned}$$

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

$$\mathcal{S}[t] = \mathcal{S}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t), \psi(t), P_2(t), Q_2(t)), \quad t \in [0,T], \tag{4.13}$$

where $(P_1(\cdot), Q_1(\cdot))$ and $(P_2(\cdot), Q_2(\cdot))$ solve the systems (3.8) with $\lambda_0 = 1$ and (4.12), respectively.

We have the following result.

Theorem 4.1. *Let $(\bar{x}, \bar{u})$ be a local minimizer for the problem (1.5) with the initial datum $\bar{x}_0$ and the control $\bar{u} \in \mathcal{V}$. Suppose (A1)–(A3) and that $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$. Let $\lambda_0 = 1$, (P_1, Q_1), ψ and λ_ℓ, $\ell \in \mathcal{I}_0^\varphi$ be as in the conclusions of Theorem 3.1 for the case that $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$.*

If $(y_1, v, \nu_0) \in \Upsilon(\bar{x}, \bar{u})$ and $\mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1) \neq \emptyset$, then for any given $y_2 \in \mathcal{G}^{(2)}(y_1, v) \cap cl_{\infty,2}(\mathcal{Q}^{(2)}(y_1)) \cap cl_{\infty,2}(\mathcal{E}^{(2)}(y_1))$ with the corresponding $\varpi_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$ and $h(\cdot) \in \mathcal{M}(\bar{u}, v)$, the following second order necessary condition holds:

$$\begin{aligned} & \langle P_1(0), \varpi_0\rangle + \frac{1}{2}\langle P_2(0)\nu_0, \nu_0\rangle + \sum_{\ell \in \mathcal{I}_0^\varphi} \mathbb{E}\Big(\big\langle \lambda_\ell \varphi_x^\ell(\bar{x}(T)), y_2(T)\big\rangle + \frac{1}{2}\mathbb{E}\ \big\langle \lambda_\ell \varphi_{xx}^\ell(\bar{x}(T))y_1(T), y_1(T)\big\rangle \Big) \\ & + \mathbb{E}\int_0^T \Big(\langle \mathscr{H}_u[t], h(t)\rangle + \frac{1}{2}\langle \mathscr{H}_{uu}[t]v(t), v(t)\rangle + \frac{1}{2}\sum_{j=1}^d \big\langle P_2(t)\sigma_u^j[t]v(t), \sigma_u^j[t]v(t)\big\rangle \\ & + \langle \mathcal{S}[t]y_1(t), v(t)\rangle \Big)dt + \mathbb{E}\int_0^T \langle y_2(t), d\psi(t)\rangle \le 0, \end{aligned} \tag{4.14}$$

where (P_2, Q_2) is defined by (4.12).

Remark 4.3. *(i) In Theorem 4.1 we assumed that the set $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)}$ is nonempty and have taken $\lambda_0 = 1$ and any (P_1, Q_1), ψ and λ_ℓ, $\ell \in \mathcal{I}_0^\varphi$ as in the conclusion of Theorem 3.1. That is our second order condition is valid for any normal multiplier appearing in the first order conditions. In Theorem 4.2 below this assumption is skipped, but, in contrast, it states that a similar second order necessary optimality condition holds true for some (P_1, Q_1) as in Theorem 3.2. That is, removing the requirement $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$, yields a weaker conclusion.*

(ii) Also, in Theorem 4.1 we assumed that $\mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1) \neq \emptyset$. This holds true under some appropriate conditions. For instance, let us assume that for $\mathcal{T}_{K_0}(\bar{x}_0) = \mathcal{C}_{K_0}(\bar{x}_0)$ and $\mathcal{T}_\Phi(\bar{u}) = \mathcal{C}_\mathcal{V}(\bar{u})$ it holds that $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$ (a sufficient condition for it will be provided in Section 5), and pick any $\bar{y}_1 \in \mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)}$. Then, for every $(y_1, v, \nu_0) \in \Upsilon(\bar{x}, \bar{u})$ such that the function $e(\cdot)$ defined by (4.9) is bounded on I_1^g, $T_{K_0}^{b(2)}(\bar{x}_0, \nu_0) \neq \emptyset$ and $T_\mathcal{V}^{b(2)}(\bar{u}, v) \neq \emptyset$, we can find $\mathcal{W}(\bar{x}_0, \nu_0)$ and $\mathcal{M}(\bar{u}, v)$ so that $\mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1) \neq \emptyset$.

Indeed, since $T_{K_0}^{b(2)}(\bar{x}_0, \nu_0)$ and $T_\mathcal{V}^{b(2)}(\bar{u}, v)$ are nonempty, they contain some nonempty convex subsets $\overline{\mathcal{W}}(\bar{x}_0, \nu_0) \subset T_{K_0}^{b(2)}(\bar{x}_0, \nu_0)$ and $\overline{\mathcal{M}}(\bar{u}, v) \subset T_\mathcal{V}^{b(2)}(\bar{u}, v)$. Fix such subsets and define nonempty convex sets $\mathcal{W}(\bar{x}_0, \nu_0) := \mathcal{C}_{K_0}(\bar{x}_0) + \overline{\mathcal{W}}(\bar{x}_0, \nu_0)$, $\mathcal{M}(\bar{u}, v) := \mathcal{C}_\mathcal{V}(\bar{u}) + \overline{\mathcal{M}}(\bar{u}, v)$. From Remark 2.1, it follows that $\mathcal{W}(\bar{x}_0, \nu_0) \subset T_{K_0}^{b(2)}(\bar{x}_0, \nu_0)$ and $\mathcal{M}(\bar{u}, v) \subset T_\mathcal{V}^{b(2)}(\bar{u}, v)$. Moreover, for every $\varpi_0 \in \overline{\mathcal{W}}(\bar{x}_0, \nu_0)$, $h \in \overline{\mathcal{M}}(\bar{u}, v)$ for any $\varrho \geq 0$ we have $\varrho\bar{\nu}_0 + \varpi_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$ and $\varrho\bar{v} + h \in \mathcal{M}(\bar{u}, v)$.

Fixing $\varrho \geq 0$ and letting y_2 be the solution of (4.1) corresponding to $\varrho\bar{\nu}_0 + \varpi_0$ and $\varrho\bar{v} + h$, we have $y_2 = \varrho\bar{y}_1 + z_2$. By Lemma 4.1,

$$\|z_2\|_{\infty,2}^2 \leq C(|\varpi_0|^2 + \|v\|_4^4 + \|h\|_2^2).$$

Since $\bar{y}_1 \in \mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)}$, and I_0^g and I_1^g are compact sets, for all sufficiently large ϱ,

$$\begin{aligned} & \mathbb{E}\langle g_x(\bar{x}(t)), y_2(t)\rangle + \frac{1}{2}\mathbb{E}\langle g_{xx}(\bar{x}(t))y_1(t), y_1(t)\rangle + e(t) \\ = \quad & \varrho\mathbb{E}\langle g_x(\bar{x}(t)), \bar{y}_1(t)\rangle + \mathbb{E}\langle g_x(\bar{x}(t)), z_2(t)\rangle + \frac{1}{2}\mathbb{E}\langle g_{xx}(\bar{x}(t))y_1(t), y_1(t)\rangle + e(t) \\ < \quad & 0, \quad \forall\, t \in I_1^g, \end{aligned}$$

and for every $\ell \in \mathcal{I}_1^\varphi$, and all ϱ sufficiently large

$$\begin{aligned} & \mathbb{E}\left\langle \varphi_x^\ell(\bar{x}(T)), y_2(T)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle \varphi_{xx}^\ell(\bar{x}(T))y_1(T), y_1(T)\right\rangle \\ = \quad & \varrho\mathbb{E}\left\langle \varphi_x^\ell(\bar{x}(T)), \bar{y}_1(T)\right\rangle + \mathbb{E}\left\langle \varphi_x^\ell(\bar{x}(T)), z_2(T)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle \varphi_{xx}^\ell(\bar{x}(T))y_1(T), y_1(T)\right\rangle < 0. \end{aligned}$$

Therefore, when ϱ is large enough, $y_2 \in \mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1)$.

Proof. (of Theorem 4.1).

Observe that if $I_1^g = \emptyset$, then $\mathcal{Q}^{(2)}(y_1) = L_\mathbb{F}^2(\Omega; C([0,T]; \mathbb{R}^n))$ and therefore, $\mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1) = \mathcal{G}^{(2)}(y_1, v) \cap \mathcal{E}^{(2)}(y_1)$. In this case, the same arguments as given below do apply by simply skipping the constraint (1.4) in the problem (1.5) and putting $\psi = 0$.

For this reason we provide the proof only when $I_1^g \neq \emptyset$.

We proceed in two steps.

Step 1: $y_2 \in \mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1)$, i.e., when y_2 is a solution of the equation (4.1) corresponding to some $(\varpi_0, h) \in \mathcal{W}(\bar{x}_0, \nu_0) \times \mathcal{M}(\bar{u}, v)$ such that

$$\mathbb{E}\langle g_x(\bar{x}(t)), y_2(t)\rangle + \frac{1}{2}\mathbb{E}\langle g_{xx}(\bar{x}(t))y_1(t), y_1(t)\rangle + e(t) < 0, \quad \forall\, t \in I_1^g.$$

and

$$\mathbb{E}\left\langle \varphi_x^\ell(\bar{x}(T)), y_2(T)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle \varphi_{xx}^\ell(\bar{x}(T))y_1(T), y_1(T)\right\rangle < 0, \quad \forall\, \ell \in \mathcal{I}_1^\varphi,$$

where I_1^g and $\mathcal{I}_1^\varphi$ are defined respectively by (4.6) and (4.7).

Let $\mu(\varepsilon) \in \mathbb{R}^n$ with $|\mu(\varepsilon)| = o(\varepsilon^2)$ and $\eta(\varepsilon) \in L_\mathbb{F}^4(\Omega; L^4(0,T; \mathbb{R}^m))$ with $\|\eta(\varepsilon)\|_4 = o(\varepsilon^2)$ be such that $x_0^\varepsilon := \bar{x}_0 + \varepsilon\nu_0 + \varepsilon^2\varpi_0 + \mu(\varepsilon) \in K_0$ and $u^\varepsilon := \bar{u} + \varepsilon v + \varepsilon^2 h + \eta(\varepsilon) \in \mathcal{V}$. Denote by x^ε the solution to (1.1)

corresponding to x_0^ε and u^ε. By the assumptions (A2)–(A3) and Lemma 4.1, for any $t \in [0,T]$ we have the following expansion

$$\begin{aligned}\mathbb{E}\ g(x^\varepsilon(t)) &= \mathbb{E}\ g(\bar{x}(t)) + \varepsilon\mathbb{E}\,\langle g_x(\bar{x}(t)), y_1(t)\rangle + \varepsilon^2\mathbb{E}\,\langle g_x(\bar{x}(t)), y_2(t)\rangle \\ &\quad + \frac{\varepsilon^2}{2}\mathbb{E}\,\langle g_{xx}(\bar{x}(t))y_1(t), y_1(t)\rangle + o(\varepsilon^2). \end{aligned} \tag{4.15}$$

Similarly to the proof of [18, Proposition 4.5], we first check that $(x^\varepsilon, u^\varepsilon) \in \mathscr{P}_{ad}$ when ε is sufficiently small.

Our first goal is to show that there exist a $\delta_0 > 0$ and an $\varepsilon_0 > 0$ such that

$$\mathbb{E}\ g(x^\varepsilon(s)) \le 0, \quad \forall\, s \in (t-\delta_0, t+\delta_0) \cap [0,T],\ \forall\, t \in I_0^g,\ \forall\, \varepsilon \in [0,\varepsilon_0]. \tag{4.16}$$

Fix an arbitrary $\hat{t} \in I_1^g$. We claim that there exist $\delta(\hat{t}) > 0$ and $\alpha(\hat{t}) > 0$ such that

$$\mathbb{E}\ g(x^\varepsilon(s)) \le 0, \quad \forall\, s \in (\hat{t}-\delta(\hat{t}), \hat{t}+\delta(\hat{t})) \cap [0,T],\ \forall\, \varepsilon \in [0,\alpha(\hat{t})]. \tag{4.17}$$

Indeed, otherwise, for any $i \in \mathbb{N}$, there would exist $\varepsilon_i \in [0, \frac{1}{i}]$, $s_i \in (\hat{t} - \frac{1}{i}, \hat{t} + \frac{1}{i}) \cap [0,T]$ such that

$$\mathbb{E}\ g(x^{\varepsilon_i}(s_i)) > 0. \tag{4.18}$$

Assume for a moment that there exists a subsequence $\{s_{i_\kappa}\}_{\kappa=1}^\infty$ of $\{s_i\}_{i=1}^\infty$ with corresponding subsequence $\{\varepsilon_{i_\kappa}\}_{\kappa=1}^\infty$ of $\{\varepsilon_i\}_{i=1}^\infty$ satisfying

$$\mathbb{E}\ g(\bar{x}(s_{i_\kappa})) < 0 \text{ and } \mathbb{E}\,\langle g_x(\bar{x}(s_{i_\kappa})), y_1(s_{i_\kappa})\rangle > 0,\ \forall\, \kappa = 1,2,\ldots. \tag{4.19}$$

Then, by (4.15),

$$\begin{aligned}\mathbb{E}\ g(x^{\varepsilon_{i_\kappa}}(s_{i_\kappa})) &= \varepsilon_{i_\kappa}^2\Big(\mathbb{E}\,\langle g_x(\bar{x}(s_{i_\kappa})), y_2(s_{i_\kappa})\rangle + \frac{1}{2}\mathbb{E}\,\langle g_{xx}(\bar{x}(s_{i_\kappa}))y_1(s_{i_\kappa}), y_1(s_{i_\kappa})\rangle \\ &\quad - \frac{\big|\mathbb{E}\,\langle g_x(\bar{x}(s_{i_\kappa})), y_1(s_{i_\kappa})\rangle\big|^2}{4\mathbb{E}\ g(\bar{x}(s_{i_\kappa}))} + \frac{o(\varepsilon_{i_\kappa}^2)}{\varepsilon_{i_\kappa}^2}\Big) \\ &\quad + \mathbb{E}\ g(\bar{x}(s_{i_\kappa}))\Big(1 + \frac{\varepsilon_{i_\kappa}\mathbb{E}\,\langle g_x(\bar{x}(s_{i_\kappa})), y_1(s_{i_\kappa})\rangle}{2\mathbb{E}\ g(\bar{x}(s_{i_\kappa}))}\Big)^2. \end{aligned} \tag{4.20}$$

Since $\hat{t} \in I_1^g$ and $y_2 \in \mathcal{Q}^{(2)}(y_1)$, there exists $\rho_0 > 0$ such that

$$\mathbb{E}\,\big\langle g_x(\bar{x}(\hat{t})), y_2(\hat{t})\big\rangle + \frac{1}{2}\mathbb{E}\,\big\langle g_{xx}(\bar{x}(\hat{t}))y_1(\hat{t}), y_1(\hat{t})\big\rangle + e(\hat{t}) < -\rho_0.$$

Therefore, when κ is large enough,

$$\mathbb{E}\,\langle g_x(\bar{x}(s_{i_\kappa})), y_2(s_{i_\kappa})\rangle + \frac{1}{2}\mathbb{E}\,\langle g_{xx}(\bar{x}(s_{i_\kappa}))y_1(s_{i_\kappa}), y_1(s_{i_\kappa})\rangle + \frac{\big|\mathbb{E}\,\langle g_x(\bar{x}(s_{i_\kappa})), y_1(s_{i_\kappa})\rangle\big|^2}{4\big|\mathbb{E}\ g(\bar{x}(s_{i_\kappa}))\big|} < -\frac{\rho_0}{2}.$$

This, together with (4.20), implies that, when κ is large enough, $\mathbb{E}\ g(x^{\varepsilon_{i_\kappa}}(s_{i_\kappa})) \le 0$ contradicting to (4.18).

Now, assume that there is no any subsequence of $\{s_i\}_{i=1}^\infty$ such that (4.19) hold. Then, $\mathbb{E}\ g(\bar{x}(s_i)) = 0$ or $\mathbb{E}\,\langle g_x(\bar{x}(s_i)), y_1(s_i)\rangle \le 0$ for all sufficiently large i. Then, if $s_i \notin I_0^g$, we have $\mathbb{E}\ g(\bar{x}(s_i)) < 0$ and hence $\mathbb{E}\,\langle g_x(\bar{x}(s_i)), y_1(s_i)\rangle \le 0$. On the other hand, if $s_i \in I_0^g$, then $\mathbb{E}\ g(\bar{x}(s_i)) = 0$ and, since $y_1 \in cl_{\infty,2}\big(\mathcal{Q}^{(1)}\big)$, $\mathbb{E}\,\langle g_x(\bar{x}(s_i)), y_1(s_i)\rangle \le 0$. In both cases,

$$\mathbb{E}\ g(\bar{x}(s_i)) + \varepsilon_i\mathbb{E}\,\langle g_x(\bar{x}(s_i)), y_1(s_i)\rangle \le 0. \tag{4.21}$$

Noting that $e(t) \ge 0$ for all $t \in [0,T]$, and I_1^g is compact, there exists $\rho_2 > 0$ such that

$$\mathbb{E}\,\langle g_x(\bar{x}(t)), y_2(t)\rangle + \frac{1}{2}\mathbb{E}\,\langle g_{xx}(\bar{x}(t))y_1(t), y_1(t)\rangle < -\rho_2, \quad \forall\, t \in I_1^g.$$

Since $s_i \to \hat{t}$ and $\hat{t} \in I_1^g$, when i is large enough,

$$\mathbb{E}\,\langle g_x(\bar{x}(s_i)), y_2(s_i)\rangle + \frac{1}{2}\mathbb{E}\,\langle g_{xx}(\bar{x}(s_i))y_1(s_i), y_1(s_i)\rangle < -\frac{\rho_2}{2}.$$

Then, by (4.15) and (4.21), for any sufficiently large i,

$$\begin{aligned}\mathbb{E}\; g(x^{\varepsilon_i}(s_i)) &\leq \varepsilon_i^2\mathbb{E}\,\langle g_x(\bar{x}(s_i)), y_2(s_i)\rangle + \frac{\varepsilon_i^2}{2}\mathbb{E}\,\langle g_{xx}(\bar{x}(s_i))y_1(s_i), y_1(s_i)\rangle + o(\varepsilon_i^2)\\ &\leq \varepsilon_i^2\big(-\frac{\rho_2}{2} + \frac{o(\varepsilon_i^2)}{\varepsilon_i^2}\big) \leq 0,\end{aligned}$$

which also contradicts (4.18). This proves (4.17).

Since I_1^g is compact, there exist finitely many $t_i \in I_1^g$, $i = 1, 2, \ldots, N$ ($N \in \mathbb{N}$) such that $I_1^g \subset \bigcup_{i=1}^N (t_i - \delta(t_i), t_i + \delta(t_i))$. Choosing $\varepsilon_1 := \min\{\alpha(t_i),\ i = 1, 2, \ldots, N\}$, we have

$$\mathbb{E}\; g(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in \bigcup_{i=1}^N (t_i - \delta(t_i), t_i + \delta(t_i)) \cap [0, T],\ \forall\, \varepsilon \in [0, \varepsilon_1]. \tag{4.22}$$

Let $I_1^c := I_0^g \setminus \bigcup_{i=1}^N (t_i - \delta(t_i), t_i + \delta(t_i))$. By the compactness of I_1^c, there exist a $\tilde{\delta} > 0$ and a $\rho_3 > 0$ (independent of t) such that

$$\mathbb{E}\,\langle g_x(\bar{x}(s)), y_1(s)\rangle < -\rho_3, \quad \forall\, s \in (t - \tilde{\delta}, t + \tilde{\delta}) \cap [0, T],\ t \in I_1^c. \tag{4.23}$$

Then, by (4.15), there exists an ε_2 such that

$$\mathbb{E}\; g(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in (t - \tilde{\delta}, t + \tilde{\delta}) \cap [0, T],\ t \in I_1^c,\ \forall\, \varepsilon \in [0, \varepsilon_2]. \tag{4.24}$$

Let $\varepsilon_0 = \min\{\varepsilon_1, \varepsilon_2\}$. Clearly,

$$I_0^g \subset \big(\bigcup_{i=1}^N (t_i - \delta(t_i), t_i + \delta(t_i))\big) \bigcup \big(\bigcup_{t \in I_1^c} (t - \tilde{\delta}, t + \tilde{\delta})\big).$$

Choosing a sufficiently small $\delta_0 > 0$ such that

$$I_0^g \subset \bigcup_{t \in I_0^g} (t - \delta_0, t + \delta_0) \subset \big(\bigcup_{i=1}^N (t_i - \delta(t_i), t_i + \delta(t_i))\big) \bigcup \big(\bigcup_{t \in I_1^c} (t - \tilde{\delta}, t + \tilde{\delta})\big),$$

from the above discussion (especially, (4.22) and (4.24)), we obtain that

$$\mathbb{E}\; g(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in (t - \delta_0, t + \delta_0) \cap [0, T],\ \forall\, t \in I_0^g,\ \forall\, \varepsilon \in [0, \varepsilon_0].$$

This proves (4.16).

Next, define

$$I_< := [0, T] \setminus \Big[\bigcup_{t \in I_0^g} (t - \delta_0, t + \delta_0)\Big].$$

It is clear that $I_<$ is also a compact set and consequently, there exists a $\rho_4 > 0$ such that

$$\mathbb{E}\; g(\bar{x}(t)) < -\rho_4, \quad \forall\, t \in I_<. \tag{4.25}$$

By (4.15), we have, for sufficiently small $\varepsilon > 0$,

$$\mathbb{E}\; g(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in I_<. \tag{4.26}$$

Combining (4.16) with (4.26), we conclude that, when ε is small enough, x^ε satisfies the constraint (1.4).

In addition, using a similar method as above at time T, we have, when ε is small enough, x^ε satisfies the constraint (1.3) (see also [16, Theorem 4.2] for a detailed discussion). This proves that $(x^\varepsilon, u^\varepsilon) \in \mathscr{P}_{ad}$.

By the optimality of $(\bar{x}, \bar{u})$ and the equality $\mathbb{E}\langle \phi_x(\bar{x}(T)), y_1(T)\rangle = 0$, we have

$$\begin{aligned} 0 \quad \leq \quad & \frac{\mathbb{E}\ \phi(x^\varepsilon(T)) - \mathbb{E}\ \phi(\bar{x}(T))}{\varepsilon^2} \\ = \quad & \frac{1}{\varepsilon}\mathbb{E}\langle \phi_x(\bar{x}(T)), y_1(T)\rangle + \mathbb{E}\langle \phi_x(\bar{x}(T)), y_2(T)\rangle \\ & + \frac{1}{2}\mathbb{E}\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle + \frac{o(\varepsilon^2)}{\varepsilon^2} \\ & \to \mathbb{E}\langle \phi_x(\bar{x}(T)), y_2(T)\rangle + \frac{1}{2}\mathbb{E}\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle, \quad (\text{as } \varepsilon \to 0^+). \end{aligned} \tag{4.27}$$

By (3.24), we have

$$\phi_x(\bar{x}(T)) = -P_1(T) - \sum_{\ell \in \mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) - \psi(T). \tag{4.28}$$

By Itô's formula,

$$\begin{aligned} & \mathbb{E}\ \langle P_1(T), y_2(T)\rangle \\ = \quad & \langle P_1(0), \varpi_0\rangle + \mathbb{E}\int_0^T \Big(\langle P_1(t), b_u[t]h(t)\rangle + \frac{1}{2}\left\langle P_1(t), y_1(t)^\top b_{xx}[t]y_1(t)\right\rangle \\ & + \left\langle P_1(t), v(t)^\top b_{xu}[t]y_1(t)\right\rangle + \frac{1}{2}\left\langle P_1(t), v(t)^\top b_{uu}[t]v(t)\right\rangle - \langle \psi(t), b_x[t]y_2(t)\rangle \\ & + \sum_{j=1}^d \left\langle Q_1^j(t), \sigma_u^j[t]h(t)\right\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle Q_1^j(t), y_1(t)^\top \sigma_{xx}^j[t]y_1(t)\right\rangle \\ & + \sum_{j=1}^d \left\langle Q_1^j(t), v(t)^\top \sigma_{xu}^j[t]y_1(t)\right\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle Q_1^j(t), v(t)^\top \sigma_{uu}^j[t]v(t)\right\rangle \Big) dt, \end{aligned} \tag{4.29}$$

and

$$\begin{aligned} & \mathbb{E}\ \langle \psi(T), y_2(T)\rangle \\ = \quad & \mathbb{E}\int_0^T \Big(\langle \psi(t), b_x[t]y_2(t)\rangle + \langle \psi(t), b_u[t]h(t)\rangle + \frac{1}{2}\left\langle \psi(t), y_1(t)^\top b_{xx}[t]y_1(t)\right\rangle \\ & + \left\langle \psi(t), v(t)^\top b_{xu}[t]y_1(t)\right\rangle + \frac{1}{2}\left\langle \psi(t), v(t)^\top b_{uu}[t]v(t)\right\rangle \Big) dt + \mathbb{E}\int_0^T \langle y_2(t), d\psi(t)\rangle. \end{aligned} \tag{4.30}$$

Combining (4.28) with (4.29) and (4.30) we obtain that

$$\begin{aligned} & \mathbb{E}\langle \phi_x(\bar{x}(T)), y_2(T)\rangle = \mathbb{E}\left\langle -P_1(T) - \sum_{\ell \in \mathcal{I}_0^\varphi} \lambda_\ell \varphi_x^\ell(\bar{x}(T)) - \psi(T), y_2(T)\right\rangle \\ = \quad & -\langle P_1(0), \varpi_0\rangle - \sum_{\ell \in \mathcal{I}_0^\varphi} \mathbb{E}\left\langle \lambda_\ell \varphi_x^\ell(\bar{x}(T)), y_2(T)\right\rangle - \mathbb{E}\int_0^T \langle y_2(t), d\psi(t)\rangle \\ & -\mathbb{E}\int_0^T \Big(\langle P_1(t) + \psi(t), b_u[t]h(t)\rangle + \frac{1}{2}\left\langle P_1(t) + \psi(t), y_1(t)^\top b_{xx}[t]y_1(t)\right\rangle \\ & \qquad + \left\langle P_1(t) + \psi(t), v(t)^\top b_{xu}[t]y_1(t)\right\rangle + \frac{1}{2}\left\langle P_1(t) + \psi(t), v(t)^\top b_{uu}[t]v(t)\right\rangle \\ & \qquad + \sum_{j=1}^d \left\langle Q_1^j(t), \sigma_u^j[t]h(t)\right\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle Q_1^j(t), y_1(t)^\top \sigma_{xx}^j[t]y_1(t)\right\rangle \\ & \qquad + \sum_{j=1}^d \left\langle Q_1^j(t), v(t)^\top \sigma_{xu}^j[t]y_1(t)\right\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle Q_1^j(t), v(t)^\top \sigma_{uu}^j[t]v(t)\right\rangle \Big) dt \end{aligned}$$

$$
\begin{aligned}
= \quad & -\langle P_1(0), \varpi_0\rangle - \sum_{\ell\in\mathcal{I}_0^\varphi} \mathbb{E}\left\langle \lambda_\ell \varphi_x^\ell(\bar{x}(T)), y_2(T)\right\rangle \\
& -\mathbb{E}\int_0^T \langle y_2(t), d\psi(t)\rangle - \mathbb{E}\int_0^T \langle \mathscr{H}_u[t], h(t)\rangle\, dt \\
& -\frac{1}{2}\mathbb{E}\int_0^T \Big(\langle \mathscr{H}_{xx}[t]y_1(t), y_1(t)\rangle + 2\langle \mathscr{H}_{xu}[t]y_1(t), v(t)\rangle + \langle \mathscr{H}_{uu}[t]v(t), v(t)\rangle\Big)dt. \qquad (4.31)
\end{aligned}
$$

In addition, by the choice of $P_2(T)$, we have

$$
\phi_{xx}(\bar{x}(T)) = -P_2(T) - \sum_{\ell\in\mathcal{I}_0^\varphi} \lambda_\ell \varphi_{xx}^\ell(\bar{x}(T)). \qquad (4.32)
$$

Also, by Itô's formula,

$$
\begin{aligned}
& \mathbb{E}\ \langle P_2(T)y_1(T), y_1(T)\rangle \qquad (4.33) \\
= \ & \langle P_2(0)\nu_0, \nu_0\rangle + \mathbb{E}\int_0^T \Big(2\langle P_2(t)y_1(t), b_u[t]v(t)\rangle + 2\sum_{j=1}^d \left\langle P_2(t)\sigma_x^j[t]y_1(t), \sigma_u^j[t]v(t)\right\rangle \\
& + \sum_{j=1}^d \left\langle P_2(t)\sigma_u^j[t]v(t), \sigma_u^j[t]v(t)\right\rangle + 2\sum_{j=1}^d \left\langle Q_2^j(t)y_1(t), \sigma_u^j[t]v(t)\right\rangle - \langle \mathscr{H}_{xx}[t]y_1(t), y_1(t)\rangle\Big)dt.
\end{aligned}
$$

Substituting (4.31), (4.32) and (4.33) into (4.27), we obtain (4.14).

Step 2: Let $y_2 \in \mathcal{G}^{(2)}(y_1, v) \cap cl_{\infty,2}(\mathcal{Q}^{(2)}(y_1)) \cap cl_{\infty,2}(\mathcal{E}^{(2)}(y_1))$ with the corresponding $\varpi_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$ and $h(\cdot) \in \mathcal{M}(\bar{u}, v)$. Since we assume that $\mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1) \neq \emptyset$, we may choose a $\hat{y}_2 \in \mathcal{G}^{(2)}(y_1, v) \cap \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1)$ and let $\hat{\varpi}_0 (\in \mathcal{W}(\bar{x}_0, \nu_0)$ and $\hat{h}(\cdot)$ $(\in \mathcal{M}(\bar{u}, v))$ be its initial datum and control in the equation (4.1). For any $\theta \in (0, 1)$, define $y_2^\theta = (1-\theta)y_2 + \theta\hat{y}_2$. Since $\mathcal{W}(\bar{x}_0, \nu_0)$ and $\mathcal{M}(\bar{u}, v)$ are convex set, y_2^θ is the solution to the equation (4.1) with initial datum $\varpi_0^\theta := (1-\theta)\varpi_0 + \theta\hat{\varpi}_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$ and control $h^\theta(\cdot) := (1-\theta)h(\cdot) + \theta\hat{h}(\cdot) \in \mathcal{M}(\bar{u}, v)$. Then, $y_2^\theta \to y_2$ in $L^2_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n))$ as $\theta \to 0$. Moreover, since $\hat{y}_2 \in \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1)$ we have $y_2^\theta \in \mathcal{Q}^{(2)}(y_1) \cap \mathcal{E}^{(2)}(y_1)$. From Step 1, we deduce that

$$
\begin{aligned}
& \left\langle P_1(0), \varpi_0^\theta\right\rangle + \frac{1}{2}\langle P_2(0)\nu_0, \nu_0\rangle \\
& + \sum_{\ell\in\mathcal{I}_0^\varphi} \mathbb{E}\Big(\left\langle \lambda_\ell \varphi_x^\ell(\bar{x}(T)), y_2^\theta(T)\right\rangle + \frac{1}{2}\mathbb{E}\ \left\langle \lambda_\ell \varphi_{xx}^\ell(\bar{x}(T))y_1(T), y_1(T)\right\rangle\Big) \\
& + \mathbb{E}\int_0^T \Big(\left\langle \mathscr{H}_u[t], h^\theta(t)\right\rangle + \frac{1}{2}\langle \mathscr{H}_{uu}[t]v(t), v(t)\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle P_2(t)\sigma_u^j[t]v(t), \sigma_u^j[t]v(t)\right\rangle \\
& + \langle \mathcal{S}[t]y_1(t), v(t)\rangle\Big)dt + \mathbb{E}\int_0^T \left\langle y_2^\theta(t), d\psi(t)\right\rangle \leq 0.
\end{aligned}
$$

Letting $\theta \to 0$, we obtain (4.14).

This completes the proof of Theorem 4.1. □

The above theorem excludes the case of abnormal multipliers. Our next aim is to state second order conditions that may be abnormal but allow to dispense some assumptions.

Theorem 4.2. *Let $(\bar{x}, \bar{u})$ be a local minimizer for the problem (1.5) with the initial datum $\bar{x}_0$ and the control $\bar{u} \in \mathcal{V}$. Assume (A1)–(A3) and that $\mathbb{E}\ |g_x(\bar{x}(t))| \neq 0$ for any $t \in I_0^g$. Let $(y_1, v, \nu_0) \in \Upsilon(\bar{x}, \bar{u})$ and suppose that $e(\cdot)$ (defined by (4.9)) is bounded on I_1^g. Consider convex sets $\mathcal{W}(\bar{x}_0, \nu_0) \subset T_{K_0}^{b(2)}(\bar{x}_0, \nu_0)$ and $\mathcal{M}(\bar{u}, v) \subset T_{\mathcal{V}}^{b(2)}(\bar{u}, v)$. Then there exist $\lambda_0 \in \{0, 1\}$, $\lambda_\ell \geq 0$ for all $\ell \in \mathcal{I}_1^\varphi$ and $\psi \in \left(\mathcal{Q}^{(1)}\right)^-$ such that the solution (P_1, Q_1) to (3.8) with $\mathcal{I}_0^\varphi$ replaced by $\mathcal{I}_1^\varphi$ satisfies the first order condition (3.13), and for any $y_2 \in \mathcal{G}^{(2)}(y_1, v)$ with the corresponding $\varpi_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$ and $h(\cdot) \in \mathcal{M}(\bar{u}, v)$ the following second order necessary condition holds true:*

$$
\langle P_1(0), \varpi_0\rangle + \frac{1}{2}\langle P_2(0)\nu_0, \nu_0\rangle + \mathbb{E}\int_0^T \Big(\langle \mathscr{H}_u[t], h(t)\rangle + \frac{1}{2}\langle \mathscr{H}_{uu}[t]v(t), v(t)\rangle + \langle \mathcal{S}[t]y_1(t), v(t)\rangle
$$

$$+\frac{1}{2}\sum_{j=1}^{d}\left\langle P_2(t)\sigma_u^j[t]v(t),\sigma_u^j[t]v(t)\right\rangle\Big)dt+\sup_{\alpha\in\mathcal{Q}^{(2)}(y_1)}\mathbb{E}\int_0^T\langle\alpha(t),d\psi(t)\rangle\le 0, \tag{4.34}$$

where (P_2,Q_2) *solves (4.12) with*

$$P_2(T)=-\lambda_0\phi_{xx}(\bar{x}(T))-\sum_{\ell\in\mathcal{I}_1^{\varphi}}\lambda_\ell\varphi_{xx}^{\ell}(\bar{x}(T)). \tag{4.35}$$

Proof. If $\mathcal{W}(\bar{x}_0,\nu_0)$ or $\mathcal{M}(\bar{u},v)$ is empty, then it is enough to apply Theorem 3.2. Assume next that these two sets are nonempty. We replace them by larger convex sets as follows $\overline{\mathcal{W}}(\bar{x}_0,\nu_0):=\mathcal{C}_{K_0}(\bar{x}_0)+\mathcal{W}(\bar{x}_0,\nu_0)$, $\overline{\mathcal{M}}(\bar{u},v):=\mathcal{T}_\Phi(\bar{u})+\mathcal{M}(\bar{u},v)$, where

$$\mathcal{T}_\Phi(\bar{u}):=\{u\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))\mid u(t,\omega)\in\mathcal{C}_U(\bar{u}(t,\omega))\ \text{ a.e. in }[0,T]\times\Omega\}.$$

By [16, Lemma 2.7], $\mathcal{T}_\Phi(\bar{u})\subset\mathcal{C}_{\mathcal{V}}(\bar{u})$. Thus, by Remark 2.1, $\overline{\mathcal{M}}(\bar{u},v)\subset T_{\mathcal{V}}^{b(2)}(\bar{u},v)$. For simplicity, we keep the same notations $\mathcal{W}(\bar{x}_0,\nu_0)$ and $\mathcal{M}(\bar{u},v)$ for these larger sets.

If $I_1^g=\emptyset$, then $\mathcal{Q}^{(2)}(y_1)=L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ and $\mathcal{G}^{(2)}(y_1,v)\cap\mathcal{Q}^{(2)}(y_1)\cap\mathcal{E}^{(2)}(y_1)=\mathcal{G}^{(2)}(y_1,v)\cap\mathcal{E}^{(2)}(y_1)$. Then [16, Theorem 4.2] implies the result with $\psi=0$.

Assume next that $I_1^g\neq\emptyset$. Since $\mathbb{E}\,|g_x(\bar{x}(t))|\neq 0$ for any $t\in I_0^g$ and $e(\cdot)$ is bounded on I_1^g, we have $-g_x(\bar{x}(\cdot))\in\mathcal{Q}^{(1)}$ and $-\varrho g_x(\bar{x}(\cdot))\in\mathcal{Q}^{(2)}(y_1)$ when $\varrho\ (>0)$ is large enough. Hence $\mathcal{Q}^{(1)}\neq\emptyset$ and $\mathcal{Q}^{(2)}(y_1)\neq\emptyset$.

Let $y_2\in\mathcal{G}^{(2)}(y_1,v)$ and (P_1,Q_1) be a solution of (3.8). Then using (4.29) and (4.30), in the same way as in (4.31), we deduce that

$$\begin{aligned}&\mathbb{E}\,\langle P_1(T)+\psi(T),y_2(T)\rangle\\ =\ &\langle P_1(0),\varpi_0\rangle+\mathbb{E}\int_0^T\langle y_2(t),d\psi(t)\rangle+\mathbb{E}\int_0^T\langle\mathscr{H}_u[t],h(t)\rangle\,dt\\ &+\frac{1}{2}\mathbb{E}\int_0^T\Big(\langle\mathscr{H}_{xx}[t]y_1(t),y_1(t)\rangle+2\,\langle\mathscr{H}_{xu}[t]y_1(t),v(t)\rangle+\langle\mathscr{H}_{uu}[t]v(t),v(t)\rangle\Big)dt.\end{aligned} \tag{4.36}$$

Set $\varphi_0:=\phi$ and define for $\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}$ the open convex sets

$$M_\ell=\left\{z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))\ \Big|\ \mathbb{E}\left\langle\varphi_x^{\ell}(\bar{x}(T)),z(T)\right\rangle+\frac{1}{2}\mathbb{E}\left\langle\varphi_{xx}^{\ell}(\bar{x}(T))y_1(T),y_1(T)\right\rangle<0\right\}.$$

If $M_{\bar{\ell}}=\emptyset$ for some $\bar{\ell}\in\{0\}\cup\mathcal{I}_1^{\varphi}$, then $\varphi_x^{\bar{\ell}}(\bar{x}(T))=0$ a.s. and $\mathbb{E}\left\langle\varphi_{xx}^{\bar{\ell}}(\bar{x}(T))y_1(T),y_1(T)\right\rangle\ge 0$. Setting $\psi(t)\equiv 0$, $\lambda_{\bar{\ell}}=1$ and $\lambda_\ell=0$ for all $\bar{\ell}\neq\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}$, we get $(P_1(t),Q_1(t))\equiv 0$, $P_2(T)=-\varphi_{xx}^{\bar{\ell}}(\bar{x}(T))$ and

$$\frac{1}{2}\mathbb{E}\ \langle P_2(T)y_1(T),y_1(T)\rangle\le 0. \tag{4.37}$$

Since $(P_1(t),Q_1(t))\equiv 0$ and $\psi(t)\equiv 0$, we obtain that

$$\frac{1}{2}\mathbb{E}\int_0^T\langle\mathscr{H}_{xx}[t]y_1(t),y_1(t)\rangle\,dt=0.$$

This and (4.33) imply (4.34).

Assume next that $M_\ell\neq\emptyset$ for every $\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}$ and observe that $\bigcap_{\ell\in\mathcal{I}_1^{\varphi}}M_\ell=\mathcal{E}^{(2)}(y_1)$. Thus, by Step 1 of the proof of Theorem 4.1,

$$\mathcal{G}^{(2)}(y_1,v)\cap\mathcal{Q}^{(2)}(y_1)\cap\bigcap_{\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}}M_\ell=\emptyset.$$

By Lemma 2.3, there exist $x^*,x_\ell^*\in L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))^*$ for all $\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}$ not vanishing simultaneously such that for $p^*=-(x^*+\sum_{\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}}x_\ell^*)$ we have

$$\inf p^*(\mathcal{G}^{(2)}(y_1,v))+\inf x^*(\mathcal{Q}^{(2)}(y_1))+\sum_{\ell\in\{0\}\cup\mathcal{I}_1^{\varphi}}\inf x_\ell^*(M_\ell)\ge 0. \tag{4.38}$$

If $\varphi_x^\ell(\bar{x}(T))=0$ for some $\ell\in\{0\}\cup\mathcal{I}_1^\varphi$, then $M_\ell=L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ and (4.38) implies that $x_\ell^*=0$.

For any $\ell\in\{0\}\cup\mathcal{I}_1^\varphi$ with $\varphi_x^\ell(\bar{x}(T))\neq 0$, consider the closed convex cone

$$Q_\ell:=\Big\{z_T\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)\ \Big|\ \mathbb{E}\left\langle\varphi_x^\ell(\bar{x}(T)),z_T\right\rangle\leq 0\ \Big\}. \tag{4.39}$$

Then $(Q_\ell)^-=\mathbb{R}_+\varphi_x^\ell(\bar{x}(T))$.

Let γ be given by (3.10) and γ^* denote its adjoint (see (3.11)). It is not difficult to realize that $\gamma^{-1}(Q_\ell)+M_\ell\subset M_\ell$ for every $\ell\in\{0\}\cup\mathcal{I}_1^\varphi$ and that $\gamma^{-1}(Q_\ell)$ is a cone. Hence, by (4.38), $-x_\ell^*\in\big(\gamma^{-1}(Q_\ell)\big)^-$. We already know that γ is surjective. Thus, by the well known result of the convex analysis, $\big(\gamma^{-1}(Q_\ell)\big)^-=\gamma^*(Q_\ell^-)$ (see for instance [2, Corollary 22, p. 144] applied to the closed convex cone Q_ℓ and the set-valued map γ^{-1} whose graph is a closed subspace of $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)\times L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$). Consequently, $-x_\ell^*=\gamma^*(\lambda_\ell\varphi_x^\ell(\bar{x}(T)))$ for some $\lambda_\ell\geq 0$. If $x_\ell^*=0$, set $\lambda_\ell=0$. By normalizing, we may assume that $\lambda_0\in\{0,1\}$.

Define

$$\Gamma_\ell=\big\{z_T\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)\ \big|\ \mathbb{E}\left\langle\varphi_x^\ell(\bar{x}(T)),z_T\right\rangle+\frac{1}{2}\mathbb{E}\left\langle\varphi_{xx}^\ell(\bar{x}(T))y_1(T),y_1(T)\right\rangle<0\big\}.$$

By the surjectivity of γ,

$$\sup_{m\in M_\ell}\left(-x_\ell^*\right)(m)=\sup_{m\in M_\ell}\mathbb{E}\langle\lambda_\ell\varphi_x^\ell(\bar{x}(T)),\gamma(m)\rangle=\sup_{z_T\in\Gamma_\ell}\mathbb{E}\langle\lambda_\ell\varphi_x^\ell(\bar{x}(T)),z_T\rangle.$$

On the other hand, by the definition of M_ℓ, for any $\ell\in\{0\}\cup\mathcal{I}_1^\varphi$ with $\varphi_x^\ell(\bar{x}(T))\neq 0$,

$$\sup_{z_T\in\Gamma_\ell}\mathbb{E}\langle\lambda_\ell\varphi_x^\ell(\bar{x}(T)),z_T\rangle=-\frac{\lambda_\ell}{2}\mathbb{E}\left\langle\varphi_{xx}^\ell(\bar{x}(T))y_1(T),y_1(T)\right\rangle.$$

From (4.38), and setting $\psi=-x^*,q^*=-p^*$, we deduce that

$$\sup_{y_2\in\mathcal{G}^{(2)}(y_1,v)}q^*(y_2)+\sup_{\alpha\in\mathcal{Q}^{(2)}(y_1)}\mathbb{E}\int_0^T\langle\alpha(t),d\psi(t)\rangle-\frac{1}{2}\mathbb{E}\sum_{\ell\in\{0\}\cup\mathcal{I}_1^\varphi}\left\langle\lambda_\ell\varphi_{xx}^\ell(\bar{x}(T))y_1(T),y_1(T)\right\rangle\leq 0. \tag{4.40}$$

Observe that that $\mathcal{Q}^{(1)}+\mathcal{Q}^{(2)}(y_1)\subset\mathcal{Q}^{(2)}(y_1)$ and $\mathcal{G}^{(1)}+\mathcal{G}^{(2)}(y_1,v)\subset\mathcal{G}^{(2)}(y_1,v)$. Thus (4.40) implies that $\psi\in(\mathcal{Q}^{(1)})^-$ and $q^*\in(\mathcal{G}^{(1)})^-$.

Since $q^*=x^*-\sum_{\ell\in\{0\}\cup\mathcal{I}_1^\varphi}\gamma^*(\lambda_\ell\varphi_x^\ell(\bar{x}(T)))$, for every $y_1\in\mathcal{G}^{(1)}$

$$-\mathbb{E}\int_0^T\langle y_1(t),d\psi(t)\rangle-\sum_{\ell\in\{0\}\cup\mathcal{I}_1^\varphi}\mathbb{E}\langle\lambda_\ell\varphi_x^\ell(\bar{x}(T)),y_1(T)\rangle\leq 0.$$

Define $P_1(T)=-\sum_{\ell\in\{0\}\cup\mathcal{I}_1^\varphi}\lambda_\ell\varphi_x^\ell(\bar{x}(T))-\psi(T)$ and consider the solution (P_1,Q_1) of (3.8) with $\mathcal{I}_0^\varphi$ replaced by $\mathcal{I}_1^\varphi$. Applying (3.22) and (3.19) we obtain (3.12). Then (3.13) follows in the same way as before.

Now, for the solution (P_2,Q_2) to (4.12) with $P_2(T)$ given by (4.35), by (4.40), for every $y_2\in\mathcal{G}^{(2)}(y_1,v)$,

$$\begin{aligned}-\mathbb{E}\int_0^T\langle y_2(t),d\psi(t)\rangle&-\sum_{\ell\in\{0\}\cup\mathcal{I}_1^\varphi}\mathbb{E}\langle\lambda_\ell\varphi_x^\ell(\bar{x}(T)),y_2(T)\rangle\\&+\frac{1}{2}\mathbb{E}\langle P_2(T)y_1(T),y_1(T)\rangle+\sup_{\alpha\in\mathcal{Q}^{(2)}(y_1)}\int_0^T\langle\alpha(t),d\psi(t)\rangle\leq 0.\end{aligned}$$

From the above inequality, using (4.29), (4.30) and (4.33), we deduce (4.34).

□

5. A sufficient condition for normality

As in the deterministic optimal control problems with state constraints, we call the first order necessary condition (3.12) *normal* if the Lagrange multiplier $\lambda_0\neq 0$. By Theorem 3.1, this is the case when $\mathcal{G}^{(1)}\cap\mathcal{Q}^{(1)}\cap\mathcal{E}^{(1)}\neq\emptyset$. In this section, we first give a sufficient condition for $\mathcal{G}^{(1)}\cap\mathcal{Q}^{(1)}\neq\emptyset$ and then use the obtained results to investigate a sufficient condition guaranteeing that $\mathcal{G}^{(1)}\cap\mathcal{Q}^{(1)}\cap\mathcal{E}^{(1)}\neq\emptyset$.

In addition to (A2) (i), we assume that

(A4) *$g(\cdot)$ is continuously differentiable up to the third order, and there exists a nonnegative $\eta \in L^\infty_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R})$ such that*

$$\big|g_x(\bar{x}(t))\big| + \big|g_{xx}(\bar{x}(t))b[t]\big| + \sum_{j=1}^d \big|g_{xx}(\bar{x}(t))\sigma^j[t]\big| + \frac{1}{2}\big|g_{xxx}(\bar{x}(t))\sigma[t]\sigma[t]^\top\big| \le \eta(t), \quad \text{a.e. } t\in[0,T], \text{ a.s.}$$

(A5) *There exist a $\delta>0$ and a $\rho>0$ such that*

$$\min_{v\in\Phi(t,\omega)\cap B_{\mathbb{R}^m}} \Big[\langle g_x(\bar{x}(t)), b_u[t]v\rangle + \sum_{j=1}^d \big\langle g_{xx}(\bar{x}(t))\sigma^j[t], \sigma^j_u[t]v\big\rangle\Big] \le -\rho, \text{ a.e. } (t,\omega)\in I^g_\delta\times\Omega,$$

where, Φ is a set-valued map satisfying (3.1), $B_{\mathbb{R}^m}$ is the unit ball in $\mathbb{R}^m$ and

$$I^g_\delta := \big\{t\in[0,T]\,\big|\,\mathbb{E}\, g(\bar{x}(t))\in[-\delta,0]\big\}. \tag{5.1}$$

Theorem 5.1. *Assume (A2)(i) and (A4)–(A5), and that $g(x_0)<0$ and $\mathcal{T}_{K_0}(\bar{x}_0)$ contains a non-zero element. Then, $\mathcal{Q}^{(1)}\cap\mathcal{G}^{(1)}\neq\emptyset$.*

Proof. We follow the proof of [7, Theorem 3.10]. We only need to prove $\mathcal{Q}^{(1)}\cap\mathcal{G}^{(1)}\neq\emptyset$ when $I^g_0\neq\emptyset$ (I^g_0 is defined by (3.4)).

By (A5) and Lemma 2.4 (using a completion argument if necessary), there exists a $v\in L^\infty_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^m))$,

$$v(t,\omega)\in\Phi(t,\omega)\cap B_{\mathbb{R}^m}, \quad \text{a.e. } (t,\omega)\in[0,T]\times\Omega$$

such that

$$\langle g_x(\bar{x}(t)), b_u[t]v(t)\rangle + \sum_{j=1}^d \big\langle g_{xx}(\bar{x}(t))\sigma^j[t], \sigma^j_u[t]v(t)\big\rangle \le -\rho, \quad \text{a.e. } (t,\omega)\in I^g_\delta\times\Omega. \tag{5.2}$$

In the following we shall construct a continuous stochastic process $y\in\mathcal{Q}^{(1)}\cap\mathcal{G}^{(1)}$.

Step 1: Define

$$t_1 = \inf\big\{t\,\big|\,t\in I^g_0\big\}.$$

Since I^g_0 is nonempty and compact, $t_1\in I^g_0$. Since $g(x_0)<0$, $t_1>0$.

We shall consider below two different cases.

(a) If $\mathbb{E}\, g(\bar{x}(t))\in(-\delta,0)$ for any $t\in[0,t_1)$, let $C_1=\dfrac{L+2}{\rho}$, $C_2=\dfrac{|g_x(\bar{x}_0)|+1}{\rho}$ and define y on $[0,t_1]\times\Omega$ as the solution of

$$\begin{cases} dy(t) = \big(b_x[t]y(t) + C_1\eta(t)|y(t)|b_u[t]v(t)\big)dt + C_2\dfrac{|\nu_0|}{t_1}b_u[t]v(t)dt \\ \qquad\qquad + \sum\limits_{j=1}^d \big(\sigma^j_x[t]y(t) + C_1\eta(t)|y(t)|\sigma^j_u[t]v(t)\big)dW^j(t) \\ \qquad\qquad + C_2\dfrac{|\nu_0|}{t_1}\sum\limits_{j=1}^d \sigma^j_u[t]v(t)dW^j(t), \ t\in[0,t_1], \\ y(0)=\nu_0, \end{cases} \tag{5.3}$$

where $\nu_0\in\mathcal{T}_{K_0}(\bar{x}_0)$, $\nu_0\neq 0$. By Itô's formula,

$$\begin{aligned} g_x(\bar{x}(t)) &= g_x(x_0) + \int_0^t \big(g_{xx}(\bar{x}(t))b[t] + \frac{1}{2}g_{xxx}(\bar{x}(t))\sigma[t]\sigma[t]^\top\big)dt \\ &\quad + \sum_{j=1}^d \int_0^t g_{xx}(\bar{x}(t))\sigma^j[t]dW^j(t), \quad \text{a.s.}, \forall\, t\in[0,T]. \end{aligned} \tag{5.4}$$

For $y(\cdot)$ and $g_x(\bar{x}(\cdot))$ as above, using Itô's formula again, we have

$$
\begin{aligned}
& \mathbb{E}\left\langle g_x(\bar{x}(t_1)), y(t_1)\right\rangle \\
= \; & \langle g_x(x_0), \nu_0\rangle + \mathbb{E}\int_0^{t_1} \langle g_x(\bar{x}(t)), b_x[t]y(t) + C_1\eta(t)|y(t)|b_u[t]v(t)\rangle\, dt \\
& + \mathbb{E}\int_0^{t_1} \left\langle g_x(\bar{x}(t)), C_2\frac{|\nu_0|}{t_1} b_u[t]v(t)\right\rangle dt \\
& + \mathbb{E}\int_0^{t_1} \sum_{j=1}^d \left\langle g_{xx}(\bar{x}(t))\sigma^j[t], \sigma_x^j[t]y(t) + C_1\eta(t)|y(t)|\sigma_u^j[t]v(t)\right\rangle dt \\
& + \mathbb{E}\int_0^{t_1} \sum_{j=1}^d \left\langle g_{xx}(\bar{x}(t))\sigma^j[t], C_2\frac{|\nu_0|}{t_1}\sigma_u^j[t]v(t)\right\rangle dt \\
& + \mathbb{E}\int_0^{t_1} \langle g_{xx}(\bar{x}(t))b[t], y(t)\rangle\, dt + \frac{1}{2}\mathbb{E}\int_0^{t_1} \left\langle g_{xxx}(\bar{x}(t))\sigma[t]\sigma[t]^\top, y(t)\right\rangle dt \\
\leq \; & |g_x(x_0)||\nu_0| + \mathbb{E}\int_0^{t_1} \left\langle g_x(\bar{x}(t)), C_2\frac{|\nu_0|}{t_1} b_u[t]v(t)\right\rangle dt \\
& + \mathbb{E}\int_0^{t_1} \sum_{j=1}^d \left\langle g_{xx}(\bar{x}(t))\sigma^j[t], C_2\frac{|\nu_0|}{t_1}\sigma_u^j[t]v(t)\right\rangle dt \\
& + \mathbb{E}\int_0^{t_1} \Big(\langle g_x(\bar{x}(t)), b_x[t]y(t)\rangle + \sum_{j=1}^d \left\langle g_{xx}(\bar{x}(t))\sigma^j[t], \sigma_x^j[t]y(t)\right\rangle \\
& + \langle g_{xx}(\bar{x}(t))b[t], y(t)\rangle + \frac{1}{2}\left\langle g_{xxx}(\bar{x}(t))\sigma[t]\sigma[t]^\top, y(t)\right\rangle \Big) dt \\
& + \mathbb{E}\int_0^{t_1} C_1\eta(t)|y(t)|\Big(\langle g_x(\bar{x}(t)), b_u[t]v(t)\rangle + \sum_{j=1}^d \left\langle g_{xx}(\bar{x}(t))\sigma^j[t], \sigma_u^j[t]v(t)\right\rangle \Big) dt \\
\leq \; & |g_x(x_0)||\nu_0| - \rho C_2|\nu_0| + \mathbb{E}\int_0^{t_1} \Big[|g_x(\bar{x}(t))||b_x[t]| + \sum_{j=1}^d |g_{xx}(\bar{x}(t))\sigma^j[t]||\sigma_x^j[t]| \\
& + |g_{xx}(\bar{x}(t))b[t]| + \frac{1}{2}|g_{xxx}(\bar{x}(t))\sigma[t]\sigma[t]^\top| \Big] |y(t)|dt - \rho C_1\mathbb{E}\int_0^{t_1} \eta(t)|y(t)|dt \\
\leq \; & -|\nu_0| - \mathbb{E}\int_0^{t_1} \eta(t)|y(t)|dt < 0. \qquad (5.5)
\end{aligned}
$$

(b) If there is a $t \in [0, t_1)$ such that $\mathbb{E}\; g(\bar{x}(t)) \leq -\delta$ then let

$$
t_0 = \sup\Big\{ t \in [0, t_1) \,\Big|\, \mathbb{E}\; g(\bar{x}(t)) \leq -\delta \Big\}.
$$

By the continuity of $\mathbb{E}\; g(\bar{x}(\cdot))$, $t_0 < t_1$ and $\mathbb{E}\; g(\bar{x}(t_0)) = -\delta$.

Let $\nu_0 \in \mathcal{T}_{K_0}(\bar{x}_0)$, $\nu_0 \neq 0$ and define y on $[0, t_0] \times \Omega$ by the following stochastic differential equation:

$$
\begin{cases}
dy(t) = b_x[t]y(t)dt + \sum\limits_{j=1}^d \sigma_x^j[t]y(t)dW^j(t), \; t \in [0, t_0], \\
y(0) = \nu_0.
\end{cases} \qquad (5.6)
$$

Letting $S(\cdot)$ be the solution to the matrix-valued stochastic differential equation:

$$
\begin{cases}
dS(t) = b_x[t]S(t)dt + \sum\limits_{j=1}^d \sigma_x^j[t]S(t)dW^j(t), \; t \in [0, t_0], \\
S(0) = I,
\end{cases}
$$

by [27, Theorem 6.14, p.47], for a.e. $(t,\omega)\in[0,t_0]\times\Omega$, $S(t)$ is invertible and

$$y(t)=S(t)\nu_0,\quad \text{a.e. } t\in[0,t_0],\ \text{a.s.}$$

Since $\nu_0\neq 0$, we have $y(t)\neq 0$ for any $t\in[0,t_0]$. In particular, $y(t_0)\neq 0$. Denote $\xi_{t_0}=y(t_0)$ and let y be the solution of the following stochastic differential equation on $[t_0,t_1]\times\Omega$:

$$\begin{cases} dy(t)=\big(b_x[t]y(t)+\dfrac{L+2}{\rho}\eta(t)|y(t)|b_u[t]v(t)\big)dt \\ \qquad\qquad +\dfrac{\big[\big(\mathbb{E}|g_x(\bar{x}(t_0))|^2\big)^{\frac12}+1\big]\big(\mathbb{E}|\xi_{t_0}|^2\big)^{\frac12}}{\rho(t_1-t_0)}b_u[t]v(t)dt \\ \qquad\qquad +\sum\limits_{j=1}^{d}\big(\sigma_x^j[t]y(t)+\dfrac{L+2}{\rho}\eta(t)|y(t)|\sigma_u^j[t]v(t)\big)dW^j(t) \\ \qquad\qquad +\dfrac{\big[\big(\mathbb{E}|g_x(\bar{x}(t_0))|^2\big)^{\frac12}+1\big]\big(\mathbb{E}|\xi_{t_0}|^2\big)^{\frac12}}{\rho(t_1-t_0)}\sum\limits_{j=1}^{d}\sigma_u^j[t]v(t)dW^j(t),\ t\in[t_0,t_1], \\ y(t_0)=\xi_{t_0}. \end{cases} \tag{5.7}$$

Similarly to (5.5), we show that

$$\mathbb{E}\ \langle g_x(\bar{x}(t_1)),y(t_1)\rangle\leq-\big(\mathbb{E}\ |\xi_{t_0}|^2\big)^{\frac12}-\mathbb{E}\int_{t_0}^{t_1}\eta(t)|y(t)|dt<0.$$

Step 2: Denote by ξ_{t_1} the random variable $y(t_1)$ obtained in Step 1. Obviously, $\xi_{t_1}\neq 0$. Define

$$t_2=\max\Big\{t\in(t_1,T]\ \Big|\ \mathbb{E}\ g(\bar{x}(s))\in[-\delta,0],\ \forall\ s\in[t_1,t]\Big\}.$$

In the sequel, we consider two different cases.

(a) If $t_2=T$, we define y on $[t_1,T]\times\Omega$ by

$$\begin{cases} dy(t)=\big(b_x[t]y(t)+\dfrac{L+2}{\rho}\eta(t)|y(t)|b_u[t]v(t)\big)dt \\ \qquad\qquad +\sum\limits_{j=1}^{d}\big(\sigma_x^j[t]y(t)+\dfrac{L+2}{\rho}\eta(t)|y(t)|\sigma_u^j[t]v(t)\big)dW^j(t),\quad t\in[t_1,T], \\ y(t_1)=\xi_{t_1}. \end{cases} \tag{5.8}$$

Then, by Itô's formula, for any $t\in[t_1,T]$,

$$\begin{aligned} &\mathbb{E}\,\langle g_x(\bar{x}(t)),y(t)\rangle \\ =\ &\mathbb{E}\,\langle g_x(\bar{x}(t_1)),y(t_1)\rangle+\mathbb{E}\int_{t_1}^{t}\Big\langle g_x(\bar{x}(s)),b_x[s]y(s)+\frac{L+2}{\rho}\eta(s)|y(s)|b_u[s]v(s)\Big\rangle ds \\ &+\mathbb{E}\int_{t_1}^{t}\sum_{j=1}^{d}\Big\langle g_{xx}(\bar{x}(s))\sigma^j[s],\sigma_x^j[s]y(s)+\frac{L+2}{\rho}\eta(s)|y(s)|\sigma_u^j[s]v(s)\Big\rangle ds \\ &+\mathbb{E}\int_{t_1}^{t}\langle g_{xx}(\bar{x}(s))b[s],y(s)\rangle\,ds+\frac12\mathbb{E}\int_{t_1}^{t}\big\langle g_{xxx}(\bar{x}(s))\sigma[s]\sigma[s]^\top,y(s)\big\rangle\,ds \\ <\ &\mathbb{E}\int_{t_1}^{t}\Big(\langle g_x(\bar{x}(s)),b_x[s]y(s)\rangle+\sum_{j=1}^{d}\big\langle g_{xx}(\bar{x}(s))\sigma^j[s],\sigma_x^j[s]y(s)\big\rangle \\ &+\langle g_{xx}(\bar{x}(s))b[s],y(s)\rangle+\frac12\big\langle g_{xxx}(\bar{x}(s))\sigma[s]\sigma[s]^\top,y(s)\big\rangle\Big)ds \\ &+\mathbb{E}\int_{t_1}^{t}\frac{L+2}{\rho}\eta(s)|y(s)|\Big(\langle g_x(\bar{x}(s)),b_u[s]v(s)\rangle+\sum_{j=1}^{d}\big\langle g_{xx}(\bar{x}(s))\sigma^j[s],\sigma_u^j[s]v(s)\big\rangle\Big)ds \end{aligned}$$

$$< \quad -\mathbb{E}\int_{t_1}^{t}\eta(s)|y(s)|ds \le 0. \tag{5.9}$$

(b) If $t_2 \neq T$, then $-\delta = \mathbb{E}\ g(\bar{x}(t_2))$ and there exists an $s \in (t_2, T]$ such that $\mathbb{E}\ g(\bar{x}(s)) < -\delta$. Let y be the solution to the equation (5.8) on $[t_1, t_2] \times \Omega$, and denote $\xi_{t_2} = y(t_2)$. By (5.9), $\xi_{t_2} \neq 0$.

If $\mathbb{E}g(\bar{x}(t)) < 0$ for any $t \in [t_2, T]$, then on $[t_2, T] \times \Omega$ we simply let $y(\cdot)$ be the solution to (3.2) with initial condition $y(t_2) = \xi_{t_2}$ and control $v \equiv 0$ (on $[t_2, T] \times \Omega$). Otherwise, there exists an $s \in (t_2, T]$ such that $\mathbb{E}\ g(\bar{x}(s)) = 0$. Define

$$t_3 = \min\Big\{t \in [t_2, T]\ \Big|\ \mathbb{E}\ g(\bar{x}(t)) = 0\Big\}, \qquad t_4 = \max\Big\{t \in [t_2, t_3]\ \Big|\ \mathbb{E}\ g(\bar{x}(s)) \le -\delta\Big\}.$$

By the definition of t_2, t_3 and t_4 we have $t_2 < t_4 < t_3$. Similarly to Step 1, we construct y on $[t_2, t_3] \times \Omega$ such that $\mathbb{E}\ \langle g_x(\bar{x}(t_3)), y(t_3)\rangle < 0$ as follows:

On $[t_1, t_2] \times \Omega$, let y be the solution to the equation (5.8) (with $y(t_2) = \xi_{t_2}$) and define y on $[t_2, t_4] \times \Omega$ by the following stochastic differential equation:

$$\begin{cases} dy(t) = b_x[t]y(t)dt + \sum\limits_{j=1}^{d}\sigma_x^j[t]y(t)dW^j(t),\ t \in [t_2, t_4], \\ y(t_2) = \xi_{t_2}, \end{cases} \tag{5.10}$$

and denote $\xi_{t_4} = y(t_4)$. By (5.9) (with t replaced by t_2), $\xi_{t_2} \neq 0$, $\xi_{t_4} \neq 0$. Then, define y on $[t_4, t_3] \times \Omega$ by the following stochastic differential equation:

$$\begin{cases} dy(t) = \big(b_x[t]y(t) + \dfrac{L+2}{\rho}\eta(t)|y(t)|b_u[t]v(t)\big)dt \\ \qquad\qquad + \dfrac{\big[\big(\mathbb{E}|g_x(\bar{x}(t_4))|^2\big)^{\frac{1}{2}} + 1\big]\big(\mathbb{E}|\xi_{t_4}|^2\big)^{\frac{1}{2}}}{\rho(t_3 - t_4)} b_u[t]v(t)dt \\ \qquad\qquad + \sum\limits_{j=1}^{d}\big(\sigma_x^j[t]y(t) + \dfrac{L+2}{\rho}\eta(t)|y(t)|\sigma_u^j[t]v(t)\big)dW^j(t) \\ \qquad\qquad + \dfrac{\big[\big(\mathbb{E}|g_x(\bar{x}(t_4))|^2\big)^{\frac{1}{2}} + 1\big]\big(\mathbb{E}|\xi_{t_4}|^2\big)^{\frac{1}{2}}}{\rho(t_3 - t_4)} \sum\limits_{j=1}^{d}\sigma_u^j[t]v(t)dW^j(t),\ t \in [t_4, t_3], \\ y(t_4) = \xi_{t_4}. \end{cases} \tag{5.11}$$

Then, $\mathbb{E}\,\langle g_x(\bar{x}(t_3)), y(t_3)\rangle < 0$ and we return to the beginning of the Step 2.

Since $[0, T]$ is a finite interval and $\mathbb{E}\ g(\bar{x}(\cdot))$ is continuous, repeating the above arguments a finite number of times we construct $y \in \mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)}$. □

In the following we give an example in which the assumptions (A4)–(A5) are satisfied.

Example 5.1. *Let $n = 1$, $m = 2$, U be a subset in $\mathbb{R}^2$ defined by*

$$U := \Big\{(u_1, u_2)^\top \in \mathbb{R}^2\ \Big|\ u_1^2 + u_2^2 \ge \frac{1}{2}, u_1 + u_2 \le 1, u_1 \ge 0, u_2 \ge 0\Big\},$$

$B = (\frac{1}{2}, 1) \in \mathbb{R}^2$, $D^j = (0, \sigma^j) \in \mathbb{R}^2$, $j = 1, \ldots, d$, $\theta \in \mathbb{R}$, $x_0 \in \mathbb{R}$ and $x_0 > 0$. Consider the control system

$$\begin{cases} dx(t) = x(t)Bu(t)dt + \sum\limits_{j=1}^{d} x(t)D^j u(t)dW^j(t), \quad t \in [0, T], \\ x(0) = x_0, \end{cases}$$

and the state constraint

$$\mathbb{E}\ x(t) \ge \theta x_0, \quad a.e.\ t \in [0, T].$$

Let $\widehat{U} := \big\{(u_1, u_2)^\top \in \mathbb{R}^2\ \big|\ u_1 + u_2 = 1, u_1 > 0,\ u_2 \ge 0\big\}(\subset U)$. Define

$$\widehat{\mathcal{U}} := \big\{u \in \mathcal{U}\ \big|\ u(t, \omega) \in \widehat{U},\ a.e.\ (t, \omega) \in [0, T] \times \Omega\big\}.$$

Clearly, for this problem,

$$g(x) = \theta x_0 - x, \quad x \in \mathbb{R}, \tag{5.12}$$

and for any control $\hat{u} \in \widehat{\mathcal{U}}$ *with the corresponding state* $\hat{x}$*,*

$$g_x(\hat{x}(t)) \equiv -1, \quad g_{xx}(\hat{x}(t)) = g_{xxx}(\hat{x}(t)) \equiv 0.$$

Letting $\eta \equiv 1$*, we find that (A4) holds true.*

It is easy to check that, for any $\hat{u} \in \widehat{\mathcal{U}}$*,*

$$T_U^b(\hat{u}(t,\omega)) = \begin{cases} \big\{\alpha(-1,0)^\top + \beta(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^\top \;\big|\; \alpha \geq 0,\ \beta \geq 0\big\}, & \hat{u}(t,\omega) = (1,0)^\top; \\ \big\{(v^1, v^2)^\top \in \mathbb{R}^2 \;\big|\; v^1 + v^2 = 0\big\}, & \hat{u}(t,\omega) = (\frac{1}{2}, \frac{1}{2})^\top; \\ \big\{(v^1, v^2)^\top \in \mathbb{R}^2 \;\big|\; v^1 + v^2 \leq 0\big\}, & \text{otherwise.} \end{cases} \tag{5.13}$$

By (5.13), $T_U^b(\hat{u}(\cdot))$ *is* $\mathcal{B}([0,T]) \otimes \mathcal{F}$*-measurable and* $\mathbb{F}$*-adapted, and for a.e.*$(t,\omega) \in [0,T] \times \Omega$*,* $T_U^b(\hat{u}(t,\omega))$ *is a nonempty closed convex cone.*

For any $t \in [0,T]$*, we have*

$$\hat{x}(t) = x_0 \exp\Big[\int_0^t \big(B\hat{u}(s) - \frac{1}{2}\sum_{j=1}^d |D^j\hat{u}(s)|^2\big)ds + \sum_{j=1}^d \int_0^t D^j\hat{u}(s)dW^j(s)\Big] > 0, \quad a.s.$$

Let I_δ *be defined by (5.1) (with* $\bar{x}$ *replaced by* $\hat{x}$ *and* g *defined by (5.12)). Assume that there exists a constant* C *such that*

$$\hat{x}(t) \geq C > 0, \quad a.e.\ (t,\omega) \in [0,T] \times \Omega. \tag{5.14}$$

(Clearly, the condition (5.14) is trivially satisfied whenever $\hat{u}_2 = 0$*. On the other hand, by means of the martingale property, it is easy to see that this condition is equivalent to the essential boundedness (with respect to the probability measure* P*) of* $\sum_{j=1}^d \int_0^T D^j\hat{u}(s)dW^j(s)$*. Note that, for any* $\xi \in L^\infty_{\mathcal{F}_T}(\Omega;\mathbb{R})$ *with mean 0, by Martingale Representation Theorem, there exists a* $\tilde{u} \in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}^2))$ *so that* $\xi = \sum_{j=1}^d \int_0^T D^j\tilde{u}(s)dW^j(s)$*. If, further, this* $\tilde{u}$ *belongs to* $\widehat{\mathcal{U}}$*, then (5.14) is satisfied with the corresponding* $\hat{u}$ *replaced by* $\tilde{u}$*.) If*

$$I(\hat{x}) := \big\{t \in [0,T] \;\big|\; \mathbb{E}\ \hat{x}(t) = \theta x_0\big\} = \emptyset$$

(e.g., $C > \theta x_0$*), there is no need to verify the condition (A5). If* $I(\hat{x}) \neq \emptyset$*, letting* $\Phi(\cdot) = T_U^b(\hat{u}(\cdot))$ *and choosing* $v = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^\top$*, we have*

$$-\hat{x}(t)Bv = -\frac{\hat{x}(t)}{2\sqrt{2}} \leq -\frac{C}{2\sqrt{2}} < 0,\ a.e.\ (t,\omega) \in I_\delta \times \Omega,$$

i.e., (A5) holds true for $\rho = \frac{C}{2\sqrt{2}}$*.*

Note that the inward pointing condition (A5) is only a sufficient condition but not a necessary condition for $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \neq \emptyset$. We give next an example in which the condition (A5) does not hold but $\mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \neq \emptyset$.

Example 5.2. *Let* $T = 3$*,* $m = n = 3$ *and* $K_0 = \{0\}(\subset \mathbb{R}^3)$*. Define*

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let us consider the following control system

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t),\ t \in [0,3], \\ x(0) = 0 \end{cases} \tag{5.15}$$

with the control set $U=\mathbb{R}^3$, the cost functional

$$\mathbb{E}\ \phi(x(3))=\frac{1}{2}\mathbb{E}\,|x^1(3)-2|^2$$

and the state constraint

$$g(x(t))=x^2(t)-1\le 0,\quad \forall\, t\in[0,T].$$

Let $\bar u(\cdot)=(\bar u^1(\cdot),\bar u^2(\cdot),\bar u^3(\cdot))^\top$ be defined by

$$\bar u^1(t)=\bar u^2(t)\equiv 0,\quad \bar u^3(t)=\begin{cases}1, & t\in[0,1);\\ -1, & t\in[1,2);\\ 0, & t\in[2,3].\end{cases}$$

Then, the corresponding state $\bar x(\cdot)=(\bar x^1(\cdot),\bar x^2(\cdot),\bar x^3(\cdot))^\top$ is given by

$$\bar x^1(t)=\begin{cases}\frac{t^3}{6}, & t\in[0,1);\\ -\frac{t^3}{6}+t^2-t+\frac{1}{3}, & t\in[1,2);\\ t-1, & t\in[2,3];\end{cases}$$

$$\bar x^2(t)=\begin{cases}\frac{t^2}{2}, & t\in[0,1);\\ -\frac{t^2}{2}+2t-1, & t\in[1,2);\\ 1, & t\in[2,3];\end{cases}\qquad \bar x^3(t)=\begin{cases}t, & t\in[0,1);\\ 2-t, & t\in[1,2);\\ 0, & t\in[2,3].\end{cases}$$

Clearly, $(\bar x,\bar u)$ is a global minimizer and $I_0^g=[2,3]$.

Let $\Phi(t)\equiv\mathbb{R}^3$. For the above $(\bar x,\bar u)$ the condition (A5) becomes:

$$\min_{v\in B_{\mathbb{R}^3}}\langle g_x(\bar x(t)),b_u[t]v\rangle\le-\rho,\quad \forall\, t\in I_\delta,$$

where $I_\delta:=\big\{t\in[0,T]\,|\,g(\bar x(t))\in[-\delta,0]\big\}$ (for some $\rho>0$ and $\delta>0$).

Obviously, $g_x(\bar x(t))=(0,1,0)^\top$, $b_u[t]=B$, and $\langle g_x(\bar x(t)),b_u[t]v\rangle=0$ for any $v\in B_{\mathbb{R}^3}$, i.e., the condition (A5) does not hold in this case.

Next we show that $\mathcal{G}^{(1)}\cap\mathcal{Q}^{(1)}\neq\emptyset$. Clearly, the corresponding first order linearized equation is:

$$\begin{cases}\dot y(t)=Ay(t)+Bv(t),\ t\in[0,3],\\ y(0)=0,\end{cases}\tag{5.16}$$

where $v(\cdot)\in L^2(0,T;\mathbb{R}^3)$.

Let $y(0)=0$, $v(t)\equiv(0,0,-1)^\top$. Then, we have $y^2(t)=-\dfrac{t^2}{2}$ and

$$\langle g_x(\bar x(t)),y(t)\rangle=-\frac{t^2}{2}<0,\quad\forall\, t\in(0,3],$$

implying that $\mathcal{G}^{(1)}\cap\mathcal{Q}^{(1)}\neq\emptyset$.

We propose next a sufficient condition for $\mathcal{G}^{(1)}\cap\mathcal{Q}^{(1)}\cap\mathcal{E}^{(1)}\neq\emptyset$.

For any $(t,\omega)\in[0,T]\times\Omega$, let Φ be a set-valued map satisfying (3.1) and denote

$$\Phi_\rho^S(t,\omega):=\Big\{v\in\Phi(t,\omega)\big|\,\langle g_x(\bar x(t)),b_u[t]v\rangle+\sum_{j=1}^d\big\langle g_{xx}(\bar x(t))\sigma^j[t],\sigma_u^j[t]v\big\rangle\le-\rho\Big\},$$

and

$$\Phi_\rho^F(t,\omega):=\Big\{v\in\Phi(t,\omega)\big|\,\langle\varphi_x^\ell(\bar x(t)),b_u[t]v\rangle+\sum_{j=1}^d\big\langle\varphi_{xx}^\ell(\bar x(t))\sigma^j[t],\sigma_u^j[t]v\big\rangle\le-\rho,\ \ell\in\mathcal{I}_0^\varphi\Big\}.$$

Let I_δ^g be defined by (5.1) and

$$\Theta(t,\omega) := \begin{cases} \Phi_\rho^S(t,\omega) \cap B_{\mathbb{R}^m}, & \text{if } (t,\omega) \in I_\delta^g \times \Omega, \\ \Phi(t,\omega) \cap B_{\mathbb{R}^m}, & \text{otherwise.} \end{cases} \tag{5.17}$$

By the definition of Φ and the assumption (A5), $\Theta(t,\omega) \neq \emptyset$ for a.e. $(t,\omega) \in [0,T] \times \Omega$.

We assume that

(A6) *Functions $\varphi^\ell(\cdot): \mathbb{R}^n \to \mathbb{R}$, $\ell \in \mathcal{I}_0^\varphi$ are continuously differentiable up to the third order, and, for some $\epsilon > 0$, $\eta \in L_{\mathbb{F}}^\infty([0,T] \times \Omega; \mathbb{R}^n)$ as in (A4),*

$$\begin{aligned} &\left|\varphi_x^\ell(\bar{x}(t))\right| + \left|\varphi_{xx}^\ell(\bar{x}(t))b[t]\right| + \sum_{j=1}^d \left|\varphi_{xx}^\ell(\bar{x}(t))\sigma^j[t]\right| \\ &+\frac{1}{2}\left|\varphi_{xxx}^\ell(\bar{x}(t))\sigma[t]\sigma[t]^\top\right| \leq \eta(t), \ \text{ a.e. } t \in [T-\epsilon, T], \text{ a.s. }, \quad \forall\, \ell \in \mathcal{I}_0^\varphi. \end{aligned}$$

(A7) There exists an $\epsilon > 0$ such that, for a.e. $(t,\omega) \in [T-\epsilon, T] \times \Omega$,

$$\Theta(t,\omega) \bigcap \Phi_\rho^F(t,\omega) \neq \emptyset.$$

Theorem 5.2. *In Theorem 5.1, assume that also (A6)–(A7) hold true. Then, $\mathcal{Q}^{(1)} \cap \mathcal{G}^{(1)} \cap \mathcal{E}^{(1)} \neq \emptyset$.*

Proof. We only need to consider the case $\mathcal{I}_0^\varphi \neq \emptyset$.

By the proof of Theorem 5.1, for any $\tau \in [0,T)$, there exists a solution y to (3.2) on $[0,\tau] \times \Omega$ satisfying

$$y(\tau) \neq 0, \quad \mathbb{E}\ \langle g_x(\bar{x}(s)), y(s)\rangle < 0, \quad \forall\, s \in [0,\tau] \cap I_0^g. \tag{5.18}$$

Let $\tilde{v} \in L_{\mathbb{F}}^\infty([T-\varepsilon, T] \times \Omega; \mathbb{R}^m)$ be such that

$$\tilde{v}(t,\omega) \in \Theta(t,\omega) \bigcap \Phi_\rho^F(t,\omega), \quad \forall\, (t,\omega) \in [T-\varepsilon, T] \times \Omega.$$

By the assumption (A7), such $\tilde{v}$ exists.

Now, we shall prove the conclusion by considering two different cases.

case a: $\mathbb{E}\ g(\bar{x}(T)) < 0$. Then for sufficiently small $\varepsilon \in [0,\epsilon]$, $\mathbb{E}g(\bar{x}(s)) < 0$ for all $s \in [T-\varepsilon, T]$, i.e. $I_0^g \cap [T-\varepsilon, T] = \emptyset$.

Let y be the stochastic process satisfying (5.18) with τ replaced by $T-\varepsilon$ and let $\bar{y} = y$ on $[0, T-\varepsilon] \times \Omega$. Define $\xi_{T-\varepsilon} = y(T-\varepsilon)$, $\alpha_{T-\varepsilon} := \sum_{\ell \in \mathcal{I}_0^\varphi} \left(\mathbb{E}\ |\varphi_x^\ell(\bar{x}(T-\varepsilon))|^2\right)^{\frac{1}{2}} + 1$, and, on $[T-\varepsilon, T] \times \Omega$, consider the following stochastic differential equation

$$\begin{cases} d\bar{y}(t) = \big(b_x[t]\bar{y}(t) + \dfrac{L+2}{\rho}\eta(t)|\bar{y}(t)|b_u[t]\tilde{v}(t)\big)dt + \dfrac{\alpha_{T-\varepsilon}\left(\mathbb{E}\ |\xi_{T-\varepsilon}|^2\right)^{\frac{1}{2}}}{\rho\varepsilon} b_u[t]\tilde{v}(t)dt \\ \qquad + \sum\limits_{j=1}^d \big(\sigma_x^j[t]\bar{y}(t) + \dfrac{L+2}{\rho}\eta(t)|\bar{y}(t)|\sigma_u^j[t]\tilde{v}(t)\big)dW^j(t) \\ \qquad + \dfrac{\alpha_{T-\varepsilon}\left(\mathbb{E}\ |\xi_{T-\varepsilon}|^2\right)^{\frac{1}{2}}}{\rho\varepsilon} \sum\limits_{j=1}^d \sigma_u^j[t]\tilde{v}(t)dW^j(t), \ t \in [T-\varepsilon, T], \\ \bar{y}(T-\varepsilon) = \xi_{T-\varepsilon}. \end{cases} \tag{5.19}$$

Similarly to (5.5), we show that

$$\mathbb{E}\left\langle \varphi_x^\ell(\bar{x}(T)), \bar{y}(T)\right\rangle < 0, \quad \forall\, \ell \in \mathcal{I}_0^\varphi.$$

case b: $\mathbb{E}\ g(\bar{x}(T)) = 0$. Then, there exists an $\varepsilon > 0$ such that $(T-\varepsilon, T] \subset I_\delta^g$ (Recall that I_δ^g is defined by (5.1)). Without loss of generality, we assume that $\varepsilon \leq \epsilon$.

Thus $\mathbb{E}\, g(\bar{x}(t)) \in [-\delta, 0]$ for any $t \in [T-\varepsilon, T]$).

If $I_0^g \cap [T-\varepsilon, T) = \emptyset$, then, let y be the stochastic process satisfying (5.18) with $\tau = T-\varepsilon$. Then, $y(T-\varepsilon) \neq 0$. Denote $\xi_{T-\varepsilon} = y(T-\varepsilon)$. Define

$$\alpha_{T-\varepsilon} := 1 + \big(\mathbb{E}|g_x(\bar{x}(T-\varepsilon))|^2\big)^{\frac{1}{2}} + \sum_{\ell \in \mathcal{I}_0^\varphi} \big(\mathbb{E}\, |\varphi_x^\ell(\bar{x}(T-\varepsilon))|^2\big)^{\frac{1}{2}} \tag{5.20}$$

and, on $[T-\varepsilon, T] \times \Omega$, let $\bar{y}$ be the solution to (5.19) with $\alpha_{T-\varepsilon}$ defined by (5.20). By Itô's formula, as before, we show that

$$\begin{aligned} \mathbb{E}\, \langle g_x(\bar{x}(T)), \bar{y}(T)\rangle &\leq \mathbb{E}\, \langle g_x(\bar{x}(T-\varepsilon)), \xi_{T-\varepsilon}\rangle - \alpha_{T-\varepsilon}\big(\mathbb{E}|\xi_{T-\varepsilon}|^2\big)^{\frac{1}{2}} - \mathbb{E}\int_{T-\varepsilon}^T \eta(s)|\bar{y}(s)|ds \\ &< 0, \end{aligned} \tag{5.21}$$

and for all $\ell \in \mathcal{I}_0^\varphi$,

$$\mathbb{E}\, \big\langle \varphi_x^\ell(\bar{x}(T)), \bar{y}(T)\big\rangle \leq \mathbb{E}\, \big\langle \varphi_x^\ell(\bar{x}(T-\varepsilon)), \xi_{T-\varepsilon}\big\rangle - \alpha_{T-\varepsilon}\big(\mathbb{E}|\xi_{T-\varepsilon}|^2\big)^{\frac{1}{2}} - \mathbb{E}\int_{T-\varepsilon}^T \eta(s)|\bar{y}(s)|ds < 0$$

which implies that $\bar{y} \in \mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)}$.

If $I_0^g \cap [T-\varepsilon, T) \neq \emptyset$. Define

$$t_F := \min\{t \in [T-\varepsilon, T] \,|\, \mathbb{E}\; g(\bar{x}(t)) = 0\}, \qquad \alpha_{t_F} := 1 + \sum_{\ell \in \mathcal{I}_0^\varphi} \big(\mathbb{E}\, |\varphi_x^\ell(\bar{x}(t_F))|^2\big)^{\frac{1}{2}}.$$

Let y be the stochastic process satisfying (5.18) with $\tau = t_F$. Then, $y(t_F) \neq 0$ and $\mathbb{E}\, \langle g_x(\bar{x}(t_F)), y(t_F)\rangle < 0$. Denote $\xi_{t_F} = y(t_F)$, and, on $[t_F, T] \times \Omega$, let $\bar{y}$ be the solution to the following equation:

$$\begin{cases} d\bar{y}(t) = \big(b_x[t]\bar{y}(t) + \dfrac{L+2}{\rho}\eta(t)|\bar{y}(t)|b_u[t]\tilde{v}(t)\big)dt + \dfrac{\alpha_{t_F}\big(\mathbb{E}|\xi_{t_F}|^2\big)^{\frac{1}{2}}}{\rho(T-t_F)} b_u[t]\tilde{v}(t)dt \\ \qquad\qquad + \sum\limits_{j=1}^d \big(\sigma_x^j[t]\bar{y}(t) + \dfrac{L+2}{\rho}\eta(t)|\bar{y}(t)|\sigma_u^j[t]\tilde{v}(t)\big)dW^j(t) \\ \qquad\qquad + \dfrac{\alpha_{t_F}\big(\mathbb{E}|\xi_{t_F}|^2\big)^{\frac{1}{2}}}{\rho(T-t_F)} \sum\limits_{j=1}^d \sigma_u^j[t]\tilde{v}(t)dW^j(t), \;\; t \in [t_F, T], \\ \bar{y}(t_F) = \xi_{t_F}. \end{cases} \tag{5.22}$$

By Itô's formula, for any $t \in [t_F, T]$

$$\mathbb{E}\, \langle g_x(\bar{x}(t)), \bar{y}(t)\rangle \leq \mathbb{E}\, \langle g_x(\bar{x}(t_F)), \xi_{t_F}\rangle - \frac{\alpha_{t_F}(t-t_F)\big(\mathbb{E}|\xi_{t_F}|^2\big)^{\frac{1}{2}}}{T-t_F} - \mathbb{E}\int_{t_F}^t \eta(s)|\bar{y}(s)|ds < 0,$$

and

$$\mathbb{E}\, \big\langle \varphi_x^\ell(\bar{x}(T)), \bar{y}(T)\big\rangle \leq \mathbb{E}\, \big\langle \varphi_x^\ell(\bar{x}(t_F)), \xi_{t_F}\big\rangle - \alpha_{t_F}\big(\mathbb{E}|\xi_{t_F}|^2\big)^{\frac{1}{2}} - \mathbb{E}\int_{t_F}^T \eta(s)|\bar{y}(s)|ds < 0, \quad \forall\, \ell \in \mathcal{I}_0^\varphi,$$

implying that $\bar{y} \in \mathcal{G}^{(1)} \cap \mathcal{Q}^{(1)} \cap \mathcal{E}^{(1)}$.

□

References

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

[2] J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, 1984.

[3] T. R. Bielecki, H. Jin, S. R. Pliska and X. Y. Zhou, *Continuous-time mean-variance portfolio selection with bankruptcy prohibition*, Math. Finance, 15 (2005), pp. 213–244.

[4] J. F. Bonnans, *Second order Pontryagin's principle for stochastic control problems*, Inria Saclay Ile de France. 2015, https://hal.inria.fr/hal-01205854.

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

[6] A.Ya. Dubovitskii and A.A. Milyutin, *Extremum problems in the presence of restrictions*, Zh. Vychisl. Mat. Mat. Fiz. (USSR Comput. Math. and Math. Phys.), 5 (1965), pp. 1–80.

[7] A. Cernea and H. Frankowska, *A connection between the maximum principle and dynamic programming for constrained control problems*, SIAM J. Control Optim. 44 (2005), pp. 673–703.

[8] J. Cvitanić and I. Karatzas, *Convex duality in constrained portfolio optimization*, Ann. Appl. Probab., 2 (1992), pp. 767–818.

[9] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential. B. Theory of Martingales*, Translated from the French by J. P. Wilson. North-Holland Mathematics Studies, 72. North-Holland Publishing Co., Amsterdam, 1982.

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

[11] N. El Karoui, S. Peng and M. C. Quenez, *A dynamic maximum principle for the optimization of recursive utilities under constraints*, Ann. Appl. Probab., 11 (2001), pp. 664–693.

[12] H. Frankowska, D. Hoehener and D. Tonon, *A second-order maximum principle in optimal control under state constraints*, Serdica Math. J., 39 (2013), pp. 233–270.

[13] Frankowska H. & Osmolovskii N. P. *Second-order necessary conditions for a strong local minimum in a problem with state and general control constraints*, SIAM J. Control Optim., in press.

[14] H. Frankowska and D. Tonon, *Inward pointing trajectories, normality of the maximum principle and the non occurrence of the Lavrentieff phenomenon in optimal control under state constraints*, J. Convex Anal., 20 (2013), pp. 1147–1180.

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

[16] H. Frankowska, H. Zhang and X. Zhang, *Stochastic optimal control problems with control and initial-final states constraints*, SIAM J. Control Optim., in press.

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

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

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

[20] M. Kisielewicz, *Stochastic Differential Inclusions and Applications*, Springer, 2013.

[21] B. Øksendal, *Stochastic Differential Equations. An Introduction with Applications*, Springer-Verlag, Berlin-Heidelberg-. New York 2000 (5th ed.)

[22] Z. Páles and V. Zeidan, *First- and second-order necessary conditions for control problems with constraints*, Trans. Amer. Math. Soc., 346 (1994), pp. 421–453.

[23] Z. Páles and V. Zeidan, *Optimum problems with certain lower semicontinuous set-valued constraints*, SIAM J. Optim. 8 (1998), pp. 707–727.

[24] P. Protter, *Stochastic Integration and Differential equations*, Second edition, Corrected third printing, Springer-Verlag, Berlin, 2005.

[25] R. Vinter, *Optimal Control*, Birkhäuser Boston, Inc., Boston, MA, 2000.

[26] T. Wang and H. Zhang, *Optimal control problems for forward-backward stochastic Volterra integral equations with closed control regions*, SIAM J. Control Optim. 55 (2017), pp. 2574–2602.

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