STOCHASTIC OPTIMAL CONTROL PROBLEMS WITH CONTROL AND INITIAL-FINAL STATES CONSTRAINTS

Hélène Frankowska, Haisen Zhang, X. Zhang

To cite this version:

Hélène Frankowska, Haisen Zhang, X. Zhang. STOCHASTIC OPTIMAL CONTROL PROBLEMS WITH CONTROL AND INITIAL-FINAL STATES CONSTRAINTS. SIAM Journal on Control and Optimization, 2018, 56, pp.1823-1855. hal-02126110

HAL Id: hal-02126110
https://hal.science/hal-02126110v1

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.

STOCHASTIC OPTIMAL CONTROL PROBLEMS WITH CONTROL AND INITIAL-FINAL STATES CONSTRAINTS

HÉLÈNE FRANKOWSKA*, HAISEN ZHANG†, AND XU ZHANG‡

Abstract. In this paper, the first and second order necessary optimality conditions are established for stochastic optimal control problems with control and initial-final states constraints. The control regions are allowed to be nonconvex, the diffusion terms contain the control variable and the final state constraints are defined by finitely many inequality constraints. In the difference with the existing literatures, the second order variations of the control set are used to derive the second order necessary conditions. This leads to stronger results under less restrictive, than usual, assumptions.

Key words. Stochastic optimal control, normal first order necessary optimality conditions, second order necessary conditions, second order tangents.

AMS subject classifications. Primary 93E20; Secondary 49J53, 60H07, 60H10.

1. Introduction. Let $d, m, n, k \in \mathbb{N}$, $T > 0$ 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 $W(\cdot)$ is defined such that $\mathbb{F}$ is the natural filtration generated by $W(\cdot)$ (augmented by all the P-null sets). Denote by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ respectively the inner product and norm in $\mathbb{R}^m$ or $\mathbb{R}^n$, which can be identified from the context, and by $\mathcal{B}(X)$ the Borel σ-field of a metric space X.

Let us consider the following controlled stochastic differential equation

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

with the Mayer-type cost functional

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

and end points constraints

$$x_0 \in K_0, \quad \mathbb{E}\ g^i(x(T)) \le 0,\ \forall\ i = 1, \cdots, k. \tag{1.3}$$

Here $u(\cdot) \in \mathcal{U}$ is the control variable, $\mathcal{U}$ is the set of $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted stochastic processes $u(\cdot)$ with values in a given closed nonempty subset U of $\mathbb{R}^m$ such that $\mathbb{E} \int_0^T |u(t, \cdot)|^2 dt < \infty$, $x(\cdot)$ solves (1.1), $b : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^n$,

*CNRS, IMJ-PRG, UMR 7586, Sorbonne Universités, UPMC Univ Paris 06, 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 the grant PGMO 2015-2832H and CNRS-NSFC PRC Project under the grant 271392. *E-mail:* helene.frankowska@imj-prg.fr.

†School of Mathematics and Statistics, Southwest University, Chongqing 400715, China and CNRS, IMJ-PRG, UMR 7586, UPMC Univ Paris 06, case 247, 4 place Jussieu, 75252 Paris, France. The research of this author is partially supported by NSF of China under grants 11401404 and 11471231, the State Scholarship Fund of China Scholarship Council under grant [2016]3035 and the fundamental research funds for the central universities under grant XDJK2015C142. *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.

$\sigma = (\sigma^1, \cdots, \sigma^d) : [0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\Omega \to \mathbb{R}^{n\times d}$, $\phi : \mathbb{R}^n\times\Omega \to \mathbb{R}$, and $g^i : \mathbb{R}^n\times\Omega \to \mathbb{R}$, $i = 1, \cdots, k$ are given functions (satisfying suitable conditions to be stated later), and K_0 is a nonempty closed set in $\mathbb{R}^n$. As usual, when the context is clear, we omit explicit writing of the variable ω ($\in \Omega$).

A state-control pair $(x(\cdot), u(\cdot))$ is called admissible if $u(\cdot) \in \mathcal{U}$ and $x(\cdot)$ is the corresponding solution of (1.1) satisfying the end points constraints (1.3). In this case, we call $u(\cdot)$ an admissible control. Denote by $\mathscr{P}_{ad}$ the set of all admissible pairs. The optimal control problem considered in this paper is to find a $(\bar{x}(\cdot), \bar{u}(\cdot)) \in \mathscr{P}_{ad}$ such that

$$J(\bar{x}(\cdot), \bar{u}(\cdot)) = \inf_{(x(\cdot),u(\cdot))\in\mathscr{P}_{ad}} J(x(\cdot), u(\cdot)). \tag{1.4}$$

Any $(\bar{x}(\cdot), \bar{u}(\cdot)) \in \mathscr{P}_{ad}$ satisfying (1.4) is called an optimal pair, $\bar{x}(\cdot)$ is called an optimal state and $\bar{u}(\cdot)$ is called an optimal control. It is well known that the Bolza type optimal control problems (involving also an integral cost) can be reduced to the Mayer problem by adding an extra variable. For this reason we investigate here the Mayer problem and state the corresponding results for the Bolza problem as well.

Stochastic optimal control problems with end points constraints have many applications. Here we give a simple example from mathematical finance. Let us consider a mean-variance portfolio selection problem as follows: Suppose there are $m + 1$ assets whose price processes $S_i(\cdot)$, $i = 0, 1, \dots, m$ are described by the differential equations:

$$\begin{cases} dS_0(t) = rS_0(t)dt, \quad t \in [0,T], \\ S_0(0) = s_0, \end{cases} \tag{1.5}$$

and, for $i = 1, \dots, m$,

$$\begin{cases} dS_i(t) = b_iS_i(t)dt + \sum_{j=1}^{d} \sigma^{ij} S_i(t)dW^j(t), \quad t \in [0,T], \\ S_i(0) = s_i, \end{cases} \tag{1.6}$$

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

$$\begin{cases} dx(t) = \big[rx(t) + \sum_{i=1}^{m}(b_i - r)u_i(t)\big]dt + \sum_{i=1}^{m}\sum_{j=1}^{d} \sigma^{ij}u_i(t)dW^j(t), \quad t \in [0,T], \\ x(0) = x_0, \end{cases} \tag{1.7}$$

where x_0 is the initial wealth.

We call $u(\cdot) = (u_1(\cdot), \dots, u_m(\cdot))^\top$ a portfolio of the investor. Let $U = \mathbb{R}^m$. The objective of the investor is to maximize the mean terminal wealth, $\mathbb{E}\, x(T)$, and at the same time to minimize the variance $\mathrm{Var}\, x(T) = \mathbb{E}\, x(T)^2 - (\mathbb{E}\, x(T))^2$ of the terminal wealth. This can be formulated as the following multi-objective optimization problem:

$$\text{Minimize } (J_1(u(\cdot)), J_2(u(\cdot))) := (-\mathbb{E}\, x(T), \mathrm{Var}\, x(T)),$$

$$\text{subject to } \begin{cases} u(\cdot) \in \mathcal{U}, \\ (x(\cdot), u(\cdot)) \text{ satisfies the equation (1.7).} \end{cases}$$

Following Markowitz's portfolio management theory, we call a portfolio $u^*(\cdot)$ an efficient portfolio if there exists no portfolio $u(\cdot)$ such that

$$J_1(u(\cdot)) \leq J_1(u^*(\cdot)), \quad J_2(u(\cdot)) \leq J_2(u^*(\cdot))$$

and at least one of the two inequalities is strict. In this case, $(-J_1(u^*(\cdot)), J_2(u^*(\cdot)))$ is called an efficient point and the set of all efficient points is called the efficient frontier.

Obviously, every rational investor will choose a portfolio belonging to the efficient frontier. However, different investors may select different portfolios on the efficient frontier, depending on their individual risk preferences. If an investor wishes to find a minimal risk portfolio with the mean terminal wealth no less than a given constant α, the corresponding optimal portfolio can be find by solving the following optimal control problem:

$$\text{Minimize Var } x(T)$$

subject to

$$\begin{cases} u(\cdot) \in \mathcal{U}, \\ \mathbb{E}\ x(T) \geq \alpha, \\ (x(\cdot), u(\cdot)) \text{ satisfies the equation (1.7).} \end{cases} \tag{1.8}$$

Though Var $x(T)$ is not the cost functional treated in this paper, it can be shown, similarly to [23, Theorem 8.2, p. 338], that there exists a constant λ such that any solution of the problem (1.8) is also a solution to the optimal control problem

$$\text{Minimize } \mathbb{E}\ x(T)^2 - \lambda \mathbb{E}\ x(T)$$

subject to (1.8). Clearly, this new problem is a special case of the optimal control problem (1.4).

Stochastic optimal control problems with end points constraints have been studied from the very beginning of the foundation of the stochastic control theory, see [11, 14] and the references cited therein. However, the early works on this subject considered only the cases when the controls are absent from the diffusion terms of the control systems. A stochastic maximum principle (which is a first order necessary condition for optimal controls) when both the drift and the diffusion are control-dependent and the control region may be nonconvex was proven in [18], by introducing two adjoint processes and with the arguments being based on the Ekeland variational principle. Similarly to its counterpart in the deterministic optimal control, in addition to the first order necessary condition, some second order necessary conditions should be established to distinguish better the optimal controls from other admissible controls. The integral type second order necessary conditions for optimal controls in the presence of state constraints have been extensively studied in the deterministic control problems, see the monograph [16] and the references cited therein for the problems with end points constrains and [12, 15, 17, 20] and so on for more general pointwise state constraints. However, to the best of our knowledge, there are only two papers considering the second order necessary conditions for stochastic optimal control problems with final state constraints. In [4], the authors discussed the stochastic optimal controls with mixed final state constraints but without the control constraints. While, in [3], the author considered the stochastic optimal controls with linear inequality control constraints and mixed final state constraints but with the diffusion terms independent from the controls. Moreover, both [3] and [4] require the convexity of the control regions.

The main purpose of this work is to establish some first and second order necessary optimality conditions for the problem (1.4) with the control-dependent diffusion term and in the absence of convexity of U. Some techniques from the classical variational analysis are introduced to treat the nonconvex control constraints. The main novelties of this paper are twofold. Firstly, the second order tangent vectors to the control set are used to formulate the second order necessary conditions. Conditions of this type are more efficient than those without the second order tangent vectors even when the control regions are convex, see examples in [10]. Secondly, instead of using the mathematical programming theory in infinite dimensional vector spaces (as done in [3, 4]), we derive the desired results from the separation theorem, which makes the proofs much simpler and direct.

The rest of the 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 conditions for stochastic optimal controls, and finally in Section 4 we establish the second order necessary conditions.

2. Preliminaries. Denote by $\mathbb{R}_+$ the set of all nonnegative numbers and by $\mathbb{R}^{n\times m}$ the space of all $n\times m$-real matrices. For any $A\in\mathbb{R}^{n\times m}$, denote by $A^\top$ its transpose and by $|A|=\sqrt{tr(AA^\top)}$ the norm of A. We use the notation $\mathbf{S}^n := \big\{A\in\mathbb{R}^{n\times n}\big|\ A^\top=A\big\}$.

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

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

Let us recall that on a given filtered probability space, any $\mathbb{F}$-progressively measurable process is $\mathcal{B}([0;T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted, and every $\mathcal{B}([0;T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted process has an $\mathbb{F}$-progressively measurable modification (see [23, Proposition 2.8, p. 17]).

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

Let X be a Banach space with a norm $\|\cdot\|_X$ and the dual X^*. Denote by B_X the closed unit ball in 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 $dist\,(x,K) := \inf_{y\in K} \|y-x\|_X$.

DEFINITION 2.1. *Let $K \subset X$. 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 $T_K^b(x)$ to K at x is defined by

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

It is well known that $\mathcal{C}_K(x)$ is a closed convex cone in X and $\mathcal{C}_K(x) \subset T_K^b(x)$. When K is convex, $\mathcal{C}_K(x) = T_K^b(x) = cl\Big\{ \alpha(y-x) \;\Big|\; \alpha \geq 0,\ y \in K \Big\}$.

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

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

By [9, Lemma 2.4], if $T_K^{b(2)}(x,v) \neq \emptyset$, then

$$\mathcal{C}_K(x) + T_K^{b(2)}(x,v) = T_K^{b(2)}(x,v).$$

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

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

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

DEFINITION 2.3. *Let X, Y be Banach spaces, $F : X \to Y$ be a given map and $x \in X$. The first order contingent variation of F at x is defined by*

$$F^{(1)}(x) := \Big\{ v \in Y \;\Big|\; \liminf_{\varepsilon \to 0^+} dist\Big(v, \frac{F(x+\varepsilon B_X) - F(x)}{\varepsilon}\Big) = 0 \Big\}.$$

When F is Fréchet differentiable at x, $F^{(1)}(x) = cl(F_x(x)(B_X))$.

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

LEMMA 2.4. *Let $K_1, \cdots, K_q$ (for some $q \in \mathbb{N}$) be convex cones in X and $\bigcap_{i=1}^q intK_i \neq \emptyset$. Then for any convex cone K_0 such that $K_0 \bigcap \big(\bigcap_{i=1}^q intK_i\big) \neq \emptyset$, we have $\big(\bigcap_{i=0}^q K_i\big)^- = \sum_{i=0}^q K_i^-$.*

Proof. By the induction argument, it is sufficient to prove this lemma for $q = 1$. To show that $(K_0 \cap K_1)^- = K_0^- + K_1^-$, let $\bar{x} \in K_0 \cap intK_1$ and $\varepsilon > 0$ be such that $\bar{x} + \varepsilon B_X \subset intK_1$. Then, $\varepsilon B_X \subset K_0 - K_1$ and by [1, page 72], it follows that $(clK_0 \cap clK_1)^- = (clK_0)^- + (clK_1)^- = K_0^- + K_1^-$.

Obviously $(clK_0 \cap clK_1)^- \subset (K_0 \cap K_1)^-$. It remains to show that $(K_0 \cap K_1)^- \subset (clK_0 \cap clK_1)^-$. Let $\xi \in (K_0 \cap K_1)^-$ and $x \in clK_0 \cap clK_1$. Since $int\, clK_1 = intK_1$ and $\bar{x} + \varepsilon B_X \subset intK_1$, we deduce that $\lambda\bar{x} + (1-\lambda)x + \lambda\varepsilon B_X \subset intK_1$ for any

$\lambda \in (0,1)$. Noting that $\lambda\bar{x} + (1-\lambda)x \in clK_0$, there exists a $y_\lambda \in K_0$ such that $y_\lambda \in \lambda\bar{x} + (1-\lambda)x + \lambda\varepsilon B_X \subset K_1$. Then $y_\lambda \in K_0 \cap K_1$ and

$$\|y_\lambda - x\|_X \le \lambda\varepsilon + \lambda\|\bar{x} - x\|_X \to 0, \text{ as } \lambda \to 0^+.$$

Consequently, $\xi(x) = \lim_{\lambda\to 0^+} \xi(y_\lambda) \le 0$. Since $x \in clK_0 \cap clK_1$ is arbitrary, we get $\xi \in (clK_0 \cap clK_1)^-$. □

LEMMA 2.5. *Let X be a Hilbert space (with an inner product $\langle\cdot,\cdot\rangle_X$), K be a nonempty closed polyhedra in X, i.e., for some $k \in \mathbb{N}$, $\{a_1,\cdots,a_k\} \subset X \setminus \{0\}$ and $\{b_1,\cdots,b_k\} \subset \mathbb{R}$,*

$$K := \left\{ x \in X \mid \langle a_i, x\rangle_X + b_i \le 0, \ \forall\, i = 1,\cdots,k \right\}.$$

If $\xi \in H\setminus\{0\}$ 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

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

Proof. Define $X_1 = Span\{a_i \mid i = 1,\cdots,k\}$. Then, $X = X_1 \oplus X_1^\perp$, where $X_1^\perp$ is the orthogonal complement of X_1. Let

$$Q := \left\{ x \in X_1 \mid \langle a_i, x\rangle_X + b_i \le 0, \ \forall\, i = 1,\cdots,k \right\}.$$

It is clear that $K = Q + X_1^\perp$ and $K \neq \emptyset$ if and only if $Q \neq \emptyset$.

We claim that $\xi \in X_1$ whenever $\sup_{x\in K}\langle \xi, x\rangle_X < +\infty$. Indeed, consider $\xi_1 \in X_1$ and $\xi_2 \in X_1^\perp$ such that $\xi = \xi_1 + \xi_2$. Let $x_1 \in Q$. Then $x^\lambda := x_1 + \lambda\xi_2 \in K$ for any $\lambda \ge 0$. Consequently,

$$\left\langle \xi, x^\lambda\right\rangle_X = \langle \xi, x_1\rangle_X + \lambda\|\xi_2\|_X^2 \le \sup_{x\in K}\langle \xi, x\rangle_X < +\infty.$$

Since $\lambda \ge 0$ is arbitrary, the above inequality yields $\xi_2 = 0$. Therefore,

$$\sup_{x\in K}\langle \xi, x\rangle_X = \sup_{x\in Q}\langle \xi, x\rangle_X .$$

By [19, Corollary 3.53], there exists an $\bar{x} \in Q \subset K$ such that $\sup_{x\in Q}\langle \xi, x\rangle_X = \langle \xi, \bar{x}\rangle_X$.

Clearly, $\bar{x} \in \partial K$ and $\langle \xi, x - \bar{x}\rangle_X \le 0$ for all $x \in K$. By the definition of the normal cone, $\xi \in N_K(\bar{x})$. It remains to prove that $N_K(\bar{x}) = \sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i$. Since for any $i \in I(\bar{x})$ and $x \in K$, $\langle a_i, x\rangle_X + b_i \le 0$ and $\langle a_i, \bar{x}\rangle_X + b_i = 0$, we have

$$\langle a_i, x - \bar{x}\rangle_X \le 0, \quad \forall\, x \in K, \ \forall\, i \in I(\bar{x}).$$

This implies that $\sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i \subset N_K(\bar{x})$. If there exists a $\zeta \in N_K(\bar{x})\setminus \sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i$, then, by the separation theorem, we can find $0 \neq y \in X$ such that

$$\sup_{\xi\in\sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i} \langle \xi, y\rangle_X < \langle \zeta, y\rangle_X .$$

Since $\sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i$ is a closed convex cone, we deduce that $\sup_{\xi\in\sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i} \langle \xi, y\rangle_X = 0$ and $\langle a_i, y\rangle_X \le 0$ for any $i \in I(\bar{x})$. Then, $\tilde{x} := \bar{x} + \lambda y \in K$ for all $\lambda > 0$. Hence, $\langle \zeta, \tilde{x} - \bar{x}\rangle_X = \lambda\langle \zeta, y\rangle_X > 0$, contradicting to $\zeta \in N_K(\bar{x})$. Therefore, $N_K(\bar{x}) = \sum_{i\in I(\bar{x})} \mathbb{R}_+ a_i$. This completes the proof of Lemma 2.5. □

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. For any $\xi\in\Xi$, $F(\xi)$ is called the value or image of F at ξ. The domain of F is defined by $Dom\,(F):=\{\xi\in\Xi\mid F(\xi)\neq\emptyset\}$. The graph of F is the subset of the product space $\Xi\times Y$ defined by

$$Graph(F):=\big\{(\xi,y)\in\Xi\times Y\;\big|\;y\in F(\xi)\big\}.$$

The inverse F^{-1} of F is the set-valued map from Y to Ξ defined by

$$F^{-1}(y):=\big\{\xi\in\Xi\;\big|\;y\in F(\xi)\big\}.$$

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

LEMMA 2.6. ([2, Theorem 8.1.4]) *Let $(\Xi,\mathscr{G},\mu)$ be a complete σ-finite measure space, and F be a set-valued map from Ξ to Y with nonempty closed images. Then F is measurable if and only if $Graph(F)\in\mathscr{G}\otimes\mathcal{B}(Y)$.*

The following result is in the same spirit as that in [2, Theorem 8.5.1], for which we shall provide a short proof for readers' convenience.

LEMMA 2.7. *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 map $\mathcal{C}_K(\varphi(\cdot))$: $\xi\rightsquigarrow\mathcal{C}_K(\varphi(\xi))$ is $\mathscr{G}$-measurable, and

$$\mathcal{T}:=\big\{v(\cdot)\in L^p(\Xi,\mathscr{G},\mu;X)\;\big|\;v(\xi)\in\mathcal{C}_K(\varphi(\xi)),\ \mu\text{–}a.e.\ \xi\in\Xi\big\}\subset\mathcal{C}_{\mathcal{K}}(\varphi(\cdot)).$$

Proof. Define $\psi:K\times X\to\mathbb{R}\cup\{+\infty\}$ as

$$\psi(x,v):=\limsup_{\substack{\varepsilon\to0^+\\ y\in K,y\to x}}\frac{dist(y+\varepsilon v,K)}{\varepsilon}.$$

Then for any $x\in K$, $\mathcal{C}_K(x)=\{v\in X\mid\psi(x,v)=0\}$ and by [5, Proposition 2.1.1], ψ is upper semicontinuous. Since $\varphi(\cdot)$ is measurable,

$$Graph(\mathcal{C}_K(\varphi(\cdot))):=\big\{(\xi,v)\in\Xi\times X\mid\psi(\varphi(\xi),v)=0\big\}\in\mathscr{G}\otimes\mathcal{B}(X).$$

Then, by Lemma 2.6, the set-valued map $\xi\rightsquigarrow\mathcal{C}_K(\varphi(\xi))$ is measurable.

To prove that $\mathcal{T}\subset\mathcal{C}_{\mathcal{K}}(\varphi(\cdot))$, it suffices to show that, for every $v(\cdot)\in\mathcal{T}$, for any sequences $\{\varepsilon_n\}\subset\mathbb{R}_+$ and $\{\varphi_n(\cdot)\}\subset\mathcal{K}$ with $\varepsilon_n\to0^+$ and $\varphi_n(\cdot)\to\varphi(\cdot)$ in $L^p(\Xi,\mathscr{G},\mu;X)$ as $n\to\infty$, there exists a sequence $\{v_n(\cdot)\}\subset L^p(\Xi,\mathscr{G},\mu;X)$ such that $v_n(\cdot)\to v(\cdot)$ in $L^p(\Xi,\mathscr{G},\mu;X)$ as $n\to\infty$ and $\varphi_n(\cdot)+\varepsilon_nv_n(\cdot)\in\mathcal{K}$ for each $n\in\mathbb{N}$.

Fix $v,\varepsilon_n,\varphi_n$ as above and set

$$\alpha_n(\xi):=dist\Big(v(\xi),\frac{K-\varphi_n(\xi)}{\varepsilon_n}\Big).$$

By [2, Theorem 8.2.11], $\alpha_n(\cdot)$ is measurable. Since $\varphi_n(\cdot)\to\varphi(\cdot)$ in $L^p(\Xi,\mathscr{G},\mu;X)$, $\varphi_n(\cdot)$ converge to $\varphi(\cdot)$ in measure. Therefore, for any subsequence $\{\varphi_{n_j}(\cdot)\}_{j=1}^\infty$ of $\{\varphi_n(\cdot)\}_{n=1}^\infty$ there exists a sub-subsequence $\{\varphi_{n_{j_k}}(\cdot)\}_{k=1}^\infty$ such that $\varphi_{n_{j_k}}(\xi)\to\varphi(\xi)$

μ-a.e. as $k \to \infty$. By the definition of the Clarke tangent cone, $\alpha_{n_{j_k}}(\xi) \to 0$, μ-a.e. as $k \to \infty$. Noting that $\alpha_n(\xi) \le \|v(\xi)\|_X$, μ-a.e., by the Lebesgue dominated convergence theorem, $\alpha_n(\cdot)$ converge to 0 in $L^p(\Xi, \mathscr{G}, \mu; X)$. By [2, Corollary 8.2.13 and Theorem 8.2.9], there exists a $\varphi_{\varepsilon_n}(\cdot) \in \mathcal{K}$ such that for any $\xi \in \Xi$,

$$\Big\| v(\xi) - \frac{\varphi_{\varepsilon_n}(\xi) - \varphi_n(\xi)}{\varepsilon_n} \Big\|_X \le \alpha_n(\xi) + \varepsilon_n.$$

Define $v_n(\cdot) = \frac{\varphi_{\varepsilon_n}(\cdot) - \varphi_n(\cdot)}{\varepsilon_n}$. Then, $v_n(\cdot)$ converge to $v(\cdot)$ in $L^p(\Xi, \mathscr{G}, \mu; X)$ as $n \to \infty$, and $\varphi_{\varepsilon_n}(\cdot) = \varphi_n(\cdot) + \varepsilon_n v_n(\cdot) \in \mathcal{K}$ for any $n \in \mathbb{N}$. This completes the proof. □

As in [13], we call a measurable set-valued map $\zeta : (\Omega, \mathcal{F}) \rightsquigarrow \mathbb{R}^m$ a set-valued random variable, and, 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. We say that Γ is $\mathbb{F}$-adapted if $\Gamma(t)$ is $\mathcal{F}_t$-measurable for any $t \in [0,T]$. Define

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

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 [13, p. 96], the following result holds.

Lemma 2.8. *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.*

3. First order necessary condition. In this section, we study the first order necessary optimality conditions for the optimal control problem (1.4). Firstly, we introduce the notion of local minimizer for (1.4).

Definition 3.1. *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 (1.4) if there exists a $\delta > 0$ such that $J(x(\cdot), u(\cdot)) \ge J(\bar{x}(\cdot), \bar{u}(\cdot))$ for any $(x(\cdot), u(\cdot)) \in \mathscr{P}_{ad}$ satisfying $|\bar{x}_0 - x_0| < \delta$ and $\|u - \bar{u}\|_2 < \delta$, where $\bar{x}_0$ and x_0 are the initial conditions of $\bar{x}(\cdot)$ and $x(\cdot)$, respectively.*

We need the following assumptions:

(A1) *The control region U is nonempty and closed.*

(A2) *The functions b, σ^j $(j = 1, \cdots, d)$, ϕ and g^i $(i = 1, \cdots, k)$ 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 $\sigma^j(\cdot, x, u, \cdot) : [0,T] \times \Omega \to \mathbb{R}^n$ $(j = 1, \cdots, d)$ 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*

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

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

are uniformly continuous in $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. There exist a constant $L \ge 0$ and a nonnegative $\eta \in L^2_{\mathbb{F}}(\Omega; L^2(0,T; \mathbb{R}))$ such that for a.e. $(t,\omega) \in [0,T] \times \Omega$, any $(x,u) \in \mathbb{R}^n \times \mathbb{R}^m$ and $j = 1, \cdots, d$,

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

(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) *For* $i = 1,\cdots,k$*, and any* $x \in \mathbb{R}^n$*, the random variable* $g^i(x,\cdot)$ *is* $\mathcal{F}_T$*-measurable,* $g^i(\cdot,\omega):\ \mathbb{R}^n \to \mathbb{R}$ *is differentiable a.s., and for any* $x,\ \tilde{x} \in \mathbb{R}^n$,

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

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

Let $(\bar{x},\bar{u})$ be a local minimizer for the problem (1.4) with $\bar{x}_0$ being the initial datum of $\bar{x}$. For $\varphi = b, \sigma^j$, $j = 1,\cdots,d$, denote

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

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

$$\begin{cases} dy_1(t) = \big(b_x[t]y_1(t) + b_u[t]v(t)\big)dt \\ \qquad\qquad + \sum\limits_{j=1}^{d} \big(\sigma^j_x[t]y_1(t) + \sigma^j_u[t]v(t)\big)dW^j(t), \quad t \in [0,T], \\ y_1(0) = \nu_0. \end{cases} \tag{3.1}$$

It is easy to see that, under the assumption (A2), the equation (3.1) admits a unique solution $y_1(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$.

Consider $\nu^\varepsilon_0 \in \mathbb{R}^n$ and $v_\varepsilon \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$ such that $\bar{x}_0 + \varepsilon\nu^\varepsilon_0 \in K_0$, $\bar{u} + \varepsilon v_\varepsilon \in \mathcal{U}$, $\nu^\varepsilon_0 \to \nu_0$ in $\mathbb{R}^n$ and $v_\varepsilon \to v$ in $L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$ as $\varepsilon \to 0^+$. For $u^\varepsilon := \bar{u} + \varepsilon v_\varepsilon$ and $x^\varepsilon_0 := \bar{x}_0 + \varepsilon\nu^\varepsilon_0$, let x^ε be the state of (1.1) corresponding to the control u^ε and the initial datum x^ε_0, and put $\delta x^\varepsilon = x^\varepsilon - \bar{x}$.

The following result, concerning the d-dimensional Wiener process, can be proved in the same way as that in [10, Lemma 3.2], originally derived for the one-dimensional Wiener process.

Lemma 3.2. *If (A2) (i) holds, then,*

$$\|y_1\|^2_{\infty,2} \le C\big(|\nu_0|^2 + \|v\|^2_2\big), \quad \|\delta x^\varepsilon\|^2_{\infty,2} = O(\varepsilon^2).$$

Furthermore,

$$\|r^\varepsilon_1\|_{\infty,2} \to 0, \quad as\ \varepsilon \to 0^+, \tag{3.2}$$

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

From now on and till the end of this section we restrict our attention to the Clarke tangent cones $\mathcal{C}_{K_0}(\bar{x}_0) \subset T^b_{K_0}(\bar{x}_0)$ and $\mathcal{C}_{\mathcal{U}}(\bar{u}) \subset T^b_{\mathcal{U}}(\bar{u})$.

Define the reachable set of the linearized control system (3.1) as follows:

$$\mathcal{R}^{(1)} := \big\{ y_1(T) \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n) \;\big|\; y_1 \text{ solves (3.1) with } (\nu_0, v) \in \mathcal{C}_{K_0}(\bar{x}_0) \times \mathcal{C}_{\mathcal{U}}(\bar{u}) \big\}.$$

Consider the set

$$\mathcal{Q}_{(1)} := \bigcap_{i \in I(\bar{x}(T))} \mathcal{Q}^i_{(1)},$$

where

$$I(\bar{x}(T)) := \big\{ i \in \{1, \cdots, k\} \;\big|\; \mathbb{E}\, g^i(\bar{x}(T)) = 0 \big\}$$

and

$$\mathcal{Q}^i_{(1)} := \big\{ z \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n) \;\big|\; \mathbb{E}\,\big\langle g^i_x(\bar{x}(T)), z \big\rangle < 0 \big\}.$$

Also, we define

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

When $\phi_x(\bar{x}(T)) = 0$ a.s., $\mathcal{L}^{(1)} = \emptyset$.

Since $\mathcal{C}_{K_0}(\bar{x}_0)$ and $\mathcal{C}_{\mathcal{U}}(\bar{u}(\cdot))$ are nonempty convex cones, $\mathcal{R}^{(1)}$ is a nonempty convex cone in $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$. Furthermore, $\mathcal{Q}_{(1)}$ and $\mathcal{L}^{(1)}$ are open convex cones in $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$.

Lemma 3.3. *If for some $A \in \mathcal{F}_T$ with $P(A) > 0$,*

$$Z(\omega) := \big\{ z \in \mathbb{R}^n \;\big|\; \big\langle g^i_x(\bar{x}(T,\omega),\omega), z \big\rangle < 0,\ \forall\, i \in I(\bar{x}(T)) \big\} \neq \emptyset, \text{ a.s. in } A, \tag{3.3}$$

then $\mathcal{Q}_{(1)} \neq \emptyset$.

Proof. Define

$$\Gamma(\omega) := \begin{cases} Z(\omega), & \text{if } \omega \in A \text{ and } Z(\omega) \neq \emptyset, \\ \mathbb{R}^n, & \text{otherwise.} \end{cases}$$

By the assumption (iii) in (A2) and (3.3), $Graph(\Gamma)$ is $\mathcal{F}_T \times \mathcal{B}(\mathbb{R}^n)$-measurable. Then, by [21, Theorem 5.8], there exists an $\mathcal{F}_T$-measurable selection $\gamma(\cdot)$ of $\Gamma(\cdot)$. Noting that $\gamma(\omega) \neq 0$ a.s. in A, we define

$$z(\omega) := \begin{cases} \dfrac{\gamma(\omega)}{|\gamma(\omega)|}, & \text{if } \omega \in A \text{ and } \gamma(\omega) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then, $z \in \mathcal{Q}_{(1)}$. ∎

Remark 3.1. *If for any $x \in \mathbb{R}^n$, the vectors $\{g^i_x(x,\omega)\}_{i \in I(x,\omega)}$ are positively independent a.s. (where $I(x,\omega) := \big\{ i \in \{1, ..., k\} \,\big|\, g^i(x,\omega) = 0 \big\}$), then the condition (3.3) holds true with $A = \Omega$.*

We associate with the first order variational equation (3.1), the following first order adjoint equation

$$\begin{cases} dP_1(t) = -\big(b_x[t]^\top P_1(t) + \sum\limits_{j=1}^{d} \sigma^j_x[t]^\top Q^j_1(t) \big) dt + \sum\limits_{j=1}^{d} Q^j_1(t) dW^j(t),\ t \in [0,T], \\ P_1(T) = \xi, \end{cases} \tag{3.4}$$

where ξ is a random variable in $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ which will be specified later. Under the assumption (A2), the backward equation (3.4) 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

$$H(t,x,u,p,q,\omega):=\langle p,b(t,x,u,\omega)\rangle+\sum_{j=1}^{d}\left\langle q^j,\sigma^j(t,x,u,\omega)\right\rangle, \tag{3.5}$$

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

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

$H_u[t]$, $H_{xx}[t]$, $H_{xu}[t]$ and $H_{uu}[t]$ are defined in a similar way.

We have the following result.

THEOREM 3.4. *Let (A1)–(A2) hold and $(\bar{x},\bar{u})$ be a local minimizer for the problem (1.4).*

(i) *If $I(\bar{x}(T))=\emptyset$ or if $I(\bar{x}(T))\neq\emptyset$ and $\mathcal{Q}_{(1)}\neq\emptyset$, then there exist $\lambda_0\in\{0,1\}$, $\lambda_i\geq 0$ for $i\in I(\bar{x}(T))$, and a solution (P_1,Q_1) to the first order adjoint equation (3.4) corresponding to $(\bar{x},\bar{u})$ such that $\lambda_0+\mathbb{E}\ |P_1(T)|\neq 0$,*

$$H_u[t]\in N_U^C(\bar{u}(t)),\quad a.e.\ t\in[0,T],\ a.s. \tag{3.6}$$

and

$$P_1(0)\in N^C_{K_0}(\bar{x}_0),\quad -P_1(T)=\lambda_0\phi_x(\bar{x}(T))+\sum_{i\in I(\bar{x}(T))}\lambda_i g^i_x(\bar{x}(T)). \tag{3.7}$$

(ii) *If $I(\bar{x}(T))\neq\emptyset$ but $\mathcal{Q}_{(1)}=\emptyset$, then, for each $i\in I(\bar{x}(T))$, there exists a $\lambda_i\geq 0$ such that $\sum_{i\in I(\bar{x}(T))}\lambda_i>0$ and $\sum_{i\in I(\bar{x}(T))}\lambda_i g^i_x(\bar{x}(T))=0$. In particular, (3.6)–(3.7) hold with $\lambda_0=0$ and $(P_1,Q_1)\equiv 0$.*

Furthermore, the above holds with $\lambda_0=1$ if $I(\bar{x}(T))=\emptyset$ or if $I(\bar{x}(T))\neq\emptyset$ and $\mathcal{Q}_{(1)}\cap\mathcal{R}^{(1)}\neq\emptyset$.

Proof. (i) If $I(\bar{x}(T))=\emptyset$, then $\mathbb{E}\ g^i(\bar{x}(T))<0$ for every $i=1,\cdots,k$, and it is easy to verify that any $(x(\cdot),u(\cdot))$ obtained by a sufficiently small perturbation of $\bar{x}_0$ and $\bar{u}(\cdot)$ is still in $\mathscr{P}_{ad}$. Recall that the Clarke tangent cone is a subset of the adjacent cone. Furthermore, results obtained in [10] for the one dimensional Wiener process can be extended to the d-dimensional Wiener process by similar proofs. Then, setting $\lambda_0=1$ and $P_1(T)=-\phi_x(\bar{x}(T))$, it follows from [10, Theorem 3.1] that $P_1(0)\in N^C_{K_0}(\bar{x}_0)$ and

$$\mathbb{E}\int_0^T\langle H_u[t],v(t)\rangle\,dt\leq 0,\quad\forall\ v(\cdot)\in\mathcal{C}_{\mathcal{U}}(\bar{u}(\cdot)). \tag{3.8}$$

By Lemma 2.7, using the same method in [22, Lemma 4.6] we conclude that

$$\langle H_u[t],v\rangle\leq 0,\quad\forall\ v\in\mathcal{C}_U(\bar{u}(t)),\ \text{a.e. } t\in[0,T],\ a.s.,$$

i.e., (3.6) holds true.

If $I(\bar{x}(T))\neq\emptyset$ and $\mathcal{Q}_{(1)}\neq\emptyset$, then

$$\mathcal{R}^{(1)}\cap\mathcal{Q}_{(1)}\cap\mathcal{L}^{(1)}=\emptyset.$$

Indeed, otherwise, there would exist a solution $\bar{y}_1$ of (3.1) such that $\bar{y}_1(T) \in \mathcal{Q}_{(1)} \cap \mathcal{L}^{(1)}$. Let $\bar{\nu} \in \mathcal{C}_{K_0}(\bar{x}_0)$ and $\bar{v}(\cdot) \in \mathcal{C}_{\mathcal{U}}(\bar{u}(\cdot))$ be the initial datum and the 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 $x_0^\varepsilon := \bar{x}_0 + \varepsilon\bar{\nu} + \mu(\varepsilon) \in K_0$ 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. By Lemma 3.2, one can find a $\rho < 0$ such that for any sufficiently small $\varepsilon > 0$,

$$\begin{aligned}\mathbb{E}\; g^i(x^\varepsilon(T)) &= \mathbb{E}\; g^i(\bar{x}(T)) + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), \bar{y}_1(T)\right\rangle + o(\varepsilon)\\ &= \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), \bar{y}_1(T)\right\rangle + o(\varepsilon) < \varepsilon\rho + o(\varepsilon) \le 0, \quad \forall\, i \in I(\bar{x}(T)),\end{aligned}$$

and, for any $i \notin I(\bar{x}(T))$,

$$\begin{aligned}\mathbb{E}\; g^i(x^\varepsilon(T)) &= \mathbb{E}\; g^i(\bar{x}(T)) + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), \bar{y}_1(T)\right\rangle + o(\varepsilon)\\ &< \rho + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), \bar{y}_1(T)\right\rangle + o(\varepsilon) \le 0.\end{aligned}$$

Moreover,

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

which contradicts to the local optimality of $\bar{u}$.

Now we consider two different subcases.

Case a: $\mathcal{Q}_{(1)} \cap \mathcal{R}^{(1)} = \emptyset$. Since $\mathcal{Q}_{(1)}$ is a nonempty open convex set and $\mathcal{R}^{(1)}$ is nonempty and convex, by the separation theorem there exists a nonzero $\xi \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ such that

$$\sup_{\alpha\in\mathcal{Q}_{(1)}} \mathbb{E}\left\langle \xi, \alpha\right\rangle \le \inf_{\beta\in\mathcal{R}^{(1)}} \mathbb{E}\left\langle \xi, \beta\right\rangle.$$

Since $\mathcal{Q}_{(1)}$ and $\mathcal{R}^{(1)}$ are cones, $0 = \sup_{\alpha\in\mathcal{Q}_{(1)}} \mathbb{E}\left\langle \xi, \alpha\right\rangle = \inf_{\beta\in\mathcal{R}^{(1)}} \mathbb{E}\left\langle \xi, \beta\right\rangle$. Therefore, $\xi \in \mathcal{Q}_{(1)}^-$ and $-\xi \in \big(\mathcal{R}^{(1)}\big)^-$. By Lemma 2.4, $\mathcal{Q}_{(1)}^- = \sum_{i\in I(\bar{x}(T))}\big(\mathcal{Q}_{(1)}^i\big)^-$, implying that, for each $i \in I(\bar{x}(T))$, there exists a $\lambda_i \ge 0$ such that $\xi = \sum_{i\in I(\bar{x}(T))} \lambda_i g_x^i(\bar{x}(T))$. Set $\lambda_0 = 0$ and $P_1(T) = -\xi$. Then, $\lambda_0 + \mathbb{E}|P_1(T)| \neq 0$, and, since $-\xi \in \big(\mathcal{R}^{(1)}\big)^-$,

$$\mathbb{E}\left\langle P_1(T), y_1(T)\right\rangle \le 0, \quad \forall\, y_1(T) \in \mathcal{R}^{(1)}. \tag{3.9}$$

By the duality between (3.1) and the (3.4),

$$\begin{aligned}&\mathbb{E}\left\langle P_1(T), y_1(T)\right\rangle \\ &= \left\langle P_1(0), \nu_0\right\rangle + \mathbb{E}\int_0^T \Big(\left\langle P_1(t), b_x[t]y_1(t)\right\rangle + \left\langle P_1(t), b_u[t]v(t)\right\rangle \\ &\quad - \left\langle b_x[t]^\top P_1(t), y_1(t)\right\rangle - \sum_{j=1}^d \left\langle \sigma_x^j[t]^\top Q_1^j(t), y_1(t)\right\rangle \\ &\quad + \sum_{j=1}^d \left\langle Q_1^j(t), \sigma_x^j[t]y_1(t)\right\rangle + \sum_{j=1}^d \left\langle Q_1^j(t), \sigma_u^j[t]v(t)\right\rangle \Big)dt \\ &= \left\langle P_1(0), \nu_0\right\rangle + \mathbb{E}\int_0^T \Big(\left\langle P_1(t), b_u[t]v(t)\right\rangle + \sum_{j=1}^d \left\langle Q_1^j(t), \sigma_u^j[t]v(t)\right\rangle \Big)dt.\end{aligned} \tag{3.10}$$

From (3.9)–(3.10), it follows that

$$\langle P_1(0), \nu_0\rangle + \mathbb{E}\int_0^T \langle H_u[t], v(t)\rangle\, dt \le 0, \quad \forall\, \nu_0 \in \mathcal{C}_{K_0}(\bar{x}_0),\ \forall\, v(\cdot) \in \mathcal{C}_{\mathcal{U}}(\bar{u}(\cdot)). \tag{3.11}$$

Letting $v = 0$, we have $P_1(0) \in N^C_{K_0}(\bar{x}_0)$. On the other hand, for $\nu_0 = 0$, we get (3.8) from (3.11). By the same arguments as before, we obtain (3.6).

Case b: $\mathcal{Q}_{(1)} \cap \mathcal{R}^{(1)} \neq \emptyset$. Since $\mathcal{R}^{(1)} \cap \mathcal{Q}_{(1)} \cap \mathcal{L}^{(1)} = \emptyset$, we have $\mathbb{E}\,\langle \phi_x(\bar{x}(T)), \kappa\rangle \ge 0$ for all $\kappa \in \mathcal{Q}_{(1)} \cap \mathcal{R}^{(1)}$. Therefore, $-\phi_x(\bar{x}(T)) \in \big(\mathcal{Q}_{(1)} \cap \mathcal{R}^{(1)}\big)^-$. Noting that $\mathcal{Q}_{(1)}$ is an open convex cone, by Lemma 2.4, $\big(\mathcal{Q}_{(1)} \cap \mathcal{R}^{(1)}\big)^- = \sum_{i\in I(\bar{x}(T))} \big(\mathcal{Q}^i_{(1)}\big)^- + \big(\mathcal{R}^{(1)}\big)^-$. Consequently, there exist $\xi \in \big(\mathcal{R}^{(1)}\big)^- \subset L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ and $\lambda_i \ge 0$ for each $i \in I(\bar{x}(T))$ such that $-\phi_x(\bar{x}(T)) = \xi + \sum_{i\in I(\bar{x}(T))} \lambda_i g^i_x(\bar{x}(T))$. Set $\lambda_0 = 1$ and define

$$P_1(T) = \xi = -\phi_x(\bar{x}(T)) - \sum_{i\in I(\bar{x}(T))} \lambda_i g^i_x(\bar{x}(T)). \tag{3.12}$$

Then, $\lambda_0 + \mathbb{E}\,|P_1(T)| \neq 0$. Since $\xi \in \big(\mathcal{R}^{(1)}\big)^-$, it follows that $\mathbb{E}\,\langle P_1(T), y_1(T)\rangle \le 0$ for every $y_1(T) \in \mathcal{R}^{(1)}$. Then, using arguments similar to the **Case a**, we obtain the first order necessary condition (3.6) and the transversality condition $P_1(0) \in N^C_{K_0}(\bar{x}_0)$.

(ii) $I(\bar{x}(T)) \neq \emptyset$ but $\mathcal{Q}_{(1)} = \emptyset$.

We prove the desired results in two different cases.

Case 1: There exists an $i \in I(\bar{x}(T))$ such that $g^i_x(\bar{x}(T)) = 0$ a.s. Set $\lambda_i = 1$ and $\lambda_j = 0$ for any $j \neq i$. Then

$$0 = \lambda_i g^i_x(\bar{x}(T)) + \sum_{j\in I(\bar{x}(T))\setminus\{i\}} \lambda_j g^j_x(\bar{x}(T)).$$

Case 2: For any $i \in I(\bar{x}(T))$, $g^i_x(\bar{x}(T)) \neq 0$ (in the space $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$). Then, $\mathcal{Q}^i_{(1)} \neq \emptyset$, for every $i \in I(\bar{x}(T))$. Since $\mathcal{Q}_{(1)} = \emptyset$, one can find an $i \in I(\bar{x}(T))$ and a subset $J \subset I(\bar{x}(T))\setminus\{i\}$ such that $\cap_{j\in J}\mathcal{Q}^j_{(1)}$ is nonempty and

$$\mathcal{Q}^i_{(1)} \bigcap \Big(\bigcap_{j\in J} \mathcal{Q}^j_{(1)}\Big) = \emptyset.$$

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

$$\sup_{\alpha\in\mathcal{Q}^i_{(1)}} \mathbb{E}\,\langle \xi, \alpha\rangle \le \inf_{\beta\in\bigcap_{j\in J}\mathcal{Q}^j_{(1)}} \mathbb{E}\,\langle \xi, \beta\rangle\,.$$

Since the involved sets are cones, $\xi \in \big(\mathcal{Q}^i_{(1)}\big)^-$ and $-\xi \in \big(\bigcap_{j\in J}\mathcal{Q}^j_{(1)}\big)^-$. By Lemma 2.4, $\xi = \lambda_i g^i_x(\bar{x}(T))$ for some $\lambda_i > 0$ and for each $j \in J$, there exists a $\lambda_j \ge 0$ such that $-\xi = \sum_{j\in J} \lambda_j g^j_x(\bar{x}(T))$. Setting $\lambda_j = 0$ for $j \in I(\bar{x}(T))\setminus(J \cup \{i\})$, we arrive at the desired result. The proof of Theorem 3.4 is complete. □

Remark 3.2. *When the diffusion in the control system depends on the control variable, to establish the stochastic maximum principle using the needle variations, more smoothness of data has to be assumed and two adjoint processes have to be introduced, see [18, 23] . When $K_0 = \{x_0\}$, the first order necessary condition (3.6)* *can be deduced from the maximum principle [23, Theorem 6.1]. The advantage of our approach is due to the fact that it leads to the first order necessary condition (3.6) using only the first order adjoint process* (P_1, Q_1)*, even if the control region is not convex. In addition, the proof of [23, Theorem 6.1] does not allow to get a geometric meaning of the finial datum* $P_1(T)$ *of the first order adjoint process, while, from the proof of Theorem 3.4 we know that,* $P_1(T)$ *is a normal vector at zero to the reachable set* $\mathcal{R}^{(1)}$ *(of the first order linearized system (3.1)).*

To end this section, we give two simple examples to show how to use the first order necessary condition to distinguish the local minimizer from other admissible controls.

Example 3.1. *Let* $T = 1$*,* $m = 2$*,* $n = 1$*,* $k = 1$*,* $K_0 = (-\infty, -1] \cup [1, +\infty)$ *and*

$$U = \{(u_1, u_2) \in \mathbb{R}^2 \,\big|\, |u_1| \geq 1, |u_2| \geq 1\}.$$

Consider the control system

$$\left\{ \begin{array}{l} dx(t) = u_1(t)dt + u_2(t)dW(t), \quad t \in [0,1], \\ x(0) = x_0 \end{array} \right. \tag{3.13}$$

with the cost functional

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

and the end points constraints

$$x_0 \in K_0, \quad \mathbb{E}\ g(x(1)) = \mathbb{E}\ x(1) \leq 0.$$

Let $\bar{x}_0 = 1$ *and* $\bar{u}(t) = (\bar{u}_1(t), \bar{u}_2(t)) \equiv (-1, 1)$*. Then, the corresponding solution of (3.13) is* $\bar{x}(t) = 1 - t + W(t)$*. In particular,* $\mathbb{E}\ \bar{x}(1) = \mathbb{E}\ W(1) = 0$*. In this case,*

$$\mathcal{C}_{K_0}(\bar{x}_0) = [0, +\infty), \quad \mathcal{C}_U((-1,1)) = \{(v_1, v_2) \in \mathbb{R}^2 \,\big|\, v_1 \leq 0, v_2 \geq 0\}$$

and

$$\mathcal{Q}_{(1)} = \{z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}) \,\big|\, \mathbb{E}\ z < 0\}.$$

Letting $\nu_0 = 0$*,* $v_1(t) < 0$*, a.e.* $t \in [0,1]$*, a.s.,* $v_2(t) \equiv 0$ *and* $y_1(\cdot)$ *be the solution to the first order linearized control system*

$$\left\{ \begin{array}{l} dy_1(t) = v_1(t)dt + v_2(t)dW(t), \quad t \in [0,1], \\ y_1(0) = \nu_0. \end{array} \right. \tag{3.14}$$

we have $y_1(1) \in \mathcal{R}^{(1)} \cap \mathcal{Q}_{(1)}$*.*

Define the Hamiltonian for this optimal control problem:

$$H(t, x, (u_1, u_2), p, q) := pu_1 + qu_2, \ (t, x, (u_1, u_2), p, q) \in [0,1] \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}. \tag{3.15}$$

Assume for a moment that $(\bar{x}, \bar{u})$ *is a local minimizer. By Theorem 3.4, there exists* $\lambda_1 \geq 0$ *such that the solution* (P_1, Q_1) *to the backward stochastic differential equation*

$$\left\{ \begin{array}{l} dP_1(t) = Q_1(t)dW(t), \quad t \in [0,1], \\ P_1(1) = -\phi_x(\bar{x}(1)) - \lambda_1 g_x(\bar{x}(1)) = -2W(1) - \lambda_1 \end{array} \right. \tag{3.16}$$

satisfies $P_1(0) \in N^C_{K_0}(\bar{x}_0)$ *and*

$$H_u[t] = (P_1(t), Q_1(t))^\top \in N^C_U(\bar{u}(t)),\ a.e.\ t \in [0,1],\ a.s. \tag{3.17}$$

By (3.16), $(P_1(t), Q_1(t)) = (-2W(t) - \lambda_1, -2)$*. Then, for any* $t \in [0,1]$*, choosing* $v = (v_1, v_2) \in \mathcal{C}_U(\bar{u}(t))$ *such that* $v_1 < 0$ *and* $v_2 = 0$*, we have*

$$P(\langle H_u[t], v\rangle = -2W(t)v_1 - \lambda_1 v_1 > 0) \neq 0,$$

contradicting to (3.17). Therefore, $\bar{u}(t) \equiv (-1,1)$ *is not a local minimizer.*

Example 3.2. *Consider the optimal control problem from Example 3.1 with* U *replaced by*

$$\widehat{U} = \{(u_1, u_2) \in \mathbb{R}^2 \,|\, |u_1| \geq 1\}.$$

Let $\hat{x}_0 = 1$ *and* $\hat{u}(t) \equiv (-1, 0)$*. Then,*

$$\mathcal{C}_{K_0}(\bar{x}_0) = [0, +\infty), \quad \mathcal{C}_{\widehat{U}}((-1,0)) = \{(v_1, v_2) \in \mathbb{R}^2 \,|\, v_1 \leq 0\}$$

and the corresponding solution of (3.13) is $\hat{x}(t) = 1 - t$*. Furthermore,* $\hat{x}(1) = 0$ *and therefore,* $\hat{u}(\cdot)$ *is a global minimizer.*

Next, let us verify that $\bar{u}(\cdot)$ *satisfies the conditions (3.6)–(3.7).*

Similarly to Example 3.1, choosing $\nu_0 = 0$*,* $v_1(t) < 0$*, a.e.* $t \in [0,1]$*, a.s.,* $v_2(t) \equiv 0$ *and letting* $y_1(\cdot)$ *be the solution to (3.14), we have* $y_1(1) \in \mathcal{R}^{(1)} \cap \mathcal{Q}_{(1)}$*. Let the Hamiltonian be defined by (3.15),* $\lambda_0 = 1$*,* $\lambda_1 = 0$ *and* (P_1, Q_1) *be the solution to the backward stochastic differential equation*

$$\begin{cases} dP_1(t) = Q_1(t)dW(t), & t \in [0,1], \\ P_1(1) = -\phi_x(\hat{x}(1)) = 0. \end{cases}$$

Clearly, $(P_1(t), Q_1(t)) \equiv (0,0)$ *and the conditions (3.6)–(3.7) trivially hold.*

4. Second order necessary conditions. We investigate next second order necessary conditions for a local minimizer $(\bar{x}, \bar{u})$ for the problem (1.4). Throughout this section, we assume that $\bar{u} \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ and define

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

In addition to the assumptions (A1) and (A2), we suppose that

(A3) *The functions* b*,* σ^j *(*$j = 1, \cdots, d$*),* ϕ *and* g^i *(*$i = 1, \cdots, k$*) satisfy the following:*

(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* $\sigma^j(t,\cdot,\cdot,\omega):$ $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ *(*$j = 1, \cdots, d$*) are twice differentiable and*

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

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

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

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

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

(iii) *For any* $i = 1, \cdots, k$, $g^i(\cdot,\omega): \ \mathbb{R}^n \to \mathbb{R}$ *are twice differentiable a.s., and for any* $x,\ \tilde{x} \in \mathbb{R}^n$,

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

For $\varphi = b,\ \sigma^j\ (j = 1, \cdots, d)$, write

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

Let $\nu_0 \in T^b_{K_0}(\bar{x}_0)$, $v \in T^b_{\mathcal{V}}(\bar{u})$, $\mathcal{W}(\bar{x}_0,\nu_0)$ be a convex subset of $T^{b(2)}_{K_0}(\bar{x}_0,\nu_0)$ and $\mathcal{M}(\bar{u},v)$ be a convex subset of $T^{b(2)}_{\mathcal{V}}(\bar{u},v)$ (Here, 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 $L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$). For any $\varpi_0 \in \mathcal{W}(\bar{x}_0,\nu_0)$ and $h \in \mathcal{M}(\bar{u},v)$, similarly to [12], we introduce the following second-order variational equation:

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

where y_1 is the solution of the linearized system (3.1).

Then 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 $\bar{x}_0 + \varepsilon\nu_0 + \varepsilon^2\varpi^\varepsilon_0 \in K_0$, $\bar{u} + \varepsilon v + \varepsilon^2 h_\varepsilon \in \mathcal{V}$, $\varpi^\varepsilon_0 \to \varpi_0$ in $\mathbb{R}^n$ and h_ε converges to h in $L^4_{\mathbb{F}}(\Omega; L^4(0,T; \mathbb{R}^m))$ as $\varepsilon \to 0^+$. Set $u^\varepsilon := \bar{u} + \varepsilon v + \varepsilon^2 h_\varepsilon$, $x^\varepsilon_0 := \bar{x}_0 + \varepsilon\nu_0 + \varepsilon^2\varpi^\varepsilon_0$ and denote by x^ε the solution of (1.1) corresponding to the initial datum x^ε_0 and the control u^ε. Put $\delta x^\varepsilon = x^\varepsilon - \bar{x}$. The next result for d-dimensional Wiener process follows by the same arguments as those used to prove [10, Lemma 4.1].

Lemma 4.1. *Let the assumptions (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* $\nu_0, \varpi_0,\ \varpi^\varepsilon_0 \in \mathbb{R}^n$ *as above, we have*

$$\|y_2\|^2_{\infty,2} \le C(|\varpi_0|^2 + |\nu_0|^4 + \|v\|^4_4 + \|h\|^2_2).$$

Furthermore,

$$(4.2)\qquad \|r^\varepsilon_2\|_{\infty,2} \to 0, \quad \varepsilon \to 0^+,$$

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

Denote

$$(4.3)\qquad \mathcal{Y}(\bar{x},\bar{u}) := \Big\{(y_1(\cdot), v(\cdot), \nu_0) \in L^4_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n)) \times T^b_{\mathcal{V}}(\bar{u}) \times T^b_{K_0}(\bar{x}_0) \ \Big| \ y_1 \text{ solves (3.1) and } \mathbb{E}\left\langle g^i_x(\bar{x}(T)), y_1(T)\right\rangle \le 0,\ \forall\, i \in I(\bar{x}(T))\Big\}$$

and define the critical cone by

$$(4.4)\qquad \Upsilon(\bar{x},\bar{u}) := \Big\{(y_1(\cdot), v(\cdot), \nu_0) \in \mathcal{Y}(\bar{x},\bar{u}) \ \Big|\ \mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_1(T)\right\rangle = 0\Big\}.$$

For any fixed $(y_1(\cdot), v(\cdot), \nu_0) \in \Upsilon(\bar{x},\bar{u})$, consider the reachable set of the affine control system (4.1), defined by

$$(4.5)\quad \mathcal{R}^{(2)}(y_1,v) := \Big\{y_2(T) \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n) \ \Big|\ y_2 \text{ is the solution of (4.1)}$$

$$\text{corresponding to some } (\varpi_0, h) \in \mathcal{W}(x_0, \nu_0) \times \mathcal{M}(\bar{u}, v)\Big\}.$$

Let

$$I(\bar{x}(T), y_1(T)) := \big\{ i \in I(\bar{x}(T)) \;\big|\; \mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_1(T)\right\rangle = 0 \big\}$$

and consider the set

$$\mathcal{Q}_{(2)}(y_1(T)) := \bigcap_{i \in I(\bar{x}(T), y_1(T))} \mathcal{Q}^i_{(2)}(y_1(T)), \tag{4.6}$$

where

$$\mathcal{Q}^i_{(2)}(y_1(T)) := \big\{ z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n) \;\big|\; \mathbb{E}\left\langle g_x^i(\bar{x}(T)), z\right\rangle + \frac{1}{2}\mathbb{E}\left\langle g^i_{xx}(\bar{x}(T)) y_1(T), y_1(T)\right\rangle < 0 \big\}.$$

Furthermore, we introduce the set

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

Remark 4.1. *For any $(y_1(\cdot), v(\cdot), \nu_0) \in \Upsilon(\bar{x}, \bar{u})$, if $I(\bar{x}(T), y_1(T)) \neq \emptyset$ and*

$$\big\{ g_x^i(\bar{x}(T)) \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n) \;\big|\; i \in I(\bar{x}(T), y_1(T)) \big\}$$

are positively independent, then $\mathcal{Q}_{(2)}(y_1(T)) \neq \emptyset$.

Indeed, under the positive independence assumption, it follows that

$$0 \notin co\big\{ g_x^i(\bar{x}(T)) \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n) \;\big|\; i \in I(\bar{x}(T), y_1(T)) \big\}.$$

Let $I(\bar{x}(T), y_1(T)) = \{i_1, ..., i_q\}$ for an integer $q \geq 1$. Define $A : L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n) \to \mathbb{R}^q$ by

$$Az = \big(\mathbb{E}\left\langle g_x^{i_1}(\bar{x}(T)), z\right\rangle, ..., \mathbb{E}\left\langle g_x^{i_q}(\bar{x}(T)), z\right\rangle\big)^\top, \quad \forall\, z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n),$$

and denote $\mathbb{R}^q_- = \{(x_1, ..., x_q) \in \mathbb{R}^q \mid x_i < 0,\ i = 1, ..., q\}$,

$$b = \frac{1}{2}\big(\mathbb{E}\left\langle g_{xx}^{i_1}(\bar{x}(T)) y_1(T), y_1(T)\right\rangle, ..., \mathbb{E}\left\langle g_{xx}^{i_q}(\bar{x}(T)) y_1(T), y_1(T)\right\rangle\big)^\top.$$

Assume otherwise that $\mathcal{Q}_{(2)}(y_1(T)) = \emptyset$. Then $\{Az + b \mid z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)\} \cap \mathbb{R}^q_- = \emptyset$. By the separation theorem, there would exist a $\xi = (\xi_1, ..., \xi_q) \neq 0$ such that

$$\sup_{\alpha \in \mathbb{R}^d_-} \langle \alpha, \xi\rangle \leq \inf_{z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)} \langle Az + b, \xi\rangle\,.$$

Hence $\xi_l \geq 0$, $l = 1, ..., q$ and $0 = \sup_{\alpha \in \mathbb{R}^d_-} \langle \alpha, \xi\rangle \leq \inf_{z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)} \langle Az + b, \xi\rangle$. Then, $0 \leq \langle Az + b, \xi\rangle = \mathbb{E}\left\langle \sum_{l=1}^q \xi_l g_x^{i_l}(\bar{x}(T)), z\right\rangle + \langle b, \xi\rangle$ for all $z \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$, implying that $\sum_{l=1}^q \xi_l g_x^{i_l}(\bar{x}(T)) = 0$ and leading to a contradiction.

For $\zeta \in L^2_{\mathcal{F}_T}(\Omega; \mathbf{S}^n)$ (which will be specified later), we introduce the following adjoint equation to (4.1):

$$\begin{cases} dP_2(t) = -\big(b_x[t]^\top P_2(t) + P_2(t) b_x[t] + \sum\limits_{j=1}^d \sigma_x^j[t]^\top P_2(t) \sigma_x^j[t] \\ \qquad\qquad + \sum\limits_{j=1}^d \sigma_x^j[t]^\top Q_2^j(t) + \sum\limits_{j=1}^d Q_2^j(t) \sigma_x^j[t] + H_{xx}[t]\big) dt \\ \qquad\qquad + \sum\limits_{j=1}^d Q_2^j(t) dW^j(t), \quad t \in [0, T], \\ P_2(T) = \zeta. \end{cases} \tag{4.7}$$

Clearly, under the assumptions (A2) (i) and (A3) (i), the equation (4.7) admits a unique strong solution $(P_2(\cdot),Q_2(\cdot))\in L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbf{S}^n))\times\big(L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbf{S}^n))\big)^d$.

To simplify the notation, we define

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

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

$$\mathbb{S}[t]=\mathbb{S}(t,\bar{x}(t),\bar{u}(t),P_1(t),Q_1(t),P_2(t),Q_2(t)),\quad t\in[0,T],\tag{4.8}$$

where $(P_1(\cdot),Q_1(\cdot))$ and $(P_2(\cdot),Q_2(\cdot))$ solve the equations (3.4) and (4.7), respectively.

Theorem 4.2. *Let (A1)–(A3) hold and $(\bar{x},\bar{u})$ be a local minimizer for the problem (1.4) with the initial datum $\bar{x}_0$ and the control $\bar{u}\in\mathcal{U}\cap L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$. Then for any $(y_1(\cdot),v(\cdot),\nu_0)\in\Upsilon(\bar{x},\bar{u})$ with $\mathcal{W}(\bar{x}_0,\nu_0)\neq\emptyset$ and $\mathcal{M}(\bar{x}_0,\nu_0)\neq\emptyset$, there exist $\lambda_0\in\{0,1\}$, $\lambda_i\geq 0$ (for each $i\in I(\bar{x}(T),y_1(T))$) not vanishing simultaneously, and adjoint processes (P_1,Q_1), (P_2,Q_2) corresponding to $(\bar{x},\bar{u})$ and the final datum*

$$P_1(T)=-\lambda_0\phi_x(\bar{x}(T))-\sum_{i\in I(\bar{x}(T),y_1(T))}\lambda_i g_x^i(\bar{x}(T)),\tag{4.9}$$

$$P_2(T)=-\lambda_0\phi_{xx}(\bar{x}(T))-\sum_{i\in I(\bar{x}(T),y_1(T))}\lambda_i g_{xx}^i(\bar{x}(T))\tag{4.10}$$

such that for any $\varpi_0\in\mathcal{W}(\bar{x}_0,\nu_0)$ and $h(\cdot)\in\mathcal{M}(\bar{u},v)$,

$$\begin{aligned}&\langle P_1(0),\varpi_0\rangle+\frac{1}{2}\langle P_2(0)\nu_0,\nu_0\rangle+\mathbb{E}\int_0^T\Big(\langle H_u[t],h(t)\rangle+\frac{1}{2}\langle H_{uu}[t]v(t),v(t)\rangle\\&+\frac{1}{2}\sum_{j=1}^d\big\langle P_2(t)\sigma_u^j[t]v(t),\sigma_u^j[t]v(t)\big\rangle+\langle\mathbb{S}[t]y_1(t),v(t)\rangle\Big)dt\leq 0.\end{aligned}\tag{4.11}$$

Moreover, the first order adjoint process (P_1,Q_1) with the finial datum defined by (4.9) also satisfies the first order necessary condition (3.6) and the transversality condition $P_1(0)\in N^C_{K_0}(\bar{x}_0)$.

Furthermore, $\lambda_0=1$ if $I(\bar{x}(T),y_1(T))=\emptyset$.

Proof. Fix $(y_1(\cdot),v(\cdot),\nu_0)\in\Upsilon(\bar{x},\bar{u})$. We first prove the desired second order necessary condition (4.11) by considering several cases.

Case 1: $I(\bar{x}(T),y_1(T))=\emptyset$. In this case, $\mathbb{E}\langle\phi_x(\bar{x}(T)),y_1(T)\rangle=0$ and for any $i\in I(\bar{x}(T))$, $\mathbb{E}\big\langle g_x^i(\bar{x}(T)),y_1(T)\big\rangle<0$. Then, for any $\varpi_0\in\mathcal{W}(\bar{x}_0,\nu_0)$, $h\in\mathcal{M}(\bar{u},v)$ and $\varepsilon>0$, there exist $\mu(\varepsilon)\in\mathbb{R}^n$ with $|\mu(\varepsilon)|=o(\varepsilon^2)$ and $\eta(\varepsilon)\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$ with $\|\eta(\varepsilon)\|_4=o(\varepsilon^2)$ such that $x_0^\varepsilon=\bar{x}_0+\varepsilon\nu_0+\varepsilon^2\varpi_0+\mu(\varepsilon)\in K_0$ and $u^\varepsilon=\bar{u}+\varepsilon v+\varepsilon^2 h+\eta(\varepsilon)\in\mathcal{V}$. Let x^ε be the solution to the control system (1.1) corresponding to the initial datum x_0^ε and the control u^ε. Then, by Lemma 4.1, for some $\rho<0$ and all sufficiently small $\varepsilon>0$ and for any $i\in I(\bar{x}(T))$,

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

$$
\begin{aligned}
&= \varepsilon\big(\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_1(T)\right\rangle + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_2(T)\right\rangle \\
&\quad + \frac{\varepsilon}{2}\mathbb{E}\left\langle g_{xx}^i(\bar{x}(T))y_1(T), y_1(T)\right\rangle + o(\varepsilon)\big) < \varepsilon\rho \le 0,
\end{aligned}
\tag{4.12}
$$

while for any $i \notin I(\bar{x}(T))$,

$$
\begin{aligned}
\mathbb{E}\ g^i(x^\varepsilon(T)) &= \mathbb{E}\ g^i(\bar{x}(T)) + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_1(T)\right\rangle + \varepsilon^2\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_2(T)\right\rangle \\
&\quad + \frac{\varepsilon^2}{2}\mathbb{E}\left\langle g_{xx}^i(\bar{x}(T))y_1(T), y_1(T)\right\rangle + o(\varepsilon^2) \\
&< \rho + \varepsilon\big(\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_1(T)\right\rangle + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_2(T)\right\rangle \\
&\quad + \frac{\varepsilon}{2}\mathbb{E}\left\langle g_{xx}^i(\bar{x}(T))y_1(T), y_1(T)\right\rangle + o(\varepsilon)\big) \le 0,
\end{aligned}
\tag{4.13}
$$

i.e., $(x^\varepsilon, u^\varepsilon) \in \mathscr{P}_{ad}$. Since $(\bar{x}, \bar{u})$ is locally optimal, for all sufficiently small $\varepsilon > 0$,

$$
\begin{aligned}
0 \le \frac{\mathbb{E}\ \phi(x^\varepsilon(T)) - \mathbb{E}\ \phi(\bar{x}(T))}{\varepsilon^2} &= \frac{1}{\varepsilon}\mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_1(T)\right\rangle + \mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_2(T)\right\rangle \\
+\frac{1}{2}\mathbb{E}\left\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\right\rangle &+ \frac{o(\varepsilon^2)}{\varepsilon^2} \\
\to \mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_2(T)\right\rangle &+ \frac{1}{2}\mathbb{E}\left\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\right\rangle, \quad (\text{as } \varepsilon \to 0^+).
\end{aligned}
\tag{4.14}
$$

On the other hand, for any solutions (P_1, Q_1) and (P_2, Q_2) of (3.4) and (4.7) respectively, by Itô's formula, we have

$$
\begin{aligned}
&\mathbb{E}\ \left\langle P_1(T), y_2(T)\right\rangle \\
&= \left\langle P_1(0), \varpi_0\right\rangle + \mathbb{E}\int_0^T \Big(\left\langle P_1(t), b_u[t]h(t)\right\rangle + \frac{1}{2}\left\langle P_1(t), y_1(t)^\top b_{xx}[t]y_1(t)\right\rangle \\
&+ \left\langle P_1(t), v(t)^\top b_{xu}[t]y_1(t)\right\rangle + \frac{1}{2}\left\langle P_1(t), v(t)^\top b_{uu}[t]v(t)\right\rangle \\
&+ \sum_{j=1}^d \left\langle Q_1^j(t), \sigma_u^j[t]h(t)\right\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle Q_1^j(t), y_1(t)^\top \sigma_{xx}^j[t]y_1(t)\right\rangle \\
&+ \sum_{j=1}^d \left\langle Q_1^j(t), v(t)^\top \sigma_{xu}^j[t]y_1(t)\right\rangle + \frac{1}{2}\sum_{j=1}^d \left\langle Q_1^j(t), v(t)^\top \sigma_{uu}^j[t]v(t)\right\rangle\Big)dt,
\end{aligned}
\tag{4.15}
$$

and

$$
\begin{aligned}
\mathbb{E}\ \left\langle P_2(T)y_1(T), y_1(T)\right\rangle &= \left\langle P_2(0)\nu_0, \nu_0\right\rangle + \mathbb{E}\int_0^T \Big(2\left\langle P_2(t)y_1(t), b_u[t]v(t)\right\rangle \\
&+2\sum_{j=1}^d \left\langle P_2(t)\sigma_x^j[t]y_1(t), \sigma_u^j[t]v(t)\right\rangle + \sum_{j=1}^d \left\langle P_2(t)\sigma_u^j[t]v(t), \sigma_u^j[t]v(t)\right\rangle \\
&+2\sum_{j=1}^d \left\langle Q_2^j(t)y_1(t), \sigma_u^j[t]v(t)\right\rangle - \left\langle H_{xx}[t]y_1(t), y_1(t)\right\rangle\Big)dt.
\end{aligned}
\tag{4.16}
$$

Letting $P_1(T) = -\phi_x(\bar{x}(T))$, $P_2(T) = -\phi_{xx}(\bar{x}(T))$ (i.e., $\lambda_0 = 1$, $\lambda_i = 0$ for each $i \in I(\bar{x}(T))$) and substituting (4.15)–(4.16) into (4.14), we obtain (4.11).

Case 2: $I(\bar{x}(T), y_1(T)) \neq \emptyset$. Firstly, we claim that

$$
\mathcal{L}^{(2)}(y_1(T)) \cap \mathcal{Q}_{(2)}(y_1(T)) \cap \mathcal{R}^{(2)}(y_1, v) = \emptyset.
$$

Indeed, if the above intersection was nonempty, then there would exist $\varpi_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$ and $h \in \mathcal{M}(\bar{u}, v)$ such that the corresponding solution $y_2(\cdot)$ of the second order variational equation (4.1) satisfies

$$\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_2(T)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle g_{xx}^i(\bar{x}(T))y_1(T), y_1(T)\right\rangle < 0, \quad \forall\, i \in I(\bar{x}(T), y_1(T))$$

and

$$\mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_2(T)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\right\rangle < 0.$$

Let $\mu(\varepsilon) \in \mathbb{R}^n$ with $|\mu(\varepsilon)| = o(\varepsilon^2)$ and $\eta(\varepsilon) \in L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$ with $\|\eta(\varepsilon)\|_4 = o(\varepsilon^2)$ satisfy $x_0^\varepsilon = \bar{x}_0 + \varepsilon\nu_0 + \varepsilon^2\varpi_0 + \mu(\varepsilon) \in K_0$ and $u^\varepsilon = \bar{u} + \varepsilon v + \varepsilon^2 h + \eta(\varepsilon) \in \mathcal{V}$, and let x^ε be the solution to the control system (1.1) corresponding to the initial datum x_0^ε and the control u^ε. Similarly to *Case 1*, one has for all $\varepsilon > 0$ small enough $\mathbb{E}\ g^i(x^\varepsilon(T)) \le 0$ for every $i \notin I(\bar{x}(T), y_1(T))$. Also, by Lemma 4.1, for any $i \in I(\bar{x}(T), y_1(T))$, there exists a $\rho < 0$ such that for all sufficiently small ε,

$$\begin{aligned}
\mathbb{E}\ g^i(x^\varepsilon(T)) =& \mathbb{E}\ g^i(\bar{x}(T)) + \varepsilon\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_1(T)\right\rangle + \varepsilon^2\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_2(T)\right\rangle \\
&+ \frac{\varepsilon^2}{2}\mathbb{E}\left\langle g_{xx}^i(\bar{x}(T))y_1(T), y_1(T)\right\rangle + o(\varepsilon^2) \\
=& \varepsilon^2\Big(\mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_2(T)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle g_{xx}^i(\bar{x}(T))y_1(T), y_1(T)\right\rangle + \frac{o(\varepsilon^2)}{\varepsilon^2}\Big) \\
<& \varepsilon^2\rho \le 0,
\end{aligned}$$

This proves that $(x^\varepsilon, u^\varepsilon) \in \mathscr{P}_{ad}$. On the other hand,

$$\begin{aligned}
\mathbb{E}\ \phi(x^\varepsilon(T)) =& \mathbb{E}\ \phi(\bar{x}(T)) + \varepsilon\mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_1(T)\right\rangle + \varepsilon^2\mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_2(T)\right\rangle \\
&+ \frac{\varepsilon^2}{2}\mathbb{E}\left\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\right\rangle + o(\varepsilon^2) \\
=& \mathbb{E}\ \phi(\bar{x}(T)) + \varepsilon^2\Big(\mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_2(T)\right\rangle \\
&+ \frac{1}{2}\mathbb{E}\left\langle \phi_{xx}(\bar{x}(T))y_1(T), y_1(T)\right\rangle + \frac{o(\varepsilon^2)}{\varepsilon^2}\Big) \\
<& \mathbb{E}\ \phi(\bar{x}(T)) + \varepsilon^2\rho < \mathbb{E}\ \phi(\bar{x}(T)),
\end{aligned}$$

contradicting the local optimality of $(\bar{x}, \bar{u})$ and proving our claim.

Next, we consider two subcases.

(i): $\mathcal{L}^{(2)}(y_1(T)) \bigcap \mathcal{Q}_{(2)}(y_1(T)) \neq \emptyset$.

Since $\mathcal{L}^{(2)}(y_1(T)) \cap \mathcal{Q}_{(2)}(y_1(T)) \cap \mathcal{R}^{(2)}(y_1, v) = \emptyset$, by the separation theorem, there exists a nonzero $\xi \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$ such that

$$\sup_{\alpha \in \mathcal{L}^{(2)}(y_1(T)) \bigcap \mathcal{Q}_{(2)}(y_1(T))} \mathbb{E}\left\langle \xi, \alpha\right\rangle \le \inf_{\beta \in \mathcal{R}^{(2)}(y_1, v)} \mathbb{E}\left\langle \xi, \beta\right\rangle.$$

By the first conclusion of Lemma 2.5, for some

$$\begin{aligned}
\alpha_0 \in& \, cl\Big(\mathcal{L}^{(2)}(y_1(T)) \bigcap \mathcal{Q}_{(2)}(y_1(T))\Big) \\
=& \, cl\Big(\mathcal{L}^{(2)}(y_1(T))\Big) \bigcap \Big(\bigcap_{i \in I(\bar{x}(T), y_1(T))} cl\Big(\mathcal{Q}^i_{(2)}(y_1(T))\Big)\Big)
\end{aligned}$$

we have $\mathbb{E}\langle\xi,\alpha_0\rangle=\sup_{\alpha\in\mathcal{L}^{(2)}(y_1(T))\bigcap\mathcal{Q}_{(2)}(y_1(T))}\mathbb{E}\langle\xi,\alpha\rangle$.

Denote by $I_0(\bar{x}(T),y_1(T))$ the set of all indices $i\in I(\bar{x}(T),y_1(T))$ such that

$$\mathbb{E}\left\langle g^i_x(\bar{x}(T)),\alpha_0\right\rangle+\frac{1}{2}\mathbb{E}\left\langle g^i_{xx}(\bar{x}(T))y_1(T),y_1(T)\right\rangle=0.$$

By the second conclusion of Lemma 2.5, there exist $\lambda_i\geq 0$ (for each $i\in I_0(\bar{x}(T),y_1(T))$) and $\lambda_0\geq 0$ ($\lambda_0=0$ if $\mathbb{E}\langle\phi_x(\bar{x}(T)),\alpha_0\rangle+\frac{1}{2}\mathbb{E}\langle\phi_{xx}(\bar{x}(T))y_1(T),y_1(T)\rangle<0$) such that

$$\xi=\lambda_0\phi_x(\bar{x}(T))+\sum_{i\in I_0(\bar{x}(T),y_1(T))}\lambda_i g^i_x(\bar{x}(T)).$$

Consequently,

$$\mathbb{E}\langle\xi,\alpha_0\rangle=-\frac{1}{2}\Big(\lambda_0\mathbb{E}\langle\phi_{xx}(\bar{x}(T))y_1(T),y_1(T)\rangle+\sum_{i\in I_0(\bar{x}(T),y_1(T))}\lambda_i\mathbb{E}\left\langle g^i_{xx}(\bar{x}(T))y_1(T),y_1(T)\right\rangle\Big).$$

Setting

$$P_1(T)=-\lambda_0\phi_x(\bar{x}(T))-\sum_{i\in I_0(\bar{x}(T),y_1(T))}\lambda_i g^i_x(\bar{x}(T)),$$

$$P_2(T)=-\lambda_0\phi_{xx}(\bar{x}(T))-\sum_{i\in I_0(\bar{x}(T),y_1(T))}\lambda_i g^i_{xx}(\bar{x}(T)),$$

we find that, for any $y_2(T)\in\mathcal{R}^{(2)}(y_1,v)$,

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\langle P_2(T)y_1(T),y_1(T)\rangle\\&=-\frac{1}{2}\Big(\lambda_0\mathbb{E}\langle\phi_{xx}(\bar{x}(T))y_1(T),y_1(T)\rangle+\sum_{i\in I_0(\bar{x}(T),y_1(T))}\lambda_i\mathbb{E}\left\langle g^i_{xx}(\bar{x}(T))y_1(T),y_1(T)\right\rangle\Big)\\&=\mathbb{E}\langle\xi,\alpha_0\rangle\leq-\mathbb{E}\langle P_1(T),y_2(T)\rangle.\end{aligned}$$

This combined with (4.15)–(4.16) implies the second order necessary condition (4.11).

(ii): $\mathcal{L}^{(2)}(y_1(T))\bigcap\mathcal{Q}_{(2)}(y_1(T))=\emptyset$.

To simplify the notation set $g^0(\cdot)=\phi(\cdot)$, $J=\{0\}\cup I(\bar{x}(T),y_1(T))$ and

$$\mathcal{Q}^0_{(2)}(y_1(T))=\big\{z\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)\mid\mathbb{E}\langle\phi_x(\bar{x}(T)),z\rangle+\frac{1}{2}\mathbb{E}\langle\phi_{xx}(\bar{x}(T))y_1(T),y_1(T)\rangle<0\big\}.$$

Clearly, $\mathcal{Q}^0_{(2)}(y_1(T))=\mathcal{L}^{(2)}(y_1(T))$.

If there exists an $i\in J$ such that $\mathcal{Q}^i_{(2)}(y_1(T))=\emptyset$, then, $g^i_x(\bar{x}(T))=0$ a.s. and

$$\mathbb{E}\left\langle g^i_{xx}(\bar{x}(T))y_1(T),y_1(T)\right\rangle\geq 0.\tag{4.17}$$

Let $\lambda_i=1$ and $\lambda_j=0$ for $j\in J\backslash\{i\}$. Then $0=\lambda_i g^i_x(\bar{x}(T))+\sum_{j\in J\backslash\{i\}}\lambda_j g^j_x(\bar{x}(T))$. Let $P_1(T)=0$, $P_2(T)=-g^i_{xx}(\bar{x}(T))$, it is easy to verify that $(P_1(t),Q_1(t))\equiv 0$, $H[t]\equiv 0$, $H_{xx}[t]\equiv 0$ and, by (4.17), $\mathbb{E}\langle P_2(T)y_1(T),y_1(T)\rangle\leq 0$. Then, by Itô's formula, the condition (4.11) holds true and it reduces to

$$\langle P_2(0)\nu_0,\nu_0\rangle+\mathbb{E}\int_0^T\Big(\sum_{j=1}^d\left\langle P_2(t)\sigma^j_u[t]v(t),\sigma^j_u[t]v(t)\right\rangle+2\langle\mathbb{S}[t]y_1(t),v(t)\rangle\Big)dt\leq 0,$$

where, in this case,

$$\mathbb{S}[t]=b_u[t]^\top P_2(t)+\sum_{j=1}^d \sigma_u^j[t]^\top Q_2^j(t)+\sum_{j=1}^d \sigma_u^j[t]^\top P_2(t)\sigma_x^j[t].$$

If $\mathcal{Q}^i_{(2)}(y_1(T))\neq\emptyset$ for any $i\in J$, then, one can find an $i_0\in J$ and a subset $J^0\subset J$ with $i_0\notin J^0$, such that $\cap_{j\in J^0}\mathcal{Q}^j_{(2)}(y_1(T))$ is nonempty and

$$\mathcal{Q}^{i_0}_{(2)}(y_1(T))\bigcap\Big(\bigcap_{j\in J^0}\mathcal{Q}^j_{(2)}(y_1(T))\Big)=\emptyset.$$

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

$$\sup_{\alpha\in\mathcal{Q}^{i_0}_{(2)}(y_1(T))}\mathbb{E}\left\langle \xi,\alpha\right\rangle\leq \inf_{\beta\in\bigcap\limits_{j\in J^0}\mathcal{Q}^j_{(2)}(y_1(T))}\mathbb{E}\left\langle \xi,\beta\right\rangle.$$

By the first conclusion of Lemma 2.5, for some $\alpha_0\in cl\big(\mathcal{Q}^{i_0}_{(2)}(y_1(T))\big)$ and $\beta_0\in\bigcap\limits_{j\in J^0}cl\big(\mathcal{Q}^j_{(2)}(y_1(T))\big)$ we have

$$\mathbb{E}\left\langle \xi,\alpha_0\right\rangle=\sup_{\alpha\in\mathcal{Q}^{i_0}_{(2)}(y_1(T))}\mathbb{E}\left\langle \xi,\alpha\right\rangle\leq \inf_{\beta\in\bigcap\limits_{j\in J^0}\mathcal{Q}^j_{(2)}(y_1(T))}\mathbb{E}\left\langle \xi,\beta\right\rangle=\mathbb{E}\left\langle \xi,\beta_0\right\rangle. \tag{4.18}$$

Moreover, by the second conclusion of Lemma 2.5, there exists a $\lambda_{i_0}>0$ such that $\xi=\lambda_{i_0}g_x^{i_0}(\bar{x}(T))$,

$$0=\mathbb{E}\left\langle g_x^{i_0}(\bar{x}(T)),\alpha_0\right\rangle+\frac{1}{2}\mathbb{E}\left\langle g_{xx}^{i_0}(\bar{x}(T))y_1(T),y_1(T)\right\rangle \tag{4.19}$$

and, for some $j\in J^0$,

$$0=\mathbb{E}\left\langle g_x^{j}(\bar{x}(T)),\beta_0\right\rangle+\frac{1}{2}\mathbb{E}\left\langle g_{xx}^{j}(\bar{x}(T))y_1(T),y_1(T)\right\rangle. \tag{4.20}$$

Denote by J^1 the set of all indices $j\in J^0$ satisfying (4.20). Then, by the second conclusion in Lemma 2.5 again, for each $j\in J^1$, there exists a $\lambda_j\geq 0$ such that

$$-\xi=-\lambda_{i_0}g_x^{i_0}(\bar{x}(T))=\sum_{j\in J^1}\lambda_j g_x^j(\bar{x}(T)). \tag{4.21}$$

Combining (4.18)–(4.21), we arrive at

$$0\leq \lambda_{i_0}\mathbb{E}\left\langle g_{xx}^{i_0}(\bar{x}(T))y_1(T),y_1(T)\right\rangle+\sum_{j\in J^1}\lambda_j\mathbb{E}\left\langle g_{xx}^{j}(\bar{x}(T))y_1(T),y_1(T)\right\rangle.$$

Set $P_1(T)=0$ and $P_2(T)=-\lambda_{i_0}g_{xx}^{i_0}(\bar{x}(T))-\sum\limits_{j\in J^1}\lambda_j g_{xx}^j(\bar{x}(T))$. Then, $(P_1(t),Q_1(t))\equiv 0$, $H[t]\equiv 0$, $H_{xx}[t]\equiv 0$ and, $\mathbb{E}\left\langle P_2(T)y_1(T),y_1(T)\right\rangle\leq 0$. Proceeding as before, we obtain the second order necessary condition (4.11).

Finally, we prove that λ_i for $i\in I(\bar{x}(T),y_1(T)))\cup\{0\}$ can be chosen so that, the first order adjoint process (P_1,Q_1) also satisfies the first order necessary condition (3.6) and the transversality condition $P_1(0)\in N^C_{K_0}(\bar{x}_0)$.

Since the Clarke tangent cone is convex, $\mathcal{C}_{K_0}(\bar{x}_0) + \mathcal{W}(\bar{x}_0, \nu_0)$ $(\subset T^{b(2)}_{K_0}(\bar{x}_0,\nu_0))$ and $\mathcal{C}_{\mathcal{V}}(\bar{u}) + \mathcal{M}(\bar{u}, v)$ $(\subset T^{b(2)}_{\mathcal{V}}(\bar{u}, v))$ are also convex. The above proof applied to $\mathcal{C}_{K_0}(\bar{x}_0) + \mathcal{W}(\bar{x}_0, \nu_0)$, $\mathcal{C}_{\mathcal{V}}(\bar{u}) + \mathcal{M}(\bar{u}, v)$ (instead of $\mathcal{W}(\bar{x}_0, \nu_0)$, $\mathcal{M}(\bar{u}, v)$) implies that $\{\lambda_i\}$ can be chosen so that, the second order necessary condition (4.11) holds true for any $\varpi_0 \in \mathcal{C}_{K_0}(\bar{x}_0) + \mathcal{W}(\bar{x}_0, \nu_0)$ and any $h \in \mathcal{C}_{\mathcal{V}}(\bar{u}) + \mathcal{M}(\bar{u}, v)$. Therefore, for all $\hat{\nu}_0 \in \mathcal{C}_{K_0}(\bar{x}_0)$, $\varpi_0 \in \mathcal{W}(\bar{x}_0, \nu_0)$, $\hat{v} \in \mathcal{C}_{\mathcal{V}}(\bar{u})$, $h \in \mathcal{M}(\bar{u}, v)$ and $\varrho > 0$,

$$
\begin{aligned}
(4.22)\qquad & \langle P_1(0), \varrho\hat{\nu}_0 + \varpi_0\rangle + \frac{1}{2}\langle P_2(0)\nu_0, \nu_0\rangle \\
& +\mathbb{E}\int_0^T \Big(\big\langle H_u[t], \varrho\hat{v}(t) + h(t)\big\rangle + \frac{1}{2}\langle H_{uu}[t]v(t), v(t)\rangle \\
& +\frac{1}{2}\sum_{j=1}^d \big\langle P_2(t)\sigma_u^j[t]v(t), \sigma_u^j[t]v(t)\big\rangle + \langle \mathbb{S}[t]y_1(t), v(t)\rangle\Big)dt \le 0,
\end{aligned}
$$

where (P_1, Q_1) and (P_2, Q_2) are solutions to the adjoint equations (3.4) and (4.7) with the final data given by (4.9) and (4.10), respectively. Dividing by ϱ the both sides of (4.22) and letting $\varrho \to +\infty$, we get

$$
(4.23)\qquad \langle P_1(0), \hat{\nu}_0\rangle + \mathbb{E}\int_0^T \langle H_u[t], \hat{v}(t)\rangle\, dt \le 0.
$$

By the arbitrariness of $\hat{\nu}_0$ and $\hat{v}$, we deduce from (4.23) that $P_1(0) \in N^C_{K_0}(\bar{x}_0)$ and

$$
\mathbb{E}\int_0^T \langle H_u[t], \hat{v}(t)\rangle\, dt \le 0, \quad \forall\, \hat{v} \in \mathcal{C}_{\mathcal{V}}(\bar{u}).
$$

This, combined with Lemma 2.7, yields (3.6). □

Remark 4.2. *Similarly to the deterministic framework, the second order necessary condition is valid only for critical directions. In particular, $\mathcal{Y}(\bar{x}, \bar{u})$ (defined by (4.3)) has to be a nonempty set. If $\mathcal{Q}_{(1)} \neq \emptyset$, $U = \mathbb{R}^m$ and the linear control system (3.1) (in which $v(\cdot)$ is regarded as a control variable) is exactly controllable, then $\mathcal{Y}(\bar{x}, \bar{u}) \neq \emptyset$.*

Remark 4.3. *Since the second order adjacent set of the convex set is convex, when both K_0 and U are convex, for any $\nu_0 \in T^b_{K_0}(\bar{x}_0)$ and $v \in T^b_{\mathcal{V}}(\bar{u})$, we can choose $\mathcal{W}(\bar{x}_0, \nu_0) = T^{b(2)}_{K_0}(\bar{x}_0, \nu_0)$ and $\mathcal{M}(\bar{u}, v) = T^{b(2)}_{\mathcal{V}}(\bar{u}, v)$.*

When both K_0 and U are convex, for any $x_0 \in K_0$ and $u \in \mathcal{V}$, $x_0 - \bar{x}_0 \in T^b_{K_0}(\bar{x}_0)$, $0 \in T^{b(2)}_{K_0}(\bar{x}_0, x_0 - \bar{x}_0)$, $u - \bar{u} \in T^b_{\mathcal{V}}(\bar{u})$ and $0 \in T^{b(2)}_{\mathcal{V}}(\bar{u}, u - \bar{u})$. In particular, choosing $\mathcal{W}(\bar{x}_0, x_0 - \bar{x}_0) = \{0\}$, $\mathcal{M}(\bar{u}, u - \bar{u}) = \{0\}$, as a consequence of Theorem 4.2, we obtain immediately the following result.

Corollary 4.3. *Let (A1)–(A3) hold and $(\bar{x}, \bar{u})$ be a local minimizer for the problem (1.4) with the initial datum $\bar{x}_0$ and the control $\bar{u} \in \mathcal{U} \cap L^4_{\mathbb{F}}(\Omega; L^4(0, T; \mathbb{R}^m))$. If the initial state constraint set K_0 and the control set U are closed and convex, then, for any $u \in \mathcal{V}$ and $x_0 \in K_0$ with the corresponding y_1 satisfying $(y_1(\cdot), u(\cdot) - \bar{u}(\cdot), x_0 - \bar{x}_0) \in \Upsilon(\bar{x}, \bar{u})$ (as in (4.4)), there exist $\lambda_0 \in \{0, 1\}$, $\lambda_i \ge 0$ (for each $i \in I(\bar{x}(T), y_1(T))$) not vanishing simultaneously, and adjoint processes (P_1, Q_1) and (P_2, Q_2) corresponding to $(\bar{x}, \bar{u})$ and the final datum given by (4.9)–(4.10) such that*

$$
\langle P_2(0)(x_0 - \bar{x}_0), (x_0 - \bar{x}_0)\rangle + \mathbb{E}\int_0^T \Big(\langle H_{uu}[t](u(t) - \bar{u}(t)), u(t) - \bar{u}(t)\rangle
$$

$$+\sum_{j=1}^{d}\Big\langle P_2(t)\sigma_u^j[t](u(t)-\bar u(t)),\sigma_u^j[t](u(t)-\bar u(t))\Big\rangle+2\left\langle \mathbb{S}[t]y_1(t),u(t)-\bar u(t)\right\rangle\Big)dt\le 0.$$

In what follows, we refine the second order necessary condition (4.11) by using first and second order adjacent vectors to U at $\bar u(t,\omega)$, which is more convenience in applications than the ones defined in the function space.

Let $v\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$. 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

$$dist(\bar u(t,\omega)+\varepsilon v(t,\omega),U)\le \varepsilon^2\eta(t,\omega),\ \text{a.e. }(t,\omega)\in[0,T]\times\Omega,\ \forall\,\varepsilon\in[0,\varepsilon_0],\tag{4.24}$$

then, similarly to the proof of [10, Theorem 4.1], we deduce that $v\in T^b_{\mathcal{V}}(\bar u)$ and that every $h(\cdot)\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$ satisfying $h(t,\omega)\in T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$ for a.e. $(t,\omega)\in[0,T]\times\Omega$ belongs to $T_{\mathcal{V}}^{b(2)}(\bar u,v)$.

Let $\Psi:[0,T]\times\Omega\rightsquigarrow\mathbb{R}^m$ be a $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable, $\mathbb{F}$-adapted set-valued map with nonempty closed convex values satisfying

$$\Psi(t,\omega)\subset T_U^{b(2)}(\bar u(t,\omega),v(t,\omega)),\ \text{a.e. }(t,\omega)\in[0,T]\times\Omega.$$

We underline that for an arbitrary U, in general, such Ψ may not exist, for instance when the sets $T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$ are empty on a subset of $[0,T]\times\Omega$ of positive measure (When U is equal to the intersection of a finite family of sets desribed by the equality and inequality constraints, then, under the Mangasarian-Fromowitz type constraints qualification assumptions, $\Psi(t,\omega):=T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$ is as requested, see the proof of Corollary 4.5 below). Define

$$\widetilde{\mathcal{M}}(\bar u,v):=\big\{h(\cdot)\in L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))\ \big|\ h(t,\omega)\in\Psi(t,\omega),\ \text{a.e.}\,(t,\omega)\in[0,T]\times\Omega\big\}.\tag{4.25}$$

Clearly, $\widetilde{\mathcal{M}}(\bar u,v)$ is a nonempty convex subset of $T_{\mathcal{V}}^{b(2)}(\bar u,v)$, providing that Ψ has a measurable selection belonging to $L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$.

Remark 4.4. *More generally, when $T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$ is a nonempty closed convex subset of $\mathbb{R}^m$ for a.e. $(t,\omega)\in[0,T]\times\Omega$, we can select $\Psi(\cdot)=T_U^{b(2)}(\bar u(\cdot),v(\cdot))$. Indeed, by arguments similar to the proof of [2, Theorem 8.5.1], one can show that $T_U^{b(2)}(\bar u(\cdot),v(\cdot))$ is $\mathscr{A}^*$-measurable, where $\mathscr{A}^*$ is the completion of $\mathscr{A}$ (defined in (2.1)). Then, for any selection $h^*(\cdot)$ of $T_U^{b(2)}(\bar u(\cdot),v(\cdot))$ there exists a $\mathscr{A}$-measurable modification $h(\cdot)$ of $h^*(\cdot)$. Therefore, $\widetilde{\mathcal{M}}(\bar u,v)$ defined by (4.25) with $\Psi(\cdot)$ replaced by $T_U^{b(2)}(\bar u(\cdot),v(\cdot))$ is a convex subset (might be empty) of $T_{\mathcal{V}}^{b(2)}(\bar u,v)$.*

The following result is another immediate consequence of Theorem 4.2.

Corollary 4.4. *Let (A1)–(A3) hold and $(\bar x,\bar u)$ be a local minimizer for the problem (1.4) with the initial datum $\bar x_0$ and the control $\bar u\in\mathcal{U}\cap L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$. Then, for any $(y_1(\cdot),v(\cdot),\nu_0)\in\Upsilon(\bar x,\bar u)$ such that v satisfies the condition (4.24), the conclusion of Theorem 4.2 is valid for any nonempty convex subset $\mathcal{W}(\bar x_0,\nu_0)$ of $T_{K_0}^{b(2)}(\bar x_0,\nu_0)$ and a nonempty convex subset $\widetilde{\mathcal{M}}(\bar u,v)$ of $T_{\mathcal{V}}^{b(2)}(\bar u,v)$ as in (4.25), provided that there exists Ψ as described above.*

We give next a sufficient condition for some $v\in T^b_{\mathcal{U}}(\bar u)$ to satisfy (4.24) when the control set U is described by finitely many mixed constraints.

Example 4.1. *Let $p,r\in\mathbb{N}$ and*

$$U=\big\{u\in\mathbb{R}^m\,\big|\,\varphi^\theta(u)=0,\ \forall\,\theta=1,\dots,p,\ \psi^\ell(u)\le 0,\ \forall\,\ell=1,\dots,r\big\},\tag{4.26}$$

where $\varphi^1,\dots,\varphi^p:\mathbb{R}^m\to\mathbb{R}$ *and* $\psi^1,\dots,\psi^r:\mathbb{R}^m\to\mathbb{R}$ *are twice continuously differentiable functions. We admit that either equality or inequality constraints may be absent and then they should be simply skipped in the expressions below.*

Set $F_w:=(\varphi^1_u(w),\dots,\varphi^p_u(w))^\top$ *(i.e.,* F_w *is a* $p\times m$*-matrix whose rows are* $\varphi^\theta_u(w)$, $\theta=1,\dots,p$*). Let* $\bar u\in\mathcal{U}$ *and assume that*

(B1) *For any* $w\in U$,

$$\sum_{\theta=1}^{p}\big(|\varphi^\theta_u(w)|+|\varphi^\theta_{uu}(w)|\big)\ +\ \sum_{\ell=1}^{r}\big(|\psi^\ell_u(w)|+|\psi^\ell_{uu}(w)|\big)\le L.$$

(B2) *There exists a* $\rho>0$ *such that for a.e.* $(t,\omega)\in[0,T]\times\Omega$

$$\rho B_{\mathbb{R}^p}\subset F_{\bar u(t,\omega)}B_{\mathbb{R}^m}.$$

(B3) *For some* $\delta>0$, $e>0$ *and for a.e.* $(t,\omega)\in[0,T]\times\Omega$, *we can find a* $\bar v_{t,\omega}\in B_{\mathbb{R}^m}$ *satisfying*

$$\big\langle\varphi^\theta_u(\bar u(t,\omega)),\bar v_{t,\omega}\big\rangle=0,\ \forall\,\theta=1,\dots,p,\ \ \big\langle\psi^\ell_u(\bar u(t,\omega)),\bar v_{t,\omega}\big\rangle\le -e,\ \forall\,\ell\in I_\delta(\bar u(t,\omega)),$$

where $I_\delta(\bar u(t,\omega)):=\{\ell\mid\psi^\ell(\bar u(t,\omega))\in[-\delta,0],\ell=1,\dots,r\}$.

We claim that any $v(\cdot)\in L^\infty_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^m)$ *such that*

$$\big\langle\varphi^\theta_u(\bar u(t,\omega)),v(t,\omega)\big\rangle=0,\ \text{a.e. }(t,\omega)\in[0,T]\times\Omega,\ \forall\,\theta=1,\dots,p,\tag{4.27}$$

$$\big\langle\psi^\ell_u(\bar u(t,\omega),v(t,\omega)\big\rangle\le 0,\ \text{a.e. }(t,\omega)\in[0,T]\times\Omega,\ \forall\,\ell\in I_\delta(\bar u(t,\omega))\tag{4.28}$$

satisfies the condition (4.24).

Clearly, (B2) implies that $F_{\bar u(t,\omega)}$ is surjective and (B3) is a uniform Mangasarian-Fromowitz condition. We do not request any measurability of the selection $\bar v_{t,\omega}$.

We first deduce from (B1)–(B3) that for some $\tilde\varepsilon>0$ and for a.e. $(t,\omega)\in[0,T]\times\Omega$,

$$\frac{3\rho}{4}B_{\mathbb{R}^p}\subset F_uB_{\mathbb{R}^m},\ \forall\,u\in\bar u(t,\omega)+\tilde\varepsilon B_{\mathbb{R}^m},\tag{4.29}$$

and for any $u\in\bar u(t,\omega)+\tilde\varepsilon B_{\mathbb{R}^m}$ one can find a $\bar v\in B_{\mathbb{R}^m}$ satisfying

$$\big\langle\varphi^\theta_u(u),\bar v\big\rangle=0,\ \forall\,\theta=1,\dots,p\quad\text{and}\quad\big\langle\psi^\ell_u(u),\bar v\big\rangle\le-\frac{e}{4},\ \forall\,\ell\in I_\delta(\bar u(t,\omega)).\tag{4.30}$$

Indeed, we may assume that $\rho<1<L$ and $e<1$. By (B1) and (B2), we can find $0<\tilde\varepsilon<e\rho/8L^2(1+e)$ independent of (t,ω), such that (4.29) holds true for a.e. $(t,\omega)\in[0,T]\times\Omega$. Fix a $(t,\omega)\in[0,T]\times\Omega$ such that (4.29) is satisfied. Also, we fix a $u\in\bar u(t,\omega)+\tilde\varepsilon B_{\mathbb{R}^m}$. Define a map $G:\mathbb{R}^m\to\mathbb{R}^p$ by $G(v)=F_uv$ for any $v\in\mathbb{R}^m$. Noting that $G^{(1)}(v)=F_uB_{\mathbb{R}^m}$ for any $v\in\mathbb{R}^m$ (with $G^{(1)}(\cdot)$ as in Definition 2.3) and applying [6, Theorem 3.1] to G, we obtain that

$$dist(\bar v_{t,\omega},G^{-1}(0))\le\frac{4}{3\rho}|G(\bar v_{t,\omega})-0|.$$

Then, there exists a v' such that $G(v')=0$ and

$$|\bar v_{t,\omega}-v'|\le\frac{2}{\rho}|G(\bar v_{t,\omega})-0|=\frac{2}{\rho}|F_u\bar v_{t,\omega}-F_{\bar u(t,\omega)}\bar v_{t,\omega}|\le\frac{2L}{\rho}|u-\bar u(t,\omega)|\le\frac{e}{4L}.$$

Thus $|v'| \le 1+e$ and for every $\ell \in I_\delta(\bar u(t,\omega))$,

$$\begin{aligned}\big\langle \psi^\ell_u(u), v'\big\rangle &= \big\langle \psi^\ell_u(u)-\psi^\ell_u(\bar u(t,\omega)), v'\big\rangle + \big\langle \psi^\ell_u(\bar u(t,\omega)), \bar v_{t,\omega}\big\rangle + \big\langle \psi^\ell_u(\bar u(t,\omega)), v'-\bar v_{t,\omega}\big\rangle\\ &\le L|u-\bar u(t,\omega)|(1+e) - e + \frac{e}{4} \le -\frac{e}{2}.\end{aligned}$$

Hence $\bar v := \frac{1}{2}v' \in B_{\mathbb{R}^m}$ satisfies (4.30).

Let $v(\cdot) \in L^\infty_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^m)$ be as in (4.27)–(4.28). We derive next from (4.29)–(4.30) that $v(\cdot)$ satisfies the condition (4.24) provided that $I_\delta(\bar u(t,\omega)) \neq \emptyset$. Let $I_\delta(\bar u(t,\omega)) = \{\ell_1,\dots,\ell_{k'}\}$ (for some $k'\in\mathbb{N}$). Observe that for any $\ell \in I_\delta(\bar u(t,\omega))$, $[-e/4, e/4] \subset \big\langle \psi^\ell_u(u), \bar v\big\rangle + [0,2L]$. Let $\tilde v := \frac{1}{4Lk'}\bar v$. Then, $|\tilde v| \le \frac{1}{4Lk'}$. By (4.30) we have, for any $\theta = 1,\dots,p$,

$$\text{(4.31)}\quad \big\langle \varphi^\theta_u(u), \tilde v\big\rangle = 0,\ \Big[-\frac{e}{16Lk'}, \frac{e}{16Lk'}\Big] \subset \big\langle \psi^\ell_u(u), \tilde v\big\rangle + \Big[0, \frac{1}{2k'}\Big] \quad \forall\, \ell \in I_\delta(\bar u(t,\omega)).$$

Fix a $(t,\omega)\in[0,T]\times\Omega$ such that (4.29) is satisfied, and let $\tilde\varepsilon > 0$ and $\tilde v$ be as above. Define a map $\widetilde G : \mathbb{R}^m \to \mathbb{R}^{p+k'}$ by $\widetilde G(w) := (\varphi^1(w),\dots,\varphi^p(w),\psi^{\ell_1}(w),\dots,\psi^{\ell_{k'}}(w))^\top$ and a map $\widehat G : \mathbb{R}^m\times\mathbb{R}^{k'}_+ \to \mathbb{R}^{p+k'}$ by $\widehat G(w,q) := \widetilde G(w) + (0_{\mathbb{R}^p}, q)$. Then (4.31) implies that for all $u \in \bar u(t,\omega) + \tilde\varepsilon B_{\mathbb{R}^m}$ and any $q \in \mathbb{R}^{k'}_+$,

$$\text{(4.32)}\qquad \{0_{\mathbb{R}^p}\}\times[-e/16Lk', e/16Lk']^{k'} \subset \widehat G^{(1)}(u,q).$$

Clearly for every $\varrho \in (0,1]$ and u, q as above,

$$\text{(4.33)}\qquad \varrho \widetilde G_u(u) B_{\mathbb{R}^m} \subset \widehat G^{(1)}(u,q).$$

Hence, taking the convex combination of (4.32) and (4.33), we obtain

$$\text{(4.34)}\qquad \frac{\varrho}{2}\widetilde G_u(u)B_{\mathbb{R}^m} + \{0_{\mathbb{R}^p}\}\times\Big[-\frac{e}{32Lk'}, \frac{e}{32Lk'}\Big]^{k'} \subset co\,\widehat G^{(1)}(u,q).$$

Setting $\varrho = e/32L^2k'$, from (4.29) and (4.34) we deduce that

$$\text{(4.35)}\qquad \frac{3\varrho\rho}{8}B_{\mathbb{R}^p}\times\frac{e}{64Lk'}B_{\mathbb{R}^{k'}} \subset co\,\widehat G^{(1)}(u,q).$$

Observe that ϱ and $\tilde\varepsilon$, do not depend on (t,ω).

Let $\bar q$ be such that $\widehat G(\bar u(t,\omega),\bar q) = 0$. Applying [6, Theorem 3.2] at $(\bar u(t,\omega),\bar q)$, we deduce that, there is a constant $c > 0$ depending only on ϱ, ρ, L and e but independent of (t,ω) so that for every $(u,q) \in (\bar u(t,\omega),\bar q) + \dfrac{\tilde\varepsilon}{4}B_{\mathbb{R}^{m\times k'}}$ with $|\widehat G(u,q)| < \dfrac{\tilde\varepsilon c}{4}$ we have

$$\text{(4.36)}\qquad dist((u,q),\widehat G^{-1}(0)) \le \frac{1}{c}|\widehat G(u,q) - 0|.$$

It is sufficient to consider the case $\|v\|_\infty > 0$. By (B1), for all $\varepsilon > 0$, $\widetilde G(\bar u(t,\omega) + \varepsilon v(t,\omega)) = \widetilde G(\bar u(t,\omega)) + \varepsilon \widetilde G_u(\bar u(t,\omega))v(t,\omega) + O(\varepsilon^2)$, where $|O(\varepsilon^2)| \le L\varepsilon^2\|v\|^2_\infty$. Let $(0_{\mathbb{R}^p}, q_\varepsilon) = -\widetilde G(\bar u(t,\omega)) - \varepsilon\widetilde G_u(\bar u(t,\omega))v(t,\omega) \in \{0_{\mathbb{R}^p}\}\times\mathbb{R}^{k'}_+$. Then $|\widehat G(\bar u(t,\omega) + \varepsilon v(t,\omega), q_\varepsilon)| = |O(\varepsilon^2)|$.

Let $\varepsilon>0$ be small enough so that $\varepsilon\|v\|_\infty(1+L)<\frac{\tilde{\varepsilon}}{8}$ and $L\varepsilon^2\|v\|_\infty^2<\frac{\tilde{\varepsilon}c}{4}$. Then, $(\bar u(t,\omega)+\varepsilon v(t,\omega),q_\varepsilon)\in(\bar u(t,\omega),\bar q)+\dfrac{\tilde{\varepsilon}}{4}B_{\mathbb{R}^{m\times k'}}$ and $|\widehat{G}(\bar u(t,\omega)+\varepsilon v(t,\omega),q_\varepsilon)|<\dfrac{\tilde{\varepsilon}c}{4}$, and, by (4.36), there exists a $(u_\varepsilon,\tilde q_\varepsilon)\in\mathbb{R}^m\times\mathbb{R}_+^{k'}$ such that

$$\widehat{G}(u_\varepsilon,\tilde q_\varepsilon)=0,\ \ |\bar u(t,\omega)+\varepsilon v(t,\omega)-u_\varepsilon|\le\frac{1}{c}|\widehat{G}(\bar u(t,\omega)+\varepsilon v(t,\omega),q_\varepsilon)|\le\frac{L}{c}\varepsilon^2\|v\|_\infty^2.$$

Consequently, in order to prove that v satisfies the condition (4.24) with $\eta(t,\omega)=\frac{L}{c}\|v\|_\infty^2$, we only need to check that $u_\varepsilon\in U$. Since $\tilde q_\varepsilon\in\mathbb{R}_+^{k'}$, we have

$$\varphi^\theta(u_\varepsilon)=0,\ \theta=1,\cdots,p\ \text{ and }\ \psi^\ell(u_\varepsilon)\le 0,\ \ell\in I_\delta(\bar u(t,\omega)). \tag{4.37}$$

Further, for all $\ell\notin I_\delta(\bar u(t,\omega))$,

$$\begin{aligned}\psi^\ell(u_\varepsilon)&\le\psi^\ell(\bar u(t,\omega))+|\psi^\ell(u_\varepsilon)-\psi^\ell(\bar u(t,\omega))|<-\delta+L|u_\varepsilon-\bar u(t,\omega)|\\&\le-\delta+L\varepsilon\|v\|_\infty+L|u_\varepsilon-\bar u(t,\omega)-\varepsilon v(t,\omega)|\\&\le-\delta+L\varepsilon\|v\|_\infty+\frac{L^2\varepsilon^2}{c}\|v\|_\infty^2.\end{aligned}$$

Therefore, $u_\varepsilon\in U$ whenever ε is so that $L\varepsilon\|v\|_\infty+\frac{L^2\varepsilon^2}{c}\|v\|_\infty^2<\delta$. This proves that there exist constants $C>0$ and $\varepsilon_0>0$ independent of (t,ω) such that for all $\varepsilon\in(0,\varepsilon_0)$ and for almost all (t,ω) satisfying $I_\delta(\bar u(t,\omega))\neq\emptyset$,

$$dist(\bar u(t,\omega)+\varepsilon v(t,\omega),U)\le C\varepsilon^2.$$

It remains to consider the case $I_\delta(\bar u(t,\omega))=\emptyset$. For this aim it is enough to apply the same arguments as above omitting functions ψ^ℓ, $\ell=1,\dots,r$. This completes the proof of our claim.

In what follows, we shall give a consequence of Theorem 4.2 for the case when U is represented by (4.26). We need the following assumption.

(B4) For any $w\in U$ with the active indices $I(w):=\{\ell\in\{1,\cdots,r\}\,|\,\psi^\ell(w)=0\}=\{\ell_1,\dots,\ell_{k_0}\}$ (for some $k_0\in\mathbb{N}$),

(4.38)
$$\{\varphi_u^1(w),...,\varphi_u^p(w)\}\cup\{\psi_u^{\ell_j}(w)\,|\,j=1,\dots,k_0\}\text{ are linearly independent.}$$

Furthermore, there exists a constant $\rho>0$ such that for every $w\in U$,

$$\rho B_{Im(\Gamma_{I(w)})}\subset\Gamma_{I(w)}B_{\mathbb{R}^{p+k_0}}, \tag{4.39}$$

where $\Gamma_{I(w)}:=(\varphi_u^1(w),\dots,\varphi_u^p(w),\psi_u^{\ell_1}(w),\dots,\psi_u^{\ell_{k_0}}(w))$, and, $B_{Im(\Gamma_{I(w)})}$ is the closed unit ball in the image space of $\Gamma_{I(w)}$.

Corollary 4.5. *Let U be given by (4.26), the assumptions (A1)–(A3), (B1) and (B4) hold and $(\bar x,\bar u)$ be a local minimizer for the problem (1.4) with the initial datum $\bar x_0$ and the control $\bar u\in\mathcal{U}\cap L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$. Then, for any $(y_1(\cdot),v(\cdot),\nu_0)\in\Upsilon(\bar x,\bar u)$ such that $v\in L^\infty_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^m)$ satisfies (4.24) and any nonempty convex subset $\mathcal{W}(\bar x_0,\nu_0)$ of $T^{b(2)}_{K_0}(\bar x_0,\nu_0)$, there exist $\mu^\theta(\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{R}))$, $\theta=1,\dots,p$, $\gamma^\ell(\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;\ \mathbb{R}_+))$, $\ell=1,\dots,r$ and $\lambda_0\in\{0,1\}$, $\lambda_i\ge 0$ for each $i\in I(\bar x(T),y_1(T))$ not vanishing simultaneously such that for every $\varpi_0\in\mathcal{W}(\bar x_0,\nu_0)$,*

$$\langle P_1(0),\varpi_0\rangle+\frac{1}{2}\langle P_2(0)\nu_0,\nu_0\rangle+\mathbb{E}\int_0^T\Big(\frac{1}{2}\langle H_{uu}[t]v(t),v(t)\rangle \tag{4.40}$$

$$+\frac{1}{2}\sum_{j=1}^{d}\left\langle P_2(t)\sigma_u^j[t]v(t),\sigma_u^j[t]v(t)\right\rangle+\langle\mathbb{S}[t]y_1(t),v(t)\rangle$$
$$-\frac{1}{2}\sum_{\theta=1}^{p}\mu^\theta(t)\left\langle\varphi_{uu}^\theta(\bar u(t))v(t),v(t)\right\rangle-\frac{1}{2}\sum_{\ell\in I_v(\bar u(t))}\gamma^\ell(t)\left\langle\psi_{uu}^\ell(\bar u(t))v(t),v(t)\right\rangle\Big)dt\le 0.$$

Here, (P_1,Q_1) and (P_2,Q_2) are respectively the adjoint processes corresponding to $(\bar x,\bar u)$ with the final datum (4.9) and (4.10), and

$$I_v(\bar u(t))=\left\{\ell\in I(\bar u(t))\;\middle|\;\left\langle\psi_u^\ell(\bar u(t)),v(t)\right\rangle=0\right\}.$$

Proof. By the linear independence assumption (4.38),

$$T_U^b(\bar u(t,\omega))=\big\{v\in\mathbb{R}^m\;\big|\;\left\langle\varphi_u^\theta(\bar u(t,\omega)),v\right\rangle=0,\ \theta=1,\dots,p,$$
$$\text{and }\left\langle\psi^\ell(\bar u(t,\omega)),v\right\rangle\le 0,\ \ \ell\in I(\bar u(t,\omega))\big\},\quad\text{a.e. }(t,\omega)\in[0,T]\times\Omega.$$

Furthermore, for any $v(t,\omega)\in T_U^b(\bar u(t,\omega))$, we have $T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))\neq\emptyset$, and,

$$T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$$
$$=\Big\{h\in\mathbb{R}^m\;\Big|\;\langle\varphi_u^\theta(\bar u(t,\omega)),h\rangle+\frac{1}{2}\langle\varphi_{uu}^\theta(\bar u(t,\omega))v(t,\omega),v(t,\omega)\rangle=0,\ \theta=1,\dots,p,$$
$$\text{and }\ \langle\psi_u^\ell(\bar u(t,\omega)),h\rangle+\frac{1}{2}\langle\psi_{uu}^\ell(\bar u(t,\omega))v(t,\omega),v(t,\omega)\rangle\le 0,\ \ell\in I_v(\bar u(t,\omega))\Big\}.$$

Obviously, $T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$ is nonempty and convex. Let $\widetilde{\mathcal{M}}(\bar u,v)$ be the set defined by (4.25) with $\Psi(t,\omega)=T_U^{b(2)}(\bar u(t,\omega),v(t,\omega))$. Under condition (B1) and (B4), using a similar discussion in the below, we have $\widetilde{\mathcal{M}}(\bar u,v)$ is nonempty. By Corollary 4.4, there exist $\lambda_0\in\{0,1\}$, $\lambda_i\ge 0$ (for each $i\in I(\bar x(T),y_1(T))$) not vanishing simultaneously, and adjoint processes (P_1,Q_1) and (P_2,Q_2) corresponding to $(\bar x,\bar u)$ and the final datum defined respectively by (4.9) and (4.10) such that, for every $\varpi_0\in\mathcal{W}(\bar x_0,\nu_0)$ and $h\in\widetilde{\mathcal{M}}(\bar u,v)$,

$$\text{(4.41)}\quad\langle P_1(0),\varpi_0\rangle+\frac{1}{2}\langle P_2(0)\nu_0,\nu_0\rangle+\mathbb{E}\int_0^T\Big(\langle H_u[t],h(t)\rangle+\frac{1}{2}\langle H_{uu}[t]v(t),v(t)\rangle$$
$$+\frac{1}{2}\sum_{j=1}^{d}\left\langle P_2(t)\sigma_u^j[t]v(t),\sigma_u^j[t]v(t)\right\rangle+\langle\mathbb{S}[t]y_1(t),v(t)\rangle\Big)dt\le 0$$

and

$$\text{(4.42)}\qquad P_1(0)\in N_{K_0}^C(\bar x_0)\text{ and }H_u[t]\in N_U^C(\bar u(t)),\ \text{a.e. }t\in[0,T],\ a.s.$$

Since in this special case $\mathcal{C}_U(\bar u(t,\omega))=T_U^b(\bar u(t,\omega))$,

$$H_u[t]\in\sum_{\theta=1}^{p}\mathbb{R}\varphi_u^\theta(\bar u(t))+\sum_{\ell\in I(\bar u(t))}\mathbb{R}_+\psi_u^\ell(\bar u(t)),\ \text{a.e. }t\in[0,T],\ \text{a.s.}$$

By the condition (4.39) and using the same arguments as that in [10, Corollary 4.1], we deduce that, there exist processes $\mu^\theta(\cdot) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}))$, $\theta = 1, \dots, p$ and $\gamma^\ell(\cdot) \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}_+))$, $\ell = 1, \dots, r$ such that

$$H_u[t] = \sum_{\theta=1}^{p} \mu^\theta(t)\varphi_u^\theta(\bar{u}(t)) + \sum_{\ell\in I(\bar{u}(t))} \gamma^\ell(t)\psi_u^\ell(\bar{u}(t)), \text{ a.e. } t\in[0,T], \text{ a.s.} \tag{4.43}$$

On the other hand, by the definition of $\Upsilon(\bar{x},\bar{u})$,

$$\mathbb{E}\left\langle \phi_x(\bar{x}(T)), y_1(T)\right\rangle = 0, \quad \mathbb{E}\left\langle g_x^i(\bar{x}(T)), y_1(T)\right\rangle \le 0, \quad \forall\, i\in I(\bar{x}(T)).$$

Then,

$$\langle P_1(T), y_1(T)\rangle = -\left\langle \lambda_0\phi_x(\bar{x}(T)) + \sum_{i\in I(\bar{x}(T),y_1(T))} \lambda_i g_x^i(\bar{x}(T)), y_1(T)\right\rangle \ge 0,$$

which, by the Ito formula, implies that

$$\langle P_1(0), \nu_0\rangle + \mathbb{E}\int_0^T \langle H_u[t], v(t)\rangle\, dt \ge 0.$$

This, together with (4.42) and Lemma 2.7, gives

$$\langle P_1(0), \nu_0\rangle = 0, \text{ and } \langle H_u[t], v(t)\rangle = 0, \quad \text{a.e. } t\in[0,T], \text{ a.s.} \tag{4.44}$$

Combining (4.43) with (4.44), we obtain that,

$$\sum_{\ell\in I(\bar{u}(t))} \gamma^\ell(t)\left\langle \psi_u^\ell(\bar{u}(t), v(t)\right\rangle = 0, \text{ a.e. } t\in[0,T], \text{ a.s.}$$

Therefore, for any $\ell \notin I_v(\bar{u}(t))$, $\gamma^\ell(t) = 0$, a.e. $t\in[0,T]$, a.s.

For every $(t,\omega)\in[0,T]\times\Omega$, let $\ell_1,\dots,\ell_\tau$ be all the elements of $I_v(\bar{u}(t,\omega))$ (for some integer $\tau \le r$). Denote

$$\begin{aligned}\zeta(t,\omega) := -\frac{1}{2}\big(&\left\langle \varphi_{uu}^1(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle, \dots, \left\langle \varphi_{uu}^p(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle,\\ &\left\langle \psi_{uu}^{\ell_1}(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle, \dots, \left\langle \psi_{uu}^{\ell_\tau}(\bar{u}(t,\omega))v(t,\omega), v(t,\omega)\right\rangle\big),\end{aligned}$$

$$A := \big\{(t,\omega)\in[0,T]\times\Omega \mid \zeta(t,\omega)\ne 0\big\},$$

and,

$$\Gamma_{I_v(\bar{u}(t,\omega))} := (\varphi_u^1(\bar{u}(t,\omega)), \dots, \varphi_u^p(\bar{u}(t,\omega)), \psi_u^{\ell_1}(\bar{u}(t,\omega)), \dots, \psi_u^{\ell_\tau}(\bar{u}(t,\omega))).$$

It is clear that the set $A\in\mathscr{A}$ with $\mathscr{A}$ defined by (2.1). Define a set-valued map by

$$\Phi(t,\omega) := \left\{ h\in B_{\mathbb{R}^m} \;\middle|\; h^\top\Gamma_{I_v(\bar{u}(t,\omega))} = \frac{\rho\zeta(t,\omega)}{|\zeta(t,\omega)|}\right\}, \quad (t,\omega)\in A.$$

By the condition (4.39), for any $(t,\omega)\in A$, $\Phi(t,\omega)\ne\emptyset$ and the graph of Φ is $\mathscr{A}\otimes\mathcal{B}(\mathbb{R}^m)$-measurable. Therefore, there exists an $\mathscr{A}$-measurable map $h(\cdot)$ (making a completion argument if necessary) such that, $h(t,\omega)\in\Phi(t,\omega)$, a.e. $(t,\omega)\in A$. Set

$$\tilde{h}(t,\omega) := \begin{cases} \dfrac{|\zeta(t,\omega)|h(t,\omega)}{\rho}, & \text{if } (t,\omega)\in A,\\ 0, & \text{otherwise.}\end{cases}$$

Then $\tilde{h}$ is $\mathscr{A}$-measurable and, by Lemma 2.8, $\tilde{h}$ is $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted. Moreover, $\zeta(t,\omega)=\tilde{h}(t,\omega)^\top\Gamma_{I_v(\bar{u}(t,\omega))}$, i.e.,

$$\left\langle \varphi_u^\theta(\bar{u}(t,\omega)),\tilde{h}(t,\omega)\right\rangle=-\frac{1}{2}\left\langle \varphi_{uu}^\theta(\bar{u}(t,\omega))v(t,\omega),v(t,\omega)\right\rangle,\ \forall\, i=1,\dots,p, \tag{4.45}$$

and

$$\left\langle \psi_u^\ell(\bar{u}(t,\omega)),\tilde{h}(t,\omega)\right\rangle=-\frac{1}{2}\left\langle \psi_{uu}^\ell(\bar{u}(t,\omega))v(t,\omega),v(t,\omega)\right\rangle,\ \forall\, \ell\in I_v(\bar{u}(t,\omega)). \tag{4.46}$$

In addition, by the condition (B1),

$$|\tilde{h}(t,\omega)|=\Big|\frac{|\zeta(t,\omega)|h(t,\omega)}{\rho}\Big|\le\frac{|\zeta(t,\omega)|}{\rho}\le C\|v\|_\infty^2,\quad \text{a.e. } (t,\omega)\in[0,T]\times\Omega.$$

Therefore, $\tilde{h}(\cdot)\in L^\infty_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^m)\subset L^4_{\mathbb{F}}(\Omega;L^4(0,T;\mathbb{R}^m))$. This yields $h\in\widetilde{\mathcal{M}}(\bar{u},v)$. Moreover, by (4.45) and (4.46),

$$\left\langle H_u[t],\tilde{h}(t)\right\rangle=-\frac{1}{2}\sum_{\theta=1}^{p}\mu^\theta(t)\left\langle \varphi_{uu}^\theta(\bar{u}(t))v(t),v(t)\right\rangle-\frac{1}{2}\sum_{\ell\in I_v(\bar{u}(t))}\gamma^\ell(t)\left\langle \psi_{uu}^\ell(\bar{u}(t))v(t),v(t)\right\rangle.$$

Substituting this equality into (4.41), we obtain (4.40). □

Because the Bolza problem can be reformulated as the Mayer one, the first and second order necessary conditions obtained in this paper can be extended to the Bolza problem. Actually, for the Bolza type cost function

$$J(x(\cdot),u(\cdot))=\mathbb{E}\Big[\int_0^T f(t,x(t),u(t))dt+\phi(x(T))\Big], \tag{4.47}$$

where the function $\phi:\mathbb{R}^n\times\Omega\to\mathbb{R}$ satisfies the conditions (A2) (ii) and (A3) (ii), and $f:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\Omega\to\mathbb{R}$ is a given function satisfying :

(A4) *For any* $(x,u)\in\mathbb{R}^n\times\mathbb{R}^m$, $f(\cdot,x,u,\cdot):\ [0,T]\times\Omega\to\mathbb{R}$ *is* $\mathcal{B}([0,T])\otimes\mathcal{F}$-*measurable and* $\mathbb{F}$-*adapted. For a.e.* $(t,\omega)\in[0,T]\times\Omega$, $f(t,\cdot,\cdot,\omega):\ \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ *is twice continuously differentiable, and for a.e.* $(t,\omega)\in[0,T]\times\Omega$, *any* $x,\ \tilde{x}\in\mathbb{R}^n$ *and* $u,\ \tilde{u}\in\mathbb{R}^m$,

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

As usual, if we introduce the following extended control system

$$\begin{cases} dx(t)=b(t,x(t),u(t))dt+\sigma(t,x(t),u(t))dW(t), \quad t\in[0,T],\\ dy(t)=f(t,x(t),u(t))dt, \quad t\in[0,T],\\ x(0)=x_0,\ y(0)=0. \end{cases} \tag{4.48}$$

Then the cost function (4.47) can be written in the Mayer form:

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

where $\hat{\phi}(x,y)=\phi(x)+y$. Under the assumptions (A1)–(A4), the new control system is well-defined and so does the cost function.

Let $\widehat{K}_0 = K_0 \times \{0\}$. For $\lambda_0 \in \{0,1\}$, define the generalized Hamiltonian

$$H^{\lambda_0}(t,x,u,p,q,\omega) := \langle p, b(t,x,u,\omega)\rangle + \sum_{j=1}^{d} \big\langle q^j, \sigma^j(t,x,u,\omega)\big\rangle - \lambda_0 f(t,x,u,\omega), \tag{4.49}$$

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

Let $\bar{u} \in \mathcal{U}$ and (P_1, Q_1) be the solution to the following adjoint equation:

$$\begin{cases} dP_1(t) = -\big(b_x[t]^\top P_1(t) + \sum\limits_{j=1}^{d} \sigma_x^j[t]^\top Q_1^j(t) - \lambda_0 f_x[t]\big)dt + \sum\limits_{j=1}^{d} Q_1^j(t)dW^j(t) \\ P_1(T) = -\lambda_0 \phi_x(\bar{x}(T)) - \sum\limits_{i=1}^{k} \lambda_i g_x^i(\bar{x}(T)), \end{cases} \tag{4.50}$$

where $\lambda_i \geq 0$, $i = 1, \dots, k$ (which will be specified later), and denote

$$H^{\lambda_0}[t] = H^{\lambda_0}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t)),$$

$H_u^{\lambda_0}[t]$, $H_{xx}^{\lambda_0}[t]$, $H_{xu}^{\lambda_0}[t]$ and $H_{uu}^{\lambda_0}[t]$ are defined in a similar way.

Similarly, we introduce the following second order adjoint equation

$$\begin{cases} dP_2(t) = -\big(b_x[t]^\top P_2(t) + P_2(t) b_x[t] + \sum\limits_{j=1}^{d} \sigma_x^j[t]^\top P_2(t) \sigma_x^j[t] \\ \qquad + \sum\limits_{j=1}^{d} \sigma_x^j[t]^\top Q_2^j(t) + \sum\limits_{j=1}^{d} Q_2^j(t)\sigma_x^j[t] + H_{xx}^{\lambda_0}[t]\big)dt \\ \qquad + \sum\limits_{j=1}^{d} Q_2^j(t) dW^j(t), \quad t \in [0,T], \\ P_2(T) = -\lambda_0 \phi_{xx}(\bar{x}(T)) - \sum\limits_{i=1}^{k} \lambda_i g_{xx}^i(\bar{x}(T)), \end{cases} \tag{4.51}$$

and set $\mathbb{S}^{\lambda_0}[t] := H_{xu}^{\lambda_0}[t] + b_u[t]^\top P_2(t) + \sum_{j=1}^{d} \sigma_u^j[t]^\top Q_2^j(t) + \sum_{j=1}^{d} \sigma_u^j[t]^\top P_2(t) \sigma_x^j[t]$.

Let $\mathcal{Y}(\bar{x}, \bar{u})$ be as in (4.3) and define

$$\widehat{\Upsilon}(\bar{x},\bar{u}) := \Big\{ (y_1(\cdot), v(\cdot), \nu_0) \in \mathcal{Y}(\bar{x},\bar{u}) \,\Big|\, \mathbb{E}\int_0^T \big(\langle f_x[t], y_1(t)\rangle + \langle f_u[t], v(t)\rangle\big)dt + \mathbb{E}\,\langle \phi_x(\bar{x}(T)), y_1(T)\rangle = 0 \Big\}.$$

As a consequence of Theorems 3.4 and 4.2, we have the following first and second order necessary condition for the Bolza optimal control problem, where we used notations of Theorems 3.4 and 4.2:

Corollary 4.6. *Let (A1)–(A4) hold and $(\bar{x}, \bar{u})$ be a locally optimal pair for the problem (1.4) with the control system (1.1) and the cost function (4.47).*

(i) *If $I(\bar{x}(T)) = \emptyset$ or if $I(\bar{x}(T)) \neq \emptyset$ and $\mathcal{Q}_{(1)} \neq \emptyset$, then there exist $\lambda_0 \in \{0,1\}$ and $\lambda_i \geq 0$ for $i \in I(\bar{x}(T))$, and the solution (P_1, Q_1) to the first order adjoint equation (4.50) corresponding to $(\bar{x},\bar{u})$ such that $\lambda_0 + \mathbb{E}\,|P_1(T)| \neq 0$,*

$$P_1(0) \in N_{K_0}^C(\bar{x}_0) \text{ and } H_u^{\lambda_0}[t] \in N_U^C(\bar{u}(t)), \quad a.e.\ t \in [0,T],\ a.s., \tag{4.52}$$

(ii) *If $I(\bar{x}(T)) \neq \emptyset$ but $\mathcal{Q}_{(1)} = \emptyset$, then for each $i \in I(\bar{x}(T))$, there exists a $\lambda_i \geq 0$ such that $\sum_{i\in I(\bar{x}(T))} \lambda_i > 0$, $\sum_{i \in I(\bar{x}(T))} \lambda_i g_x^i(\bar{x}(T)) = 0$ and the same relations as above hold with $\lambda_0 = 0$ and $(P_1, Q_1) \equiv 0$.*

Moreover the above holds true with $\lambda_0 = 1$ *if* $I(\bar{x}(T)) = \emptyset$ *or if* $I(\bar{x}(T)) \neq \emptyset$ *and* $\mathcal{Q}_{(1)} \cap \mathcal{R}^{(1)} \neq \emptyset$.

Furthermore, if $\bar{u} \in \mathcal{U} \cap L^4_{\mathbb{F}}(\Omega; L^4(0,T;\mathbb{R}^m))$*, then, for any* $(y_1(\cdot), v(\cdot), \nu_0) \in \widehat{\Upsilon}(\bar{x},\bar{u})$ *with* $\mathcal{W}(\bar{x}_0,\nu_0) \neq \emptyset$*,* $\mathcal{M}(\bar{u},v) \neq \emptyset$*, there exist* $\lambda_0 \in \{0,1\}$*,* $\lambda_i \geq 0$ *(for each* $i \in I(\bar{x}(T), y_1(T))$*) not vanishing simultaneously and adjoint processes* (P_1,Q_1) *and* (P_2,Q_2) *defined by (4.50)–(4.51) corresponding to* $(\bar{x},\bar{u})$ *such that (4.52) holds true and for every* $\varpi_0 \in \mathcal{W}(\bar{x}_0,\nu_0)$ *and* $h(\cdot) \in \mathcal{M}(\bar{u},v)$,

$$\begin{aligned}&\langle P_1(0),\varpi_0\rangle + \frac{1}{2}\langle P_2(0)\nu_0,\nu_0\rangle + \mathbb{E}\int_0^T \Big(\big\langle H_u^{\lambda_0}[t], h(t)\big\rangle + \frac{1}{2}\big\langle H_{uu}^{\lambda_0}[t]v(t), v(t)\big\rangle \\ &\quad + \frac{1}{2}\sum_{j=1}^d \big\langle P_2(t)\sigma_u^j[t]v(t), \sigma_u^j[t]v(t)\big\rangle + \big\langle \mathbb{S}^{\lambda_0}[t]y_1(t), v(t)\big\rangle \Big)dt \leq 0.\end{aligned}$$

REFERENCES

[1] J. -P. Aubin, *Optima and Equilibria. An Introduction to Nonlinear Analysis*, Second edition, Springer-Verlag, Berlin, 1998.
[2] J. -P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Berlin, 1990.
[3] J. F. Bonnans, *Second order Pontryagin's principle for stochastic control problems*, Inria Saclay Ile de France. 2015, https://hal.inria.fr/hal-01205854.
[4] 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.
[5] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
[6] H. Frankowska, *High order inverse function theorems*, Ann. Inst. Henri Poincaré, Analyse Non Linéaire, 6 (1989), pp. 283–303.
[7] H. Frankowska and N. Osmolovskii, *Second-order necessary optimality conditions for the Mayer problem subject to a general control constraint*, in *Analysis and Geometry in Control theory and its Applications*, Springer INdAM Ser., 11, Springer, Cham, 2015, pp. 171–207.
[8] H. Frankowska and D. Hoehener, *Pointwise second-order necessary optimality conditions and second-order sensitivity relations in optimal control*, J. Differential Equations, 262 (2017), pp. 5735–5772.
[9] 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.
[10] 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.
[11] U. G. Haussmann, *General necessary conditions for optimal control of stochastic systems*, Math. Program. Study, 6 (1976), pp. 30–48.
[12] 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.
[13] M. Kisielewicz, *Stochastic Differential Inclusions and Applications*, Springer, 2013.
[14] H. J. Kushner, *Necessary conditions for continuous parameter stochastic optimization problems*, SIAM J. Control Optim., 10 (1972), pp. 550–565.
[15] H. Maurer, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Mathematical programming at Oberwolfach (Proc. Conf., Math. Forschungsinstitut, Oberwolfach, 1979), Math. Programming Stud., 14 (1981), pp. 163–177.
[16] N. P. Osmolovskii and H. Maurer, *Applications to Regular and Bang-Bang Control. Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control*, SIAM, Philadelphia, PA, 2012.
[17] 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.
[18] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim., 28 (1990), pp. 966–979.
[19] R. T. Rockafellar and R. J.-B. Wets, (1998), *Variational Analysis*, Gründlehren der Mathematischen Wissensschaften, vol. 317, Springer Verlag, New York.
[20] J. Warga, *A second-order condition that strengthens Pontryagin's maximum principle*, J. Differential Equations , 28 (1978), pp. 284–307.

[21] D. H. Wagner, *Survey of measurable selection theorems*, SIAM J. Control Optim., 15 (1977), pp. 859–903.

[22] 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), 2574–2602.

[23] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, Berlin, 2000.