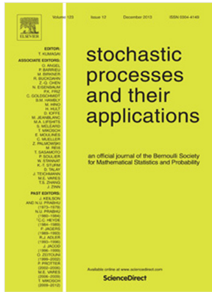
Journal Pre-proof



Necessary conditions for stochastic optimal control problems in infinite dimensions

Hélène Frankowska, Xu Zhang

PII: S0304-4149(18)30759-2
DOI: https://doi.org/10.1016/j.spa.2019.11.010
Reference: SPA 3605

To appear in: *Stochastic Processes and their Applications*

Received date : 26 December 2018
Revised date : 24 September 2019
Accepted date : 26 November 2019

Necessary Conditions for Stochastic Optimal Control Problems in Infinite Dimensions

Hélène Frankowska* and Xu Zhang†

Abstract

The purpose of this paper is to establish the first and second order necessary conditions for stochastic optimal controls in infinite dimensions. The control system is governed by a stochastic evolution equation, in which both drift and diffusion terms may contain the control variable and the set of controls is allowed to be nonconvex. Only one adjoint equation is introduced to derive the first order necessary optimality condition either by means of the classical variational analysis approach or, under an additional assumption, by using differential calculus of set-valued maps. More importantly, in order to avoid the essential difficulty with the well-posedness of higher order adjoint equations, using again the classical variational analysis approach, only the first and the second order adjoint equations are needed to formulate the second order necessary optimality condition, in which the solutions to the second order adjoint equation are understood in the sense of the relaxed transposition.

2010 Mathematics Subject Classification: 93E20, 49J53, 60H15

Key Words: Stochastic optimal control; First and second order necessary optimality conditions; Variational equation; Adjoint equation; Transposition solution; Maximum principle.

1 Introduction and notations

Let $T > \tau \geq 0$, and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be a complete filtered probability space (satisfying the usual conditions). Write $\mathbf{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$, and denote by $\mathbb{F}$ the progressive σ-field (in $[0,T]\times\Omega$) with respect to $\mathbf{F}$. For a Banach space X with the norm $|\cdot|_X$, X^* stands for the dual space of X. For any $t \in [0,T]$ and $r \in [1,\infty)$, denote by $L^r_{\mathcal{F}_t}(\Omega; X)$ the Banach space of all $\mathcal{F}_t$-measurable random variables $\xi : \Omega \to X$ such that $\mathbb{E}|\xi|^r_X < \infty$, with the canonical norm. Also, denote by $D_{\mathbb{F}}([\tau,T]; L^r(\Omega; X))$ the vector space of all X-valued, r-th power integrable $\mathbf{F}$-adapted processes $\phi(\cdot)$ such that $\phi(\cdot) : [\tau,T] \to L^r(\Omega, \mathcal{F}_T, \mathbb{P}; X)$ is càdlàg, i.e., right continuous with left limits. Then $D_{\mathbb{F}}([\tau,T]; L^r(\Omega; X))$ is a Banach space with the norm

$$|\phi(\cdot)|_{D_{\mathbb{F}}([\tau,T];L^r(\Omega;X))} \overset{\triangle}{=} \sup_{t\in[\tau,T)} \left(\mathbb{E}|\phi(t)|^r_X\right)^{\frac{1}{r}}.$$

*CNRS, IMJ-PRG, Sorbonne Université, case 247, 4 place Jussieu, 75252 Paris, France. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0254. It also partially benefitted from "FMJH Program Gaspard Monge in optimization and operation research", and from the support to this program from EDF under grant PGMO 2018-0047H, and the CNRS-NSFC PRC Project under grant 271392. *E-mail:* helene.frankowska@imj-prg.fr.

†School of Mathematics, Sichuan University, Chengdu 610064, China. The research of this author is partially supported by the NSF of China under grants 11931011 and 11221101, the NSFC-CNRS Joint Research Project under grant 11711530142 and the PCSIRT under grant IRT_16R53 from the Chinese Education Ministry. *E-mail:* zhang_xu@scu.edu.cn.

Similarly, denote by $C_{\mathbb{F}}([\tau,T];L^r(\Omega;X))$ the Banach space of all X-valued $\mathbf{F}$-adapted processes $\phi(\cdot)$ so that $\phi(\cdot):[\tau,T]\to L^r_{\mathcal{F}_T}(\Omega;X)$ is continuous, with the norm inherited from $D_{\mathbb{F}}([\tau,T];L^r(\Omega;X))$.

Fix $r_1,r_2,r_3,r_4\in[1,\infty]$ and define

$$L^{r_1}_{\mathbb{F}}(\Omega;L^{r_2}(\tau,T;X))=\Big\{\varphi:(\tau,T)\times\Omega\to X\;\Big|\;\varphi(\cdot)\text{ is }\mathbf{F}\text{-adapted and }\mathbb{E}\Big(\int_\tau^T|\varphi(t)|^{r_2}_X dt\Big)^{\frac{r_1}{r_2}}<\infty\Big\},$$

$$L^{r_2}_{\mathbb{F}}(\tau,T;L^{r_1}(\Omega;X))=\Big\{\varphi:(\tau,T)\times\Omega\to X\;\Big|\;\varphi(\cdot)\text{ is }\mathbf{F}\text{-adapted and }\int_\tau^T\Big(\mathbb{E}|\varphi(t)|^{r_1}_X\Big)^{\frac{r_2}{r_1}}dt<\infty\Big\}.$$

Both $L^{r_1}_{\mathbb{F}}(\Omega;L^{r_2}(\tau,T;X))$ and $L^{r_2}_{\mathbb{F}}(\tau,T;L^{r_1}(\Omega;X))$ are Banach spaces with the canonical norms. If $r_1=r_2$, we simply denote the above spaces by $L^{r_1}_{\mathbb{F}}(\tau,T;X)$. For a Banach space Y, $\mathcal{L}(X;Y)$ stands for the (Banach) space of all bounded linear operators from X to Y, with the usual operator norm (when $Y=X$, we simply write $\mathcal{L}(X)$ instead of $\mathcal{L}(X,Y)$).

Let H be an infinite dimensional separable Hilbert space, V be another separable Hilbert space, and denote by $\mathcal{L}_2^0$ the Hilbert space of all Hilbert-Schmidt operators from V to H. Let $W(\cdot)$ be a V-valued, standard Q-Brownian motion or a cylindrical Brownian motion (see Subsection 2.2 for their definitions). In the sequel, to simplify the presentation, we only consider the case of a cylindrical Brownian motion.

Let A be an unbounded linear operator (with domain $D(A)\subset H$), which generates a C_0-semigroup $\{S(t)\}_{t\geq0}$ on H. Denote by A^* the dual operator of A. It is well-known that $D(A)$ is a Hilbert space with the usual graph norm, and A^* generates a C_0-semigroup $\{S^*(t)\}_{t\geq0}$ on H, which is the dual C_0-semigroup of $\{S(t)\}_{t\geq0}$.

Let $p\geq1$ and denote by p' the Hölder conjugate of p, i.e., $p'=\frac{p}{p-1}$ if $p\in(1,\infty)$; and $p'=\infty$ if $p=1$. Consider a nonempty, closed subset U of another separable Hilbert space $\widetilde{H}$. Put

$$\mathcal{U}_{ad}\triangleq\Big\{u(\cdot)\in L^p_{\mathbb{F}}(\Omega;L^2(0,T;\widetilde{H}))\;\Big|\;u(t,\omega)\in U\text{ a.e. }(t,\omega)\in[0,T]\times\Omega\Big\}.$$

Let us consider the following controlled stochastic evolution equation:

$$\begin{cases}dx(t)=\big(Ax(t)+a(t,x(t),u(t))\big)dt+b(t,x(t),u(t))dW(t) & \text{in }(0,T],\\ x(0)=x_0,\end{cases}\tag{1.1}$$

where $a(\cdot,\cdot,\cdot):[0,T]\times H\times\widetilde{H}\times\Omega\to H$ and $b(\cdot,\cdot,\cdot):[0,T]\times H\times\widetilde{H}\times\Omega\to\mathcal{L}_2^0$, $u(\cdot)\in\mathcal{U}_{ad}$ and $x_0\in H$. In (1.1), $u(\cdot)$ is called the control, while $x(\cdot)=x(\cdot;x_0,u(\cdot))$ is the corresponding state process. As usual, when the context is clear, we omit the $\omega(\in\Omega)$ argument in the defined functions.

Let $f(\cdot,\cdot,\cdot,\cdot):[0,T]\times H\times\widetilde{H}\times\Omega\to\mathbb{R}$ and $g(\cdot,\cdot):H\times\Omega\to\mathbb{R}$. Define the cost functional $\mathcal{J}(\cdot)$ (for the controlled system (1.1)) as follows:

$$\mathcal{J}(u(\cdot))\triangleq\mathbb{E}\Big(\int_0^T f(t,x(t),u(t))dt+g(x(T))\Big),\qquad\forall\,u(\cdot)\in\mathcal{U}_{ad},\tag{1.2}$$

where $x(\cdot)$ is the solution to (1.1) corresponding to the control $u(\cdot)$.

We associate with these data the following optimal control problem for (1.1):

Problem (OP) *Find a control $\bar{u}(\cdot)\in\mathcal{U}_{ad}$ such that*

$$\mathcal{J}\big(\bar{u}(\cdot)\big)=\inf_{u(\cdot)\in\mathcal{U}_{ad}}\mathcal{J}\big(u(\cdot)\big).\tag{1.3}$$

Any $\bar{u}(\cdot)$ satisfying (1.3) is called an optimal control. The corresponding state process $\bar{x}(\cdot)$ is called an optimal state process, and $\big(\bar{x}(\cdot),\bar{u}(\cdot)\big)$ is called an optimal pair.

The purpose of this paper is to establish first and second order necessary optimality conditions for Problem (OP). We refer to [2, 8, 16, 18, 21] and the references cited therein for some early works on the first order necessary optimality condition. Nevertheless, all of these works on the necessary conditions for optimal controls of infinite dimensional stochastic evolution equations addressed only some very special cases, as the case with the diffusion term not depending on the control variable, or when the control set U is convex, and so on. For the general case, stimulated by the work [17] addressing the same problem when $\dim H < \infty$, and because presently $\dim H = \infty$, one has to handle a difficult problem of the well-posedness of the second order adjoint equation, which is an operator-valued backward stochastic evolution equation (2.11) (see the next section). This problem was solved at almost the same time in [4, 6, 13], where the Pontryagin-type maximum principles were established for optimal controls of general infinite dimensional nonlinear stochastic systems. Nevertheless, the techniques used in [13] and [4, 6] are quite different. Indeed, since the most difficult part, i.e., the correction term "$Q_2(t)$" in (2.11) does not appear in the final formulation of the Pontryagin-type maximum principle, the strategy in both [4] and [6] is to ignore completely $Q_2(t)$ in the well-posedness result for (2.11). By contrast, [13] characterized the above mentioned $Q_2(t)$ because it was anticipated that this term should be useful somewhere else. More precisely, a new concept of solution, i.e., relaxed transposition solution was introduced in [13] to prove the well-posedness of (2.11).

On the other hand, to the best of our knowledge, the only works on second order necessary optimality condition for Problem (OP) are [9, 11], in which the control set U is supposed to be convex so that the classical convex variation technique can be employed. For the general stochastic optimal control problems in the finite dimensional framework, when nonconvex control regions are considered and spike variations are used as perturbations, as shown in [19, 20], to derive the second order necessary optimality conditions, the cost functional needs to be expanded up to the forth order and four adjoint equations have to be introduced. Consequently, an essential difficulty would arise in the present infinite dimensional situation. Indeed, the infinite dimensional counterparts of multilinear function-valued backward stochastic differential equations introduced in [19] are multilinear operator-valued backward stochastic evolution equations. So far the well-posedness of these sort of equations is completely unknown.

In this paper, in order to avoid the above mentioned difficulty with the well-posedness, as in our earlier work dealing with the finite dimensional framework [5], we use the classical variational analysis approach to establish the second order necessary optimality condition for Problem (OP) in the general setting. The main advantage of this approach is that only two adjoint equations (i.e., (2.6) and (2.11) from the next section) are needed to state the desired second order condition. It deserves mentioning that, though $Q_2(t)$ in (2.11) is useless in the first order optimality condition, it plays a key role in the statement of our second order necessary optimality condition (see the last two terms of (5.9) in Theorem 5.1). Of course, as in [5], the classical variational analysis approach can also be used to derive the first order necessary optimality condition for Problem (OP). Furthermore, we also show a "true" Pontryagin-type maximum principle for Problem (OP) under some mild assumptions.

The rest of this paper is organized as follows. In Section 2, some preliminaries on set-valued analysis and mild stochastic processes are collected. Section 3 is devoted to the first order necessary optimality condition for Problem (OP); while in Section 5, we shall establish the second order necessary optimality condition. In Section 4 we prove the maximum principle under a convexity assumption using again a variational technique.

2 Preliminaries

2.1 Some elements of set-valued analysis

For a nonempty subset $K \subset X$, the distance between a point $x \in X$ and K is defined by $dist\,(x, K) := \inf_{y \in K} |y - x|_X$. For any $\delta > 0$, write $B(x, \delta) = \{y \in X \mid |y - x|_X < \delta\}$.

Definition 2.1 *The adjacent cone $T_K^b(x)$ to K at x is defined by*

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

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

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

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

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

Let $G: \ X \rightsquigarrow Y$ be a set-valued map with nonempty values. We denote by $graph(G)$ the set of all $(x, y) \in X \times Y$ such that $y \in G(x)$. Recall that G is called Lipschitz around x if there exist $c \ge 0$ and $\delta > 0$ such that

$$G(x_1) \subset G(x_2) + B(0, c|x_1 - x_2|), \quad \forall\, x_1,\ x_2 \in B(x, \delta).$$

The adjacent directional derivative $D^\flat G(x, y)(v)$ of G at $(x, y) \in graph(G)$ in a direction $v \in X$ is a subset of Y defined by

$$w \in D^\flat G(x, y)(v) \Leftrightarrow (v, w) \in T^b_{graph(G)}(x, y).$$

If G is locally Lipschitz at x, then

$$D^\flat G(x, y)(v) = \mathrm{Liminf}_{\varepsilon \to 0+} \frac{G(x + \varepsilon v) - y}{\varepsilon},$$

where Liminf stands for the Peano-Kuratowski limit (see for instance [1]).

We shall need the following result.

Proposition 2.1 ([1, Proposition 5.2.6]) *Consider a set-valued map $G : X \rightsquigarrow Y$ with nonempty convex values and assume that G is Lipschitz around x. Then, for any $y \in G(x)$, the values of $D^\flat G(x, y)$ are convex and*

$$D^\flat G(x, y)(0) = T^b_{G(x)}(y) \supset G(x) - y,$$

$$D^\flat G(x, y)(v) + D^\flat G(x, y)(0) = D^\flat G(x, y)(v).$$

2.2 Vector-valued Brownian motions

In this subsection, we collect some preliminaries on vector-valued Brownian motions. We refer to [3, Section 4.1 of Chapter 4] and [15, Section 2.9 of Chapter 2] for more material on this subject (in [3], Brownian motions are called alternatively Wiener processes).

Let $Q \in \mathcal{L}(V)$ be a positive definite, trace-class operator on V. Then there is an orthonormal basis $\{e_j\}_{j=1}^{\infty}$ in V and a sequence $\{\lambda_j\}_{j=1}^{\infty}$ of positive numbers satisfying $\sum_{j=1}^{\infty} \lambda_j < \infty$ and $Qe_j = \lambda_j e_j$ for $j = 1, 2, \cdots$ (λ_j and e_j are respectively eigenvalue and eigenvector of Q). When V is an m-dimensional Hilbert space (for some $m \in \mathbb{N}$), the sequence $\{\lambda_j\}_{j=1}^{\infty}$ is reduced to $\{\lambda_j\}_{j=1}^{m}$.

We begin with the following notion of V-valued, standard Q-Brownians.

Definition 2.3 *A continuous, V-valued, $\mathbf{F}$-adapted process $W(\cdot) \equiv \{W(t)\}_{t\geq 0}$ is called a V-valued, standard Q-Brownian motion if*

1) $\mathbb{P}(\{W(0) = 0\}) = 1$*;*

2) *For any $0 \leq s < t < \infty$, the random variable $W(t) - W(s)$ is independent of $\mathcal{F}_s$, and $W(t) - W(s) \sim \mathcal{N}(0, (t-s)Q)$.*

In particular, $W(\cdot)$ is called a real valued, standard Brownian motion if $V = \mathbb{R}$.

Next, we consider another kind of vector-valued Brownian motions. Let $\{w_j(\cdot)\}_{j=1}^{\infty}$ be a sequence of independent, real valued, standard Brownian motions.

Definition 2.4 *The following formal series (in V)*

$$W(t) = \sum_{j=1}^{\infty} w_j(t) e_j, \quad t \geq 0, \tag{2.1}$$

is called a (V-valued) cylindrical Brownian motion.

Clearly, the series in (2.1) does not converge in V. Nevertheless, this series does converge in a larger space, as we shall see below.

Fix an arbitrary sequence $\{\mu_j\}_{j=1}^{\infty}$ of positive numbers such that $\sum_{j=1}^{\infty} \mu_j^2 < \infty$. Let V_1 be the completion of V w.r.t. the following norm:

$$|f|_{V_1} = \sqrt{\sum_{j=1}^{\infty} \mu_j^2 |\langle f, e_j\rangle_V|^2}, \qquad \forall\, f \in V.$$

Then, V_1 is a separable Hilbert space, $V \subset V_1$ and the embedding map $J:\ V \to V_1$ is a Hilbert-Schmidt operator. Let $Q_1 = JJ^*$. Then $Q_1 \in \mathcal{L}(V_1)$ is a positive definite, trace-class operator on V_1. Moreover, one can show that the following series

$$W(t) = \sum_{j=1}^{\infty} w_j(t) J e_j, \quad t \geq 0, \tag{2.2}$$

converges in $L^2_{\mathbb{F}}(\Omega; C([0,T]; V_1))$ for any $T > 0$ and defines a V_1-valued, standard Q_1-Brownian motion.

From now on, to simplify the presentation, we only consider the case of a cylindrical Brownian motion.

2.3 Notions of solutions to some stochastic evolution equations and backward stochastic evolution equations

In the rest of this paper, we shall denote by C a generic constant, depending on T, A, p and L, which may be different from one place to another.

First, we recall that $x(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega; H))$ is called a mild solution to the equation (1.1) if

$$x(t) = S(t)x_0 + \int_0^t S(t-s)a(s,x(s),u(s))ds + \int_0^t S(t-s)b(s,x(s),u(s))dW(s), \qquad \forall\, t \in [0,T].$$

Throughout this section, we assume the following:

(A1) *$a(\cdot,\cdot,\cdot,\cdot) : [0,T] \times H \times \widetilde{H} \times \Omega \to H$ and $b(\cdot,\cdot,\cdot,\cdot) : [0,T] \times H \times \widetilde{H} \times \Omega \to \mathcal{L}_2^0$ are two (vector-valued) functions such that*

i) *For any $(x,u) \in H \times \widetilde{H}$, the functions $a(\cdot,x,u,\cdot) : [0,T] \times \Omega \to H$ and $b(\cdot,x,u,\cdot) : [0,T] \times \Omega \to \mathcal{L}_2^0$ are $\mathbb{F}$-measurable;*

ii) *For any $(t,x) \in [0,T] \times H$, the functions $a(t,x,\cdot) : \widetilde{H} \to H$ and $b(t,x,\cdot) : \widetilde{H} \to \mathcal{L}_2^0$ are continuous a.s.;*

iii) *There exist a constant $L > 0$ and a nonnegative $\eta \in L^p(\Omega; L^1(0,T;\mathbb{R}))$ such that, for a.e. $(t,\omega) \in [0,T] \times \Omega$ and any $(x_1,x_2,u_1,u_2,u) \in H \times H \times \widetilde{H} \times \widetilde{H} \times \widetilde{H}$,*

$$\begin{cases} |a(t,x_1,u_1,\omega) - a(t,x_2,u_2,\omega)|_H + |b(t,x_1,u_1,\omega) - b(t,x_2,u_2,\omega)|_{\mathcal{L}_2^0} \\ \quad \leq L(|x_1 - x_2|_H + |u_1 - u_2|_{\widetilde{H}}), \\ |a(t,0,u,\omega)|_H + |b(t,0,u,\omega)|^2_{\mathcal{L}_2^0} \leq \eta(t,\omega). \end{cases}$$

Further, we impose the following assumptions on the cost functional:

(A2) *Suppose that $f(\cdot,\cdot,\cdot,\cdot) : [0,T] \times H \times \widetilde{H} \times \Omega \to \mathbb{R}$ and $g(\cdot,\cdot) : H \times \Omega \to \mathbb{R}$ are two functions such that*

i) *For any $(x,u) \in H \times \widetilde{H}$, the function $f(\cdot,x,u,\cdot) : [0,T] \times \Omega \to \mathbb{R}$ is $\mathbb{F}$-measurable and $g(x,\cdot) : \Omega \to \mathbb{R}$ is $\mathcal{F}_T$-measurable;*

ii) *For any $(t,x) \in [0,T] \times H$, the function $f(t,x,\cdot) : \widetilde{H} \to \mathbb{R}$ is continuous a.s.;*

iii) *For all $(t,x_1,x_2,u) \in [0,T] \times H \times H \times \widetilde{H}$,*

$$\begin{cases} |f(t,x_1,u) - f(t,x_2,u)|_H + |g(x_1) - g(x_2)|_H \leq L|x_1 - x_2|_H, \quad a.s., \\ |f(t,0,u)|_H + |g(0)|_H \leq L, \quad a.s. \end{cases}$$

It is easy to show the following result.

Lemma 2.1 *Assume (A1). Then, for any $x_0 \in H$ and $u \in L^p_{\mathbb{F}}(\Omega;\, L^2(0,T;\widetilde{H}))$, the state equation (1.1) admits a unique mild solution $x \in C_{\mathbb{F}}([0,T]; L^p(\Omega;H))$, and for any $t \in [0,T]$ the following estimate holds:*

$$\sup_{s\in[0,t]} \mathbb{E}|x(s,\cdot)|^p_H \leq C\mathbb{E}\Big[|x_0|^p_H + \Big(\int_0^t |a(s,0,u(s),\cdot)|_H ds\Big)^p + \Big(\int_0^t |b(s,0,u(s),\cdot)|^2_{\mathcal{L}_2^0} ds\Big)^{\frac{p}{2}}\Big]. \quad (2.3)$$

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

$$\sup_{s\in[0,t]} \mathbb{E}|x(s,\cdot) - \tilde{x}(s,\cdot)|^p_H \leq C\mathbb{E}\ \Big[|x_0 - \tilde{x}_0|^p_H + \Big(\int_0^t |u(s,\cdot) - \tilde{u}(s,\cdot)|^2_{\widetilde{H}} ds\Big)^{\frac{p}{2}}\Big]. \quad (2.4)$$

Assume that a, b, f are continuously differentiable with respect to x, u and g is continuously differentiable with respect to x.

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

$$\varphi[t] = \varphi(t, \bar{x}(t), \bar{u}(t), \omega), \quad \varphi_x[t] = \varphi_x(t, \bar{x}(t), \bar{u}(t), \omega), \quad \varphi_u[t] = \varphi_u(t, \bar{x}(t), \bar{u}(t), \omega) \tag{2.5}$$

and consider the following H-valued backward stochastic evolution equation[1]:

$$\begin{cases} dP_1(t) = -A^* P_1(t)dt - \big(a_x[t]^* P_1(t) + b_x[t]^* Q_1(t) - f_x[t]\big)dt + Q_1(t)dW(t), \quad t \in [0, T], \\ P_1(T) = -g_x(\bar{x}(T)). \end{cases} \tag{2.6}$$

Let us recall below the well-posedness result for the equation (2.6) in the transposition sense, developed in [12, 13, 14].

We consider the following (forward) stochastic evolution equation:

$$\begin{cases} dz = (Az + \psi_1(s))ds + \psi_2(s)dW(s) \quad \text{in } (t, T], \\ z(t) = \eta, \end{cases} \tag{2.7}$$

where $t \in [0, T]$, $\psi_1 \in L^1_{\mathbb{F}}(t, T; L^2(\Omega; H))$, $\psi_2 \in L^2_{\mathbb{F}}(t, T; L^2(\Omega; \mathcal{L}_2^0))$ and $\eta \in L^2_{\mathcal{F}_t}(\Omega; H)$. We call $(P_1(\cdot), Q_1(\cdot)) \in D_{\mathbb{F}}([0, T]; L^2(\Omega; H)) \times L^2_{\mathbb{F}}(0, T; \mathcal{L}_2^0))$ a transposition solution to (2.6) if for any $t \in [0, T]$, $\psi_1(\cdot) \in L^1_{\mathbb{F}}(t, T; L^2(\Omega; H))$, $\psi_2(\cdot) \in L^2_{\mathbb{F}}(t, T; L^2(\Omega; \mathcal{L}_2^0))$, $\eta \in L^2_{\mathcal{F}_t}(\Omega; H)$ and the corresponding mild solution $z \in C_{\mathbb{F}}([t, T]; L^2(\Omega; H))$ to (2.7), the following is satisfied

$$\begin{aligned} &\mathbb{E}\big\langle z(T), -g_x\big(\bar{x}(T)\big)\big\rangle_H + \mathbb{E}\int_t^T \big\langle z(s), a_x[s]^* P_1(s) + b_x[s]^* Q_1(s) - f_x[s]\big\rangle_H ds \\ &= \mathbb{E}\big\langle \eta, P_1(t)\big\rangle_H + \mathbb{E}\int_t^T \big\langle \psi_1(s), P_1(s)\big\rangle_H ds + \mathbb{E}\int_t^T \big\langle \psi_2(s), Q_1(s)\big\rangle_{\mathcal{L}_2^0} ds. \end{aligned} \tag{2.8}$$

We have the following result for the well-posedness of (2.6).

Lemma 2.2 ([13, 14]) *The equation (2.6) admits one and only one transposition solution* $(P_1(\cdot), Q_1(\cdot)) \in D_{\mathbb{F}}([0, T]; L^2(\Omega; H)) \times L^2_{\mathbb{F}}(0, T; L^2(\Omega; \mathcal{L}_2^0))$. *Furthermore,*

$$|(P_1(\cdot), Q_1(\cdot))|_{D_{\mathbb{F}}([0,T];L^2(\Omega;H))\times L^2_{\mathbb{F}}(0,T;\mathcal{L}_2^0))} \leq C\big(|g_x\big(\bar{x}(T)\big)|_{L^2_{\mathcal{F}_T}(\Omega;H)} + |f_x[\cdot]|_{L^1_{\mathbb{F}}(0,T;L^2(\Omega;H))}\big). \tag{2.9}$$

The proof of Lemma 2.2 is based on the following Riesz-type Representation Theorem established in [10].

Theorem 2.1 *Suppose* $1 \leq q < \infty$, *and that* X^* *has the Radon-Nikodým property. Then*

$$L^p_{\mathbb{F}}(0, T; L^q(\Omega; X))^* = L^{p'}_{\mathbb{F}}(0, T; L^{q'}(\Omega; X^*)).$$

Fix any $r_1, r_2, r_3, r_4 \in [1, \infty]$ and denote by $\mathcal{L}_{pd}\big(L^{r_1}_{\mathbb{F}}(0, T; L^{r_2}(\Omega; X)); L^{r_3}_{\mathbb{F}}(0, T; L^{r_4}(\Omega; Y))\big)$ the vector space of all bounded, pointwise defined linear operators $\mathcal{L}$ from $L^{r_1}_{\mathbb{F}}(0, T; L^{r_2}(\Omega; X))$ to $L^{r_3}_{\mathbb{F}}(0, T; L^{r_4}(\Omega; Y))$, i.e., for a.e. $(t, \omega) \in (0, T) \times \Omega$, there exists an $L(t, \omega) \in \mathcal{L}(X; Y)$ satisfying $\big(\mathcal{L}u(\cdot)\big)(t, \omega) = L(t, \omega)u(t, \omega)$, $\forall\, u(\cdot) \in L^{r_1}_{\mathbb{F}}(0, T; L^{r_2}(\Omega; X))$.

[1] Throughout this paper, for any operator-valued process (*resp.* random variable) R, we denote by R^* its pointwise dual operator-valued process (*resp.* random variable). In particular, if $R \in L^1_{\mathbb{F}}(0, T; L^2(\Omega; \mathcal{L}(H)))$, then $R^* \in L^1_{\mathbb{F}}(0, T; L^2(\Omega; \mathcal{L}(H)))$, and $|R|_{L^1_{\mathbb{F}}(0,T;L^2(\Omega;\mathcal{L}(H)))} = |R^*|_{L^1_{\mathbb{F}}(0,T;L^2(\Omega;\mathcal{L}(H)))}$.

Consider the Hamiltonian

$$\mathbb{H}(t,x,u,v,w,\omega) = \langle v, a(t,x,u,\omega) \rangle_H + \langle w, b(t,x,u,\omega) \rangle_{\mathcal{L}_2^0} - f(t,x,u,\omega), \tag{2.10}$$

where $(t,x,u,v,w,\omega) \in [0,T] \times H \times \widetilde{H} \times H \times \mathcal{L}_2^0 \times \Omega$. Assume next that a, b, f are twice continuously differentiable with respect to x, u with uniformly bounded second derivatives and that g is twice continuously differentiable with respect to x with uniformly bounded second derivatives.

We then consider the following $\mathcal{L}(H)$-valued backward stochastic evolution equation:

$$\begin{cases} dP_2(t) = -\Big(A^*P_2(t) + P_2(t)A + a_x[t]^*P_2(t) + P_2(t)a_x[t] + b_x[t]^*P_2(t)b_x[t] \\ \qquad\qquad + b_x[t]^*Q_2(t) + Q_2(t)b_x[t] + \mathbb{H}_{xx}[t]\Big)dt + Q_2(t)dW(t), \qquad t \in [0,T], \\ P_2(T) = -g_{xx}(\bar{x}(T)), \end{cases} \tag{2.11}$$

where $\mathbb{H}_{xx}[t] = \mathbb{H}_{xx}(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t))$ with $(P_1(\cdot), Q_1(\cdot))$ given by (2.6).

Let us introduce the following two (forward) stochastic evolution equations:

$$\begin{cases} dx_1 = (A + a_x[s])x_1 ds + u_1(s)ds + b_x[s]x_1 dW(s) + v_1(s)dW(s) \quad \text{in } (t,T], \\ x_1(t) = \xi_1 \end{cases} \tag{2.12}$$

and

$$\begin{cases} dx_2 = (A + a_x[s])x_2 ds + u_2(s)ds + b_x[s]x_2 dW(s) + v_2(s)dW(s) \quad \text{in } (t,T], \\ x_2(t) = \xi_2, \end{cases} \tag{2.13}$$

where $\xi_1, \xi_2 \in L^{2p'}_{\mathcal{F}_t}(\Omega; H)$, $u_1, u_2 \in L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; H))$, and $v_1, v_2 \in L^2_{\mathbb{F}}(t,T;\, L^{2p'}(\Omega; \mathcal{L}_2^0))$. Also, we need to define the solution space for (2.11). For this purpose, for $p > 1$, write

$$\begin{aligned} & L^p_{\mathbb{F},w}(\Omega; D([0,T]; \mathcal{L}(H)) \\ & \stackrel{\triangle}{=} \Big\{ P(\cdot,\cdot) \;\Big|\; P(\cdot,\cdot) \in \mathcal{L}_{pd}\big(L^2_{\mathbb{F}}(0,T; L^{2p'}(\Omega; H));\; L^2_{\mathbb{F}}(0,T; L^{\frac{2p}{p+1}}(\Omega; H))\big), \\ & \quad P(\cdot,\cdot)\xi \in D_{\mathbb{F}}([t,T]; L^{\frac{2p}{p+1}}(\Omega; H)) \text{ and} \\ & \quad |P(\cdot,\cdot)\xi|_{D_{\mathbb{F}}([t,T]; L^{\frac{2p}{p+1}}(\Omega;H))} \le C|\xi|_{L^{2p'}_{\mathcal{F}_t}(\Omega;H)} \text{ for every } t \in [0,T] \text{ and } \xi \in L^{2p'}_{\mathcal{F}_t}(\Omega; H) \Big\}, \end{aligned} \tag{2.14}$$

$$\mathcal{H}_t \stackrel{\triangle}{=} L^{2p'}_{\mathcal{F}_t}(\Omega; H) \times L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; H)) \times L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; \mathcal{L}_2^0)), \quad \forall\, t \in [0,T),$$

and[2]

$$\begin{aligned} \mathcal{Q}^p[0,T] \stackrel{\triangle}{=} \Big\{ \big(Q^{(\cdot)}, \widehat{Q}^{(\cdot)}\big) \;\Big|\; & Q^{(t)}, \widehat{Q}^{(t)} \in \mathcal{L}\big(\mathcal{H}_t;\; L^2_{\mathbb{F}}(t,T; L^{\frac{2p}{p+1}}(\Omega; \mathcal{L}_2^0))\big) \\ & \text{and } Q^{(t)}(0,0,\cdot)^* = \widehat{Q}^{(t)}(0,0,\cdot) \text{ for any } t \in [0,T) \Big\}. \end{aligned} \tag{2.15}$$

The notion of relaxed transposition solution to (2.11) (cf. [13, 14]): We call $\big(P_2(\cdot),\ Q_2^{(\cdot)},\ \widehat{Q}_2^{(\cdot)}\big) \in L^p_{\mathbb{F},w}(\Omega;\ D([0,T];\ \mathcal{L}(H))) \times \mathcal{Q}^p[0,T]$ a relaxed transposition solution to the equation

[2] By Theorem 2.1, since $\mathcal{L}_2^0$ is a Hilbert space, we deduce that $Q^{(t)}(0,0,\cdot)^*$ is a bounded linear operator from $L^2_{\mathbb{F}}(t,T;\ L^{\frac{2p}{p+1}}(\Omega; \mathcal{L}_2^0))^* = L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; \mathcal{L}_2^0))$ to $L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; \mathcal{L}_2^0))^* = L^2_{\mathbb{F}}(t,T; L^{\frac{2p}{p+1}}(\Omega; \mathcal{L}_2^0))$. Hence, $Q^{(t)}(0,0,\cdot)^* = \widehat{Q}^{(t)}(0,0,\cdot)$ makes sense.

(2.11) if for any $t \in [0,T]$, $\xi_1, \xi_2 \in L^{2p'}_{\mathcal{F}_t}(\Omega; H)$, $u_1(\cdot), u_2(\cdot) \in L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; H))$ and $v_1(\cdot), v_2(\cdot) \in L^2_{\mathbb{F}}(t,T; L^{2p'}(\Omega; \mathcal{L}_2^0))$, the following is satisfied

$$\begin{aligned}
&\mathbb{E}\big\langle -g_{xx}(\bar{x}(T))x_1(T), x_2(T)\big\rangle_H + \mathbb{E}\int_t^T \big\langle \mathbb{H}_{xx}[s]x_1(s), x_2(s)\big\rangle_H ds \\
&= \mathbb{E}\big\langle P_2(t)\xi_1, \xi_2\big\rangle_H + \mathbb{E}\int_t^T \big\langle P_2(s)u_1(s), x_2(s)\big\rangle_H ds \\
&\quad + \mathbb{E}\int_t^T \big\langle P_2(s)x_1(s), u_2(s)\big\rangle_H ds + \mathbb{E}\int_t^T \big\langle P_2(s)b_x(s)x_1(s), v_2(s)\big\rangle_{\mathcal{L}_2^0} ds \\
&\quad + \mathbb{E}\int_t^T \big\langle P_2(s)v_1(s), b_x(s)x_2(s) + v_2(s)\big\rangle_{\mathcal{L}_2^0} ds + \mathbb{E}\int_t^T \big\langle v_1(s), \widehat{Q}_2^{(t)}(\xi_2, u_2, v_2)(s)\big\rangle_{\mathcal{L}_2^0} ds \\
&\quad + \mathbb{E}\int_t^T \big\langle Q_2^{(t)}(\xi_1, u_1, v_1)(s), v_2(s)\big\rangle_{\mathcal{L}_2^0} ds,
\end{aligned} \tag{2.16}$$

where $x_1(\cdot)$ and $x_2(\cdot)$ solve respectively (2.12) and (2.13).

We have the following well-posedness result for the equation (2.11) in the sense of relaxed transposition solution.

Theorem 2.2 ([13, 14]) *Assume that $p \in (1,2]$ and the Banach space $L^p_{\mathcal{F}_T}(\Omega; \mathbb{R})$ is separable. Then, the equation (2.11) admits one and only one relaxed transposition solution $\big(P_2(\cdot), Q_2^{(\cdot)}, \widehat{Q}_2^{(\cdot)}\big) \in L^p_{\mathbb{F},w}(\Omega;\ D([0,T]; \mathcal{L}(H))) \times \mathcal{Q}^p[0,T]$. Furthermore,*

$$\begin{aligned}
&|P_2|_{\mathcal{L}(L^2_{\mathbb{F}}(0,T;L^{2p'}(\Omega;H));\ L^2(0,T;L^{2p/(p+1)}_{\mathbb{F}}(\Omega;H)))} + \sup_{t\in[0,T)} \big|\big(Q_2^{(t)}, \widehat{Q}_2^{(t)}\big)\big|_{\mathcal{L}(\mathcal{H}_t;\ L^2_{\mathbb{F}}(t,T;L^{2p/(p+1)}(\Omega;\mathcal{L}_2^0)))^2} \\
&\le C\big(|\mathbb{H}_{xx}[\cdot]|_{L^1_{\mathbb{F}}(0,T;\ L^p(\Omega;\mathcal{L}(H)))} + |g_{xx}(\bar{x}(T))|_{L^p_{\mathcal{F}_T}(\Omega;\ \mathcal{L}(H))}\big).
\end{aligned} \tag{2.17}$$

3 First order necessary optimality condition

We impose the following further assumptions on $a(\cdot,\cdot,\cdot,\cdot)$, $b(\cdot,\cdot,\cdot,\cdot)$, $f(\cdot,\cdot,\cdot,\cdot)$ and $g(\cdot,\cdot)$.

(A3) *The functions $a(\cdot,\cdot,\cdot,\cdot)$, $b(\cdot,\cdot,\cdot,\cdot)$ and $f(\cdot,\cdot,\cdot,\cdot)$ are C^1 with respect to the second and third variables, while $g(\cdot,\cdot)$ is C^1 with respect to the first variable.*

Note that (A1)–(A3) imply that for a.e. $(t,\omega) \in [0,T] \times \Omega$ and any $(x,u) \in H \times U$,

$$\begin{cases}
|a_x(t,x,u,\omega)|_{\mathcal{L}(H)} + |b_x(t,x,u,\omega)|_{\mathcal{L}(H;\mathcal{L}_2^0)} + |f_x(t,x,u,\omega)|_H + |g_x(x,\omega)|_H \le L, \\
|a_u(t,x,u,\omega)|_{\mathcal{L}(\widetilde{H};H)} + |b_u(t,x,u,\omega)|_{\mathcal{L}(\widetilde{H};\mathcal{L}_2^0)} + |f_u(t,x,u,\omega)|_{\widetilde{H}} \le L.
\end{cases} \tag{3.1}$$

Below, to simplify the notations, sometimes we will write "a.e. $t \in [0,T]$, $a.s.$" instead of "for a.e. $(t,\omega) \in [0,T] \times \Omega$".

Now, let us introduce the classical first order variational control system. Let $\bar{u}, v, v_\varepsilon \in L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\widetilde{H}))$ ($\beta \ge 1$) satisfy $v_\varepsilon \to v$ in $L^\beta_{\mathbb{F}}(\Omega; L^2(0,T;\widetilde{H}))$ as $\varepsilon \to 0^+$. For $u^\varepsilon := \bar{u} + \varepsilon v_\varepsilon$, let x^ε be the state of (1.1) corresponding to the control u^ε, and put

$$\delta x^\varepsilon = x^\varepsilon - \bar{x}.$$

Consider the following linearized stochastic control system (recall (2.5) for the notations $a_x[\cdot]$, $a_u[\cdot]$, $b_x[\cdot]$ and $b_u[\cdot]$):

$$\begin{cases} dy_1(t) = \big(Ay_1(t) + a_x[t]y_1(t) + a_u[t]v(t)\big)dt + \big(b_x[t]y_1(t) + b_u[t]v(t)\big)dW(t), & t \in [0,T], \\ y_1(0) = 0. \end{cases} \tag{3.2}$$

The system (2.6) is the first order adjoint equation associated with (3.2) and the cost function f. Similarly to [5], it is easy to establish the following estimates.

Lemma 3.1 *If (A1) and (A3) hold, and $\beta \geq 2$, then, for any $\bar{u}, v, v_\varepsilon$ and δx^ε as above*

$$|y_1|_{L^\infty_{\mathbb{F}}(0,T;L^\beta(\Omega;H))} \leq C|v|_{L^\beta_{\mathbb{F}}(\Omega;L^2(0,T;\widetilde{H}))}, \quad |\delta x^\varepsilon|_{L^\infty_{\mathbb{F}}(0,T;L^\beta(\Omega;H))} = O(\varepsilon), \quad \text{as } \varepsilon \to 0^+.$$

Furthermore,

$$|r_1^\varepsilon|_{L^\infty_{\mathbb{F}}(0,T;L^\beta(\Omega;H))} \to 0, \quad \text{as } \varepsilon \to 0^+, \tag{3.3}$$

where

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

Further, we have the following result.

Theorem 3.1 *Let the assumptions (A1) with $p = 2$, (A2) and (A3) hold and let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be an optimal pair of Problem (OP). Consider the transposition solution $(P_1(\cdot), Q_1(\cdot))$ to (2.6). Then,*

$$\mathbb{E}\int_0^T \langle\, \mathbb{H}_u[t], v(t)\,\rangle_{\widetilde{H}} dt \leq 0, \qquad \forall\, v \in T^b_{\mathcal{U}_{ad}}(\bar{u}), \tag{3.4}$$

where $\mathbb{H}_u[t] = \mathbb{H}_u(t, \bar{x}(t), \bar{u}(t), P_1(t), Q_1(t))$.

Proof. Let $v \in T^b_{\mathcal{U}_{ad}}(\bar{u}(\cdot))$. Then, for any $\varepsilon > 0$, there exists a $v_\varepsilon \in L^2_{\mathbb{F}}(\Omega; L^2(0,T;\widetilde{H}))$ such that $\bar{u} + \varepsilon v_\varepsilon \in \mathcal{U}_{ad}$ and

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

Expanding the cost functional $J(\cdot)$ at $\bar{u}(\cdot)$, we have

$$\begin{aligned} 0 \;&\leq\; \frac{J(u^\varepsilon(\cdot)) - J(\bar{u}(\cdot))}{\varepsilon} \\ &=\; \mathbb{E}\int_0^T \Big(\int_0^1 \langle\, f_x(t, \bar{x}(t) + \theta\delta x^\varepsilon(t), \bar{u}(t) + \varepsilon v_\varepsilon(t)), \frac{\delta x^\varepsilon(t)}{\varepsilon} \rangle_H d\theta \\ &\qquad\quad + \int_0^1 \langle\, f_u(t, \bar{x}(t), \bar{u}(t) + \theta\varepsilon v_\varepsilon(t)), v_\varepsilon(t) \rangle_{\widetilde{H}} d\theta \Big) dt \\ &\quad + \mathbb{E}\int_0^1 \langle\, g_x(\bar{x}(T) + \theta\delta x^\varepsilon(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} \rangle_H d\theta \\ &=\; \mathbb{E}\int_0^T \big(\langle\, f_x[t], y_1(t) \rangle_H + \langle\, f_u[t], v(t)\,\rangle_{\widetilde{H}} \big) dt + \mathbb{E}\langle\, g_x(\bar{x}(T)), y_1(T)\,\rangle_H + \rho_1^\varepsilon, \end{aligned} \tag{3.5}$$

where

$$\begin{aligned}
\rho_1^\varepsilon \;\; = \;\; & \mathbb{E}\int_0^T \Big(\int_0^1 \langle\, f_x(t,\bar{x}(t)+\theta\delta x^\varepsilon(t),\bar{u}(t)+\varepsilon v_\varepsilon(t)) - f_x[t], \frac{\delta x^\varepsilon(t)}{\varepsilon}\rangle_H d\theta \\
& \qquad + \int_0^1 \langle\, f_u(t,\bar{x}(t),\bar{u}(t)+\theta\varepsilon v_\varepsilon(t)) - f_u[t], v_\varepsilon(t)\,\rangle_{\widetilde{H}} d\theta \\
& \qquad + \langle\, f_x[t], \frac{\delta x^\varepsilon(t)}{\varepsilon} - y_1(t)\rangle_H + \langle\, f_u[t], v_\varepsilon(t) - v(t)\,\rangle_{\widetilde{H}}\Big) dt \\
& + \mathbb{E}\int_0^1 \langle\, g_x(\bar{x}(T)+\theta\delta x^\varepsilon(T)) - g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon}\rangle_H d\theta \\
& \qquad + \mathbb{E}\,\langle\, g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} - y_1(T)\rangle_H.
\end{aligned} \tag{3.6}$$

By Lemma 3.1 (applied with $\beta = 2$) and (A2)–(A3), using the Dominated Convergence Theorem, we conclude that

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

Similarly, we have

$$\mathbb{E}\int_0^T\int_0^1 \langle\, f_u(t,\bar{x}(t),\bar{u}(t)+\theta\varepsilon v_\varepsilon(t)) - f_u[t], v_\varepsilon(t)\,\rangle_{\widetilde{H}} d\theta dt \to 0, \qquad \text{as } \varepsilon \to 0^+,$$

and

$$\mathbb{E}\int_0^1 \langle\, g_x(\bar{x}(T)+\theta\delta x^\varepsilon(T)) - g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon}\rangle_H d\theta \to 0, \qquad \text{as } \varepsilon \to 0^+.$$

Then, by (A2)–(A3) and Lemma 3.1, we obtain that

$$\begin{aligned}
\lim_{\varepsilon\to 0^+} \big|\rho_1^\varepsilon\big| \;\; \leq \;\; & \overline{\lim_{\varepsilon\to 0^+}} \Big|\mathbb{E}\int_0^T \langle\, f_x[t], \frac{\delta x^\varepsilon(t)}{\varepsilon} - y_1(t)\rangle_H dt\Big| \\
& + \overline{\lim_{\varepsilon\to 0^+}} \Big|\mathbb{E}\int_0^T \langle\, f_u[t], v_\varepsilon(t) - v(t)\,\rangle_{\widetilde{H}} dt\Big| \\
& + \overline{\lim_{\varepsilon\to 0^+}} \Big|\mathbb{E}\,\langle\, g_x(\bar{x}(T)), \frac{\delta x^\varepsilon(T)}{\varepsilon} - y_1(T)\rangle_H\Big| = 0.
\end{aligned} \tag{3.7}$$

Therefore, from (3.5) and (3.7), we conclude that

$$0 \;\; \leq \;\; \mathbb{E}\int_0^T \big(\langle\, f_x[t], y_1(t)\,\rangle_H + \langle\, f_u[t], v(t)\,\rangle_{\widetilde{H}}\big) dt + \mathbb{E}\langle\, g_x(\bar{x}(T)), y_1(T)\,\rangle_H. \tag{3.8}$$

By means of the definition of transposition solution to (2.6), we have

$$
\begin{aligned}
&-\mathbb{E}\langle\, g_x(\bar{x}(T)), y_1(T)\,\rangle_H = \mathbb{E}\langle\, P_1(T), y_1(T)\,\rangle_H\\
&= \mathbb{E}\int_0^T \big(\langle\, P_1(t), a_x[t]y_1(t)\,\rangle_H + \langle\, P_1(t), a_u[t]v(t)\,\rangle_H\\
&\qquad +\langle\, Q_1(t), b_x[t]y_1(t)\,\rangle_{\mathcal{L}_2^0} + \langle\, Q_1(t), b_u[t]v(t)\,\rangle_{\mathcal{L}_2^0}\\
&\qquad -\langle\, a_x[t]^*P_1(t), y_1(t)\,\rangle_H - \langle\, b_x[t]^*Q_1(t), y_1(t)\,\rangle_H + \langle\, f_x[t], y_1(t)\,\rangle_H\big)dt\\
&= \mathbb{E}\int_0^T \big(\langle\, P_1(t), a_u[t]v(t)\,\rangle_H + \langle\, Q_1(t), b_u[t]v(t)\,\rangle_{\mathcal{L}_2^0} + \langle\, f_x[t], y_1(t)\,\rangle_H\big)dt.
\end{aligned}
\tag{3.9}
$$

Substituting (3.9) in (3.8), and recalling (2.10), we obtain that

$$
\begin{aligned}
0 &\le -\mathbb{E}\int_0^T \big(\langle\, P_1(t), a_u[t]v(t)\,\rangle_H + \langle\, Q_1(t), b_u[t]v(t)\,\rangle_{\mathcal{L}_2^0} - \langle\, f_u[t], v(t)\,\rangle_{\widetilde{H}}\big)dt\\
&= -\mathbb{E}\int_0^T \langle\, \mathbb{H}_u[t], v(t)\,\rangle_{\widetilde{H}}dt,
\end{aligned}
\tag{3.10}
$$

which gives (3.4). This completes the proof of Theorem 3.1. □

As a consequence of Theorem 3.1, one has the following pointwise first order necessary condition for optimal pairs of Problem (OP) that follows easily from measurable selection theorems.

Theorem 3.2 *Under the assumptions in Theorem 3.1, it holds that*

$$
\mathbb{H}_u[t] \in N_U^b(\bar{u}(t)),\ \text{a.e. } t\in[0,T],\ \text{a.s.} \tag{3.11}
$$

4 Maximum Principle

In this section, we address the Pontryagin-type maximum principle for Problem (OP). For this purpose, we introduce the following assumption.

(A4) *For a.e.* $(t,\omega)\in[0,T]\times\Omega$ *and for all* $x\in H$*, the set*

$$
F(t,x,\omega) := \big\{(a(t,x,u,\omega), b(t,x,u,\omega), f(t,x,u,\omega)+r)\ \big|\ u\in U,\ r\ge 0\big\}
$$

is closed and convex in $H\times H\times\mathbb{R}$.

The above assumption is familiar in the deterministic optimal control where it is very useful to guarantee the existence of optimal controls. In particular, it holds true whenever a, b are affine in the control, U is convex and compact and f is convex with respect to u.

Also, we impose the following assumption, which is weaker than (A1) because Lipschitz continuity with respect to u is no longer required.

(A5) *Suppose that* $a : [0,T]\times H\times U\times\Omega \to H$ *and* $b : [0,T]\times H\times U\times\Omega\to\mathcal{L}_2^0$ *are two (vector-valued) functions satisfying the conditions* i) *and* ii) *in* (A1)*, and such that for some* $L\ge 0$*,* $\eta\in L^2([0,T]\times\Omega;\mathbb{R})$ *and for a.e.* $(t,\omega)\in[0,T]\times\Omega$ *and any* $(x_1,x_2,u)\in H\times H\times U$*,*

$$
\begin{cases}
|a(t,x_1,u,\omega)-a(t,x_2,u,\omega)|_H + |b(t,x_1,u,\omega)-b(t,x_2,u,\omega)|_{\mathcal{L}_2^0}\le L|x_1-x_2|_H,\\
|a(t,0,u,\omega)|_H + |b(t,0,u,\omega)|^2_{\mathcal{L}_2^0}\le\eta(t,\omega).
\end{cases}
$$

Under the assumption (A5), the system (1.1) is still well-posed.

We have the following Pontryagin-type maximum principle for Problem (OP) (recall (2.10) for the definition of $\mathbb{H}(\cdot)$):

Theorem 4.1 *Assume (A2), (A4) and (A5) and that the functions $a(\cdot,\cdot,\cdot,\cdot)$, $b(\cdot,\cdot,\cdot,\cdot)$ and $f(\cdot,\cdot,\cdot,\cdot)$ are C^1 with respect to the second variable, while $g(\cdot,\cdot)$ is C^1 with respect to the first variable. If $(\bar{x},\bar{u})$ is an optimal pair for Problem (OP), then (P_1,Q_1) defined in Lemma 2.2 verifies the maximality condition*

$$\max_{u\in U}\mathbb{H}(t,\bar{x}(t),u,P_1(t),Q_1(t))=\mathbb{H}(t,\bar{x}(t),\bar{u}(t),P_1(t),Q_1(t)),\quad \text{for a.e. } t\in[0,T]\ a.s.$$

Proof. Fix $u(\cdot)\in\mathcal{U}_{ad}$ and consider the following linearized system:

$$\begin{cases} d\tilde{y}_1(t)=\big(A\tilde{y}_1(t)+a_x[t]\tilde{y}_1(t)+\alpha(t)\big)dt+\big(b_x[t]\tilde{y}_1(t)+\beta(t)\big)dW(t), & t\in[0,T],\\ \tilde{y}_1(0)=0, \end{cases}\tag{4.1}$$

where

$$\alpha(t)=a(t,\bar{x}(t),u(t))-a[t],\qquad \beta(t)=b(t,\bar{x}(t),u(t))-b[t].$$

Then

$$|\alpha(t,\omega)|_H+|\beta(t,\omega)|_{\mathcal{L}_2^0}\le C(|\bar{x}(t,\omega)|+|\eta(t,\omega)|+1).$$

Denote by $D_x^\flat F$ the adjacent derivative of F with respect to x. One can easily check that for a.e. $t\in[0,T]$,

$$(a_x[t]\tilde{y}_1(t),b_x[t]\tilde{y}_1(t),f_x[t]\tilde{y}_1(t))\in D_x^\flat F(t,\bar{x}(t),a[t],b[t],f[t])(\tilde{y}_1(t)),\quad a.s.$$

Define

$$\gamma(t)=f(t,\bar{x}(t),u(t))-f[t].$$

By Proposition 2.1, we have

$$\big(a_x[t]\tilde{y}_1(t)+\alpha(t),b_x[t]\tilde{y}_1(t)+\beta(t),f_x[t]\tilde{y}_1(t)+\gamma(t)\big)\in D_x^\flat F(t,\bar{x}(t),a[t],b[t],f[t])(\tilde{y}_1(t)).$$

Recall that $\mathbb{F}$ stands for the progressive σ-field (in $[0,T]\times\Omega$) with respect to $\mathbf{F}$. Let us consider the measure space $([0,T]\times\Omega,\mathbb{F},dt\times d\mathbb{P})$ and its completion $([0,T]\times\Omega,\mathcal{A},\mu)$.

Fix any sequence $\varepsilon_i\downarrow 0$ and define for every $j\ge 1$ the sets

$$\begin{aligned}\widehat{H}_j^i(t,\omega)=\big\{(u,r)\in U\times\mathbb{R}_+\ \big|\ & |a(t,\bar{x}(t,\omega)+\varepsilon_i\tilde{y}_1(t,\omega),u,\omega)-(a[t]+\varepsilon_i(a_x[t]\tilde{y}_1(t,\omega)+\alpha(t,\omega)))|_H\le\varepsilon_i/2^j\\ & |b(t,\bar{x}(t,\omega)+\varepsilon_i\tilde{y}_1(t,\omega),u,\omega)-(b[t]+\varepsilon_i(b_x[t]\tilde{y}_1(t,\omega)+\beta(t,\omega)))|_{\mathcal{L}_2^0}\le\varepsilon_i/2^j,\\ & |f(t,\bar{x}(t,\omega)+\varepsilon_i\tilde{y}_1(t,\omega),u,\omega)+r-(f[t]+\varepsilon_i(f_x[t]\tilde{y}_1(t,\omega)+\gamma(t,\omega)))|\le\varepsilon_i/2^j\big\}\end{aligned}$$

and

$$H_j^i(t,\omega)=\begin{cases}\widehat{H}_j^i(t,\omega) & \text{if } \widehat{H}_j^i(t,\omega)\neq\emptyset\\ \{(\bar{u}(t,\omega),0)\} & \text{otherwise.}\end{cases}$$

By the definition of set-valued derivative, we can find a subsequence i_j and a decreasing family of measurable sets $A_j\subset[0,T]\times\Omega$, such that $\lim_{j\to\infty}\mu(A_j)=0$ and $\widehat{H}_j^{i_j}(t,\omega)\neq\emptyset$ for $(t,\omega)\in([0,T]\times\Omega)\setminus A_j$.

By [1, Theorem 8.2.9], $H_j^{i_j}$ is μ-measurable. Since it also has closed nonempty values, it admits a measurable selection. Modifying this selection on a set of measure zero, we obtain $u_j\in\mathcal{U}_{ad}$

and an $\mathbf{F}$-adapted, real-valued process r_j satisfying for a.e. (t,ω) with $\widehat{H}_j^{i_j}(t,\omega) \neq \emptyset$ the following inequalities:

$$
\begin{aligned}
&|a(t,\bar{x}(t,\omega)+\varepsilon_{i_j}\tilde{y}_1(t,\omega),u_j(t,\omega),\omega)-(a[t]+\varepsilon_{i_j}(a_x[t]\tilde{y}_1(t,\omega)+\alpha(t,\omega)))|_H \le \varepsilon_{i_j}/2^j,\\
&|b(t,\bar{x}(t,\omega)+\varepsilon_{i_j}\tilde{y}_1(t,\omega),u_j(t,\omega),\omega)-(b[t]+\varepsilon_{i_j}(b_x[t]\tilde{y}_1(t,\omega)+\beta(t,\omega)))|_{\mathcal{L}_2^0} \le \varepsilon_{i_j}/2^j,\\
&|f(t,\bar{x}(t,\omega)+\varepsilon_{i_j}\tilde{y}_1(t,\omega),u_j(t,\omega),\omega)+r_j(t,\omega)-(f[t]+\varepsilon_{i_j}(f_x[t]\tilde{y}_1(t,\omega)+\gamma(t,\omega)))| \le \varepsilon_{i_j}/2^j.
\end{aligned}
$$

Define

$$
\begin{aligned}
&a_j(t)=a(t,\bar{x}(t)+\varepsilon_{i_j}\tilde{y}_1(t),u_j(t)), \qquad b_j(t)=b(t,\bar{x}(t)+\varepsilon_{i_j}\tilde{y}_1(t),u_j(t)),\\
&f_j(t)=f(t,\bar{x}(t)+\varepsilon_{i_j}\tilde{y}_1(t),u_j(t)).
\end{aligned}
$$

Observe that if $\widehat{H}_j^{i_j}(t,\omega)=\emptyset$, then by (A2), (A5), it follows that

$$
\begin{aligned}
&|a_j(t,\omega)-(a[t,\omega]+\varepsilon_{i_j}(a_x[t,\omega]\tilde{y}_1(t,\omega)+\alpha(t,\omega)))|_H \le C\varepsilon_{i_j}(|\tilde{y}_1(t,\omega)|+|\bar{x}(t,\omega)|+|\eta(t,\omega)|),\\
&|b_j(t,\omega)-(b[t,\omega]+\varepsilon_{i_j}(b_x[t,\omega]\tilde{y}_1(t,\omega)+\beta(t,\omega)))|_{\mathcal{L}_2^0} \le C\varepsilon_{i_j}(|\tilde{y}_1(t,\omega)|+|\bar{x}(t,\omega)|+|\eta(t,\omega)|),\\
&|f_j(t,\omega)-(f[t,\omega]+\varepsilon_{i_j}(f_x[t,\omega]\tilde{y}_1(t,\omega)+\gamma(t,\omega)))| \le C\varepsilon_{i_j}(|\tilde{y}_1(t,\omega)|+|\bar{x}(t,\omega)|+L).
\end{aligned}
$$

Consider the solution x_j of the following stochastic equation:

$$
\begin{cases}
dx_j(t)=\big(Ax_j(t)+a(t,x_j(t),u_j(t))\big)dt+b(t,x_j(t),u_j(t))dW(t) & \text{in } (0,T],\\
x_j(0)=x_0.
\end{cases}
\tag{4.2}
$$

Then, by (4.2) we have

$$
\begin{aligned}
&\mathbb{E}|x_j(t)-\bar{x}(t)-\varepsilon_{i_j}\tilde{y}_1(t)|_H^2\\
&\le 2\mathbb{E}\left|\int_0^t S(t-s)(a(s,x_j(s),u_j(s))-a[s]-\varepsilon_{i_j}(a_x[s]\tilde{y}_1(s)+\alpha(s)))ds\right|_H^2\\
&\quad+2\mathbb{E}\left|\int_0^t S(t-s)(b(s,x_j(s),u_j(s))-b[s]-\varepsilon_{i_j}(b_x[s]\tilde{y}_1(s)+\beta(s)))dW(s)\right|_H^2\\
&\le C\mathbb{E}\int_0^t |a_j(s)-a[s]-\varepsilon_{i_j}(a_x[s]\tilde{y}_1(s)+\alpha(s))|_H^2 ds\\
&\quad+C\mathbb{E}\int_0^t |x_j(s)-\bar{x}(s)-\varepsilon_{i_j}\tilde{y}_1(s)|_H^2 ds\\
&\quad+C\mathbb{E}\int_0^t |b(s,x_j(s),u_j(s))-b[s]-\varepsilon_{i_j}(b_x[s]\tilde{y}_1(s)+\beta(s))|_{\mathcal{L}_2^0}^2 ds)\\
&\le C\varepsilon_{i_j}^2 2^{-j}+C\varepsilon_{i_j}^2\int_{A_j}\left(|\tilde{y}_1(s)|_H^2+|\bar{x}(s)|_H^2+|\eta(s)|^2\right)ds\,d\mathbb{P}\\
&\quad+C\int_0^t \mathbb{E}|x_j(s)-\bar{x}(s)-\varepsilon_{i_j}\tilde{y}_1(s)|_H^2 ds.
\end{aligned}
\tag{4.3}
$$

This and the Gronwall inequality imply that for a constant $C>0$, a sequence $\delta_j \downarrow 0$ and for all j

$$
\mathbb{E}|x_j(t)-\bar{x}(t)-\varepsilon_{i_j}\tilde{y}_1(t)|_H^2 \le C\varepsilon_{i_j}^2(2^{-j}+\delta_j).
$$

Consequently

$$
\lim_{j\to\infty}\sup_{t\in[0,T]}\mathbb{E}\left|\frac{x_j(t)-\bar{x}(t)}{\varepsilon_{i_j}}-\tilde{y}_1(t)\right|_H^2=0.
$$

On the other hand,

$$
\begin{aligned}
&\mathbb{E}\left|\int_0^T (f(s,x_j(s),u_j(s)) + r_j(s) - f[s] - \varepsilon_{i_j}(f_x[s]\tilde{y}_1(s) + \gamma(s)))ds\right| \\
&\leq \mathbb{E}\int_0^T \big|f_j(s) + r_j(s) - f[s] - \varepsilon_{i_j}(f_x[s]\tilde{y}_1(s) + \gamma(s))\big|\, ds + C\mathbb{E}\int_0^T \big|\bar{x}(s) + \varepsilon_{i_j}\tilde{y}_1(s) - x_j(s)\big|_H\, ds.
\end{aligned}
\tag{4.4}
$$

Thus, by the similar arguments as above,

$$
\lim_{j\to\infty}\frac{1}{\varepsilon_{i_j}}\mathbb{E}\left|\int_0^T (f(s,x_j(s),u_j(s)) + r_j(s) - f[s] - \varepsilon_{i_j}(f_x[s]\tilde{y}_1(s) + \gamma(s)))ds\right| = 0.
$$

By the optimality of $(\bar{x},\bar{u})$, we arrive at

$$
\begin{aligned}
0 &\leq \mathbb{E}\left[\int_0^T (f(s,x_j(s),u_j(s)) + r_j(s) - f[s])\, ds + g(x_j(T)) - g(\bar{x}(T))\right] \\
&\leq \varepsilon_{i_j}\mathbb{E}\int_0^T (\langle\, f_x[s], \tilde{y}_1(s)\,\rangle_H + \gamma(s))ds \\
&\quad +\mathbb{E}\left|\int_0^T \big(f(s,x_j(s),u_j(s)) + r_j(s) - f[s] - \varepsilon_{i_j}(\langle\, f_x[s], \tilde{y}_1(s)\,\rangle_H + \gamma(s))\big)\, ds\right| \\
&\quad +\mathbb{E}\big(g(x_j(T)) - g(\bar{x}(T))\big).
\end{aligned}
\tag{4.5}
$$

Dividing by ε_{i_j} and taking the limit yields

$$
0 \;\leq\; \mathbb{E}\int_0^T (\langle\, f_x[t], \tilde{y}_1(t)\,\rangle_H + \gamma(t))dt + \mathbb{E}\langle\, g_x(\bar{x}(T)), \tilde{y}_1(T)\,\rangle_H. \tag{4.6}
$$

By means of the definition of transposition solution to (2.6), in the same way as that in (3.9) we finally obtain that

$$
\begin{aligned}
&\mathbb{E}\int_0^T \big(\langle\, P_1(t), a(t,\bar{x}(t),u(t)) - a[t]\,\rangle_H + \langle\, Q_1(t), b(t,\bar{x}(t),u(t)) - b[t]\,\rangle_{\mathcal{L}_2^0} \\
&\qquad -(f(t,\bar{x}(t),u(t)) - f[t])\big)dt \\
&\leq 0.
\end{aligned}
\tag{4.7}
$$

For all $(t,\omega)\in[0,T]\times\Omega$ and integer $i\geq 1$ define the closed sets

$$
\begin{aligned}
K_i(t,\omega) = \big\{u\in U \;\big|\; &\langle\, P_1(t,\omega), a(t,\bar{x}(t,\omega),u,\omega) - a[t]\,\rangle_H + \langle\, Q_1(t,\omega), b(t,\bar{x}(t,\omega),u,\omega) - b[t]\,\rangle_{\mathcal{L}_2^0} \\
&-(f(t,\bar{x}(t,\omega),u,\omega) - f[t]) \geq 2^{-i}\big\}
\end{aligned}
$$

and

$$
\widehat{K}_i(t,\omega) = \begin{cases} K_i(t,\omega) & \text{if } K_i(t,\omega)\neq\emptyset, \\ \{\bar{u}(t,\omega)\} & \text{otherwise.}\end{cases}
$$

Let A_i denote the set of all (t,ω) such that $K_i(t,\omega)\neq\emptyset$. To end the proof it is enough to show that $\mu-$measure of the following set

$$
\begin{aligned}
\big\{(t,\omega) \;\big|\; \exists\, u\in U \text{ so that } &\langle\, P_1(t,\omega), a(t,\bar{x}(t,\omega),u,\omega) - a[t]\,\rangle_H \\
&+\langle\, Q_1(t,\omega), b(t,\bar{x}(t,\omega),u,\omega) - b[t]\,\rangle_{\mathcal{L}_2^0} - (f(t,\bar{x}(t,\omega),u,\omega) - f[t]) > 0\big\}
\end{aligned}
$$

is equal to zero. Indeed otherwise there would exist an $i \geq 1$ such that $\mu(A_i) > 0$. By [1, Theorem 8.2.9] $K_i(\cdot,\cdot)$ admits a $\mu-$measurable selection. Modifying it on a set of measure zero, we obtain a control $u(\cdot) \in \mathcal{U}_{ad}$ such that for a.e. $(t,\omega) \in A_i$,

$$\begin{aligned}&\langle\, P_1(t,\omega), a(t,\bar{x}(t,\omega),u(t,\omega),\omega) - a[t]\,\rangle_H + \langle\, Q_1(t,\omega), b(t,\bar{x}(t,\omega),u(t,\omega),\omega) - b[t]\,\rangle_{\mathcal{L}_2^0}\\ &\geq f(t,\bar{x}(t,\omega),u(t,\omega),\omega) - f[t] + 2^{-i},\end{aligned}$$

and for a.e. $(t,\omega) \notin A_i$,

$$\begin{aligned}&\langle\, P_1(t,\omega), a(t,\bar{x}(t,\omega),u(t,\omega),\omega) - a[t]\,\rangle_H + \langle\, Q_1(t,\omega), b(t,\bar{x}(t,\omega),u(t,\omega),\omega) - b[t]\,\rangle_{\mathcal{L}_2^0}\\ &= f(t,\bar{x}(t,\omega),u(t,\omega),\omega) - f[t].\end{aligned}$$

Integrating the above two relations, we get a contradiction with (4.7). Consequently $\mu(A_i) = 0$ for all i. □

Remark 4.1 *In [14] under stronger regularity assumptions (but without assuming (A4)) and using the spike variation technique, the following maximality condition was proved :*

$$\begin{aligned}&\langle P_1(t), a(t,\bar{x}(t),u) - a[t])\rangle_H + \langle Q_1(t), b(t,\bar{x}(t),u) - b[t])\rangle_{\mathcal{L}_2^0} + f[t] - f(t,\bar{x}(t),u)\\ &\quad+\frac{1}{2}\langle P_2(t)(b(t,\bar{x}(t),u) - b[t]), b(t,\bar{x}(t),u) - b[t]\rangle_{\mathcal{L}_2^0}\\ &\leq 0, \qquad \text{a.e. in } [0,T]\times\Omega,\ \forall u \in U,\end{aligned} \tag{4.8}$$

where (P_1, Q_1) is given by Lemma 2.2, and $P_2(\cdot)$ is provided by Theorem 2.2. Note that, when $\dim H < \infty$ *and $A = 0$, the inequality (4.8) gives the classical maximality condition established in [17].*

We claim that the maximality condition (4.8) implies the conclusion of Theorem 4.1 whenever the sets $F(t,x,\omega)$ defined above are convex. Indeed, fix (t,ω) such that the inequality (4.8) is verified and let $u \in U$. Then for all small $\varepsilon > 0$, the vector

$$(a[t], b[t], f[t]) + \varepsilon(a(t,\bar{x}(t,\omega),u,\omega) - a[t], b(t,\bar{x}(t,\omega),u,\omega) - b[t], f(t,\bar{x}(t,\omega),u,\omega) - f[t])$$

is an element of $F(t,\bar{x}(t),\omega)$. Fix $\varepsilon > 0$ sufficiently small. Then there exist $u_1 \in U$ and $r \geq 0$ such that

$$\begin{aligned}a(t,\bar{x}(t,\omega),u_1,\omega) &= a[t] + \varepsilon(a(t,\bar{x}(t,\omega),u,\omega) - a[t]),\\ b(t,\bar{x}(t,\omega),u_1,\omega) &= b[t] + \varepsilon(b(t,\bar{x}(t,\omega),u,\omega) - b[t]),\\ f(t,\bar{x}(t,\omega),u_1,\omega) + r &= f[t] + \varepsilon(f(t,\bar{x}(t,\omega),u,\omega) - f[t]).\end{aligned}$$

Hence, by the maximality condition and noting $r \geq 0$ we obtain that

$$\begin{aligned}&\langle P_1(t,\omega), a(t,\bar{x}(t,\omega),u_1,\omega) - a[t]\rangle_H + \langle Q_1(t,\omega), b(t,\bar{x}(t,\omega),u_1,\omega) - b[t]\rangle_{\mathcal{L}_2^0} + f[t]\\ &\quad -f(t,\bar{x}(t,\omega),u_1,\omega) - r + \frac{1}{2}\langle P_2(t,\omega)(b(t,\bar{x}(t,\omega),u_1,\omega) - b[t]), b(t,\bar{x}(t,\omega),u_1,\omega) - b[t]\rangle_{\mathcal{L}_2^0}\\ &\leq 0.\end{aligned}$$

This implies that

$$\begin{aligned}&\varepsilon\langle P_1(t,\omega), a(t,\bar{x}(t,\omega),u,\omega) - a[t]\rangle_H + \langle Q_1(t,\omega), b(t,\bar{x}(t,\omega),u,\omega) - b[t]\rangle_{\mathcal{L}_2^0}\\ &\quad -\varepsilon(f(t,\bar{x}(t,\omega),u,\omega) - f[t]) + \frac{\varepsilon^2}{2}\langle P_2(t,\omega)(b(t,\bar{x}(t,\omega),u,\omega) - b[t]), b(t,\bar{x}(t,\omega),u,\omega) - b[t]\rangle_{\mathcal{L}_2^0}\\ &\leq 0.\end{aligned}$$

Since the above is valid for any small $\varepsilon > 0$, dividing both sides of the above inequality by ε and taking the limit when $\varepsilon \to 0+$ we get

$$\langle P_1(t,\omega), a(t,\bar{x}(t,\omega),u,\omega) - a[t]\rangle_H + \langle Q_1(t,\omega), b(t,\bar{x}(t,\omega),u,\omega) - b[t]\rangle_{\mathcal{L}_2^0}$$
$$-(f(t,\bar{x}(t,\omega),u,\omega) - f[t]) \leq 0.$$

Because $u \in U$ is arbitrary, we end the proof of our claim.

On the other hand, Theorem 4.1 does NOT imply the maximality condition (4.8) because the later does not need the assumption (A4) and also because (4.8) contains an additional term involving $P_2(t)$. We would like to underline here that the classical technique of convexification of dynamics, which works well in the deterministic setting, does not have its equivalent for stochastic control systems.

5 Second order necessary optimality condition

In this section, we investigate the second order necessary conditions for the local minimizers $(\bar{x},\bar{u})$ of (1.3). In addition to the assumptions (A1)–(A3), we suppose that

(A6) *The functions a, b, f and g satisfy the following:*

i) *For a.e. $(t,\omega) \in [0,T]\times\Omega$, the functions $a(t,\cdot,\cdot,\omega): H\times\widetilde{H} \to H$ and $b(t,\cdot,\cdot,\omega): H\times\widetilde{H} \to \mathcal{L}_2^0$ are twice differentiable and*

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

is uniformly continuous in $x \in H$ and $u \in \widetilde{H}$, and,

$$|a_{(x,u)^2}(t,x,u,\omega)|_{\mathcal{L}((H,\widetilde{H})\times(H,\widetilde{H});H)} + |b_{(x,u)^2}(t,x,u,\omega)|_{\mathcal{L}((H,\widetilde{H})\times(H,\widetilde{H});\mathcal{L}_2^0)} \leq L,\ \forall\,(x,u) \in H\times U;$$

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

$$\begin{cases} |f_{(x,u)^2}(t,x,u,\omega)|_{\mathcal{L}((H,\widetilde{H})\times(H,\widetilde{H});\mathbb{R})} \leq L, \\ |f_{(x,u)^2}(t,x,u,\omega) - f_{(x,u)^2}(t,\tilde{x},\tilde{u},\omega)|_{\mathcal{L}((H,\widetilde{H})\times(H,\widetilde{H});\mathbb{R})} \leq L(|x-\tilde{x}|_H + |u-\tilde{u}|_{\widetilde{H}}), \\ |g_{xx}(x,\omega)|_{\mathcal{L}(H\times H;\mathbb{R})} \leq L,\ |g_{xx}(x,\omega) - g_{xx}(\tilde{x},\omega)|_{\mathcal{L}(H\times H;\mathbb{R})} \leq L|x-\tilde{x}|_H. \end{cases}$$

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

$$\varphi_{xx}[t] = \varphi_{xx}(t,\bar{x}(t),\bar{u}(t)), \quad \varphi_{xu}[t] = \varphi_{xu}(t,\bar{x}(t),\bar{u}(t)), \quad \varphi_{uu}[t] = \varphi_{uu}(t,\bar{x}(t),\bar{u}(t)). \tag{5.1}$$

Let $\bar{u}, v, h, h_\varepsilon \in L_{\mathbb{F}}^{2\beta}(\Omega; L^4(0,T;\widetilde{H}))$ $(\beta \geq 1)$ be such that

$$|h_\varepsilon - h|_{L_{\mathbb{F}}^{2\beta}(\Omega;L^4(0,T;\widetilde{H}))} \to 0, \quad \text{as } \varepsilon \to 0^+. \tag{5.2}$$

Set

$$u^\varepsilon := \bar{u} + \varepsilon v + \varepsilon^2 h_\varepsilon. \tag{5.3}$$

Denote by x^ε the solution of (1.1) with u replaced by u^ε. Put

$$\delta x^\varepsilon = x^\varepsilon - \bar{x}, \qquad \delta u^\varepsilon = \varepsilon v + \varepsilon^2 h_\varepsilon. \tag{5.4}$$

Similarly to [7], we introduce the following second-order variational equation[3]:

$$\begin{cases} dy_2(t) = \Big(Ay_2(t) + a_x[t]y_2(t) + 2a_u[t]h(t) + a_{xx}[t]\big(y_1(t), y_1(t)\big) + 2a_{xu}[t]\big(y_1(t), v(t)\big) \\ \qquad\qquad + a_{uu}[t]\big(v(t), v(t)\big)\Big)dt + \Big(b_x[t]y_2(t) + 2b_u[t]h(t) + b_{xx}[t]\big(y_1(t), y_1(t)\big) \\ \qquad\qquad + 2b_{xu}[t]\big(y_1(t), v(t)\big) + b_{uu}[t]\big(v(t), v(t)\big)\Big)dW(t), \qquad t \in [0,T], \\ y_2(0) = 0, \end{cases} \tag{5.5}$$

where y_1 is the solution to the first variational equation (3.2) (for $v(\cdot)$ as above). The adjoint equation for (5.5) is given by (2.11) (See Theorem 2.2 for its well-posedness in the sense of relaxed transposition solution).

Similarly to [5, Lemma 4.1], we have the following estimates for solutions to (5.5).

Lemma 5.1 *Assume (A1)–(A3), (A6) and let $\beta \geq 1$. Then, for $\bar u, v, h, h_\varepsilon \in L^{2\beta}_{\mathbb{F}}(\Omega; L^4(0,T;\widetilde H))$ so that (5.2) holds and for δx^ε given by (5.4), we have*

$$\|y_2\|_{L^\infty_{\mathbb{F}}(0,T;L^\beta(\Omega;H))} \leq C\left(\|v\|_{L^{2\beta}_{\mathbb{F}}(\Omega;L^4(0,T;\widetilde H))} + \|h\|_{L^{\beta}_{\mathbb{F}}(\Omega;L^2(0,T;\widetilde H))}\right).$$

Furthermore,

$$\|r_2^\varepsilon\|_{L^\infty_{\mathbb{F}}(0,T;L^\beta(\Omega;H))} \to 0, \qquad \text{as } \varepsilon \to 0^+, \tag{5.6}$$

where

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

To simplify the notation, we define[4] (Recall (2.10) for the definition of $\mathbb{H}(\cdot)$):

$$\begin{aligned} \mathbb{S}(t,x,u,y_1,z_1,y_2,\omega) := \ & \mathbb{H}_{xu}(t,x,u,y_1,z_1,\omega) + a_u(t,x,u,\omega)^* y_2 \\ & + b_u(t,x,u,\omega)^* y_2 b_x(t,x,u,\omega), \end{aligned} \tag{5.7}$$

where $(t,x,u,y_1,z_1,y_2,\omega) \in [0,T] \times H \times \widetilde H \times H \times \mathcal{L}_2^0 \times \mathcal{L}(H) \times \Omega$, and denote

$$\mathbb{S}[t] = \mathbb{S}(t, \bar x(t), \bar u(t), P_1(t), Q_1(t), P_2(t)), \quad t \in [0,T]. \tag{5.8}$$

Let $\bar u \in \mathcal{U}_{ad} \cap L^4_{\mathbb{F}}(0,T;\widetilde H)$. Define the critical set

$$\Upsilon_{\bar u} := \left\{ v \in L^2_{\mathbb{F}}(0,T;\widetilde H)) \ \Big| \ \langle \mathbb{H}_u(t,\omega), v(t,\omega)\rangle_{\widetilde H} = 0 \ \text{ a.e. } (t,\omega) \in [0,T] \times \Omega \right\},$$

and the set of admissible second order variations by

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

We have the following result.

[3]Recall that, for any C^2-function $F(\cdot) : X \to Y$ and $x_0 \in X$, $F_{xx}(x_0) \in \mathcal{L}(X \times X; Y)$. This means that, for any $x_1, x_2 \in X$, $F_{xx}(x_0)(x_1,x_2) \in Y$. Hence, by (5.1), $a_{xx}[t]\big(y_1(t), y_1(t)\big)$ (in (5.5)) stands for $a_{xx}(t,\bar x(t),\bar u(t))\big(y_1(t), y_1(t)\big)$. One has a similar meaning for $a_{uu}[t]\big(v(t), v(t)\big)$ and so on.

[4]Note that the definition of $\mathbb{S}(t,x,u,y_1,z_1,y_2,\omega)$ in (5.7) is different from the one in [5, p. 3708]. The main reason for this is due to the fact that the characterization of $Q(\cdot)$ in Theorem 2.2 is much weaker than the one in the finite dimensions.

Theorem 5.1 *Let (A1)–(A3) and (A6) hold, $(\bar{x}, \bar{u})$ be an optimal pair for Problem (OP) and $\bar{u} \in L^4_{\mathbb{F}}(0,T;\widetilde{H})$. Then, for all $(v,h) \in \mathcal{A}_{\bar{u}}$ with $v \in \Upsilon_{\bar{u}}$, we have*

$$\begin{aligned}
&\mathbb{E}\int_0^T \Big[2\left\langle \mathbb{H}_u[s], h(s)\right\rangle_{\widetilde{H}} + 2\left\langle \mathbb{S}[s]y_1(s), v(s)\right\rangle_{\widetilde{H}} + \left\langle \mathbb{H}_{uu}[s]v(s), v(s)\right\rangle_{\widetilde{H}} \\
&\quad + \big\langle P_2(s)b_u[s]v(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} + \big\langle b_u[s]v(s), \widehat{Q}_2^{(0)}(0, a_u v, b_u v)(s)\big\rangle_{\mathcal{L}_2^0} \\
&\quad + \big\langle Q_2^{(0)}(0, a_u v, b_u v)(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0}\Big] ds \\
&\le 0.
\end{aligned} \tag{5.9}$$

Proof. Since $\bar{u} \in L^4_{\mathbb{F}}(0,T;\widetilde{H})$, for any $(v,h) \in \mathcal{A}_{\bar{u}}$ with $v \in \Upsilon_{\bar{u}}$, similarly to the proof of [5, Theorem 4.1], we may choose $h_\varepsilon \in L^4_{\mathbb{F}}(0,T;\widetilde{H}))$ so that (5.2) holds for $\beta = 2$ and $u^\varepsilon \in \mathcal{U}_{ad}$, where u^ε and the corresponding x^ε, δx^ε and δu^ε are as in (5.3)–(5.4). Denote

$$\tilde{f}^\varepsilon_{xx}(t) := \int_0^1 (1-\theta) f_{xx}(t, \bar{x}(t) + \theta\delta x^\varepsilon(t), \bar{u}(t) + \theta\delta u^\varepsilon(t)) d\theta.$$

Mappings $\tilde{f}^\varepsilon_{xu}(t)$, $\tilde{f}^\varepsilon_{uu}(t)$ and $\tilde{g}^\varepsilon_{xx}(T)$ are defined in a similar way.

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

$$\begin{aligned}
&\frac{J(u^\varepsilon) - J(\bar{u})}{\varepsilon^2} \\
&= \frac{1}{\varepsilon^2}\mathbb{E}\int_0^T \Big(\langle f_x[t], \delta x^\varepsilon(t)\rangle_H + \langle f_u[t], \delta u^\varepsilon(t)\rangle_{\widetilde{H}} + \langle \tilde{f}^\varepsilon_{xx}(t)\delta x^\varepsilon(t), \delta x^\varepsilon(t)\rangle_H \\
&\quad + 2\langle \tilde{f}^\varepsilon_{xu}(t)\delta x^\varepsilon(t), \delta u^\varepsilon(t)\rangle_{\widetilde{H}} + \langle \tilde{f}^\varepsilon_{uu}(t)\delta u^\varepsilon(t), \delta u^\varepsilon(t)\rangle_{\widetilde{H}}\Big) dt \\
&\quad + \frac{1}{\varepsilon^2}\mathbb{E}\Big(\langle g_x(\bar{x}(T)), \delta x^\varepsilon(T)\rangle_H + \langle \tilde{g}^\varepsilon_{xx}(\bar{x}(T))\delta x^\varepsilon(T), \delta x^\varepsilon(T)\rangle_H\Big) \\
&= \mathbb{E}\int_0^T \Big[\frac{1}{\varepsilon}\langle f_x[t], y_1(t)\rangle_H + \frac{1}{2}\langle f_x[t], y_2(t)\rangle_H + \frac{1}{\varepsilon}\langle f_u[t], v(t)\rangle_{\widetilde{H}} + \langle f_u[t], h(t)\rangle_{\widetilde{H}} \\
&\quad + \frac{1}{2}\Big(\langle f_{xx}[t]y_1(t), y_1(t)\rangle_H + 2\langle f_{xu}[t]y_1(t), v(t)\rangle_{\widetilde{H}} + \langle f_{uu}[t]v(t), v(t)\rangle_{\widetilde{H}}\Big)\Big] dt \\
&\quad + \mathbb{E}\Big(\frac{1}{\varepsilon}\langle g_x(\bar{x}(T)), y_1(T)\rangle_H + \frac{1}{2}\langle g_x(\bar{x}(T)), y_2(T)\rangle_H \\
&\quad + \frac{1}{2}\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle_H\Big) + \rho_2^\varepsilon,
\end{aligned}$$

where

$$\begin{aligned}
\rho_2^\varepsilon \;=\; &\mathbb{E}\int_0^T \Big(\langle f_x[t], r_2^\varepsilon(t)\rangle_H + \langle f_u[t], h_\varepsilon(t) - h(t)\rangle_{\widetilde{H}}\Big) dt + \mathbb{E}\langle g_x(\bar{x}(T)), r_2^\varepsilon(T)\rangle_H \\
&+ \mathbb{E}\int_0^T \Big[\Big(\Big\langle \tilde{f}^\varepsilon_{xx}(t)\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta x^\varepsilon(t)}{\varepsilon}\Big\rangle_H - \frac{1}{2}\langle f_{xx}[t]y_1(t), y_1(t)\rangle_H\Big) \\
&\qquad + \Big(2\Big\langle \tilde{f}^\varepsilon_{xu}(t)\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta u^\varepsilon(t)}{\varepsilon}\Big\rangle_{\widetilde{H}} - \langle f_{xu}[t]y_1(t), v(t)\rangle_{\widetilde{H}}\Big) \\
&\qquad + \Big(\Big\langle \tilde{f}^\varepsilon_{uu}(t)\frac{\delta u^\varepsilon(t)}{\varepsilon}, \frac{\delta u^\varepsilon(t)}{\varepsilon}\Big\rangle_{\widetilde{H}} - \frac{1}{2}\langle f_{uu}[t]v(t), v(t)\rangle_{\widetilde{H}}\Big)\Big] dt \\
&+ \mathbb{E}\Big(\Big\langle \tilde{g}^\varepsilon_{xx}(\bar{x}(T))\frac{\delta x^\varepsilon(T)}{\varepsilon}, \frac{\delta x^\varepsilon(T)}{\varepsilon}\Big\rangle_H - \frac{1}{2}\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle_H\Big).
\end{aligned}$$

As in Lemma 5.1, we find that $\lim_{\varepsilon\to 0^+}\rho_2^\varepsilon = 0$. On the other hand, by calculations done in (3.8)–(3.10) and the definition of $\mathbb{H}(\cdot))$ in (2.10), and recalling that $v \in \Upsilon_{\bar{u}}$, we have

$$\begin{aligned}
&\frac{1}{\varepsilon}\mathbb{E}\int_0^T \Big(\langle f_x[t], y_1(t)\rangle_H + \langle f_u[t], v(t)\rangle_{\widetilde{H}}\Big)dt + \frac{1}{\varepsilon}\mathbb{E}\langle g_x(\bar{x}(T)), y_1(T)\rangle_H \\
&= -\frac{1}{\varepsilon}\mathbb{E}\int_0^T \langle \mathbb{H}_u[t], v(t)\rangle_H dt = 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
0 &\le \lim_{\varepsilon\to 0^+}\frac{J(u^\varepsilon(\cdot)) - J(\bar{u}(\cdot))}{\varepsilon^2} \\
&= \mathbb{E}\int_0^T \Big[\frac{1}{2}\langle f_x[t], y_2(t)\rangle_H + \langle f_u[t], h(t)\rangle_{\widetilde{H}} \\
&\quad + \frac{1}{2}\Big(\langle f_{xx}[t]y_1(t), y_1(t)\rangle_H + 2\langle f_{xu}[t]y_1(t), v(t)\rangle_{\widetilde{H}} + \langle f_{uu}[t]v(t), v(t)\rangle_{\widetilde{H}}\Big)\Big]dt \\
&\quad + \frac{1}{2}\mathbb{E}\Big(\langle g_x(\bar{x}(T)), y_2(T)\rangle_H + \langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle_H\Big).
\end{aligned} \tag{5.10}$$

From (2.8) and (5.5), it follows that

$$\begin{aligned}
&\mathbb{E}\langle g_x(\bar{x}(T)), y_2(T)\rangle_H \\
&= \mathbb{E}\int_0^T \big\langle a_x[s]^*P_1(s) + b_x[s]^*Q_1(s) - f_x[s], y_2(s)\big\rangle_H ds \\
&\quad - \mathbb{E}\int_0^T \big\langle P_1(s), a_x[s]y_2(s) + 2a_u[s]h(s) + a_{xx}[s]\big(y_1(s), y_1(s)\big) \\
&\quad + 2a_{xu}[s]\big(y_1(s), v(s)\big) + a_{uu}[s]\big(v(s), v(s)\big)\big\rangle_H ds \\
&\quad - \mathbb{E}\int_0^T \big\langle Q_1(s), b_x[s]y_2(s) + 2b_u[s]h(s) + b_{xx}[s]\big(y_1(s), y_1(s)\big) \\
&\quad + 2b_{xu}[s]\big(y_1(s), v(s)\big) + b_{uu}[s]\big(v(s), v(s)\big)\big\rangle_{\mathcal{L}_2^0} ds \\
&= -\mathbb{E}\int_0^T \Big[\big\langle f_x[s], y_2(s)\big\rangle_H + \big\langle P_1(s), 2a_u[s]h(s) + a_{xx}[s]\big(y_1(s), y_1(s)\big) \\
&\quad + 2a_{xu}[s]\big(y_1(s), v(s)\big) + a_{uu}[s]\big(v(s), v(s)\big)\big\rangle_H + \big\langle Q_1(s), 2b_u[s]h(s) + b_{xx}[s]\big(y_1(s), y_1(s)\big) \\
&\quad + 2b_{xu}[s]\big(y_1(s), v(s)\big) + b_{uu}[s]\big(v(s), v(s)\big)\big\rangle_{\mathcal{L}_2^0}\Big] ds.
\end{aligned} \tag{5.11}$$

On the other hand, by (2.16) and (3.2), we find that

$$
\begin{aligned}
&\mathbb{E}\big\langle g_{xx}(\bar{x}(T))y_1(T), y_1(T)\big\rangle_H \\
&= \mathbb{E}\int_0^T \big\langle \mathbb{H}_{xx}[s]y_1(s), y_1(s)\big\rangle_H ds - \mathbb{E}\int_0^T \big\langle P_2(s)a_u[s]v(s), y_1(s)\big\rangle_H ds \\
&\quad -\mathbb{E}\int_0^T \big\langle P_2(s)y_1(s), a_u[s]v(s)\big\rangle_H ds - \mathbb{E}\int_0^T \big\langle P_2(s)b_x[s]y_1(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} ds \\
&\quad -\mathbb{E}\int_0^T \big\langle P_2(s)b_u[s]v(s), b_x[s]y_1(s) + b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} ds \\
&\quad -\mathbb{E}\int_0^T \big\langle b_u[s]v(s), \widehat{Q}_2^{(0)}(0, a_u v, b_u v)(s)\big\rangle_{\mathcal{L}_2^0} ds - \mathbb{E}\int_0^T \big\langle Q_2^{(0)}(0, a_u v, b_u v)(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} ds \\
&= \mathbb{E}\int_0^T \Big[\big\langle \mathbb{H}_{xx}[s]y_1(s), y_1(s)\big\rangle_H - 2\big\langle P_2(s)a_u[s]v(s), y_1(s)\big\rangle_H - 2\big\langle P_2(s)b_x[s]y_1(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} \\
&\quad -\big\langle P_2(s)b_u[s]v(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} - \big\langle b_u[s]v(s), \widehat{Q}_2^{(0)}(0, a_u v, b_u v)(s)\big\rangle_{\mathcal{L}_2^0} \\
&\quad -\big\langle Q_2^{(0)}(0, a_u v, b_u v)(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0}\Big] ds.
\end{aligned} \tag{5.12}
$$

Substituting (5.11) and (5.12) into (5.10) yields

$$
\begin{aligned}
0 &\geq -\mathbb{E}\int_0^T \Big[\langle f_u[s], h(s)\rangle_{\widetilde{H}} + \langle f_{xu}[s]y_1(s), v(s)\rangle_{\widetilde{H}} + \frac{1}{2}\langle f_{uu}[s]v(s), v(s)\rangle_{\widetilde{H}}\Big] ds \\
&\quad +\mathbb{E}\int_0^T \Big[\big\langle P_1(s), a_u[s]h(s) + a_{xu}[s]\big(y_1(s), v(s)\big) + \frac{1}{2}a_{uu}[s]\big(v(s), v(s)\big)\big\rangle_H \\
&\quad +\big\langle Q_1(s), b_u[s]h(s) + b_{xu}[s]\big(y_1(s), v(s)\big) + \frac{1}{2}b_{uu}[s]\big(v(s), v(s)\big)\big\rangle_{\mathcal{L}_2^0}\Big] ds \\
&\quad +\frac{1}{2}\mathbb{E}\int_0^T \Big[2\big\langle P_2(s)a_u[s]v(s), y_1(s)\big\rangle_H + 2\big\langle P_2(s)b_x[s]y_1(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} \\
&\quad +\big\langle P_2(s)b_u[s]v(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} + \big\langle b_u[s]v(s), \widehat{Q}_2^{(0)}(0, a_u v, b_u v)(s)\big\rangle_{\mathcal{L}_2^0} \\
&\quad +\big\langle Q_2^{(0)}(0, a_u v, b_u v)(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0}\Big] ds \\
&= \frac{1}{2}\mathbb{E}\int_0^T \Big[2\,\langle \mathbb{H}_u[s], h(s)\rangle_{\widetilde{H}} + 2\,\langle \mathbb{S}[s]y_1(s), v(s)\rangle_{\widetilde{H}} + \langle \mathbb{H}_{uu}[s]v(s), v(s)\rangle_{\widetilde{H}} \\
&\quad +\big\langle P_2(s)b_u[s]v(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0} + \big\langle b_u[s]v(s), \widehat{Q}_2^{(0)}(0, a_u v, b_u v)(s)\big\rangle_{\mathcal{L}_2^0} \\
&\quad +\big\langle Q_2^{(0)}(0, a_u v, b_u v)(s), b_u[s]v(s)\big\rangle_{\mathcal{L}_2^0}\Big] ds.
\end{aligned}
$$

Then, we obtain the desired inequality (5.9). This completes the proof of Theorem 5.1. □

Acknowledgment

The authors are grateful to two anonymous referees for their helpful comments on the first version of this paper.

References

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

[2] A. Bensoussan. *Stochastic maximum principle for distributed parameter systems. J. Franklin Inst.* **315** (1983), 387–406.

[3] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions, Second Edition.* Cambridge University Press, Cambridge, 2014.

[4] K. Du and Q. Meng. *A maximum principle for optimal control of stochastic evolution equations. SIAM J. Control Optim.* **51** (2013), 4343–4362.

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

[6] M. Fuhrman, Y. Hu and G. Tessitore. *Stochastic maximum principle for optimal control of SPDEs. Appl. Math. Optim.* **68** (2013), 181–217.

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

[8] Y. Hu and S. Peng. *Maximum principle for semilinear stochastic evolution control systems. Stochastics Stochastics Rep.* **33** (1990), 159–180.

[9] Q. Lü. *Second order necessary conditions for optimal control problems of stochastic evolution equations. Proceedings of the 35th Chinese Control Conference.* Chengdu, China, July 27-29, (2016), 2620–2625.

[10] Q. Lü, J. Yong and X. Zhang. *Representation of Itô integrals by Lebesgue/Bochner integrals. J. Eur. Math. Soc.* **14** (2012), 1795–1823 (Erratum: *J. Eur. Math. Soc.* **20** (2018), 259–260).

[11] Q. Lü, H. Zhang and X. Zhang. *Second order optimality conditions for optimal control problems of stochastic evolution equations.* submitted.

[12] Q. Lü and X. Zhang. *Well-posedness of backward stochastic differential equations with general filtration. J. Differential Equations.* **254** (2013), 3200–3227.

[13] Q. Lü and X. Zhang. *General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions.* Springer Briefs in Mathematics, Springer, New York, 2014. (See also http://arxiv.org/abs/1204.3275)

[14] Q. Lü and X. Zhang. *Transposition method for backward stochastic evolution equations revisited, and its application. Math. Control Relat. Fields.* **5** (2015), 529–555.

[15] Q. Lü and X. Zhang. *Mathematical Theory for Stochastic Distributed Parameter Control Systems.* Book in preparation.

[16] N. I. Mahmudova and M. A. McKibben. *On backward stochastic evolution equations in Hilbert spaces and optimal control. Nonlinear Anal.* **67** (2007), 1260–1274.

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

[18] S. Tang and X. Li. *Maximum principle for optimal control of distributed parameter stochastic systems with random jumps.* In: *Differential Equations, Dynamical Systems, and Control Science.* Edited by K. D. Elworthy, W. N. Everitt and E. B. Lee. Lecture Notes in Pure and Appl. Math. vol. **152**. Dekker, New York, 1994, 867–890.

[19] H. Zhang and X. Zhang. *Pointwise second-order necessary conditions for stochastic optimal controls, Part II: The general cases. SIAM J. Control Optim.* **55** (2017), 2841–2875.

[20] H. Zhang and X. Zhang. *Second-order necessary conditions for stochastic optimal control problems. SIAM Rev.* **60** (2018), 139–178.

[21] X. Y. Zhou. *On the necessary conditions of optimal controls for stochastic partial differential equations. SIAM J. Control Optim.* **31** (1993), 1462–1478.