First and second order necessary optimality conditions for controlled stochastic evolution equations with control and state constraints

Hélène Frankowska, Qi Lü

First and Second Order Necessary Optimality Conditions for Controlled Stochastic Evolution Equations with Control and State Constraints

Hélène Frankowska * and Qi Lü†

Abstract

The purpose of this paper is to establish first and second order necessary optimality conditions for optimal control problems of stochastic evolution equations with control and state constraints. The control acts both in the drift and diffusion terms and the control region is a nonempty closed subset of a separable Hilbert space. We employ some classical set-valued analysis tools and theories of the transposition solution of vector-valued backward stochastic evolution equations and the relaxed-transposition solution of operator-valued backward stochastic evolution equations to derive these optimality conditions. The correction part of the second order adjoint equation, which does not appear in the first order optimality condition, plays a fundamental role in the second order optimality condition.

Key words: Stochastic optimal control, necessary optimality conditions, set-valued analysis.

AMS subject classifications: Primary 93E20; Secondary 60H15.

1. Introduction

Let $T > 0$ and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ a complete filtered probability space with the càdlàg (right continuous with left limits) filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$, on which a cylindrical Brownian motion $\{W(t)\}_{t\in[0,T]}$ taking values in a separable Hilbert space V is defined. Let H be a separable Hilbert space and A be an unbounded linear operator generating a contractive C_0-semigroup $\{S(t)\}_{t\geq 0}$ on H. For a nonempty closed subset U of a separable Hilbert space H_1 define

$$\mathcal{U}^p \triangleq \Big\{u(\cdot) : [0,T] \to U \;\Big|\; u(\cdot) \in L^p_{\mathbb{F}}(0,T;H_1)\Big\}$$

and consider the following controlled stochastic evolution equation (SEE for short):

$$\left\{\begin{array}{ll} dx(t) = \big(Ax(t) + a(t,x(t),u(t))\big)dt + b(t,x(t),u(t))dW(t) & \text{in } (0,T],\\ x(0) = \nu_0 \in H, \end{array}\right. \tag{1.1}$$

where $u \in \mathcal{U}^2$. A process $x(\cdot) \equiv x(\cdot\,;\nu_0,u) \in L^2_{\mathbb{F}}(\Omega;C([0,T];H))$ is called a mild solution of (1.1) if

*CNRS, IMJ-PRG, UMR 7586, Sorbonne Université, case 247, 4 place Jussieu, 75252 Paris, France. The research of this author is partially supported by the Gaspard Monge Program for Optimisation and Operational Research, Jacques Hadamard Mathematical Foundation (FMJH) and the AFOSR grant FA 9550-18-1-0254. *E-mail:* helene.frankowska@imj-prg.fr.

†School of Mathematics, Sichuan University, Chengdu, 610064, China. The research of this author is supported by the NSF of China under grant 11471231, the Fundamental Research Funds for the Central Universities in China under grant 2015SCU04A02 and Grant MTM2014-52347 of the MICINN, Spain. *E-mail:* lu@scu.edu.cn. The authors are grateful to Dr. Haisen Zhang for the fruitful discussion which played an important role in proving the main results of this paper.

$$x(t) = S(t)\nu_0 + \int_0^t S(t-s)a(s,x(s),u(s))ds + \int_0^t S(t-s)b(s,x(s),u(s))dW(s),\ \mathbb{P}\text{-a.s.}, \forall\, t\in[0,T].$$

Many controlled stochastic partial differential equations, such as controlled stochastic wave/heat/-Schrödinger equations, can be regarded as a special case of the system (1.1).

Let $\mathcal{V}$ be a nonempty closed subset of H, and $h:\Omega\times H\to\mathbb{R}$, $g^j: H\to\mathbb{R}$ $(j=0,\cdots,n)$. Define a Mayer type cost functional $\mathcal{J}(\cdot)$ (for the control system (1.1)) as

$$\mathcal{J}(u(\cdot),\nu_0) = \mathbb{E}h(x(T)) \tag{1.2}$$

with the state constraint

$$\mathbb{E}g^0(x(t)) \le 0, \qquad \text{for all } t\in[0,T], \tag{1.3}$$

and the initial-final states constraints

$$\nu_0 \in \mathcal{V}, \qquad \mathbb{E}g^j(x(T)) \le 0, \quad j=1,\cdots,n. \tag{1.4}$$

The set of admissible controls at the initial datum ν_0 is given by

$$\mathcal{U}_{ad}^{\nu_0} \triangleq \Big\{ u\in\mathcal{U}^2 \ \big|\ \text{the corresponding solution } x(\cdot) \text{ of (1.1) satisfies (1.3) and (1.4)} \Big\}$$

and the one of admissible trajectory-control pairs by

$$\mathcal{P}_{ad} \triangleq \Big\{ (x(\cdot),u(\cdot)) \ \big|\ u\in\mathcal{U}_{ad}^{\nu_0} \text{ for some } \nu_0\in\mathcal{V} \Big\}.$$

Remark 1.1. *Here $\mathcal{U}_{ad}^{\nu_0}$ depends on the choice of ν_0. Different ν_0 may give different $\mathcal{U}_{ad}^{\nu_0}$.*

Under the usual assumptions, (1.1) has exactly one (mild) solution $x(\cdot,\nu_0)$ with initial value $\nu_0\in\mathcal{V}$, which is called an admissible state.

We pose the optimal control problem for the system (1.1) as follows:

Problem (OP) *Find $(\bar\nu_0,\bar u(\cdot))\in\mathcal{V}\times\mathcal{U}_{ad}^{\nu_0}$ such that*

$$\mathcal{J}(\bar\nu_0,\bar u(\cdot)) = \inf_{(\nu_0,u(\cdot))\in\mathcal{V}\times\mathcal{U}_{ad}^{\nu_0}} \mathcal{J}(\nu_0,u(\cdot)). \tag{1.5}$$

In(1.5), $\bar u(\cdot)$ is said to be an optimal control and $\bar x(\cdot)$ the corresponding optimal state. $(\bar x(\cdot),\bar u(\cdot))$ is called an *optimal pair* and $(\bar\nu_0,\bar x(\cdot),\bar u(\cdot))$ is called an *optimal triple*.

Our purpose is to establish first and second order necessary optimality conditions for **Problem (OP)**.

We could also consider a more general Bolza-type cost functional

$$\mathcal{J}(u(\cdot),\nu_0) = \mathbb{E}\Big[\int_0^T \tilde h(t,x(t),u(t))dt + h(x(T))\Big].$$

However, it is well known that such optimal control problem can be reduced to **Problem (OP)** by considering an extended control system:

$$\begin{cases} dx(t) = \big(Ax(t)+a(t,x(t),u(t))\big)dt + b(t,x(t),u(t))dW(t) & \text{in } (0,T],\\ d\tilde x(t) = \tilde h(t,x(t),u(t))dt & \text{in } [0,T],\\ x(0)=\nu_0\in H, \quad \tilde x(0)=0 \end{cases} \tag{1.6}$$

with the Mayer type cost functional

$$\mathcal{J}(u(\cdot),\nu_0) = h(x(T)) + \tilde x(T),$$

under constraints

$$\mathbb{E}g^0(x(t)) \leq 0, \quad \text{for all } t \in [0,T], \quad \nu_0 \in \mathcal{V}, \quad \mathbb{E}g^j(x(T)) \leq 0, \quad j = 1, \cdots, n.$$

It is one of the important issues in optimal control theory to establish necessary optimality conditions for optimal controls, which is useful for characterizing optimal controls or solving the optimal control problems numerically. Since the seminal work [34], necessary optimality conditions are studied extensively for different kinds of control systems. We refer the readers to [15, 17, 19, 23, 38, 40, 41] and the rich references therein for the first and second order necessary optimality conditions for systems governed by ordinary differential equations, by partial differential equations and by stochastic differential equations.

It is natural to seek to extend the theory of necessary optimality conditions to those infinite dimensional SEEs. The main motivation is to study the optimal control of systems governed by stochastic partial differential equations, which are useful models for many processes in natural sciences (see [5, 22] and the rich references therein).

We refer to [3] for a pioneering work on first order necessary optimality condition (Pontryagin-type maximum principle) and subsequent extensions [19, 37, 42] and so on. Nevertheless, for a long time, almost all of the works on the necessary conditions for optimal controls of infinite dimensional SEEs addressed only the case that the diffusion term does NOT depend on the control variable (i.e., the function $b(\cdot,\cdot,\cdot)$ in (1.1) is independent of u). As far as we know, the stochastic maximum principle for general infinite dimensional nonlinear stochastic systems with control-dependent diffusion coefficients and possibly nonconvex control domains had been a long standing problem till the very recent papers ([10, 18, 29, 30, 31]). In these papers first order necessary optimality conditions for controlled SEEs are established by several authors with no constraint on the state. Further, in [27, 28], some second order necessary optimality conditions for controlled SEEs are obtained, provided that there is no constraint on the state and U is convex. As far as we know, there are no results on first or second order necessary optimality conditions for controlled SEEs with state constraints and for a nonconvex set U.

Compared with [10, 18, 27, 28, 29, 30, 31], the main novelty of the present work is in employing some sharp tools of set-valued analysis with the following advantages:

- only one adjoint equation is needed to get a first order necessary optimality condition even when the diffusion term is control dependent and U is nonconvex;
- two second order necessary optimality conditions are obtained by using two adjoint equations;
- state constraints are presented.

The rest of this paper is organized as follows: in Section 2, we introduce some notations and assumptions and recall some concepts and results from the set-valued analysis to be used in this paper; Section 3 is devoted to establishing first order necessary optimality conditions; at last, in Section 4, we obtain two integral-type second order necessary optimality conditions.

2. Preliminaries

2.1. Notations and assumptions

Let X be a Banach space. For each $t \in [0,T]$ and $r \in [1,\infty)$, denote by $L^r_{\mathcal{F}_t}(\Omega;X)$ the Banach space of all (strongly) $\mathcal{F}_t$-measurable random variables $\xi : \Omega \to X$ such that $\mathbb{E}|\xi|^r_X < \infty$, with the norm $|\xi|_{L^r_{\mathcal{F}_t}(\Omega;X)} \overset{\triangle}{=} \big(\mathbb{E}|\xi|^r_X\big)^{1/r}$. Write $D_{\mathbb{F}}([0,T];L^r(\Omega;X))$ for the Banach space of all X-valued, rth power integrable $\mathbb{F}$-adapted processes $\varphi(\cdot)$ such that $\varphi : [0,T] \to L^r_{\mathcal{F}_T}(\Omega;X)$ is càdlàg, with the norm $|\varphi(\cdot)|_{D_{\mathbb{F}}([0,T];L^r(\Omega;X))} = \sup_{t\in[0,T]} \big(\mathbb{E}|\varphi(t)|_X^r\big)^{1/r}$. Write $C_{\mathbb{F}}([0,T];L^r(\Omega;X))$ for the Banach space of all X-valued, $\mathbb{F}$-adapted processes $\varphi(\cdot)$ such that $\varphi:[0,T]\to L^r_{\mathcal{F}_T}(\Omega;X)$ is continuous, with the norm inherited from $D_{\mathbb{F}}([0,T];\,L^r(\Omega;X))$.

Denote by $D([0,T];X)$ the Banach space of all X-valued càdlàg functions $\varphi(\cdot)$ such that $\sup_{t\in[0,T]}|\varphi(t)|_X<\infty$, with the norm $|\varphi|_{D([0,T];X)}=\sup_{t\in[0,T]}|\varphi(t)|_X$; by $L^2_{\mathbb{F}}(\Omega;D([0,T];X))$ the Banach space of all X-valued $\mathbb{F}$-adapted càdlàg processes $\varphi(\cdot)$ such that $\mathbb{E}\big(\sup_{t\in[0,T]}|\varphi(t)|_X\big)^2<\infty$, with the norm $|\varphi|_{L^2_{\mathbb{F}}(\Omega;D([0,T];X))}=\big[\mathbb{E}\big(\sup_{t\in[0,T]}|\varphi(t)|_X\big)^2\big]^{\frac{1}{2}}$; by $L^2_{\mathbb{F}}(\Omega;C([0,T];X))$ the space of all $\mathbb{F}$-adapted continuous processes $\varphi(\cdot)$ such that $\mathbb{E}\big(\sup_{t\in[0,T]}|\varphi(t)|_X\big)^2<\infty$, with the norm inherited from $L^2_{\mathbb{F}}(\Omega;D([0,T];X))$; by $L^2_{\mathbb{F}}(\Omega;BV([0,T];X))$ the Banach space of all X-valued, $\mathbb{F}$-adapted processes $\varphi(\cdot)$ whose sample paths have bounded variation, $\mathbb{P}$-a.s., such that $\mathbb{E}\,|\varphi(\cdot,\omega)|^2_{BV[0,T]}<+\infty$, with the norm $\big(\mathbb{E}|\varphi(\cdot,\omega)|^2_{BV[0,T]}\big)^{\frac{1}{2}}$; and by $L^2_{\mathbb{F}}(\Omega;BV_0([0,T];X))$ the space of processes $\varphi\in L^2_{\mathbb{F}}(\Omega;BV([0,T];X))$ satisfying $\varphi(0)=0$, with the norm inherited from $L^2_{\mathbb{F}}(\Omega;BV([0,T];X))$. For any $\varphi\in L^2_{\mathbb{F}}(\Omega;BV_0([0,T];X))$, one can find a $\tilde\varphi\in L^2_{\mathbb{F}}(\Omega;BV_0([0,T];X))\cap L^2_{\mathbb{F}}(\Omega;D([0,T];X))$ such that $\varphi=\tilde\varphi$ for a.e. $(t,\omega)\in[0,T]\times\Omega$. Hence, in this paper, without loss of generality, any $\varphi\in L^2_{\mathbb{F}}(\Omega;BV_0([0,T];X))$ can be considered as an element in $L^2_{\mathbb{F}}(\Omega;D([0,T];X))$.

Fix any $r_1,r_2\in[1,\infty]$. Put

$$L^{r_1}_{\mathbb{F}}(\Omega;L^{r_2}(0,T;X))=\Big\{\varphi:(0,T)\times\Omega\to X\ \Big|\ \varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted and }\mathbb{E}\Big(\int_0^T|\varphi(t)|_X^{r_2}dt\Big)^{\frac{r_1}{r_2}}<\infty\Big\},$$
$$L^{r_2}_{\mathbb{F}}(0,T;L^{r_1}(\Omega;X))=\Big\{\varphi:(0,T)\times\Omega\to X\ \Big|\ \varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted and }\int_0^T\Big(\mathbb{E}|\varphi(t)|_X^{r_1}\Big)^{\frac{r_2}{r_1}}dt<\infty\Big\}.$$

Clearly, the above two sets are Banach spaces with the following norms respectively

$$|\phi(\cdot)|_{L^{r_1}_{\mathbb{F}}(\Omega;L^{r_2}(0,T;X))}\triangleq\Big[\mathbb{E}\Big(\int_0^T|\phi(t)|_X^{r_2}dt\Big)^{\frac{r_1}{r_2}}\Big]^{\frac{1}{r_1}}$$

and

$$|\phi(\cdot)|_{L^{r_2}_{\mathbb{F}}(0,T;L^{r_1}(\Omega;X))}\triangleq\Big[\int_0^T\big(\mathbb{E}|\phi(t)|_X^{r_1}dt\big)^{\frac{r_2}{r_1}}\Big]^{\frac{1}{r_2}}.$$

If $r_1=r_2$, we simply write $L^{r_1}_{\mathbb{F}}(0,T;X)$ for the above spaces. As usual, if there is no danger of confusion, we omit the ω ($\in\Omega$) argument in the notations of functions and operators.

Let H be a separable Hilbert space and A be an unbounded linear operator (with the domain $D(A)$) on H, which generates a contractive C_0-semigroup $\{S(t)\}_{t\geq0}$ on H. It is well known that $D(A)$ is a Hilbert space with the usual graph norm. By A^*, we denote the adjoint operator of A, which generates the adjoint C_0-semigroup $\{S^*(t)\}_{t\geq0}$. Denote by $\mathcal{L}_2$ the space of all Hilbert-Schmidt operators from V to H, which is a Hilbert space with the canonical norm.

Throughout this paper, we use C to denote a generic constant, which may change from line to line.

Let us introduce the following condition:

(AS1) *$a(\cdot,\cdot,\cdot,\cdot):[0,T]\times H\times H_1\times\Omega\to H$ and $b(\cdot,\cdot,\cdot,\cdot):[0,T]\times H\times H_1\times\Omega\to\mathcal{L}_2$ are two maps such that: i) For any $(x,u)\in H\times H_1$, $a(\cdot,x,u,\cdot):[0,T]\times\Omega\to H$ and $b(\cdot,x,u,\cdot):[0,T]\times\Omega\to\mathcal{L}_2$ are $\mathcal{B}([0,T])\times\mathcal{F}$ measurable and $\mathbb{F}$-adapted; ii) For any $(t,x,\omega)\in[0,T]\times H\times\Omega$, $a(t,x,\cdot,\omega):H_1\to H$ and $b(t,x,\cdot,\omega):H_1\to\mathcal{L}_2$ are continuous, and*

$$\left\{\begin{array}{ll}|a(t,x_1,u,\omega)-a(t,x_2,u,\omega)|_H+|b(t,x_1,u,\omega)-b(t,x_2,u,\omega)|_{\mathcal{L}_2}\leq C|x_1-x_2|_H,\\ \qquad\qquad\qquad\qquad\forall\,(t,x_1,x_2,u,\omega)\in[0,T]\times H\times H\times H_1\times\Omega,\\ |a(t,0,u,\omega)|_H+|b(t,0,u,\omega)|_{\mathcal{L}_2}\leq C,\qquad\forall\,(t,u,\omega)\in[0,T]\times H_1\times\Omega.\end{array}\right.\tag{2.1}$$

We have the following result:

Lemma 2.1. *Let* **(AS1)** *hold. Then the equation* (1.1) *admits a unique mild solution. Furthermore, for some* $C > 0$ *and all* $\nu_0 \in H$,

$$|x(\cdot)|_{L^2_{\mathbb{F}}(\Omega;C([0,T];H))} \leq C\big(1 + |\nu_0|_H\big).$$

The proof of Lemma 2.1 can be found in [7, Chapter 7].

2.2. Set-valued analysis

For readers' convenience, we collect some basic facts from set-valued analysis. More information can be found in [2].

Let Z (*resp.* $\widetilde{Z}$) be a Banach (*resp.* separable Banach) space with the norm $|\cdot|_Z$ (*resp.* $|\cdot|_{\widetilde{Z}}$). Denote by Z^* (*resp.* $\widetilde{Z}^*$) the dual space of Z (*resp.* $\widetilde{Z}$). For any subset $\mathcal{K} \subset Z$, denote by $\text{int}\mathcal{K}$ and $\text{cl}\mathcal{K}$ the interior and closure of $\mathcal{K}$, respectively. $\mathcal{K}$ is called a cone if $\alpha z \in \mathcal{K}$ for every $\alpha \geq 0$ and $z \in \mathcal{K}$. Define the distance between a point $z \in Z$ and $\mathcal{K}$ as

$$dist\,(z, \mathcal{K}) \overset{\triangle}{=} \inf_{y \in \mathcal{K}} |y - z|_Z$$

and the metric projection of z onto $\mathcal{K}$ as

$$\Pi_{\mathcal{K}}(z) \overset{\triangle}{=} \{y \in \mathcal{K} \mid |y - x|_Z = dist\,(z, \mathcal{K})\}.$$

Definition 2.1. *For* $z \in \mathcal{K}$*, the Clarke tangent cone* $\mathcal{C}_{\mathcal{K}}(z)$ *to* $\mathcal{K}$ *at* z *is*

$$\mathcal{C}_{\mathcal{K}}(z) \overset{\triangle}{=} \Big\{v \in Z \;\Big|\; \lim_{\varepsilon \to 0^+, y \in \mathcal{K}, y \to z} \frac{dist\,(y + \varepsilon v, \mathcal{K})}{\varepsilon} = 0\Big\}.$$

and the adjacent cone $T^b_{\mathcal{K}}(z)$ *to* $\mathcal{K}$ *at* z *is*

$$T^b_{\mathcal{K}}(z) \overset{\triangle}{=} \Big\{v \in Z \;\Big|\; \lim_{\varepsilon \to 0^+} \frac{dist\,(z + \varepsilon v, \mathcal{K})}{\varepsilon} = 0\Big\}.$$

$\mathcal{C}_{\mathcal{K}}(z)$ is a closed convex cone in Z and $\mathcal{C}_{\mathcal{K}}(z) \subset T^b_{\mathcal{K}}(z)$. When $\mathcal{K}$ is convex, $\mathcal{C}_{\mathcal{K}}(z) = T^b_{\mathcal{K}}(z) = \text{cl}\{\alpha(\hat{z} - z) | \alpha \geq 0,\ \hat{z} \in \mathcal{K}\}$.

Definition 2.2. *For* $z \in \mathcal{K}$ *and* $v \in T^b_{\mathcal{K}}(z)$*, the second order adjacent subset to* $\mathcal{K}$ *at* (z, v) *is defined by*

$$T^{b(2)}_{\mathcal{K}}(z, v) \overset{\triangle}{=} \Big\{h \in Z \;\Big|\; \lim_{\varepsilon \to 0^+} \frac{dist\,(z + \varepsilon v + \varepsilon^2 h, \mathcal{K})}{\varepsilon^2} = 0\Big\}.$$

The dual cone of the Clarke tangent cone $\mathcal{C}_{\mathcal{K}}(z)$, denoted by $\mathcal{N}_{\mathcal{K}}(z)$, is called the normal cone of $\mathcal{K}$ at z, i.e.,

$$\mathcal{N}_{\mathcal{K}}(z) \overset{\triangle}{=} \Big\{\xi \in Z^* \;\Big|\; \langle \xi, v \rangle_{Z^*, Z} \leq 0,\ \forall\, v \in \mathcal{C}_{\mathcal{K}}(z)\Big\}.$$

Definition 2.3. *Let* (Ξ, Σ) *be a measurable space, and* $F : \Xi \rightsquigarrow Z$ *be a set-valued map. For any* $\xi \in \Xi$, $F(\xi)$ *is called the value of* F *at* ξ*. The domain of* F *is* $Dom\,(F) \overset{\triangle}{=} \{\xi \in \Xi \mid F(\xi) \neq \emptyset\}$*.* F *is called measurable if* $F^{-1}(B) \overset{\triangle}{=} \{\xi \in \Xi \mid F(\xi) \cap B \neq \emptyset\} \in \Sigma$ *for any* $B \in \mathcal{B}(Z)$*, where* $\mathcal{B}(Z)$ *is the Borel* σ*-algebra on* Z*.*

Lemma 2.2. *[16, Lemma 2.7] Suppose that (Ξ,Σ,μ) is a complete finite measure space, $p\ge 1$ and $\mathcal{K}$ is a closed nonempty subset of $\widetilde{Z}$. Put*

$$\mathbb{K}\stackrel{\triangle}{=}\big\{\varphi(\cdot)\in L^p(\Xi,\Sigma,\mu;\widetilde{Z})\ \big|\ \varphi(\xi)\in\mathcal{K},\ \mu\text{-a.e. }\xi\in\Xi\big\}. \tag{2.2}$$

Then for any $\varphi(\cdot)\in\mathbb{K}$, the set-valued map $T^b_{\mathbb{K}}(\varphi(\cdot))$: $\xi\rightsquigarrow T^b_{\mathcal{K}}(\varphi(\xi))$ and $\mathcal{C}_{\mathbb{K}}(\varphi(\cdot)):\xi\rightsquigarrow\mathcal{C}_{\mathcal{K}}(\varphi(\xi))$ are Σ-measurable, and

$$\big\{v(\cdot)\in L^p(\Xi,\Sigma,\mu;Z)\,|\ v(\xi)\in T^b_{\mathcal{K}}(\varphi(\xi)),\ \mu\text{-a.e. }\xi\in\Xi\big\}\subset T^b_{\mathbb{K}}(\varphi(\cdot)),$$

$$\big\{v(\cdot)\in L^p(\Xi,\Sigma,\mu;Z)\,|\ v(\xi)\in\mathcal{C}_{\mathcal{K}}(\varphi(\xi)),\ \mu\text{-a.e. }\xi\in\Xi\big\}\subset\mathcal{C}_{\mathbb{K}}(\varphi(\cdot)).$$

The following result provides a criteria for the measurability of set-valued maps.

Lemma 2.3. *[2, Theorem 8.1.4] Let (Ξ,Σ,μ) be a complete σ-finite measure space and F be a set-valued map from Ξ to $\widetilde{Z}$ with nonempty closed images. Then F is measurable if and only if the graph of F belongs to $\Sigma\otimes\mathcal{B}(\widetilde{Z})$.*

Definition 2.4. *We call a map $\zeta:(\Omega,\mathcal{F})\rightsquigarrow Z$ a set-valued random variable if it is measurable.*

We call a map $\Psi:[0,T]\times\Omega\rightsquigarrow Z$ a measurable set-valued stochastic process if Ψ is $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable.

We say that a measurable set-valued stochastic process Ψ is $\mathbb{F}$-adapted if $\Psi(t,\cdot)$ is $\mathcal{F}_t$-measurable for all $t\in[0,T]$.

Let

$$\mathcal{G}\stackrel{\triangle}{=}\big\{B\in\mathcal{B}([0,T])\otimes\mathcal{F}\ \big|\ B_t\in\mathcal{F}_t,\ \forall\, t\in[0,T]\big\}, \tag{2.3}$$

where $B_t\stackrel{\triangle}{=}\{\omega\in\Omega\mid(t,\omega)\in B\}$ is the t-section of B. Obviously, $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{B}([0,T])\otimes\mathcal{F}$. Denote by $\mathbf{m}$ the Lebesgue measure on $[0,T]$. The measure space $([0,T]\times\Omega,\mathcal{G},\mathbf{m}\times\mathbb{P})$ may be incomplete. Let us give a completed version of it.

Let $\widetilde{\mathcal{G}}$ be the collection of $B\subset[0,T]\times\Omega$ for which there exist $B_1,\ B_2\in\mathcal{G}$ such that $B_1\subset B\subset B_2$ and $(\mathbf{m}\times\mathbb{P})(B_2\backslash B_1)=0$. One can define a function $\widetilde{\mu}$ on $\widetilde{\mathcal{G}}$ as $\widetilde{\mu}(B)=[\mathbf{m}\times\mathbb{P}](B_1)$ for any $B\in\widetilde{\mathcal{G}}$. By Proposition 1.5.1 in [6], the measure space $([0,T]\times\Omega,\widetilde{\mathcal{G}},\widetilde{\mu})$ is a completion of $([0,T]\times\Omega,\mathcal{G},\mathbf{m}\times\mathbb{P})$.

Define

$$\mathcal{L}^2_{\mathbb{F}}(0,T;H_1)\stackrel{\triangle}{=}\Big\{y:[0,T]\times\Omega\to H_1\big|y(\cdot)\text{ is }\widetilde{\mathcal{G}}\text{-measurable},\int_{[0,T]\times\Omega}|y(s,\omega)|^2_{H_1}d\widetilde{\mu}(s,\omega)<\infty\Big\},$$

$$\widetilde{\mathcal{U}}^{\nu_0}_{ad}\stackrel{\triangle}{=}\Big\{u:[0,T]\times\Omega\to H_1\ \big|\ u(\cdot)\text{ is }\widetilde{\mathcal{G}}\text{-measurable},\ u(t)\in U,\ \widetilde{\mu}\text{-a.e., the corresponding solution }x(\cdot)\text{ of (1.1) satisfies (1.3) and (1.4)}\Big\}.$$

Clearly, $\mathcal{U}^{\nu_0}_{ad}\subset\widetilde{\mathcal{U}}^{\nu_0}_{ad}$ and $L^2_{\mathbb{F}}(0,T;H_1)\subset\mathcal{L}^2_{\mathbb{F}}(0,T;H_1)$.

Let $\Xi=[0,T]\times\Omega$, $\mu=\widetilde{\mu}$ and $Z=H_1$. From Lemma 2.2, we deduce the following result.

Corollary 2.1. *For any $u(\cdot)\in\widetilde{\mathcal{U}}^{\nu_0}_{ad}$, $\mathcal{C}_U(u(\cdot)):[0,T]\times\Omega\rightsquigarrow H_1$ is $\widetilde{\mathcal{G}}$-measurable and*

$$\mathcal{T}_u\stackrel{\triangle}{=}\big\{v(\cdot)\in\mathcal{L}^2_{\mathbb{F}}(0,T;H_1)\ \big|\ v(t)\in\mathcal{C}_U(u(t)),\ \widetilde{\mu}\text{-a.e.}\big\}\subset\mathcal{C}_{\widetilde{\mathcal{U}}^{\nu_0}_{ad}}(u(\cdot)). \tag{2.4}$$

The next result concerns the completion of a measure space, which is a corollary of Proposition 1.5.1 in [6].

Lemma 2.4. *Let (Ξ,Σ,μ) be a σ-finite measure space with the completion $(\Xi,\widetilde{\Sigma},\widetilde{\mu})$, and f be a $\widetilde{\Sigma}$-measurable function from Ξ to Z. Then there exists a Σ-measurable function g such that $\widetilde{\mu}(g(\xi)\neq f(\xi))=0$.*

Due to Lemma 2.4, in what follows, we omit $\widetilde{}$ to simplify notation.

Lemma 2.5. *Let H be a separable Hilbert space. A set-valued stochastic process $F : [0,T] \times \Omega \rightsquigarrow H$ is $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted if and only if F is $\mathcal{G}$-measurable.*

Proof. Since H is separable, it has an orthonormal basis $\{e_k\}_{k=1}^{\infty}$. Denote by Γ_k the projection operator from H to $H_k \overset{\triangle}{=} \mathrm{span}\,\{e_k\}$. Let $F_k(\cdot) = \big\langle F(\cdot), e_k \big\rangle_H$. From [21, p. 96], we know that the set-valued stochastic process $F_k : [0,T] \times \Omega \rightsquigarrow \mathbb{R}$ is $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted if and only if F_k is $\mathcal{G}$-measurable. Then Lemma 2.5 follows from the fact that $F(\cdot) = \displaystyle\sum_{k=1}^{\infty} F_k(\cdot) e_k$. $\square$

Next, we recall the notion of measurable selection for a set-valued map.

Definition 2.5. *Let (Ξ, Σ) be a measurable space and $\widehat{Z}$ a complete separable metric space. Let F be a set-valued map from Ξ to $\widehat{Z}$. A measurable map $f : \Xi \to \widehat{Z}$ is called a measurable selection of F if $f(\xi) \in F(\xi)$ for all $\xi \in \Xi$.*

A result concerning the measurable selection is given below.

Lemma 2.6. *[2, Theorem 8.1.3] Let $\widehat{Z}$ be a complete separable metric space, (Ξ, Σ) a measurable space, and $F : \Xi \rightsquigarrow \widehat{Z}$ a measurable set-valued map with nonempty closed values. Then there exists a measurable selection of F.*

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

Lemma 2.7. *Suppose that (Ξ, Σ, μ) is a complete σ-finite measure space, $\mathcal{K}$ is a closed nonempty subset in $\widetilde{Z}$ and $\varphi(\cdot)$ is a Σ-measurable map from Ξ to $\widetilde{Z}$. Then the projection map $\xi \rightsquigarrow \Pi_{\mathcal{K}}(\varphi(\xi))$ is Σ-measurable. If $\Pi_{\mathcal{K}}(\varphi(\xi)) \neq \emptyset$ for all $\xi \in \Xi$, then there exists a Σ-measurable, $\widetilde{Z}$-valued selection $\psi(\cdot)$ such that $|\psi(\xi) - \varphi(\xi)|_{\widetilde{Z}} = dist\,(\varphi(\xi), \mathcal{K})$, μ-a.e.*

At last, let us recall some results concerning convex cones.

Definition 2.6. *For a cone $\mathcal{K}$ in Z, the convex closed cone $\mathcal{K}^- = \{\xi \in Z^* | \xi(z) \leq 0 \text{ for all } z \in \mathcal{K}\}$ is called the dual cone of $\mathcal{K}$.*

Lemma 2.8. *[16, Lemma 2.4] Let $m \in \mathbb{N}$. Let $\mathcal{K}_1, \cdots, \mathcal{K}_m$ be convex cones in $\widetilde{Z}$ and $\displaystyle\bigcap_{j=1}^{m} \mathrm{int}\mathcal{K}_j \neq \emptyset$. Then for any convex cone $\mathcal{K}_0$ such that $\mathcal{K}_0 \bigcap \Big(\displaystyle\bigcap_{j=1}^{m} \mathcal{K}_j \Big) \neq \emptyset$, we have $\Big(\displaystyle\bigcap_{j=0}^{m} \mathcal{K}_j \Big)^- = \displaystyle\sum_{j=0}^{m} \mathcal{K}_j^-$.*

Definition 2.7. *We call $\mathcal{K}$ a nonempty closed polyhedron in Z if for some $n \in \mathbb{N}$, $\{z_1^*, \cdots, z_n^*\} \subset Z^* \setminus \{0\}$ and $\{b_1, \cdots, b_n\} \subset \mathbb{R}$,*

$$\mathcal{K} \overset{\triangle}{=} \{y \in \widetilde{Z} |\ \langle y_j, y \rangle_{\widetilde{Z}} + b_j \leq 0,\ \forall\ j = 1, \cdots, n\}.$$

Lemma 2.9. *[16, Lemma 2.5] Let $\widehat{Z}$ be a Hilbert space. Let $\mathcal{K}$ be a nonempty closed polyhedron in $\widehat{Z}$. Then, for any $0 \neq \xi \in \widehat{Z}$ such that $\sup_{y \in \mathcal{K}} \langle \xi, y \rangle_{\widehat{Z}} < +\infty$, this supremum is attained at some $\bar{y} \in \partial\mathcal{K}$. Furthermore, $\xi \in \sum_{j \in In(\bar{y})} \mathbb{R}^+ y_j$, where*

$$In(\bar{y}) \overset{\triangle}{=} \big\{ j \in \{1, \cdots, n\} | \langle y_j, y \rangle_{\widehat{Z}} + b_j = 0 \big\}$$

and

$$\mathbb{R}^+ y_j \overset{\triangle}{=} \big\{ \alpha y_j | \alpha > 0 \big\}.$$

Lemma 2.10. *Let $M_0, M_1, \dots, M_n$ be nonempty convex subsets of Z such that M_j is open for all $j \in \{1, \cdots, n\}$. Then*

$$M_0 \cap M_1 \cap \ldots \cap M_n = \emptyset \tag{2.5}$$

if and only if there are $z_0^*, z_1^*, \cdots, z_n^* \in Z^*$, not vanishing simultaneously, such that

$$z_0^* + z_1^* + \cdots + z_n^* = 0, \quad \inf_{z \in M_0} z_0^*(z) + \inf_{z \in M_1} z_1^*(z) + \cdots + \inf_{z \in M_n} z_n^*(z) \geq 0. \tag{2.6}$$

Furthermore, if (2.6) holds true and for some $j \in \{0, \dots, n\}$ there is a nonempty cone $\mathcal{K}_j \subset Z$ and $z_j \in Z$ such that $z_j + \mathcal{K}_j \subset M_j$, then $-z_j^* \in \mathcal{K}_j^-$.

Proof of the above lemma can be found in [13].

3. First order necessary conditions

This section is devoted to establishing a first order necessary optimality condition for **Problem (OP)**. Let us first impose the following assumptions:

(AS2) *For a.e. $(t, \omega) \in [0, T] \times \Omega$, the functions $a(t, \cdot, \cdot, \omega) : H \times H_1 \to H$ and $b(t, \cdot, \cdot, \omega) : H \times H_1 \to \mathcal{L}_2$ are differentiable, and $(a_x(t, x, u, \omega), a_u(t, x, u, \omega))$ and $(b_x(t, x, u, \omega), b_u(t, x, u, \omega))$ are uniformly continuous with respect to $x \in H$ and $u \in U$. For any $p \geq 1$, there exists a nonnegative $\eta \in L^2_{\mathbb{F}}(0, T; \mathbb{R})$ such that for a.e. $(t, \omega) \in [0, T] \times \Omega$ and for all $x \in H$ and $u \in H_1$,*

$$\begin{cases} |a(t, 0, u, \omega)|_H + |b(t, 0, u, \omega)|_{\mathcal{L}_2} \leq C(\eta(t, \omega) + |u|_{H_1}), \\ |a_x(t, x, u, \omega)|_{\mathcal{L}(H)} + |a_u(t, x, u, \omega)|_{\mathcal{L}(H_1; H)} + |b_x(t, x, u, \omega)|_{\mathcal{L}(H; \mathcal{L}_2)} + |b_u(t, x, u, \omega)|_{\mathcal{L}(H_1; \mathcal{L}_2)} \leq C. \end{cases}$$

(AS3) *The functional $h(\cdot, \omega) : H \to \mathbb{R}$ is differentiable, $\mathbb{P}$-a.s., and there exists an $\eta \in L^2_{\mathcal{F}_T}(\Omega)$ such that for any $x, \tilde{x} \in H$,*

$$\begin{cases} |h(x, \omega)| \leq C(\eta(\omega)^2 + |x|_H^2), \qquad |h_x(0, \omega)|_H \leq C\eta(\omega), \quad \mathbb{P}\text{-a.s.}, \\ |h_x(x, \omega) - h_x(\tilde{x}, \omega)|_H \leq C|x - \tilde{x}|_H, \quad \mathbb{P}\text{-a.s.} \end{cases}$$

(AS4) *For $j = 0, \cdots, n$, the functional $g^j : H \to \mathbb{R}$ is differentiable, and for any $x, \tilde{x} \in H$,*

$$|g^j(x)| \leq C(1 + |x|_H^2), \qquad |g_x^j(x) - g_x^j(\tilde{x})|_H \leq C|x - \tilde{x}|_H.$$

Remark 3.1. (AS2) *is a condition about the regularity of a and b. It is used to compute the Taylor expansion of the cost functional with respect to the control u. The Fréchet differentiability can be relaxed if one assume that the semigroup $\{S(t)\}_{g \geq 0}$ has some smoothing effect. In this paper, we purpose to present the key idea in a simple way and do not consider this case.*

Remark 3.2. *Typical examples fulfill* **(AS3)** *and* **(AS4)** *are quadratic functional. For instance, $h(x, \omega) = \eta(\omega)^2 + |x|_H^2$ and $g^j(x) = |x|_H^2 - 1$ ($j = 0, \cdots, n$) for $x \in H$.*

Let Φ be a set-valued stochastic process satisfying

1. Φ is $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted;
2. for a.e. $(t, \omega) \in [0, T] \times \Omega$, $\Phi(t, \omega)$ is a nonempty closed convex cone in H_1;
3. $\Phi(t, \omega) \subset T_U^b(\bar{u}(t, \omega))$, for a.e. $(t, \omega) \in [0, T] \times \Omega$.

Let

$$\mathcal{T}_\Phi(\bar u) \triangleq \Big\{ u(\cdot) \in L^2_{\mathbb F}(0,T;H_1) \,\Big|\, u(t,\omega) \in \Phi(t,\omega), \text{ a.e. } (t,\omega) \in [0,T]\times\Omega \Big\}.$$

Clearly, $\mathcal{T}_\Phi(\bar u)$ is a closed convex cone in $L^2_{\mathbb F}(0,T;H_1)$. Since $0 \in \mathcal{T}_\Phi(\bar u)$, $\mathcal{T}_\Phi(\bar u)$ is nonempty. By Lemma 2.2, we can choose $\Phi(t,\omega) = \mathcal{C}_U(\bar u(t,\omega))$. However, in general, there may exist a $\Phi(t,\omega)$ as above such that $\mathcal{C}_U(\bar u(t,\omega)) \subsetneq \Phi(t,\omega) \subset T^b_U(\bar u(t,\omega))$.

For φ equal to a, b, f, g or h, write

$$\varphi_1[t] = \varphi_x(t,\bar x(t),\bar u(t)), \quad \varphi_2[t] = \varphi_u(t,\bar x(t),\bar u(t)).$$

Consider the following linearized stochastic control system:

$$\begin{cases} dx_1(t) = \big(Ax_1(t) + a_1[t]x_1(t) + a_2[t]u_1(t)\big)dt + \big(b_1[t]x_1(t) + b_2[t]u_1(t)\big)dW(t) \text{ in } (0,T], \\ x_1(0) = \nu_1. \end{cases} \tag{3.1}$$

It is a classical result that, under **(AS1)**, for any $u_1 \in \mathcal{T}_\Phi(\bar u)$ and $\nu_1 \in T^b_{\mathcal K}(\bar x_0)$, (3.1) admits a unique solution $x_1(\cdot) \in L^2_{\mathbb F}(\Omega; C([0,T];H))$ (e.g.[7, Chapter 6]).

By Lemma 2.2, $\mathcal{T}_\Phi(\bar u) \subset T^b_{\mathcal{U}^2}(\bar u)$. For any $\varepsilon > 0$, choose $\nu_1^\varepsilon \in H$ and $v^\varepsilon \in L^2_{\mathbb F}(0,T;H_1)$ such that

$$\nu_0^\varepsilon \triangleq \bar\nu_0 + \varepsilon\nu_1^\varepsilon \in \mathcal V, \quad u^\varepsilon \triangleq \bar u + \varepsilon v^\varepsilon \in \mathcal{U}^2$$

and

$$\nu_1^\varepsilon \to \nu_1 \text{ in } H \text{ and } v^\varepsilon \to u_1 \text{ in } L^2_{\mathbb F}(0,T;H_1) \text{ as } \varepsilon \to 0^+.$$

Let $x^\varepsilon(\cdot)$ be the solution of (1.1) corresponding to the control $u^\varepsilon(\cdot)$ and the initial datum ν_0^ε, and put

$$\delta x^\varepsilon(\cdot) = x^\varepsilon(\cdot) - \bar x(\cdot), \qquad r_1^\varepsilon(\cdot) \triangleq \frac{\delta x^\varepsilon(\cdot) - \varepsilon x_1(\cdot)}{\varepsilon}. \tag{3.2}$$

We have the following results:

Lemma 3.1. *If* **(AS1)**–**(AS2)** *hold, then for* $p \ge 2$,

$$|x_1|_{L^\infty_{\mathbb F}(0,T;L^p(\Omega;H))} \le C\big(|\nu_1|_H + |u_1|_{L^p_{\mathbb F}(\Omega;L^2(0,T;H))}\big), \tag{3.3}$$

$$|\delta x^\varepsilon|_{L^\infty_{\mathbb F}(0,T;L^p(\Omega;H))} = O(\varepsilon), \tag{3.4}$$

and

$$\lim_{\varepsilon\to 0^+} |r_1^\varepsilon|_{L^\infty_{\mathbb F}(0,T;L^p(\Omega;H))} = 0. \tag{3.5}$$

Proof of Lemma 3.1 is provided in Appendix A.

Next, we give a result which is very useful to get the first order pointwise necessary condition.

Lemma 3.2. *Let* $\tilde u(\cdot) \in \mathcal{U}^{\nu_0}_{ad}$, *and* $F : [0,T]\times\Omega \to H_1$ *be an* $\mathbb F$*-adapted process such that*

$$\mathbb E \int_0^T \langle F(t), v(t)\rangle_{H_1} dt \le 0, \quad \forall\, v(\cdot) \in \mathcal{C}_{\mathcal{U}^{\nu_0}_{ad}}(\tilde u(\cdot)).$$

Then, for a.e. $(t,\omega) \in [0,T]\times\Omega$, $\langle F(t,\omega), v\rangle_{H_1} \le 0$, $\forall\, v \in \mathcal{C}_U(\tilde u(t,\omega))$.

Proof of Lemma 3.2 is postponed to Appendix C.

Lemma 3.3. *For each bounded linear functional* Λ *on* $L^2_{\mathbb F}(\Omega; C([0,T];H))$, *there exists a process* $\psi \in L^2_{\mathbb F}(\Omega; BV_0([0,T];H))$ *such that*

$$\Lambda(z(\cdot)) = \mathbb E \int_0^T \langle z(t), d\psi(t)\rangle_H\,, \quad \forall\, z(\cdot) \in L^2_{\mathbb F}(\Omega; C([0,T];H)), \tag{3.6}$$

and

$$|\Lambda|_{L^2_{\mathbb F}(\Omega;C([0,T];H))^*} = |\psi|_{L^2_{\mathbb F}(\Omega;BV([0,T];H))}. \tag{3.7}$$

Proof of Lemma 3.3 is given in Appendix D.

Let $\mathcal{T}_{\mathcal{K}}(\bar{\nu}_0)$ be a nonempty closed convex cone contained in $T^b_{\mathcal{K}}(\bar{\nu}_0)$. Put

$$\mathcal{X}_{(1)} \overset{\triangle}{=} \big\{x_1(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))\,\big|\,x_1(\cdot) \text{ solves (3.1) with } u_1\in\mathcal{T}_{\Phi}(\bar{u}) \text{ and } \nu_1\in\mathcal{T}_{\mathcal{K}}(\bar{\nu}_0)\big\}, \tag{3.8}$$

$$\mathcal{I}_0(\bar{x}) \overset{\triangle}{=} \{t\in[0,T]\,|\,\mathbb{E}g^0(\bar{x}(t))=0\}, \tag{3.9}$$

$$\mathcal{I}(\bar{x}) = \big\{j\in\{1,\cdots,n\}\,|\,\mathbb{E}g^j(\bar{x}(T))=0\big\},$$

$$\mathcal{G}^0_{(1)} = \big\{z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))\,\big|\,\mathbb{E}\left\langle g^0_x(\bar{x}(t)),z(t)\right\rangle_H<0,\ \forall\,t\in\mathcal{I}_0(\bar{x})\big\}, \tag{3.10}$$

$$\mathcal{G}^j_{(1)} \overset{\triangle}{=} \big\{z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))\,\big|\,\mathbb{E}\left\langle g^j_x(\bar{x}(T)),z(T)\right\rangle_H<0\big\},\qquad \forall j\in\mathcal{I}(\bar{x}), \tag{3.11}$$

$$\mathcal{G}_{(1)} \overset{\triangle}{=} \bigcap_{j\in\mathcal{I}(\bar{x})}\mathcal{G}^j_{(1)}, \tag{3.12}$$

$$\mathcal{G}^j_{(1)}(T) \overset{\triangle}{=} \big\{\zeta\in L^2_{\mathcal{F}_T}(\Omega;H)\,\big|\,\mathbb{E}\left\langle g^j_x(\bar{x}(T)),\zeta\right\rangle_H<0\big\},\quad j\in\mathcal{I}(\bar{x}), \tag{3.13}$$

$$\mathcal{G}_{(1)}(T) \overset{\triangle}{=} \bigcap_{j\in\mathcal{I}(\bar{x})}\mathcal{G}^j_{(1)}(T), \tag{3.14}$$

$$\mathcal{H}_{(1)} \overset{\triangle}{=} \{z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))|\,\mathbb{E}\langle h_x(\bar{x}(T)),z(T)\rangle_H<0\},$$

and

$$\mathcal{H}_{(1)}(T) \overset{\triangle}{=} \{z\in L^2_{\mathcal{F}_T}(\Omega;H)\,|\,\mathbb{E}\langle h_x(\bar{x}(T)),z\rangle_H<0\}.$$

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

If $\mathcal{I}_0(\bar{x})=\emptyset$ (*resp.* $\mathcal{I}(\bar{x})=\emptyset$), we set $\mathcal{G}^0_{(1)}=L^2_{\mathbb{F}}(\Omega;C([0,T];H))$ (*resp.* $\mathcal{G}_{(1)}=L^2_{\mathbb{F}}(\Omega;C([0,T];H))$). If $h_x(\bar{x}(T))=0$, $\mathbb{P}$-a.s., then $\mathcal{H}_{(1)}=\emptyset$ and $\mathcal{H}_{(1)}(T)=\emptyset$.

Define a map $\Gamma: L^2_{\mathbb{F}}(\Omega;C([0,T];H))\to L^2_{\mathcal{F}_T}(\Omega;H)$ as

$$\Gamma(z)=z(T),\quad \forall\,z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H)). \tag{3.15}$$

Denote by Γ^* the adjoint operator of Γ. Clearly, Γ is surjective. From (3.11) to (3.14), we see that

$$\mathcal{G}^j_{(1)}(T)=\Gamma(\mathcal{G}^j_{(1)}),\qquad j\in\mathcal{I}(\bar{x}),\qquad \mathcal{G}_{(1)}(T)=\Gamma(\mathcal{G}_{(1)}). \tag{3.16}$$

If $\mathcal{G}^0_{(1)}$ and $\mathcal{G}_{(1)}$ are nonempty, then

$$\mathrm{cl}\mathcal{G}^0_{(1)}=\big\{z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))\,\big|\,\mathbb{E}\left\langle g_x(\bar{x}(t)),z(t)\right\rangle_H\le 0,\ \forall\,t\in\mathcal{I}_0(\bar{x})\big\},$$

and

$$\mathrm{cl}\mathcal{G}_{(1)}=\big\{z(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))\,\big|\,\mathbb{E}\left\langle g^j_x(\bar{x}(T)),z(T)\right\rangle_H\le 0,\ \forall\,j\in\mathcal{I}(\bar{x})\big\}.$$

Lemma 3.4. *$\mathcal{G}^0_{(1)}$ is an open convex cone in $L^2_{\mathbb{F}}(\Omega;C([0,T];H))$.*

Proof. Clearly, $\mathcal{G}^0_{(1)}$ is a cone. It is sufficient to prove that it is open.

Let $z(\cdot)\in\mathcal{G}^0_{(1)}$. Since $\bar{x}(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))$, $\mathcal{I}_0(\bar{x})$ is a compact subset of $[0,T]$. This, together with the fact that $\mathbb{E}\left\langle g_x(\bar{x}(\cdot)),z(\cdot)\right\rangle_H$ is continuous with respect to t, implies that there exists a constant $\rho>0$ such that

$$\mathbb{E}\left\langle g_x(\bar{x}(t)),z(t)\right\rangle_H<-\rho,\qquad \forall\,t\in\mathcal{I}_0(\bar{x}).$$

Let

$$\delta = \frac{\rho}{2|g_x(\bar{x}(\cdot))|_{L^\infty_{\mathbb{F}}(0,T;L^2(\Omega;H))}}.$$

Then for any $\eta \in L^2_{\mathbb{F}}(\Omega; C([0,T];H))$ with $\|\eta\|_{L^2_{\mathbb{F}}(\Omega;C([0,T];H))} \leq \delta$,

$$\mathbb{E}\left\langle g_x(\bar{x}(t)), z(t)+\eta(t)\right\rangle_H < -\frac{\rho}{2}, \quad \forall\, t \in \mathcal{I}_0(\bar{x}).$$

This proves that $z \in \mathrm{int}\mathcal{G}^0_{(1)}$. $\square$

Now we introduce the first order adjoint equation for (3.1):

$$\begin{cases} dy(t) = -\big(A^*y(t) + a_1[t]^*y(t) + b_1[t]^*Y(t)\big)dt + d\psi(t) + Y(t)dW(t) \quad \text{in } [0,T), \\ y(T) = y_T, \end{cases} \tag{3.17}$$

where $y_T \in L^2_{\mathcal{F}_T}(\Omega; H)$ and $\psi \in L^2_{\mathbb{F}}(\Omega; BV_0([0,T];H))$.

Since neither the usual natural filtration condition nor the quasi-left continuity is assumed for the filtration $\mathbb{F}$ in this paper, one cannot apply the existence results for mild or weak solution of infinite dimensional BSEEs (e.g. [20, 32]) to obtain the well-posedness of the equation (3.17). Thus, we use the notion of transposition solution here. To this end, consider the following (forward) SEE:

$$\begin{cases} d\phi(s) = \big(A\phi(s) + f_1(s)\big)ds + f_2(s)dW(s) \quad \text{in } (t,T], \\ \phi(t) = \eta, \end{cases} \tag{3.18}$$

where $t \in [0,T]$, $f_1 \in L^1_{\mathbb{F}}(t,T;L^2(\Omega;H))$, $f_2 \in L^2_{\mathbb{F}}(t,T;\mathcal{L}_2)$, $\eta \in L^2_{\mathcal{F}_t}(\Omega;H)$ (See [7, Chapter 6] for the well-posedness of (3.18) in the sense of mild solution). We now introduce the following notion.

Definition 3.1. *We call $(y(\cdot), Y(\cdot)) \in D_{\mathbb{F}}([0,T];L^2(\Omega;H)) \times L^2_{\mathbb{F}}(0,T;\mathcal{L}_2)$ a transposition solution of (3.17) if for any $t \in [0,T]$, $f_1(\cdot) \in L^1_{\mathbb{F}}(t,T;L^2(\Omega;H))$, $f_2(\cdot) \in L^2_{\mathbb{F}}(t,T;\mathcal{L}_2)$, $\eta \in L^2_{\mathcal{F}_t}(\Omega;H)$ and the corresponding solution $\phi \in L^2_{\mathbb{F}}(\Omega;C([t,T];H))$ to the equation (3.18), we have*

$$\begin{aligned} &\mathbb{E}\big\langle \phi(T), y(T)\big\rangle_H + \mathbb{E}\int_t^T \big\langle \phi(s), a_1[s]^*y(s) + b_1[s]^*Y(s))\big\rangle_H ds \\ &= \mathbb{E}\big\langle \eta, y(t)\big\rangle_H + \mathbb{E}\int_t^T \big\langle f_1(s), y(s)\big\rangle_H ds + \mathbb{E}\int_t^T \big\langle f_2(s), Y(s)\big\rangle_{\mathcal{L}_2} ds + \mathbb{E}\int_t^T \big\langle \phi(s), d\psi(s)\big\rangle_H. \end{aligned} \tag{3.19}$$

Lemma 3.5. *Assume that* **(AS1)**–**(AS2)** *hold and $\psi \in L^2_{\mathbb{F}}(\Omega; BV_0([0,T];H))$. Then the equation (3.17) admits a unique transposition solution $(y, Y) \in D_{\mathbb{F}}([0,T];L^2(\Omega;H)) \times L^2_{\mathbb{F}}(0,T;\mathcal{L}_2)$.*

If $\psi = 0$ and $W(\cdot)$ is a one dimensional Brownian motion, Lemma 3.5 is proved in [29, Chapter 3]. The proof for the case $\psi \neq 0$ is similar. We only give a sketch in Appendix E.

Define the Hamiltonian

$$\mathbb{H}(t,x,u,p,q,\omega) \overset{\triangle}{=} \langle p, a(t,x,u,\omega)\rangle_H + \langle q, b(t,x,u,\omega)\rangle_{\mathcal{L}_2}, \tag{3.20}$$

where $(t,x,u,p,q,\omega) \in [0,T] \times H \times H_1 \times H \times \mathcal{L}_2 \times \Omega$.

Now we state a first order necessary optimality condition in the integral form.

Theorem 3.1. *Let* **(AS1)**–**(AS4)** *hold and $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{\nu}_0)$ be an optimal triple for* **Problem (OP)**. *If $\mathbb{E}|g^0_x(\bar{x}(t))|_H \neq 0$ for any $t \in \mathcal{I}_0(\bar{x})$, then there exist $\lambda_0 \in \{0,1\}$, $\lambda_j \geq 0$ for $j \in \mathcal{I}(\bar{x})$ and $\psi \in \big(\mathcal{G}^0_{(1)}\big)^-$ with $\psi(0) = 0$ satisfying*

$$\lambda_0 + \sum_{j\in\mathcal{I}(\bar{x})} \lambda_j + |\psi|_{L^2_{\mathbb{F}}(\Omega;BV(0,T;H))} \neq 0, \tag{3.21}$$

such that the corresponding transposition solution $(y(\cdot), Y(\cdot))$ of the first order adjoint equation (3.17) with $y(T) = -\lambda_0 h_x(\bar{x}(T)) - \sum_{j\in\mathcal{I}(\bar{x})} \lambda_j g_x^j(\bar{x}(T))$ *verifies that*

$$\mathbb{E}\left\langle y(0), \nu\right\rangle_H + \mathbb{E}\int_0^T \left\langle \mathbb{H}_u[t], v(t)\right\rangle_{H_1} dt \leq 0, \;\; \forall\, \nu \in \mathcal{T}_{\mathcal{K}}(\bar{\nu}_0), \;\; \forall\, v(\cdot) \in \mathcal{T}_{\Phi}(\bar{u}), \tag{3.22}$$

where $\mathbb{H}_u[t] = \mathbb{H}_u(t, \bar{x}(t), \bar{u}(t), y(t), Y(t), \omega)$*. In addition, if* $\mathcal{G}^0_{(1)} \cap \mathcal{X}_{(1)} \cap \mathcal{G}_{(1)} \neq \emptyset$*, the above holds with* $\lambda_0 = 1$*.*

Remark 3.3. *If $\bar{u}$ takes an isolated point of U in a positive measure set of $[0,T]\times\Omega$, then* (3.22) *does not give us any information about the optimal control at these point since ar these points, $v(t,\omega) = 0$ for $v(\cdot) \in \mathcal{T}_{\Phi}(\bar{u})$. This is a drawback of Theorem 3.1. In handle such case, one should employ the spike variation technique. The cost is that one should use two adjoint equations. More details can be found in [29, 30, 31].*

Remark 3.4. *In Theorem 3.1, we assume that $\mathbb{E}|g_x^0(\bar{x}(t))|_H \neq 0$ for any $t \in \mathcal{I}_0(\bar{x})$. This can be verified by many concrete g^0. For example, let $g^0(\eta) = |\eta|_H^2 - 1$ for any $\eta \in H$. If $t \in \mathcal{I}_0(\bar{x})$, then $\mathbb{E}|\bar{x}(t)|_H^2 = 1$. Therefore, $\mathbb{E}|g_x^0(\bar{x}(t))|_H = 2\mathbb{E}|\bar{x}(t)|_H \neq 0$.*

Proof of Theorem 3.1. We first claim that

$$\mathcal{X}_{(1)} \cap \mathcal{G}^0_{(1)} \cap \mathcal{G}_{(1)} \cap \mathcal{H}_{(1)} = \emptyset. \tag{3.23}$$

If this is not the case, then there would exist $\tilde{x}_1(\cdot) \in \mathcal{X}_{(1)} \cap \mathcal{G}^0_{(1)} \cap \mathcal{G}_{(1)}$ such that

$$\mathbb{E}\left\langle h_x(\bar{x}(T)), \tilde{x}_1(T)\right\rangle_H < 0. \tag{3.24}$$

Let $\tilde{\nu}_1 \in \mathcal{T}_{\mathcal{K}}(\bar{\nu}_0)$ be the initial datum and $\tilde{u}_1(\cdot) \in \mathcal{T}_{\Phi}(\bar{u}(\cdot))$ the control corresponding to $\tilde{x}_1(\cdot)$. Let $\mu^\varepsilon \in H$ with $|\mu^\varepsilon| = o(\varepsilon)$ and $\eta^\varepsilon(\cdot) \in L^2_{\mathbb{F}}(0,T;H_1)$ with $|\eta^\varepsilon|_{L^2_{\mathbb{F}}(0,T;H_1)} = o(\varepsilon)$ be such that

$$\nu_0^\varepsilon \overset{\triangle}{=} \bar{\nu}_0 + \varepsilon\tilde{\nu}_1 + \mu^\varepsilon \in \mathcal{V}, \qquad u^\varepsilon(\cdot) \overset{\triangle}{=} \bar{u}(\cdot) + \varepsilon\tilde{u}_1(\cdot) + \eta^\varepsilon(\cdot) \in \mathcal{U}^2.$$

Let $x^\varepsilon(\cdot)$ be the solution of the control system (1.1) with the initial datum ν_0^ε and the control $u^\varepsilon(\cdot)$.

Since $\tilde{x}_1(\cdot) \in \mathcal{G}^0_{(1)}$, we know that $\mathbb{E}\left\langle g_x^0(\bar{x}(\cdot)), \tilde{x}_1(\cdot)\right\rangle_H$ is continuous with respect to t. This, together with the compactness of $\mathcal{I}^0(\bar{x})$, implies that there exists $\rho_0 > 0$ such that

$$\mathbb{E}\left\langle g_x^0(\bar{x}(t)), \tilde{x}_1(t)\right\rangle_H < -\rho_0 \text{ for every } t \in \mathcal{I}^0(\bar{x}).$$

Moreover, there exists $\delta > 0$ (independent of $t \in \mathcal{I}^0(\bar{x})$) such that

$$\mathbb{E}\left\langle g_x^0(\bar{x}(s)), \tilde{x}_1(s)\right\rangle_H < -\frac{\rho_0}{2}, \forall\, s \in (t-\delta, t+\delta) \cap [0,T] \text{ and } t \in \mathcal{I}^0(\bar{x}).$$

By Lemma 3.1, there is an $\varepsilon_0 > 0$ such that for every $\varepsilon \in [0, \varepsilon_0]$,

$$\begin{aligned} \mathbb{E}g^0(x^\varepsilon(s)) &= \mathbb{E}g(\bar{x}(s)) + \varepsilon\mathbb{E}\left\langle g_x(\bar{x}(s)), \tilde{x}_1(s)\right\rangle_H + o(\varepsilon) \\ &\leq \varepsilon\,\mathbb{E}\left\langle g_x(\bar{x}(s)), \tilde{x}_1(s)\right\rangle_H + o(\varepsilon) \\ &< -\frac{\varepsilon\rho_0}{4} < 0, \quad \forall\, s \in (t-\delta, t+\delta) \cap [0,T], \;\; t \in \mathcal{I}^0(\bar{x}). \end{aligned} \tag{3.25}$$

Since $\mathcal{I}_\delta^c \overset{\triangle}{=} [0,T] \setminus \bigcup_{t\in\mathcal{I}^0(\bar{x})}(t-\delta, t+\delta)$ is compact, there exist $\rho_1 > 0$ and $\varepsilon_1 > 0$ such that for any $\varepsilon \in [0, \varepsilon_1]$,

$$
\begin{aligned}
\mathbb{E}g(x^\varepsilon(t)) &= \mathbb{E}g(\bar{x}(t)) + \varepsilon\mathbb{E}\left\langle g_x(\bar{x}(t)), \tilde{x}_1(t)\right\rangle_H + o(\varepsilon)\\
&< -\rho_1 + \varepsilon\mathbb{E}\left\langle g_x(\bar{x}(t)), \tilde{x}_1(t)\right\rangle_H + o(\varepsilon)\\
&< -\frac{\rho_1}{2} < 0, \quad \forall\, t \in \mathcal{I}_\delta^c.
\end{aligned}
\tag{3.26}
$$

By (3.25) and (3.26), $x^\varepsilon(\cdot)$ satisfies the state constraint (1.3) for $\varepsilon < \min\{\varepsilon_0, \varepsilon_1\}$.

Since $\tilde{x}_1(T) \in \mathcal{G}_{(1)}(T)$, $\mathbb{E}\langle g_x^j(\bar{x}(T)), \tilde{x}_1(T)\rangle_H < 0$ for every $j \in \mathcal{I}(\bar{x})$. Similar to the proof of (3.26), for every sufficiently small ε, $x^\varepsilon(\cdot)$ satisfies the final state constraint (1.4), and $(x^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in \mathcal{P}_{ad}$. Following (3.24), there exists $\rho_2 > 0$ such that for all sufficiently small ε,

$$
\begin{aligned}
\mathbb{E}h(x^\varepsilon(T)) &= \mathbb{E}h(\bar{x}(T)) + \varepsilon\mathbb{E}\left\langle h_x(\bar{x}(T)), \tilde{x}_1(T)\right\rangle_H + o(\varepsilon)\\
&< \mathbb{E}h(\bar{x}(T)) - \varepsilon\rho_2 + o(\varepsilon) < \mathbb{E}h(\bar{x}(T)),
\end{aligned}
$$

contradicting the optimality of $(\bar{x}(\cdot), \bar{u}(\cdot))$. This completes the proof of (3.23).

To finish the proof, we consider three different cases.

Case 1: $\mathcal{G}^0_{(1)} \cap \mathcal{X}_{(1)} = \emptyset$.

Noting that $\mathcal{G}^0_{(1)}$ is nonempty, open and convex, and $\mathcal{X}_{(1)}$ is nonempty and convex, by the Hahn-Banach separation theorem and Lemma 3.3, there exists a nonzero $\psi(\cdot) \in L^2_{\mathbb{F}}(\Omega;\, BV_0([0,T];H))$ such that

$$
\sup_{z\in\mathcal{G}^0_{(1)}} \mathbb{E}\int_0^T \left\langle z(t), d\psi(t)\right\rangle_H \le \inf_{\tilde{z}\in\mathcal{X}_{(1)}} \mathbb{E}\int_0^T \left\langle \tilde{z}(t), d\psi(t)\right\rangle_H .
$$

Since $\mathcal{G}^0_{(1)}$ and $\mathcal{X}_{(1)}$ are cones,

$$
0 = \sup_{z\in\mathcal{G}^0_{(1)}} \mathbb{E}\int_0^T \left\langle z(t), d\psi(t)\right\rangle_H = \inf_{\tilde{z}\in\mathcal{X}_{(1)}} \mathbb{E}\int_0^T \left\langle \tilde{z}(t), d\psi(t)\right\rangle_H .
$$

Therefore, $\psi \in \left(\mathcal{G}^0_{(1)}\right)^-$ and $-\psi \in \left(\mathcal{X}_{(1)}\right)^-$. Consequently, for all $z_1(\cdot) \in \mathcal{X}_{(1)}$,

$$
\mathbb{E}\int_0^T \left\langle z_1(t), d\psi(t)\right\rangle_H \ge 0. \tag{3.27}
$$

Furthermore, it follows from the definition of the transposition solution to (3.17) that for every x_1 solving (3.1) with $u_1 \in \mathcal{T}_\Phi(\bar{u})$ and $\nu_1 \in T^b_{\mathcal{K}}(\bar{x}_0)$,

$$
\begin{aligned}
&\mathbb{E}\left\langle y(T), x_1(T)\right\rangle_H - \left\langle y(0), \nu_1\right\rangle_H\\
&= \mathbb{E}\int_0^T \Big(\left\langle y(t), a_1[t]x_1(t)\right\rangle_H + \left\langle y(t), a_2[t]u_1(t)\right\rangle_H - \left\langle a_1[t]^*y(t), x_1(t)\right\rangle_H - \left\langle b_1[t]^*Y(t), x_1(t)\right\rangle_H\\
&\qquad + \left\langle Y(t), b_1[t]x_1(t)\right\rangle_{\mathcal{L}_2} + \left\langle Y(t), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2}\Big)dt + \mathbb{E}\int_0^T \left\langle x_1(t), d\psi(t)\right\rangle_H\\
&= \mathbb{E}\int_0^T \Big(\left\langle y(t), a_2[t]u_1(t)\right\rangle_H + \left\langle Y(t), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2}\Big)dt + \mathbb{E}\int_0^T \left\langle x_1(t), d\psi(t)\right\rangle_H .
\end{aligned}
\tag{3.28}
$$

Set $\lambda_0 = 0$, $\lambda_j = 0$, $j \in \mathcal{I}(\bar{x})$ and $y(T) = 0$. Then, (3.21) holds and (3.22) follows from (3.27) and (3.28).

Case 2: $\mathcal{G}^0_{(1)} \cap \mathcal{X}_{(1)} \neq \emptyset$ and $\mathcal{G}^0_{(1)} \cap \mathcal{X}_{(1)} \cap \mathcal{G}_{(1)} = \emptyset$.

If $\mathcal{G}_{(1)} = \emptyset$, we claim that for each $j \in \mathcal{I}(\bar{x})$, there exists $\lambda_j \ge 0$ such that

$$\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j>0,\qquad \sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g_x^j(\bar{x}(T))=0. \tag{3.29}$$

Indeed, if there is a $j_0\in\mathcal{I}(\bar{x})$ such that $g_x^{j_0}(\bar{x}(T))=0$, then we can take $\lambda_{j_0}=1$ and $\lambda_j=0$ for all $j\in\mathcal{I}(\bar{x})\setminus\{j_0\}$. In this context, (3.29) hold.

If $g_x^j(\bar{x}(T))\neq 0$ for all $j\in\mathcal{I}(\bar{x})$, then $\mathcal{G}_{(1)}^j\neq\emptyset$ for all $j\in\mathcal{I}(\bar{x})$ since Γ is surjective (recall (3.11) for the definition of $\mathcal{G}_{(1)}^j$). From (3.16), we find that $\mathcal{G}_{(1)}^j(T)\neq\emptyset$ for all $j\in\mathcal{I}(\bar{x})$. On the other hand, since $\mathcal{G}_{(1)}=\bigcap_{j\in\mathcal{I}(\bar{x})}\mathcal{G}_{(1)}^j=\emptyset$, by (3.16), we get that $\mathcal{G}_{(1)}(T)=\bigcap_{j\in\mathcal{I}(\bar{x})}\mathcal{G}_{(1)}^j(T)=\emptyset$. Then one can find a $j_0\in\mathcal{I}(\bar{x})$ and a subset $\mathcal{I}_{j_0}\subset\mathcal{I}(\bar{x})\setminus\{j_0\}$ such that $\bigcap_{j\in\mathcal{I}_{j_0}}\mathcal{G}_{(1)}^j(T)\neq\emptyset$ and

$$\mathcal{E}_T^{(1,j_0)}\bigcap\Big(\bigcap_{j\in\mathcal{I}_{j_0}}\mathcal{G}_{(1)}^j(T)\Big)=\emptyset.$$

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

$$\sup_{\eta\in\mathcal{E}_T^{(1,j_0)}}\mathbb{E}\langle\xi,\eta\rangle_H\leq\inf_{\eta\in\bigcap_{j\in\mathcal{I}_{j_0}}\mathcal{G}_{(1)}^j(T)}\mathbb{E}\langle\xi,\eta\rangle_H.$$

Noting that $\mathcal{G}_{(1)}^j(T)$ $(j\in\mathcal{I}(\bar{x}))$ is a cone, $\xi\in\big(\mathcal{E}_T^{(1,j_0)}\big)^-$ and $-\xi\in\big(\bigcap_{j\in\mathcal{I}_{j_0}}\mathcal{G}_{(1)}^j(T)\big)^-$. By Lemma 2.8, $\xi=\lambda_{j_0}g_x^{j_0}(\bar{x}(T))$ for some $\lambda_{j_0}>0$. Further, for every $j\in\mathcal{I}_{j_0}$, there exists $\lambda_j\geq 0$ such that $-\xi=\sum\limits_{j\in\mathcal{I}_{j_0}}\lambda_j g_x^j(\bar{x}(T))$. Let $\lambda_j=0$ for $j\in\mathcal{I}(\bar{x})\setminus(\mathcal{I}_{j_0}\cup\{j_0\})$, we get (3.29).

By taking $\lambda_0=0$, $\psi=0$ and $y(T)=0$, we have (3.21) and the condition (3.22) holds trivially with $(y,Y)\equiv 0$.

If $\mathcal{G}_{(1)}\neq\emptyset$, then $\Gamma\big(\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)}\big)\cap\mathcal{G}_{(1)}(T)=\emptyset$. By the Hahn-Banach theorem, there exists a nonzero $\xi\in L^2_{\mathcal{F}_T}(\Omega;H)$ such that

$$\sup_{\alpha\in\Gamma(\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)})}\mathbb{E}\left\langle\xi,\alpha\right\rangle_H\leq\inf_{\beta\in\mathcal{G}_{(1)}(T)}\mathbb{E}\left\langle\xi,\beta\right\rangle_H.$$

Since both $\Gamma(\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)})$ and $\mathcal{G}_{(1)}(T)$ are cones,

$$0=\sup_{\alpha\in\Gamma(\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)})}\mathbb{E}\left\langle\xi,\alpha\right\rangle_H=\inf_{\beta\in\mathcal{G}_{(1)}(T)}\mathbb{E}\left\langle\xi,\beta\right\rangle_H.$$

Therefore, $\xi\in\big(\Gamma(\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)})\big)^-$ and $-\xi\in\big(\mathcal{G}_{(1)}(T)\big)^-$.

By Lemma 2.8, for each $j\in\mathcal{I}(\bar{x})$, there exists $\lambda_j\geq 0$ such that

$$\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j>0,\qquad -\xi=\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g_x^j(\bar{x}(T)).$$

Since $0\geq\mathbb{E}\left\langle\xi,\Gamma(z)\right\rangle_H$ for all $z\in\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)}$, we have that $\Gamma^*(\xi)\in\big(\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)}\big)^-$. By Lemma 2.8, there exists $\psi\in\big(\mathcal{G}_{(1)}^0\big)^-$ with $\psi(0)=0$ such that $\Gamma^*(\xi)-\psi\in\big(\mathcal{X}_{(1)}\big)^-$. Thus, for all $z(\cdot)\in\mathcal{X}_{(1)}$,

$$0\geq\mathbb{E}\left\langle\xi,z(T)\right\rangle_H-\mathbb{E}\int_0^T\left\langle z(t),d\psi(t)\right\rangle_H. \tag{3.30}$$

Let $\lambda_0=0$. Since $\xi\neq 0$, (3.21) holds. Set $y(T)=-\sum\limits_{j\in\mathcal{I}(\bar{x})}\lambda_j g_x^j(\bar{x}(T))$. By (3.28) and (3.30), we obtain (3.22).

Case 3: $\mathcal{G}_{(1)}^0\cap\mathcal{X}_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$.

In this case, it holds that $\Gamma\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)\cap\mathcal{G}_{(1)}(T)\neq\emptyset$. By (3.23),

$$\mathbb{E}\left\langle h_x(\bar{x}(T)),z(T)\right\rangle_H\geq 0,\quad \forall z(\cdot)\in\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\cap\mathcal{G}_{(1)}.$$

This yields that

$$\mathbb{E}\left\langle h_x(\bar{x}(T)),\zeta\right\rangle_H\geq 0,\quad \forall\,\zeta\in\Gamma\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)\cap\mathcal{G}_{(1)}(T).$$

Consequently,

$$-h_x(\bar{x}(T))\in\big[\Gamma\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)\cap\mathcal{G}_{(1)}(T)\big]^-.$$

By Lemma 2.8,

$$\big[\Gamma\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)\cap\mathcal{G}_{(1)}(T)\big]^-=\big[\Gamma\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)\big]^-+\big(\mathcal{G}_{(1)}(T)\big)^-.$$

Then, for each $j\in\mathcal{I}(\bar{x})$, there exists $\lambda_j\geq 0$ such that

$$\xi\triangleq\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g^j_x(\bar{x}(T))\in\big(\mathcal{G}_{(1)}(T)\big)^-$$

and that

$$-h_x(\bar{x}(T))-\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g^j_x(\bar{x}(T))\in\big[\Gamma\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)\big]^-.$$

Therefore,

$$\Gamma^*\Big(-h_x(\bar{x}(T))-\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g^j_x(\bar{x}(T))\Big)\in\big(\mathcal{G}^0_{(1)}\cap\mathcal{X}_{(1)}\big)^-=\big(\mathcal{G}^0_{(1)}\big)^-+\big(\mathcal{X}_{(1)}\big)^-.$$

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

$$\Gamma^*\Big(-h_x(\bar{x}(T))-\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g^j_x(\bar{x}(T))\Big)-\psi\in\big(\mathcal{X}_{(1)}\big)^-.$$

Set $\lambda_0=1$ and $y(T)=-h_x(\bar{x}(T))-\sum\limits_{j\in\mathcal{I}(\bar{x})}\lambda_j g^j_x(\bar{x}(T))$. Then, (3.21) holds and for all $z\in\mathcal{X}_{(1)}$,

$$0\geq-\mathbb{E}\left\langle h_x(\bar{x}(T))),z(T)\right\rangle_H-\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j\mathbb{E}\left\langle g^j_x(\bar{x}(T)),z(T)\right\rangle_H-\mathbb{E}\int_0^T\left\langle z(t),d\psi(t)\right\rangle_H.\tag{3.31}$$

Combining (3.31) with (3.28), we obtain (3.22). This completes the proof of Theorem 3.1. $\square$

Let $\Phi(t,\omega)=\mathcal{C}_U(\bar{u}(t,\omega))$, for a.e. $(t,\omega)\in[0,T]\times\Omega$ and $\mathcal{T}_{\mathcal{V}}(\bar{\nu}_0)=\mathcal{C}_{\mathcal{V}}(\bar{\nu}_0)$. From Theorem 3.1 and Lemma 3.2, it is easy to obtain the following pointwise first order necessary condition.

Theorem 3.2. *Let* **(AS1)**–**(AS4)** *hold and* $(\bar{x}(\cdot),\bar{u}(\cdot),\bar{\nu}_0)$ *be an optimal triple for* **Problem (OP)** *such that* $\mathbb{E}|g^0_x(\bar{x}(t))|_H\neq 0$ *for any* $t\in\mathcal{I}^0(\bar{x})$. *Then for* (y,Y) *as in Theorem 3.1,*

$$y(0)\in\mathcal{N}_{\mathcal{V}}(\bar{\nu}_0),\quad \mathbb{H}_u[t]\in\mathcal{N}_U(\bar{u}(t)),\ a.e.\ t\in[0,T],\ \mathbb{P}\text{-}a.s.\tag{3.32}$$

Remark 3.5. *If both the control set* U *and the initial state constraint set* $\mathcal{V}$ *are convex, then* $\mathcal{N}^C_U(\bar{u})$ *and* $\mathcal{N}^C_{\mathcal{V}}(\bar{x}_0)$ *are simply the normal cones of convex analysis.*

Remark 3.6. *Let*

$$\begin{aligned}\mathcal{H}(t,x,u,\omega)&=\mathbb{H}(t,x,u,y(t),Y(t),\omega)-\frac{1}{2}\left\langle P(t)b(t,\bar{x}(t),\bar{u}(t),\omega),b(t,\bar{x}(t),\bar{u}(t),\omega)\right\rangle_{\mathcal{L}_2}\\&\quad+\frac{1}{2}\big\langle P(t)\big(b(t,x,u,\omega)-b(t,\bar{x}(t),\bar{u}(t),\omega)\big),b(t,x,u,\omega)-b(t,\bar{x}(t),\bar{u}(t),\omega)\big\rangle_{\mathcal{L}_2},\end{aligned}$$

where $P(\cdot)$ is the first element of the solution of the second order adjoint process with respect to $(\bar{x}(\cdot),\bar{u}(\cdot),\bar{\nu}_0)$ (defined by (4.1) in Section 4). If there is no state constraint, the stochastic maximum principle (e.g. [29, 30]) says that, if $(\bar{x}(\cdot),\bar{u}(\cdot),\bar{\nu}_0)$ is an optimal triple, then

$$\mathcal{H}(t,\bar{x}(t),\bar{u}(t))=\max_{v\in U}\mathcal{H}(t,\bar{x}(t),v),\qquad a.e.\ t\in[0,T],\ \ \mathbb{P}\text{-}a.s. \tag{3.33}$$

This implies that

$$\langle \mathbb{H}_u(t,\omega),v\rangle_{H_1}\le 0,\quad \forall\, v\in\mathcal{C}_U(\bar{u}(t,\omega)),\ a.e.\ (t,\omega)\in[0,T]\times\Omega,$$

i.e., the second condition in (3.32) *holds. However, to derive* (3.33)*, one has to assume that a, b and h are C^2 with respect to the variable x. Therefore, in practice, under some usual structural assumptions on U, it is more convenient to use the condition* (3.32) *directly.*

As for the deterministic optimal control problems with state constraints, we call the first order necessary condition (3.22) *normal* if the Lagrange multiplier $\lambda_0\neq 0$. By Theorem 3.1, this is the case when $\mathcal{X}_{(1)}\cap\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$. Let us give some conditions to guarantee it. To this end, we first introduce the following equation:

$$\begin{cases} d\tilde{y}(t)=-\big(A^*\tilde{y}(t)+a_1[t]^*\tilde{y}(t)+b_1[t]^*\widetilde{Y}(t)+\alpha(t)\big)ds+\widetilde{Y}(t)dW(t) \quad \text{in } [0,T),\\ \tilde{y}(T)=0, \end{cases} \tag{3.34}$$

where $\alpha(\cdot)\in L^2_{\mathbb{F}}(0,T;H)$. The equation (3.34) is a special case of (3.17), where $d\psi(\cdot)=\alpha(\cdot)ds$.

Let us make the following assumptions:

(AAS1) $\alpha(\cdot)=0$ whenever $a_1(\cdot)^*\tilde{y}[\cdot]+b_1[\cdot]^*\widetilde{Y}(\cdot)=0$.

(AAS2) $\mathcal{C}_U(\bar{u}(t,\omega))=H_1$, for a.e. $(t,\omega)\in[0,T]\times\Omega$.

(AAS3) There is a $\beta(\cdot)\in C_{\mathbb{F}}([0,T];L^2(\Omega;H))$ such that

$$\begin{cases}\mathbb{E}\left\langle g^0_x(\bar{x}(t)),\beta(t)\right\rangle_H<0,\quad \forall\, t\in\mathcal{I}^0(\bar{x}),\\ \mathbb{E}\left\langle g^j_x(\bar{x}(T)),\beta(T)\right\rangle_H<0,\ \forall\, j\in\mathcal{I}(\bar{x}).\end{cases}$$

Remark 3.7. **(AAS1)** *is a condition about the unique continuation for the solution of* (3.34)*. It means that if $a_1[\cdot]^*\tilde{y}(\cdot)+b_1[\cdot]^*\widetilde{Y}(\cdot)=0$, then the nonhomogeneous term $\alpha(\cdot)$ must be zero. A sufficient condition for* **(AAS1)** *is that $a_1[\cdot]^*$ is injective and $b_1[\cdot]^*=0$.*

Remark 3.8. **(AAS2)** *means that $\mathcal{T}_\Phi(\bar{u})=L^2_{\mathbb{F}}(0,T;H_1)$. This, together with* **(AAS1)***, guarantees that the solution set of* (3.1) *is rich enough for us to choose one belonging to $\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}$.* **(AAS2)** *holds for some trivial cases. For example, $U=H_1$ or $\bar{u}(t,\omega)\in \mathrm{int}U$, $\mathbb{P}$-a.s. for a.e. $t\in[0,T]$. Note that we put state constraints* (1.3) *and* (1.4) *in the control problem. Hence, even for $U=H_1$, the optimal control problem is not trivial. We believe that for some concrete control problem, both* **(AAS1)** *and* **(AAS2)** *can be dropped. A possible way to do it is to follow the idea in the proof of Proposition 3.3 in [12]. The detailed analysis is beyond the scope of this paper and will be investigated in future work.*

Remark 3.9. *From the definition of $\mathcal{X}_{(1)}$, $\mathcal{G}^0_{(1)}$ and $\mathcal{G}_{(1)}$, it is clear that* **(AAS3)** *is necessary for $\mathcal{X}_{(1)}\cap\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$.*

Proposition 3.1. *Let* **(AS1)**–**(AS4)** *and* **(AAS1)**–**(AAS3)** *hold. Then $\mathcal{X}_{(1)}\cap\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$.*

Proof. We divide the proof into two steps.

Step 1. It follows from **(AAS2)** that $\mathcal{T}_\Phi(\bar u) = L^2_{\mathbb{F}}(0,T;H_1)$. Define a map $\Pi : \mathcal{T}_\Phi(\bar u) \to L^2_{\mathbb{F}}(0,T;H)$ in the following way:

$$\Pi(u_1)(\cdot) = x_1(\cdot),$$

where $x_1(\cdot)$ is the solution of (3.1) for some $u_1(\cdot) \in \mathcal{T}_\Phi(\bar u)$.

We claim that

$$\Pi(\mathcal{T}_\Phi(\bar u)) \text{ is dense in } L^2_{\mathbb{F}}(0,T;H). \tag{3.35}$$

Let us prove (3.35) by a contradiction argument. Without loss of generality, we assume that $\nu_1 = 0$. If (3.35) was false, then there would exist a nonzero $\beta_0(\cdot) \in L^2_{\mathbb{F}}(0,T;H)$ such that for any $u_1(\cdot) \in \mathcal{T}_\Phi(\bar u)$,

$$\mathbb{E}\int_0^T \big\langle x_1(t), \beta_0(t)\big\rangle_H dt = 0. \tag{3.36}$$

Let $\alpha = \beta_0$. By the definition of the transposition solution of (3.34), we have that for any $u_1(\cdot) \in \mathcal{T}_\Phi(\bar u)$,

$$0=\mathbb{E}\int_0^T \big\langle x_1(t), \beta_0(t)\big\rangle_H dt=\mathbb{E}\int_0^T \big\langle u_1(t), a_2(t)^*\tilde y(t)\big\rangle_{H_1} ds + \mathbb{E}\int_0^T \big\langle u_1(t), b_2(t)^*\widetilde Y(t)\big\rangle_{H_1} dt. \tag{3.37}$$

This, together with the choice of $u_1(\cdot)$, implies that $a_1[\cdot]^*\tilde y(\cdot) + b_1[\cdot]^*\widetilde Y(\cdot) = 0$ for a.e. $t \in [0,T]$. By **(AAS1)**, we see $\alpha = 0$ in $L^2_{\mathbb{F}}(0,T;H)$, a contradiction. Consequently, (3.35) holds.

Step 2. Since $\mathcal{I}^0(\bar x)$ is compact, by **(AAS3)**, one can find a $\beta(\cdot) \in C_{\mathbb{F}}([0,T];L^2(\Omega;H))$ such that, there are $\varepsilon_0 > 0$ and $M_0 > 0$ so that

$$\begin{cases} \mathbb{E}\big\langle g_x^0(\bar x(t)), \beta(t)\big\rangle_H < -\varepsilon_0, \quad |g_x^0(\bar x(t))|_{L^2_{\mathcal{F}_t}(\Omega;H)} \le M_0, \quad \forall\, t \in \mathcal{I}^0(\bar x), \\ \mathbb{E}\big\langle g_x^j(\bar x(T)), \beta(T)\big\rangle_H < -\varepsilon_0, \quad |g_x^j(\bar x(T))|_{L^2_{\mathcal{F}_t}(\Omega;H)} \le M_0, \quad \forall\, j \in \mathcal{I}(\bar x). \end{cases} \tag{3.38}$$

It follows from (3.35) that for every $k \in \mathbb{N}$, there is $u_{1,k} \in \mathcal{T}_\Phi(\bar u)$ such that the corresponding solution $x_{1,k} = \Pi(u_{1,k})$ satisfies that

$$|x_{1,k} - \beta|_{L^2_{\mathbb{F}}(0,T;H)} < \frac{1}{k}.$$

Consequently, there is a subsequence $\{u_{1,k_j}\}_{j=1}^\infty$ of $\{u_{1,k}\}_{k=1}^\infty$ such that

$$\lim_{j\to\infty} x_{1,k_j}(t) = \beta(t) \text{ in } L^2_{\mathcal{F}_T}(\Omega;H), \qquad \text{for a.e. } t \in [0,T]. \tag{3.39}$$

Since both $x_{1,k_j}(\cdot)$ and $\beta(\cdot)$ belong to $C_{\mathbb{F}}([0,T];L^2(\Omega;H))$, we get from (3.39) that

$$\lim_{j\to\infty} x_{1,k_j}(\cdot) = \beta(\cdot) \ \text{ in } \ C_{\mathbb{F}}([0,T];L^2(\Omega;H)).$$

Hence, there exists $N \in \mathbb{N}$ such that

$$|x_{1,N}(t) - \beta(t)|_{L^2_{\mathcal{F}_T}(\Omega;H)} < \frac{\varepsilon_0}{2M_0} \ \text{ for all } t \in [0,T].$$

This, together with **(AAS3)** and (3.38), implies that

$$\begin{aligned} \mathbb{E}\big\langle g_x^0(\bar x(t)), x_1(t)\big\rangle_H &= \mathbb{E}\big\langle g_x^0(\bar x(t)), x_1(t)-\beta(t)\big\rangle_H + \mathbb{E}\big\langle g_x^0(\bar x(t)), \beta(t)\big\rangle_H \\ &\le M_0 \times \frac{\varepsilon_0}{2M_0} - \varepsilon_0 < 0, \qquad\qquad \forall\, t \in \mathcal{I}^0(\bar x) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\big\langle g_x^j(\bar x(T)), x_1(T)\big\rangle_H &= \mathbb{E}\big\langle g_x^j(\bar x(T)), x_1(T)-\beta(T)\big\rangle_H + \mathbb{E}\big\langle g_x^j(\bar x(T)), \beta(T)\big\rangle_H \\ &\le M_0 \times \frac{\varepsilon_0}{2M_0} - \varepsilon_0 < 0, \qquad\qquad \forall\, j \in \mathcal{I}(\bar x). \end{aligned}$$

This completes the proof. $\square$

4. Second order necessary conditions

In this section, we establish second order necessary conditions for the optimal triple of **Problem (OP)**. In addition to **(AS1)**–**(AS4)**, we impose the following:

(AS5) *For a.e.* $(t,\omega)\in[0,T]\times\Omega$, *the operators* $a(t,\cdot,\cdot,\omega):\ H\times H_1\to H$ *and* $b(t,\cdot,\cdot,\omega):\ H\times H_1\to\mathcal{L}_2$ *are* C^2, *and* $a_{xu}(t,x,u,\omega)$ *and* $b_{xu}(t,x,u,\omega)$ *are uniformly continuous with respect to* $x\in H$ *and* $u\in H_1$, *and*

$$
\begin{aligned}
&|a_{xx}(t,x,u,\omega)|_{\mathcal{L}(H\times H;H)}+|a_{xu}(t,x,u,\omega)|_{\mathcal{L}(H\times H_1;H)}+|a_{uu}(t,x,u,\omega)|_{\mathcal{L}(H_1\times H_1;H)}\\
&+|b_{xx}(t,x,u,\omega)|_{\mathcal{L}(H\times H;\mathcal{L}_2)}+|b_{xu}(t,x,u,\omega)|_{\mathcal{L}(H\times H_1;\mathcal{L}_2)}+|b_{uu}(t,x,u,\omega)|_{\mathcal{L}(H_1\times H_1;\mathcal{L}_2)}\le C,\\
&\hfill\forall\,(x,u)\in H\times H_1.
\end{aligned}
$$

(AS6) *The functional* $h(\cdot,\omega):\ H\to\mathbb{R}$ *is* C^2, $\mathbb{P}$*-a.s., and for any* $x,\ \tilde x\in H$,

$$
|h_{xx}(x,\omega)|_{\mathcal{L}(H\times H;\mathbb{R})}\le C,\quad |h_{xx}(x,\omega)-h_{xx}(\tilde x,\omega)|_{\mathcal{L}(H\times H;\mathbb{R})}\le C|x-\tilde x|_H.
$$

(AS7) *For* $j=0,1,\cdots,n$, *the functional* $g^j(\cdot):\ H\to\mathbb{R}$ *is* C^2, *and for any* $x,\ \tilde x\in H$,

$$
|g^j_{xx}(x,\omega)|_{\mathcal{L}(H\times H;\mathbb{R})}\le C,\quad |g^j_{xx}(x,\omega)-g^j_{xx}(\tilde x,\omega)|_{\mathcal{L}(H\times H;\mathbb{R})}\le C|x-\tilde x|_H.
$$

(AS8) *The optimal control* $\bar u\in\mathcal{U}^4$.

In what follows, $\mathcal{U}^4$ is viewed as a subset of $L^4_{\mathbb{F}}(0,T;H_1)$ in the definitions of $T^b_{\mathcal{U}^4}(\bar u)$ and $T^{b(2)}_{\mathcal{U}^4}(\bar u,v)$.

(AS9) $(\Omega,\mathcal{F}_T,\mathbb{P})$ *is separable.*

Remark 4.1. *Similar to* **(AS2)**, **(AS5)** *is used to compute the Taylor expansion of the cost functional with respect to the control* u. *On the other hand, typical examples fulfill* **(AS6)** *and* **(AS7)** *are quadratic functional. For instance,* $h(x,\omega)=\eta(\omega)^2+|x|_H^2$ *and* $g^j(x)=|x|^2_H-1$ *(*$j=0,\cdots,n$*) for* $x\in H$.

Remark 4.2. *If* U *is bounded and the optimal control exists, then* **(AS7)** *holds.*

Remark 4.3. *Recall that* $(\Omega,\mathcal{F}_T,\mathbb{P})$ *is separable if there exists a countable family* $\mathcal{D}\subset\mathcal{F}_T$ *such that, for any* $\varepsilon>0$ *and* $B\in\mathcal{F}_T$ *one can find* $B_1\in\mathcal{D}$ *with* $\mathbb{P}\big((B\setminus B_1)\cup(B_1\setminus B)\big)<\varepsilon$. *Probability space enjoying such kind of property is called a standard probability space. Except some artificial examples, almost all frequently used probability spaces are standard ones(e.g. [36]). From [4, Section 13.4], if* **(AS9)** *holds, then* $L^p_{\mathcal{F}_T}(\Omega)$ *(*$1\le p<\infty$*) is separable.*

Consider the following $\mathcal{L}(H)$-valued BSEE*:

$$
\begin{cases}
dP=-(A^*+J^*)Pdt-P(A+J)dt-K^*PKdt-(K^*Q+QK)dt+Fdt+QdW(t)\ \text{ in }[0,T),\\
P(T)=P_T,
\end{cases}
\tag{4.1}
$$

where $F\in L^1_{\mathbb{F}}(0,T;L^2(\Omega;\mathcal{L}(H)))$, $P_T\in L^2_{\mathcal{F}_T}(\Omega;\mathcal{L}(H))$, $J\in L^4_{\mathbb{F}}(0,T;L^\infty(\Omega;\mathcal{L}(H)))$ and $K\in L^4_{\mathbb{F}}(0,T;L^\infty(\Omega;\mathcal{L}(H;\mathcal{L}_2)))$. In (4.1), the unknown (or solution) is a pair (P,Q).

*Throughout this paper, for any operator-valued process (*resp.* random variable) R, we denote by R^* its pointwisely dual operator-valued process (*resp.* random variable), e.g., if $R\in L^{r_1}_{\mathbb{F}}(0,T;L^{r_2}(\Omega;\mathcal{L}(H)))$, then $R^*\in L^{r_1}_{\mathbb{F}}(0,T;L^{r_2}(\Omega;\mathcal{L}(H)))$, and $\|R\|_{L^{r_1}_{\mathbb{F}}(0,T;L^{r_2}(\Omega;\mathcal{L}(H)))}=\|R^*\|_{L^{r_1}_{\mathbb{F}}(0,T;L^{r_2}(\Omega;\mathcal{L}(H)))}$.

Let us first recall the definition of the relaxed transposition solution of (4.1). To this end, consider two SEEs:

$$
\left\{\begin{array}{l} d\phi_1(s)=\big[(A+J)\phi_1(s)+\tilde f_1(s)\big]ds+\big(K\phi_1(s)+\hat f_1(s)\big)dW(s) \quad \text{ in } (t,T],\\ \phi_1(t)=\xi_1 \end{array}\right. \tag{4.2}
$$

and

$$
\left\{\begin{array}{l} d\phi_2(s)=\big[(A+J)\phi_2(s)+\tilde f_2(s)\big]ds+\big(K\phi_2(s)+\hat f_2(s)\big)dW(s) \quad \text{ in } (t,T],\\ \phi_2(t)=\xi_2. \end{array}\right. \tag{4.3}
$$

Here $t\in[0,T)$, $\xi_1,\xi_2\in L^4_{\mathcal{F}_t}(\Omega;H)$, $\tilde f_1,\tilde f_2\in L^2_{\mathbb{F}}(t,T;L^4(\Omega;H))$ and $\hat f_1,\hat f_2\in L^2_{\mathbb{F}}(t,T;L^4(\Omega;\mathcal{L}_2))$.

Write

$$
\begin{aligned}
&\mathcal{P}[0,T]\\
&\stackrel{\triangle}{=}\Big\{P(\cdot,\cdot)\ \Big|\ P(\cdot,\cdot)\in\mathcal{L}\big(L^2_{\mathbb{F}}(0,T;L^4(\Omega;H)),\ L^2_{\mathbb{F}}(0,T;L^{\frac43}(\Omega;H))\big),\ P(t,\omega)\in\mathcal{L}(H)\text{ for a.e.}\\
&\quad (t,\omega)\in[0,T]\times\Omega,\ \text{ and for every } t\in[0,T] \text{ and } \xi\in L^4_{\mathcal{F}_t}(\Omega;H),\ P(\cdot,\cdot)\xi\in D_{\mathbb{F}}([t,T];L^{\frac43}(\Omega;H))\\
&\quad \text{and } \|P(\cdot,\cdot)\xi\|_{D_{\mathbb{F}}([t,T];L^{\frac43}(\Omega;H))}\le C\|\xi\|_{L^4_{\mathcal{F}_t}(\Omega;H)}\Big\}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{Q}[0,T]\stackrel{\triangle}{=}\Big\{\big(Q^{(\cdot)},\widehat Q^{(\cdot)}\big)\ \Big|\ &\text{For any } t\in[0,T],\ \text{both } Q^{(t)} \text{ and } \widehat Q^{(t)} \text{ are bounded linear operators}\\
&\text{from } L^4_{\mathcal{F}_t}(\Omega;H)\times L^2_{\mathbb{F}}(t,T;L^4(\Omega;H))\times L^2_{\mathbb{F}}(t,T;L^4(\Omega;\mathcal{L}_2)) \text{ to } L^2_{\mathbb{F}}(t,T;L^{\frac43}(\Omega;\mathcal{L}_2))\\
&\text{and } Q^{(t)}(0,0,\cdot)^*=\widehat Q^{(t)}(0,0,\cdot)\Big\}.
\end{aligned}
$$

In what follows, for $P\in\mathcal{P}[0,T]$, we write $|P|_{\mathcal{P}[0,T]}$ for $|P|_{\mathcal{L}\left(L^2_{\mathbb{F}}(0,T;L^4(\Omega;H)),\ L^2_{\mathbb{F}}(0,T;L^{\frac43}(\Omega;H))\right)}$. Similarly, for $\big(Q^{(\cdot)},\widehat Q^{(\cdot)}\big)\in\mathbb{Q}[0,T]$, we put

$$
\big|\big(Q^{(\cdot)},\widehat Q^{(\cdot)}\big)\big|_{\mathbb{Q}[0,T]}\stackrel{\triangle}{=}\big|\big(Q^{(\cdot)},\widehat Q^{(\cdot)}\big)\big|_{\mathcal{L}\left(L^4_{\mathcal{F}_t}(\Omega;H)\times L^2_{\mathbb{F}}(t,T;L^4(\Omega;H))\times L^2_{\mathbb{F}}(t,T;L^4(\Omega;\mathcal{L}_2));L^2_{\mathbb{F}}(t,T;L^{\frac43}(\Omega;\mathcal{L}_2))\right)}.
$$

Definition 4.1. *We call $\big(P(\cdot),(Q^{(\cdot)},\widehat Q^{(\cdot)})\big)\in\mathcal{P}[0,T]\times\mathbb{Q}[0,T]$ a relaxed transposition solution of (4.1) if for every $t\in[0,T]$, $\xi_1,\xi_2\in L^4_{\mathcal{F}_t}(\Omega;H)$, $\tilde f_1(\cdot),\tilde f_2(\cdot)\in L^2_{\mathbb{F}}(t,T;\ L^4(\Omega;H))$ and $\hat f_1(\cdot),\hat f_2(\cdot)\in L^2_{\mathbb{F}}(t,T;L^4(\Omega;\mathcal{L}_2))$, the following is satisfied*

$$
\begin{aligned}
&\mathbb{E}\big\langle P_T\phi_1(T),\phi_2(T)\big\rangle_H-\mathbb{E}\int_t^T\big\langle F(s)\phi_1(s),\phi_2(s)\big\rangle_H ds\\
&=\mathbb{E}\big\langle P(t)\xi_1,\xi_2\big\rangle_H+\mathbb{E}\int_t^T\big\langle P(s)\tilde f_1(s),\phi_2(s)\big\rangle_H ds+\mathbb{E}\int_t^T\big\langle P(s)\phi_1(s),\tilde f_2(s)\big\rangle_H ds\\
&\quad+\mathbb{E}\int_t^T\big\langle P(s)K(s)\phi_1(s),\hat f_2(s)\big\rangle_{\mathcal{L}_2}ds+\mathbb{E}\int_t^T\big\langle P(s)\hat f_1(s),K(s)\phi_2(s)+\hat f_2(s)\big\rangle_{\mathcal{L}_2}ds\\
&\quad+\mathbb{E}\int_t^T\big\langle \hat f_1(s),\widehat Q^{(t)}(\xi_2,\tilde f_2,\hat f_2)(s)\big\rangle_{\mathcal{L}_2}ds+\mathbb{E}\int_t^T\big\langle Q^{(t)}(\xi_1,\tilde f_1,\hat f_1)(s),\hat f_2(s)\big\rangle_{\mathcal{L}_2}ds.
\end{aligned}\tag{4.4}
$$

Here, $\phi_1(\cdot)$ and $\phi_2(\cdot)$ solve (4.2) and (4.3), respectively.

Lemma 4.1. *Let* **(AS9)** *hold. Then the equation (4.1) admits a unique relaxed transposition solution $\big(P(\cdot),(Q^{(\cdot)},\ \widehat Q^{(\cdot)})\big)\in\mathcal{P}[0,T]\times\mathbb{Q}[0,T]$. Furthermore,*

$$
|P|_{\mathcal{P}[0,T]}+\big|\big(Q^{(\cdot)},\widehat Q^{(\cdot)}\big)\big|_{\mathbb{Q}[0,T]}\le C\big(|F|_{L^1_{\mathbb{F}}(0,T;\,L^2(\Omega;\mathcal{L}(H)))}+|P_T|_{L^2_{\mathcal{F}_T}(\Omega;\,\mathcal{L}(H))}\big).
$$

The proof is almost the same as the one of [29, Theorem 6.1]. The only difference is that one should replace the inner product of H by $\mathcal{L}_2$ for terms involving $\hat{f}_1$ and $\hat{f}_2$. Hence we omit it.

For φ equal to a or b, let

$$\varphi_{11}[t]=\varphi_{xx}(t,\bar{x}(t),\bar{u}(t)),\quad \varphi_{12}[t]=\varphi_{xu}(t,\bar{x}(t),\bar{u}(t)),\quad \varphi_{22}[t]=\varphi_{uu}(t,\bar{x}(t),\bar{u}(t)).$$

For $\nu_1\in T^b_{\mathcal{V}}(\bar{x}_0)$, $u_1\in T^b_{\mathcal{U}^4}(\bar{u})$, $\nu_2\in T^{b(2)}_{\mathcal{V}}(\bar{x}_0,\nu_1)$ and $u_2\in T^{b(2)}_{\mathcal{U}^4}(\bar{u},u_1)$, consider the following second order variational equation:

$$\begin{cases} dx_2(t)=\Big[Ax_2(t)+a_1[t]x_2(t)+a_2[t]u_2(t)+\dfrac{1}{2}a_{11}[t]\big(x_1(t),x_1(t)\big)+a_{12}[t]\big(x_1(t),u_1(t)\big)\\ \qquad\qquad +\dfrac{1}{2}a_{22}[t]\big(u_1(t),u_1(t)\big)\Big]dt+\Big[b_1[t]x_2(t)+b_2[t]u_2(t)+\dfrac{1}{2}b_{11}[t]\big(x_1(t),x_1(t)\big)\\ \qquad\qquad +b_{12}[t]\big(x_1(t),u_1(t)\big)+\dfrac{1}{2}b_{22}[t]\big(u_1(t),u_1(t)\big)\Big]dW(t) \qquad\qquad \text{in } (0,T],\\ x_2(0)=\nu_2, \end{cases} \tag{4.5}$$

where $x_1(\cdot)$ is the solution of the first order variational equation (3.1) (for $u_1(\cdot)$ and ν_1 as above). Further, from **(AS5)**, we know that $a_{11}[t]\big(x_1(t),x_1(t)\big)\in H$. So do the other terms such as $b_{11}[t]\big(x_1(t),x_1(t)\big)$.

By the definition of the second order adjacent tangent, for any $\varepsilon>0$, there exist $\nu_2^\varepsilon\in H$ and $u_2^\varepsilon(\cdot)\in L^4_{\mathbb{F}}(0,T;H_1)$ such that

$$\nu_0^\varepsilon\overset{\triangle}{=}\bar{\nu}_0+\varepsilon\nu_1+\varepsilon^2\nu_2^\varepsilon\in\mathcal{V},\qquad u^\varepsilon(\cdot)\overset{\triangle}{=}\bar{u}(\cdot)+\varepsilon u_1(\cdot)+\varepsilon^2u_2^\varepsilon(\cdot)\in\mathcal{U}^4$$

and

$$\lim_{\varepsilon\to0^+}\nu_2^\varepsilon=\nu_2 \text{ in } H,\qquad \lim_{\varepsilon\to0^+}u_2^\varepsilon=u_2 \text{ in } L^4_{\mathbb{F}}(0,T;H_1).$$

Denote by $x^\varepsilon(\cdot)$ the solution of (1.1) corresponding to the control $u^\varepsilon(\cdot)$ and the initial datum ν_0^ε. Put

$$\delta x^\varepsilon(\cdot)\overset{\triangle}{=}x^\varepsilon(\cdot)-\bar{x}(\cdot),\qquad r_2^\varepsilon(\cdot)\overset{\triangle}{=}\frac{\delta x^\varepsilon(\cdot)-\varepsilon x_1(\cdot)-\varepsilon^2x_2(\cdot)}{\varepsilon^2}.$$

We have the following result.

Lemma 4.2. *Suppose that* **(AS1)**, **(AS2)** *and* **(AS5)** *hold. Then, for* $\nu_1,\nu_2,\nu_2^\varepsilon\in H$ *and* $u_1(\cdot),u_2(\cdot)$, $u_2^\varepsilon(\cdot)\in L^4_{\mathbb{F}}(0,T;H_1)$ *as above, we have*

$$\|x_2\|_{L^\infty_{\mathbb{F}}(0,T;L^2(\Omega;H))}\le C\big(|\nu_2|_H+|\nu_1|_H^2+|u_1|^2_{L^4_{\mathbb{F}}(0,T;H_1)}+|u_2|_{L^2_{\mathbb{F}}(0,T;H_1)}\big)$$

and

$$\lim_{\varepsilon\to0^+}|r_2^\varepsilon|_{L^\infty_{\mathbb{F}}(0,T;L^2(\Omega;H))}=0. \tag{4.6}$$

Proof of Lemma 4.2 is provided in Appendix B.

Put

$$\begin{aligned}\mathcal{Y}(\bar{x},\bar{u})\overset{\triangle}{=}\big\{&(x_1(\cdot),u_1(\cdot),\nu_1)\in C_{\mathbb{F}}([0,T];L^4(\Omega;H))\times T^b_{\mathcal{U}^4}(\bar{u})\times T^b_{\mathcal{V}}(\bar{\nu}_0)\big|\ x_1(\cdot)\text{ solves (3.1)},\\ &x_1(\cdot)\in\mathrm{cl}\mathcal{G}^0_{(1)}\cap\mathrm{cl}\mathcal{G}_{(1)}\text{ and }\mathbb{E}\langle g^j_x(\bar{x}(T)),x_1(T)\rangle_H\le0,\ \forall\, j\in\mathcal{I}(\bar{x})\big\}\end{aligned} \tag{4.7}$$

and define the critical cone

$$\mathcal{Z}(\bar{x},\bar{u})\overset{\triangle}{=}\Big\{(x_1(\cdot),u_1(\cdot),\nu_1)\in\mathcal{Y}(\bar{x},\bar{u})\Big|\ \mathbb{E}\langle h_x(\bar{x}(T)),x_1(T)\rangle_H=0\Big\}. \tag{4.8}$$

For a fixed $(x_1(\cdot), u_1(\cdot), \nu_1) \in \mathcal{Z}(\bar{x}, \bar{u})$, let $\mathcal{W}(\bar{\nu}_0, \nu_1)$ and $\mathcal{M}(\bar{u}, u_1)$ be convex subsets of $T_{\mathcal{V}}^{b(2)}(\bar{\nu}_0, \nu_1)$ and $T_{\mathcal{U}^4}^{b(2)}(\bar{u}, u_1)$, respectively. Put

$$\begin{aligned}\mathcal{X}_{(2)}(x_1, u_1) \triangleq \big\{ & x_2(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T]; H)) \big| \, x_2(\cdot) \text{ is the solution of (4.5) corresponding} \\ & \text{to some } (\nu_2, u_2) \in \mathcal{W}(\bar{\nu}_0, \nu_1) \times \mathcal{M}(\bar{u}, u_1) \big\}.\end{aligned} \tag{4.9}$$

Let

$$\mathbb{I}^0(\bar{x}, x_1) \triangleq \big\{ t \in \mathcal{I}^0(\bar{x}) \big| \; \mathbb{E}\langle g_x^0(\bar{x}(t)), x_1(t) \rangle_H = 0 \big\},$$

$$\mathbb{I}(\bar{x}, x_1) \triangleq \big\{ j \in \mathcal{I}(\bar{x}) \big| \; \mathbb{E}\langle g_x^j(\bar{x}(T)), x_1(T) \rangle_H = 0 \big\},$$

$$\begin{aligned}\tau^g(\bar{x}) \triangleq \Big\{ & t \in [0,T] \; \Big| \; \exists \, \{s_k\}_{k=1}^{\infty} \subset [0,T] \text{ such that } \lim_{k\to\infty} s_k = t, \, \mathbb{E} g^0(\bar{x}(s_k)) < 0, \\ & \mathbb{E}\left\langle g_x^0(\bar{x}(s_k)), x_1(s_k) \right\rangle_H > 0, \; \forall \, k = 1, 2, \cdots \Big\},\end{aligned} \tag{4.10}$$

$$e(t) \triangleq \begin{cases} \displaystyle\varlimsup_{\substack{s\to t \\ \mathbb{E}\, g^0(\bar{x}(s))<0 \\ \mathbb{E}\langle g_x^0(\bar{x}(s)), x_1(s)\rangle_H > 0}} \frac{\left| \mathbb{E}\left\langle g_x^0(\bar{x}(s)), x_1(s) \right\rangle_H \right|^2}{4\left| \mathbb{E} g^0(\bar{x}(s)) \right|_H}, & t \in \tau^g(\bar{x}), \\ 0, & \text{otherwise}, \end{cases} \tag{4.11}$$

$$\begin{aligned}\mathcal{G}^0_{(2)}(x_1) \triangleq \Big\{ & z \in L^2_{\mathbb{F}}(\Omega; C([0,T]; H)) \Big| \text{ for all } t \in \mathbb{I}^0(\bar{x}, x_1), \\ & \mathbb{E}\langle g_x^0(\bar{x}(t)), z(t) \rangle_H + \frac{1}{2}\mathbb{E}\langle g_{xx}^0(\bar{x}(t)) x_1(t), x_1(t) \rangle_H + e(t) < 0 \Big\},\end{aligned} \tag{4.12}$$

$$\mathcal{G}^j_{(2)}(x_1) \triangleq \Big\{ z \in L^2_{\mathbb{F}}(\Omega; C([0,T]; H)) \Big| \mathbb{E}\langle g_x^j(\bar{x}(T)), z(T) \rangle_H + \frac{1}{2}\mathbb{E}\langle g_{xx}^j(\bar{x}(T)) x_1(T), x_1(T) \rangle_H < 0 \Big\},$$

$$\mathcal{G}_{(2)}(x_1) \triangleq \bigcap_{j \in \mathbb{I}(\bar{x}, x_1)} \mathcal{G}^j_{(2)}(x_1), \tag{4.13}$$

and

$$\mathcal{H}_{(2)}(x_1) \triangleq \Big\{ z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];H)) \Big| \, \mathbb{E}\langle h_x(\bar{x}(T)), z(T) \rangle_H + \frac{1}{2}\mathbb{E}\langle h_{xx}(\bar{x}(T)) x_1(T), x_1(T) \rangle_H < 0 \Big\}. \tag{4.14}$$

Remark 4.4. *If $x_1 \in \mathcal{G}^0_{(1)}$, then $\mathbb{I}^0(\bar{x}, x_1) = \emptyset$. Consequently, $\mathcal{G}^0_{(2)}(x_1) = L^2_{\mathbb{F}}(\Omega; C([0,T]; H))$. In addition, if there exists $\delta > 0$ such that*

$$\mathbb{E}\left\langle g_x(\bar{x}(s)), x_1(s) \right\rangle_H \le 0, \quad \forall \, s \in (t-\delta, t+\delta) \cap [0,T], \; t \in \mathcal{I}^0(\bar{x}),$$

then $e(t) = 0$ for any $t \in \mathbb{I}^0(\bar{x}, x_1)$. In this case,

$$\begin{aligned}\mathcal{G}^0_{(2)}(x_1) = \Big\{ & z(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T]; H)) \Big| \text{ For all } t \in \mathbb{I}^0(\bar{x}, x_1), \\ & \mathbb{E}\left\langle g_x^0(\bar{x}(t)), z(t) \right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle g_{xx}^0(\bar{x}(t)) x_1(t), x_1(t) \right\rangle_H < 0 \Big\}.\end{aligned}$$

Remark 4.5. *Let $z_1 \in \mathcal{G}^0_{(1)}$ and $z_2 \in \mathcal{G}^0_{(2)}(x_1)$. Then for every $t \in \mathbb{I}^0(\bar{x}, x_1) \subset \mathcal{I}^0(\bar{x})$, we have $\mathbb{E}\langle g_x^0(\bar{x}(t)), z_1(t) \rangle_H < 0$ and $\mathbb{E}\langle g_x^0(\bar{x}(t)), z_2(t) \rangle_H + \frac{1}{2}\mathbb{E}\langle g_{xx}^0(\bar{x}(t)) x_1(t), x_1(t) \rangle_H + e(t) < 0$. Therefore,*

$$\mathbb{E}\langle g_x^0(\bar{x}(t)), z_1(t) + z_2(t) \rangle_H + \frac{1}{2}\mathbb{E}\langle g_{xx}^0(\bar{x}(t)) x_1(t), x_1(t) \rangle_H + e(t) < 0,$$

which implies that $z_1 + z_2 \in \mathcal{G}^0_{(2)}(x_1)$. Consequently, $\mathcal{G}^0_{(1)} + \mathcal{G}^0_{(2)}(x_1) \subset \mathcal{G}^0_{(2)}(x_1)$. Similarly, if $\Phi(t,\omega) = \mathcal{C}_U(\bar{u}(t,\omega))$, then we can prove that $\mathcal{X}_{(1)} + \mathcal{X}_{(2)}(x_1, u_1) \subset \mathcal{X}_{(2)}(x_1, u_1)$.

Let (y,Y), ψ and λ_j, $j\in\mathcal{I}(\bar{x})$ be defined as in the proof of Theorem 3.1 in the case when $\mathcal{X}_{(1)}\cap\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$ (See (3.8), (3.10) and (3.12) for the definitions of $\mathcal{X}_{(1)}$, $\mathcal{G}^0_{(1)}$ and $\mathcal{G}_{(1)}$, respectively), where $y(T)=-h_x(\bar{x}(T))-\sum\limits_{j\in\mathcal{I}(\bar{x})}\lambda_j g^j_x(\bar{x}(T))$.

Let $(P(\cdot),(Q^{(\cdot)},\widehat{Q}^{(\cdot)}))$ be the relaxed transposition solution of the equation (4.1) in which P_T, $J(\cdot)$, $K(\cdot)$ and $F(\cdot)$ are given by

$$
\begin{aligned}
&P_T=-h_{xx}\big(\bar{x}(T)\big),\quad J(t)=a_1[t],\quad K(t)=b_1[t],\\
&F(t)=-\mathbb{H}_{xx}[t]\overset{\triangle}{=}-\mathbb{H}_{xx}(t,\bar{x}(t),\bar{u}(t),y(t),Y(t),\omega).
\end{aligned}
$$

We have the following result.

Theorem 4.1. *Suppose that* **(AS1)**–**(AS9)** *hold and that $\mathcal{X}_{(1)}\cap\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$ for an optimal triple $(\bar{x}(\cdot),\bar{u}(\cdot),\bar{\nu}_0)$ of* **Problem (OP)**. *If $\mathcal{X}_{(2)}(x_1,u_1)\cap\mathcal{G}^0_{(2)}(x_1)\cap\mathcal{G}_{(2)}(x_1)\neq\emptyset$, then for any $x_2(\cdot)\in\mathcal{X}_{(2)}(x_1,u_1)\cap\mathrm{cl}\mathcal{G}^0_{(2)}(x_1)\cap\mathrm{cl}\mathcal{G}_{(2)}(x_1)$ with the corresponding $\nu_2\in\mathcal{W}(\bar{\nu}_0,\nu_1)$ and $u_2(\cdot)\in\mathcal{M}(\bar{u},u_1)$, we have*

$$
\begin{aligned}
&\langle y(0),\nu_2\rangle_H+\frac{1}{2}\langle P(0)\nu_1,\nu_1\rangle_H+\sum_{j\in\mathcal{I}(\bar{x})}\mathbb{E}\big\langle\lambda_j g^j_x(\bar{x}(T)),x_2(T)\big\rangle_H\\
&+\mathbb{E}\int_0^T\Big(\big\langle\mathbb{H}_u[t],u_2(t)\big\rangle_{H_1}+\frac{1}{2}\big\langle\mathbb{H}_{uu}[t]u_1(t),u_1(t)\big\rangle_{H_1}+\frac{1}{2}\big\langle b_2[t]^*P(t)b_2[t]u_1(t),u_1(t)\big\rangle_{H_1}\\
&\qquad+\big\langle\big(\mathbb{H}_{xu}[t]+a_2[t]^*P(t)+b_2[t]^*P(t)b_1[t]\big)x_1(t),u_1(t)\big\rangle_{H_1}\\
&+\frac{1}{2}\big\langle\big(\widehat{Q}^{(0)}+Q^{(0)}\big)\big(0,a_2[t]u_1(t),b_2[t]u_1(t)\big),b_2[t]u_1(t)\big\rangle_{\mathcal{L}_2}\Big)dt+\mathbb{E}\int_0^T\langle x_2(t),d\psi(t)\rangle_H\leq 0,
\end{aligned}
\tag{4.15}
$$

where

$$
\mathbb{H}_{uu}[t]\overset{\triangle}{=}\mathbb{H}_{uu}(t,\bar{x}(t),\bar{u}(t),y(t),Y(t),\omega),\quad \mathbb{H}_{xu}[t]\overset{\triangle}{=}\mathbb{H}_{xu}(t,\bar{x}(t),\bar{u}(t),y(t),Y(t),\omega).
$$

Remark 4.6. *Similar to Theorem 3.1, if $\bar{u}$ takes an isolated point of U in a positive measure set of $[0,T]\times\Omega$, then (4.15) does not give us any information about the optimal control at these point since ar these points. This is a drawback of Theorem 4.1. As the first order necessary condition, one may use the spike variation technique. However, in such case, one has to use four adjoint equations. A detailed analysis of this is beyond of the scope of this paper.*

Remark 4.7. *In Theorem 4.1, we take $\lambda_0=1$ and (y,Y), ψ and λ_j, $j\in\mathcal{I}(\bar{x})$ as in Theorem 3.1. Accordingly, the terms $\sum\limits_{j\in\mathcal{I}(\bar{x})}\mathbb{E}\big\langle\lambda_j g^j_x(\bar{x}(T)),x_2(T)\big\rangle_H$ and $\mathbb{E}\int_0^T\langle x_2(t),d\psi(t)\rangle_H$ appear. By doing so, our second order condition is valid for any normal multiplier appearing in the first order conditions.*

In Theorem 4.1 we assumed that $\mathcal{X}_{(2)}(x_1,u_1)\cap\mathcal{G}^0_{(2)}(x_1)\cap\mathcal{G}_{(2)}(x_1)\neq\emptyset$. It seems that this condition is not easy to verify. Let us give a result concerning this below.

Proposition 4.1. *Assume that there is $(x_1,u_1,\nu_1)\in\mathcal{Z}(\bar{x},\bar{u})$ such that the function $e(\cdot)$ defined by (4.11) is bounded on $\mathbb{I}^0(\bar{x},x_1)$, and that $T^{b(2)}_{\mathcal{V}}(\bar{\nu}_0,\nu_1)$ and $T^{b(2)}_{\mathcal{U}^4}(\bar{u},u_1)$ are nonempty. If $\mathcal{X}_{(1)}\cap\mathcal{G}^0_{(1)}\cap\mathcal{G}_{(1)}\neq\emptyset$ (with $\mathcal{T}_{\mathcal{K}}(\bar{\nu}_0)$ and $\mathcal{T}_{\Phi}(\bar{u})$ being replaced by $\mathcal{C}_{\mathcal{K}}(\bar{\nu}_0)$ and $\mathcal{C}_{\mathcal{U}^4}(\bar{u})$, respectively), then $\mathcal{X}_{(2)}(x_1,u_1)\cap\mathcal{G}^0_{(2)}(x_1)\cap\mathcal{G}_{(2)}(x_1)\neq\emptyset$.*

Proof. If $\mathcal{X}_{(1)} \cap \mathcal{G}^0_{(1)} \cap \mathcal{G}_{(1)} \neq \emptyset$ (with $\mathcal{T}_{\mathcal{K}}(\bar{\nu}_0)$ and $\mathcal{T}_{\Phi}(\bar{u})$ being replaced by $\mathcal{C}_{\mathcal{K}}(\bar{\nu}_0)$ and $\mathcal{C}_{\mathcal{U}^4}(\bar{u})$, respectively), then there exists $\hat{x}_1(\cdot) \in \mathcal{X}_{(1)} \cap \mathcal{G}^0_{(1)} \cap \mathcal{G}_{(1)}$ with the initial datum $\hat{\nu}_1 \in \mathcal{C}_{\mathcal{K}}(\bar{\nu}_0)$ and the control $\hat{u}_1(\cdot) \in \mathcal{C}_{\mathcal{U}^4}(\bar{u})$.

Since $T^{b(2)}_{\mathcal{K}}(\bar{\nu}_0, \nu_1)$ and $T^{b(2)}_{\mathcal{U}^4}(\bar{u}, v)$ are nonempty, they contain some nonempty convex subsets $\mathcal{W}^1(\bar{\nu}_0, \nu_1)$ and $\mathcal{M}^1(\bar{u}, u_1)$, respectively.

Put
$$\mathcal{W}(\bar{\nu}_0, \nu_1) \overset{\triangle}{=} \mathcal{C}_{\mathcal{K}}(\bar{\nu}_0) + \mathcal{W}^1(\bar{\nu}_0, \nu_1), \qquad \mathcal{M}(\bar{u}, u_1) \overset{\triangle}{=} \mathcal{C}_{\mathcal{U}^4}(\bar{u}) + \mathcal{M}^1(\bar{u}, u_1).$$
It follows from Lemma 2.4 in [11] that $\mathcal{W}(\bar{\nu}_0, \nu_1) \subset T^{b(2)}_{\mathcal{K}}(\bar{\nu}_0, \nu_1)$ and $\mathcal{M}(\bar{u}, u_1) \subset T^{b(2)}_{\mathcal{U}^4}(\bar{u}, u_1)$. Moreover, for every $\tilde{\nu}_2 \in \mathcal{W}^1(\bar{\nu}_0, \nu_1)$, $\tilde{u}_2 \in \mathcal{M}^1(\bar{u}, u_1)$ and $\delta \geq 0$, we have $\delta\hat{\nu}_1 + \tilde{\nu}_2 \in \mathcal{W}(\bar{\nu}_0, \nu_1)$ and $\delta\hat{u}_1 + \tilde{u}_2 \in \mathcal{M}(\bar{u}, u_1)$.

Fixing $\delta \geq 0$ and letting $x_{2,\delta}(\cdot)$(*resp.* $\tilde{x}_2$) be the solution of (4.5) corresponding to $\delta\hat{\nu}_1 + \tilde{\nu}_2$ (*resp.* $\tilde{\nu}_2$) and $\delta\hat{u}_1 + \tilde{u}_2$ (*resp.* $\tilde{u}_2$), we have $x_{2,\delta}(\cdot) = \delta\hat{x}_1(\cdot) + \tilde{x}_2(\cdot)$. It follows from Lemma 4.2 that
$$|\tilde{x}_2|^2_{L^\infty_{\mathbb{F}}(0,T;L^2(\Omega;H))} \leq C\big(|\tilde{\nu}_2|^2_H + |\hat{u}_1|^4_{L^4_{\mathbb{F}}(0,T;H)} + |\tilde{u}_2|^2_{L^2_{\mathbb{F}}(0,T;H)}\big).$$
Since $\hat{x}_1(\cdot) \in \mathcal{X}_{(1)} \cap \mathcal{G}^0_{(1)} \cap \mathcal{G}_{(1)}$, and $\mathcal{I}^0(\bar{x})$ and $\mathbb{I}^0(\bar{x}, \hat{x}_1)$ are compact sets, for all sufficiently large δ,
$$\begin{aligned}
&\mathbb{E}\left\langle h_x(\bar{x}(t)), x_{2,\delta}(t)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle h_{xx}(\bar{x}(t))x_1(t), x_1(t)\right\rangle_H + e(t)\\
&= \delta\mathbb{E}\left\langle h_x(\bar{x}(t)), \hat{x}_1(t)\right\rangle_H + \mathbb{E}\left\langle h_x(\bar{x}(t)), \tilde{x}_2(t)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle h_{xx}(\bar{x}(t))x_1(t), x_1(t)\right\rangle_H + e(t)\\
&< 0, \qquad\qquad \forall\, t \in \mathbb{I}^0(\bar{x}, \hat{x}_1),
\end{aligned}$$
and for every $j \in \mathbb{I}(\bar{x}, \hat{x}_1)$, and all δ sufficiently large
$$\begin{aligned}
&\mathbb{E}\left\langle g^j_x(\bar{x}(T)), x_{2,\delta}(T)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle g^j_{xx}(\bar{x}(T))x_1(T), x_1(T)\right\rangle_H\\
&= \delta\mathbb{E}\left\langle g^j_x(\bar{x}(T)), \hat{x}_1(T)\right\rangle_H + \mathbb{E}\left\langle g^j_x(\bar{x}(T)), \tilde{x}_1(T)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle g^j_{xx}(\bar{x}(T))x_1(T), x_1(T)\right\rangle_H < 0.
\end{aligned}$$
Therefore, when δ is large enough, $x_{2,\delta}(\cdot) \in \mathcal{X}_{(2)}(x_1, u_1) \cap \mathcal{G}^0_{(2)}(x_1) \cap \mathcal{G}_{(2)}(x_1)$. This yields that $\mathcal{X}_{(2)}(x_1, u_1) \cap \mathcal{G}^0_{(2)}(x_1) \cap \mathcal{G}_{(2)}(x_1) \neq \emptyset$ □

Proof of Theorem 4.1. If $\mathbb{I}^0(\bar{x}, x_1) = \emptyset$, then $\mathcal{G}^0_{(2)}(x_1) = L^2_{\mathbb{F}}(\Omega; C([0,T]; H))$. Hence,
$$\mathcal{X}_{(2)}(x_1, u_1) \cap \mathcal{G}^0_{(2)}(x_1) \cap \mathcal{G}_{(2)}(x_1) = \mathcal{X}_{(2)}(x_1, u_1) \cap \mathcal{G}_{(2)}(x_1).$$
In such case, without loss of generality, we can ignore the constraint (1.3) and put $\psi = 0$. Thus, we only need to consider the case $\mathbb{I}^0(\bar{x}, x_1) \neq \emptyset$.

The proof is divided into five steps. In the first four steps, we deal with the special case when $x_2(\cdot) \in \mathcal{X}_{(2)}(x_1, u_1) \cap \mathcal{G}^0_{(2)}(x_1) \cap \mathcal{G}_{(2)}(x_1)$. Then, in the last step, we handle the general case.

Step 1: Since $x_2(\cdot) \in \mathcal{X}_{(2)}(x_1, u_1) \cap \mathcal{G}^0_{(2)}(x_1) \cap \mathcal{G}_{(2)}(x_1)$, $x_2(\cdot)$ is a solution of the equation (4.5) corresponding to some $(\nu_2, u_2) \in \mathcal{W}(\bar{x}_0, \nu_0) \times \mathcal{M}(\bar{u}, u_1)$ such that
$$\mathbb{E}\left\langle g^0_x(\bar{x}(t)), x_2(t)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle g^0_{xx}(\bar{x}(t))x_1(t), x_1(t)\right\rangle_H + e(t) < 0, \quad \forall\, t \in \mathbb{I}^0(\bar{x}, x_1)$$
and
$$\mathbb{E}\left\langle g^j_x(\bar{x}(T)), x_2(T)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle g^j_{xx}(\bar{x}(T))x_1(T), x_1(T)\right\rangle_H < 0, \quad \forall\, j \in \mathbb{I}(\bar{x}, x_1).$$
Let $\mu^\varepsilon \in H$ and $\eta^\varepsilon(\cdot) \in L^4_{\mathbb{F}}(0,T; H_1)$ be such that
$$\begin{aligned}
&|\mu^\varepsilon| = o(\varepsilon^2), && \nu^\varepsilon_0 \overset{\triangle}{=} \bar{x}_0 + \varepsilon\nu_1 + \varepsilon^2\nu_2 + \mu^\varepsilon \in \mathcal{V},\\
&|\eta^\varepsilon|_{L^4_{\mathbb{F}}(0,T;H_1)} = o(\varepsilon^2), && u^\varepsilon(\cdot) \overset{\triangle}{=} \bar{u}(\cdot) + \varepsilon u_1(\cdot) + \varepsilon^2 u_2(\cdot) + \eta^\varepsilon(\cdot) \in \mathcal{U}^4.
\end{aligned}$$

Denote by $x^\varepsilon(\cdot)$ the solution of (1.1) corresponding to ν_0^ε and $u^\varepsilon(\cdot)$. By (**AS1**)–(**AS7**) and Lemma 4.2, for any $t\in[0,T]$, we have

$$\begin{aligned}\mathbb{E}g^0(x^\varepsilon(t)) &= \mathbb{E}g^0(\bar x(t)) + \varepsilon\mathbb{E}\left\langle g_x^0(\bar x(t)), x_1(t)\right\rangle_H + \varepsilon^2\mathbb{E}\left\langle g_x^0(\bar x(t)), x_2(t)\right\rangle_H\\ &\quad+\frac{\varepsilon^2}{2}\mathbb{E}\left\langle g_{xx}^0(\bar x(t))x_1(t), x_1(t)\right\rangle_H + o(\varepsilon^2).\end{aligned}\tag{4.16}$$

Step 2: Fix an arbitrary $\hat t\in\mathbb{I}(\bar x, x_1)$. In this step, we prove that there exist $\delta(\hat t)>0$ and $\alpha(\hat t)>0$ such that

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

If (4.17) is false, then for any $\ell\in\mathbb{N}$, we can find $\varepsilon_\ell\in[0,1/\ell]$ and $s_\ell\in(\hat t-1/\ell,\hat t+1/\ell)\cap[0,T]$ such that

$$\mathbb{E}g^0(x^{\varepsilon_\ell}(s_\ell))>0.\tag{4.18}$$

We consider two different cases.

Case 1.1. There exists a subsequence $\{s_{\ell_k}\}_{k=1}^\infty$ of $\{s_\ell\}_{\ell=1}^\infty$ satisfying

$$\mathbb{E}g^0(\bar x(s_{\ell_k}))<0 \ \text{ and }\ \mathbb{E}\left\langle g_x^0(\bar x(s_{\ell_k})), x_1(s_{\ell_k})\right\rangle_H>0,\ \forall\, k=1,2,\cdots.\tag{4.19}$$

By (4.16),

$$\begin{aligned}&\mathbb{E}g^0(x^{\varepsilon_{\ell_k}}(s_{\ell_k}))\\ &=\varepsilon_{\ell_k}^2\Big(\mathbb{E}\left\langle g_x^0(\bar x(s_{\ell_k})), x_2(s_{\ell_k})\right\rangle_H+\frac12\mathbb{E}\left\langle g_{xx}^0(\bar x(s_{\ell_k}))x_1(s_{\ell_k}), x_1(s_{\ell_k})\right\rangle_H-\frac{\left|\mathbb{E}\left\langle g_x^0(\bar x(s_{\ell_k})), x_1(s_{\ell_k})\right\rangle_H\right|^2}{4\mathbb{E}\ g^0(\bar x(s_{\ell_k}))}\\ &\qquad+\frac{o(\varepsilon_{\ell_k}^2)}{\varepsilon_{\ell_k}^2}\Big)+\mathbb{E}g^0(\bar x(s_{\ell_k}))\Big(1+\frac{\varepsilon_{\ell_k}\mathbb{E}\left\langle g_x(\bar x(s_{\ell_k})), x_1(s_{\ell_k})\right\rangle_H}{2\mathbb{E}g^0(\bar x(s_{\ell_k}))}\Big)^2.\end{aligned}$$

Since $\hat t\in\mathbb{I}^0(\bar x,x_1)$ and $x_2(\cdot)\in\mathcal{G}^0_{(2)}(x_1)$, there exists $\rho_0>0$ such that

$$\mathbb{E}\left\langle g_x^0(\bar x(\hat t)), x_2(\hat t)\right\rangle_H+\frac12\mathbb{E}\left\langle g_{xx}^0(\bar x(\hat t))x_1(\hat t), x_1(\hat t)\right\rangle_H+e(\hat t)<-\rho_0.$$

Therefore, when k is large enough,

$$\mathbb{E}\left\langle g_x^0(\bar x(s_{\ell_k})), x_2(s_{\ell_k})\right\rangle_H+\frac12\mathbb{E}\left\langle g_{xx}^0(\bar x(s_{\ell_k}))x_1(s_{\ell_k}), x_1(s_{\ell_k})\right\rangle_H+\frac{\left|\mathbb{E}\left\langle g_x^0(\bar x(s_{\ell_k})), x_1(s_{\ell_k})\right\rangle_H\right|^2}{4\left|\mathbb{E}g^0(\bar x(s_{\ell_k}))\right|}<-\frac{\rho_0}{2},$$

which, together with (4.19), implies that $\mathbb{E}g^0(x^{\varepsilon_{\ell_k}}(s_{\ell_k}))\le 0$, provided that k is large enough. This contradicts (4.18).

Case 1.2: There is no subsequence of $\{s_\ell\}_{\ell=1}^\infty$ such that (4.19) holds.

Under this circumstance,

$$\mathbb{E}g^0(\bar x(s_\ell))=0 \text{ or } \mathbb{E}\left\langle g_x^0(\bar x(s_\ell)), x_1(s_\ell)\right\rangle_H\le 0 \text{ for all sufficiently large } \ell.$$

If $s_\ell\notin\mathcal{I}^0(\bar x)$, we have $\mathbb{E}g^0(\bar x(s_\ell))<0$. Thus, $\mathbb{E}\left\langle g_x^0(\bar x(s_\ell)), x_1(s_\ell)\right\rangle_H\le 0$. On the other hand, if $s_\ell\in\mathcal{I}^0(\bar x)$, then $\mathbb{E}g^0(\bar x(s_\ell))=0$. Since $x_1(\cdot)\in\mathrm{cl}\mathcal{G}^0_{(1)}$, $\mathbb{E}\left\langle g_x^0(\bar x(s_\ell)), x_1(s_\ell)\right\rangle_H\le 0$. In both cases,

$$\mathbb{E}g^0(\bar x(s_\ell))+\varepsilon_\ell\mathbb{E}\left\langle g_x^0(\bar x(s_\ell)), x_1(s_\ell)\right\rangle_H\le 0.\tag{4.20}$$

Noting that $e(t)\ge 0$ for all $t\in[0,T]$ and $\mathbb{I}^0(\bar x,x_1)$ is compact, there exists $\rho_2>0$ such that

$$\mathbb{E}\left\langle g_x^0(\bar x(t)), x_2(t)\right\rangle_H+\frac12\mathbb{E}\left\langle g_{xx}^0(\bar x(t))x_1(t), x_1(t)\right\rangle_H<-\rho_2,\quad\forall\, t\in\mathbb{I}^0(\bar x,x_1).$$

Since $s_\ell\to\hat t$ and $\hat t\in\mathbb{I}(\bar x,x_1)$, when ℓ is large enough,

$$\mathbb{E}\left\langle g_x^0(\bar x(s_\ell)), x_2(s_\ell)\right\rangle_H+\frac12\mathbb{E}\left\langle g_{xx}^0(\bar x(s_\ell))x_1(s_\ell), x_1(s_\ell)\right\rangle_H<-\frac{\rho_2}{2}.$$

Then, by (4.16) and (4.20), for any sufficiently large ℓ,

$$\begin{aligned}\mathbb{E}g^0(x^{\varepsilon_\ell}(s_\ell)) &\leq \varepsilon_\ell^2\mathbb{E}\left\langle g_x^0(\bar{x}(s_\ell)), x_2(s_\ell)\right\rangle_H + \frac{\varepsilon_\ell^2}{2}\mathbb{E}\left\langle g_{xx}^0(\bar{x}(s_\ell))x_1(s_\ell), x_1(s_\ell)\right\rangle_H + o(\varepsilon_\ell^2)\\ &\leq \varepsilon_\ell^2\Big(-\frac{\rho_2}{2} + \frac{o(\varepsilon_\ell^2)}{\varepsilon_\ell^2}\Big) \leq 0,\end{aligned}$$

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

Step 3: In this step, we prove that $(x^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in \mathcal{P}_{ad}$, provided that ε is sufficiently small.

By the compactness of $\mathbb{I}^0(\bar{x}, x_1)$, we can find $\{t_\ell\}_\ell^N \subset \mathbb{I}^0(\bar{x}, x_1)$ ($N \in \mathbb{N}$) such that

$$\mathbb{I}^0(\bar{x}, x_1) \subset \bigcup_{\ell=1}^{N}\big(t_\ell - \delta(t_\ell), t_\ell + \delta(t_\ell)\big).$$

Let $\varepsilon_1 \triangleq \min\{\alpha(t_\ell),\ \ell = 1, 2, \ldots, N\}$. Then we have that

$$\mathbb{E}g^0(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in \bigcup_{\ell=1}^{N}\big(t_\ell - \delta(t_\ell), t_\ell + \delta(t_\ell)\big) \cap [0,T],\ \forall\, \varepsilon \in [0, \varepsilon_1]. \tag{4.21}$$

Let $\mathcal{I}_0^c \triangleq \mathcal{I}^0(\bar{x}) \setminus \bigcup_{\ell=1}^N(t_\ell - \delta(t_\ell), t_\ell + \delta(t_\ell))$. Since $\mathcal{I}_0^c$ is compact, we can find $\tilde{\delta} > 0$ and $\rho_3 > 0$ (independent of t) such that

$$\mathbb{E}\left\langle g_x^0(\bar{x}(s)), x_1(s)\right\rangle_H < -\rho_3, \quad \forall\, s \in (t - \tilde{\delta}, t + \tilde{\delta}) \cap [0,T], \quad t \in \mathcal{I}_0^c.$$

This, together with (4.16), implies that there exists $\varepsilon_2 > 0$ such that

$$\mathbb{E}g^0(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in (t - \tilde{\delta}, t + \tilde{\delta}) \cap [0,T],\ \forall\, t \in \mathcal{I}_0^c,\ \forall\, \varepsilon \in [0, \varepsilon_2]. \tag{4.22}$$

Clearly,

$$\mathcal{I}^0(\bar{x}) \subset \Big[\bigcup_{\ell=1}^{N}(t_\ell - \delta(t_\ell), t_\ell + \delta(t_\ell))\Big] \bigcup \Big[\bigcup_{t \in \mathcal{I}_0^c}(t - \tilde{\delta}, t + \tilde{\delta})\Big].$$

Let $\delta_0 > 0$ be small enough such that

$$\mathcal{I}^0(\bar{x}) \subset \bigcup_{t \in \mathcal{I}^0(\bar{x})}(t - \delta_0, t + \delta_0) \subset \Big[\bigcup_{\ell=1}^{N}(t_\ell - \delta(t_\ell), t_\ell + \delta(t_\ell))\Big] \bigcup \Big[\bigcup_{t \in \mathcal{I}_0^c}(t - \tilde{\delta}, t + \tilde{\delta})\Big].$$

Put $\varepsilon_0 = \min\{\varepsilon_1, \varepsilon_2\}$. It follows from (4.22) that

$$\mathbb{E}g^0(x^\varepsilon(s)) \leq 0, \quad \forall\, s \in (t - \delta_0, t + \delta_0) \cap [0,T],\ \forall\, t \in \mathcal{I}^0(\bar{x}),\ \forall\, \varepsilon \in [0, \varepsilon_0]. \tag{4.23}$$

Set

$$\mathcal{I}^{cc} \triangleq [0,T] \setminus \Big[\bigcup_{t \in \mathcal{I}^0(\bar{x})}(t - \delta_0, t + \delta_0)\Big].$$

From the compactness of $\mathcal{I}^{cc}$ and the continuity if $\mathbb{E}g^0(\bar{x}(\cdot))$ with respect to t, we know that there exists $\rho_4 > 0$ such that

$$\mathbb{E}g^0(\bar{x}(t)) < -\rho_4, \quad \forall\, t \in \mathcal{I}^{cc}.$$

This, together with (4.16), implies that for all sufficiently small $\varepsilon > 0$,

$$\mathbb{E}g^0(x^\varepsilon(t)) \leq 0, \quad \forall\, t \in \mathcal{I}^{cc}. \tag{4.24}$$

Combining (4.23) and (4.24), we conclude that $x^\varepsilon(\cdot)$ satisfies the constraint (1.3), provided that ε is small enough.

By a similar argument, we can show that when ε is small enough, $x^\varepsilon(\cdot)$ satisfies the constraint (1.4). This proves that $(x^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in \mathcal{P}_{ad}$, provided that ε is sufficiently small.

Step 4: By the optimality of $(\bar{x}(\cdot), \bar{u}(\cdot))$ and the equality $\mathbb{E}\left\langle h_x(\bar{x}(T)), x_1(T)\right\rangle_H = 0$, we have

$$
\begin{aligned}
0 &\le \lim_{\varepsilon\to 0^+} \frac{\mathbb{E}h(x^\varepsilon(T)) - \mathbb{E}h(\bar x(T))}{\varepsilon^2}\\
&= \mathbb{E}\left\langle h_x(\bar x(T)), x_2(T)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle h_{xx}(\bar x(T))x_1(T), x_1(T)\right\rangle_H + \lim_{\varepsilon\to 0^+}\frac{o(\varepsilon^2)}{\varepsilon^2} \qquad (4.25)\\
&= \mathbb{E}\left\langle h_x(\bar x(T)), x_2(T)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle h_{xx}(\bar x(T))x_1(T), x_1(T)\right\rangle_H .
\end{aligned}
$$

From the definition of the transposition solution of the equation (3.17), we get that

$$
\begin{aligned}
&\mathbb{E}\left\langle y(T), x_2(T)\right\rangle_H\\
&= \left\langle y(0), \nu_2\right\rangle_H + \mathbb{E}\int_0^T \Big(\left\langle y(t), a_2[t]u_2(t)\right\rangle_H + \frac{1}{2}\left\langle y(t), a_{11}[t](x_1(t), x_1(t))\right\rangle_H + \left\langle y(t), a_{12}[t](x_1(t), u_1(t))\right\rangle_H\\
&\quad + \frac{1}{2}\left\langle y(t), a_{22}[t](u_1(t), u_1(t))\right\rangle_H + \left\langle Y(t), b_2[t]u_2(t)\right\rangle_{\mathcal{L}_2} + \frac{1}{2}\left\langle Y(t), b_{11}[t](x_1(t), x_1(t))\right\rangle_{\mathcal{L}_2} \qquad (4.26)\\
&\quad + \left\langle Y(t), b_{12}[t](x_1(t), u_1(t))\right\rangle_{\mathcal{L}_2} + \frac{1}{2}\left\langle Y(t), b_{22}[t](u_1(t), u_1(t))\right\rangle_{\mathcal{L}_2}\Big)dt + \mathbb{E}\int_0^T \left\langle x_2(t), d\psi(t)\right\rangle_H .
\end{aligned}
$$

This, together with the choice of $y(T)$, implies that

$$
\begin{aligned}
&\mathbb{E}\left\langle h_x(\bar x(T)), x_2(T)\right\rangle_H\\
&= -\left\langle y(0), \nu_2\right\rangle_H - \sum_{j\in\mathcal{I}(\bar x)} \lambda_j \left\langle g_x^j(\bar x(T)), x_2(T)\right\rangle_H - \mathbb{E}\int_0^T \left\langle x_2(t), d\psi(t)\right\rangle_H\\
&\quad - \mathbb{E}\int_0^T \Big(\left\langle y(t), a_2[t]u_2(t)\right\rangle_H + \frac{1}{2}\left\langle y(t), a_{11}[t](x_1(t), x_1(t))\right\rangle_H + \left\langle y(t), a_{12}[t](x_1(t), u_1(t))\right\rangle_H \qquad (4.27)\\
&\qquad + \frac{1}{2}\left\langle y(t), a_{22}[t](u_1(t), u_1(t))\right\rangle_H + \left\langle Y(t), b_2[t]u_2(t)\right\rangle_{\mathcal{L}_2} + \frac{1}{2}\left\langle Y(t), b_{11}[t](x_1(t), x_1(t))\right\rangle_{\mathcal{L}_2}\\
&\qquad + \left\langle Y(t), b_{12}[t](x_1(t), u_1(t))\right\rangle_{\mathcal{L}_2} + \frac{1}{2}\left\langle Y(t), b_{22}[t](u_1(t), u_1(t))\right\rangle_{\mathcal{L}_2}\Big)dt\\
&= -\left\langle y(0), \nu_2\right\rangle_H - \mathbb{E}\int_0^T \left\langle x_2(t), d\psi(t)\right\rangle_H - \mathbb{E}\int_0^T \left\langle \mathbb{H}_u[t], u_2(t)\right\rangle_{H_1} dt\\
&\quad - \frac{1}{2}\mathbb{E}\int_0^T \big(\langle \mathbb{H}_{xx}[t]x_1(t), x_1(t)\rangle_H + 2\langle \mathbb{H}_{xu}[t]x_1(t), u_1(t)\rangle_{H_1} + \langle \mathbb{H}_{uu}[t]u_1(t), u_1(t)\rangle_{H_1}\big)dt.
\end{aligned}
$$

By the definition of the relaxed transposition solution of (4.1), we have that

$$
\begin{aligned}
&\mathbb{E}\left\langle P(T)x_1(T), x_1(T)\right\rangle_H\\
&= \left\langle P(0)\nu_1, \nu_1\right\rangle_H + \mathbb{E}\int_0^T \big(2\left\langle P(t)x_1(t), a_2[t]u_1(t)\right\rangle_H + 2\left\langle P(t)b_1[t]x_1(t), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2} \qquad (4.28)\\
&\quad + \left\langle P(t)b_2[t]u_1(t), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2} + \left\langle \big(\widehat{Q}^{(0)} + Q^{(0)}\big)\big(0, a_2[t]u_1(t), b_2[t]u_1(t)\big), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2}\\
&\quad - \left\langle \mathbb{H}_{xx}[t]x_1(t), x_1(t)\right\rangle_H \big)dt.
\end{aligned}
$$

This, together with (4.27) and (4.25), implies (4.15).

Step 5: In this step, we handle the case when $x_2(\cdot) \in \mathcal{X}_{(2)}(x_1, u_1) \cap \mathrm{cl}\mathcal{G}^0_{(2)}(x_1) \cap \mathrm{cl}\mathcal{G}_{(2)}(x_1)$.

Let $\hat x_2(\cdot) \in \mathcal{X}_{(2)}(x_1, u_1) \cap \mathrm{cl}\mathcal{G}^0_{(2)}(x_1) \cap \mathrm{cl}\mathcal{G}_{(2)}(x_1)$ with the corresponding $\hat\nu_2 \in \mathcal{W}(\bar\nu_0, \nu_1)$ and $\hat u_2(\cdot) \in \mathcal{M}(\bar u, u_1)$. For $\theta \in (0, 1)$, put

$$
x_2^\theta = (1-\theta)x_2 + \theta\hat x_2.
$$

Noting that $\mathcal{W}(\bar\nu_0, \nu_1)$ and $\mathcal{M}(\bar u, u_1)$ are convex, x_2^θ is the solution of the equation (4.5) with the initial datum

$$\nu_2^\theta \triangleq (1-\theta)\nu_2 + \theta\hat{\nu}_2 \in \mathcal{W}(\bar{\nu}_0,\nu_1)$$

and the control

$$u_2^\theta(\cdot) \triangleq (1-\theta)u_2(\cdot) + \theta\hat{u}_2(\cdot) \in \mathcal{M}(\bar{u},u_1).$$

Then, it is easy to show that

$$\lim_{\theta\to 0} x_2^\theta = x_2 \text{ in } L^2_{\mathbb{F}}(\Omega;C([0,T];H)).$$

Furthermore, since $\hat{x}_2(\cdot) \in \mathcal{G}^0_{(2)}(x_1)\cap\mathcal{G}_{(2)}(x_1)$, we have $x_2^\theta(\cdot) \in \mathcal{G}^0_{(2)}(x_1)\cap\mathcal{G}_{(2)}(x_1)$ for $\theta\neq 0$. From **Step 1**, we deduce that

$$\begin{aligned}
&\langle y(0),\nu_2^\theta\rangle_H + \frac{1}{2}\langle P(0)\nu_1,\nu_1\rangle_H + \sum_{j\in\mathcal{I}(\bar{x})}\mathbb{E}\big\langle \lambda_j g_x^j(\bar{x}(T)), x_2^\theta(T)\big\rangle_H\\
&+\mathbb{E}\int_0^T \Big(\big\langle \mathbb{H}_u[t], u_2^\theta(t)\big\rangle_{H_1} + \frac{1}{2}\big\langle \mathbb{H}_{uu}[t]u_1(t),u_1(t)\big\rangle_{H_1} + \frac{1}{2}\big\langle b_2[t]^*P(t)b_2[t]u_1(t),u_1(t)\big\rangle_{H_1}\\
&\qquad+\big\langle\big(\mathbb{H}_{xu}[t]+a_2[t]^*P(t)+b_2[t]^*P(t)b_1[t]\big)x_1(t),u_1(t)\big\rangle_{H_1}\\
&\qquad+\frac{1}{2}\big\langle\big(\widehat{Q}^{(0)}+Q^{(0)}\big)\big(0,a_2[t]u_1(t),b_2[t]u_1(t)\big),b_2[t]u_1(t)\big\rangle_{\mathcal{L}_2}\Big)dt+\mathbb{E}\int_0^T\big\langle x_2^\theta(t),d\psi(t)\big\rangle_H \le 0.
\end{aligned}$$

Letting $\theta\to 0$ in the above inequality, we obtain (4.15). This completes the proof of Theorem 4.1. □

Remark 4.8. *The second order necessary condition is only valid for $\mathcal{Y}(\bar{x},\bar{u})$ (recall (4.7) for the definition) being nonempty. If $\mathcal{G}_{(1)}(T)\neq\emptyset$, $U=H_1$, (3.1) is exactly controllable and there are no state constraints, then $\mathcal{Y}(\bar{x},\bar{u})\neq\emptyset$. However, to enjoy the exact controllability property, one needs some restrictive conditions (e.g.[24, 25, 26]).*

Next, we give another second order necessary condition.

Theorem 4.2. *Suppose that* **(AS1)**–**(AS9)** *hold and $(\bar{x}(\cdot),\bar{u}(\cdot),\bar{\nu}_0)$ be an optimal triple of* **Problem (OP)**. *Let $\Phi(t,\omega)=\mathcal{C}_U(\bar{u}(t,\omega))$. Assume that $\mathbb{E}|g_x^0(\bar{x}(t))|_H\neq 0$ for all $t\in\mathcal{I}^0(\bar{x})$. Let $(x_1,u_1,\nu_1)\in\Upsilon(\bar{x},\bar{u})$ and suppose that $e(\cdot)$ (defined by (4.11)) is bounded on $\mathbb{I}^0(\bar{x},x_1)$. Let $\mathcal{W}(\bar{\nu}_0,\nu_1)\subset T^{b(2)}_{\mathcal{V}}(\bar{x}_0,\nu_1)$ and $\mathcal{M}(\bar{u},u_1)\subset T^{b(2)}_{\mathcal{U}^4}(\bar{u},u_1)$ be convex. Then there exist $\lambda_0\in\{0,1\}$, $\lambda_j\geq 0$ for all $j\in\mathcal{I}(\bar{x})$ and $\psi\in\big(\mathcal{G}^0_{(1)}\big)^-$ such that the solution (y,Y) of (3.17) with $y_T=-\lambda_0 h_x(\bar{x}(T))-\sum_{j\in\mathcal{I}(\bar{x})}\lambda_j g_x^j(\bar{x}(T))$ and $\mathcal{I}(\bar{x})$ replaced by $\mathbb{I}(\bar{x},x_1)$ satisfies the first order condition (3.32), and for any $x_2(\cdot)\in\mathcal{X}_{(2)}(x_1,u_1)$ with the corresponding $\nu_2\in\mathcal{W}(\bar{\nu}_0,\nu_1)$ and $u_2(\cdot)\in\mathcal{M}(\bar{u},u_1)$, the second order necessary condition* (4.15) *holds true, where $(P(\cdot),Q^{(\cdot)},\widehat{Q}^{(\cdot)})$ is the relaxed transposition solution of (4.1) with $P(T)=-\lambda_0 h_{xx}(\bar{x}(T))-\sum_{j=1}^n\lambda_j g_{xx}^j(\bar{x}(T))$.*

Proof. If either $\mathcal{W}(\bar{\nu}_0,\nu_1)$ or $\mathcal{M}(\bar{u},u_1)$ is empty, then by Theorem 3.2, we get the desired result. Therefore, in the rest of the proof, we assume that these two sets are nonempty. Put

$$\widetilde{\mathcal{W}}(\bar{x}_0,\nu_1)\triangleq\mathcal{C}_{\mathcal{V}}(\bar{x}_0)+\mathcal{W}(\bar{x}_0,\nu_1),\quad \widetilde{\mathcal{M}}(\bar{u},u_1)\triangleq\mathcal{T}_\Phi(\bar{u})+\mathcal{M}(\bar{u},u_1),$$

where

$$\mathcal{T}_\Phi(\bar{u})\triangleq\{u\in L^4_{\mathbb{F}}(0,T;H_1)\mid u(t,\omega)\in\mathcal{C}_U(\bar{u}(t,\omega)) \text{ a.e. in } [0,T]\times\Omega\}.$$

By Lemma 2.2, $\mathcal{T}_\Phi(\bar{u})\subset\mathcal{C}_{\mathcal{U}^4}(\bar{u})$. Thus, by Lemma 2.4 from [17], $\widetilde{\mathcal{M}}(\bar{u},u_1)\subset T^{b(2)}_{\mathcal{U}^4}(\bar{u},u_1)$.

We divide the rest of the proof into two steps. In **Step 1**, we handle the case when $\mathbb{I}^0(\bar{x},x_1)=\emptyset$. In **Step 2**, we deal with the case when $\mathbb{I}^0(\bar{x},x_1)\neq\emptyset$.

Step 1. If $\mathbb{I}^0(\bar{x},x_1)=\emptyset$, then $\mathcal{G}^0_{(2)}(x_1)=L^2_{\mathbb{F}}(\Omega;C([0,T];H))$ and

$$\mathcal{X}_{(2)}(x_1,u_1)\cap\mathcal{G}^0_{(2)}(x_1)\cap\mathcal{G}_{(2)}(x_1)=\mathcal{X}_{(2)}(x_1,u_1)\cap\mathcal{G}_{(2)}(x_1).$$

Fix $(x_1(\cdot),u_1(\cdot),\nu_1)\in\mathcal{Z}(\bar{x},\bar{u})$ (recall (4.8) for the definition of $\mathcal{Z}(\bar{x},\bar{u})$). Consider the following two different cases:

Case 1.1: $\mathbb{I}(\bar{x},x_1)=\emptyset$.

In this context,

$$\mathbb{E}\langle h_x(\bar{x}(T)),x_1(T)\rangle_H=0,\qquad \mathbb{E}\langle g^j_x(\bar{x}(T)),x_1(T)\rangle_H<0,\quad \forall\, j\in\mathcal{I}(\bar{x}).$$

Then for any $\nu_2\in\widetilde{\mathcal{W}}(\bar{\nu}_0,\nu_1)$, $u_2\in\widetilde{\mathcal{M}}(\bar{u},u_1)$ and $\varepsilon>0$, there exist $\nu^\varepsilon\in H$ and $v^\varepsilon\in L^4_{\mathbb{F}}(0,T;H_1)$ such that

$$|\nu^\varepsilon|_H=o(\varepsilon^2),\qquad \nu^\varepsilon_0\overset{\triangle}{=}\bar{\nu}_0+\varepsilon\nu_1+\varepsilon^2\nu_2+\nu^\varepsilon\in\mathcal{V}$$

and

$$|v^\varepsilon|_{L^4_{\mathbb{F}}(0,T;H_1)}=o(\varepsilon^2),\qquad u^\varepsilon\overset{\triangle}{=}\bar{u}+\varepsilon u_1+\varepsilon^2u_2+v^\varepsilon\in\mathcal{U}^4.$$

Let $x^\varepsilon(\cdot)$ be the solution of the control system (1.1) with the initial datum ν^ε_0 and the control $u^\varepsilon(\cdot)$. Put

$$h^\varepsilon_{11}(T)\overset{\triangle}{=}\int_0^1(1-\theta)h_{xx}(\bar{x}(T)+\theta\delta x^\varepsilon(T))d\theta.$$

By Lemma 4.2, there is $\rho<0$ such that for each $j\in\mathcal{I}(\bar{x})$ and all sufficiently small $\varepsilon>0$,

$$\begin{aligned}\mathbb{E}g^j(x^\varepsilon(T))&=\mathbb{E}g^j(\bar{x}(T))+\varepsilon\mathbb{E}\langle g^j_x(\bar{x}(T)),x_1(T)\rangle_H+\varepsilon^2\mathbb{E}\langle g^j_x(\bar{x}(T)),x_2(T)\rangle_H\\&\quad+\frac{\varepsilon^2}{2}\mathbb{E}\langle g^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H+o(\varepsilon^2)\\&=\varepsilon\Big(\mathbb{E}\langle g^j_x(\bar{x}(T)),x_1(T)\rangle_H+\varepsilon\mathbb{E}\langle g^j_x(\bar{x}(T)),x_2(T)\rangle_H\\&\quad+\frac{\varepsilon}{2}\mathbb{E}\langle g^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H+o(\varepsilon)\Big)<\varepsilon\rho<0\end{aligned}$$

and, for each $j\notin\mathcal{I}(\bar{x})$, $\mathbb{E}g^j(x^\varepsilon(T))=\mathbb{E}g^j(\bar{x}(T))+O(\varepsilon)\le\rho+O(\varepsilon)$. Consequently, $(x^\varepsilon(\cdot),u^\varepsilon(\cdot))\in\mathcal{P}_{ad}$.

Direct computations yield

$$\begin{aligned}&\frac{\mathcal{J}(u^\varepsilon)-\mathcal{J}(\bar{u})}{\varepsilon^2}\\&=\frac{1}{\varepsilon^2}\mathbb{E}\Big(\langle h_x(\bar{x}(T)),\delta x^\varepsilon(T)\rangle_H+\frac{1}{2}\langle h^\varepsilon_{11}(\bar{x}(T))\delta x^\varepsilon(T),\delta x^\varepsilon(T)\rangle_H\Big)\\&=\mathbb{E}\Big(\frac{1}{\varepsilon}\langle h_x(\bar{x}(T)),x_1(T)\rangle_H+\langle h_x(\bar{x}(T)),x_2(T)\rangle_H+\frac{1}{2}\langle h_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H\Big)+\rho^\varepsilon_2,\end{aligned}$$

where

$$\rho^\varepsilon_2=\mathbb{E}\Big(\frac{1}{2}\Big\langle h^\varepsilon_{11}(\bar{x}(T))\frac{\delta x^\varepsilon(T)}{\varepsilon},\frac{\delta x^\varepsilon(T)}{\varepsilon}\Big\rangle_H-\frac{1}{2}\langle h_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H\Big).$$

Similar to the proof of Lemma 4.2, we can show that $\lim\limits_{\varepsilon\to0^+}\rho^\varepsilon_2=0$. Therefore,

$$0\le\lim_{\varepsilon\to0^+}\frac{\mathcal{J}(u^\varepsilon(\cdot))-\mathcal{J}(\bar{u}(\cdot))}{\varepsilon^2}=\mathbb{E}\Big(\langle h_x(\bar{x}(T)),x_2(T)\rangle_H+\frac{1}{2}\langle h_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H\Big).\qquad(4.29)$$

It follows from the definition of the transposition solution of (3.17) that

$$
\begin{aligned}
&\mathbb{E}\left\langle y(T), x_2(T)\right\rangle_H \\
&= \mathbb{E}\left\langle y(0), \nu_2\right\rangle_H + \mathbb{E}\int_0^T \Big(\left\langle y(t), a_2[t]u_2(t)\right\rangle_H + \frac{1}{2}\left\langle y(t), a_{11}[t]\big(x_1(t), x_1(t)\big)\right\rangle_H \\
&\quad + \left\langle y(t), a_{12}[t]\big(x_1(t), u_1(t)\big)\right\rangle_{H_1} + \frac{1}{2}\left\langle y(t), a_{22}[t]\big(u_1(t), u_1(t)\big)\right\rangle_{H_1} \\
&\quad + \left\langle Y(t), b_2[t]u_2(t)\right\rangle_{\mathcal{L}_2} + \frac{1}{2}\left\langle Y(t), b_{11}[t]\big(x_1(t), x_1(t)\big)\right\rangle_{\mathcal{L}_2} \\
&\quad + \left\langle Y(t), b_{12}[t]\big(x_1(t), u_1(t)\big)\right\rangle_{\mathcal{L}_2} + \frac{1}{2}\left\langle Y(t), b_{22}[t]\big(u_1(t), u_1(t)\big)\right\rangle_{\mathcal{L}_2} \Big) dt.
\end{aligned}
\tag{4.30}
$$

By the definition of the relaxed transposition solution of (4.1), we have

$$
\begin{aligned}
&\mathbb{E}\left\langle P(T)x_1(T), x_1(T)\right\rangle_H \\
&= \mathbb{E}\left\langle P(0)\nu_1, \nu_1\right\rangle_H + \mathbb{E}\int_0^T \Big(2\left\langle P(t)x_1(t), a_2[t]u_1(t)\right\rangle_H + 2\left\langle P(t)b_1[t]x_1(t), b_2[t]u_1(t)\right\rangle_H \\
&\quad + \left\langle P(t)b_2[t]u_1(t), b_2[t]u_1(t)\right\rangle_H - \left\langle \mathbb{H}_{xx}(t)x_1(t), x_1(t)\right\rangle_H \Big) dt \\
&\quad + \mathbb{E}\int_0^T \left\langle \widehat{Q}^{(0)}(0, a_2u_1, b_2u_1)(t) + Q^{(0)}(0, a_2u_1, b_2u_1)(t), b_2(t)u_1(t)\right\rangle_{\mathcal{L}_2} dt.
\end{aligned}
\tag{4.31}
$$

Let $\lambda_0 = 1$, $\lambda_j = 0$ for all $j \in \mathcal{I}(\bar{x})$ and $\psi = 0$. It follows from (4.29)–(4.31) that

$$
\begin{aligned}
0 &\geq \mathbb{E}\left\langle y(0), \nu_2\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle P(0)\nu_1, \nu_1\right\rangle_H + \mathbb{E}\int_0^T \Big[\left\langle y(t), a_2[t]u_2(t)\right\rangle_H + \left\langle Y(t), b_2[t]u_2(t)\right\rangle_{\mathcal{L}_2} \\
&\quad + \frac{1}{2}\Big(\left\langle y(t), a_{22}[t]\big(u_1(t), u_1(t)\big)\right\rangle_H + \left\langle Y(t), b_{22}[t]\big(u_1(t), u_1(t)\big)\right\rangle_{\mathcal{L}_2} + \left\langle P(t)b_2[t]u_1(t), b_2[t]u_1(t)\right\rangle_H \Big) \\
&\quad + \left\langle y(t), a_{12}[t]\big(x_1(t), u_1(t)\big)\right\rangle_H + \left\langle Y(t), b_{12}[t]\big(x_1(t), u_1(t)\big)\right\rangle_{\mathcal{L}_2} \\
&\quad + \left\langle a_2[t]^*P(t)x_1(t), u_1(t)\right\rangle_{H_1} + \left\langle b_2[t]^*P(t)b_1[t]x_1(t), u_1(t)\right\rangle_{H_1} \Big] dt \\
&\quad + \frac{1}{2}\mathbb{E}\int_0^T \left\langle \widehat{Q}^{(0)}(0, a_2u_1, b_2u_1)(t) + Q^{(0)}(0, a_2v, b_2u_1)(t), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2} dt \\
&= \mathbb{E}\left\langle y(0), \nu_2\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle P(0)\nu_1, \nu_1\right\rangle_H \\
&\quad + \mathbb{E}\int_0^T \Big(\left\langle \mathbb{H}_u[t], u_2(t)\right\rangle_{H_1} + \frac{1}{2}\left\langle \mathbb{H}_{uu}(t)u_1(t), u_1(t)\right\rangle_{H_1} + \frac{1}{2}\left\langle b_2[t]^*P(t)b_2[t]u_1(t), u_1(t)\right\rangle_{H_1} \Big) dt \\
&\quad + \mathbb{E}\int_0^T \left\langle \big(\mathbb{H}_{xu}(t) + a_2[t]^*P(t) + b_2[t]^*P(t)b_1[t]\big)y(t), u_1(t)\right\rangle_{H_1} dt \\
&\quad + \frac{1}{2}\mathbb{E}\int_0^T \left\langle \widehat{Q}^{(0)}(0, a_2[t]u_1, b_2[t]u_1)(t) + Q^{(0)}(0, a_2[t]u_1, b_2[t]u_1)(t), b_2[t]u_1(t)\right\rangle_{\mathcal{L}_2} dt.
\end{aligned}
$$

Case 1.2: $\mathbb{I}(\bar{x}, x_1) \neq \emptyset$.

First, we claim that

$$
\mathcal{E}^2(x_1) \cap \mathcal{H}_{(2)}(x_1) \cap \mathcal{G}^2(x_1, u_1) = \emptyset. \tag{4.32}
$$

Indeed, if (4.32) was false, then there would exist $\nu_2 \in \widetilde{\mathcal{W}}(\bar{\nu}_0, \nu_1)$ and $u_2(\cdot) \in \widetilde{\mathcal{M}}(\bar{u}, u_1)$ such that for some $\rho < 0$, the corresponding solution $x_2(\cdot)$ of (4.5) satisfies

$$
\mathbb{E}\left\langle g_x^j(\bar{x}(T)), x_2(T)\right\rangle_H + \frac{1}{2}\mathbb{E}\left\langle g_{xx}^j(\bar{x}(T))x_1(T), x_1(T)\right\rangle_H < 2\rho, \qquad \forall\, j \in \mathbb{I}(\bar{x}, x_1)
$$

and

$$\mathbb{E}\big\langle h_x(\bar{x}(T)), x_2(T)\big\rangle_H + \frac{1}{2}\mathbb{E}\big\langle h_{xx}(\bar{x}(T))x_1(T), x_1(T)\big\rangle_H < 2\rho.$$

Let $\nu^\varepsilon \in H$ and $v^\varepsilon \in L^4_{\mathbb{F}}(0,T;H_1)$ be such that

$$|\nu^\varepsilon|_H = o(\varepsilon^2), \qquad \nu_0^\varepsilon \overset{\triangle}{=} \bar{\nu}_0 + \varepsilon\nu_1 + \varepsilon^2\nu_2 + \nu^\varepsilon \in \mathcal{V}$$

and

$$|v^\varepsilon|_{L^4_{\mathbb{F}}(0,T;H_1)} = o(\varepsilon^2), \qquad u^\varepsilon(\cdot) \overset{\triangle}{=} \bar{u}(\cdot) + \varepsilon u_1(\cdot) + \varepsilon^2 u_2(\cdot) + v^\varepsilon(\cdot) \in \mathcal{U}^4.$$

Let $x^\varepsilon(\cdot)$ be the solution of the control system (1.1) with the initial datum ν_0^ε and the control $u^\varepsilon(\cdot)$. Similar to **Case 1.1**, one can prove that for every $j \notin \mathbb{I}(\bar{x}, x_1)$ and for all $\varepsilon > 0$ small enough, $\mathbb{E}g^j(x^\varepsilon(T)) \le 0$. Meanwhile, by Lemma 4.2, for any $j \in \mathbb{I}(\bar{x}, x_1)$, and for all sufficiently small $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{E}g^j(x^\varepsilon(T)) &= \mathbb{E}g^j(\bar{x}(T)) + \varepsilon\mathbb{E}\big\langle g_x^j(\bar{x}(T)), x_1(T)\big\rangle_H + \varepsilon^2\mathbb{E}\big\langle g_x^j(\bar{x}(T)), x_2(T)\big\rangle_H \\
&\quad + \frac{\varepsilon^2}{2}\mathbb{E}\big\langle g_{xx}^j(\bar{x}(T))x_1(T), x_1(T)\big\rangle_H + o(\varepsilon^2) \\
&= \varepsilon^2\Big(\mathbb{E}\big\langle g_x^j(\bar{x}(T)), x_2(T)\big\rangle_H + \frac{1}{2}\mathbb{E}\big\langle g_{xx}^j(\bar{x}(T))x_1(T), x_1(T)\big\rangle_H + \frac{o(\varepsilon^2)}{\varepsilon^2}\Big) \\
&< \varepsilon^2\rho < 0.
\end{aligned}$$

This proves that $(x^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in \mathcal{P}_{ad}$.

On the other hand, for all sufficiently small $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{E}h(x^\varepsilon(T)) &= \mathbb{E}h(\bar{x}(T)) + \varepsilon\mathbb{E}\big\langle h_x(\bar{x}(T)), x_1(T)\big\rangle_H + \varepsilon^2\mathbb{E}\big\langle h_x(\bar{x}(T)), x_2(T)\big\rangle_H \\
&\quad + \frac{\varepsilon^2}{2}\mathbb{E}\big\langle h_{xx}(\bar{x}(T))x_1(T), x_1(T)\big\rangle_H + o(\varepsilon^2) \\
&= \mathbb{E}h(\bar{x}(T)) + \varepsilon^2\Big(\mathbb{E}\big\langle h_x(\bar{x}(T)), x_2(T)\big\rangle_H + \frac{1}{2}\mathbb{E}\big\langle h_{xx}(\bar{x}(T))x_1(T), x_1(T)\big\rangle_H + \frac{o(\varepsilon^2)}{\varepsilon^2}\Big) \\
&< \mathbb{E}h(\bar{x}(T)) + \varepsilon^2\rho < \mathbb{E}h(\bar{x}(T)).
\end{aligned}$$

This contradicts the optimality of $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{\nu}_0)$. Hence, (4.32) holds.

Next, we consider two subcases (recall (4.9), (4.13) and (4.14) for the definitions of $\mathcal{X}_{(2)}(x_1, u_1)$, $\mathcal{G}_{(2)}(x_1)$ and $\mathcal{H}_{(2)}(x_1)$).

Case 1.2.1. $\mathcal{H}_{(2)}(x_1) \cap \mathcal{G}_{(2)}(x_1) \neq \emptyset$.

Under these circumstances, $\Gamma\big(\mathcal{H}_{(2)}(x_1)\big) \cap \Gamma\big(\mathcal{G}_{(2)}(x_1)\big) \neq \emptyset$. Since $\Gamma\big(\mathcal{H}_{(2)}(x_1)\big) \cap \Gamma\big(\mathcal{G}_{(2)}(x_1)\big) \cap \Gamma\big(\mathcal{X}_{(2)}(x_1, u_1)\big) = \emptyset$, by the Hahn-Banach separation theorem, we can find a nonzero $\xi \in L^2_{\mathcal{F}_T}(\Omega; H)$ such that

$$\sup_{\alpha \in \Gamma(\mathcal{H}_{(2)}(x_1)) \cap \Gamma(\mathcal{G}_{(2)}(x_1))} \mathbb{E}\langle \xi, \alpha\rangle_H \le \inf_{\beta \in \Gamma(\mathcal{X}_{(2)}(x_1, u_1))} \mathbb{E}\langle \xi, \beta\rangle_H.$$

By Lemma 2.9, there exists

$$\alpha_0 \in \mathrm{cl}\big(\Gamma\big(\mathcal{H}_{(2)}(x_1)\big) \bigcap \Gamma\big(\mathcal{G}_{(2)}(x_1)\big)\big) = \mathrm{cl}\Gamma\big(\mathcal{H}_{(2)}(x_1)\big) \bigcap \Big(\bigcap_{j \in \mathbb{I}(\bar{x}, x_1)} \mathrm{cl}\Gamma\big(\mathcal{E}^{2,j}(x_1(T))\big)\Big)$$

such that

$$\mathbb{E}\langle \xi, \alpha_0\rangle_H = \sup_{\alpha \in \Gamma(\mathcal{H}_{(2)}(x_1)) \cap \Gamma(\mathcal{G}_{(2)}(x_1))} \mathbb{E}\langle \xi, \alpha\rangle_H.$$

Put

$$\mathbb{I}_0(\bar{x}, x_1) \overset{\triangle}{=} \Big\{ j \in \mathbb{I}(\bar{x}, x_1) \,\Big|\, \mathbb{E}\big\langle g_x^j(\bar{x}(T)), \alpha_0\big\rangle_H + \frac{1}{2}\mathbb{E}\big\langle g_{xx}^j(\bar{x}(T))x_1(T), x_1(T)\big\rangle_H = 0 \Big\}.$$

By Lemma 2.9, for every $j \in \mathbb{I}_0(\bar{x}, x_1)$, there exists $\lambda_j \ge 0$ such that

$$\xi = \lambda_0 h_x(\bar{x}(T)) + \sum_{j \in \mathbb{I}_0(\bar{x}, x_1)} \lambda_j g_x^j(\bar{x}(T)), \tag{4.33}$$

where $\lambda_0=0$ if $\mathbb{E}\langle h_x(\bar{x}(T)),\alpha_0\rangle_H+\dfrac{1}{2}\mathbb{E}\langle h_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H<0$. Then (4.33) yields

$$\mathbb{E}\langle\xi,\alpha_0\rangle_H=-\frac{1}{2}\Big(\lambda_0\mathbb{E}\big\langle h_{xx}(\bar{x}(T))x_1(T),x_1(T)\big\rangle_H+\sum_{j\in\mathbb{I}_0(\bar{x},x_1)}\lambda_j\mathbb{E}\big\langle g^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\big\rangle_H\Big).$$

Setting

$$y(T)=-\lambda_0h_x(\bar{x}(T))-\sum_{j\in\mathbb{I}_0(\bar{x},x_1)}\lambda_jg^j_x(\bar{x}(T))$$

and

$$P(T)=-\lambda_0h_{xx}(\bar{x}(T))-\sum_{j\in\mathbb{I}_0(\bar{x},x_1)}\lambda_jg^j_{xx}(\bar{x}(T)),$$

we find that for any $x_2(T)\in\Gamma\big(\mathcal{X}_{(2)}(x_1,u_1)\big)$,

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\langle P(T)x_1(T),x_1(T)\rangle_H\\&=-\frac{1}{2}\Big(\lambda_0\mathbb{E}\langle h_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H+\sum_{j\in\mathbb{I}_0(\bar{x},x_1)}\lambda_j\mathbb{E}\langle g^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\rangle_H\Big)\\&=\mathbb{E}\langle\xi,\alpha_0\rangle_H\le\mathbb{E}\langle y(T),x_2(T)\rangle_H.\end{aligned}$$

This, together with (4.30) and (4.31), implies (4.15).

Case 1.2.2. $\Gamma\big(\mathcal{H}_{(2)}(x_1)\big)\cap\Gamma\big(\mathcal{G}_{(2)}(x_1)\big)=\emptyset$.

For simplicity of notations, we put $g^{n+1}(\cdot)=h(\cdot)$, $\mathbf{I}=\{n+1\}\cup\mathbb{I}(\bar{x},x_1)$ and $\mathcal{G}^{n+1}_{(2)}(x_1)=\mathcal{H}_{(2)}(x_1)$.

If there exists $j\in\mathbf{I}$ such that $\Gamma\big(\mathcal{G}^j_{(2)}(x_1)\big)=\emptyset$, then $g^j_x(\bar{x}(T))=0$, $\mathbb{P}$-a.s. and

$$\mathbb{E}\big\langle g^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\big\rangle_H\ge0. \tag{4.34}$$

Let $\lambda_j=1$ and $\lambda_k=0$ for $k\in\mathbf{I}\setminus\{j\}$. Then $\lambda_jg^j_x(\bar{x}(T))+\sum_{k\in\mathbf{I}\setminus\{j\}}\lambda_kg^k_x(\bar{x}(T))=0$. Let $y(T)=0$ and $P(T)=-g^j_{xx}(\bar{x}(T))$. It is easy to see that $(y(\cdot),Y(\cdot))=(0,0)$, $\mathbb{H}(\cdot)=0$, $\mathbb{H}_{xx}[\cdot]=0$ and by (4.34), $\mathbb{E}\langle P(T)x_1(T),x_1(T)\rangle_H\le0$. Then, by the definition of the relaxed transposition solution of (4.1), (4.15) holds and it is reduced to

$$\begin{aligned}\mathbb{E}\langle P(0)\nu_1,\nu_1\rangle_H+\mathbb{E}\int_0^T\Big[&\big\langle b_2[t]^*P(t)b_2[t]u_1(t),u_1(t)\big\rangle_{H_1}+2\big\langle\big(a_2[t]^*P(t)+b_2[t]^*P(t)b_1[t]\big)x_1(t),u_1(t)\big\rangle_{H_1}\\&+\big\langle\big(\widehat{Q}^{(0)}+Q^{(0)}\big)\big(0,a_2[t]u_1(t),b_2[t]u_1(t)\big),b_2[t]u_1(t)\big\rangle_{\mathcal{L}_2}\Big]dt\le0.\end{aligned}$$

If $\Gamma\big(\mathcal{G}^j_{(2)}(x_1)\big)\ne\emptyset$ for all $j\in\mathbf{I}$, then one can find $j_0\in\mathbf{I}$ and a subset $\mathbf{I}^0\subset\mathbf{I}$ with $j_0\notin\mathbf{I}^0$ such that

$$\bigcap_{j\in\mathbf{I}^0}\Gamma\big(\mathcal{G}^j_{(2)}(x_1)\big)\ne\emptyset,\qquad\Big(\bigcap_{j\in\mathbf{I}^0}\Gamma\big(\mathcal{G}^j_{(2)}(x_1)\big)\Big)\bigcap\Gamma\big(\mathcal{G}^{j_0}_{(2)}(x_1)\big)=\emptyset.$$

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

$$\sup_{\alpha\in\mathcal{E}^{(2,j_0)}_T(x_1)}\mathbb{E}\langle\xi,\alpha\rangle_H\le\inf_{\beta\in\cap_{j\in\mathbf{I}^0}\mathcal{E}^{(2,j)}_T(x_1)}\mathbb{E}\langle\xi,\beta\rangle_H.$$

By Lemma 2.9, we can find $\alpha_0\in\mathrm{cl}\Gamma\big(\mathcal{G}^{j_0}_{(2)}(x_1)\big)$ and $\beta_0\in\bigcap_{j\in\mathbf{I}^0}\mathrm{cl}\Gamma\big(\mathcal{G}^j_{(2)}(x_1)\big)$ such that

$$\mathbb{E}\langle\xi,\alpha_0\rangle_H=\sup_{\alpha\in\Gamma(\mathcal{G}^{j_0}_{(2)}(x_1))}\mathbb{E}\langle\xi,\alpha\rangle_H\le\inf_{\beta\in\cap_{j\in\mathbf{I}^0}\Gamma(\mathcal{G}^j_{(2)}(x_1))}\mathbb{E}\langle\xi,\beta\rangle_H=\mathbb{E}\langle\xi,\beta_0\rangle_H. \tag{4.35}$$

It follows from Lemma 2.9 that there exists $\lambda_{j_0} > 0$ such that $\xi = \lambda_{j_0} g_x^{j_0}(\bar{x}(T))$ and

$$0 = \mathbb{E}\big\langle g_x^{j_0}(\bar{x}(T)), \alpha_0 \big\rangle_H + \frac{1}{2}\mathbb{E}\big\langle g_{xx}^{j_0}(\bar{x}(T))x_1(T), x_1(T) \big\rangle_H. \tag{4.36}$$

Denote by $\mathbf{I}^1$ the set of all indices $j \in \mathbf{I}^0$ satisfying

$$0 = \mathbb{E}\big\langle g_x^{j}(\bar{x}(T)), \alpha_0 \big\rangle_H + \frac{1}{2}\mathbb{E}\big\langle g_{xx}^{j}(\bar{x}(T))x_1(T), x_1(T) \big\rangle_H. \tag{4.37}$$

Then, by Lemma 2.9 once more, for each $j \in \mathbf{I}^1$, there exists $\lambda_j \geq 0$ such that

$$-\xi = -\lambda_{j_0} g_x^{j_0}(\bar{x}(T)) = \sum_{j \in \mathbf{I}^1} \lambda_j g_x^j(\bar{x}(T)). \tag{4.38}$$

Combing (4.35)–(4.38), we obtain that

$$0 \leq \lambda_{j_0}\mathbb{E}\big\langle g_{xx}^{j_0}(\bar{x}(T))x_1(T), x_1(T) \big\rangle_H + \sum_{j \in \mathbf{I}^1} \mathbb{E}\big\langle g_{xx}^{j}(\bar{x}(T))x_1(T), x_1(T) \big\rangle_H.$$

Let $y(T) = 0$ and $P(T) = -\lambda_{j_0} g_{xx}^{j_0}(\bar{x}(T)) - \sum_{j \in \mathbf{I}^1} g_{xx}^j(\bar{x}(T))$. Then

$$(y(\cdot), Y(\cdot)) = (0, 0), \quad \mathbb{H}(\cdot) = 0, \quad \mathbb{H}_{xx}[\cdot] = 0, \quad \mathbb{E}\langle P(\bar{x}(T))x_1(T), x_1(T)\rangle_H \leq 0.$$

Applying the same argument as before, we obtain (4.15) with $\psi = 0$.

Step 2. In this step, we deal with the case that $\mathbb{I}^0(\bar{x}, x_1) \neq \emptyset$.

From $\mathbb{E}|g_x(\bar{x}(t))|_H \neq 0$ for any $t \in \mathcal{I}^0(\bar{x})$ and $e(\cdot)$ (recall (4.11) for the definition of $e(\cdot)$) is bounded on $\mathbb{I}^0(\bar{x}, x_1)$, we get that $-g_x(\bar{x}(\cdot)) \in \mathcal{G}^0_{(1)}$ and $-\delta g_x(\bar{x}(\cdot)) \in \mathcal{G}^0_{(2)}(x_1)$ when δ (> 0) is large enough. Thus, $\mathcal{G}^0_{(1)} \neq \emptyset$ and $\mathcal{G}^0_{(2)}(x_1) \neq \emptyset$.

Let $x_2(\cdot) \in \mathcal{X}_{(2)}(x_1, u_1)$ and $(y(\cdot), Y(\cdot))$ be the transposition solution to (3.17). We deduce from (4.26) that

$$\begin{aligned}
&\mathbb{E}\,\langle y(T), x_2(T)\rangle_H \\
&= \mathbb{E}\,\langle y(0), \nu_2\rangle_H + \mathbb{E}\int_0^T \langle x_2(t), d\psi(t)\rangle_H + \mathbb{E}\int_0^T \langle \mathbb{H}_u[t], u_2(t)\rangle_{H_1} dt \\
&\quad + \frac{1}{2}\mathbb{E}\int_0^T \big(\langle \mathbb{H}_{xx}[t]x_1(t), x_1(t)\rangle_H + 2\,\langle \mathbb{H}_{xu}[t]x_1(t), u_1(t)\rangle_{H_1} + \langle \mathbb{H}_{uu}[t]u_1(t), u_1(t)\rangle_{H_1}\big)dt.
\end{aligned}$$

If $\mathcal{G}^{j_1}_{(2)}(x_1) = \emptyset$ for some $j_1 \in \mathbf{I}$, then $g_x^{j_1}(\bar{x}(T)) = 0$, $\mathbb{P}$-a.s. and $\mathbb{E}\langle g_{xx}^{j_1}(\bar{x}(T))x_1(T), x_1(T)\rangle_H \geq 0$. Therefore, by setting $\psi(\cdot) = 0$, $\lambda_{j_1} = 1$ and $\lambda_j = 0$ for all $j_1 \neq j \in \mathbf{I}$, we get $(y(\cdot), Y(\cdot)) = (0, 0)$, $P(T) = -g_{xx}^{j_1}(\bar{x}(T))$, $\mathbb{E}\,\langle P(T)x_1(T), x_1(T)\rangle_H \leq 0$ and

$$\mathbb{E}\int_0^T \langle \mathbb{H}_{xx}[t]x_1(t), x_1(t)\rangle_H \, dt = 0.$$

These facts, together with (4.28), imply (4.15).

Next, assume that $\mathcal{G}^j_{(2)}(x_1) \neq \emptyset$ for every $j \in \mathbf{I}$. We claim that

$$\mathcal{X}_{(2)}(x_1, u_1) \bigcap \mathcal{G}^0_{(2)}(x_1) \bigcap \Big(\bigcap_{j \in \mathbf{I}} \mathcal{G}^j_{(2)}(x_1)\Big) = \emptyset. \tag{4.39}$$

Indeed, if

$$\mathcal{X}_{(2)}(x_1, u_1) \bigcap \mathcal{G}^0_{(2)}(x_1) \bigcap \Big(\bigcap_{j \in \mathbb{I}(\bar{x}, x_1)} \mathcal{G}^j_{(2)}(x_1)\Big) = \emptyset,$$

then (4.39) holds. Otherwise, for any

$$x_2 \in \mathcal{X}_{(2)}(x_1,u_1)\bigcap \mathcal{G}^0_{(2)}(x_1)\bigcap\Big(\bigcap_{j\in\mathbb{I}(\bar{x},x_1)} \mathcal{G}^j_{(2)}(x_1)\Big),$$

from (4.14) and (4.25), we see that $x_2\notin \mathcal{H}_{(2)}(x_1)=\mathcal{G}^{n+1}_{(2)}(x_1)$. This also yields (4.39).

It follows from Lemma 2.10 that there exist $x^*, x_j^*\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))^*$ for all $j\in\mathbf{I}$, which do not vanish simultaneously, such that for $\kappa^*=-\Big(x^*+\sum_{j\in\mathbf{I}}x_j^*\Big)$,

$$\inf_{z\in\mathcal{X}_{(2)}(x_1,u_1)}\kappa^*(z)+\inf_{z\in\mathcal{G}^0_{(2)}(x_1)}x^*(z)+\sum_{j\in\mathbf{I}}\inf_{z\in\mathcal{G}^j_{(2)}(x_1)}x_j^*(z)\geq 0. \tag{4.40}$$

If $g^j_x(\bar{x}(T))=0$ for some $j\in\mathbf{I}$, then $\mathcal{G}^j_{(2)}(x_1)=L^2_{\mathbb{F}}(\Omega;C([0,T];H))$. This, together with (4.40), yields $x_j^*=0$.

For each $j\in\mathbf{I}$ with $g^j_x(\bar{x}(T))\neq 0$, put

$$R_j\overset{\triangle}{=}\big\{z_T\in L^2_{\mathcal{F}_T}(\Omega;H)\,\big|\,\mathbb{E}\left\langle g^j_x(\bar{x}(T)),z_T\right\rangle_H\leq 0\ \big\}. \tag{4.41}$$

Then R_j is a closed convex cone and $(R_j)^-=\mathbb{R}_+g^j_x(\bar{x}(T))$.

Let Γ be given by (3.15). It is easy to show that

$$\Gamma^{-1}(R_j)+\mathcal{G}^j_{(2)}(x_1)\subset\mathcal{G}^j_{(2)}(x_1)\ \text{ for every }\ j\in\mathbf{I}$$

and that $\Gamma^{-1}(R_j)$ is a cone. Hence, by (4.40), $-x_j^*\in\big(\Gamma^{-1}(R_j)\big)^-$. Noting that Γ is surjective, by the well known result of convex analysis, $\big(\Gamma^{-1}(R_j)\big)^-=\Gamma^*(R_j^-)$ (see for instance [1, Corollary 22, p. 144] applied to the closed convex cone R_j and the set-valued map Γ^{-1} whose graph is a closed subspace of $L^2_{\mathcal{F}_T}(\Omega;H)\times L^2_{\mathbb{F}}(\Omega;C([0,T];H))$). Therefore,

$$-x_j^*=\Gamma^*(\lambda_jg^j_x(\bar{x}(T)))\text{ for some }\lambda_j\geq 0.$$

If $x_j^*=0$, then we put $\lambda_j=0$. By normalizing, we may assume that $\lambda_0\in\{0,1\}$.

Since the map Γ is surjective, we have that

$$\sup_{z\in\mathcal{G}^j_{(2)}(x_1)}\left(-x_j^*\right)(z)=\sup_{z\in\mathcal{G}^j_{(2)}(x_1)}\mathbb{E}\big\langle\lambda_jg^j_x(\bar{x}(T)),\Gamma(z)\big\rangle_H=\sup_{z_T\in\mathcal{E}^{(2,j)}_T(x_1)}\mathbb{E}\big\langle\lambda_jg^j_x(\bar{x}(T)),z(T)\big\rangle_H.$$

By the definition of $\mathcal{G}^j_{(2)}(x_1)$, for any $j\in\mathbf{I}$ with $g^j_x(\bar{x}(T))\neq 0$,

$$\sup_{z_T\in\mathcal{E}^{(2,j)}_T(x_1)}\mathbb{E}\langle\lambda_jg^j_x(\bar{x}(T)),z_T\rangle_H=-\frac{\lambda_j}{2}\mathbb{E}\left\langle g^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\right\rangle_H.$$

From (4.40) (by setting $d\psi=-x^*$), we deduce that

$$\sup_{x_2\in\mathcal{X}_{(2)}(x_1,v)}(-\kappa^*)(x_2)+\sup_{\alpha\in\mathcal{G}^0_{(2)}(x_1)}\mathbb{E}\int_0^T\langle\alpha(t),d\psi(t)\rangle_H-\frac{1}{2}\mathbb{E}\sum_{j\in\mathbf{I}}\left\langle\lambda_jg^j_{xx}(\bar{x}(T))x_1(T),x_1(T)\right\rangle_H\leq 0. \tag{4.42}$$

Recalling Remark 4.5 for the inclusions $\mathcal{G}^0_{(1)}+\mathcal{G}^0_{(2)}(x_1)\subset\mathcal{G}^0_{(2)}(x_1)$ and $\mathcal{X}_{(1)}+\mathcal{X}_{(2)}(x_1,u_1)\subset\mathcal{X}_{(2)}(x_1,u_1)$, we get from (4.42) that $d\psi\in(\mathcal{G}^0_{(1)})^-$ and $-\kappa^*\in(\mathcal{X}_{(1)})^-$.

Put $y(T)=-\sum_{j\in\mathbf{I}}\lambda_jg^j_x(\bar{x}(T))$ and let $(y(\cdot),Y(\cdot))$ be the solution to (3.17) with $\mathcal{I}(\bar{x})$ replaced by $\mathbb{I}(\bar{x},x_1)$. Let $P(T)=-\lambda_0h_{xx}(\bar{x}(T))$ and $(P(\cdot),Q^{(\cdot)},\widehat{Q}^{(\cdot)})$ be the relaxed solution of (4.1). By (4.42), for every $x_2(\cdot)\in\mathcal{X}_{(2)}(\bar{x},x_1)$,

$$-\mathbb{E}\int_0^T \langle x_2(t), d\psi(t)\rangle_H - \sum_{j\in\mathbf{I}} \mathbb{E}\langle \lambda_j g_x^j(\bar{x}(T)), x_2(T)\rangle_H$$
$$+\frac{1}{2}\mathbb{E}\langle P(T)x_1(T), x_1(T)\rangle_H + \sup_{\alpha\in\mathcal{G}^0_{(2)}(x_1)} \int_0^T \langle \alpha(t), d\psi(t)\rangle_H \le 0.$$

From the above inequality, using (4.26) and (4.28), we complete the proof. □

A. Proof of Lemma 3.1

We first recall the following result. Its proof can be found in [7, Chapter 7].

Lemma A.1. *Assume that **(AS1)** holds. Then, for any $\nu_0 \in H$, $p \ge 1$ and $u(\cdot) \in L^p_{\mathbb{F}}(\Omega; L^2(0,T; H_1))$, the equation (1.1) admits a unique solution $x(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega; H))$, and for any $t \in [0,T]$,*

$$\sup_{s\in[0,t]} \mathbb{E}\Big(|x(s)|_H^p\Big) \le C\mathbb{E}\Big[|\nu_0|_H^p + \Big(\int_0^t |a(s,0,u(s))|_H ds\Big)^p + \Big(\int_0^t |b(s,0,u(s))|^2_{\mathcal{L}_2} ds\Big)^{\frac{p}{2}}\Big]. \tag{A.1}$$

Moreover, if $\tilde{x}$ is the solution of (1.1) corresponding to $(\tilde{\nu}_0, \tilde{u}) \in H \times L^p_{\mathbb{F}}(\Omega; L^2(0,T;H_1))$, then, for any $t \in [0,T]$,

$$\sup_{s\in[0,t]} \mathbb{E}\Big(|x(s)-\tilde{x}(s)|_H^p\Big) \le C\mathbb{E}\ \Big[|\nu_0-\tilde{\nu}_0|_H^p + \Big(\int_0^t |u(s)-\tilde{u}(s)|^2_{H_1} ds\Big)^{\frac{p}{2}}\Big]. \tag{A.2}$$

Proof of Lemma 3.1. From (3.1) and Lemma A.1 we deduce that

$$\begin{aligned}\mathbb{E}\Big(|x_1(t)|_H^p\Big) \le& C\mathbb{E}\Big[|\nu_1|_H^p + \Big(\int_0^t |a_2[s]u_1(s)|_H ds\Big)^p + \Big(\int_0^t |b_2[s]u_1(s)|^2_{\mathcal{L}_2} ds\Big)^{\frac{p}{2}}\Big]\\ \le& C\mathbb{E}\Big[|\nu_1|^p + \Big(\int_0^t |u_1(t)|^2_{H_1} dt\Big)^{\frac{p}{2}}\Big].\end{aligned}$$

This implies that

$$\sup_{t\in[0,T]} \mathbb{E}\Big(|x_1(t)|_H^p\Big) \le C\mathbb{E}\Big[|\nu_1|^p + \Big(\int_0^T |u_1(t)|^2_{H_1} dt\Big)^{\frac{p}{2}}\Big],$$

which yields (3.3).

Since
$$\lim_{\varepsilon\to 0^+} \nu_1^\varepsilon = \nu_1 \text{ in } H, \quad \lim_{\varepsilon\to 0^+} u_1^\varepsilon(\cdot) = u_1(\cdot) \text{ in } L^p_{\mathbb{F}}(\Omega; L^2(0,T;H_1)),$$

it follows from (A.2) that

$$\sup_{t\in[0,T]} \mathbb{E}\Big(|\delta x^\varepsilon(t)|_H^p\Big) \le C\mathbb{E}\Big(\varepsilon^p|\nu_0^\varepsilon|_H^p + \Big(\int_0^T |\varepsilon u_1^\varepsilon(t)|^2_{H_1} ds\Big)^{\frac{p}{2}}\Big) = O(\varepsilon^p).$$

This implies (3.4).

Let
$$\begin{cases}\tilde{a}_1^\varepsilon(t) \triangleq \displaystyle\int_0^1 a_x(t, \bar{x}(t)+\theta\delta x^\varepsilon(t), \bar{u}(t)+\theta\varepsilon u_1^\varepsilon(t))d\theta,\\ \tilde{a}_2^\varepsilon(t) \triangleq \displaystyle\int_0^1 a_u(t, \bar{x}(t)+\theta\delta x^\varepsilon(t), \bar{u}(t)+\theta\varepsilon u_1^\varepsilon(t))d\theta,\\ \tilde{b}_1^\varepsilon(t) \triangleq \displaystyle\int_0^1 b_x(t, \bar{x}(t)+\theta\delta x^\varepsilon(t), \bar{u}(t)+\theta\varepsilon u_1^\varepsilon(t))d\theta,\\ \tilde{b}_2^\varepsilon(t) \triangleq \displaystyle\int_0^1 b_u(t, \bar{x}(t)+\theta\delta x^\varepsilon(t), \bar{u}(t)+\theta\varepsilon u_1^\varepsilon(t))d\theta.\end{cases}$$

Then, $\delta x^\varepsilon(\cdot)$ is the solution of the following SEE:

$$\begin{cases} d\delta x^\varepsilon(t)=\big(A\delta x^\varepsilon(t)+\tilde a_1^\varepsilon(t)\delta x^\varepsilon(t)+\varepsilon\tilde a_2^\varepsilon(t)u_1^\varepsilon(t)\big)dt+\big(\tilde b_1^\varepsilon(t)\delta x^\varepsilon(t)+\varepsilon\tilde b_2^\varepsilon(t)u_1^\varepsilon(t)\big)dW(t) \quad \text{in } (0,T],\\ \delta x^\varepsilon(0)=\varepsilon\nu_1^\varepsilon, \end{cases}$$

and $r_1^\varepsilon(\cdot)$ solves

$$\begin{cases} dr_1^\varepsilon(t)=\Big[Ar_1^\varepsilon(t)+\tilde a_1^\varepsilon(t)r_1^\varepsilon(t)+\big(\tilde a_1^\varepsilon(t)-a_1[t]\big)x_1(t)+\tilde a_2^\varepsilon(t)\big(u_1^\varepsilon(t)-u_1(t)\big)\\ \qquad\qquad +\big(\tilde a_2^\varepsilon(t)-a_2[t]\big)u_1(t)\Big]dt+\Big[\tilde b_1^\varepsilon(t)r_1^\varepsilon(t)+\big(\tilde b_1^\varepsilon(t)-b_1[t]\big)x_1(t)\\ \qquad\qquad +\tilde b_2^\varepsilon(t)\big(u_1^\varepsilon(t)-u_1(t)\big)+\big(\tilde b_2^\varepsilon(t)-b_2[t]\big)u_1(t)\Big]dW(t) \qquad \text{in } (0,T],\\ r_1^\varepsilon(0)=\nu_1^\varepsilon-\nu_1. \end{cases} \tag{A.3}$$

For any sequence $\{\varepsilon_j\}_{j=1}^\infty$ of positive numbers satisfying $\lim\limits_{j\to\infty}\varepsilon_j=0$, we can find a subsequence $\{j_k\}_{k=1}^\infty\subset\mathbb{N}$ such that

$$\begin{cases} \lim\limits_{k\to\infty}\sup\limits_{t\in[0,T]}|\delta x^{\varepsilon_{j_k}}(t)|_H\to 0, \quad \mathbb{P}\text{-a.s.},\\ \lim\limits_{k\to\infty}\varepsilon_{j_k}|u_1^{\varepsilon_{j_k}}(t)|_{H_1}=0, \qquad \mathbb{P}\text{-a.s. for a.e. } t\in[0,T]. \end{cases}$$

Hence,

$$\big|(\tilde a_1^{\varepsilon_j}(\cdot)-a_1(\cdot))x_1(\cdot)\big|_H\to 0 \ \text{ in measure, as } j\to\infty.$$

From **(AS2)**, we see that

$$\lim_{k\to\infty}\big|(\tilde a_1^{\varepsilon_{j_k}}(t)-a_1[t])x_1(t)\big|_H=0, \qquad \mathbb{P}\text{-a.s. for a.e. } t\in[0,T].$$

Then, it follows from Lebesgue's dominated convergence theorem that

$$\lim_{j\to\infty}\mathbb{E}\int_0^T\big|\big(\tilde a_1^{\varepsilon_j}(t)-a_1[t]\big)x_1(t)\big|_H^p dt=0. \tag{A.4}$$

A similar argument implies that

$$\begin{aligned} \lim_{j\to\infty}\mathbb{E}\Big[\Big(\int_0^T\big|\big(\tilde a_2^{\varepsilon_j}(t)-a_2[t]\big)u_1(t)\big|_H^2dt\Big)^{\frac p2}+\Big(\int_0^T\big|\big(\tilde b_1^{\varepsilon_j}(t)-b_1[t]\big)x_1(t)\big|_{\mathcal{L}_2}^2dt\Big)^{\frac p2}\\ +\Big(\int_0^T\big|\big(\tilde b_2^{\varepsilon_j}(t)-b_2[t]\big)u_1(t)\big|_{\mathcal{L}_2}^2dt\Big)^{\frac p2}\Big]=0. \end{aligned} \tag{A.5}$$

On the other hand,

$$\begin{aligned} &\lim_{j\to\infty}\mathbb{E}\Big[\Big(\int_0^T|\tilde a_2^{\varepsilon_j}(t)\big(u_1^{\varepsilon_j}(t)-u_1(t)\big)|_{H_1}^2dt\Big)^{\frac p2}+\Big(\int_0^T|\tilde b_2^{\varepsilon_j}(t)\big(u_1^{\varepsilon_j}(t)-u_1(t)\big)|_{\mathcal{L}_2}^2dt\Big)^{\frac p2}\Big]\\ &\le C\lim_{j\to\infty}\mathbb{E}\Big(\int_0^T|u_1^{\varepsilon_j}(t)-u_1(t)|_{H_1}^2dt\Big)^{\frac p2}\to 0. \end{aligned}$$

Therefore, by Lemma A.1 and (A.3)–(A.5), we obtain that

$$\begin{aligned} &\lim_{j\to\infty}\sup_{t\in[0,T]}\mathbb{E}\Big(|r_1^{\varepsilon_j}(t)|_H^p\Big)\\ &\le C\lim_{j\to\infty}\mathbb{E}\Big[|\nu_1^{\varepsilon_{j_k}}-\nu_1|_H^p+\Big(\int_0^T\big|\big(\tilde a_1^{\varepsilon_j}(t)-a_1[t]\big)x_1(t)+\tilde a_2^{\varepsilon_j}(t)\big(u_1^{\varepsilon_j}(t)-u_1(t)\big)+\big(\tilde a_2^{\varepsilon_j}(t)-a_2[t]\big)u_1(t)\big|_Hdt\Big)^p\\ &\qquad +\Big(\int_0^T\big|\big(\tilde b_1^{\varepsilon_j}(t)-b_1[t]\big)x_1(t)+\tilde b_2^{\varepsilon_j}(t)\big(u_1^{\varepsilon_{j_k}}(t)-u_1(t)\big)+\big(\tilde b_2^{\varepsilon_j}(t)-b_2[t]\big)u_1(t)\big|_{\mathcal{L}_2}^2dt\Big)^{\frac p2}\Big]=0. \end{aligned}$$

Since the sequence $\{\varepsilon_j\}_{j=1}^\infty$ is arbitrary, the proof is complete. $\square$

B. Proof of Lemma 4.2

Proof. By Lemma 3.1 (applied with $p=4$), we obtain

$$\sup_{t\in[0,T]} \mathbb{E}\Big(|x_1(t)|_H^4\Big) \leq C\mathbb{E}\Big[|\nu_1|_H^4 + \Big(\int_0^T |u_1(t)|_{H_1}^2 dt\Big)^2\Big]. \tag{B.1}$$

By (4.5), (B.1) and Hölder's inequality, we have that

$$\begin{aligned}
&\sup_{t\in[0,T]} \mathbb{E}\Big(|x_2(t)|_H^2\Big)\\
&\leq C\mathbb{E}\Big[|\nu_2|_H^2 + \Big(\int_0^T \big|2a_2[t]u_2(t) + a_{11}[t]\big(x_1(t),x_1(t)\big) + 2a_{12}[t]\big(x_1(t),u_1(t)\big) + a_{22}[t]\big(u_1(t),u_1(t)\big)\big|_H dt\Big)^2\\
&\qquad + \int_0^T |2b_2[t]u_2(t) + b_{11}[t]\big(x_1(t),x_1(t)\big) + 2b_{12}[t]\big(x_1(t),u_1(t)\big) + b_{22}[t]\big(u_1(t),u_1(t)\big)|_{\mathcal{L}_2}^2 dt\Big]\\
&\leq C\mathbb{E}\Big(|\nu_2|_H^2 + \int_0^T |u_2(t)|_{H_1}^2 dt + \int_0^T |u_1(t)|_{H_1}^4 dt\Big) + \sup_{t\in[0,T]} \Big(\mathbb{E}|x_1(t)|_H^4 + \mathbb{E}|x_1(t)|_H^2 \mathbb{E}\int_0^T |u_1(t)|_{H_1}^2 dt\Big)\\
&\leq C\mathbb{E}\Big(|\nu_2|_H^2 + |\nu_1|_H^4 + \int_0^T |u_2(t)|_{H_1}^2 dt + \int_0^T |u_1(t)|_{H_1}^4 dt\Big).
\end{aligned}$$

Let

$$\begin{cases}
\tilde a_{11}^\varepsilon(t) \stackrel{\triangle}{=} \int_0^1 (1-\theta)a_{xx}(t,\bar x(t)+\theta\delta x^\varepsilon(t),\bar u(t)+\theta\delta u^\varepsilon(t))d\theta,\\
\tilde a_{12}^\varepsilon(t) \stackrel{\triangle}{=} \int_0^1 (1-\theta)a_{xu}(t,\bar x(t)+\theta\delta x^\varepsilon(t),\bar u(t)+\theta\delta u^\varepsilon(t))d\theta,\\
\tilde a_{22}^\varepsilon(t) \stackrel{\triangle}{=} \int_0^1 (1-\theta)a_{uu}(t,\bar x(t)+\theta\delta x^\varepsilon(t),\bar u(t)+\theta\delta u^\varepsilon(t))d\theta,\\
\tilde b_{11}^\varepsilon(t) \stackrel{\triangle}{=} \int_0^1 (1-\theta)b_{xx}(t,\bar x(t)+\theta\delta x^\varepsilon(t),\bar u(t)+\theta\delta u^\varepsilon(t))d\theta,\\
\tilde b_{12}^\varepsilon(t) \stackrel{\triangle}{=} \int_0^1 (1-\theta)b_{xu}(t,\bar x(t)+\theta\delta x^\varepsilon(t),\bar u(t)+\theta\delta u^\varepsilon(t))d\theta,\\
\tilde b_{22}^\varepsilon(t) \stackrel{\triangle}{=} \int_0^1 (1-\theta)b_{uu}(t,\bar x(t)+\theta\delta x^\varepsilon(t),\bar u(t)+\theta\delta u^\varepsilon(t))d\theta.
\end{cases}$$

Then, δx^ε solves

$$\begin{cases}
d\delta x^\varepsilon(t) = \Big[A\delta x^\varepsilon(t) + a_1[t]\delta x^\varepsilon(t) + a_2[t]\delta u^\varepsilon(t) + \tilde a_{11}^\varepsilon(t)\big(\delta x^\varepsilon(t),\delta x^\varepsilon(t)\big)\\
\qquad\qquad +2\delta\tilde a_{12}^\varepsilon(t)\big(x^\varepsilon(t),\delta u^\varepsilon(t)\big) + \tilde a_{22}^\varepsilon(t)\big(\delta u^\varepsilon(t),\delta u^\varepsilon(t)\big)\Big]dt\\
\qquad\qquad +\Big[b_1[t]\delta x^\varepsilon(t) + b_2[t]\delta u^\varepsilon(t) + \tilde b_{11}^\varepsilon(t)\big(\delta x^\varepsilon(t),\delta x^\varepsilon(t)\big)\\
\qquad\qquad +2\tilde b_{12}^\varepsilon(t)\big(\delta x^\varepsilon(t),\delta u^\varepsilon(t)\big) + \tilde b_{22}^\varepsilon(t)\big(\delta u^\varepsilon(t),\delta u^\varepsilon(t)\big)\Big]dW(t) \qquad \text{in } (0,T],\\
\delta x^\varepsilon(0) = \varepsilon\nu_1 + \varepsilon^2\nu_2^\varepsilon.
\end{cases} \tag{B.2}$$

Consequently, r_2^ε solves

$$\begin{cases} dr_2^\varepsilon(t) = \Big\{ Ar_2^\varepsilon(t) + a_1[t]r_2^\varepsilon(t) + a_2[t]\big(u_2^\varepsilon(t) - u_2(t)\big) + \Big[\tilde{a}_{11}^\varepsilon(t)\Big(\dfrac{\delta x^\varepsilon(t)}{\varepsilon}, \dfrac{\delta x^\varepsilon(t)}{\varepsilon}\Big) \\ \qquad -\dfrac{1}{2}a_{11}[t]\big(x_1(t), x_1(t)\big)\Big] + \Big[2\tilde{a}_{12}^\varepsilon(t)\big(\dfrac{\delta x^\varepsilon(t)}{\varepsilon}, \dfrac{\delta u^\varepsilon(t)}{\varepsilon}\big) - a_{12}[t]\big(x_1(t), u_1(t)\big)\Big] \\ \qquad +\Big[\tilde{a}_{22}^\varepsilon(t)\Big(\dfrac{\delta u^\varepsilon(t)}{\varepsilon}, \dfrac{\delta u^\varepsilon(t)}{\varepsilon}\Big) - \dfrac{1}{2}a_{22}[t]\big(u_1(t), u_1(t)\big)\Big]\Big\}dt \\ \qquad +\Big\{ b_1[t]r_2^\varepsilon(t) + b_2[t]\big(h_\varepsilon(t) - h(t)\big) + \big[\tilde{b}_{11}^\varepsilon(t)\big(\dfrac{\delta x^\varepsilon(t)}{\varepsilon}, \dfrac{\delta x^\varepsilon(t)}{\varepsilon}\big) - \dfrac{1}{2}b_{11}[t]\big(x_1(t), x_1(t)\big)\big] \\ \qquad +\big[2\tilde{b}_{12}^\varepsilon(t)\big(\dfrac{\delta x^\varepsilon(t)}{\varepsilon}, \dfrac{\delta u^\varepsilon(t)}{\varepsilon}\big) - b_{12}[t]\big(x_1(t), u_1(t)\big)\big] \\ \qquad +\big[\tilde{b}_{22}^\varepsilon(t)\big(\dfrac{\delta u^\varepsilon(t)}{\varepsilon}, \dfrac{\delta u^\varepsilon(t)}{\varepsilon}\big) - \dfrac{1}{2}b_{22}[t]\big(u_1(t), u_1(t)\big)\big]\Big\}dW(t) \qquad \text{in } (0,T], \\ r_2^\varepsilon(0) = \nu_2^\varepsilon - \nu_2. \end{cases} \tag{B.3}$$

Since $u_2^\varepsilon(\cdot)$ converges to $u_2(\cdot)$ in $L^4_{\mathbb{F}}(0,T;H_1)$, we have

$$\lim_{\varepsilon\to 0^+} \mathbb{E}\Big(\int_0^T \Big|a_2[t]\big(u_2^\varepsilon(t) - u_2(t)\big)\Big|_H dt\Big)^2 + \lim_{\varepsilon\to 0^+} \mathbb{E}\Big(\int_0^T \Big|b_2[t]\big(u_2^\varepsilon(t) - u_2(t)\big)\Big|^2_{\mathcal{L}_2} dt\Big) = 0. \tag{B.4}$$

By Hölder's inequality,

$$\begin{aligned} &\mathbb{E}\Big(\int_0^T \Big|\tilde{a}_{11}^\varepsilon(t)\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta x^\varepsilon(t)}{\varepsilon}\Big) - \frac{1}{2}a_{11}[t]\big(x_1(t), x_1(t)\big)\Big|_H dt\Big)^2 \\ &\leq C\mathbb{E}\Big(\int_0^T \Big|\tilde{a}_{11}^\varepsilon(t)\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta x^\varepsilon(t)}{\varepsilon}\Big) - \frac{1}{2}a_{11}[t]\big(x_1(t), x_1(t)\big)\Big|^2_H dt\Big) \\ &\leq C\mathbb{E}\Big[\int_0^T \Big|\big(\tilde{a}_{11}^\varepsilon(t) - \frac{1}{2}a_{11}[t]\big)\Big(\frac{\delta x^\varepsilon(t)}{\varepsilon}, \frac{\delta x^\varepsilon(t)}{\varepsilon}\Big)\Big|^2_H dt\Big] \\ &\quad + C\mathbb{E}\Big[\sup_{t\in[0,T]} \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon} - x_1(t)\Big|^2_H \Big(\sup_{t\in[0,T]} \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^2_H + \sup_{t\in[0,T]} |x_1(t)|^2_H\Big)\Big] \\ &\leq C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]} \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^4_H\Big)\Big]^{1/2} \Big[\mathbb{E}\Big(\int_0^T \Big|\tilde{a}_{11}^\varepsilon(t) - \frac{1}{2}a_{11}[t]\Big|^4_{\mathcal{L}(H\times H;H)} \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^4_H dt\Big)\Big]^{1/2} \\ &\quad + C\Big[\mathbb{E}\Big(\sup_{t\in[0,T]} \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon} - x_1(t)\Big|^4_H\Big)\Big]^{\frac{1}{2}} \Big[\mathbb{E}\Big(\sup_{t\in[0,T]} \Big|\frac{\delta x^\varepsilon(t)}{\varepsilon}\Big|^4_H + \sup_{t\in[0,T]} |x_1(t)|^4_H\Big)\Big]^{\frac{1}{2}}. \end{aligned} \tag{B.5}$$

Since

$$\lim_{\varepsilon\to 0^+} \nu_2^\varepsilon = \nu_2 \text{ in } H, \qquad \lim_{\varepsilon\to 0^+} u_2^\varepsilon(\cdot) = u_2(\cdot) \text{ in } L^4_{\mathbb{F}}(0,T;H_1),$$

by Lemma A.1,

$$\sup_{t\in[0,T]} \mathbb{E}\Big(|\delta x^\varepsilon(t)|^4_H\Big) \leq C\mathbb{E}\Big[|\varepsilon\nu_1 + \varepsilon^2\nu_2^\varepsilon|^4_H + \Big(\int_0^T |\varepsilon u_1(t) + \varepsilon^2 u_2^\varepsilon(t)|^2_H dt\Big)^2\Big] = O(\varepsilon^4).$$

As the proof of (3.5) in Lemma 3.1, we obtain that

$$\lim_{\varepsilon\to 0^+} \sup_{t\in[0,T]} \mathbb{E}\Big|\frac{\delta x^\varepsilon(t)}{\varepsilon} - x_1(t)\Big|^4_H = 0.$$

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

$$a_{xx}(\cdot,\bar{x}(\cdot)+\theta\delta x^{\varepsilon_j}(\cdot),\bar{u}(\cdot)+\theta\delta u^{\varepsilon_j}(\cdot))-a_{11}[\cdot]\to 0,\quad \text{in measure, as } j\to\infty. \tag{B.6}$$

Since

$$\tilde{a}_{11}^{\varepsilon_j}(t)-\frac{1}{2}a_{11}[t]=\int_0^1(1-\theta)\big(a_{xx}(t,\bar{x}(t)+\theta\delta x^{\varepsilon_j}(t),\bar{u}(t)+\theta\delta u^{\varepsilon_j}(t))-a_{11}[t]\big)d\theta,$$

it follows from (B.5), (B.6) and the Lebesgue dominated convergence theorem that

$$\lim_{j\to\infty}\mathbb{E}\Big(\int_0^T\Big|\tilde{a}_{11}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}a_{11}[t]\big(x_1(t),x_1(t)\big)\Big|_H dt\Big)^2=0. \tag{B.7}$$

Since,

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^T\Big|2\tilde{a}_{12}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-a_{12}[t]\big(x_1(t),u_1(t)\big)\Big|_H dt\Big)^2\\
&\le C\mathbb{E}\Big(\int_0^T\Big|2\tilde{a}_{12}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-a_{12}[t]\big(x_1(t),u_1(t)\big)\Big|_H^2 dt\Big)\\
&\le C\sup_{t\in[0,T]}\Big(\mathbb{E}\Big|\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big|_H^4\Big)^{\frac{1}{2}}\Big(\mathbb{E}\int_0^T\Big|\tilde{a}_{12}^{\varepsilon_j}(t)-\frac{1}{2}a_{12}[t]\Big|_{\mathcal{L}(H\times H_1;H_1)}^4\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|_{H_1}^4 dt\Big)^{\frac{1}{2}}\\
&\quad+C\sup_{t\in[0,T]}\Big(\mathbb{E}\Big|\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}-x_1(t)\Big|_H^4\Big)^{\frac{1}{2}}\Big(\mathbb{E}\int_0^T\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|_{H_1}^4 dt\Big)^{\frac{1}{2}}\\
&\quad+C\sup_{t\in[0,T]}\Big(\mathbb{E}|x_1(t)|_H^4\Big)^{\frac{1}{2}}\Big(\mathbb{E}\int_0^T\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}-u_1(t)\Big|_{H_1}^4 dt\Big)^{\frac{1}{2}}.
\end{aligned}$$

Similar to the proof of (B.7), we have that

$$\lim_{j\to\infty}\mathbb{E}\Big(\int_0^T\Big|2\tilde{a}_{12}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-a_{12}[t]\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},u_1(t)\Big)\Big|_H dt\Big)^2=0. \tag{B.8}$$

Similarly,

$$\begin{aligned}
&\lim_{j\to\infty}\mathbb{E}\Big(\int_0^T\Big|\tilde{a}_{22}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}a_{22}[t]\big(u_1(t),u_1(t)\big)\Big|_H dt\Big)^2\\
&\le C\lim_{j\to\infty}\mathbb{E}\Big(\int_0^T\Big|\Big(\tilde{a}_{22}^{\varepsilon_j}(t)-\frac{1}{2}a_{22}[t]\Big)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)\Big|_H^2 dt\Big)\\
&\quad+C\lim_{j\to\infty}\mathbb{E}\Big[\int_0^T\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}-u_1(t)\Big|_{H_1}^2\Big(\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|_{H_1}^2+|u_1(t)|_{H_1}^2\Big)dt\Big]\\
&\le C\lim_{j\to\infty}\mathbb{E}\int_0^T\Big|\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big|_{H_1}^4\Big|\tilde{a}_{22}^{\varepsilon_j}(t)-\frac{1}{2}a_{22}[t]\Big|_{\mathcal{L}(H_1\times H_1;H_1)}^2 dt\\
&\quad+C\lim_{j\to\infty}\mathbb{E}\int_0^T\big|\varepsilon_j u_2^{\varepsilon_j}(t)\big|_{H_1}^2\Big(\big|u_1(t)+\varepsilon_j u_2^{\varepsilon_j}(t)\big|_{H_1}^2+|u_1(t)|_{H_1}^2\Big)dt=0.
\end{aligned} \tag{B.9}$$

Similar to the above argument, we obtain

$$\lim_{j\to\infty}\mathbb{E}\int_0^T\Big|\tilde{b}_{11}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}b_{11}[t]\big(x_1(t),x_1(t)\big)\Big|_{\mathcal{L}_2}^2 dt=0, \tag{B.10}$$

$$\lim_{j\to\infty}\mathbb{E}\int_0^T\Big|2\tilde{b}_{12}^{\varepsilon_j}(t)\Big(\frac{\delta x^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-b_{12}[t]\big(x_1(t),u_1(t)\big)\Big|_{\mathcal{L}_2}^2 dt=0, \tag{B.11}$$

and

$$\lim_{j\to\infty}\mathbb{E}\int_0^T\Big|\tilde b_{22}^{\varepsilon_j}(t)\Big(\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j},\frac{\delta u^{\varepsilon_j}(t)}{\varepsilon_j}\Big)-\frac{1}{2}b_{22}[t]\big(u_1(t),u_1(t)\big)\Big|_{\mathcal{L}_2}^2dt=0. \tag{B.12}$$

By Lemma A.1, and using (B.3), (B.4) and (B.7)–(B.12),

$$\lim_{j\to\infty}\sup_{t\in[0,T]}\mathbb{E}|r_2^{\varepsilon_j}(t)|_H^2=0.$$

The desired result follows from the fact that the sequence $\{\varepsilon_j\}_{j=1}^{\infty}$ is arbitrary. $\square$

C. Proof of Lemma 3.2

Proof of Lemma 3.2. We borrow some idea from [39]. The proof is divided into three steps.

Step 1. For any $\tilde v(\cdot)\in\mathcal{C}_{\widetilde{\mathcal{U}}_{ad}^{\nu_0}}(\bar u(\cdot))$, we know that $\tilde v(\cdot)\in\mathcal{L}_{\mathbb{F}}^2(0,T;H_1)$. By Lemma 2.4, there exists a $\mathcal{G}$-measurable function $v(\cdot)$ on $[0,T]\times\Omega$ such that $\tilde v(s,\omega)=v(s,\omega)$, $\widetilde{\mu}$-a.e. Therefore,

$$\int_{[0,T]\times\Omega}|\tilde v(s,\omega)-v(s,\omega)|_{H_1}^2d\widetilde{\mu}(s,\omega)=0 \tag{C.1}$$

and

$$|v(\cdot)|^2_{L^2_{\mathbb{F}}(0,T;H_1)}=\int_{[0,T]\times\Omega}|v(s,\omega)|_{H_1}^2d\widetilde{\mu}(s,\omega)=\int_{[0,T]\times\Omega}|\tilde v(s,\omega)|_{H_1}^2d\widetilde{\mu}(s,\omega)<\infty.$$

Since $\tilde v(\cdot)\in\mathcal{C}_{\widetilde{\mathcal{U}}_{ad}^{\nu_0}}(\bar u(\cdot))$, we have

$$\lim_{\hat v\to\bar u,\varepsilon\to0^+}\frac{1}{\varepsilon}\inf_{\tilde u\in\widetilde{\mathcal{U}}_{ad}^{\nu_0}}\Big(\mathbb{E}\int_0^T|\hat v(t)+\varepsilon\tilde v(t)-\tilde u(t)|_{H_1}^2dt\Big)^{\frac{1}{2}}=0.$$

This, together with (C.1), implies that

$$\begin{aligned}&\lim_{\hat v\to\bar u,\varepsilon\to0^+}\frac{1}{\varepsilon}\inf_{\tilde u\in\widetilde{\mathcal{U}}_{ad}^{\nu_0}}\Big(\mathbb{E}\int_0^T|\hat v(t)+\varepsilon v(t)-\tilde u(t)|_{H_1}^2dt\Big)^{\frac{1}{2}}\\ &\le\lim_{\hat v\to\bar u,\varepsilon\to0^+}\frac{1}{\varepsilon}\inf_{\tilde u\in\widetilde{\mathcal{U}}_{ad}^{\nu_0}}\Big(\mathbb{E}\int_0^T|\hat v(t)+\varepsilon\tilde v(t)-\tilde u(t)|_{H_1}^2dt\Big)^{\frac{1}{2}}=0.\end{aligned} \tag{C.2}$$

For any $\tilde u\in\widetilde{\mathcal{U}}_{ad}^{\nu_0}\subset\mathcal{L}_{\mathbb{F}}^2(0,T;H_1)$, by Lemma 2.4, there exists a $\mathcal{G}$-measurable function $u(\cdot)$ on $[0,T]\times\Omega$ such that $\tilde u(s,\omega)=u(s,\omega)$, $\widetilde{\mu}$-a.e. Hence,

$$\int_{[0,T]\times\Omega}|\tilde u(s,\omega)-u(s,\omega)|_{H_1}^2d\widetilde{\mu}(s,\omega)=0. \tag{C.3}$$

Consequently, $u\in\mathcal{U}_{ad}^{\nu_0}$. This, together with (C.2) and (C.3), implies that

$$\begin{aligned}&\lim_{\hat v\to\bar u,\varepsilon\to0^+}\frac{1}{\varepsilon}\inf_{u\in\mathcal{U}_{ad}^{\nu_0}}\Big(\mathbb{E}\int_0^T|\hat v(t)+\varepsilon v(t)-u(t)|_{H_1}^2dt\Big)^{\frac{1}{2}}\\ &\le\lim_{\hat v\to\bar u,\varepsilon\to0^+}\frac{1}{\varepsilon}\inf_{\tilde u\in\widetilde{\mathcal{U}}_{ad}^{\nu_0}}\Big(\mathbb{E}\int_0^T|\hat v(t)+\varepsilon v(t)-\tilde u(t)|_{H_1}^2dt\Big)^{\frac{1}{2}}=0.\end{aligned} \tag{C.4}$$

Therefore, $v(\cdot)\in\mathcal{C}_{\mathcal{U}_{ad}^{\nu_0}}(\bar u(\cdot))$ and

$$
\begin{aligned}
&\int_{[0,T]\times\Omega}\big\langle F(t,\omega),\tilde v(t,\omega)\big\rangle_{H_1} d\widetilde\mu(t,\omega)\\
&=\int_{[0,T]\times\Omega}\big\langle F(t,\omega),v(t,\omega)\big\rangle_{H_1} d\widetilde\mu(t,\omega)=\mathbb{E}\int_0^T\big\langle F(t),v(t)\big\rangle_{H_1}dt\le 0.
\end{aligned}\tag{C.5}
$$

Step 2. In this step, we prove that the set

$$
\mathcal{A}_{\bar u}\triangleq\big\{(t,\omega)\in[0,T]\times\Omega\;\big|\;\big\langle F(t),v\big\rangle_{H_1}\le 0,\quad\forall\, v\in\mathcal{C}_U(\bar u(t))\big\}\in\widetilde{\mathcal{G}}.\tag{C.6}
$$

We achieve this goal by showing that

$$
\mathcal{A}^c_{\bar u}=\big\{(t,\omega)\in[0,T]\times\Omega\;\big|\;\exists\, v\in\mathcal{C}_U(\bar u(t)),\ \big\langle F(t),v\big\rangle_{H_1}>0\big\}\in\widetilde{\mathcal{G}}.\tag{C.7}
$$

For $k\in\mathbb{N}$, let

$$
\mathcal{B}_{\bar u,k}\triangleq\Big\{(t,\omega)\in[0,T]\times\Omega\;\Big|\;\exists\, v\in\mathcal{C}_U(\bar u(t)),\ \big\langle F(t),v\big\rangle_{H_1}\ge\frac{1}{k}\Big\}.
$$

Clearly,

$$
\mathcal{A}^c_{\bar u}=\bigcup_{k=1}^{\infty}\mathcal{B}_{\bar u,k}.\tag{C.8}
$$

By Corollary 2.1 the set-valued map $\mathcal{C}_U(\bar u(\cdot)):[0,T]\times\Omega\rightsquigarrow H_1$ is $\widetilde{\mathcal{G}}$-measurable. It follows from Lemma 2.3 that

$$
\big\{(t,\omega,v)\in[0,T]\times\Omega\times H_1\big|\ v\in\mathcal{C}_U(\bar u(t,\omega))\big\}\in\widetilde{\mathcal{G}}\otimes\mathcal{B}(H_1).
$$

By the assumption on $F(\cdot)$, we have that

$$
\Big\{(t,\omega,v)\in[0,T]\times\Omega\times H_1\Big|\ \langle F(t),v\rangle_{H_1}\ge\frac{1}{k},\ v\in\mathcal{C}_U(\bar u(t,\omega))\Big\}\in\widetilde{\mathcal{G}}\otimes\mathcal{B}(H_1).\tag{C.9}
$$

Define a set-valued map $\Lambda_k(\cdot,\cdot):[0,T]\times\Omega\rightsquigarrow H_1$ as

$$
\Lambda_k(t,\omega)\triangleq\Big\{v\in H_1\Big|\ \langle F(t),v\rangle_{H_1}\ge\frac{1}{k},\ v\in\mathcal{C}_U(\bar u(t,\omega))\Big\},\qquad(t,\omega)\in[0,T]\times\Omega.
$$

It follows from Lemma 2.3 and (C.9) that Λ_k is $\widetilde{\mathcal{G}}$-measurable. Then $\mathcal{B}_{\bar u,k}=\Lambda_k^{-1}(H_1)\in\widetilde{\mathcal{G}}$. This, together with (C.8), implies (C.7). Consequently, we have (C.6).

Step 3. In this step we prove that $\widetilde\mu(\mathcal{A}_{\bar u})=T$.

For $k,\ m=1,2,\cdots$, let

$$
B(0,m)\triangleq\{v\in H_1||v|_{H_1}\le m\}
$$

and

$$
\mathcal{B}_{\bar u,k,m}\triangleq\Big\{(t,\omega)\in[0,T]\times\Omega\,\Big|\,\exists\, v\in\mathcal{C}_U(\bar u(t))\cap B(0,m),\ s.t.\ \ \langle F(t),v\rangle_{H_1}\ge\frac{1}{k}\Big\}.
$$

It is clear that

$$
\mathcal{A}^c_{\bar u}=\bigcup_{k\ge1}\bigcup_{m\ge1}\mathcal{B}_{\bar u,k,m}.
$$

Similar to the proof of $\mathcal{B}_{\bar u,k}\in\widetilde{\mathcal{G}}$, one can show that $\mathcal{B}_{\bar u,k,m}\in\widetilde{\mathcal{G}}$.

Now we only need to prove that $\widetilde\mu(\mathcal{B}_{\bar u,k,m})=0$ for every $k,\ m\ge1$. Let us do this by a contradiction argument.

Suppose that there exist k and m such that $\widetilde\mu(\mathcal{B}_{\bar u,k,m})>0$. Define the set-valued map $\Upsilon^{k,m}:\mathcal{B}_{\bar u,k,m}\rightsquigarrow H_1$ by

$$
\Upsilon^{k,m}(t,\omega)\triangleq\Big\{v\in\mathcal{C}_U(\bar u(t))\cap B(0,m)\;\Big|\;\langle F(t),v\rangle_{H_1}\ge\frac{1}{k}\Big\}.
$$

Obviously, $\Upsilon^{k,m}(t,\omega)$ is closed-valued. Similar to (C.9),

$$\Big\{(t,\omega,v)\in[0,T]\times\Omega\times H_1\Big|\ v\in\mathcal{C}_U(\bar u(t,\omega))\cap B(0,m),\ \langle F(t),v\rangle_{H_1}\geq\frac{1}{k}\Big\}\in\widetilde{\mathcal{G}}\otimes\mathcal{B}(H_1). \tag{C.10}$$

This, together with Lemma 2.3, implies that $\Upsilon^{k,m}$ is $\widetilde{\mathcal{G}}$-measurable. Then by Lemma 2.6 there exists a $\widetilde{\mathcal{G}}$-measurable selection $v^{k,m}(\cdot)$ on $\mathcal{B}_{\bar u,k,m}$, i.e.,

$$v^{k,m}(t,\omega)\in\Upsilon^{k,m}(t,\omega)\subset\big[\mathcal{C}_U(\bar u(t))\cap B(0,m)\big],\quad\forall\,(t,\omega)\in\mathcal{B}_{\bar u,k,m}.$$

By Lemma 2.2,

$$\big\{v(\cdot)\in\mathcal{L}^2_{\mathbb{F}}(0,T;H_1)\ \big|\ v(t)\in\mathcal{C}_U(u(t)),\ \widetilde\mu\text{-a.e.}\big\}\subset\mathcal{C}_{\widetilde{\mathcal{U}}^{\nu_0}_{ad}}(u(\cdot)).$$

Let $\widetilde v^{k,m}(\cdot)\overset{\triangle}{=}v^{k,m}(\cdot)\chi_{\mathcal{B}_{\bar u,k,m}}(\cdot)$. Then

$$\widetilde\mu\Big\{(t,\omega)\in[0,T]\ \Big|\ \big\langle F(t),\widetilde v^{k,m}(t)\big\rangle_{H_1}\geq\frac{1}{k}\Big\}\geq\widetilde\mu(\mathcal{B}_{\bar u,k,m})>0. \tag{C.11}$$

Therefore,

$$\int_{[0,T]}\int_\Omega\big\langle F(t,\omega),\widetilde v^{k,m}(t,\omega)\big\rangle_{H_1}d\widetilde\mu(t,\omega)\geq\frac{1}{k}\widetilde\mu(\mathcal{B}_{\bar u,k,m})>0. \tag{C.12}$$

On the other hand, by Corollary 2.1, one has $v^{k,m}(\cdot)\in\mathcal{T}_{\bar u}\subset\mathcal{C}_{\widetilde{\mathcal{U}}^{\nu_0}_{ad}}(\bar u(\cdot))$. It follows from (C.5) that

$$\int_{[0,T]}\int_\Omega\big\langle F(t,\omega),\widetilde v^{k,m}(t,\omega)\big\rangle_{H_1}d\widetilde\mu(t,\omega)\leq 0,$$

which contradicts to (C.12). Therefore, $\widetilde\mu(\mathcal{B}_{\bar u,k,m})=0$. Consequently, $\widetilde\mu(\mathcal{A}^c_{\bar u})=0$. Since $\mathcal{A}^c_{\bar u}\in\widetilde{\mathcal{G}}$, there exists a $\mathcal{G}$-measurable set $\mathcal{E}_{\bar u}$ satisfying $\mathcal{A}^c_{\bar u}\subset\mathcal{E}_{\bar u}$ and $\widetilde\mu(\mathcal{A}^c_{\bar u})=\mu(\mathcal{E}_{\bar u})=0$. Thus, $\mathcal{E}^c_{\bar u}\subset\mathcal{A}_{\bar u}$ and $[\mathbf{m}\times\mathbb{P}](\mathcal{E}^c_{\bar u})=T$. This completes the proof. $\square$

D. Proof of Lemma 3.3

The case that H is finite dimensional was studied in [17]. The proof for the general case is similar. We give it here for the sake of completeness.

Proof of Lemma 3.3. Obviously, $L^2_{\mathbb{F}}(\Omega;C([0,T];H))$ is a linear subspace of $L^2_{\mathbb{F}}(\Omega;D([0,T];H))$. For a given $\Lambda\in L^2_{\mathbb{F}}(\Omega;C([0,T];H))^*$, by the Hahn-Banach theorem, there is an extension $\widetilde\Lambda\in L^2_{\mathbb{F}}(\Omega;D([0,T];H))^*$ such that

$$|\widetilde\Lambda|_{L^2_{\mathbb{F}}(\Omega;D([0,T];H))^*}=|\Lambda|_{L^2_{\mathbb{F}}(\Omega;C([0,T];H))^*} \tag{D.1}$$

and

$$\widetilde\Lambda(x(\cdot))=\Lambda(x(\cdot)),\qquad\forall\,x(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];H)). \tag{D.2}$$

Recall that $\{e_k\}_{k=1}^\infty$ is an orthonormal basis of H and Γ_k the projective operator from H to $H_k\overset{\triangle}{=}\operatorname{span}\{e_k\}$. Let $\Lambda_k=\Lambda\Gamma_k$ and $\widetilde\Lambda_k=\widetilde\Lambda\Gamma_k$. Clearly,

$$\begin{aligned}\Lambda_k&\in L^2_{\mathbb{F}}(\Omega;C([0,T];H_k))^*=L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}))^*,\\ \widetilde\Lambda_k&\in L^2_{\mathbb{F}}(\Omega;D([0,T];H_k))^*=L^2_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}))^*,\end{aligned}$$

and

$$\widetilde\Lambda(x(\cdot))=\lim_{m\to\infty}\sum_{k=1}^m\widetilde\Lambda_k(x(\cdot)),\quad\forall\,x(\cdot)\in L^2_{\mathbb{F}}(\Omega;D([0,T];H)). \tag{D.3}$$

For each $k\in\mathbb{N}$, from the proof of [9, Theorem 65, p. 254], we deduce that, there exist two $\mathbb{R}$-valued processes $\psi_k^+(\cdot)$ and $\psi_k^-(\cdot)$ of bounded variations such that $\psi_k^+(\cdot)$ is optional and purely discontinuous, $\psi_k^-(\cdot)$ is predictable with $\psi_k^-(0)=0$,

$$
|\widetilde{\Lambda}_k|^2_{L^2_{\mathbb{F}}(\Omega;D([0,T];\mathbb{R}))^*} = \mathbb{E}\Big|\int_{(0,T]} d|\psi_k^-(t)| + \int_{[0,T)} d|\psi_k^+(t)|\Big|^2 \tag{D.4}
$$

and, for any $x(\cdot) \in L^2_{\mathbb{F}}(\Omega; D([0,T];H))$,

$$
\widetilde{\Lambda}_k(x(\cdot)) = \mathbb{E}\Big(\int_{(0,T]} \Gamma_k x_-(t) d\psi_k^-(t) + \int_{[0,T)} \Gamma_k x(t) d\psi_k^+(t)\Big), \tag{D.5}
$$

where $x_-(\cdot)$ is the predictable modification of $x(\cdot)$, which equals $x(\cdot)$ when $x(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];H))$.

Define two H-valued processes $\psi^+(\cdot)$ and $\psi^-(\cdot)$ as follows:

$$
\psi^+(\cdot) = \sum_{k=1}^{\infty} \psi_k^+(\cdot) e_k, \qquad \psi^-(\cdot) = \sum_{k=1}^{\infty} \psi_k^-(\cdot) e_k.
$$

Then

$$
\int_{(0,T]} \langle x_-(t), d\psi^-(t)\rangle_H = \sum_{k=1}^{\infty} \int_{(0,T]} \Gamma_k x_-(t) d\psi_k^-(t)
$$

and

$$
\int_{[0,T)} \langle x(t), d\psi^+(t)\rangle_H = \sum_{k=1}^{\infty} \int_{[0,T)} \Gamma_k x(t) d\psi_k^+(t).
$$

It follows from (D.3) and (D.5) that

$$
\widetilde{\Lambda}(x(\cdot)) = \int_{(0,T]} \langle x_-(t), d\psi^-(t)\rangle_H + \int_{[0,T)} \langle x(t), d\psi^+(t)\rangle_H, \quad \forall\, x(\cdot) \in L^2_{\mathbb{F}}(\Omega; D([0,T];H)). \tag{D.6}
$$

This, together with the arbitrariness of $x(\cdot) \in L^2_{\mathbb{F}}(\Omega; D([0,T];H))$, implies that $\psi^+(\cdot)$ and $\psi^-(\cdot)$ are functions of bounded variation and

$$
|\widetilde{\Lambda}|^2_{L^2_{\mathbb{F}}(\Omega;D([0,T];H))^*} = \mathbb{E}\Big|\int_{(0,T]} d|\psi^-(t)|_H + \int_{[0,T)} d|\psi^+(t)|_H\Big|^2. \tag{D.7}
$$

Put $\psi^* \overset{\triangle}{=} \psi^- + \psi^+$. By (D.2), we have

$$
\Lambda(x(\cdot)) = \mathbb{E}\int_0^T \langle x(t), d\psi^*(t)\rangle_H\,, \qquad \forall\, x(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T];H)).
$$

Letting $\psi = \psi^* - \psi^*(0)$, we obtain (3.6). (3.7) follows from (3.6). $\square$

E. Proof of Lemma 3.5

Before proving Lemma 3.5, we first recall the following Riesz-type Representation Theorem (See [26, Corollary 2.3 and Remark 2.4]).

Lemma E.1. *Fix t_1 and t_2 satisfying $0 \le t_2 < t_1 \le T$. Assume that $\mathcal{Y}$ is a reflexive Banach space. Then, for any $r, s \in [1,\infty)$, it holds that*

$$
(L^r_{\mathbb{F}}(t_2,t_1;L^s(\Omega;\mathcal{Y})))^* = L^{r'}_{\mathbb{F}}(t_2,t_1;L^{s'}(\Omega;\mathcal{Y}^*)),
$$

where

$$
s' = \begin{cases} s/(s-1), & \text{if } s \ne 1, \\ \infty & \text{if } s = 1; \end{cases} \qquad r' = \begin{cases} r/(r-1), & \text{if } r \ne 1, \\ \infty & \text{if } r = 1. \end{cases}
$$

Next, we recall the following result.

Lemma E.2. *[29, Lemma 2.5] Assume that $f_1 \in L^2_{\mathbb{F}}(0,T;H)$ and $f_2 \in L^2_{\mathbb{F}}(0,T;H)$. Then there exists a decreasing sequence $\{\varepsilon_n\}_{n=1}^{\infty}$ of positive numbers such that $\lim_{n\to\infty} \varepsilon_n = 0$, and*

$$\lim_{n\to\infty} \frac{1}{\varepsilon_n} \int_t^{t+\varepsilon_n} \mathbb{E}\langle f_1(t), f_2(\tau)\rangle_H d\tau = \mathbb{E}\langle f_1(t), f_2(t)\rangle_H, \qquad \text{a.e. } t \in [0,T]. \tag{E.1}$$

Proof of Lemma 3.5. It suffices to consider a particular case for (3.17):

$$\begin{cases} dy(t) = -A^* y(t)dt + f(t)dt + d\psi(t) + Y(t)dW(t) & \text{in } [0,T), \\ y(T) = y_T, \end{cases} \tag{E.2}$$

where $y_T \in L^p_{\mathcal{F}_T}(\Omega;H)$ and $f(\cdot) \in L^1_{\mathbb{F}}(0,T;L^2(\Omega;H))$. The general case follows from the well-posedness of (E.2) and the standard fixed point technique.

We divide the proof into several steps. Since the proof is very similar to that of [29, Theorem 3.1], we give below only a sketch.

Step 1. For any $t \in [\tau,T]$, let us define a linear functional $\mathfrak{F}$ (depending on t) on the Banach space $L^1_{\mathbb{F}}(t,T;L^2(\Omega;H)) \times L^2_{\mathbb{F}}(t,T;\mathcal{L}_2) \times L^2_{\mathcal{F}_t}(\Omega;H)$ as follows:

$$\begin{aligned} &\mathfrak{F}\big(f_1(\cdot), f_2(\cdot), \eta\big) = \mathbb{E}\big\langle \phi(T), y_T\big\rangle_H - \mathbb{E}\int_t^T \big\langle \phi(s), f(s)\big\rangle_H ds - \mathbb{E}\int_t^T \langle \phi(s), d\psi(s)\rangle_H, \\ &\qquad\qquad \forall\, \big(f_1(\cdot), f_2(\cdot), \eta\big) \in L^1_{\mathbb{F}}(t,T;L^2(\Omega;H)) \times L^2_{\mathbb{F}}(t,T;\mathcal{L}_2) \times L^2_{\mathcal{F}_t}(\Omega;H), \end{aligned} \tag{E.3}$$

where $\phi(\cdot) \in L^2_{\mathbb{F}}(\Omega;C([t,T];H))$ is the mild solution of the equation (3.18). It is an easy matter to show that $\mathfrak{F}$ is a bounded linear functional. By Lemma E.1, there exists a triple

$$\big(y^t(\cdot), Y^t(\cdot), \xi^t\big) \in L^{\infty}_{\mathbb{F}}(t,T;L^2(\Omega;H)) \times L^2_{\mathbb{F}}(t,T;\mathcal{L}_2) \times L^2_{\mathcal{F}_t}(\Omega;H)$$

such that

$$\begin{aligned} &\mathbb{E}\big\langle \phi(T), y_T\big\rangle_H - \mathbb{E}\int_t^T \big\langle \phi(s), f(s)\big\rangle_H ds - \mathbb{E}\int_t^T \langle \phi(s), d\psi(s)\rangle_H \\ &= \mathbb{E}\int_t^T \big\langle f_1(s), y^t(s)\big\rangle_H ds + \mathbb{E}\int_t^T \big\langle f_2(s), Y^t(s)\big\rangle_{\mathcal{L}_2} ds + \mathbb{E}\big\langle \eta, \xi^t\big\rangle_H. \end{aligned} \tag{E.4}$$

It is clear that $\xi^T = y_T$. Furthermore,

$$\begin{aligned} &|(y^t(\cdot), Y^t(\cdot), \xi^t)|_{L^{\infty}_{\mathbb{F}}(t,T;L^2(\Omega;H)) \times L^2_{\mathbb{F}}(t,T;H) \times L^2_{\mathcal{F}_t}(\Omega;H)} \\ &\leq C\big(|f(\cdot)|_{L^1_{\mathbb{F}}(t,T;L^2(\Omega;H))} + |y_T|_{L^2_{\mathcal{F}_T}(\Omega;H)} + |\psi|_{L^2_{\mathbb{F}}(\Omega;BV(0,T;H))}\big), \qquad \forall\, t \in [\tau,T]. \end{aligned} \tag{E.5}$$

Step 2. Note that $(y^t(\cdot), Y^t(\cdot))$ obtained in Step 1 may depend on t. Now we show the time consistency of $(y^t(\cdot), Y^t(\cdot))$, that is, for any t_1 and t_2 satisfying $0 \leq t_2 \leq t_1 \leq T$, it holds that

$$\big(y^{t_2}(s,\omega), Y^{t_2}(s,\omega)\big) = \big(y^{t_1}(s,\omega), Y^{t_1}(s,\omega)\big), \qquad \text{a.e. } (s,\omega) \in [t_1,T] \times \Omega, \tag{E.6}$$

for a suitable choice of the η, f_1 and f_2 in (3.18). In fact, for any fixed $\varrho_1(\cdot) \in L^1_{\mathbb{F}}(t_1,T;L^2(\Omega;H))$ and $\varrho_2(\cdot) \in L^2_{\mathbb{F}}(t_1,T;\mathcal{L}_2)$, we choose first $t = t_1$, $\eta = 0$, $f_1(\cdot) = \varrho_1(\cdot)$ and $f_2(\cdot) = \varrho_2(\cdot)$ in (3.18). From (E.4), we get that

$$\begin{aligned} &\mathbb{E}\big\langle \phi^{t_1}(T), y_T\big\rangle_H - \mathbb{E}\int_{t_1}^T \big\langle \phi^{t_1}(s), f(s)\big\rangle_H ds - \mathbb{E}\int_{t_1}^T \big\langle \phi^{t_1}(s), d\psi(s)\big\rangle_H \\ &= \mathbb{E}\int_{t_1}^T \big\langle \varrho_1(s), y^{t_1}(s)\big\rangle_H ds + \mathbb{E}\int_{t_1}^T \big\langle \varrho_2(s), Y^{t_1}(s)\big\rangle_{\mathcal{L}_2} ds. \end{aligned} \tag{E.7}$$

Next, we choose $t=t_2$, $\eta=0$, $f_1(\cdot)=\chi_{[t_1,T]}(\cdot)\varrho_1(\cdot)$ and $f_2(\cdot)=\chi_{[t_1,T]}(\cdot)\varrho_2(\cdot)$ in (3.18). It follows from (E.4) that

$$
\begin{aligned}
&\mathbb{E}\big\langle \phi^{t_1}(T), y_T\big\rangle_H - \mathbb{E}\int_{t_1}^T \big\langle \phi^{t_1}(s), f(s)\big\rangle_H ds - \mathbb{E}\int_{t_1}^T \big\langle \phi^{t_1}(s), d\psi(s)\big\rangle_H \\
&= \mathbb{E}\int_{t_1}^T \big\langle \varrho_1(s), y^{t_2}(s)\big\rangle_H ds + \mathbb{E}\int_{t_1}^T \big\langle \varrho_2(s), Y^{t_2}(s)\big\rangle_{\mathcal{L}_2} ds.
\end{aligned}
\tag{E.8}
$$

Combining (E.7) and (E.8), we get

$$
\begin{aligned}
&\mathbb{E}\int_{t_1}^T \big\langle \varrho_1(s), y^{t_1}(s)-y^{t_2}(s)\big\rangle_H ds + \mathbb{E}\int_{t_1}^T \big\langle \varrho_2(s), Y^{t_1}(s)-Y^{t_2}(s)\big\rangle_{\mathcal{L}_2} ds = 0,\\
&\qquad\qquad \forall\ \varrho_1(\cdot)\in L^1_{\mathbb{F}}(t_1,T;L^2(\Omega;H)),\quad \varrho_2(\cdot)\in L^2_{\mathbb{F}}(t_1,T;\mathcal{L}_2).
\end{aligned}
$$

This yields the desired equality (E.6).

Put

$$
y(t,\omega)=y^{\tau}(t,\omega),\qquad Y(t,\omega)=Y^{\tau}(t,\omega),\qquad \forall\ (t,\omega)\in[\tau,T]\times\Omega. \tag{E.9}
$$

From (E.6), we see that

$$
\big(y^t(s,\omega),Y^t(s,\omega)\big)=\big(y(s,\omega),Y(s,\omega)\big),\qquad \text{a.e. } (s,\omega)\in[t,T]\times\Omega. \tag{E.10}
$$

Combining (E.4) and (E.10), we deduce that

$$
\begin{aligned}
&\mathbb{E}\big\langle \phi(T), y_T\big\rangle_H - \mathbb{E}\int_t^T \big\langle \phi(s), f(s)\big\rangle_H ds - \mathbb{E}\int_t^T \big\langle \phi(s), d\psi(s)\big\rangle_H \\
&= \mathbb{E}\big\langle \eta, \xi^t\big\rangle_H + \mathbb{E}\int_t^T \big\langle f_1(s), y(s)\big\rangle_H ds + \mathbb{E}\int_t^T \big\langle f_2(s), Y(s)\big\rangle_{\mathcal{L}_2} ds,\\
&\qquad \forall\ \big(f_1(\cdot),f_2(\cdot),\eta\big)\in L^1_{\mathbb{F}}(t,T;L^2(\Omega;H))\times L^2_{\mathbb{F}}(t,T;\mathcal{L}_2)\times L^2_{\mathcal{F}_t}(\Omega;H).
\end{aligned}
\tag{E.11}
$$

Step 3. We show in this step that ξ^t has a càdlàg modification. The detail is lengthy and very similar to Step 3 in the proof of [29, Theorem 3.1], and hence we omit it here.

First of all, we claim that, for each $t\in[0,T]$,

$$
\mathbb{E}\Big(S^*(T-t)y_T-\int_t^T S^*(s-t)f(s)ds-\int_t^T S^*(s-t)d\psi(s)\ \Big|\ \mathcal{F}_t\Big)=\xi^t,\quad \mathbb{P}\text{-a.s.} \tag{E.12}
$$

To prove this, we note that for any $\eta\in L^2_{\mathcal{F}_t}(\Omega;H)$, $f_1=0$ and $f_2=0$, the corresponding solution of (3.18) is given by $\phi(s)=S(s-t)\eta$ for $s\in[t,T]$. Hence, by (E.11), we obtain that

$$
\mathbb{E}\big\langle S(T-t)\eta, y_T\big\rangle_H - \mathbb{E}\langle \eta,\xi^t\rangle_H = \mathbb{E}\int_t^T \big\langle S(s-t)\eta, f(s)\big\rangle_H ds + \mathbb{E}\int_t^T \big\langle S(s-t)\eta, d\psi(s)\big\rangle_H. \tag{E.13}
$$

Noting that

$$
\mathbb{E}\big\langle S(T-t)\eta, y_T\big\rangle_H = \mathbb{E}\big\langle \eta, S^*(T-t)y_T\big\rangle_H = \mathbb{E}\big\langle \eta, \mathbb{E}(S^*(T-t)y_T\mid\mathcal{F}_t)\big\rangle_H,
$$

$$
\mathbb{E}\int_t^T \big\langle S(s-t)\eta, f(s)\big\rangle_H ds = \mathbb{E}\Big\langle \eta, \int_t^T S^*(s-t)f(s)ds\Big\rangle_H = \mathbb{E}\Big\langle \eta, \mathbb{E}\Big(\int_t^T S^*(s-t)f(s)ds\ \Big|\ \mathcal{F}_t\Big)\Big\rangle_H,
$$

and

$$
\mathbb{E}\int_t^T \big\langle S(s-t)\eta, d\psi(s)\big\rangle_H = \mathbb{E}\Big\langle \eta, \int_t^T S^*(s-t)d\psi(s)\Big\rangle_H = \mathbb{E}\Big\langle \eta, \mathbb{E}\Big(\int_t^T S^*(s-t)d\psi(s)\ \Big|\ \mathcal{F}_t\Big)\Big\rangle_H,
$$

by (E.13), we conclude that for every $\eta\in L^2_{\mathcal{F}_t}(\Omega;H)$,

$$\mathbb{E}\Big\langle \eta, \mathbb{E}\Big(S^*(T-t)y_T - \int_t^T S^*(s-t)f(s)ds - \int_t^T S^*(s-t)d\psi(s) \,\Big|\, \mathcal{F}_t\Big) - \xi^t \Big\rangle_H = 0. \tag{E.14}$$

Clearly, (E.12) follows from (E.14) immediately.

In the rest of this step, we show that the process

$$\Big\{\mathbb{E}\Big(S^*(T-t)y_T - \int_t^T S^*(s-t)f(s)ds \,\Big|\, \mathcal{F}_t\Big)\Big\}_{t\in[0,T]}$$

has a càdlàg modification.

Recall that for any $\lambda \in \rho(A)$, the bounded operator A_λ (resp. A_λ^*) generates a C_0-group $\{S_\lambda(t)\}_{t\in\mathbb{R}}$ (resp. $\{S_\lambda^*(t)\}_{t\in\mathbb{R}}$) on H.

For each $t \in [0,T]$, put

$$\xi_\lambda^t \stackrel{\triangle}{=} \mathbb{E}\Big(S_\lambda^*(T-t)y_T - \int_t^T S_\lambda^*(s-t)f(s)ds - \int_t^T S_\lambda^*(s-t)d\psi(s) \,\Big|\, \mathcal{F}_t\Big) \tag{E.15}$$

and

$$\Phi_\lambda(t) \stackrel{\triangle}{=} S_\lambda^*(t)\xi_\lambda^t - \int_0^t S_\lambda^*(s)f(s)ds - \int_t^T S_\lambda^*(s-t)d\psi(s). \tag{E.16}$$

We claim that $\{\Phi_\lambda(t)\}$ is an H-valued $\mathbb{F}$-martingale. In fact, for any $\tau_1, \tau_2 \in [0,T]$ with $\tau_1 \le \tau_2$, it follows from (E.15) and (E.16) that

$$\begin{aligned}
&\mathbb{E}(\Phi_\lambda(\tau_2) \mid \mathcal{F}_{\tau_1})\\
&= \mathbb{E}\Big(S_\lambda^*(\tau_2)\xi_\lambda^{\tau_2} - \int_0^{\tau_2} S_\lambda^*(s)f(s)ds - \int_0^{\tau_2} S_\lambda^*(s)d\psi(s) \,\Big|\, \mathcal{F}_{\tau_1}\Big)\\
&= \mathbb{E}\Big[\mathbb{E}\Big(S_\lambda^*(T)y_T - \int_{\tau_2}^T S_\lambda^*(s)f(s)ds - \int_{\tau_2}^T S_\lambda^*(s)d\psi(s)\Big|\mathcal{F}_{\tau_2}\Big) - \int_0^{\tau_2} S_\lambda^*(s)f(s)ds - \int_0^{\tau_2} S_\lambda^*(s)d\psi(s)\Big|\mathcal{F}_{\tau_1}\Big]\\
&= \mathbb{E}\Big(S_\lambda^*(T)y_T - \int_0^T S_\lambda^*(s)f(s)ds - \int_0^T S_\lambda^*(s)d\psi(s) \,\Big|\, \mathcal{F}_{\tau_1}\Big)\\
&= S_\lambda^*(\tau_1)\mathbb{E}\Big(S_\lambda^*(T-\tau_1)y_T - \int_{\tau_1}^T S_\lambda^*(s-\tau_1)f(s)ds - \int_{\tau_1}^T S_\lambda^*(s-\tau_1)d\psi(s) \,\Big|\, \mathcal{F}_{\tau_1}\Big)\\
&\quad - \int_0^{\tau_1} S_\lambda^*(s)f(s)ds - \int_0^{\tau_1} S_\lambda^*(s)d\psi(s)\\
&= S_\lambda^*(\tau_1)\xi_\lambda^{\tau_1} - \int_0^{\tau_1} S_\lambda^*(s)f(s)ds - \int_0^{\tau_1} S_\lambda^*(s)d\psi(s)\\
&= X_\lambda(\tau_1), \quad \mathbb{P}\text{-a.s.},
\end{aligned}$$

as desired.

Now, since $\{X_\lambda(t)\}_{0\le t\le T}$ is an H-valued $\mathbb{F}$-martingale, it enjoys a càdlàg modification, and hence so does the following process

$$\{\xi_\lambda^t\}_{0\le t\le T} = \Big\{S_\lambda^*(-t)\Big[X_\lambda(t) + \int_0^t S_\lambda^*(s)f(s)ds + \int_0^t S_\lambda^*(s)d\psi(s)ds\Big]\Big\}_{0\le t\le T}.$$

Here we have used the fact that $\{S_\lambda^*(t)\}_{t\in\mathbb{R}}$ is a C_0-group on H. We still use $\{\xi_\lambda^t\}_{0\le t\le T}$ to stand for its càdlàg modification.

From (E.12) and (E.15), it follows that

$$\begin{aligned}
&\lim_{\lambda\to\infty} |\xi^\cdot - \xi_\lambda^\cdot|_{L^\infty_{\mathbb{F}}(0,T;L^2(\Omega;H))}\\
&= \lim_{\lambda\to\infty} \Big|\mathbb{E}\Big(S^*(T-\cdot))y_T - \int_\cdot^T S^*(s-\cdot))f(s)ds - \int_\cdot^T S^*(s-\cdot))d\psi(s) \,\Big|\, \mathcal{F}_\cdot\Big)
\end{aligned}$$

$$
\begin{aligned}
&-\mathbb{E}\Big(S_\lambda^*(T-\cdot))y_T-\int_\cdot^T S_\lambda^*(s-\cdot)f(s)ds-\int_\cdot^T S_\lambda^*(s-\cdot)d\psi(s)\;\Big|\;\mathcal{F}_\cdot\Big)\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))} \\
&\le \lim_{\lambda\to\infty}\Big|S^*(T-\cdot)y_T-S_\lambda^*(T-\cdot)y_T\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))} \\
&\quad+\lim_{\lambda\to\infty}\Big|\int_\cdot^T S^*(s-\cdot)f(s)ds-\int_\cdot^T S_\lambda^*(s-\cdot)f(s)ds\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))} \\
&\quad+\lim_{\lambda\to\infty}\Big|\int_\cdot^T S^*(s-\cdot)d\psi(s)-\int_\cdot^T S_\lambda^*(s-\cdot)d\psi(s)\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))}.
\end{aligned}
\tag{E.17}
$$

Let us prove the right hand side of (E.17) equals zero. First, we prove

$$
\lim_{\lambda\to\infty}\Big|S^*(T-\cdot)y_T-S_\lambda^*(T-\cdot)y_T\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))}=0. \tag{E.18}
$$

By the property of Yosida approximations, we deduce that for any $\alpha\in H$, it holds that

$$
\lim_{\lambda\to\infty}|S^*(T-\cdot)\alpha-S_\lambda^*(T-\cdot)\alpha|_{L^\infty(0,T;H)}=0
$$

and that

$$
\big|S^*(T-\cdot)y_T-S_\lambda^*(T-\cdot)y_T\big|_H\le C|y_T|_H.
$$

Thus, by Lebesgue's dominated convergence, we obtain (E.18).

Similarly, we can prove that

$$
\lim_{\lambda\to\infty}\Big|\int_\cdot^T S^*(s-\cdot)f(s)ds-\int_\cdot^T S_\lambda^*(s-\cdot)f(s)ds\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))}=0 \tag{E.19}
$$

and

$$
\lim_{\lambda\to\infty}\Big|\int_\cdot^T S^*(s-\cdot)d\psi(s)-\int_\cdot^T S_\lambda^*(s-\cdot)d\psi(s)\Big|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))}=0. \tag{E.20}
$$

By (E.17), (E.18), (E.19) and (E.20), we obtain that $\lim\limits_{m\to\infty}\lim\limits_{\lambda\to\infty}|\xi^\cdot-\xi^\cdot_{\lambda,m}|_{L_{\mathbb{F}}^\infty(0,T;L^2(\Omega;H))}=0$. Recalling that $\xi^\cdot_\lambda\in D_{\mathbb{F}}([0,T];L^2(\Omega;H))$, we deduce that $\xi^\cdot$ enjoys a cádlág modification.

Step 4. In this step, we show that, for a.e. $t\in[0,T]$,

$$
\xi^t=y(t),\quad \mathbb{P}\text{-a.s.} \tag{E.21}
$$

Choosing $t=t_2$, $f_1(\cdot)=0$, $f_2(\cdot)=0$ and $\eta=(t_1-t_2)\gamma$ in (3.18), utilizing (E.11), we obtain that

$$
\begin{aligned}
&\mathbb{E}\big\langle S(T-t_2)(t_1-t_2)\gamma,y_T\big\rangle_H-\mathbb{E}\big\langle (t_1-t_2)\gamma,\xi^{t_2}\big\rangle_H \\
&=\mathbb{E}\int_{t_2}^T\big\langle S(\tau-t_2)(t_1-t_2)\gamma,f(\tau)\big\rangle_H d\tau+\mathbb{E}\int_{t_2}^T\big\langle S(\tau-t_2)(t_1-t_2)\gamma,d\psi(\tau)\big\rangle_H.
\end{aligned}
\tag{E.22}
$$

Choosing $t=t_2$, $f_1(\tau,\omega)=\chi_{[t_2,t_1]}(\tau)\gamma(\omega)$, $f_2(\cdot)=0$ and $\eta=0$ in (3.18), utilizing (E.11) again, we find that

$$
\begin{aligned}
&\mathbb{E}\Big\langle\int_{t_2}^T S(T-s)\chi_{[t_2,t_1]}(s)\gamma ds,y_T\Big\rangle_H \\
&=\mathbb{E}\int_{t_2}^{t_1}\Big\langle\int_{t_2}^\tau S(\tau-s)\gamma ds,f(\tau)\Big\rangle_H d\tau+\mathbb{E}\int_{t_1}^T\Big\langle S(\tau-t_1)\int_{t_2}^{t_1}S(t_1-s)\gamma ds,f(\tau)\Big\rangle_H d\tau \\
&\quad+\mathbb{E}\int_{t_2}^{t_1}\Big\langle\int_{t_2}^\tau S(\tau-s)\gamma ds,d\psi(\tau)\Big\rangle_H+\mathbb{E}\int_{t_1}^T\Big\langle S(\tau-t_1)\int_{t_2}^{t_1}S(t_1-s)\gamma ds,d\psi(\tau)\Big\rangle_H \\
&\quad+\mathbb{E}\int_{t_2}^{t_1}\langle\gamma,y(\tau)\rangle_H d\tau.
\end{aligned}
\tag{E.23}
$$

It follows from (E.22) and (E.23) that

$$
\begin{aligned}
&\mathbb{E}\langle \gamma, \xi^{t_2}\rangle_H\\
&= \frac{1}{t_1-t_2}\int_{t_2}^{t_1} \mathbb{E}\langle \gamma, y(\tau)\rangle_H d\tau + \mathbb{E}\big\langle S(T-t_2)\gamma, y_T\big\rangle_H - \frac{1}{t_1-t_2}\mathbb{E}\Big\langle \int_{t_2}^{T} S(T-\tau)\chi_{[t_2,t_1]}(\tau)\gamma d\tau, y_T\Big\rangle_H\\
&\quad -\mathbb{E}\int_{t_2}^{T}\langle S(\tau-t_2)\gamma, f(\tau)\rangle_H d\tau + \frac{1}{t_1-t_2}\mathbb{E}\int_{t_2}^{t_1}\Big\langle \int_{t_2}^{\tau} S(\tau-s)\gamma, f(\tau)\Big\rangle_H d\tau\\
&\quad +\frac{1}{t_1-t_2}\mathbb{E}\int_{t_1}^{T}\Big\langle S(\tau-t_1)\int_{t_2}^{t_1} S(t_1-s)\gamma ds, f(\tau)\Big\rangle_H d\tau - \mathbb{E}\int_{t_2}^{T}\langle S(\tau-t_2)\gamma, d\psi(\tau)\rangle_H\\
&\quad +\frac{1}{t_1-t_2}\mathbb{E}\int_{t_2}^{t_1}\Big\langle \int_{t_2}^{\tau} S(\tau-s)\gamma, d\psi(\tau)\Big\rangle_H + \frac{1}{t_1-t_2}\mathbb{E}\int_{t_1}^{T}\Big\langle S(\tau-t_1)\int_{t_2}^{t_1} S(t_1-s)\gamma ds, d\psi(\tau)\Big\rangle_H.
\end{aligned}
\tag{E.24}
$$

Now we analyze the terms in the right hand side of (E.24). First, it is easy to show that

$$
\begin{aligned}
&\lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\mathbb{E}\int_{t_2}^{t_1}\Big\langle \int_{t_2}^{\tau} S(s-t_2)\gamma, f(\tau)\Big\rangle_H d\tau\\
&+\lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\mathbb{E}\int_{t_2}^{t_1}\Big\langle \int_{t_2}^{\tau} S(\tau-s)\gamma, d\psi(\tau)\Big\rangle_H = 0, \qquad \forall\, \gamma\in L^2_{\mathcal{F}_{t_2}}(\Omega;H).
\end{aligned}
\tag{E.25}
$$

Further,

$$
\begin{aligned}
&\lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\mathbb{E}\Big\langle \int_{t_2}^{T} S(T-\tau)\chi_{[t_2,t_1]}(\tau)\gamma d\tau, y_T\Big\rangle_H\\
&= \lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\mathbb{E}\Big\langle \int_{t_2}^{t_1} S(T-\tau)\gamma d\tau, y_T\Big\rangle_H = \mathbb{E}\big\langle S(T-t_2)\gamma, y_T\big\rangle_H.
\end{aligned}
\tag{E.26}
$$

Utilizing the semigroup property of $\{S(t)\}_{t\geq 0}$, we have

$$
\begin{aligned}
&\lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\mathbb{E}\Big[\int_{t_1}^{T}\Big\langle S(\tau-t_1)\int_{t_2}^{t_1} S(t_1-s)\gamma ds, f(\tau)\Big\rangle_H d\tau\\
&\qquad\qquad + \int_{t_1}^{T}\Big\langle S(\tau-t_1)\int_{t_2}^{t_1} S(t_1-s)\gamma ds, d\psi(\tau)\Big\rangle_H\Big]\\
&= \mathbb{E}\int_{t_2}^{T}\big\langle S(\tau-t_2)\gamma, f(\tau)\big\rangle_H d\tau + \mathbb{E}\int_{t_2}^{T}\big\langle S(\tau-t_2)\gamma, d\psi(\tau)\big\rangle_H.
\end{aligned}
\tag{E.27}
$$

From (E.24), (E.25), (E.26) and (E.27), we arrive at

$$
\lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\int_{t_2}^{t_1}\mathbb{E}\langle \gamma, y(\tau)\rangle_H d\tau = \mathbb{E}\big\langle \gamma, \xi^{t_2}\big\rangle_H, \quad \forall\, \gamma\in L^2_{\mathcal{F}_{t_2}}(\Omega;H),\ t_2\in[0,T).
\tag{E.28}
$$

Now, by (E.28), we conclude that, for a.e. $t_2\in(0,T)$

$$
\lim_{t_1\to t_2+0}\frac{1}{t_1-t_2}\int_{t_2}^{t_1}\mathbb{E}\big\langle \xi^{t_2}-y(t_2), y(\tau)\big\rangle_H d\tau = \mathbb{E}\big\langle \xi^{t_2}-y(t_2), \xi^{t_2}\big\rangle_H.
\tag{E.29}
$$

By Lemma E.2, we can find a monotonic sequence $\{\varepsilon_n\}_{n=1}^{\infty}$ of positive numbers with $\lim\limits_{n\to\infty}\varepsilon_n = 0$, such that

$$
\lim_{n\to\infty}\frac{1}{\varepsilon_n}\int_{t_2}^{t_2+\varepsilon_n}\mathbb{E}\langle \xi^{t_2}-y(t_2), y(\tau)\rangle_H d\tau = \mathbb{E}\langle \xi^{t_2}-y(t_2), y(t_2)\rangle_H, \qquad \text{a.e. } t_2\in[0,T).
\tag{E.30}
$$

By (E.29)–(E.30), we arrive at

$$
\mathbb{E}\langle \xi^{t_2}-y(t_2), \xi^{t_2}\rangle_H = \mathbb{E}\langle \xi^{t_2}-y(t_2), y(t_2)\rangle_H, \qquad \text{a.e. } t_2\in[0,T].
\tag{E.31}
$$

By (E.31), we find that $\mathbb{E}\big|\xi^{t_2}-y(t_2)\big|_H^2 = 0$ for $t_2\in[0,T]$ a.e., which implies (E.21).

This completes the proof of Lemma 3.5.

□

References

[1] J. -P. Aubin and I. Ekeland. *Applied nonlinear analysis.* John Wiley & Sons, Inc., New York, 1984.

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

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

[4] A. M. Bruckner, J. B. Bruckner and B. S. Thomson. *Real Analysis.* Prentice Hall (Pearson), Upper Saddle River, 1997

[5] R. A. Carmona and B. Rozovskii, eds. *Stochastic Partial Differential Equations: Six Perspectives.* Math. Surveys Monogr. 64, AMS, Providence, RI, 1999.

[6] D. L. Cohn. *Measure theory.* Birkhäuser/Springer, New York, 2013.

[7] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge, 1992.

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

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

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

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

[12] H. Frankowska, E. M. Marchini and M. Mazzola. *Necessary optimality conditions for infinite dimensional state constrained control problems. J. Differential Equations* **264** (2018), 7294–7327.

[13] H. Frankowska and N. P. Osmolovskii. *Second-order necessary conditions for a strong local minimum in a problem with state and general control constraints. SIAM J. Control Optim.* **56** (2018), 2353–2376.

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

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

[16] H. Frankowska, H. Zhang and X. Zhang. *Stochastic optimal control problems with control and initial-final states constraints. SIAM J. Control Optim.* **56** (2018), 1823–1855.

[17] H. Frankowska, H. Zhang and X. Zhang. *Necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints. Trans. Amer. Math. Soc.* In press.

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

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

[20] Y. Hu and S. Peng. *Adapted solution of backward semilinear stochastic evolution equations. Stoch. Anal. & Appl.* **9** (1991), 445–459.

[21] M. Kisielewicz, *Stochastic differential inclusions and applications.* Springer Optimization and Its Applications, 80. Springer, New York, 2013.

[22] P. Kotelenez. *Stochastic ordinary and stochastic partial differential equations. Transition from microscopic to macroscopic equations.* Springer, New York, 2008.

[23] X. Li and J. Yong. *Optimal Control Theory for Infinite-Dimensional Systems.* Birkhäuser Boston, Inc., Boston, MA, 1995.

[24] Q. Lü. *Exact controllability for stochastic Schrödinger equations. J. Differential Equations.* **255** (2013), 2484–2504.

[25] Q. Lü. *Exact controllability for stochastic transport equations. SIAM J. Control Optim.* **52** (2014), 397–419.

[26] Q. Lü, J. Yong and X. Zhang. *Representation of Itô integrals by Lebesgue/ Bochner integrals. J. Eur. Math. Soc.* **14** (2012), 1795–1823.

[27] 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.

[28] Q. Lü, H. Zhang and X. Zhang. *Second order optimality conditions for optimal control problems of stochastic evolution equations.* arXiv:1811.07337

[29] Q. Lü and X. Zhang. *General Pontryagin-Type Stochastic Maximum Principle and Backward Stochastic Evolution Equations in Infinite Dimensions.* Springer Briefs in Mathematics. Springer, Cham, 2014.

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

[31] Q. Lü and X. Zhang. *Operator-valued backward stochastic Lyapunov equations in infinite dimensions, and its application. Math. Control Relat. Fields.* **8** (2018), 337–381.

[32] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and Their Applications.* Lecture Notes in Math. vol. **1702**. Springer-Verlag, New York, 1999.

[33] A. P. Meyer. *Probability and potentials.* Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London, 1966.

[34] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mischenko. *Mathematical Theory of Optimal Processes.* Wiley, New York, 1962.

[35] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis.* Springer Verlag, New York, 1998.

[36] V. A. Rohlin. *On the fundamental ideas of measure theory. Amer. Math. Soc. Translation 1952*, (1952). 55 pp.

[37] 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.* Lecture Notes in Pure and Appl. Math. vol. **152**. Dekker, New York, 1994, 867–890.

[38] R. Tröltzsch. *Optimal control of partial differential equations. Theory, methods and applications.* Graduate Studies in Mathematics, 112. American Mathematical Society, Providence, RI, 2010.

[39] T. Wang and H. Zhang. *Optimal Control Problems of Forward-Backward Stochastic Volterra Integral Equations with Closed Control Regions. SIAM J. Control Optim.* **55** (2017), 2574–2602.

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

[41] H. Zhang and X. Zhang, *Pointwise second-order necessary conditions for stochastic optimal controls, Part I: The case of convex control constraint. SIAM J. Control Optim.* **53**(2015), 2267–2296.

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