A Mini-Course on Stochastic Control^{*}

Qi Lü[†] and Xu Zhang[‡]

Abstract

This course is addressed to giving a short introduction to control theory of stochastic systems, governed by stochastic differential equations in both finite and infinite dimensions. We will mainly explain the new phenomenon and difficulties in the study of controllability and optimal control problems for these sort of equations. In particular, we will show by some examples that both the formulation of stochastic control problems and the tools to solve them may differ considerably from their deterministic counterpart.

Contents

Τ	Introduction			
2	Some preliminary results from probability theory and stochastic analysis			
	2.1	Probability, random variables and expectation	3	
	2.2	Stochastic processes	5	
	2.3	Itô's integral and its properties	7	
	2.4	Stochastic evolution equations		
	2.5	Backward stochastic evolution equations	11	
3	3 Controllability of stochastic (ordinary) differential equations			
4	Pontryagin-type maximum principle for controlled stochastic (ordinary)			
	differential equations			

^{*}This is a lecture notes of a short introduction to stochastic control. It was written for the LIASFMA (Sino-French International Associated Laboratory for Applied Mathematics) Autumn School "Control and Inverse Problems of Partial Differential Equations" at Zhejiang University, Hangzhou, China from October 17 to October 22, 2016. The second named author thanks Professors Jean-Michel Coron and Tatsien Li for their kind invitation, and Professor Gang Bao and his team for their hospitality during the teaching of this course.

[†]School of Mathematics, Sichuan University, Chengdu 610064, Sichuan Province, China. The research of this author is partially supported by NSF of China under grants 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.

[‡]School of Mathematics, Sichuan University, Chengdu 610064, Sichuan Province, China. The research of this author is partially supported by the NSF of China under grants 11231007, the PCSIRT under grant IRT_15R53 and the Chang Jiang Scholars Program from the Chinese Education Ministry. *E-mail:* zhang_xu@scu.edu.cn.

5	Cor	Controllability of stochastic differential equations in infinite dimensions:						
	$\mathbf{A}\mathbf{n}$	analys	sis of a typical equation	26				
	5.1	Formu	lation of the problem	27				
5.2 Controllability of a class of stochastic parabolic systems				28				
		5.2.1	Some preliminaries	30				
		5.2.2	Proof of the null controllability result	35				
		5.2.3	Proof of the approximate controllability result	38				
	5.3	Null c	ontrollability of stochastic parabolic systems	40				
		5.3.1	A weighted identity and Carleman estimate for a stochastic parabolic-					
			like operator	41				
		5.3.2	Global Carleman estimate for backward stochastic parabolic equations	48				
		5.3.3	Proof of the observability estimate for backward stochastic parabolic					
			equations	50				
6	Pontryagin-type maximum principle for controlled stochastic evolution							
	equations in infinite dimensions							
	6.1	Formu	lation of the problem	51				
	6.2	Pontr	yagin-type maximum principle for convex control domain	53				
6.3 Pontryagin-type maximum principle for the general case				58				
		6.3.1	Relaxed transposition solution to operator-valued backward stochastic					
			evolution equations	58				
		6.3.2	Statement of the Pontryagin-type maximum principle	61				
		6.3.3	Proof of the Pontryagin-type stochastic maximum principle	61				

1 Introduction

It is well-known that control theory was founded by N. Wiener in 1948. After that, this theory was greatly extended to various complicated settings and widely used in sciences and technologies.

Clearly, "control" means a suitable manner for people to change the dynamics of a system under consideration. There are two fundamental issues in control theory. One is "feasibility", or in the language of control theory, controllability, which means that one may find at least one way to achieve a goal. Another is "optimality", or optimal control, which indicates that, one hopes to find the best way, in some sense, to achieve the goal.

Roughly speaking, control theory can be divided into two parts. The first part is control theory for deterministic systems, and the second part is that for stochastic systems. Of course, these two parts are not completely separated but rather they are inextricably linked each other.

Control theory for deterministic systems can be again divided into two parts. The first part is control theory for finite dimensional systems, mainly governed by ordinary differential equations, and the second part is that for (deterministic) distributed parameter systems, mainly described by differential equations in infinite dimensional spaces, typically by partial differential equations. Control theory for finite dimensional systems is by now relatively

mature. There exist a huge list of works on control theory for distributed parameter systems but it is still quite active.

Likewise, control theory for stochastic systems can be divided into two parts. The first part is control theory for stochastic finite dimensional systems, governed by stochastic (ordinary) differential equations, and the second part is that for stochastic distributed parameter systems, described by stochastic differential equations in infinite dimensions, typically by stochastic partial differential equations.

One can find a huge list of publications on control theory for stochastic finite dimensional systems and its applications, say, in mathematical finance. Nevertheless, most of the existing works in this respect are mainly addressed/related to the optimal control problems. As we shall see later in this course, so far controllability theory for stochastic finite dimensional systems is NOT well-developed.

Control theory for stochastic distributed parameter systems, is, in our opinion, still at its very beginning stage. This is actually a rather new branch of mathematical control theory, which is indeed the main concern of this course (See [14, 15] for more material).

One of the most essential difficulties in the study of control theory for stochastic distributed parameter systems is that, compared to the deterministic setting, people know very little about stochastic evolution equation (and in particular, about stochastic partial differential equations) although significant progresses have been made there, especially in recent years. On the other hand, as we shall show in this course, both the formulation of stochastic control problems in infinite dimensions and the tools to solve them may differ considerably from their deterministic/finite-dimensional counterparts. Because of this, one has to develop new mathematical tools to solve some problems in this field.

The rest of this course is organized as follows. In Section 2, we collect some preliminary results (without proofs) from probability theory and stochastic analysis. In Sections 3 and 4, we analyze respectively the controllability and optimal controls for stochastic differential equations in finite dimensions; while in Sections 5–6, we consider the same problems but for stochastic evolution equations in infinite dimensions.

2 Some preliminary results from probability theory and stochastic analysis

For the proofs of the results presented in this section, we refer to [1, 14]. In what follows, we shall denote by C a generic positive constant, which may change from one place to another.

2.1 Probability, random variables and expectation

Fix a nonempty set Ω and a σ -field \mathcal{F} on Ω . Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, i.e. a complete measure space for which $\mathbb{P}(\Omega) = 1$. Any point $\omega \in \Omega$ is called a sample, any $A \in \mathcal{F}$ is called an event and $\mathbb{P}(A)$ represents the probability of event A. If an event $A \in \mathcal{F}$ is such that $\mathbb{P}(A) = 1$, then we may alternatively say that A holds, \mathbb{P} -a.s., or simply A holds a.s.

Let H be a Hilbert space. Each H-valued, strongly measurable function $f: (\Omega, \mathcal{F}) \to (H, \mathcal{B}(H))$ is called an (H-valued) random variable. Clearly, $f^{-1}(\mathcal{B}(H))$ is a sub- σ -field of

 \mathcal{F} , which is called the σ -field generated by f, denoted by $\sigma(f)$. Further, if f is Bochner integrable w.r.t. the measure \mathbb{P} , i.e. the integral

$$\mathbb{E}f \equiv \int_{\Omega} f(\omega) d\, \mathbb{P}(\omega)$$

exists, then we say that f has a mean. We also call $\mathbb{E}f$ the (mathematical) expectation of f. For a given index set Λ and a family of H-valued, random variables $\{f_{\lambda}\}_{{\lambda}\in\Lambda}$ (defined on (Ω, \mathcal{F})), we denote by $\sigma(f_{\lambda}; {\lambda} \in \Lambda)$ the σ -field generated by $\cup_{{\lambda}\in\Lambda}\sigma(f_{\lambda})$.

For any $p \in [1, \infty)$, denote by $L^p_{\mathcal{F}}(\Omega; H) \equiv L^p(\Omega, \mathcal{F}, \mathbb{P}; H)$ the set of all random variables f such that $|f|_H^p$ has means. It is a Banach space with the norm $|f|_{L^p_{\mathcal{F}}(\Omega)} = \left(\int_{\Omega} |f|_H^p d\mathbb{P}\right)^{1/p}$. In particular, $L^2_{\mathcal{F}}(\Omega; H)$ is a Hilbert space. We simply denote $L^p_{\mathcal{F}}(\Omega; \mathbb{R})$ by $L^p_{\mathcal{F}}(\Omega)$. For any $f \in L^2_{\mathcal{F}}(\Omega)$, we define the variance of f by

$$Var f = \mathbb{E}(f - \mathbb{E}f)^2.$$

Let $A, B \in \mathcal{F}$. We say that A and B are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Let \mathcal{J}_1 and \mathcal{J}_2 be two subsets of \mathcal{F} . We say that \mathcal{J}_1 and \mathcal{J}_2 are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for any $A \in \mathcal{J}_1$ and $B \in \mathcal{J}_2$. Let $f, g: (\Omega, \mathcal{F}) \to (H, \mathcal{B}(H))$ be two random variables. We say that f and g (resp. f and \mathcal{J}_1) are independent if $\sigma(f)$ and $\sigma(g)$ (resp. $\sigma(f)$ and \mathcal{J}_1) are independent.

Let $X: (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a random variable. We call $F(x) \equiv \mathbb{P}\{X \leq x\}$ the distribution function of X. If for some function $p(\cdot)$, one has

$$F(x) = \int_{-\infty}^{x} p(\xi)d\xi,$$

then the function $p(\cdot)$ is called the density of X. If $p(\cdot)$ is of the following form:

$$p(x) = (2\pi\mu)^{-1/2} \exp\left\{-\frac{1}{2\mu}(x-\lambda)^2\right\},$$

where $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}^+$, then X is called a normally distributed random variable (or X is a normal distribution). Clearly, λ and μ are the mean and variance of f, respectively.

Assume that $\mathcal{J} \subset \mathcal{F}$ is a given sub- σ -field and $f \in L^1_{\mathcal{F}}(\Omega; H)$. Define a function on \mathcal{J} by

$$\nu(B) = \int_B f d\mathbb{P}, \quad \forall B \in \mathcal{J}.$$

It is easy to see that ν is an (H-valued) vector measure of bounded variation on (Ω, \mathcal{J}) , and $\nu(B) = 0$ whenever $\mathbb{P}(B) = 0$. Hence, there is a (unique) function in $L^1_{\mathcal{J}}(\Omega; H)$, denoted by $\mathbb{E}(f \mid \mathcal{J})$, such that

$$\int_{B} \mathbb{E}(f \mid \mathcal{J}) d\mathbb{P} = \int_{B} f d\mathbb{P}, \qquad \forall B \in \mathcal{J}.$$
(2.1)

This function is called the conditional expectation of f given σ -field \mathcal{J} .

Example 2.1 Let $B_1, B_2 \subset \mathcal{F}$ such that $B_1 \cup B_2 = \Omega$, $B_1 \cap B_2 = \emptyset$, and $\mathbb{P}(B_k) > 0$ for all k = 1, 2. Let $f \in L^1_{\mathcal{F}}(\Omega; H)$ and $\mathcal{J} = \{\emptyset, \Omega, B_1, B_2\}$. Then

$$\mathbb{E}(f \mid \mathcal{J})(\omega) = \sum_{k=1}^{2} \frac{1}{\mathbb{P}(B_k)} \int_{B_k} f d\mathbb{P} \chi_{B_k}(\omega).$$

We collect some basic properties of conditional expectation as follows.

Theorem 2.1 Let \mathcal{J} be a sub- σ -field of \mathcal{F} and $f \in L^1_{\mathcal{F}}(\Omega; H)$. It holds that:

- 1) The map $\mathbb{E}(\cdot \mid \mathcal{J}): L^1_{\mathcal{F}}(\Omega; H) \to L^1_{\mathcal{J}}(\Omega; H)$ is linear and continuous;
- 2) $\mathbb{E}(a \mid \mathcal{J}) = a$, $\mathbb{P}|_{\mathcal{J}}$ -a.s., $\forall a \in H$;
- 3) If $\alpha \in L^1_{\mathcal{I}}(\Omega)$ satisfies $\alpha f \in L^1_{\mathcal{F}}(\Omega; H)$, then

$$\mathbb{E}(\alpha f \mid \mathcal{J}) = \alpha \mathbb{E}(f \mid \mathcal{J}), \qquad \mathbb{P}|_{\mathcal{J}}\text{-a.s.}$$

In particular, $\mathbb{E}(\alpha \mid \mathcal{J}) = \alpha$, $\mathbb{P}|_{\mathcal{J}} - \text{a.s.}$;

4) If f is independent of \mathcal{J} , then

$$\mathbb{E}(f \mid \mathcal{J}) = \mathbb{E}f, \qquad \mathbb{P}|_{\mathcal{J}}\text{-a.s.};$$

5) Let \mathcal{J}' be a sub- σ -field of \mathcal{J} . Then

$$\mathbb{E}(\mathbb{E}(f \mid \mathcal{J}) \mid \mathcal{J}') = \mathbb{E}(\mathbb{E}(f \mid \mathcal{J}') \mid \mathcal{J}) = \mathbb{E}(f \mid \mathcal{J}'), \qquad \mathbb{P}|_{\mathcal{J}'}\text{-a.s.};$$

6) (Jensen's inequality) Let $\phi: H \to \mathbb{R}$ be a convex function such that $\phi(f) \in L^1_{\mathcal{F}}(\Omega)$. Then

$$\phi(\mathbb{E}(f \mid \mathcal{J})) \leq \mathbb{E}(\phi(f) \mid \mathcal{J}), \qquad \mathbb{P}|_{\mathcal{J}}\text{-a.s.}$$

In particular, for any $p \geq 1$,

$$\left| \mathbb{E}(f \mid \mathcal{J}) \right|_{H}^{p} \leq \mathbb{E}(|f|_{H}^{p} \mid \mathcal{J}), \qquad \mathbb{P}|_{\mathcal{J}}$$
-a.s.

provided that $\mathbb{E}|f|_H^p$ exists.

2.2Stochastic processes

Let $\mathcal{I} = [0, T]$ with T > 0. A family of H-valued random variables $\{X(t)\}_{t \in \mathcal{I}}$ is called a stochastic process. For any $\omega \in \Omega$, the map $t \mapsto X(t,\omega)$ is called a sample path (of X). We will interchangeably use $\{X(t)\}_{t\in\mathcal{I}}$, $X(\cdot)$ or even X to denote a (stochastic) process.

An (H-valued) process $X(\cdot)$ is said to be continuous (resp., cádlàg, i.e., right-continuous)with left limits) if there is a \mathbb{P} -null set $N \in \mathcal{F}$, such that for any $\omega \in \Omega \setminus N$, the sample path $X(\cdot,\omega)$ is continuous (resp., cádlàg) in H. In a similar way, one can define rightcontinuous stochastic processes, etc. Two (H-valued) processes $X(\cdot)$ and $X(\cdot)$ are said to be stochastically equivalent if $\mathbb{P}(\{X(t) = \overline{X}(t)\}) = 1$ for any $t \in \mathcal{I}$. In this case, one is said to be a modification of the other.

We call a family of sub- σ -fields $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$ in \mathcal{F} a filtration if $\mathcal{F}_{t_1}\subset\mathcal{F}_{t_2}$ for all $t_1,t_2\in\mathcal{F}_{t_2}$ \mathcal{I} with $t_1 \leq t_2$. For any $t \in \mathcal{I}$, we put

$$\mathcal{F}_{t+} \stackrel{\triangle}{=} \bigcap_{s \in (t, +\infty) \cap \mathcal{I}} \mathcal{F}_{s}, \qquad \mathcal{F}_{t-} \stackrel{\triangle}{=} \bigcup_{s \in [0, t) \cap \mathcal{I}} \mathcal{F}_{s}.$$

If $\mathcal{F}_{t+} = \mathcal{F}_t$ (resp. $\mathcal{F}_{t-} = \mathcal{F}_t$), then $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$ is said to be right (resp. left) continuous. In the sequel, for simplicity, we write $\mathbf{F} = \{\mathcal{F}_t\}_{t\in\mathcal{I}}$ unless we want to emphasize what \mathcal{F}_t or I exactly is. We call $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ a filtered probability space.

We say that $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ satisfies the usual condition if $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, \mathcal{F}_0 contains all \mathbb{P} -null sets in \mathcal{F} , and \mathbf{F} is right continuous. We shall keep these assumptions in what follows unless stated otherwise.

Definition 2.1 Let $X(\cdot)$ be an H-valued process.

- 1) $X(\cdot)$ is said to be measurable if the map $(t,\omega) \mapsto X(t,\omega)$ is strongly $(\mathcal{B}(\mathcal{I}) \times \mathcal{F})/\mathcal{B}(H)$ -measurable;
- 2) $X(\cdot)$ is said to be **F**-adapted if it is measurable, and for each $t \in \mathcal{I}$, the map $\omega \mapsto X(t,\omega)$ is strongly $\mathcal{F}_t/\mathcal{B}(H)$ -measurable;
- 3) $X(\cdot)$ is said to be **F**-progressively measurable if for each $t \in \mathcal{I}$, the map $(s, \omega) \mapsto X(s, \omega)$ from $[0, t] \times \Omega$ to H is strongly $(\mathcal{B}([0, t]) \times \mathcal{F}_t)/\mathcal{B}(H)$ -measurable.

A set $A \in \mathcal{I} \times \Omega$ is called progressively measurable w.r.t. **F** if the process $\chi_A(\cdot)$ is progressive. The class of all progressively measurable sets is a σ -field, called the progressive σ -field w.r.t. **F**, denoted by \mathbb{F} . One can show that, an (H-valued) process $\varphi : [0, T] \times \Omega \to H$ is **F**-progressively measurable if and only if it is strongly \mathbb{F} -measurable.

It is clear that if $X(\cdot)$ is **F**-progressively measurable, it must be **F**-adapted. Conversely, it can be proved that, for any **F**-adapted process $X(\cdot)$, there is an **F**-progressively measurable process $\widetilde{X}(\cdot)$ which is stochastically equivalent to $X(\cdot)$. For this reason, in the sequel, by saying that a process $X(\cdot)$ is **F**-adapted, we mean that it is **F**-progressively measurable.

For any $p, q \in [1, \infty)$, write

$$L_{\mathbb{F}}^{p}(\Omega; L^{q}(0, T; H)) \stackrel{\triangle}{=} \Big\{ \varphi : (0, T) \times \Omega \to H \, \Big| \, \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted and } \mathbb{E}\Big(\int_{0}^{T} |\varphi(t)|_{H}^{q} dt \Big)^{\frac{p}{q}} < \infty \Big\},$$

$$L_{\mathbb{F}}^{q}(0, T; L^{p}(\Omega; H)) \stackrel{\triangle}{=} \Big\{ \varphi : (0, T) \times \Omega \to H \, \Big| \, \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted and } \int_{0}^{T} \Big(\mathbb{E}|\varphi(t)|_{H}^{p} \Big)^{\frac{q}{p}} dt < \infty \Big\}.$$

Similarly, we may also define (for $1 \le p, q < \infty$)

$$\left\{ \begin{array}{ll} L_{\mathbb{F}}^{\infty}(\Omega;L^{q}(0,T;H)), & L_{\mathbb{F}}^{p}(\Omega;L^{\infty}(0,T;H)), & L_{\mathbb{F}}^{\infty}(\Omega;L^{\infty}(0,T;H)), \\ L_{\mathbb{F}}^{\infty}(0,T;L^{p}(\Omega;H)), & L_{\mathbb{F}}^{q}(0,T;L^{\infty}(\Omega;H)), & L_{\mathbb{F}}^{\infty}(0,T;L^{\infty}(\Omega;H)). \end{array} \right.$$

All these spaces are Banach spaces (with the canonical norms). In the sequel, we shall simply denote $L^p_{\mathbb{F}}(\Omega; L^p(0, T; H)) \equiv L^p_{\mathbb{F}}(0, T; L^p(\Omega; H))$ by $L^p_{\mathbb{F}}(0, T; H)$; and further simply denote $L^p_{\mathbb{F}}(0, T; \mathbb{R})$ by $L^p_{\mathbb{F}}(0, T)$.

For any $p \in [1, \infty)$, set

$$L_{\mathbb{F}}^{p}(\Omega; C([0,T]; H)) \stackrel{\triangle}{=} \left\{ \varphi : [0,T] \times \Omega \to H \, \middle| \, \varphi(\cdot) \text{ is continuous,} \right.$$

$$\mathbf{F}\text{-adapted and } \mathbb{E}\left(|\varphi(\cdot)|_{C([0,T]; H)}^{p} \right) < \infty \right\}$$

and

$$\begin{split} C_{\mathbb{F}}([0,T];L^p(\Omega;H)) & \stackrel{\triangle}{=} \Big\{ \varphi : [0,T] \times \Omega \to H \, \Big| \, \varphi(\cdot) \text{ is } \mathbf{F}\text{-adapted} \\ & \text{and } \varphi(\cdot) : [0,T] \to L^p_{\mathcal{F}_T}(\Omega;H) \text{ is continuous} \Big\}. \end{split}$$

One can show that both $L^p_{\mathbb{F}}(\Omega; C([0,T];H))$ and $C_{\mathbb{F}}([0,T];L^p(\Omega;H))$ are Banach spaces with canonical norms $|\varphi(\cdot)|_{L^p_{\mathbb{F}}(\Omega;C([0,T];H))} = \left(\mathbb{E}(|\varphi(\cdot)|^p_{C([0,T];H)})\right)^{1/p}$ and $|\varphi(\cdot)|_{C_{\mathbb{F}}([0,T];L^p(\Omega;H))} = \max_{t \in [0,T]} \left(\mathbb{E}(|\varphi(t)|^p_H)\right)^{1/p}$, respectively. Also, we denote by $D_{\mathbb{F}}([0,T];L^p(\Omega;H))$ the Banach space of all processes such that X(t) is càdlàg in $L^p_{\mathcal{F}_T}(\Omega;H)$, w.r.t. $t \in [0,T]$, such that $|\mathbb{E}|X(\cdot)|^p_H|^{1/p}_{L^\infty(0,T)} < \infty$, with the canonical norm.

We need to introduce two important classes of stochastic processes, i.e., Brownian motion and martingale.

Definition 2.2 A continuous **F**-adapted process $W(\cdot)$ is called a 1-dimensional Brownian motion (over \mathcal{I}), if for all $s, t \in \mathcal{I}$ with $0 \le s < t < T$, W(t) - W(s) is independent of \mathcal{F}_s , and normally distributed with mean 0 and variance t - s. In addition, if $\mathbb{P}(W(0) = 0) = 1$, then $W(\cdot)$ is called a 1-dimensional standard Brownian motion.

In the seugel, we fix a 1-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$. Write

$$\mathcal{F}_t^W \stackrel{\triangle}{=} \sigma(W(s); \ s \in [0, t]) \subset \mathcal{F}_t, \qquad \forall \ t \in \mathcal{I}.$$
 (2.2)

Generally, the filtration $\{\mathcal{F}_t^W\}_{t\in\mathcal{I}}$ is left-continuous, but not necessarily right-continuous. Nevertheless, the augmentation $\{\hat{\mathcal{F}}_t^W\}_{t\in\mathcal{I}}$ of $\{\mathcal{F}_t^W\}_{t\in\mathcal{I}}$ by adding all \mathbb{P} -null sets is continuous, and $W(\cdot)$ is still a Brownian motion on the (augmented) filtered probability space $(\Omega, \mathcal{F}, \{\hat{\mathcal{F}}_t^W\}_{t\in\mathcal{I}}, \mathbb{P})$. In the sequel, by saying that \mathbf{F} is the natural filtration generated by $W(\cdot)$, we mean that \mathbf{F} is generated as in (2.2) with the above augmentation, and hence in this case \mathbf{F} is continuous.

Definition 2.3 An H-valued, \mathbf{F} -adapted process $X = \{X(t)\}_{t \in \mathcal{I}}$ is called an \mathbf{F} -martingale, if X(t) is Bochner integrable for each $t \in \mathcal{I}$, and $E(X(t) \mid \mathcal{F}_s) = X(s)$ a.s., for any $t, s \in \mathcal{I}$ with s < t.

Clearly, for any $f \in L^1_{\mathcal{F}}(\Omega; H)$, the process $\{\mathbb{E}(f \mid \mathcal{F}_t)\}_{t \in \mathcal{I}}$ is an \mathbb{F} -martingale. Write

 $\mathcal{M}^2[0,T] = \big\{ \, X \in L^2_{\mathbb{F}}(0,T;H) \, \big| \, X \text{ is a right-continuous, } \mathbb{F}\text{-martingale with } X(0) = 0, \mathbb{P}\text{-a.s.} \, \big\},$ $\mathcal{M}^2_c[0,T] = \big\{ \, X \in \mathcal{M}^2[0,T] \, \, \big| \, \, X \text{ is continuous} \, \big\} \, .$

Define

$$|X|_{\mathcal{M}^2[0,T]} = \sqrt{\mathbb{E}|X(T)|_H^2}, \qquad \forall X \in \mathcal{M}^2[0,T].$$

Then, $(\mathcal{M}^2[0,T], |\cdot|_{\mathcal{M}^2[0,T]})$ is a Hilbert space, and $\mathcal{M}_c^2[0,T]$ is a closed subspace of $\mathcal{M}^2[0,T]$.

2.3 Itô's integral and its properties

We now define the Itô integral

$$\int_0^T X(t)dW(t) \tag{2.3}$$

of an H-valued, \mathbf{F} -adapted stochastic process $X(\cdot)$ (satisfying suitable conditions) w.r.t. a Brownian motion W(t). Note that one cannot define (2.3) to be a Lebesgue-Stieltjes type

integral by regarding ω as a parameter. Indeed, the map $t \ni [0,T] \mapsto W(t,\cdot)$ is nowhere differentiable, \mathbb{P} -a.s.

Denote by \mathcal{L}_0 the class of simple processes $f \in L^2_{\mathbb{F}}(0,T;H)$ of the forms:

$$f(t,\omega) = \sum_{j=0}^{n} f_j(\omega) \chi_{[t_j,t_{j+1})}(t), \qquad (t,\omega) \in [0,T] \times \Omega, \tag{2.4}$$

where $0 = t_0 < t_1 < \dots < t_{n+1} = T$, f_j is \mathcal{F}_{t_j} -measurable with $\sup \{|f_j(\omega)|_H \mid j \in \{0,\dots,n\}, \omega \in \Omega\} < \infty$. One can show that \mathcal{L}_0 is dense in $L^2_{\mathbb{F}}(0,T;H)$.

We now define the Itô integral (2.3) as a mapping $f \in L^2_{\mathbb{F}}(0,T) \mapsto I(f) \in \mathcal{M}^2_c[0,T]$. First, assume that $f \in \mathcal{L}_0$ takes the form of (2.4). Then we set

$$I(f)(t,\omega) = \sum_{j=0}^{n} f_j(\omega)[W(t \wedge t_{j+1}, \omega) - W(t \wedge t_j, \omega)].$$
 (2.5)

It is easy to show that $I(f) \in \mathcal{M}_c^2[0,T]$ and the following Itô isometry holds:

$$|I(f)|_{\mathcal{M}^2[0,T]} = |f|_{L^2_{\mathbb{R}}(0,T;H)}.$$
 (2.6)

Generally, for $f \in L^2_{\mathbb{F}}(0,T;H)$, one can find a sequence of $\{f_k\} \subset \mathcal{L}_0$ such that $|f_k - f|_{L^2_{\mathbb{F}}(0,T;H)} \to 0$ as $k \to \infty$. Since $|I(f_k) - I(f_j)|_{\mathcal{M}^2[0,T]} = |f_k - f_j|_{L^2_{\mathbb{F}}(0,T;H)}$, one deduces that $\{I(f_k)\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathcal{M}^2[0,T]$ and therefore, it converges to a unique element $X \in \mathcal{M}^2[0,T]$. Clearly, X is determined uniquely by f and is independent of the particular choice of $\{f_k\}_{k=1}^{\infty}$. This process is called the Itô integral of $f \in L^2_{\mathbb{F}}(0,T;H)$ w.r.t. the Brownian Motion $W(\cdot)$. We shall denote it by $\int_0^t f(s)dW(s)$ or simply $\int_0^t fdW$.

Theorem 2.2 Let $f, g \in L^2_{\mathbb{F}}(0, T; H), a, b \in L^2_{\mathcal{F}_s}(\Omega), T \ge t > s \ge 0$. Then

- 1) $\int_{s}^{t} (af + bg)dW = a \int_{s}^{t} fdW + b \int_{s}^{t} gdW$, P-a.s.;
- 2) $\mathbb{E}\left(\int_{s}^{t} f dW \mid \mathcal{F}_{s}\right) = 0, \mathbb{P}\text{-a.s.};$
- 3) $\mathbb{E}\left(\langle \int_{s}^{t} f dW, \int_{s}^{t} g dW \rangle_{H} \mid \mathcal{F}_{s}\right) = \mathbb{E}\left(\int_{s}^{t} \langle f(r, \cdot), g(r, \cdot) \rangle_{H} dr \mid \mathcal{F}_{s}\right)$, \mathbb{P} -a.s.;
- 4) The stochastic process $\{\int_0^t f(s)dW(s)\}_{t\in[0,T]}$ is a martingale.

For any $p \in (0, \infty)$, denote by $L_{\mathbb{F}}^{p,loc}(0, T; H)$ the set of **F**-adapted stochastic processes $f(\cdot)$ satisfying only $\int_0^T |f(t)|_H^p dt < \infty$, \mathbb{P} -a.s. One can define the Itô integral $\int_0^t \Phi dW$ for $\Phi \in L_{\mathbb{F}}^{2,loc}(0,T;H)$ (See [14, 23] for more details).

Definition 2.4 An H-valued, \mathbf{F} -adapted process $X(\cdot)$ is called an $It\hat{o}$ process if there exist two H-valued stochastic processes $\phi(\cdot) \in L^{1,loc}_{\mathbb{F}}(0,T;H)$ and $\Phi(\cdot) \in L^{2,loc}_{\mathbb{F}}(0,T;H)$ such that

$$X(t) = X(0) + \int_0^t \phi(s)ds + \int_0^t \Phi(s)dW(s), \quad \mathbb{P}\text{-}a.s., \quad \forall \ t \in [0, T].$$
 (2.7)

The following fundamental result is known as *Itô's formula*.

Theorem 2.3 Let $X(\cdot)$ be given by (2.7). Let $F:[0,T]\times H\to\mathbb{R}$ be a function such that its partial derivatives F_t , F_x and F_{xx} are uniformly continuous on any bounded subset of $[0,T]\times H$. Then,

$$F(t, X(t)) - F(0, X(0))$$

$$= \int_0^t F_x(s, X(s)) \Phi(s) dW(s) + \int_0^t \left[F_t(s, X(s)) + \left\langle F_x(s, X(s)), \phi(s) \right\rangle_H \right] ds, \quad \mathbb{P}\text{-}a.s., \quad \forall \ t \in [0, T].$$

$$(2.8)$$

The following deep result, known as the *Burkholder-Davis-Gundy inequality*, links Itô's integral to the Lebesgue/Bochner integral.

Theorem 2.4 For any p > 0, there exists a constant $C_p > 0$ such that for any T > 0 and $f \in L_{\mathbb{F}}^{p,loc}(0,T;H)$,

$$\frac{1}{\mathcal{C}_p} \mathbb{E}\left(\int_0^T |f(s)|_H^2 ds\right)^{\frac{p}{2}} \le \mathbb{E}\left(\sup_{t \in [0,T]} \left|\int_0^t f(s) dW(s)\right|_H^p\right) \le \mathcal{C}_p \mathbb{E}\left(\int_0^T |f(s)|_H^2 ds\right)^{\frac{p}{2}}. \tag{2.9}$$

2.4 Stochastic evolution equations

In what follows, we shall always assume that H is a separable Hilbert space, and A is an unbounded linear operator (with domain D(A) on H), which is the infinitesimal generator of a C_0 -semigroup $\{S(t)\}_{t\geq 0}$. Denote by A^* the dual operator of A. Clearly, D(A) is a Hilbert space with the usual graph norm, and A^* is the infinitesimal generator of $\{S^*(t)\}_{t\geq 0}$, the dual C_0 -semigroup of $\{S(t)\}_{t\geq 0}$.

Let us consider the following stochastic evolution equation:

$$\begin{cases}
 dX(t) = \left[AX(t) + F(t, X(t)) \right] dt + \widetilde{F}(t, X(t)) dW(t) & \text{in } (0, T], \\
 X(0) = X_0.
\end{cases}$$
(2.10)

Here $X_0 \in L^p_{\mathcal{F}_0}(\Omega; H)$ (for some $p \geq 2$), and $F(\cdot, \cdot)$ and $\widetilde{F}(\cdot, \cdot)$ are measurable functions from $[0, T] \times \Omega \times H$ to H, satisfying the following conditions:

Condition 2.1

$$\begin{cases}
|F(t,y) - F(t,z)|_{H} \leq \mathcal{C}|y - z|_{H}, & \forall y, z \in H, \text{ a.e. } t \in [0,T], \mathbb{P}\text{-a.s.}, \\
|\widetilde{F}(t,y) - \widetilde{F}(t,z)|_{H} \leq \mathcal{C}|y - z|_{H}, & \forall y, z \in H, \text{ a.e. } t \in [0,T], \mathbb{P}\text{-a.s.}, \\
F(\cdot,0) \in L_{\mathbb{F}}^{p}(\Omega; L^{1}(0,T;H)), & \widetilde{F}(\cdot,0) \in L_{\mathbb{F}}^{p}(\Omega; L^{2}(0,T;H)).
\end{cases} (2.11)$$

First, we give the notion of strong solution to the equation (2.10).

Definition 2.5 An H-valued stochastic process $X(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega;H))$ is called a strong solution to (2.10) if $X(t,\omega) \in D(A)$ for a.e. $(t,\omega) \in [0,T] \times \Omega$, $AX(\cdot) \in L^{1,loc}_{\mathbb{F}}(0,T;H)$, and for all $t \in [0,T]$,

$$X(t) = X_0 + \int_0^t \left[AX(s) + F(s, X(s)) \right] ds + \int_0^t \widetilde{F}(s, X(s)) dW(s), \text{ \mathbb{P}-a.s.}$$

Generally speaking, one needs very strong conditions to guarantee the existence of a strong solution. Thus, people introduce two types of "weak" solutions.

Definition 2.6 An H-valued stochastic process $X(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega;H))$ is called a weak solution to (2.10) if for any $t \in [0,T]$ and $\xi \in D(A^*)$,

$$\langle X(t), \xi \rangle_{H} = \langle X_{0}, \xi \rangle_{H} + \int_{0}^{t} (\langle X(s), A^{*}\xi \rangle_{H} + \langle F(s, X(s)), \xi \rangle_{H}) ds$$

$$+ \int_{0}^{t} \langle \widetilde{F}(s, X(s)), \xi \rangle_{H} dW(s), \qquad \mathbb{P}\text{-a.s.}$$

Definition 2.7 An H-valued stochastic process $X(\cdot) \in C_{\mathbb{F}}([0,T];L^p(\Omega;H))$ is called a mild solution to (2.10) if for any $t \in [0,T]$,

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(s,X(s))ds + \int_0^t S(t-s)\widetilde{F}(s,X(s))dW(s), \quad \mathbb{P}\text{-}a.s.$$

It is easiest to show the well-posedness of (2.10) in the framework of mild solution among the above three kinds of solutions. Indeed, we have the following result.

Theorem 2.5 Let $p \geq 2$. Then, there is a unique mild solution $X(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega; H))$ to (2.10). Moreover,

$$|X(\cdot)|_{C_{\mathbb{F}}([0,T];L^{p}(\Omega;H))} \leq \mathcal{C}(|X_{0}|_{L^{p}_{\mathcal{F}_{0}}(\Omega;H)} + |F(\cdot,0)|_{L^{p}_{\mathbb{F}}(\Omega;L^{1}(0,T;H))} + |\widetilde{F}(\cdot,0)|_{L^{p}_{\mathbb{F}}(\Omega;L^{2}(0,T;H))}). \tag{2.12}$$

If p > 2 or $\{S(t)\}_{t\geq 0}$ is a contraction semigroup, then one can get a better regularity for the mild solution with respect to time, i.e., $X(\cdot, \omega) \in C([0, T]; H)$, \mathbb{P} -a.s. Here we only consider the latter case.

Theorem 2.6 If A generates a contraction semigroup and $p \geq 1$, then (2.10) admits a unique mild solution $X(\cdot) \in L^p_{\mathbb{F}}(\Omega; C([0,T];H))$. Moreover,

$$|X(\cdot)|_{L_{\mathbb{F}}^{p}(\Omega;C([0,T];H))} \leq \mathcal{C}(|X_{0}|_{L_{\mathcal{F}_{0}}^{p}(\Omega;H)} + |F(\cdot,0)|_{L_{\mathbb{F}}^{p}(\Omega;L^{1}(0,T;H))} + |\widetilde{F}(\cdot,0)|_{L_{\mathbb{F}}^{p}(\Omega;L^{2}(0,T;H))}). \tag{2.13}$$

The following result indicates the space smoothing effect of mild solutions to a class of stochastic evolutions equations, say the stochastic parabolic equation.

Theorem 2.7 Let $p \geq 1$. Assume that A is a self-adjoint, negative definite (unbounded linear) operator on H. Then, the equation (2.10) admits a unique mild solution $X(\cdot) \in L^p_{\mathbb{F}}(\Omega; C([0,T];H)) \cap L^p_{\mathbb{F}}(\Omega; L^2(0,T;D((-A)^{\frac{1}{2}})))$. Moreover,

$$|X(\cdot)|_{L_{\mathbb{F}}^{p}(\Omega;C([0,T];H))} + |X(\cdot)|_{L_{\mathbb{F}}^{p}(\Omega;L^{2}(0,T;D((-A)^{\frac{1}{2}})))}$$

$$\leq C(|X_{0}|_{L_{\mathcal{F}_{0}}^{p}(\Omega;H)} + |F(\cdot,0)|_{L_{\mathbb{F}}^{p}(\Omega;L^{1}(0,T;H))} + |\widetilde{F}(\cdot,0)|_{L_{\mathbb{F}}^{p}(\Omega;L^{2}(0,T;H))}).$$
(2.14)

Next result gives the relationship between mild and weak solutions to (2.10).

Theorem 2.8 Any weak solution to (2.10) is also a mild solution and vice versa.

Usually, the mild solution does not have enough regularity. For example, when establishing the pointwise identity for Carleman estimate, we need the functions to be second order differentiable in the sense of weak derivative with respect to the spatial variable. Nevertheless, these problems can be solved by the following strategy:

- 1. Introduce some approximating equations with strong solutions such that the limit of these strong solutions is the mild or weak solution of the original equation.
- 2. Obtain the desired properties for these strong solutions.
- 3. Utilize the density argument to establish the desired properties for the mild/weak solutions.

There are many methods to implement the above three steps in the setting of deterministic partial differential equations. Roughly speaking, any of these methods, which does not destroy the adaptedness of the solution, can be applied to stochastic partial differential equations. Here we only present one approach. Introduce an approximating system of (2.10) as follows:

$$\begin{cases}
 dX^{\lambda}(t) = AX^{\lambda}(t)dt + R(\lambda)F(t, X^{\lambda}(t))dt + R(\lambda)\widetilde{F}(t, X^{\lambda}(t))dW(t) & \text{in } (0, T], \\
 X^{\lambda}(0) = R(\lambda)X_{0} \in D(A).
\end{cases}$$
(2.15)

Here $\lambda \in \rho(A)$, the resolvent set of A, and $R(\lambda) \stackrel{\triangle}{=} \lambda(\lambda I - A)^{-1}$ with I being the identity operator on H.

Theorem 2.9 For each $X_0 \in L^p_{\mathcal{F}_0}(\Omega; H)$ with $p \geq 2$ and $\lambda \in \rho(A)$, the equation (2.15) admits a unique strong solution $X^{\lambda}(\cdot) \in C_{\mathbb{F}}([0,T]; L^p(\Omega; H))$. Moreover, as $\lambda \to \infty$, the solution $X^{\lambda}(\cdot)$ converges to $X(\cdot)$ in $C_{\mathbb{F}}([0,T]; L^p(\Omega; H))$, where $X(\cdot)$ solves (2.10) in the sense of the mild solution.

2.5 Backward stochastic evolution equations

Backward stochastic differential equations and more generally, backward stochastic evolution equations are by-products in the study of stochastic control theory, both of which have independent interest and been applied in other places.

Let us consider the following H-valued, backward stochastic evolution equation

$$\begin{cases}
 dy(t) = -[Ay(t) + F(t, y(t), Y(t))]dt - Y(t)dW(t) & \text{in } [0, T), \\
 y(T) = \xi.
\end{cases}$$
(2.16)

Here $\xi \in L^p_{\mathcal{F}_T}(\Omega; H)$ (for some $p \geq 1$), $F : [0, T] \times \Omega \times H \times H \to H$ is a measurable functionm satisfying that

$$\begin{cases}
F(\cdot,0,0) \in L_{\mathbb{F}}^{p}(\Omega; L^{1}(0,T;H)), \\
|F(t,y_{1},z_{1}) - F(t,y_{2},z_{2})|_{H} \leq \mathcal{C}(|y_{1} - y_{2}|_{H} + |z_{1} - z_{2}|_{H}), \\
\forall y_{1}, y_{2}, z_{1}, z_{2} \in H, \text{ a.e. } t \in [0,T], \mathbb{P}\text{-a.s.}
\end{cases} (2.17)$$

Similarly to the case of stochastic evolution equations, one introduces below notions of strong, weak and mild solutions to the equation (2.16).

Definition 2.8 A stochastic process $(y(\cdot), Y(\cdot)) \in L^p_{\mathbb{F}}(\Omega; C([0,T]; H)) \times L^p_{\mathbb{F}}(\Omega; L^2(0,T; H))$ is called a strong solution to (2.16) if $y(t) \in D(A)$ for a.e. $(t,\omega) \in [0,T] \times \Omega$, $Ay(\cdot) \in L^{1,loc}_{\mathbb{F}}(0,T; H)$, and for all $t \in [0,T]$,

$$y(t) = \xi + \int_t^T \left[Ay(s) + F(s, y(s), Y(s)) \right] ds + \int_t^T Y(s) dW(s), \quad \mathbb{P}\text{-a.s.}$$

Definition 2.9 A stochastic process $(y(\cdot), Y(\cdot)) \in L^p_{\mathbb{F}}(\Omega; C([0, T]; H)) \times L^p_{\mathbb{F}}(\Omega; L^2(0, T; H))$ is called a weak solution to (2.16) if for any $t \in [0, T]$ and $\eta \in D(A^*)$,

$$\begin{split} \langle\, y(t),\eta\,\rangle_H &= \langle\, \xi,\eta\,\rangle_H + \int_t^T \langle\, y(s),A^*\eta\,\rangle_H ds \\ &- \int_t^T \langle\, F(s,y(s),Y(s)),\eta\,\rangle_H ds - \int_t^T \langle\, Y(s),\eta\,\rangle_H dW(s), \quad \mathbb{P}\text{-a.s.} \end{split}$$

Definition 2.10 A stochastic process $(y(\cdot), Y(\cdot)) \in L^p_{\mathbb{F}}(\Omega; C([0,T]; H)) \times L^p_{\mathbb{F}}(\Omega; L^2(0,T; H))$ is called a mild solution to (2.16) if for any $t \in [0,T]$,

$$y(t) = S(T-t)\xi + \int_t^T S(s-t)F(s,y(s),Y(s))ds + \int_t^T S(s-t)Y(s)dW(s), \quad \mathbb{P}\text{-a.s.}$$

Similar to Theorem 2.5 (but here one needs that the filtration \mathbf{F} is natural), one can get the well-posedness of (2.16) in the sense of mild solution.

Theorem 2.10 Assume that **F** is the natural filtration generated by $W(\cdot)$. Then, for any $p \geq 1$ and $\xi \in L^p_{\mathcal{F}_T}(\Omega; H)$, the equation (2.16) admits a unique mild solution $(y(\cdot), Y(\cdot)) \in L^p_{\mathbb{F}}(\Omega; C([0,T]; H)) \times L^p_{\mathbb{F}}(\Omega; L^2(0,T; H))$ satisfying that

$$|(y,Y)|_{L_{\mathbb{F}}^{p}(\Omega;C([0,T];H))\times L_{\mathbb{F}}^{p}(\Omega;L^{2}(0,T;H))} \leq \mathcal{C}(|\xi|_{L_{\mathcal{F}_{T}}^{p}(\Omega;H)} + |F(\cdot,0,0)|_{L_{\mathbb{F}}^{p}(0,T;H)}). \tag{2.18}$$

Also, similar to Theorem 2.8, we have the following relationship between the weak and mild solutions to (2.16).

Theorem 2.11 A stochastic process (y, Y) is a weak solution to (2.16) if and only if it is a mild solution to the same equation.

Similarly to Theorem 2.7, the following result describes the the smoothing effect of mild solutions to a class of backward stochastic evolution equations.

Theorem 2.12 Let **F** be the natural filtration generated by $W(\cdot)$, $F(\cdot,0,0) \in L^1_{\mathbb{F}}(0,T;L^2(\Omega;H))$, and A be a self-adjoint, negative definite (unbounded linear) operator on H. Then, for any $\xi \in L^2_{\mathcal{F}_T}(\Omega;H)$, the equation (2.16) admits a unique mild solution $(y(\cdot),Y(\cdot)) \in (L^2_{\mathbb{F}}(\Omega;C([0,T];H)) \cap L^2_{\mathbb{F}}(0,T;D((-A)^{\frac{1}{2}}))) \times L^2_{\mathbb{F}}(0,T;H)$. Moreover,

$$|y(\cdot)|_{L_{\mathbb{F}}^{2}(\Omega;C([0,T];H))} + |y(\cdot)|_{L_{\mathbb{F}}^{2}(0,T;D((-A)^{\frac{1}{2}}))} + |Y(\cdot)|_{L_{\mathbb{F}}^{2}(0,T;H)}$$

$$\leq \mathcal{C}(|\xi|_{L_{\mathcal{F}_{T}}^{2}(\Omega;H)} + |F(\cdot,0,0)|_{L_{\mathbb{F}}^{1}(0,T;L^{2}(\Omega;H))}).$$
(2.19)

Similarly to (2.15), we introduce an approximating equation of (2.16) as follows:

$$\begin{cases}
dy^{\lambda}(t) = -\left[Ay^{\lambda}(t) + R(\lambda)F(t, y^{\lambda}(t), Y(t))\right]dt - R(\lambda)Y^{\lambda}(t)dW(t) & \text{in } (0, T], \\
y^{\lambda}(T) = R(\lambda)\xi \in D(A).
\end{cases} (2.20)$$

Similarly to Theorem 2.9, we have the following result.

Theorem 2.13 Assume that **F** is the natural filtration generated by $W(\cdot)$, and $F(\cdot,0,0) \in L^1_{\mathbb{F}}(0,T;L^2(\Omega;H))$. Then, for each $\xi \in L^2_{\mathcal{F}_T}(\Omega;H)$ and $\lambda \in \rho(A)$, the equation (2.20) admits a unique strong solution $(y^{\lambda}(\cdot),Y^{\lambda}(\cdot)) \in L^2_{\mathbb{F}}(\Omega;C([0,T];D(A))) \times L^2_{\mathbb{F}}(0,T;H)$. Moreover, as $\lambda \to \infty$, $(y^{\lambda}(\cdot),Y^{\lambda}(\cdot))$ converges to $(y(\cdot),Y(\cdot))$ (in $L^2_{\mathbb{F}}(\Omega;C([0,T];H)) \times L^2_{\mathbb{F}}(0,T;H)$), the mild solution to (2.16).

Note that, in Theorems 2.10 and 2.12–2.13, we need the filtration \mathbf{F} to be natural. For the general filtration, as we shall see later, we need to employ the stochastic transposition method (developed in [11, 12, 13]) to show the well-posedness of the equation (2.16).

3 Controllability of stochastic (ordinary) differential equations

In this section, we assume **F** the natural filtration generated by $W(\cdot)$.

We begin with the following controlled system governed by a deterministic linear ordinary differential equation:

$$\begin{cases}
\frac{dy(t)}{dt} = Ay(t) + Bu(t), & t > 0, \\
y(0) = y_0.
\end{cases}$$
(3.1)

In (3.1), $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ $(n, m \in \mathbb{N})$, $y(\cdot)$ is the state variable, $u(\cdot)$ is the control variable, \mathbb{R}^n and \mathbb{R}^m are respectively the state and control spaces.

Definition 3.1 The system (3.1) is called exactly controllable at time T if for any $y_0, y_T \in \mathbb{R}^n$, there is a control $u(\cdot) \in L^1(0,T;\mathbb{R}^m)$ such that the solution $y(\cdot)$ to (3.1) satisfies $y(T) = y_T$.

One has the following result:

Theorem 3.1 The system (3.1) is exactly controllable at time T if and only if the Kalman rank condition holds

$$rank [B, AB, \cdots, A^{n-1}B] = n.$$

Write $G_T = \int_0^T e^{At} B B^{\mathsf{T}} e^{A^{\mathsf{T}} t} dt$. Further, one can show the following result:

Theorem 3.2 If the system (3.1) is exactly controllable at time T, then $\det G_T \neq 0$. Moreover, for any $y_0, y_T \in \mathbb{R}^n$, the control

$$u^*(t) = -B^{\mathsf{T}} e^{A^{\mathsf{T}}(T-t)} G_T^{-1} (e^{AT} y_0 - y_T)$$

transfers y_0 to y_T at time T.

Remark 3.1 From Theorem 3.2, it is easy to see that, if (3.1) is exactly controllable at time T (by means of L^1 -(in time) controls), then the same controllability can be achieved by using analytic-(in time) controls. Actually the same can be said for the case that the control class $L^1(0,T;\mathbb{R}^m)$ in Definition 3.1 is replaced by $L^p(0,T;\mathbb{R}^m)$ for any $p \in [1,\infty]$. However, we shall see a completely different phenomenon ever in the simplest stochastic situation.

Now, let us consider the following controlled system governed by a stochastic linear ordinary differential equation:

$$\begin{cases} dy = (Ay + Bu)dt + (Cy + Du)dW(t) & \text{in } [0, T], \\ y(0) = y_0, \end{cases}$$
 (3.2)

where $C \in \mathbb{R}^{n \times n}$ and $D \in \mathbb{R}^{n \times m}$, $u(\cdot)$ is the control (valued in \mathbb{R}^m) and $x(\cdot)$ is the state (valued in \mathbb{R}^n).

Definition 3.2 The system (3.2) is called exactly controllable (at time T) if for any $y_0 \in \mathbb{R}^n$ and $y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$, there exists a control $u(\cdot) \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^m)$ such that the corresponding solution $y(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n))$ to (3.2) satisfies that $y(T) = y_T$.

Define a (deterministic) function $\eta(\cdot)$ on [0,T] by

$$\eta(t) = \begin{cases}
1, & \text{for } t \in \left[\left(1 - \frac{1}{2^{2i}} \right) T, \left(1 - \frac{1}{2^{2i+1}} \right) T \right), & i = 0, 1, 2, \dots, \\
-1, & \text{otherwise.}
\end{cases}$$
(3.3)

One can show that ([18]) there exists a constant $\beta > 0$ such that

$$\int_{t}^{T} |\eta(s) - c|^{2} ds \ge 4\beta(T - t), \quad \text{for any } (c, t) \in \mathbb{R} \times [0, T].$$
(3.4)

One has the following result, which provides a necessary condition for the exact controllability of (3.2). **Proposition 3.1** ([18]) If the system (3.2) is exactly controllable, then rank D = n.

Proof: We use the contradiction argument. Assume that the system (3.2) was exactly controllable for some matrix D with rankD < n. Then, we would find a vector $v \in \mathbb{R}^n$ with $|v|_{\mathbb{R}^n} = 1$ such that $v \cdot D = 0$.

Let $y_T = \int_0^T \eta(t) dW(t) v$ (recall (3.3) for $\eta(\cdot)$). Since (3.2) was exactly controllable, there would exist a control $u \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ such that

$$y_T = y_0 + \int_0^T [Ay(t) + Bu(t)]dt + \int_0^T [Cy(t) + Du(t)]dW(t),$$

which implies that

$$\int_{0}^{T} \eta(t)dW(t) = v \cdot y_{0} + \int_{0}^{T} v \cdot \left[Ay(t) + Bu(t) \right] dt + \int_{0}^{T} v \cdot Cy(t) dW(t).$$
 (3.5)

Hence,

$$\int_0^T \left[\eta(t) - v \cdot Cy(t) \right] dW(t) = v \cdot y_0 + \int_0^T v \cdot \left[Ay(t) + Bu(t) \right] dt.$$

Therefore,

$$\int_0^t \left[\eta(s) - v \cdot Cy(s) \right] dW(s)$$

$$= v \cdot y_0 + \int_0^t v \cdot \left[Ay(s) + Bu(s) \right] ds + \mathbb{E} \left(\int_t^T v \cdot \left[Ay(s) + Bu(s) \right] ds \middle| \mathcal{F}_t \right).$$

This gives that

$$\int_{t}^{T} \left[\eta(s) - v \cdot Cy(s) \right] dW(s)$$

$$= \int_{t}^{T} v \cdot \left[Ay(s) + Bu(s) \right] ds - \mathbb{E} \left(\int_{t}^{T} v \cdot \left[Ay(s) + Bu(s) \right] ds \middle| \mathcal{F}_{t} \right),$$

which implies that

$$\mathbb{E} \int_{t}^{T} \left| \eta(s) - v \cdot Cy(s) \right|^{2} ds$$

$$= \mathbb{E} \left[\int_{t}^{T} v \cdot \left[Ay(s) + Bu(s) \right] ds - \mathbb{E} \left(\int_{t}^{T} v \cdot \left[Ay(s) + Bu(s) \right] ds \middle| \mathcal{F}_{t} \right) \right]^{2}$$

$$\leq \mathbb{E} \left[\int_{t}^{T} v \cdot \left[Ay(s) + Bu(s) \right] ds \right]^{2}$$

$$\leq (T - t) \int_{t}^{T} \left| v \cdot \left[Ay(s) + Bu(s) \right] \right|^{2} ds.$$
(3.6)

On the other hand, by the inequality (3.4), we have that

$$\mathbb{E} \int_{t}^{T} \left| \eta(s) - v \cdot Cy(s) \right|^{2} ds$$

$$\geq \frac{1}{2} \mathbb{E} \int_{t}^{T} \left| \eta(s) - v \cdot Cy(T) \right|^{2} ds - \mathbb{E} \int_{t}^{T} \left| v \cdot Cy(T) - v \cdot Cy(s) \right|^{2} ds \qquad (3.7)$$

$$\geq 2\beta (T - t) - \mathbb{E} \int_{t}^{T} \left| v \cdot Cy(T) - v \cdot Cy(s) \right|^{2} ds.$$

By virtue of that $y(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n))$, there is a $\tilde{t} \in [0,T)$ such that

$$\mathbb{E}\big|v\cdot Cy(T)-v\cdot Cy(s)\big|\leq \beta, \quad \text{ for all } s\in [\tilde{t},T).$$

This, together with (3.7) implies that

$$\mathbb{E} \int_{t}^{T} \left| \eta(s) - v \cdot Cy(s) \right|^{2} ds \ge \beta(T - t), \text{ for all } t \in [\tilde{t}, T).$$
 (3.8)

From (3.6) and (3.8), we have that

$$\beta \le \int_{t}^{T} \left| v \cdot \left[Ay(s) + Bu(s) \right] \right|^{2} ds$$
, for all $t \in [\tilde{t}, T)$,

which leads to a contradiction.

Proposition 3.2 If the system (3.2) is exactly controllable at time T, then (A, B) fulfills the Kalman rank condition.

Proof: Let $\tilde{y} = \mathbb{E}y$, where y is a solution to (3.2) with some $y_0 \in \mathbb{R}^n$ and $u(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$. Then \tilde{y} solves

$$\begin{cases} \frac{d\tilde{y}}{dt} = A\tilde{y} + B\mathbb{E}u & \text{in } [0, T], \\ \tilde{y}(0) = y_0. \end{cases}$$
 (3.9)

Since (3.2) is exactly controllable, we see that (3.9) is exactly controllable. Hence, (A, B) fulfills the Kalman rank condition.

By means of Propositions 3.1–3.2, it follows that we should assume that $\operatorname{rank} D = n$ and (A, B) fulfills the Kalman rank condition if we expect the exact controllability of the system (3.2) in the sense of Definition 3.2.

Since rank D = n, it is easy to see that $n \leq m$, and we can find two matrices $K_1 \in \mathbb{R}^{m \times m}$ and $K_2 \in \mathbb{R}^{m \times n}$ such that $DK_1 = (I_n, 0)$ and that $DK_2 = -C$. Introducing a simple linear transformation

$$u = K_1 \left(\begin{array}{c} v_2 \\ v_1 \end{array} \right) + K_2 y,$$

where $v_1 \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^{m-n})$ and $v_2 \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$, we see that the system (3.2) is reduced to the following system

$$\begin{cases}
dy = (A_1 y + A_2 v_2 + B_1 v_1) dt + v_2 dW(t) & \text{in } [0, T], \\
y(0) = y_0,
\end{cases}$$
(3.10)

where

$$A_1 = A + BK_2, \ A_2 \in \mathbb{R}^{n \times n}, \ B_1 \in \mathbb{R}^{n \times (m-n)} \text{ and } A_2v_2 + B_1v_1 = BK_1 \begin{pmatrix} v_2 \\ v_1 \end{pmatrix}.$$

In order to deal with the exact controllability problem for (3.10), we consider the following controlled backward stochastic differential system:

$$\begin{cases} dy = (A_1y + A_2Y + B_1v)dt + YdW(t) & \text{in } [0, T], \\ y(T) = y_T, \end{cases}$$
(3.11)

where $y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$, $v \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^{m-n})$ is the control variable.

Definition 3.3 The system (3.11) is called exactly controllable (at time 0) if for any $y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$ and $y_0 \in \mathbb{R}^n$, there is a control $v \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^{n \times (m-n)})$ such that the corresponding solution $(y(\cdot), Y(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C([0, T]; \mathbb{R}^n)) \times L^2_{\mathbb{F}}(0, T; \mathbb{R}^n)$ to (3.11) satisfies $y(0) = y_0$.

It is easy to show the following result:

Proposition 3.3 The system (3.10) is exactly controllable at time T if and only if the system (3.11) is exactly controllable at time 0.

The dual equation of the system (3.11) is the following stochastic ordinary differential equation:

$$\begin{cases} dz = -A_1^{\top} z dt - A_2^{\top} z dW(t) & \text{in } [0, T], \\ z(0) = z_0 \in \mathbb{R}^n. \end{cases}$$
 (3.12)

Similar to Theorem 3.1, one can show the following result:

Theorem 3.3 The following statements are equivalent:

- 1) The system (3.11) is exactly controllable at time 0;
- 2) All solutions to (3.12) satisfy the following observability estimate:

$$|z_0|^2 \le \mathcal{C}\mathbb{E} \int_0^T |B_1^\top z(t)|^2 dt, \qquad \forall \ z_0 \in \mathbb{R}^n;$$
(3.13)

3) Solutions to (3.12) enjoy the following observability:

$$B_1^{\top} z(\cdot) \equiv 0 \text{ in } (0, T), \text{ a.s. } \Rightarrow z_0 = 0;$$
 (3.14)

4) The following rank condition holds:

$$rank [B_1, A_1B_1, A_2B_1, A_1^2B_1, A_1A_2B_1, A_2^2B, A_2A_1B_1, \cdots] = n.$$
(3.15)

Proof: By means of the classical duality argument, it is easy to show that "1) \iff 2)". The proof of "2) \iff 3)" is easy.

"4) \Longrightarrow 3)". We use an idea from the proof of [18, Theorem 3.2]. Let us assume that $B_1^{\mathsf{T}}z(\cdot)\equiv 0$ in (0,T), a.s. for some $z_0\in\mathbb{R}^n$. Then,

$$B_1^{\top} z(t) = B_1^{\top} z_0 + \int_0^t B_1^{\top} A_1^{\top} z(s) ds + \int_0^t B_1^{\top} A_2^{\top} z(s) dW(s) = 0, \qquad \forall \ t \in (0, T).$$

Therefore, we have that

$$B_1^{\mathsf{T}} z_0 = 0, \quad B_1^{\mathsf{T}} A_1^{\mathsf{T}} z \equiv 0, \quad B_1^{\mathsf{T}} A_2^{\mathsf{T}} z \equiv 0.$$
 (3.16)

Hence $B_1^{\top} A_1^{\top} z_0 = B_1^{\top} A_2^{\top} z_0 = 0$.

Noticing that $z(\cdot)$ solves (3.12), we have that

$$z(t) = z_0 + \int_0^t A_1^{\mathsf{T}} z(s) ds + \int_0^t A_2^{\mathsf{T}} z(s) dW(s).$$

This together with (3.16) implies that

$$B_1^{\top} A_1^{\top} z = B_1^{\top} A_1^{\top} z_0 + \int_0^t B_1^{\top} A_1^{\top} A_1^{\top} z(s) ds + \int_0^t B_1^{\top} A_1^{\top} A_2^{\top} z(s) dW(s) = 0,$$

and

$$B_1^{\top} A_2^{\top} z = B_1^{\top} A_2^{\top} z_0 + \int_0^t B_1^{\top} A_2^{\top} A_1^{\top} z(s) ds + \int_0^t B_1^{\top} A_2^{\top} A_2^{\top} z(s) dW(s) = 0,$$

which are equivalent to

$$B_1^{\mathsf{T}} A_1^{\mathsf{T}} A_1^{\mathsf{T}} z \equiv B_1^{\mathsf{T}} A_1^{\mathsf{T}} A_2^{\mathsf{T}} z \equiv B_1^{\mathsf{T}} A_2^{\mathsf{T}} A_1^{\mathsf{T}} z \equiv B_1^{\mathsf{T}} A_2^{\mathsf{T}} A_2^{\mathsf{T}} z \equiv 0,$$

and implies that $B_1^{\top} A_1^{\top} A_2^{\top} z_0 = B_1^{\top} A_1^{\top} A_1^{\top} z_0 = B_1^{\top} A_2^{\top} A_1^{\top} z_0 = B_1^{\top} A_2^{\top} A_2^{\top} z_0 = 0$.

Utilizing the above argument, by induction, we can conclude that

$$z_0^{\top}[B_1, A_1B_1, A_2B_1, A_1^2B_1, A_1A_2B_1, A_2^2B, A_2A_1B_1, \cdots] = 0.$$
 (3.17)

By (3.15) and (3.17), it follows that $z_0 = 0$.

"3) \Longrightarrow 4)". We use the contradiction argument. Assume that (3.15) was false. Then, we could find a nonzero $z_0 \in \mathbb{R}^n$ satisfying (3.17). For this z_0 , denote by $z(\cdot)$ the corresponding solution to (3.12). Clearly, $z(\cdot)$ can be approximated (in $L^2_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$) by the Picard sequence $\{z_k(\cdot)\}_{k=0}^{\infty}$ defined as follows

$$\begin{cases}
z_0(\cdot) = z_0, \\
z_k(\cdot) = z_0 + \int_0^{\cdot} A_1^{\top} z_{k-1}(s) ds + \int_0^{\cdot} A_2^{\top} z_{k-1}(s) dW(s), & k \in \mathbb{N}.
\end{cases}$$
(3.18)

By (3.17) and (3.18), via a direct computation, one can show that

$$B_1^{\mathsf{T}} z_k(\cdot) = 0, \qquad k = 0, 1, 2, \cdots.$$
 (3.19)

By (3.19), we deduce that $B_1^{\top}z(\cdot)\equiv 0$ in (0,T). Hence, by (3.14), it follows that $z_0=0$, which is a contradiction.

As a consequence of Theorem 3.3, we have the following characterization for the exact controllability of (3.10) (and hence also for that of (3.2)).

Corollary 3.1 ([18]) The system (3.10) is exactly controllable at time T if and only if the rank condition (3.15) holds.

In the above, we introduce two (different) controls v_1 and v_2 in the system (3.10), and both v_1 and v_2 are L^2 -(in time). Is it possible to introduce only one control or to use other class of controls?

We consider the simplest one-dimensional controlled "stochastic" differential equation as follows

$$\begin{cases} dy(t) = u(t)dt, \\ y(0) = y_0. \end{cases}$$
 (3.20)

We say that the system (3.20) is exactly controllable if for any $y_0 \in \mathbb{R}$ and $y_T \in L^2_{\mathcal{F}_T}(\Omega)$, there exists a control $u(\cdot) \in L^1_{\mathbb{F}}(0,T;L^2(\Omega))$ such that the corresponding solution $y(\cdot)$ satisfies $y(T) = y_T$.

It is showed in [10] that the system (3.20) is exactly controllable at any time T > 0 (by means of $L^1_{\mathbb{R}}(0,T;L^2(\Omega))$ -controls).

On the other hand, surprisingly, in virtue of Proposition 3.1, the system (3.20) is NOT exactly controllable if one is confined to use admissible controls $u(\cdot)$ in $L^2_{\mathbb{F}}(0,T;L^2(\Omega))!$ Further, the authors in [10] showed that the system (3.20) is NOT exactly controllable, either provided that one uses admissible controls $u(\cdot)$ in $L^p_{\mathbb{F}}(0,T;L^2(\Omega))$ for any $p \in (1,\infty]$.

To the best of our knowledge, unlike the deterministic case, there exists no universally accepted notion for stochastic controllability so far. Motivated by the above example, we introduced a corrected formulation for the exact controllability of stochastic differential equations.

Definition 3.4 The system (3.2) is called exactly controllable if for any $y_0 \in \mathbb{R}^n$ and $y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$, one can find a control $u(\cdot) \in L^1_{\mathbb{R}}(0, T; L^2(\Omega; \mathbb{R}^m))$ such that $Du(\cdot, \omega) \in L^2(0, T; \mathbb{R}^n)$, a.e. $\omega \in \Omega$, and the corresponding solution $y(\cdot)$ to (3.2) satisfies $y(T) = y_T$.

The above definition seems to be a reasonable notion for exact controllability of stochastic differential equations. Nevertheless, a complete study on this problem is still under consideration and it does not seem to be easy.

One may think that the requirement of exact controllability for (3.2) is too strong. How about the null/approximate controllability? Consider the following two weaker notions of controllability.

Definition 3.5 The system (3.2) is called null controllable (at time T) if for any $y_0 \in \mathbb{R}^n$, there exists a control $u(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ such that the corresponding solution $y(\cdot)$ to (3.2) satisfies y(T) = 0.

Definition 3.6 The system (3.2) is called approximately controllable (at time T) if for any $y_0 \in \mathbb{R}^n$, $y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$ and $\varepsilon > 0$, there exists a control $u(\cdot) \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^m)$ such that the corresponding solution $y(\cdot)$ to (3.2) satisfies $|y(T) - y_T|_{L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)} < \varepsilon$.

We shall show below that there exists no any rank condition for the null/approximate controllability of (3.2). In fact, if there is a such kind of rank condition, then it should has the following properties:

- It is robust with respect to perturbations small enough;
- The system (3.2) is null/approximately controllable at time T for any T > 0.

However, as pointed in [15], such properties cannot be held. In fact, consider the following 2-dimensional stochastic differential system:

$$\begin{cases} dy_1 = y_2 dt + \varepsilon y_2 dW(t) & \text{in } [0, T], \\ dy_2 = u dt & \text{in } [0, T], \\ y_1(0) = y_{10}, \ y_2(0) = y_{20}, \end{cases}$$
(3.21)

where $(y_{10}, y_{20}) \in \mathbb{R}^2$, $u(\cdot) \in L^1_{\mathbb{F}}(0, T; L^2(\Omega))$ is the control variable, ε is a parameter. Clearly, if $\varepsilon = 0$, then (3.21) is null controllable. If the above two properties held, then there would exist an $\varepsilon_0 > 0$ such that for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ and T > 0, (3.21) is null controllable. Let us take $y_{10} = 0$, $y_{20} = 1$, $\varepsilon = \varepsilon_0$ and $T = \frac{\varepsilon_0^2}{2}$. Since (3.21) is null controllable at $T = \frac{\varepsilon_0^2}{2}$, then

$$y_1\left(\frac{\varepsilon_0^2}{2}\right) = \int_0^{\frac{\varepsilon_0^2}{2}} y_2 dt + \varepsilon_0 \int_0^{\frac{\varepsilon_0^2}{2}} y_2 dW(t) = 0.$$

Thus,

$$\mathbb{E}\Big|\int_0^{\frac{\varepsilon_0^2}{2}} y_2 dt\Big|^2 = \mathbb{E}\Big|\varepsilon_0 \int_0^{\frac{\varepsilon_0^2}{2}} y_2 dW(t)\Big|^2 = \varepsilon_0^2 \int_0^{\frac{\varepsilon_0^2}{2}} \mathbb{E}|y_2|^2 dt. \tag{3.22}$$

On the other hand,

$$\mathbb{E} \Big| \int_{0}^{\frac{\varepsilon_{0}^{2}}{2}} y_{2} dt \Big|^{2} \leq \mathbb{E} \Big| \Big(\int_{0}^{\frac{\varepsilon_{0}^{2}}{2}} 1 dt \Big) \Big(\int_{0}^{\frac{\varepsilon_{0}^{2}}{2}} |y_{2}|^{2} dt \Big) \Big| \leq \frac{\varepsilon_{0}^{2}}{2} \int_{0}^{\frac{\varepsilon_{0}^{2}}{2}} \mathbb{E} |y_{2}|^{2} dt. \tag{3.23}$$

It follows from (3.22) and (3.23) that $\int_0^{\frac{\varepsilon_0^2}{2}} \mathbb{E}|y_2|^2 dt = 0$, which contradicts the choice of $y_2(0)$. Next, we consider the approximate controllability. For this purpose, we introduce the following backward stochastic differential equation:

$$\begin{cases}
dz_1 = Z_1 dW(t) & \text{in } [0, T], \\
dz_2 = -(z_1 + \varepsilon Z_1) dt + Z_2 dW(t) & \text{in } [0, T], \\
z_1(T) = z_{1T}, \ z_2(T) = z_{2T},
\end{cases}$$
(3.24)

where $(z_{1T}, z_{2T}) \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^2)$. By the classical duality argument, it is easy to show that the approximate controllability of (3.21) is equivalent to the following observability of (3.24): If $z_2(\cdot) = 0$, then $(z_1(\cdot), Z_1(\cdot), z_2(\cdot), Z_2(\cdot)) = (0, 0, 0, 0)$.

If $\varepsilon = 0$ and $z_2(\cdot) = 0$, then we

$$-\int_0^t z_1(s)ds + \int_0^t Z_2(s)dW(s) = 0, \quad \text{for all } t \in [0, T].$$

This, together with the uniqueness of the decomposition of semimartingale (See [21, page 358]), implies that $z_1(\cdot) = Z_2(\cdot) = 0$. Then, by the first equation in (3.24), we see that $Z_1(\cdot) = 0$. Therefore, we conclude that (3.21) is approximately controllable if $\varepsilon = 0$.

However, if $\varepsilon \neq 0$, then it is easy to check that

$$(z_1(t), Z_1(t), z_2(t), Z_2(t)) = \left(\exp\left\{-\frac{W(t)}{\varepsilon} - \frac{t}{2\varepsilon^2}\right\}, -\frac{1}{\varepsilon}\exp\left\{-\frac{W(t)}{\varepsilon} - \frac{t}{2\varepsilon^2}\right\}, 0, 0\right)$$

is a solution to (3.24) with $(z_{1T}, z_{2T}) = \left(\exp\left\{-\frac{W(T)}{\varepsilon} - \frac{T}{2\varepsilon^2}\right\}, -\frac{1}{\varepsilon}\exp\left\{-\frac{W(T)}{\varepsilon} - \frac{T}{2\varepsilon^2}\right\}\right)$. Hence, the above observability of (3.24) does not hold. Therefore, (3.21) is not approximately controllable whenever $\varepsilon \neq 0$.

Generally speaking, when n > 1, the controllability for the linear system (3.2) is far from well-understood. Actually, in our opinion, compared to the deterministic case, the controllability/observability for stochastic differential equations is at its enfant stage.

4 Pontryagin-type maximum principle for controlled stochastic (ordinary) differential equations

The first order necessary optimality condition, i.e., Pontryagin-type maximum principle, for optimal control problems for stochastic (ordinary) differential equations is by now well-understood (at least when there exist no endpoint constraints). When \mathbf{F} is the natural filtration, the general stochastic maximum principle was established in [18]. In this section, we do not assume that \mathbf{F} is the natural filtration. Thus, we cannot use the classical well-posedness theory of backward stochastic differential equations. A key point is that we need to use the stochastic transposition method, developed in [11].

Let U be a separable metric space with its metric $\mathbf{d}(\cdot,\cdot)$. Put

$$\mathcal{U}[0,T] \triangleq \Big\{ u(\cdot) : [0,T] \to U \mid u(\cdot) \text{ is } \mathbf{F}\text{-adapted} \Big\}.$$

We assume the following condition.

(A1) Suppose that $a(\cdot,\cdot,\cdot):[0,T]\times\mathbb{R}^n\times U\to\mathbb{R}^n$ and $b(\cdot,\cdot,\cdot):[0,T]\times\mathbb{R}^n\times U\to\mathbb{R}^n$ are two functions satisfying: i) For any $(x,u)\in\mathbb{R}^n\times U$, the functions $a(\cdot,x,u):[0,T]\to\mathbb{R}^n$ and $b(\cdot,x,u):[0,T]\to\mathbb{R}^n$ are Lebesgue measurable; ii) For any $(t,x)\in[0,T]\times\mathbb{R}^n$, the functions $a(t,x,\cdot):U\to\mathbb{R}^n$ and $b(t,x,\cdot):U\to\mathbb{R}^n$ are continuous; and iii) There is a constant $\mathcal{C}_L>0$ such that for all $(t,x_1,x_2,u)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times U$,

$$\begin{cases}
|a(t, x_1, u) - a(t, x_2, u)|_{\mathbb{R}^n} + |b(t, x_1, u) - b(t, x_2, u)|_{\mathbb{R}^n} \le \mathcal{C}_L |x_1 - x_2|_{\mathbb{R}^n}, \\
|a(t, 0, u)|_{\mathbb{R}^n} + |b(t, 0, u)|_{\mathbb{R}^n} \le \mathcal{C}_L.
\end{cases}$$
(4.1)

Let us consider the following controlled stochastic differential equation:

$$\begin{cases} dx = a(t, x, u)dt + b(t, x, u)dW(t) & \text{in } [0, T], \\ x(0) = x_0, \end{cases}$$
(4.2)

where $u \in \mathcal{U}[0,T]$ and $x_0 \in L^p_{\mathcal{F}_0}(\Omega;\mathbb{R}^n)$ for a given $p \geq 2$. Under the assumption (A1), it is easy to show that the equation (4.2) is well-posed in the sense of adapted solutions in the space $L^p_{\mathbb{F}}(\Omega; C([0,T];\mathbb{R}^n))$.

Also, we need the following condition:

(A2) Suppose that $g(\cdot,\cdot,\cdot):[0,T]\times\mathbb{R}^n\times U\to\mathbb{R}$ and $h(\cdot):\mathbb{R}^n\to\mathbb{R}$ are two functions satisfying: i) For any $(x,u)\in\mathbb{R}^n\times U$, the function $g(\cdot,x,u):[0,T]\to\mathbb{R}$ is Lebesgue measurable; ii) For any $(t,x)\in[0,T]\times\mathbb{R}^n$, the function $g(t,x,\cdot):U\to\mathbb{R}$ is continuous; and iii) For all $(t,x_1,x_2,u)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times U$,

$$\begin{cases}
|g(t, x_1, u) - g(t, x_2, u)|_{\mathbb{R}^n} + |h(x_1) - h(x_2)|_{\mathbb{R}^n} \le \mathcal{C}_L |x_1 - x_2|_{\mathbb{R}^n}, \\
|g(t, 0, u)|_{\mathbb{R}^n} + |h(0)|_{\mathbb{R}^n} \le \mathcal{C}_L.
\end{cases}$$
(4.3)

Define a cost functional $\mathcal{J}(\cdot)$ (for the controlled system (4.2)) as follows:

$$\mathcal{J}(u(\cdot)) \triangleq \mathbb{E}\Big[\int_0^T g(t, x(t), u(t)) dt + h(x(T))\Big], \quad \forall u(\cdot) \in \mathcal{U}[0, T], \tag{4.4}$$

where $x(\cdot)$ is the corresponding solution to (4.2).

Let us consider the following optimal control problem for the system (4.2):

Problem (OPF) Find a $\bar{u}(\cdot) \in \mathcal{U}[0,T]$ such that

$$\mathcal{J}(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0,T]} \mathcal{J}(u(\cdot)). \tag{4.5}$$

Any $\bar{u}(\cdot)$ satisfying (4.5) is called an optimal control. The corresponding state process $\bar{x}(\cdot)$ is called an optimal state (process), and $(\bar{x}(\cdot), \bar{u}(\cdot))$ is called an optimal pair.

Furthermore, we impose the following assumption.

(A3) The functions a(t, x, u), b(t, x, u), g(t, x, u) and h(x) are C^2 in x, and for $\varphi(t, x, u) = b(t, x, u)$, $\sigma(t, x, u)$, f(t, x, u), h(x) and any $t \in [0, T]$, $x, \widehat{x} \in \mathbb{R}^n$ and $u, \widehat{u} \in U$, it holds that

$$\begin{cases}
|\varphi(t, x, u) - \varphi(t, \widehat{x}, \widehat{u})| \leq C_L (|x - \widehat{x}| + \mathbf{d}(u, \widehat{u})), \\
|\varphi(t, 0, u)| \leq C_L, \\
|\varphi_x(t, x, u) - \varphi_x(t, \widehat{x}, \widehat{u})| \leq C_L (|x - \widehat{x}| + \mathbf{d}(u, \widehat{u})), \\
|\varphi_{xx}(t, x, u) - \varphi_{xx}(t, \widehat{x}, \widehat{u})| \leq C_L (|x - \widehat{x}| + \mathbf{d}(u, \widehat{u})).
\end{cases}$$

Suppose that $(\bar{x}(\cdot), \bar{u}(\cdot))$ is a given optimal pair. Similar to the corresponding deterministic setting, one introduces the following first order adjoint equation (which is however a backward stochastic differential equation in the stochastic case):

$$\begin{cases}
dy(t) = -\left[a_x(t, \bar{x}(t), \bar{u}(t))^{\top} y(t) + b_x(t, \bar{x}(t), \bar{u}(t))^{\top} Y(t) - g_x(t, \bar{x}(t), \bar{u}(t))\right] dt \\
+ Y(t) dW(t) & \text{in } [0, T], \\
y(T) = -h_x(\bar{x}(T)).
\end{cases} (4.6)$$

Next, to establish the desired maximum principle for stochastic controlled systems with control-dependent diffusion and possibly nonconvex control domains, one has to introduce an additional second order adjoint equation as follows:

$$\begin{cases}
dP(t) = -\left[a_{x}(t, \bar{x}(t), \bar{u}(t))^{\top} P(t) + P(t) a_{x}(t, \bar{x}(t), \bar{u}(t)) + b_{x}(t, \bar{x}(t), \bar{u}(t))^{\top} P(t) b_{x}(t, \bar{x}(t), \bar{u}(t)) \\
+ b_{x}(t, \bar{x}(t), \bar{u}(t))^{\top} Q(t) + Q(t) b_{x}(t, \bar{x}(t), \bar{u}(t)) + \mathbb{H}_{xx}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t))\right] dt \\
+ Q(t) dW(t) & \text{in } [0, T), \\
P(T) = -h_{xx}(\bar{x}(T)).
\end{cases}$$
(4.7)

In (4.7), the *Hamiltonian* $\mathbb{H}(\cdot,\cdot,\cdot,\cdot,\cdot)$ is defined by

$$\mathbb{H}(t, x, u, y_1, y_2) = \langle y_1, a(t, x, u) \rangle_{\mathbb{R}^n} + \langle y_2, b(t, x, u) \rangle_{\mathbb{R}^n} - g(t, x, u),$$
$$(t, x, u, y_1, y_2) \in [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^n.$$

Since we do not assume that **F** is the natural filtration, the equations (4.6)/(4.7) may not have classical adapted solutions. We need to introduce below the notion of transposition solutions to the following backward stochastic differential equation:

$$\begin{cases} dy(t) = f(t, y(t), Y(t))dt + Y(t)dW(t) & \text{in } [0, T], \\ y(T) = y_T, \end{cases}$$
where $y_T \in L^p_{\mathcal{F}_T}(\Omega; \mathbb{R}^n), f(\cdot, \cdot, \cdot)$ satisfies $f(\cdot, 0, 0) \in L^p_{\mathbb{F}}(\Omega; L^1(0, T; \mathbb{R}^n)),$ and

$$|f(t, p_1, q_1) - f(t, p_2, q_2)|_{\mathbb{R}^n} \le \mathcal{C}_L(|p_1 - p_2|_{\mathbb{R}^n} + |q_1 - q_2|_{\mathbb{R}^n}),$$

$$t \in [0, T], \ \mathbb{P}\text{-a.s.}, \forall \ p_1, p_2, q_1, q_2 \in \mathbb{R}^n.$$

$$(4.9)$$

In order to define the transposition solution to (4.8), for any $t \in [0, T]$, we consider the following linear stochastic differential equation

$$\begin{cases}
dz(\tau) = u(\tau)d\tau + v(\tau)dW(\tau), & \tau \in (t, T], \\
z(t) = \eta.
\end{cases} (4.10)$$

For any given $u(\cdot) \in L^1_{\mathbb{F}}(t,T;L^q(\Omega;\mathbb{R}^n)), \ v(\cdot) \in L^q_{\mathbb{F}}(\Omega;L^2(t,T;\mathbb{R}^n))$ and $\eta \in L^q_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$, the equation (4.10) admits a unique adapted solution $z(\cdot) \in L^q_{\mathbb{R}}(\Omega; C([t,T];\mathbb{R}^n))$. Now, if the equation (4.8) admits an adapted solution $(y(\cdot), Y(\cdot)) \in L^p_{\mathbb{F}}(\Omega; C([0,T]; \mathbb{R}^n)) \times L^p_{\mathbb{F}}(0,T; L^2(\Omega; \mathbb{R}^n))$ \mathbb{R}^n), then, applying Itô's formula to $\langle z(t), y(t) \rangle_{\mathbb{R}^n}$, it is easy to check that

$$\mathbb{E} \langle z(T), y_T \rangle_{\mathbb{R}^n} - \mathbb{E} \langle \eta, y(t) \rangle_{\mathbb{R}^n}$$

$$= \mathbb{E} \int_t^T \langle z(\tau), f(\tau, y(\tau), Y(\tau)) \rangle_{\mathbb{R}^n} d\tau + \mathbb{E} \int_t^T \langle u(\tau), y(\tau) \rangle_{\mathbb{R}^n} d\tau + \mathbb{E} \int_t^T \langle v(\tau), Y(\tau) \rangle_{\mathbb{R}^n} d\tau.$$
(4.11)

This inspires us to introduce the following new notion of solution to the equation (4.8).

Definition 4.1 We call $(y(\cdot), Y(\cdot)) \in D_{\mathbb{F}}([0, T]; L^p(\Omega; \mathbb{R}^n)) \times L^p_{\mathbb{F}}(\Omega; L^2(0, T; \mathbb{R}^n))$ a transposition solution to (4.8) if the identity (4.11) holds for any $t \in [0, T]$, $u(\cdot) \in L^1_{\mathbb{F}}(t, T; L^q(\Omega; \mathbb{R}^n))$, $v(\cdot) \in L^q_{\mathbb{F}}(\Omega; L^2(t, T; \mathbb{R}^n))$ and $\eta \in L^q_{\mathcal{F}_t}(\Omega; \mathbb{R}^n)$.

We have the following well-posedness result for (4.8) in the sense of transposition solution.

Theorem 4.1 ([11]) For any given $y_T \in L^p_{\mathcal{F}_T}(\Omega; \mathbb{R}^m)$, the equation (4.8) admits a unique transposition solution $(y(\cdot), Y(\cdot)) \in D_{\mathbb{F}}([0, T]; L^p(\Omega; \mathbb{R}^m)) \times L^p_{\mathbb{F}}(\Omega; L^2(0, T; \mathbb{R}^m))$. Furthermore,

$$|(y(\cdot), Y(\cdot))|_{D_{\mathbb{F}}([0,T];L^{p}(\Omega;\mathbb{R}^{m}))\times L_{\mathbb{F}}^{p}(\Omega;L^{2}(0,T;\mathbb{R}^{m}))}$$

$$\leq C\left[|f(\cdot,0,0)|_{L_{\mathbb{F}}^{p}(\Omega;L^{1}(0,T;\mathbb{R}^{m}))} + |y_{T}|_{L_{\mathcal{F}_{T}}^{p}(\Omega;\mathbb{R}^{m})}\right].$$

$$(4.12)$$

By means of the transposition solutions $(y(\cdot), Y(\cdot))$ and $(P(\cdot), Q(\cdot))$ respectively to (4.6) and (4.7) (guaranteed by Theorem 4.1), we can establish the following Pontryagin-type maximum principle for Problem (OPF).

Theorem 4.2 ([11]) Let (A1)–(A3) hold and $x_0 \in \mathbb{R}^n$. Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be an optimal pair of Problem (OPF). Then

$$\mathbb{H}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)) - \mathbb{H}(t, \bar{x}(t), u, y(t), Y(t))$$

$$-\frac{1}{2} \langle P(t) \left[b(t, \bar{x}(t), \bar{u}(t)) - b(t, \bar{x}(t), u) \right], b(t, \bar{x}(t), \bar{u}(t)) - b(t, \bar{x}(t), u) \rangle_{\mathbb{R}^{n}}$$

$$\geq 0, \qquad \forall u \in U, \quad \text{a.e. } t \in [0, T], \quad \mathbb{P}\text{-a.s.}$$

$$(4.13)$$

Sketch of the proof of Theorem 4.2: Since the detailed proof of Theorem 4.2 is too lengthy, we shall give below only a sketch to show some key points for establishing the stochastic maximum principle.

Fix any $u(\cdot) \in \mathcal{U}[0,T]$ and $\varepsilon > 0$, let

$$u^{\varepsilon}(t) = \begin{cases} \bar{u}(t), & t \in [0, T] \setminus E_{\varepsilon}, \\ u(t), & t \in E_{\varepsilon}, \end{cases}$$

where $E_{\varepsilon} \subseteq [0,T]$ is a measurable set with Lebesgue measure $|E_{\varepsilon}| = \varepsilon$. For $\varphi = a, b$ and f, we set

$$\begin{cases}
\varphi_x(t) = \varphi_x(t, \bar{x}(t), \bar{u}(t)), & \varphi_{xx}(t) = \varphi_{xx}(t, \bar{x}(t), \bar{u}(t)), \\
\delta\varphi(t) = \varphi(t, \bar{x}(t), u(t)) - \varphi(t, \bar{x}(t), \bar{u}(t)).
\end{cases} (4.14)$$

Let $x_1^{\varepsilon}(\cdot)$ and $x_2^{\varepsilon}(\cdot)$ solve respectively the following stochastic differential equations

$$\begin{cases}
 dx_1^{\varepsilon}(t) = a_x(t)x_1^{\varepsilon}(t)dt + \left[b_x(t)x_1^{\varepsilon}(t) + \chi_{E_{\varepsilon}}(t)\delta b(t)\right]dW(t) & \text{in } [0, T], \\
 x_1^{\varepsilon}(0) = 0,
\end{cases}$$
(4.15)

and

$$\begin{cases}
dx_2^{\varepsilon}(t) = \left[a_x(t) x_2^{\varepsilon}(t) + \chi_{E_{\varepsilon}}(t) \delta a(t) + \frac{1}{2} a_{xx}(t) \left(x_1^{\varepsilon}(t), x_1^{\varepsilon}(t) \right) \right] dt \\
+ \left[b_x(t) x_2^{\varepsilon}(t) + \chi_{E_{\varepsilon}}(t) \delta b_x(t) x_1^{\varepsilon}(t) + \frac{1}{2} b_{xx}(t) \left(x_1^{\varepsilon}(t), x_1^{\varepsilon}(t) \right) \right] dW(t) & \text{in } [0, T], \\
x_2^{\varepsilon}(0) = 0.
\end{cases}$$
(4.16)

Then, by some lengthy but direct computations, one can obtain that

$$\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\overline{u}(\cdot)) \\
= \mathbb{E} \left\langle h_{x}(\overline{x}(T)), x_{1}^{\varepsilon}(T) + x_{2}^{\varepsilon}(T) \right\rangle_{\mathbb{R}^{n}} + \frac{1}{2} \mathbb{E} \left\langle h_{xx}(\overline{x}(T)) x_{1}^{\varepsilon}(T), x_{1}^{\varepsilon}(T) \right\rangle_{\mathbb{R}^{n}} \\
+ \mathbb{E} \int_{0}^{T} \left\{ \left\langle g_{x}(t), x_{1}^{\varepsilon}(t) + x_{2}^{\varepsilon}(t) \right\rangle_{\mathbb{R}^{n}} + \frac{1}{2} \left\langle g_{xx}(t) x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) \right\rangle_{\mathbb{R}^{n}} + \chi_{E_{\varepsilon}}(t) \delta g(t) \right\} dt + o(\varepsilon).$$
(4.17)

By means of the fact that $(y(\cdot), Y(\cdot))$ is the transposition solution to the equation (4.6) with p = 2, we find that

$$-\mathbb{E}\langle h_x(\bar{x}(T)), x_1^{\varepsilon}(T)\rangle_{\mathbb{R}^n} = \mathbb{E}\int_0^T \left[\langle g_x(t), x_1^{\varepsilon}(t)\rangle_{\mathbb{R}^n} + \chi_{E_{\varepsilon}}(t)\langle \delta b(t), Y(t)\rangle_{\mathbb{R}^n}\right] dt, \tag{4.18}$$

and

$$-\mathbb{E}\langle h_{x}(\bar{x}(T)), x_{2}^{\varepsilon}(T)\rangle_{\mathbb{R}^{n}}$$

$$= \mathbb{E}\int_{0}^{T} \left\{ \langle g_{x}(t), x_{2}^{\varepsilon}(t)\rangle_{\mathbb{R}^{n}} + \frac{1}{2} \left[\langle y(t), a_{xx}(t) \left(x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) \right) \rangle_{\mathbb{R}^{n}} + \langle Y(t), b_{xx}(t) \left(x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) \right) \rangle_{\mathbb{R}^{n}} \right] + \chi_{E_{\varepsilon}}(t) \left[\langle y(t), \delta a(t) \rangle_{\mathbb{R}^{n}} + \langle Y(t), \delta b_{x}(t) x_{1}^{\varepsilon}(t) \rangle_{\mathbb{R}^{n}} \right] \right\} dt.$$

$$(4.19)$$

Further, put $x_3^{\varepsilon}(t) = x_1^{\varepsilon}(t)x_1^{\varepsilon}(t)^{\top} (\in \mathbb{R}^{n \times n})$. A direct computation shows that $x_3^{\varepsilon}(\cdot)$ solves

$$\begin{cases}
dx_3^{\varepsilon}(t) = \left\{ a_x(t) x_3^{\varepsilon}(t) + x_3^{\varepsilon}(t) a_x(t)^{\top} + b_x(t) x_3^{\varepsilon}(t) b_x(t)^{\top} + \chi_{E_{\varepsilon}}(t) \delta b(t) \delta b(t)^{\top} \\
+ \chi_{E_{\varepsilon}}(t) \left[b_x(t) x_1^{\varepsilon}(t) \delta b(t)^{\top} + \delta b(t) x_1^{\varepsilon}(t)^{\top} b_x(t)^{\top} \right] \right\} dt \\
+ \left[b_x(t) x_3^{\varepsilon}(t) + x_3^{\varepsilon}(t) b_x(t)^{\top} + \chi_{E_{\varepsilon}}(t) \left(\delta b(t) x_1^{\varepsilon}(t)^{\top} + x_1^{\varepsilon}(t) \delta b(t)^{\top} \right) \right] dW(t) & \text{in } (0, T], \\
x_3^{\varepsilon}(0) = 0.
\end{cases}$$
(4.20)

Utilizing the fact that $(P(\cdot), Q(\cdot))$ is the transposition solution to the equation (4.7) with p = 4, and noting that the inner product defined in $\mathbb{R}^{n \times n}$ is $\operatorname{tr}(P_1 P_2^{\mathsf{T}})$ for $P_1, P_2 \in \mathbb{R}^{n \times n}$, we find that

$$-\mathbb{E}\operatorname{tr}\left[h_{xx}(\bar{x}(T))x_3^{\varepsilon}(T)\right]$$

$$=\mathbb{E}\int_0^T\operatorname{tr}\left[\chi_{E_{\varepsilon}}(t)\delta b(t)^{\top}P(t)\delta b(t)-\mathbb{H}_{xx}(t,\bar{x}(t),\bar{u}(t),y(t),Y(t))x_3^{\varepsilon}(t)\right]dt+o(\varepsilon),$$

which gives that

$$-\mathbb{E}\langle h_{xx}(\bar{x}(T))x_{1}^{\varepsilon}(T), x_{1}^{\varepsilon}(T)\rangle_{\mathbb{R}^{n}}$$

$$= \mathbb{E}\int_{0}^{T} \left[\chi_{E_{\varepsilon}}(t)\langle P(t)\delta b(t), \delta b(t)\rangle_{\mathbb{R}^{n}} - \langle \mathbb{H}_{xx}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t))x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t)\rangle_{\mathbb{R}^{n}}\right] dt + o(\varepsilon).$$
(4.21)

From (4.17)–(4.21), we obtain that

$$\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \\
= \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \Big\{ \Big[\mathbb{H}(t, \bar{x}(t), u(t), y(t), Y(t)) - \mathbb{H}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)) \Big] \\
- \frac{1}{2} \Big\langle P(t) \Big[b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \Big], b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \Big\rangle_{\mathbb{R}^{n}} \Big\} dt + o(\varepsilon).$$
(4.22)

Since $\bar{u}(\cdot)$ is the optimal control, we have $\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \geq 0$. This, together with (4.22), yields that

$$\mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \Big\{ \Big[\mathbb{H}(t, \bar{x}(t), u(t), y(t), Y(t)) - \mathbb{H}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)) \Big] \\
- \frac{1}{2} \Big\langle P(t) \big[b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \big], b(t, \bar{x}(t), u(t)) - b(t, \bar{x}(t), \bar{u}(t)) \Big\rangle_{\mathbb{R}^{n}} \Big\} dt \\
\geq o(\varepsilon), \tag{4.23}$$

which leads to (4.13).

For some optimal controls, it may happen that the first-order necessary conditions turn out to be trivial. When an optimal control is singular, the first-order necessary condition cannot provide enough information for the theoretical analysis and numerical computation, and therefore one needs to study the second order necessary conditions. Quite different from the deterministic setting, there exist some essential difficulties in deriving the pointwise second-order necessary condition from an integral-type one when the diffusion term of the control system contains the control variable, even for the case of convex control constraint (see the first four paragraphs of subsection 3.2 in [24] for a detailed explanation). In [24, 25], these difficulties were overcome by means of some technique from the Malliavin calculus, and some pointwise second-order necessary conditions for stochastic optimal controls were established, even for the general case when the control region is nonconvex but the full picture is still quite unclear (see [2] for some recent progresses).

5 Controllability of stochastic differential equations in infinite dimensions: An analysis of a typical equation

This section is devoted to studying the controllability of stochastic differential equations in infinite dimensions. Since the stochastic controllability problem is even less understood in finite dimensions, we shall concentrate only on a typical equation, i.e., a stochastic parabolic system. Our main results can be described as follows:

- When the coefficients of the underlying system are space-independent, using the spectral method, we show the null/approximate controllability using only one control applied to the drift term;
- The null/approximate controllability of general stochastic parabolic systems with two controls are shown by means of duality argument.

In each of the above cases, we shall explain the main differences between the deterministic problem and its stochastic counterpart.

5.1 Formulation of the problem

Throughout this section, we assume that \mathbf{F} is the natural filtration generated by $W(\cdot)$, $G \subset \mathbb{R}^n$ $(n \in \mathbb{N})$ is a given bounded domain with a C^{∞} boundary Γ , and G_0 is a given nonempty open subset of G. Denote by χ_{G_0} the characteristic function of G_0 in G. Put $Q \stackrel{\triangle}{=} (0,T) \times G$, $\Sigma \stackrel{\triangle}{=} (0,T) \times \Gamma$ and $Q_0 \stackrel{\triangle}{=} (0,T) \times G_0$. Also, we assume that $a^{jk} : \overline{G} \to \mathbb{R}^{n \times n}$ $(j,k=1,2,\cdots,n)$ satisfies $a^{jk} \in C^1(\overline{G})$, $a^{jk} = a^{kj}$, and for some $s_0 > 0$,

$$\sum_{j,k=1}^{n} a^{jk}(x)\xi_{j}\xi_{k} \ge s_{0}|\xi|^{2}, \quad \forall (x,\xi) \equiv (x_{1},\cdots,x_{n},\xi_{1},\cdots,\xi_{n}) \in G \times \mathbb{R}^{n}.$$
 (5.1)

Let us fix an $m \in \mathbb{N}$ and consider the following controlled stochastic parabolic system:

$$\begin{cases} dy - \sum_{j,k=1}^{n} (a^{jk}y_{x_j})_{x_k} dt = \left(\sum_{j=1}^{n} a_{1j}y_{x_j} + a_2y + \chi_{G_0}u\right) dt + (a_3y + v) dW(t) & \text{in } Q, \\ y = 0 & \text{on } \Sigma, \\ y(0) = y_0 & \text{in } G, \end{cases}$$
(5.2)

where

$$\begin{cases}
 a_{1j} \in L_{\mathbb{F}}^{\infty}(0, T; W^{1,\infty}(G; \mathbb{R}^{m \times m})), & j = 1, 2, \cdots, n, \\
 a_{2} \in L_{\mathbb{F}}^{\infty}(0, T; L^{\infty}(G; \mathbb{R}^{m \times m})), & a_{3} \in L_{\mathbb{F}}^{\infty}(0, T; L^{\infty}(G; \mathbb{R}^{m \times m})).
\end{cases}$$
(5.3)

In the system (5.2), the initial state $y_0 \in L^2_{\mathcal{F}_0}(\Omega; L^2(G; \mathbb{R}^m)), y$ is the state variable, and the control variable consists of a pair $(u, v) \in L^2_{\mathbb{F}}(0, T; L^2(G_0; \mathbb{R}^m)) \times L^2_{\mathbb{F}}(0, T; L^2(G; \mathbb{R}^m)).$

We first recall the following well-posedness result for the equation (5.2). The proof can be found in [1, Chapter 6] and [14, Chapter 3].

Lemma 5.1 Let $a_{1j} \in L^{\infty}_{\mathbb{F}}(0,T;L^{\infty}(G;\mathbb{R}^{m\times m}))$ for $j=1,2,\cdots,n$, and a_2 and a_3 be given as in (5.3). Then, for any $y_0 \in L^2(G;\mathbb{R}^m)$ and $(u,v) \in L^2_{\mathbb{F}}(0,T;L^2(G_0;\mathbb{R}^m)) \times L^2_{\mathbb{F}}(0,T;L^2(G;\mathbb{R}^m))$, the system (5.2) admits a unique weak solution $y \in L^2_{\mathbb{F}}(\Omega;C([0,T];L^2(G;\mathbb{R}^m))) \cap L^2_{\mathbb{F}}(0,T;H^1_0(G;\mathbb{R}^m))$. Moreover,

$$|y|_{L_{\mathbb{F}}^{2}(\Omega;C([0,T];L^{2}(G;\mathbb{R}^{m})))\cap L_{\mathbb{F}}^{2}(0,T;H_{0}^{1}(G;\mathbb{R}^{m}))} \leq C(|y_{0}|_{L^{2}(G;\mathbb{R}^{m})} + |(u,v)|_{L_{\mathbb{F}}^{2}(0,T;L^{2}(G_{0};\mathbb{R}^{m}))\times L_{\mathbb{F}}^{2}(0,T;L^{2}(G;\mathbb{R}^{m}))}).$$

$$(5.4)$$

Definition 5.1 The system (5.2) is said to be null controllable if for any $y_0 \in L^2(G; \mathbb{R}^m)$, there exists a pair of $(u, v) \in L^2_{\mathbb{F}}(0, T; L^2(G_0; \mathbb{R}^m)) \times L^2_{\mathbb{F}}(0, T; L^2(G; \mathbb{R}^m))$ such that the corresponding solution to (5.2) fulfills that y(T) = 0, \mathbb{P} -a.s.

Note that we introduce two controls u and v in (5.2). In view of the controllability result for the deterministic parabolic equation, it is more natural to use only one control and

consider the following controlled stochastic parabolic system (which is a special case of (5.2) with $v \equiv 0$):

$$\begin{cases} dy - \sum_{j,k=1}^{n} (a^{jk}y_{x_j})_{x_k} dt = \left(\sum_{j=1}^{n} a_{1j}y_{x_j} + a_2y + \chi_{G_0}u\right) dt + a_3y dW(t) & \text{in } Q, \\ y = 0 & \text{on } \Sigma, \\ y(0) = y_0 & \text{in } G. \end{cases}$$
(5.5)

It is easy to see that, the dual system of both (5.2) and (5.5) is the following backward stochastic parabolic system:

$$\begin{cases}
dz + \sum_{j,k=1}^{n} (a^{jk} z_{x_j})_{x_k} dt = \left[\sum_{j=1}^{n} (a_{1j}^{\top} z)_{x_j} - a_2^{\top} z - a_3^{\top} Z \right] dt + Z dW(t) & \text{in } Q, \\
z = 0 & \text{on } \Sigma, \\
z(T) = z_T & \text{in } G.
\end{cases} (5.6)$$

We have the following well-posedness result for the equation (5.6) (See [14, Chapter 4] for example).

Proposition 5.1 Under the condition (5.3), for any $z_T \in L^2_{\mathcal{F}_T}(\Omega; L^2(G; \mathbb{R}^m))$, the system (5.6) admits one and only one weak solution $(z, Z) \in (L^2_{\mathbb{F}}(\Omega; C([0, T]; L^2(G; \mathbb{R}^m))) \cap L^2_{\mathbb{F}}(0, T; H^1_0(G; \mathbb{R}^m))) \times L^2_{\mathbb{F}}(0, T; L^2(G; \mathbb{R}^m))$. Moreover, for any $t \in [0, T]$,

$$|(z(\cdot), Z(\cdot))|_{\left(L_{\mathbb{F}}^{2}(\Omega; C([0,t]; L^{2}(G; \mathbb{R}^{m}))) \cap L_{\mathbb{F}}^{2}(0,t; H_{0}^{1}(G; \mathbb{R}^{m}))\right) \times L_{\mathbb{F}}^{2}(0,t; L^{2}(G; \mathbb{R}^{m}))} \leq C|z(t)|_{L_{\mathcal{F}_{t}}^{2}(\Omega; L^{2}(G; \mathbb{R}^{m}))}.$$
(5.7)

In order to obtain the null controllability of (5.5), we need to prove that solutions to the system (5.6) satisfy the following observability estimate:

$$|z(0)|_{L^{2}_{\mathcal{F}_{0}}(\Omega; L^{2}(G; \mathbb{R}^{m}))} \leq \mathcal{C}|z|_{L^{2}_{\mathbb{F}}(0, T; L^{2}(G_{0}; \mathbb{R}^{m}))}, \quad \forall \ z_{T} \in L^{2}_{\mathcal{F}_{T}}(\Omega; L^{2}(G; \mathbb{R}^{m})).$$
 (5.8)

Unfortunately, at this moment, we are not able to prove the observability estimate (5.8) for the general case. Instead, we obtain a weak version of (5.8), i.e., a weak observability estimate (for the system (5.6)) in Theorem 5.4 (See Subsection 5.3). By duality, Theorem 5.4 implies the null controllability of (5.2).

There exists a main difficulty to establish (5.8), that is, though the correction term "Z" plays a "coercive" role for the well-posedness of (5.6), it seems to be a "bad" (non-homogeneous) term when one tries to prove (5.8) using the global Carleman estimate.

Nevertheless, based on the spectral method, for some special case, we are able to show the controllability of (5.5).

5.2 Controllability of a class of stochastic parabolic systems

In this subsection, we show that when the coefficients of the stochastic parabolic system are space-independent, it is null/approximately controllable using only one control applied to the drift term. These results were first proved in [8].

We consider the following stochastic parabolic system:

$$\begin{cases} dy - \sum_{j,k=1}^{n} (a^{jk}y_{x_j})_{x_k} dt = [a(t)y + \chi_E(t)\chi_{G_0}(x)u]dt + b(t)ydW(t) & \text{in } Q, \\ y = 0 & \text{on } \Sigma, \\ y(0) = y_0 & \text{in } G, \end{cases}$$
(5.9)

where $a(\cdot) \in L^{\infty}_{\mathbb{F}}(0,T;\mathbb{R}^{m\times m})$ and $b(\cdot) \in L^{\infty}_{\mathbb{F}}(0,T;\mathbb{R}^{m\times m})$ are given, E is a fixed Lebesgue measurable subset in (0,T) with a positive Lebesgue measure (i.e., $\mathbf{m}(E) > 0$). In (5.9), y is the state variable (valued in $L^2(G;\mathbb{R}^m)$), $y_0 \in L^2(G;\mathbb{R}^m)$ is the initial state, u is the control variable, and the control space is $L^{\infty}_{\mathbb{F}}(0,T;L^2(\Omega;L^2(G;\mathbb{R}^m)))$.

Definition 5.2 The system (5.9) is said to be null controllable at time T if for any $y_0 \in L^2(G; \mathbb{R}^m)$, there exists a $u \in L^\infty_{\mathbb{F}}(0, T; L^2(\Omega; L^2(G; \mathbb{R}^m)))$ such that the corresponding solution to (5.9) fulfills that y(T) = 0, \mathbb{P} -a.s.

We have the following null controllability result for the system (5.9).

Theorem 5.1 The system (5.9) is null controllable at time T.

Remark 5.1 When E = (0,T), one can use the global Carleman estimate to prove the corresponding null controllability result for the deterministic counterpart of (5.9). However, at least at this moment we do not know how to use a similar method to prove Theorem 5.1 even for the same case that E = (0,T).

Next, we consider the approximate controllability for the system (5.9) under a stronger assumption on the controller $E \times G_0$ than that for the null controllability.

Definition 5.3 The system (5.9) is said to be approximately controllable at time T if for any initial datum $y_0 \in L^2(G; \mathbb{R}^m)$, any final state $y_T \in L^2_{\mathcal{F}_T}(\Omega; L^2(G; \mathbb{R}^m))$ and any $\varepsilon > 0$, there exists a control $u \in L^\infty_{\mathbb{F}}(0, T; L^2(\Omega; L^2(G; \mathbb{R}^m)))$ such that the corresponding solution to (5.9) satisfies that $|y(T) - y_T|_{L^2_{\mathcal{F}_T}(\Omega; L^2(G; \mathbb{R}^m))} \leq \varepsilon$.

Theorem 5.2 The system (5.9) is approximately controllable at time T if and only if $\mathbf{m}((s,T)\cap E)>0$ for any $s\in[0,T)$.

At the first glance, it seems that Theorem 5.2 is unreasonable. If $b(\cdot) \equiv 0$, then the system (5.9) is like a deterministic parabolic equation with a random parameter. The readers may guess that one can obtain the approximate controllability by only assuming that $\mathbf{m}(E) > 0$. However, this is not the case. The reason for this comes from our definition of the approximate controllability for the system (5.9). We expect that any element belonging to $L^2_{\mathcal{F}_T}(\Omega; L^2(G))$ rather than $L^2_{\mathcal{F}_s}(\Omega; L^2(G))$ (s < T) can be attached as close as one wants. Hence we need to put the control u to be active until the time T.

In some sense, it is surprising that one needs a little more assumption in Theorem 5.2 for the approximate controllability of (5.9) than that in Theorem 5.1 for the null controllability. Indeed, it is well-known that in the deterministic setting, the null controllability is

usually stronger than the approximate controllability. But this does not remain to be true in the stochastic case. Actually, from Theorem 5.2, we see that the additional condition (compared to the null controllability) that $\mathbf{m}((s,T) \cap E) > 0$ for any $s \in [0,T)$ is not only sufficient but also necessary for the approximate controllability of (5.9). Therefore, in the setting of stochastic distributed parameter systems, the null controllability does NOT imply the approximate controllability. This indicates that there exists some essential difference between the controllability theory of the deterministic parabolic equations and its stochastic counterpart.

5.2.1 Some preliminaries

Before proving Theorems 5.1 and 5.2, we give some preliminary results. To begin with, we recall the following known property about Lebesgue measurable sets.

Lemma 5.2 ([7, pp. 256–257]) For a.e. $\tilde{t} \in E$, there exists a sequence of numbers $\{t_i\}_{i=1}^{\infty} \subset (0,T)$ such that

$$t_1 < t_2 < \dots < t_i < t_{i+1} < \dots < \tilde{t}, \qquad t_i \to \tilde{t} \text{ as } i \to \infty,$$
 (5.10)

$$\mathbf{m}(E \cap [t_i, t_{i+1}]) \ge \rho_1(t_{i+1} - t_i), \quad i = 1, 2, \cdots,$$
 (5.11)

$$\frac{t_{i+1} - t_i}{t_{i+2} - t_{i+1}} \le \rho_2, \quad i = 1, 2, \cdots,$$
(5.12)

where ρ_1 and ρ_2 are two positive constants which are independent of i.

Next, we give the following result (which is a Riesz-type Representation Theorem for the dual of space $L^p_{\mathbb{F}}(0,T;L^q(\Omega;H))$). Its proof can be found in [10] or [14, Chapter 1].

Lemma 5.3 Suppose $1 \leq p, q < \infty$, and that H is a Hilbert space. Then

$$L_{\mathbb{F}}^{p}(0,T;L^{q}(\Omega;H))^{*} = L_{\mathbb{F}}^{p'}(0,T;L^{q'}(\Omega;H)). \tag{5.13}$$

Here, p' and q' are respectively the (usual Hölder) conjugate numbers of p and q.

Next, let us define an unbounded operator A on $L^2(G)$ as follows:

$$\begin{cases}
D(A) = H^{2}(G) \cap H_{0}^{1}(G), \\
Ah = -\sum_{j,k=1}^{n} (a^{jk} h_{x_{j}})_{x_{k}}, \quad \forall h \in D(A).
\end{cases}$$
(5.14)

Let $\{\lambda_i\}_{i=1}^{\infty}$ be the eigenvalues of A, and $\{e_i\}_{i=1}^{\infty}$ be the corresponding eigenfunctions satisfying $|e_i|_{L^2(G)} = 1$, $i = 1, 2, 3 \cdots$. It holds that $0 < \lambda_1 \le \lambda_2 \le \lambda_3 \le \cdots \le \lambda_k \le \cdots \to \infty$. For any $r \ge \lambda_1$, write $\Lambda_r = \{i \in \mathbb{N} \mid \lambda_i \le r\}$. We recall the following observability estimate (for partial sums of the eigenfunctions of A), established in [9, Theorem 1.2] (See also [6, Theorem 3] for a special case of this result).

Lemma 5.4 There exist two positive constants $C_1 \geq 1$ and $C_2 \geq 1$ such that

$$\sum_{i \in \Lambda_r} |a_i|^2 \le \mathcal{C}_1 e^{\mathcal{C}_2 \sqrt{r}} \int_{G_0} \left| \sum_{i \in \Lambda_r} a_i e_i(x) \right|^2 dx \tag{5.15}$$

holds for any $r \geq \lambda_1$ and $a_i \in \mathbb{C}$ with $i \in \Lambda_r$.

Further, for any s_1 and s_2 satisfying $0 \le s_1 < s_2 \le T$, we introduce the following backward stochastic parabolic system:

$$\begin{cases}
dz + \sum_{j,k=1}^{n} (a^{jk} z_{x_j})_{x_k} dt = -[a(t)^{\top} z + b(t)^{\top} Z] dt + Z dW(t) & \text{in } (s_1, s_2) \times G, \\
z = 0 & \text{on } (s_1, s_2) \times \Gamma, \\
z(s_2) = \eta & \text{in } G,
\end{cases}$$
(5.16)

where $\eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; L^2(G; \mathbb{R}^m)).$

Put

$$r_0 = 2|a|_{L_{\mathbb{F}}^{\infty}(0,T;\mathbb{R}^{m\times m})} + |b|_{L_{\mathbb{F}}^{\infty}(0,T;\mathbb{R}^{m\times m})}^{2}.$$

For each $r \geq \lambda_1$, we set $H_r = \text{span}\{e_i \mid \lambda_i \leq r\}$ and denote by Π_r the orthogonal projection from $L^2(G)$ to H_r . Write

$$H_r^m = \overbrace{H_r \times H_r \times \dots \times H_r}^{m \text{ times}}.$$
 (5.17)

To simplify the notation, we also denote by Π_r the orthogonal projection from $L^2(G; \mathbb{R}^m)$ to H_r^m . We need the following observability result for (5.16) with the final data belonging to $L^2_{\mathcal{F}_{s_2}}(\Omega; H_r^m)$, a proper subspace of $L^2_{\mathcal{F}_{s_2}}(\Omega; L^2(G; \mathbb{R}^m))$.

Proposition 5.2 For each $r \geq \lambda_1$, the solution to the system (5.16) with $\eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; H_r^m)$ satisfies that

$$\mathbb{E}|z(s_1)|_{L^2(G;\mathbb{R}^m)}^2 \le \frac{C_1 e^{C_2\sqrt{r} + r_0(s_2 - s_1)}}{(\mathbf{m}(E \cap [s_1, s_2]))^2} |\chi_E \chi_{G_0} z|_{L_{\mathbb{F}}^1(s_1, s_2; L^2(\Omega; L^2(G;\mathbb{R}^m)))}^2, \tag{5.18}$$

whenever $\mathbf{m}(E \cap [s_1, s_2]) \neq 0$.

Proof: Each $\eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; H^m_r)$ can be written as $\eta = \sum_{i \in \Lambda_r} \eta_i e_i(x)$ for some $\eta_i \in L^2_{\mathcal{F}_{s_2}}(\Omega; \mathbb{R}^m)$

with $i \in \Lambda_r$. The solution (z, Z) to (5.16) can be expressed as

$$z = \sum_{i \in \Lambda_r} z_i(t)e_i, \qquad Z = \sum_{i \in \Lambda_r} Z_i(t)e_i,$$

where $z_i(\cdot) \in C_{\mathbb{F}}([s_1, s_2]; L^2(\Omega; \mathbb{R}^m))$ and $Z_i(\cdot) \in L^2_{\mathbb{F}}(s_1, s_2; \mathbb{R}^m)$, and satisfy the following equation

$$\begin{cases} dz_i - \lambda_i z_i dt = -[a(t)^{\top} z_i + b(t)^{\top} Z_i] dt + Z_i dW(t) & \text{in } [s_1, s_2], \\ z_i(T) = \eta_i. \end{cases}$$

By Lemma 5.4, for any $t \in [s_1, s_2]$, we have

$$\mathbb{E} \int_{G} |z(t)|_{\mathbb{R}^{m}}^{2} dx = \mathbb{E} \sum_{i \in \Lambda_{r}} |z_{i}(t)|_{\mathbb{R}^{m}}^{2} \leq C_{1} e^{C_{2}\sqrt{r}} \mathbb{E} \int_{G_{0}} \left| \sum_{i \in \Lambda_{r}} z_{i}(t) e_{i} \right|_{\mathbb{R}^{m}}^{2} dx$$

$$= C_{1} e^{C_{2}\sqrt{r}} \mathbb{E} \int_{G_{0}} |z(t)|_{\mathbb{R}^{m}}^{2} dx.$$
(5.19)

By Itô's formula, we find that

$$d(e^{r_0t}|z|_{\mathbb{R}^m}^2) = r_0e^{r_0t}|z|_{\mathbb{R}^m}^2 + e^{r_0t}(\langle dz, z \rangle_{\mathbb{R}^m} + \langle z, dz \rangle_{\mathbb{R}^m}) + e^{r_0t}|dz|_{\mathbb{R}^m}^2.$$

Hence,

$$\mathbb{E}\left(e^{r_{0}t} \int_{G} |z(t)|_{\mathbb{R}^{m}}^{2} dx\right) - \mathbb{E}\left(e^{r_{0}s_{1}} \int_{G} |z(s_{1})|_{\mathbb{R}^{m}}^{2} dx\right)
= r_{0} \mathbb{E} \int_{s_{1}}^{t} \int_{G} e^{r_{0}s} |z(s)|_{\mathbb{R}^{m}}^{2} dx ds + 2 \sum_{i \in \Lambda_{r}} \mathbb{E} \int_{s_{1}}^{t} e^{r_{0}s} \lambda_{i} |z_{i}(s)|_{\mathbb{R}^{m}}^{2} ds
+ \mathbb{E} \int_{s_{1}}^{t} \int_{G} e^{r_{0}s} \left(-\langle a(s)^{\top} z(s) + b(s)^{\top} Z(s), z(s)\rangle_{\mathbb{R}^{m}} - \langle z(s), a(s)^{\top} z(s) + b(s)^{\top} Z(s)\rangle_{\mathbb{R}^{m}} + |Z(s)|_{\mathbb{R}^{m}}^{2}\right) dx ds
\geq 2 \sum_{i \in \Lambda_{r}} \mathbb{E} \int_{s_{1}}^{t} e^{r_{0}s} \lambda_{i} |z_{i}(s)|_{\mathbb{R}^{m}}^{2} ds \geq 0.$$
(5.20)

From (5.19) and (5.20), we obtain that, for any $t \in [s_1, s_2]$,

$$\mathbb{E} \int_{G} |z(s_{1}, x)|_{\mathbb{R}^{m}}^{2} dx \leq C_{1} e^{C_{2}\sqrt{r} + r_{0}(s_{2} - s_{1})} \mathbb{E} \int_{G_{0}} |z(t, x)|_{\mathbb{R}^{m}}^{2} dx.$$
 (5.21)

By (5.21), it follows that

$$\int_{E \cap [s_1, s_2]} \left[\mathbb{E} \int_G |z(s_1, x)|_{\mathbb{R}^m}^2 dx \right]^{\frac{1}{2}} dt
\leq \left(\mathcal{C}_1 e^{\mathcal{C}_2 \sqrt{r} + r_0(s_2 - s_1)} \right)^{\frac{1}{2}} \int_{E \cap [s_1, s_2]} \left[\mathbb{E} \int_{G_0} |z(t, x)|_{\mathbb{R}^m}^2 dx \right]^{\frac{1}{2}} dt.$$

Hence, when $\mathbf{m}(E \cap [s_1, s_2]) \neq 0$, we obtain that for each $\eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; H^m_r)$,

$$\mathbb{E} \int_{G} |z(s_{1}, x)|_{\mathbb{R}^{m}}^{2} dx
\leq \frac{C_{1}e^{C_{2}\sqrt{r} + r_{0}(s_{2} - s_{1})}}{(\mathbf{m}(E \cap [s_{1}, s_{2}]))^{2}} \Big\{ \int_{s_{1}}^{s_{2}} \Big[\mathbb{E} \int_{G} |\chi_{E}(t)\chi_{G_{0}}(x)z(t, x)|_{\mathbb{R}^{m}}^{2} dx \Big]^{\frac{1}{2}} dt \Big\}^{2}
= \frac{C_{1}e^{C_{2}\sqrt{r} + r_{0}(s_{2} - s_{1})}}{(\mathbf{m}(E \cap [s_{1}, s_{2}]))^{2}} |\chi_{E}\chi_{G_{0}}z|_{L_{\mathbb{F}}^{1}(s_{1}, s_{2}; L^{2}(\Omega; L^{2}(G; \mathbb{R}^{m})))}^{2},$$

which gives (5.18).

By means of the usual duality argument, Proposition 5.2 yields a partial controllability result for the following controlled system:

$$\begin{cases} dy - \sum_{j,k=1}^{n} (a^{jk}y_{x_j})_{x_k} dt = [a(t)y + \chi_E \chi_{G_0} u] dt + b(t)y dW(t) & \text{in } (s_1, s_2) \times G, \\ y = 0 & \text{on } (s_1, s_2) \times \Gamma, \\ y(s_1) = y_{s_1} & \text{in } G, \end{cases}$$
(5.22)

where $y_{s_1} \in L^2_{\mathcal{F}_{s_1}}(\Omega; L^2(G; \mathbb{R}^m))$. That is, we have the following result.

Proposition 5.3 If $\mathbf{m}(E \cap [s_1, s_2]) \neq 0$, then for every $r \geq \lambda_1$ and $y_{s_1} \in L^2_{\mathcal{F}_{s_1}}(\Omega; L^2(G; \mathbb{R}^m))$, there exists a control $u_r \in L^\infty_{\mathbb{F}}(s_1, s_2; L^2(\Omega; L^2(G; \mathbb{R}^m)))$ such that the solution y to the system (5.22) with $u = u_r$ satisfies that $\Pi_r(y(s_2)) = 0$, \mathbb{P} -a.s. Moreover, u_r verifies that

$$|u_r|_{L_{\mathbb{F}}^{\infty}(s_1, s_2; L^2(\Omega; L^2(G; \mathbb{R}^m)))}^2 \le \frac{C_1 e^{C_2 \sqrt{r} + r_0(s_2 - s_1)}}{(\mathbf{m}(E \cap [s_1, s_2]))^2} |y_{s_1}|_{L_{\mathcal{F}_{s_1}}^2(\Omega; L^2(G; \mathbb{R}^m))}^2.$$
(5.23)

Proof: Define a subspace H of $L^1_{\mathbb{F}}(s_1, s_2; L^2(\Omega; L^2(G; \mathbb{R}^m)))$:

$$H = \left\{ f = \chi_E \chi_{G_0} z \mid (z, Z) \text{ solves (5.16) for some } \eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; H_r^m) \right\}$$

and a linear functional \mathcal{L} on H:

$$\mathcal{L}(f) = -\mathbb{E} \int_{G} \langle y_{s_1}, z(s_1) \rangle_{\mathbb{R}^m} dx.$$

By Proposition 5.2, it is easy to check that \mathcal{L} is a bounded linear functional on H and

$$|\mathcal{L}|^2 \le \frac{C_1 e^{C_2 \sqrt{r} + r_0(s_2 - s_1)}}{(\mathbf{m}(E \cap [s_1, s_2]))^2} |y_{s_1}|^2_{L^2_{\mathcal{F}_{s_1}}(\Omega; L^2(G; \mathbb{R}^m))}.$$

By the Hahn-Banach Theorem, \mathcal{L} can be extended to a bounded linear functional $\widetilde{\mathcal{L}}$ (satisfying $|\widetilde{\mathcal{L}}|=|\mathcal{L}|$) on $L^1_{\mathbb{F}}(s_1,s_2;L^2(\Omega;L^2(G;\mathbb{R}^m)))$. By Lemma 5.3, there exists a control $u_r \in L^\infty_{\mathbb{F}}(s_1,s_2;L^2(\Omega;L^2(G;\mathbb{R}^m)))$ such that

$$\mathbb{E} \int_{s_1}^{s_2} \int_G \langle u_r, f \rangle_{\mathbb{R}^m} dx dt = \widetilde{\mathcal{L}}(f), \quad \forall f \in L^1_{\mathbb{F}}(s_1, s_2; L^2(\Omega; L^2(G; \mathbb{R}^m))).$$

In particular, for any $\eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; H_r^m)$, the corresponding solution (z, Z) to (5.16) satisfies

$$\mathbb{E} \int_{s_1}^{s_2} \int_G \langle u_r, \chi_E \chi_{G_0} z \rangle_{\mathbb{R}^m} dx dt = -\mathbb{E} \int_G \langle y_{s_1}, z(s_1) \rangle_{\mathbb{R}^m} dx.$$
 (5.24)

Applying Itô's formula to $\langle y, z \rangle_{\mathbb{R}^m}$, where y solves the system (5.22) with $u = u_r$, we obtain that

$$\mathbb{E} \int_{G} \langle y(s_2), \eta \rangle_{\mathbb{R}^m} dx - \mathbb{E} \int_{G} \langle y_{s_1}, z(s_1) \rangle_{\mathbb{R}^m} dx = \mathbb{E} \int_{s_1}^{s_2} \int_{G} \langle \chi_E \chi_{G_0} u_r, z \rangle_{\mathbb{R}^m} dx dt.$$
 (5.25)

Combining (5.24) and (5.25), we arrive at

$$\mathbb{E} \int_{G} \langle y(s_2), \eta \rangle_{\mathbb{R}^m} dx = 0, \qquad \forall \, \eta \in L^2_{\mathcal{F}_{s_2}}(\Omega; H_r^m),$$

which implies that $\Pi_r(y(s_2)) = 0$, \mathbb{P} -a.s. Moreover, $|u_r|_{L_{\mathbb{F}}^{\infty}(s_1, s_2; L^2(\Omega; L^2(G; \mathbb{R}^m)))} = |\mathcal{L}|$, which yields (5.23).

Finally, for any $s \in [0, T)$, we consider the following equation:

$$\begin{cases} dy - \sum_{j,k=1}^{n} (a^{jk}y_{x_j})_{x_k} dt = a(t)ydt + b(t)ydW(t) & \text{in } (s,T) \times G, \\ y = 0 & \text{on } (s,T) \times \Gamma, \\ y(s) = y_s & \text{in } G, \end{cases}$$

$$(5.26)$$

where $y_s \in L^2_{\mathcal{F}_s}(\Omega; L^2(G; \mathbb{R}^m))$. Let us show the following decay result for the system (5.26).

Proposition 5.4 Let $r \geq \lambda_1$. Then, for any $y_s \in L^2_{\mathcal{F}_s}(\Omega; L^2(G; \mathbb{R}^m))$ with $\Pi_r(y_s) = 0$, \mathbb{P} -a.s., the corresponding solution y to (5.26) satisfies that

$$\mathbb{E}|y(t)|_{L^{2}(G;\mathbb{R}^{m})}^{2} \leq e^{-(2r-r_{0})(t-s)}|y_{s}|_{L^{2}_{\mathcal{F}_{s}}(\Omega;L^{2}(G;\mathbb{R}^{m}))}^{2}, \quad \forall t \in [s,T].$$
(5.27)

Proof: Since $y_s \in L^2_{\mathcal{F}_s}(\Omega; L^2(G; \mathbb{R}^m))$ satisfying $\Pi_r(y_s) = 0$, we see that $y_s = \sum_{i \in \mathbb{N} \setminus \Lambda_r} y_s^i e_i$ for some $y_s^i \in L^2_{\mathcal{F}_s}(\Omega; \mathbb{R}^m)$ with $i \in \mathbb{N} \setminus \Lambda_r$. Clearly, the solution y to (5.26) can be expressed as $y = \sum_{i \in \mathbb{N} \setminus \Lambda_r} y^i(t) e_i$, where $y^i(\cdot) \in C_{\mathbb{F}}([s,T]; L^2(\Omega; \mathbb{R}^m))$ solves the following stochastic differential equation:

$$\begin{cases} dy^{i} + \lambda_{i}y^{i}dt = a(t)y^{i}dt + b(t)y^{i}dW(t) & \text{in } [s, T], \\ y^{i}(s) = y_{s}^{i}. \end{cases}$$

By Itô's formula, we have that

$$d(e^{(2r-r_0)(t-s)}|y|_{\mathbb{R}^m}^2) = e^{(2r-r_0)(t-s)} (\langle dy, y \rangle_{\mathbb{R}^m} + \langle y, dy \rangle_{\mathbb{R}^m})$$

$$+ e^{(2r-r_0)(t-s)} |dy|_{\mathbb{R}^m}^2 + (2r-r_0)e^{(2r-r_0)(t-s)} |y|_{\mathbb{R}^m}^2.$$

Hence, by $\lambda_i > r$ for each $i \in \mathbb{N} \setminus \Lambda_r$ and recalling that $r_0 = 2|a|_{L^{\infty}_{\mathbb{F}}(0,T;\mathbb{R}^{m \times m})} + |b|^2_{L^{\infty}_{\mathbb{F}}(0,T;\mathbb{R}^{m \times m})}$, we arrive at

$$\begin{split} & \mathbb{E} \int_{G} e^{(2r-r_{0})(t-s)} |y(t)|_{\mathbb{R}^{m}}^{2} dx - \mathbb{E} \int_{G} |y(s)|_{\mathbb{R}^{m}}^{2} dx \\ & = -2 \sum_{i \in \mathbb{N} \backslash \Lambda_{r}} \lambda_{i} \mathbb{E} \int_{s}^{t} e^{(2r-r_{0})(\sigma-s)} |y^{i}(\sigma)|_{\mathbb{R}^{m}}^{2} d\sigma + \mathbb{E} \int_{s}^{t} \int_{G} e^{(2r-r_{0})(\sigma-s)} \left(\langle \, ay,y \, \rangle_{\mathbb{R}^{m}} + \langle \, y,ay \, \rangle_{\mathbb{R}^{m}} \right) dx d\sigma \\ & + \mathbb{E} \int_{s}^{t} \int_{G} e^{(2r-r_{0})(\sigma-s)} |b(\sigma)y(\sigma)|_{\mathbb{R}^{m}}^{2} dx d\sigma + (2r-r_{0})\mathbb{E} \int_{s}^{t} e^{(2r-r_{0})(\sigma-s)} |y(\sigma)|_{\mathbb{R}^{m}}^{2} dx d\sigma \\ & \leq 0, \end{split}$$

which gives the desired estimate (5.27) immediately.

5.2.2 Proof of the null controllability result

Now we are in a position to prove Theorem 5.1.

Proof of Theorem 5.1: We borrow some idea in [5]. For simplicity, we assume that m = 1. By Lemma 5.2, we may take a number $\tilde{t} \in E$ with $\tilde{t} < T$ and a sequence $\{t_N\}_{N=1}^{\infty} \subset (0, T)$ such that (5.10)–(5.12) hold for some positive numbers ρ_1 and ρ_2 .

Write $\tilde{y}_0 = \psi(t_1)$, where $\psi(\cdot)$ solves the following stochastic parabolic system:

$$\begin{cases} d\psi - \sum_{j,k=1}^{n} (a^{jk}\psi_{x_j})_{x_k} dt = a(t)\psi dt + b(t)\psi dW(t) & \text{in } (0,t_1) \times G, \\ \psi = 0 & \text{on } (0,t_1) \times \Gamma, \\ \psi(0) = y_0 & \text{in } G. \end{cases}$$

Let us consider the following controlled stochastic parabolic system:

$$\begin{cases}
d\tilde{y} - \sum_{j,k=1}^{n} (a^{jk} \tilde{y}_{x_j})_{x_k} dt = [a(t)\tilde{y} + \chi_E \chi_{G_0} \tilde{u}] dt + b(t) \tilde{y} dW(t) & \text{in } (t_1, \tilde{t}) \times G, \\
\tilde{y} = 0 & \text{on } (t_1, \tilde{t}) \times \Gamma, \\
\tilde{y}(t_1) = \tilde{y}_0 & \text{in } G.
\end{cases} (5.28)$$

It suffices to find a control $\tilde{u} \in L_{\mathbb{F}}^{\infty}(t_1, \tilde{t}; L^2(\Omega; L^2(G)))$ with

$$|\tilde{u}|_{L_{\mathbb{F}}^{\infty}(t_{1},\tilde{t};L^{2}(\Omega;L^{2}(G)))}^{2} \le \mathcal{C}\mathbb{E}|\tilde{y}_{0}|_{L^{2}(\Omega)}^{2},$$

$$(5.29)$$

such that the solution \tilde{y} to (5.28) satisfies $\tilde{y}(\tilde{t}) = 0$ in G, \mathbb{P} -a.s.

Set $I_N = [t_{2N-1}, t_{2N}]$ and $J_N = [t_{2N}, t_{2N+1}]$ for $N \in \mathbb{N}$. Then $[t_1, \tilde{t}] = \bigcup_{N=1}^{\infty} (I_N \cup J_N)$. Clearly, $\mathbf{m}(E \cap I_N) > 0$ and $\mathbf{m}(E \cap J_N) > 0$. We will introduce a suitable control on each I_N and allow the system to evolve freely on every J_N . Also, we fix a suitable, strictly increasing sequence $\{r_N\}_{N=1}^{\infty}$ of positive integers (to be given later) satisfying that $\lambda_1 \leq r_1 < r_2 < \cdots < r_N \to \infty$ as $N \to \infty$.

We consider first the controlled stochastic parabolic system on the interval $I_1 = [t_1, t_2]$ as follows:

$$\begin{cases} dy_1 - \sum_{j,k=1}^n (a^{jk}y_{1,x_j})_{x_k} dt = [a(t)y_1 + \chi_E \chi_{G_0} u_1] dt + b(t)y_1 dW(t) & \text{in } (t_1, t_2) \times G, \\ y_1 = 0 & \text{on } (t_1, t_2) \times \Gamma, \\ y_1(t_1) = \tilde{y}_0 & \text{in } G. \end{cases}$$
(5.30)

By Proposition 5.3, there exists a control $u_1 \in L^{\infty}_{\mathbb{F}}(t_1, t_2; L^2(\Omega; L^2(G)))$ with the estimate:

$$|u_1|_{L_{\mathbb{F}}^{\infty}(t_1,t_2;L^2(\Omega;L^2(G)))}^2 \leq \frac{C_1 e^{C_2\sqrt{r_1}+r_0T}}{(\mathbf{m}(E\cap[t_1,t_2]))^2} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2,$$

such that $\Pi_{r_1}(y(t_2)) = 0$ in G, \mathbb{P} -a.s. By (5.11), we see that

$$|u_1|_{L_{\mathbb{F}}^{\infty}(t_1,t_2;L^2(\Omega;L^2(G)))}^2 \le \frac{C_1 e^{C_2\sqrt{r_1}+r_0T}}{\rho_1^2(t_2-t_1)^2} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2.$$
(5.31)

Applying Itô's formula to $e^{-(r_0+1)t}|y_1(t)|^2_{L^2(G)}$, similar to the proof of (5.20), we obtain that

$$\begin{split} &e^{-(r_0+1)t_2}\mathbb{E}|y_1(t_2)|^2_{L^2(G)}\\ &=e^{-(r_0+1)t_1}\mathbb{E}|y_1(t_1)|^2_{L^2(G)}-(r_0+1)\mathbb{E}\int_{t_1}^{t_2}e^{-(r_0+1)s}\int_G|y_1|^2dxds\\ &-2\sum_{j,k=1}^n\mathbb{E}\int_{t_1}^{t_2}e^{-(r_0+1)s}\int_Ga^{jk}y_{1,x_j}y_{1,x_k}dxds\\ &+\mathbb{E}\int_{t_1}^{t_2}e^{-(r_0+1)s}\int_G[2a(s)|y_1|^2+|b(s)y_1|^2]dxds+2\mathbb{E}\int_{t_1}^{t_2}e^{-(r_0+1)s}\int_G\chi_E\chi_{G_0}u_1y_1dxds\\ &\leq e^{-(r_0+1)t_1}\mathbb{E}|y_1(t_1)|^2_{L^2(G)}+\mathbb{E}\int_{t_1}^{t_2}e^{-(r_0+1)s}\int_G|u_1|^2dxds\\ &\leq e^{-(r_0+1)t_1}\mathbb{E}|\tilde{y}_0|^2_{L^2(G)}+\frac{e^{-(r_0+1)t_1}-e^{-(r_0+1)t_2}}{r_0+1}|u_1|^2_{L^\infty_F(t_1,t_2;L^2(\Omega;L^2(G)))}. \end{split}$$

Hence, in view of (5.31),

$$\mathbb{E}|y_1(t_2)|_{L^2(G)}^2 \le \frac{C_3 e^{C_3 \sqrt{r_1}}}{(t_2 - t_1)^2} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2.$$
(5.32)

where $C_3 = \max(2\rho_1^{-2}C_1e^{(2r_0+1)T}, C_2)$.

Then, on the interval $J_1 \equiv [t_2, t_3]$, we consider the following stochastic parabolic system without control:

$$\begin{cases} dz_1 - \sum_{j,k=1}^n (a^{jk} z_{1,x_j})_{x_k} dt = a(t) z_1 dt + b(t) z_1 dW(t) & \text{in } (t_2, t_3) \times G, \\ z_1 = 0 & \text{on } (t_2, t_3) \times \Gamma, \\ z_1(t_2) = y_1(t_2) & \text{in } G. \end{cases}$$

Since $\Pi_{r_1}(y_1(t_2)) = 0$, P-a.s., by Proposition 5.4, we have

$$\mathbb{E}|z_1(t_3)|_{L^2(G)}^2 \le e^{(-2r_1+r_0)(t_3-t_2)} \mathbb{E}|y_1(t_2)|_{L^2(G)}^2 \le \frac{C_3 e^{C_3\sqrt{r_1}}}{(t_2-t_1)^2} e^{(-2r_1+r_0)(t_3-t_2)} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2. \tag{5.33}$$

Generally, on the interval I_N with $N \in \mathbb{N} \setminus \{1\}$, we consider a controlled stochastic parabolic system as follows:

$$\begin{cases} dy_N - \sum_{j,k=1}^n (a^{jk}y_{N,x_j})_{x_k} dt = [a(t)y_N + \chi_E \chi_{G_0} u_N] dt + b(t)y_N dW(t) & \text{in } (t_{2N-1}, t_{2N}) \times G, \\ y_N = 0 & \text{on } (t_{2N-1}, t_{2N}) \times \Gamma, \\ y_N(t_{2N-1}) = z_{N-1}(t_{2N-1}) & \text{in } G. \end{cases}$$

Similar to the above argument (See the proof of (5.31) and (5.32)), one can find a control $u_N \in L_{\mathbb{F}}^{\infty}(t_{2N-1}, t_{2N}; L^2(\Omega; L^2(G)))$ with the estimate:

$$|u_N|_{L_{\mathbb{F}}^{\infty}(t_{2N-1},t_{2N};L^2(\Omega;L^2(G)))}^2 \le \frac{C_1 e^{C_2\sqrt{r_N}+r_0T}}{\rho_1^2(t_{2N}-t_{2N-1})^2} \mathbb{E}|z_{N-1}(t_{2N-1})|_{L^2(G)}^2.$$
(5.34)

such that $\Pi_{r_N}(y_N(t_{2N})) = 0$ in G, \mathbb{P} -a.s. Moreover,

$$\mathbb{E}|y_N(t_{2N})|_{L^2(G)}^2 \le \frac{C_3 e^{C_3 \sqrt{r_N}}}{(t_{2N} - t_{2N-1})^2} \mathbb{E}|z_{N-1}(t_{2N-1})|_{L^2(G)}^2.$$
(5.35)

On the interval J_N , we consider the following stochastic parabolic system without control:

$$\begin{cases} dz_N - \sum_{j,k=1}^n (a^{jk} z_{N,x_j})_{x_k} dt = a(t) z_N dt + b(t) z_N dW(t) & \text{in } (t_{2N}, t_{2N+1}) \times G, \\ z_N = 0 & \text{on } (t_{2N}, t_{2N+1}) \times \Gamma, \\ z_N(t_{2N}) = y_N(t_{2N}) & \text{in } G. \end{cases}$$

Since $\Pi_{r_N}(y_N(t_{2N})) = 0$, \mathbb{P} -a.s., by Proposition 5.4 and similar to (5.33), and recalling that $y_N(t_{2N-1}) = z_{N-1}(t_{2N-1})$ in G, we have

$$\mathbb{E}|z_{N}(t_{2N+1})|_{L^{2}(G)}^{2} \leq \frac{C_{3}e^{C_{3}\sqrt{r_{N}}}}{(t_{2N}-t_{2N-1})^{2}}e^{(-2r_{N}+r_{0})(t_{2N+1}-t_{2N})}\mathbb{E}|y_{N}(t_{2N-1})|_{L^{2}(G)}^{2}
\leq \frac{C_{4}e^{C_{4}\sqrt{r_{N}}}}{(t_{2N}-t_{2N-1})^{2}}e^{-2r_{N}(t_{2N+1}-t_{2N})}\mathbb{E}|z_{N-1}(t_{2N-1})|_{L^{2}(G)}^{2},$$
(5.36)

where $C_4 = C_3 e^{r_0 T}$.

Inductively, by (5.12) and (5.36), we conclude that, for all $N \ge 1$,

$$\mathbb{E}|z_{N}(t_{2N+1})|_{L^{2}(G)}^{2} \leq \frac{C_{4}^{N}e^{C_{4}(\sqrt{r_{N}}+\sqrt{r_{N-1}}+\cdots+\sqrt{r_{1}})}}{(t_{2N}-t_{2N-1})^{2}(t_{2N-2}-t_{2N-3})^{2}\cdots(t_{2}-t_{1})^{2}} \\
\times \exp\left\{-2r_{N}(t_{2N+1}-t_{2N})-2r_{N-1}(t_{2N-1}-t_{2N-2})-\cdots-2r_{1}(t_{3}-t_{2})\right\}\mathbb{E}|\tilde{y}_{0}|_{L^{2}(G)}^{2} \\
\leq \frac{C_{4}^{N}\exp\left\{C_{4}N\sqrt{r_{N}}-2r_{N}(t_{2N+1}-t_{2N})\right\}}{(t_{2N}-t_{2N-1})^{2}(t_{2N-2}-t_{2N-3})^{2}\cdots(t_{2}-t_{1})^{2}}\mathbb{E}|\tilde{y}_{0}|_{L^{2}(G)}^{2} \\
\leq \frac{C_{4}^{N}\rho_{2}^{2N(N-1)}\exp\left\{C_{4}N\sqrt{r_{N}}-2(t_{2}-t_{1})\rho_{2}^{1-2N}r_{N}\right\}}{(t_{2}-t_{1})^{2N}}\mathbb{E}|\tilde{y}_{0}|_{L^{2}(G)}^{2}. \tag{5.37}$$

By (5.12), (5.34)–(5.35) and (5.37), we see that

$$|u_N|^2_{L^\infty_{\mathbb{F}}(t_{2N-1},t_{2N};L^2(\Omega;L^2(G)))}$$

$$\leq \frac{C_1 C_4^{N-1} \rho_2^{2N(N-1)}}{\rho_1^2 (t_2 - t_1)^{2N}} \exp\left\{C_2 \sqrt{r_N} + r_0 T + C_4 (N-1) \sqrt{r_{N-1}} - 2(t_2 - t_1) \rho_2^{3-2N} r_{N-1}\right\} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2. \tag{5.38}$$

and

$$\mathbb{E}|y_{N}(t_{2N})|_{L^{2}(G)}^{2} \leq \frac{C_{3}C_{4}^{N-1}\rho_{2}^{2N(N-1)}}{(t_{2}-t_{1})^{2N}} \exp\left\{C_{3}\sqrt{r_{N}} + C_{4}(N-1)\sqrt{r_{N-1}} - 2(t_{2}-t_{1})\rho_{2}^{3-2N}r_{N-1}\right\} \mathbb{E}|\tilde{y}_{0}|_{L^{2}(G)}^{2}.$$
(5.39)

We now choose $r_N = \max(2^{N^2}, [\lambda_1] + 1)$. From (5.38)–(5.39), it is easy to see that, whenever N is large enough,

$$|u_N|_{L_{\mathbb{F}}^{\infty}(t_{2N-1}, t_{2N}; L^2(\Omega; L^2(G)))}^2 \le \frac{1}{2^N} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2 \tag{5.40}$$

and

$$\mathbb{E}|y_N(t_{2N})|_{L^2(G)}^2 \le \frac{1}{2^N} \mathbb{E}|\tilde{y}_0|_{L^2(G)}^2. \tag{5.41}$$

We now construct a control \tilde{u} by setting

$$\tilde{u}(t,x) = \begin{cases} u_N(t,x), & (t,x) \in I_N \times G, & N \ge 1, \\ 0, & (t,x) \in J_N \times G, & N \ge 1. \end{cases}$$
 (5.42)

By (5.40), we see that $\tilde{u} \in L_{\mathbb{F}}^{\infty}(t_1, \tilde{t}; L^2(\Omega; L^2(G)))$ satisfies (5.29). Let \tilde{y} be the solution to the system (5.28) corresponding to the control constructed in (5.42). Then $\tilde{y}(\cdot) = y_N(\cdot)$ on $I_N \times G$. By (5.41) and recalling that $t_{2N} \to \tilde{t}$ as $N \to \infty$, we deduce that $\tilde{y}(\tilde{t}) = 0$, \mathbb{P} -a.s. This completes the proof of Theorem 5.1.

5.2.3 Proof of the approximate controllability result

To begin with, we show the following two preliminary results, which have some independent interests.

Proposition 5.5 If $\mathbf{m}((s,T) \cap E) > 0$ for any $s \in [0,T)$, then for any given $\eta \in L^2_{\mathcal{F}_T}(\Omega; L^2(G; \mathbb{R}^m))$, the corresponding solution to (5.16) with $s_1 = 0$ and $s_2 = T$ satisfies

$$|z(s)|_{L^{2}_{\mathcal{F}_{s}}(\Omega; L^{2}(G; \mathbb{R}^{m}))} \le \mathcal{C}(s)|\chi_{E}\chi_{G_{0}}z|_{L^{1}_{\mathbb{F}}(s, T; L^{2}(\Omega; L^{2}(G; \mathbb{R}^{m})))}.$$
(5.43)

Here and henceforth, C(s) > 0 is a generic constant depending on s.

Proof: We consider the following controlled stochastic parabolic system:

$$\begin{cases} dy - \sum_{j,k=1}^{n} (a^{jk}y_{x_j})_{x_k} dt = [a(t)y + \chi_{(s,T)\cap E}\chi_{G_0}u]dt + b(t)ydW(t) & \text{in } (s,T) \times G, \\ y = 0 & \text{on } (s,T) \times \Gamma, \\ y(s) = y_s & \text{in } G, \end{cases}$$

$$(5.44)$$

where y is the state variable, u is the control variable, the initial state $y_s \in L^2_{\mathcal{F}_s}(\Omega; L^2(G; \mathbb{R}^m))$ and the control $u(\cdot) \in L^\infty_{\mathbb{F}}(s, T; L^2(\Omega; L^2(G; \mathbb{R}^m)))$. By the proof of Theorem 5.1, it is easy to show that the system (5.44) is null controllable, i.e., for any $y_s \in L^2_{\mathcal{F}_s}(\Omega; L^2(G; \mathbb{R}^m))$, there exists a control $u \in L^\infty_{\mathbb{F}}(s, T; L^2(\Omega; L^2(G; \mathbb{R}^m)))$ such that y(T) = 0 in G, \mathbb{P} -a.s., and

$$|u|_{L_{\mathbb{F}}^{\infty}(s,T;L^{2}(\Omega;L^{2}(G;\mathbb{R}^{m})))}^{2} \leq \mathcal{C}(s)|y_{s}|_{L_{\mathcal{F}_{a}}^{2}(\Omega;L^{2}(G;\mathbb{R}^{m}))}^{2}.$$
(5.45)

Applying Itô's formula to $\langle y, z \rangle_{\mathbb{R}^m}$, where y and (z, Z) solve respectively (5.44) and (5.16) with $s_1 = 0$ and $s_2 = T$, and noting that y(T) = 0 in G, \mathbb{P} -a.s., we obtain that

$$-\mathbb{E} \int_{G} \langle y_{s}, z(s) \rangle_{\mathbb{R}^{m}} dx = \mathbb{E} \int_{(s,T) \cap E} \int_{G_{0}} \langle u, z \rangle_{\mathbb{R}^{m}} dx dt.$$

Choosing $y_s = -z(s)$ in (5.44), we then have

$$\mathbb{E} \int_{G} |z(s)|_{\mathbb{R}^{m}}^{2} dx = \mathbb{E} \int_{(s,T)\cap E} \int_{G_{0}} \langle u,z \rangle_{\mathbb{R}^{m}} dx dt$$

$$\leq |u|_{L_{\mathbb{F}}^{\infty}(s,T;L^{2}(\Omega;L^{2}(G;\mathbb{R}^{m})))} |\chi_{(s,T)\cap E}\chi_{G_{0}}z|_{L_{\mathbb{F}}^{1}(s,T;L^{2}(\Omega;L^{2}(G;\mathbb{R}^{m})))}$$

$$\leq \mathcal{C}(s) \Big(\mathbb{E} \int_{G} |z(s)|_{\mathbb{R}^{m}}^{2} dx \Big)^{\frac{1}{2}} |\chi_{(s,T)\cap E}\chi_{G_{0}}z|_{L_{\mathbb{F}}^{1}(s,T;L^{2}(\Omega;L^{2}(G;\mathbb{R}^{m})))},$$

which gives immediately the desired estimate (5.43).

As an easy consequence of Proposition 5.5, we have the following unique continuation property for solutions to (5.16) with $s_1 = 0$ and $s_2 = T$.

Corollary 5.1 If $\mathbf{m}((s,T) \cap E) > 0$ for any $s \in [0,T)$, then any solution (z,Z) to (5.16) with $s_1 = 0$ and $s_2 = T$ vanishes identically in Q, \mathbb{P} -a.s. provided that z = 0 in $G_0 \times E$, \mathbb{P} -a.s.

Proof: Since z=0 in $G_0 \times E$, \mathbb{P} -a.s., by Proposition 5.5, we see that z(s)=0 in G, \mathbb{P} -a.s., for any $s \in [0,T)$. Therefore, $z \equiv 0$ in Q, \mathbb{P} -a.s.

Remark 5.2 If the condition $\mathbf{m}((s,T) \cap E) > 0$ for any $s \in [0,T)$ was not assumed, the conclusion in Corollary 5.1 might fail to be true. This can be shown by the following counterexample. Let E satisfy that $\mathbf{m}(E) > 0$ and $\mathbf{m}((s_0,T) \cap E) = 0$ for some $s_0 \in [0,T)$. Let $(z_1,Z_1) = 0$ in $(0,s_0) \times G$, \mathbb{P} -a.s. and ξ_2 be a nonzero process in $L^2_{\mathbb{F}}(s_0,T;\mathbb{R}^m)$ (Then $Z_2 \equiv \xi_2 e_1$ is a nonzero process in $L^2_{\mathbb{F}}(s_0,T;L^2(G;\mathbb{R}^m))$). Solving the following forward stochastic differential equation:

$$\begin{cases} d\zeta_1 - \lambda_1 \zeta_1 dt = -[a(t)^{\top} \zeta_1 + b(t)^{\top} \xi_2] dt + \xi_2 dW(t) & in [s_0, T], \\ \zeta_1(s_0) = 0, \end{cases}$$

we find a nonzero $\zeta_1 \in L^2_{\mathbb{F}}(\Omega; C([s_0, T]; \mathbb{R}^m))$. In this way, we find a nonzero solution $(z_2, Z_2) \equiv (\zeta_1 e_1, \xi_2 e_1) \in L^2_{\mathbb{F}}(\Omega; C([s_0, T]; L^2(G; \mathbb{R}^m))) \times L^2_{\mathbb{F}}(s_0, T; L^2(G; \mathbb{R}^m))$ to the following forward stochastic partial differential equation:

$$\begin{cases}
dz_{2} + \sum_{j,k=1}^{n} (a^{jk} z_{2,x_{j}})_{x_{k}} dt = -[a(t)^{\top} z_{2} + b(t)^{\top} Z_{2}] dt + Z_{2} dW(t) & in (s_{0}, T) \times G, \\
z_{2} = 0 & on (s_{0}, T) \times \Gamma, \\
z_{2}(s_{0}) = 0 & in G.
\end{cases}$$
(5.46)

(Note however that one cannot solve the system (5.46) directly because this system is not well-posed). Put

$$(z, Z) = \begin{cases} (z_1, Z_1), & in (0, s_0) \times G, \\ (z_2, Z_2), & in (s_0, T) \times G. \end{cases}$$

Then, (z, Z) is a nonzero solution to (5.16) with $s_1 = 0$ and $s_2 = T$, and z = 0 in $G_0 \times E$, \mathbb{P} -a.s. Note also that, the nonzero solution constructed for the system (5.46) indicates that, in general, the forward uniqueness does NOT hold for backward stochastic differential equations.

We are now in a position to prove Theorem 5.2.

Proof of Theorem 5.2: The "if" part follows from Corollary 5.1. To prove the "only if" part, we use the contradiction argument. Assume that $\mathbf{m}((s_0,T)\cap E)=0$ for some $s_0\in[0,T)$. Since the system (5.9) is approximately controllable at time T, we deduce that any solution (z,Z) to (5.16) with $s_1=0$ and $s_2=T$ vanishes identically in Q provided that z=0 in $G_0\times E$, \mathbb{P} -a.s. This contradicts the counterexample in Remark 5.2.

5.3 Null controllability of stochastic parabolic systems

In this subsection, we deal with the null controllability for (5.2). The results in this subsection are taken from [22].

We have the following result.

Theorem 5.3 Let the condition (5.3) be satisfied. Then the system (5.2) is null controllable at time T.

In order to prove Theorem 5.3, by means of the standard duality argument, it suffices to establish the following observability result for (5.6):

Theorem 5.4 Let the condition (5.3) be satisfied. Then, for all $z_T \in L^2_{\mathcal{F}_T}(\Omega; L^2(G; \mathbb{R}^m))$, solutions $(z, Z) \in L^2_{\mathbb{F}}(\Omega; C([0, T]; L^2(G; \mathbb{R}^m))) \times L^2_{\mathbb{F}}(0, T; L^2(G; \mathbb{R}^m))$ to the system (5.6) satisfy

$$|z(0)|_{L^{2}_{\mathcal{F}_{0}}(\Omega;L^{2}(G;\mathbb{R}^{m}))} \leq \mathcal{C}(|z|_{L^{2}_{\mathbb{F}}(0,T;L^{2}(G_{0};\mathbb{R}^{m}))} + |Z|_{L^{2}_{\mathbb{F}}(0,T;L^{2}(G;\mathbb{R}^{m}))}). \tag{5.47}$$

Remark 5.3 In Theorem 5.4, we assume that $a_{1j} \in L_{\mathbb{F}}^{\infty}(0,T;W^{1,\infty}(G;\mathbb{R}^{m\times m}))$ for $j=1,2,\cdots,n$ (See the condition (5.3)). It seems that this assumption can be weakened as $a_{1j} \in L_{\mathbb{F}}^{\infty}(0,T;L^{\infty}(G;\mathbb{R}^{m\times m}))$.

The rest of this subsection is devoted to giving a proof of Theorem 5.4. For simplicity, we consider only the case m = 1.

5.3.1 A weighted identity and Carleman estimate for a stochastic parabolic-like operator

In order to prove Theorem 5.4, we need to derive a weighted identity and Carleman estimate for a stochastic parabolic-like operator.

We assume that

$$b^{jk} = b^{kj} \in L^2_{\mathbb{F}}(\Omega; C^1([0, T]; W^{2, \infty}(G))), \qquad j, k = 1, 2, \dots, n,$$
 (5.48)

 $\ell \in C^{1,3}(Q)$ and $\Psi \in C^{1,2}(Q)$. Write

$$\begin{cases}
\mathcal{A} = -\sum_{j,k=1}^{n} \left(b^{jk} \ell_{x_{j}} \ell_{x_{k}} - b^{jk}_{x_{k}} \ell_{x_{j}} - b^{jk} \ell_{x_{j}x_{k}} \right) - \Psi - \ell_{t}, \\
\mathcal{B} = 2 \left[\mathcal{A} \Psi - \sum_{j,k=1}^{n} \left(\mathcal{A} b^{jk} \ell_{x_{j}} \right)_{x_{k}} \right] - \mathcal{A}_{t} - \sum_{j,k=1}^{n} \left(b^{jk} \Psi_{x_{k}} \right)_{x_{j}}, \\
c^{jk} = \sum_{j',k'=1}^{n} \left[2 b^{jk'} \left(b^{j'k} \ell_{x_{j'}} \right)_{x_{k'}} - \left(b^{jk} b^{j'k'} \ell_{x_{j'}} \right)_{x_{k'}} \right] - \frac{b^{jk}_{t}}{2} + \Psi b^{jk}.
\end{cases} (5.49)$$

First, we establish a fundamental weighted identity for the stochastic parabolic-like operator " $dh - \sum_{j,k=1}^{n} (b^{jk} h_{x_j})_{x_k} dt$ ".

Theorem 5.5 Let h be an $H^2(G)$ -valued continuous semi-martingale. Set $\theta = e^{\ell}$ and $w = \theta h$. Then, for any $t \in [0,T]$ and a.e. $(x,\omega) \in G \times \Omega$,

$$2\theta \left[-\sum_{j,k=1}^{n} (b^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \right] \left[dh - \sum_{j,k=1}^{n} (b^{jk} h_{x_{j}})_{x_{k}} dt \right] + 2\sum_{j,k=1}^{n} (b^{jk} w_{x_{j}} dw)_{x_{k}}$$

$$+2\sum_{j,k=1}^{n} \left[\sum_{j',k'=1}^{n} (2b^{jk} b^{j'k'} \ell_{x_{j'}} w_{x_{j}} w_{x_{k'}} - b^{jk} b^{j'k'} \ell_{x_{j}} w_{x_{j'}} w_{x_{k'}} \right) + \Psi b^{jk} w_{x_{j}} w - b^{jk} \left(\mathcal{A} \ell_{x_{j}} + \frac{\Psi_{x_{j}}}{2} \right) w^{2} \right]_{x_{k}} dt$$

$$= 2\sum_{j,k=1}^{n} c^{jk} w_{x_{j}} w_{x_{k}} dt + \mathcal{B}w^{2} dt + d \left(\sum_{j,k=1}^{n} b^{jk} w_{x_{j}} w_{x_{k}} + \mathcal{A}w^{2} \right) + 2 \left[-\sum_{j,k=1}^{n} \left(b^{jk} w_{x_{j}} \right)_{x_{k}} + \mathcal{A}w \right]^{2} dt$$

$$-\theta^{2} \sum_{j,k=1}^{n} b^{jk} (dh_{x_{j}} + \ell_{x_{j}} dh) (dh_{x_{k}} + \ell_{x_{k}} dh) - \theta^{2} \mathcal{A} (dh)^{2}.$$

$$(5.50)$$

Proof: The proof is divided into four steps.

Step 1. Recalling that $\theta = e^{\ell}$ and $w = \theta h$, one has $dh = \theta^{-1}(dw - \ell_t w dt)$ and $h_{x_j} = \theta^{-1}(w_{x_j} - \ell_{x_j} w)$ for $i = 1, 2, \dots, m$. By (5.48), it is easy to see that $\sum_{j,k=1}^n b^{jk}(\ell_{x_j} w_{x_k} + \ell_{x_j} w_{x_k})$

¹Since only the symmetry condition (5.48) is assumed for the coefficient matrix (b^{jk}) , we call " $dh - \sum_{j,k=1}^{n} (b^{jk} h_{x_j})_{x_k} dt$ " a stochastic parabolic-like operator.

$$\ell_{x_k} w_{x_j} = 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k}$$
. Hence,

$$\theta \sum_{j,k=1}^{n} (b^{jk} h_{x_{j}})_{x_{k}} = \theta \sum_{j,k=1}^{n} [\theta^{-1} b^{jk} (w_{x_{j}} - \ell_{x_{j}} w)]_{x_{k}}$$

$$= \sum_{j,k=1}^{n} [b^{jk} (w_{x_{j}} - \ell_{x_{j}} w)]_{x_{k}} - \sum_{j,k=1}^{n} b^{jk} (w_{x_{j}} - \ell_{x_{j}} w) \ell_{x_{k}}$$

$$= \sum_{j,k=1}^{n} \left[(b^{jk} w_{x_{j}})_{x_{k}} - b^{jk} (\ell_{x_{j}} w_{x_{k}} + \ell_{x_{k}} w_{x_{j}}) + (b^{jk} \ell_{x_{j}} \ell_{x_{k}} - b^{jk}_{x_{k}} \ell_{x_{j}} - b^{jk} \ell_{x_{j}x_{k}}) w \right]$$

$$= \sum_{j,k=1}^{n} \left[(b^{jk} w_{x_{j}})_{x_{k}} - 2b^{jk} \ell_{x_{j}} w_{x_{k}} + (b^{jk} \ell_{x_{j}} \ell_{x_{k}} - b^{jk}_{x_{k}} \ell_{x_{j}} - b^{jk} \ell_{x_{j}x_{k}}) w \right].$$
(5.51)

$$\begin{cases}
I \stackrel{\triangle}{=} -\sum_{j,k=1}^{n} (b^{jk} w_{x_j})_{x_k} + \mathcal{A}w, & I_1 \stackrel{\triangle}{=} \left[-\sum_{j,k=1}^{n} (b^{jk} w_{x_j})_{x_k} + \mathcal{A}w \right] dt, \\
I_2 \stackrel{\triangle}{=} dw + 2\sum_{j,k=1}^{n} b^{jk} \ell_{x_j} w_{x_k} dt, & I_3 \stackrel{\triangle}{=} \Psi w dt.
\end{cases} (5.52)$$

By (5.51) and (5.52), it follows that

$$\theta \left[dh - \sum_{j,k=1}^{n} (b^{jk} h_{x_j})_{x_k} dt \right] = I_1 + I_2 + I_3.$$

Hence,

$$2\theta \left[-\sum_{j,k=1}^{n} (b^{jk} w_{x_j})_{x_k} + \mathcal{A}w \right] \left[dh - \sum_{j,k=1}^{n} (b^{jk} h_{x_j})_{x_k} dt \right] = 2I(I_1 + I_2 + I_3).$$
 (5.53)

Step 2. Let us compute $2II_2$. Utilizing (5.48) again, and noting that

$$\sum_{j,k,j',k'=1}^{n} (b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k}})_{x_{k'}} = \sum_{j,k,j',k'=1}^{n} (b^{jk}b^{j'k'}\ell_{x_{j}}w_{x_{j'}}w_{x_{k'}})_{x_{k}},$$

$$2\sum_{j,k,j',k'=1}^{n}b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k}x_{k'}}$$

$$=\sum_{j,k,j',k'=1}^{n}b^{jk}b^{j'k'}\ell_{x_{j'}}(w_{x_{j}}w_{x_{k}x_{k'}}+w_{x_{k}}w_{x_{j}x_{k'}})=\sum_{j,k,j',k'=1}^{n}b^{jk}b^{j'k'}\ell_{x_{j'}}(w_{x_{j}}w_{x_{k}})_{x_{k'}}$$

$$=\sum_{j,k,j',k'=1}^{n}(b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k}})_{x_{k'}}-\sum_{j,k,j',k'=1}^{n}(b^{jk}b^{j'k'}\ell_{x_{j'}})_{x_{k'}}w_{x_{j}}w_{x_{k}}$$

$$=\sum_{j,k,j',k'=1}^{n}(b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j'}}w_{x_{k'}})_{x_{k}}-\sum_{j,k,j',k'=1}^{n}(b^{jk}b^{j'k'}\ell_{x_{j'}})_{x_{k'}}w_{x_{j}}w_{x_{k}}.$$

$$(5.54)$$

Hence, by (5.54), and noting that

$$\sum_{j,k,j',k'=1}^n b^{jk} (b^{j'k'} \ell_{x_{j'}})_{x_k} w_{x_j} w_{x_{k'}} = \sum_{j,k,j',k'=1}^n b^{jk'} (b^{j'k} \ell_{x_{j'}})_{x_{k'}} w_{x_j} w_{x_k},$$

we obtain that

$$4\left[-\sum_{j,k=1}^{n}(b^{jk}w_{x_{j}})_{x_{k}}+\mathcal{A}w\right]\sum_{j,k=1}^{n}b^{jk}\ell_{x_{j}}w_{x_{k}}$$

$$=-4\sum_{j,k,j',k'=1}^{n}(b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k'}})_{x_{k}}+4\sum_{j,k,j',k'=1}^{n}b^{jk}(b^{j'k'}\ell_{x_{j'}})_{x_{k}}w_{x_{j}}w_{x_{k'}}$$

$$+4\sum_{j,k,j',k'=1}^{n}b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k}x_{k'}}+2\mathcal{A}\sum_{j,k=1}^{n}b^{jk}\ell_{x_{j}}(w^{2})_{x_{k}}$$

$$=-2\sum_{j,k=1}^{n}\left[\sum_{j',k'=1}^{n}\left(2b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k'}}-b^{jk}b^{j'k'}\ell_{x_{j}}w_{x_{j'}}w_{x_{k'}}\right)-\mathcal{A}b^{jk}\ell_{x_{j}}w^{2}\right]_{x_{k}}$$

$$+2\sum_{j,k,j',k'=1}^{n}\left[2b^{jk'}(b^{j'k}\ell_{x_{j'}})_{x_{k'}}-(b^{jk}b^{j'k'}\ell_{x_{j'}})_{x_{k'}}\right]w_{x_{j}}w_{x_{k}}-2\sum_{j,k=1}^{n}\left(\mathcal{A}b^{jk}\ell_{x_{j}})_{x_{k}}w^{2}\right).$$

Using Itô's formula, we have

$$2\left[-\sum_{j,k=1}^{n}(b^{jk}w_{x_{j}})_{x_{k}} + \mathcal{A}w\right]dw$$

$$= -2\sum_{j,k=1}^{n}(b^{jk}w_{x_{j}}dw)_{x_{k}} + 2\sum_{j,k=1}^{n}b^{jk}w_{x_{j}}dw_{x_{k}} + 2\mathcal{A}wdw$$

$$= -2\sum_{j,k=1}^{n}(b^{jk}w_{x_{j}}dw)_{x_{k}} + d\left(\sum_{j,k=1}^{n}b^{jk}w_{x_{j}}w_{x_{k}} + \mathcal{A}w^{2}\right)$$

$$-\sum_{j,k=1}^{n}b^{jk}_{t}w_{x_{j}}w_{x_{k}}dt - \mathcal{A}_{t}w^{2}dt - \sum_{j,k=1}^{n}b^{jk}dw_{x_{j}}dw_{x_{k}} - \mathcal{A}(dw)^{2}.$$
(5.56)

Now, from (5.52), (5.55) and (5.56), we arrive at

$$2II_{2} = -2\sum_{j,k=1}^{n} \left[\sum_{j',k'=1}^{n} \left(2b^{jk}b^{j'k'}\ell_{x_{j'}}w_{x_{j}}w_{x_{k'}} - b^{jk}b^{j'k'}\ell_{x_{j}}w_{x_{j'}}w_{x_{k'}} \right) - \mathcal{A}b^{jk}\ell_{x_{j}}w^{2} \right]_{x_{k}} dt$$

$$-2\sum_{j,k=1}^{n} \left(b^{jk}w_{x_{j}}dw \right)_{x_{k}} + d\left(\sum_{j,k=1}^{n} b^{jk}w_{x_{j}}w_{x_{k}} + \mathcal{A}w^{2} \right)$$

$$+2\sum_{j,k=1}^{n} \left[\sum_{j',k'=1}^{n} \left(2b^{jk'}(b^{j'k}\ell_{x_{j'}})_{x_{k'}} - (b^{jk}b^{j'k'}\ell_{x_{j'}})_{x_{k'}} \right) - \frac{b_{t}^{jk}}{2} \right] w_{x_{j}}w_{x_{k}} dt$$

$$-\left[\mathcal{A}_{t} + 2\sum_{j,k=1}^{n} \left(\mathcal{A}b^{jk}\ell_{x_{j}} \right)_{x_{k}} \right] w^{2} dt - \sum_{j,k=1}^{n} b^{jk} dw_{x_{j}} dw_{x_{k}} - \mathcal{A}(dw)^{2}.$$

$$(5.57)$$

Step 3. Let us compute $2II_3$. By (5.52), we get

$$2II_{3} = 2\left[-\sum_{j,k=1}^{n} (b^{jk}w_{x_{j}})_{x_{k}} + \mathcal{A}w\right]\Psi w dt$$

$$= \left[-2\sum_{j,k=1}^{n} (\Psi b^{jk}w_{x_{j}}w)_{x_{k}} + 2\Psi \sum_{j,k=1}^{n} b^{jk}w_{x_{j}}w_{x_{k}} + \sum_{j,k=1}^{n} b^{jk}\Psi_{x_{k}}(w^{2})_{x_{j}} + 2\mathcal{A}\Psi w^{2}\right] dt$$

$$= \left\{-\sum_{j,k=1}^{n} \left(2\Psi b^{jk}w_{x_{j}}w - b^{jk}\Psi_{x_{j}}w^{2}\right)_{x_{k}} + 2\Psi \sum_{j,k=1}^{n} b^{jk}w_{x_{j}}w_{x_{k}} + \left[-\sum_{j,k=1}^{n} (b^{jk}\Psi_{x_{k}})_{x_{j}} + 2\mathcal{A}\Psi\right]w^{2}\right\} dt.$$

$$(5.58)$$

Step 4. Finally, combining the equalities (5.53), (5.57) and (5.58), and noting that

$$\sum_{j,k=1}^{n} b^{jk} dw_{x_j} dw_{x_k} + \mathcal{A}(dw)^2 = \theta^2 \sum_{j,k=1}^{n} b^{jk} (dh_{x_j} + \ell_{x_j} dh) (dh_{x_k} + \ell_{x_k} dh) + \theta^2 \mathcal{A}(dh)^2,$$

we conclude the desired equality (5.50) immediately.

Next, we shall derive a Carleman estimate for the stochastic parabolic-like operator " $dh - \sum_{j,k=1}^{n} (b^{jk} h_{x_j})_{x_k} dt$ ",

For any fixed nonnegative and nonzero function $\psi \in C^4(\overline{G})$, and (large) parameters $\lambda > 1$ and $\mu > 1$, we choose

$$\theta = e^{\ell}, \quad \ell = \lambda \alpha, \quad \alpha(t, x) = \frac{e^{\mu \psi(x)} - e^{2\mu |\psi|_{C(\overline{G})}}}{t(T - t)}, \quad \varphi(t, x) = \frac{e^{\mu \psi(x)}}{t(T - t)}, \tag{5.59}$$

and

$$\Psi = 2\sum_{j,k=1}^{n} b^{jk} \ell_{x_j x_k}.$$
 (5.60)

In what follows, for a positive integer r, we denote by $O(\mu^r)$ a function of order μ^r for large μ (which is independent of λ); by $O_{\mu}(\lambda^r)$ a function of order λ^r for fixed μ and for large λ . In a similar way, we use the notation $O(e^{\mu|\psi|_{C(\overline{G})}})$ and so on. For $j, k = 1, 2, \dots, n$, it is easy to check that

$$\ell_t = \lambda \alpha_t, \quad \ell_{x_i} = \lambda \mu \varphi \psi_{x_i}, \quad \ell_{x_i x_k} = \lambda \mu^2 \varphi \psi_{x_i} \psi_{x_k} + \lambda \mu \varphi \psi_{x_i x_k}$$
 (5.61)

and that

$$\alpha_t = \varphi^2 O(e^{2\mu|\psi|_{C(\overline{G})}}), \qquad \varphi_t = \varphi^2 O(e^{\mu|\psi|_{C(\overline{G})}}). \tag{5.62}$$

We have the following result.

Theorem 5.6 Assume that either $(b^{jk})_{n\times n}$ or $-(b^{jk})_{n\times n}$ is a uniformly positive definite matrix, and its smallest eigenvalue is bigger than a constant $s_0 > 0$. Let h and $w = \theta h$ be that in Theorem 5.5 with θ being given in (5.59). Then, the equality (5.50) holds for any $t \in [0,T]$ and a.e. $(x,\omega) \in G \times \Omega$. Moreover, for A, B and c^{jk} appeared in (5.50) (and given by (5.49)), when $|\nabla \psi(x)| > 0$, λ and μ are large enough, it holds that

$$\begin{cases}
A = -\lambda^{2} \mu^{2} \varphi^{2} \sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}} \psi_{x_{k}} + \lambda \varphi^{2} O(e^{2\mu|\psi|_{C(\overline{G})}}), \\
B \ge 2s_{0}^{2} \lambda^{3} \mu^{4} \varphi^{3} |\nabla \psi|^{4} + \lambda^{3} \varphi^{3} O(\mu^{3}) + \lambda^{2} \varphi^{3} O(\mu^{2} e^{2\mu|\psi|_{C(\overline{G})}}) + \lambda \varphi^{3} O(e^{2\mu|\psi|_{C(\overline{G})}}), \\
\sum_{j,k=1}^{n} c^{jk} w_{x_{j}} w_{x_{k}} \ge [s_{0}^{2} \lambda \mu^{2} \varphi |\nabla \psi|^{2} + \lambda \varphi O(\mu)] |\nabla w|^{2}
\end{cases} (5.63)$$

for any $t \in [0, T]$, \mathbb{P} -a.s.

Proof: By Theorem 5.5, it remains to prove the estimates in (5.63). Noting (5.60)–(5.61), from (5.49), we have $\ell_{x_jx_k} = \lambda \mu^2 \varphi \psi_{x_j} \psi_{x_k} + \lambda \varphi O(\mu)$ and that

$$\begin{split} &\sum_{j,k=1}^{n} c^{jk} w_{x_{j}} w_{x_{k}} \\ &= \sum_{j,k=1}^{n} \Big\{ \sum_{j',k'=1}^{n} \Big[2b^{jk'} b^{j'k} \ell_{x_{j'}x_{k'}} + b^{jk} b^{j'k'} \ell_{x_{j'}x_{k'}} + 2b^{jk'} b^{j'k}_{x_{k'}} \ell_{x_{j'}} - (b^{jk} b^{j'k'})_{x_{k'}} \ell_{x_{j'}} \Big] - \frac{b^{jk}_{t}}{2} \Big\} w_{x_{j}} w_{x_{k}} \\ &= \sum_{j,k=1}^{n} \Big\{ \sum_{j',k'=1}^{n} \Big[2\lambda \mu^{2} \varphi b^{jk'} b^{j'k} \psi_{x_{j'}} \psi_{x_{k'}} + \lambda \mu^{2} \varphi b^{jk} b^{j'k'} \psi_{x_{j'}} \psi_{x_{k'}} + \lambda \varphi O(\mu) \Big] \Big\} w_{x_{j}} w_{x_{k}} \\ &= 2\lambda \mu^{2} \varphi \Big(\sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}} w_{x_{k}} \Big)^{2} + \lambda \mu^{2} \varphi \Big(\sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}} \psi_{x_{k}} \Big) \Big(\sum_{j,k=1}^{n} b^{jk} w_{x_{j}} w_{x_{k}} \Big) + \lambda \varphi |\nabla w|^{2} O(\mu) \\ &\geq \left[s_{0}^{2} \lambda \mu^{2} \varphi |\nabla \psi|^{2} + \lambda \varphi O(\mu) \right] |\nabla w|^{2}, \end{split}$$

which gives the last inequality in (5.63).

Similarly, by the definition of \mathcal{A} in (5.49), and noting (5.62), we see that

$$\begin{split} \mathcal{A} &= -\sum_{j,k=1}^{n} \left(b^{jk} \ell_{x_j} \ell_{x_k} - b^{jk}_{x_k} \ell_{x_j} + b^{jk} \ell_{x_j x_k} \right) - \ell_t \\ &= -\lambda \mu \sum_{j,k=1}^{n} \left[b^{jk} \lambda \mu \varphi^2 \psi_{x_j} \psi_{x_k} - b^{jk}_{x_k} \varphi \psi_{x_j} + b^{jk} \left(\mu \varphi \psi_{x_j} \psi_{x_k} + \varphi \psi_{x_j x_k} \right) \right] + \lambda \varphi^2 O(e^{2\mu |\psi|_{C(\overline{G})}}) \\ &= -\lambda^2 \mu^2 \varphi^2 \sum_{j,k=1}^{n} b^{jk} \psi_{x_j} \psi_{x_k} + \lambda \varphi^2 O(e^{2\mu |\psi|_{C(\overline{G})}}). \end{split}$$

Hence, we get the first estimate in (5.63).

Now, let us estimate \mathcal{B} (recall (5.49) for the definition of \mathcal{B}). For this, by (5.61), and recalling the definitions of Ψ (in (5.60)), we see that

$$\begin{split} \Psi &= 2\lambda\mu \sum_{j,k=1}^{n} b^{jk} (\mu\varphi\psi_{x_{j}}\psi_{x_{k}} + \varphi\psi_{x_{j}x_{k}}) = 2\lambda\mu^{2}\varphi \sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}}\psi_{x_{k}} + \lambda\varphi O(\mu); \\ &\qquad \qquad \ell_{x_{j'}x_{k'}x_{k}} = \lambda\mu^{3}\varphi\psi_{x_{j'}}\psi_{x_{k'}}\psi_{x_{k}} + \lambda\varphi O(\mu^{2}), \\ &\qquad \qquad \ell_{x_{j'}x_{k'}x_{j}x_{k}} = \lambda\mu^{4}\varphi\psi_{x_{j'}}\psi_{x_{k'}}\psi_{x_{j}}\psi_{x_{k}} + \lambda\varphi O(\mu^{3}), \\ &\qquad \qquad \Psi_{x_{k}} = 2\sum_{j',k'=1}^{n} \left(b^{j'k'}\ell_{x_{j'}x_{k'}}\right)_{x_{k}} = 2\sum_{j',k'=1}^{n} \left(b^{j'k'}\ell_{x_{j'}x_{k'}} + b^{j'k'}\ell_{x_{j'}x_{k'}} + b^{j'k'}\ell_{x_{j'}x_{k'}x_{k}}\right) \\ &\qquad \qquad = 2\lambda\mu^{3}\varphi \sum_{j',k'=1}^{n} b^{j'k'}\psi_{x_{j'}}\psi_{x_{k'}}\psi_{x_{k}} + \lambda\varphi O(\mu^{2}), \\ &\qquad \qquad \Psi_{x_{j}x_{k}} = 2\sum_{j',k'=1}^{n} \left(b^{j'k'}\ell_{x_{j'}x_{k'}} + b^{j'k'}\ell_{x_{j'}x_{k'}x_{j}x_{k}} + 2b^{j'k'}\ell_{x_{j'}x_{k'}x_{j}}\right) \\ &\qquad \qquad = 2\lambda\mu^{4}\varphi \sum_{j',k'=1}^{n} b^{j'k'}\psi_{x_{j'}}\psi_{x_{k'}}\psi_{x_{j}}\psi_{x_{k}} + \lambda\varphi O(\mu^{3}), \\ &\qquad \qquad \qquad - \sum_{j,k=1}^{n} \left(b^{jk}\Psi_{x_{k}}\right)_{x_{j}} = -\sum_{j,k=1}^{n} \left(b^{jk}\Psi_{x_{k}} + b^{jk}\Psi_{x_{j}x_{k}}\right) = -2\lambda\mu^{4}\varphi \left(\sum_{j,k=1}^{n} b^{jk}\psi_{x_{j}}\psi_{x_{k}}\right)^{2} + \lambda\varphi O(\mu^{3}). \end{split}$$

Hence, recalling the definition of A (in (5.49)), and using (5.61) and (5.62), we have that

$$\mathcal{A}\Psi = -2\lambda^{3}\mu^{4}\varphi^{3} \left(\sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}} \psi_{x_{k}}\right)^{2} + \lambda^{3}\varphi^{3} O(\mu^{3}) + \lambda^{2}\varphi^{3} O(\mu^{2} e^{2\mu|\psi|_{C(\overline{G})}}),$$

$$\begin{split} \mathcal{A}_{x_{k}} &= -\sum_{j',k'=1}^{n} \left(b_{x_{k}}^{j'k'} \ell_{x_{j'}} \ell_{x_{k'}} + 2b^{j'k'} \ell_{x_{j'}} \ell_{x_{k'}x_{k}} - b_{x_{k'}x_{k}}^{j'k'} \ell_{x_{j'}} - b_{x_{k'}}^{j'k'} \ell_{x_{j'}x_{k}} \right. \\ &\quad + b_{x_{k}}^{j'k'} \ell_{x_{j'}x_{k'}} + b^{j'k'} \ell_{x_{j'}x_{k'}x_{k}} \right) - \ell_{tx_{k}} \\ &= -\sum_{j',k'=1}^{n} \left(2b^{j'k'} \ell_{x_{j'}} \ell_{x_{k'}x_{k}} + b^{j'k'} \ell_{x_{j'}x_{k'}x_{k}} \right) - \ell_{tx_{k}} + \left(\lambda \varphi + \lambda^{2} \varphi^{2} \right) O(\mu^{2}) \\ &= -2\lambda^{2} \mu^{3} \varphi^{2} \sum_{j',k'=1}^{n} b^{j'k'} \psi_{x_{j'}} \psi_{x_{k'}} \psi_{x_{k}} + \lambda^{2} \varphi^{2} O(\mu^{2}) + \lambda \varphi^{2} O(\mu e^{2\mu|\psi|_{C(\overline{G})}}), \\ \sum_{j,k=1}^{n} \mathcal{A}_{x_{k}} b^{jk} \ell_{x_{j}} = -2\lambda^{3} \mu^{4} \varphi^{3} \left(\sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}} \psi_{x_{k}} \right)^{2} + \lambda^{3} \varphi^{3} O(\mu^{3}) + \lambda^{2} \varphi^{3} O(\mu^{2} e^{2\mu|\psi|_{C(\overline{G})}}), \\ \sum_{j,k=1}^{n} \left(\mathcal{A} b^{jk} \ell_{x_{j}} \right)_{x_{k}} = \sum_{j,k=1}^{n} \mathcal{A}_{x_{k}} b^{jk} \ell_{x_{j}} + \mathcal{A} \sum_{j,k=1}^{n} \left(b_{x_{k}}^{jk} \ell_{x_{j}} + b^{jk} \ell_{x_{j}x_{k}} \right) \\ = -3\lambda^{3} \mu^{4} \varphi^{3} \left(\sum_{j,k=1}^{n} b^{jk} \psi_{x_{j}} \psi_{x_{k}} \right)^{2} + \lambda^{3} \varphi^{3} O(\mu^{3}) + \lambda^{2} \varphi^{3} O(\mu^{2} e^{2\mu|\psi|_{C(\overline{G})}}), \end{split}$$

and that

$$\mathcal{A}_{t} = -\sum_{j,k=1}^{n} \left(b^{jk} \ell_{x_{j}} \ell_{x_{k}} - b^{jk}_{x_{k}} \ell_{x_{j}} + b^{jk} \ell_{x_{j}x_{k}} - \ell_{t} \right)_{t}$$

$$= -\sum_{j,k=1}^{n} \left[b^{jk} \left(\ell_{x_{j}} \ell_{x_{k}} \right)_{t} - b^{jk}_{x_{k}} \ell_{x_{j}t} + b^{jk} \ell_{x_{j}x_{k}t} \right] + \lambda^{2} \varphi^{2} O(\mu^{2}) + \lambda \varphi^{3} O(e^{2\mu|\psi|_{C(\overline{G})}})$$

$$= \lambda^{2} \varphi^{3} O(\mu^{2} e^{2\mu|\psi|_{C(\overline{G})}}) + \lambda \varphi^{3} O(e^{2\mu|\psi|_{C(\overline{G})}}).$$

From the definition of \mathcal{B} (See (5.49)), we have that

$$\begin{split} \mathcal{B} &= -4\lambda^{3}\mu^{4}\varphi^{3}\Big(\sum_{j,k=1}^{n}b^{jk}\psi_{x_{j}}\psi_{x_{k}}\Big)^{2} + \lambda^{3}\varphi^{3}O(\mu^{3}) + \lambda^{2}\varphi^{3}O(\mu^{2}e^{2\mu|\psi|_{C(\overline{G})}}) \\ &+ 6\lambda^{3}\mu^{4}\varphi^{3}\Big(\sum_{j,k=1}^{n}b^{jk}\psi_{x_{j}}\psi_{x_{k}}\Big)^{2} + \lambda^{3}\varphi^{3}O(\mu^{3}) + \lambda^{2}\varphi^{3}O(\mu^{2}e^{2\mu|\psi|_{C(\overline{G})}}) \\ &+ \lambda^{2}\varphi^{3}O(\mu^{2}e^{2\mu|\psi|_{C(\overline{G})}}) + \lambda\varphi^{3}O(e^{2\mu|\psi|_{C(\overline{G})}}) - 2\lambda\mu^{4}\varphi\Big(\sum_{j,k=1}^{n}b^{jk}\psi_{x_{j}}\psi_{x_{k}}\Big)^{2} + \lambda\varphi O(\mu^{3}) \\ &= 2\lambda^{3}\mu^{4}\varphi^{3}\Big(\sum_{ij}b^{jk}\psi_{x_{j}}\psi_{x_{k}}\Big)^{2} + \lambda^{3}\varphi^{3}O(\mu^{3}) + \lambda^{2}\varphi^{3}O(\mu^{2}e^{2\mu|\psi|_{C(\overline{G})}}) + \lambda\varphi^{3}O(e^{2\mu|\psi|_{C(\overline{G})}}), \end{split}$$

which leads to the second estimate in (5.63).

5.3.2 Global Carleman estimate for backward stochastic parabolic equations

As a key preliminary to prove Theorem 5.4, we need to establish a global Carleman estimate for the following backward stochastic parabolic equation:

$$\begin{cases}
dz + \sum_{j,k=1}^{n} (a^{jk} z_{x_j})_{x_k} dt = f dt + Z dW(t) & \text{in } Q, \\
z = 0 & \text{on } \Sigma, \\
z(T) = z_T & \text{in } G.
\end{cases}$$
(5.64)

We begin with the following known technical result (See [3, p. 4, Lemma 1.1] and [9] for its proof), which shows the existence of a nonnegative function with an arbitrarily given critical point location in G.

Lemma 5.5 For any nonempty open subset G_1 of G, there is a $\psi \in C^{\infty}(\overline{G})$ such that $\psi > 0$ in G, $\psi = 0$ on Γ , and $|\nabla \psi(x)| > 0$ for all $x \in \overline{G \setminus G_1}$.

Let us choose θ and ℓ as that in (5.59), and ψ given by Lemma 5.5 with G_1 being any fixed nonempty open subset of G such that $\overline{G_1} \subset G_0$. The desired global Carleman estimate for (5.64) is stated as follows:

Theorem 5.7 There is a constant $\mu_0 = \mu_0(G, G_0, (a^{jk})_{n \times n}, T) > 0$ such that for all $\mu \ge \mu_0$, one can find two constants $\mathcal{C} = \mathcal{C}(\mu) > 0$ and $\lambda_0 = \lambda_0(\mu) > 0$ such that for all $\lambda \ge \lambda_0$, $f \in L^2_{\mathbb{F}}(0, T; L^2(G))$ and $z_T \in L^2_{\mathcal{F}_T}(\Omega; L^2(G))$, the solution $(z, Z) \in C_{\mathbb{F}}([0, T]; L^2(\Omega; L^2(G))) \times L^2_{\mathbb{F}}(0, T; L^2(G))$ to (5.64) satisfies that

$$\lambda^{3} \mu^{4} \mathbb{E} \int_{Q} \theta^{2} \varphi^{3} z^{2} dx dt + \lambda \mu^{2} \mathbb{E} \int_{Q} \theta^{2} \varphi |\nabla z|^{2} dx dt$$

$$\leq \mathcal{C} \Big(\lambda^{3} \mu^{4} \mathbb{E} \int_{Q_{0}} \theta^{2} \varphi^{3} z^{2} dx dt + \mathbb{E} \int_{Q} \theta^{2} f^{2} dx dt + \lambda^{2} \mu^{2} \mathbb{E} \int_{Q} \theta^{2} \varphi^{2} Z^{2} dx dt \Big).$$

$$(5.65)$$

Proof: We use Theorem 5.6 with b^{jk} and h replaced respectively by $-a^{jk}$ and z (and hence $w = \theta z$).

Integrating the equality (5.50) (with b^{jk} replaced by $-a^{jk}$) on G, taking mean value in

both sides, and noting (5.63), we conclude that

$$2\mathbb{E} \int_{Q} \theta \Big[\sum_{j,k=1}^{n} (a^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \Big] \Big[dz + \sum_{j,k=1}^{n} (a^{jk} z_{x_{j}})_{x_{k}} dt \Big] dx - 2\mathbb{E} \int_{Q} \sum_{j,k=1}^{n} (a^{jk} w_{x_{j}} dw)_{x_{k}} dx$$

$$+ 2\mathbb{E} \int_{Q} \sum_{j,k=1}^{n} \Big[\sum_{j',k'=1}^{n} \Big(2a^{jk} a^{j'k'} \ell_{x_{j'}} w_{x_{j}} w_{x_{k'}} - a^{jk} a^{j'k'} \ell_{x_{j}} w_{x_{j'}} w_{x_{k'}} \Big) - \Psi a^{jk} w_{x_{j}} w$$

$$+ a^{jk} \Big(\mathcal{A} \ell_{x_{j}} + \frac{\Psi_{x_{j}}}{2} \Big) w^{2} \Big]_{x_{k}} dx dt$$

$$\geq 2s_{0}^{2} \mathbb{E} \int_{Q} \Big[\varphi \Big(\lambda \mu^{2} |\nabla \psi|^{2} + \lambda O(\mu) \Big) |\nabla w|^{2} + \varphi^{3} \Big(\lambda^{3} \mu^{4} |\nabla \psi|^{4} + \lambda^{3} O(\mu^{3}) \Big)$$

$$+ \lambda^{2} O(\mu^{2} e^{2\mu|\psi|_{C(\overline{G})}}) + \lambda O(e^{2\mu|\psi|_{C(\overline{G})}}) \Big) w^{2} \Big] dx dt + 2\mathbb{E} \int_{Q} \Big| \sum_{j,k=1}^{n} (a^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \Big|^{2} dx dt$$

$$+ \mathbb{E} \int_{Q} \theta^{2} \sum_{j,k=1}^{n} a^{jk} (dz_{x_{j}} + \ell_{x_{j}} dz) (dz_{x_{k}} + \ell_{x_{k}} dz) dx - \mathbb{E} \int_{Q} \theta^{2} \mathcal{A} (dz)^{2} dx,$$

$$(5.66)$$

where

$$\mathcal{A} = \sum_{j,k=1}^{n} (a^{jk} \ell_{x_j} \ell_{x_k} - a^{jk}_{x_k} \ell_{x_j} + a^{jk} \ell_{x_j x_k}) - \ell_t, \qquad \Psi = -2 \sum_{j,k=1}^{n} a^{jk} \ell_{x_j x_k}.$$

It follows from (5.64) that

$$2\mathbb{E} \int_{Q} \theta \Big[\sum_{j,k=1}^{n} (a^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \Big] \Big[dz + \sum_{j,k=1}^{n} (a^{jk} z_{x_{j}})_{x_{k}} dt \Big] dx$$

$$= 2\mathbb{E} \int_{Q} \theta \Big[\sum_{j,k=1}^{n} (a^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \Big] (fdt + ZdW(t)) dx$$

$$= 2\mathbb{E} \int_{Q} \theta \Big[-\sum_{j,k=1}^{n} (a^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \Big] fdt dx$$

$$\leq \mathbb{E} \int_{Q} \Big| \sum_{j,k=1}^{n} (a^{jk} w_{x_{j}})_{x_{k}} + \mathcal{A}w \Big|^{2} dt dx + \mathbb{E} \int_{Q} \theta^{2} f^{2} dt dx.$$

$$(5.67)$$

It is clear that the term " $\mathbb{E} \int_Q \theta^2 \sum_{j,k=1}^n a^{jk} (dz_{x_j} + \ell_{x_j} dz) (dz_{x_k} + \ell_{x_k} dz) dx$ " in (5.66) is nonnegative. Hence, by (5.66)–(5.67), one can show that

$$2s_0^2 \mathbb{E} \int_Q \left[\varphi \left(\lambda \mu^2 |\nabla \psi|^2 + \lambda O(\mu) \right) |\nabla w|^2 + \varphi^3 \left(\lambda^3 \mu^4 |\nabla \psi|^4 + \lambda^3 O(\mu^3) \right) \right. \\ \left. + \lambda^2 O(\mu^2 e^{2\mu|\psi|_{C(\overline{G})}}) + \lambda O(e^{2\mu|\psi|_{C(\overline{G})}}) \right) w^2 \right] dxdt$$

$$\leq \mathbb{E} \int_Q \theta^2 (f^2 + \mathcal{A}Z^2) dxdt.$$

$$(5.68)$$

From (5.68), we conclude that there is a $\mu_0 > 0$ such that for all $\mu \ge \mu_0$, one can find a constant $\lambda_0 = \lambda_0(\mu)$ so that for any $\lambda \ge \lambda_0$, it holds that

$$\lambda \mu^{2} \mathbb{E} \int_{Q} \theta^{2} \varphi \left(|\nabla z|^{2} + \lambda^{2} \mu^{2} \varphi^{2} z^{2} \right) dx dt$$

$$\leq \mathcal{C} \left[\mathbb{E} \int_{Q} \theta^{2} (f^{2} + \lambda^{2} \mu^{2} \varphi^{2} Z^{2}) dx dt + \lambda \mu^{2} \mathbb{E} \int_{0}^{T} \int_{G_{1}} \theta^{2} \varphi \left(|\nabla z|^{2} + \lambda^{2} \mu^{2} \varphi^{2} z^{2} \right) dx dt \right].$$
(5.69)

Choose a cut-off function $\zeta \in C_0^{\infty}(G_0; [0, 1])$ so that $\zeta \equiv 1$ in G_1 . By $d(\theta^2 \varphi h^2) = h^2(\theta^2 \varphi)_t dt + 2\theta^2 \varphi h dh + \theta^2 \varphi (dh)^2$, recalling $\lim_{t\to 0^+} \varphi(t, \cdot) = \lim_{t\to T^-} \varphi(t, \cdot) \equiv 0$ and using (5.64), we find that

$$0 = \mathbb{E} \int_{Q_0} \theta^2 \Big[\zeta^2 z^2 (\varphi_t + 2\lambda \varphi \eta_t) + 2\zeta^2 \varphi \sum_{j,k=1}^n a^{jk} z_{x_j} z_{x_k} + 2\mu \zeta^2 \varphi (1 + 2\lambda \varphi) z \sum_{j,k=1}^n a^{jk} z_{x_j} \psi_{x_k} + 4\zeta \varphi z \sum_{j,k=1}^n a^{jk} z_{x_j} \zeta_{x_k} + 2\zeta^2 \varphi f z + \zeta^2 \varphi Z^2 \Big] dx dt.$$

Therefore, for any $\varepsilon > 0$, one has

$$2\mathbb{E} \int_{Q_0} \theta^2 \zeta^2 \varphi \sum_{j,k=1}^n a^{jk} z_{x_j} z_{x_k} dx dt + \mathbb{E} \int_{Q_0} \theta^2 \zeta^2 \varphi Z^2 dx dt$$

$$\leq \varepsilon \mathbb{E} \int_{Q_0} \theta^2 \zeta^2 \varphi |\nabla z|^2 dx dt + \frac{\mathcal{C}}{\varepsilon} \mathbb{E} \int_{Q_0} \theta^2 \left(\frac{1}{\lambda^2 \mu^2} f^2 + \lambda^2 \mu^2 \varphi^3 z^2\right) dx dt.$$
(5.70)

Since the matrix $(a^{jk})_{1 \leq i,j \leq n}$ is uniformly positive definite, we conclude from (5.70) that

$$\mathbb{E} \int_0^T \int_{G_1} \theta^2 \varphi |\nabla z|^2 dx dt \le \mathcal{C} \mathbb{E} \int_{Q_0} \theta^2 \left(\frac{1}{\lambda^2 \mu^2} f^2 + \lambda^2 \mu^2 \varphi^3 z^2 \right) dx dt. \tag{5.71}$$

Combining (5.69) and (5.71), we obtain (5.65). This completes the proof of Theorem 5.7. $\hfill\Box$

5.3.3 Proof of the observability estimate for backward stochastic parabolic equations

We are now in a position to prove Theorem 5.4.

Proof of Theorem 5.4: Applying Theorem 5.7 to the equation (5.6), we deduce that, for all $\mu \geq \mu_0$ and $\lambda \geq \lambda_0(\mu)$,

$$\lambda^{3}\mu^{4}\mathbb{E}\int_{Q}\theta^{2}\varphi^{3}z^{2}dxdt + \lambda\mu^{2}\mathbb{E}\int_{Q}\theta^{2}\varphi|\nabla z|^{2}dxdt$$

$$\leq \mathcal{C}\Big\{\lambda^{3}\mu^{4}\mathbb{E}\int_{Q_{0}}\theta^{2}\varphi^{3}z^{2}dxdt + \mathbb{E}\int_{Q}\theta^{2}\Big[\sum_{j=1}^{n}\left(a_{1j}z\right)_{x_{j}} - a_{2}z - a_{3}Z\Big]^{2}dxdt + \lambda^{2}\mu^{2}\mathbb{E}\int_{Q}\theta^{2}\varphi^{2}Z^{2}dxdt\Big\}$$

$$\leq \mathcal{C}\Big[\lambda^{3}\mu^{4}\mathbb{E}\int_{Q_{0}}\theta^{2}\varphi^{3}z^{2}dxdt + \mathbb{E}\int_{Q}\theta^{2}\left(|\nabla z|^{2} + \lambda^{2}\mu^{2}\varphi^{2}z^{2} + Z^{2}\right)dxdt + \lambda^{2}\mu^{2}\mathbb{E}\int_{Q}\theta^{2}\varphi^{2}Z^{2}dxdt\Big].$$

$$(5.72)$$

Choosing $\mu = \mu_0$ and $\lambda = \mathcal{C}$, from (5.72), we obtain that

$$\mathbb{E} \int_{Q} \theta^{2} \varphi^{3} z^{2} dx dt \leq \mathcal{C} \Big(\mathbb{E} \int_{Q_{0}} \theta^{2} \varphi^{3} z^{2} dx dt + \mathbb{E} \int_{Q} \theta^{2} \varphi^{2} Z^{2} dx dt \Big).$$
 (5.73)

Recalling (5.59), it follows from (5.73) that

$$\mathbb{E} \int_{T/4}^{3T/4} \int_{G} z^{2} dx dt$$

$$\leq C \frac{\max_{(t,x)\in Q} \left(\theta^{2}(t,x)\varphi^{3}(t,x) + \theta^{2}(t,x)\varphi^{2}(t,x)\right)}{\min_{x\in G} \left(\theta^{2}(T/4,x)\varphi^{3}(T/2,x)\right)} \left(\mathbb{E} \int_{Q_{0}} z^{2} dx dt + \mathbb{E} \int_{Q} Z^{2} dx dt\right) \qquad (5.74)$$

$$\leq C \left(\mathbb{E} \int_{Q_{0}} z^{2} dx dt + \mathbb{E} \int_{Q} Z^{2} dx dt\right).$$

By (5.7) in Proposition 5.1, it follows that

$$\mathbb{E} \int_{G} z^{2}(0)dx \le \mathcal{C}\mathbb{E} \int_{G} z^{2}(t)dx, \qquad \forall t \in [0, T].$$
 (5.75)

Combining (5.74) and (5.75), we conclude that, the solution (z, Z) to the equation (5.6) satisfies (5.47). This completes the proof of Theorem 5.4.

6 Pontryagin-type maximum principle for controlled stochastic evolution equations in infinite dimensions

This section is addressed to studying the first order necessary optimality condition, i.e., Pontryagin-type maximum principle, for optimal control problems for nonlinear stochastic evolution equations in infinite dimensions, in which both drift and diffusion terms can contain the control variables, and the control domain is allowed to be nonconvex. The results in this part are taken from [12, 13].

6.1 Formulation of the problem

In this section, U and $\mathcal{U}[0,T]$ are the same as that in Section 4. To simplify the presentation, we assume that H is a separable, real Hilbert space.

Let us impose the following condition.

(B1) For $\varphi = a, b$, suppose that $\varphi(\cdot, \cdot, \cdot) : [0, T] \times H \times U \to H$ satisfies : i) For any $(x, u) \in H \times U$, the functions $\varphi(\cdot, x, u) : [0, T] \to H$ is Lebesgue measurable; ii) For any $(t, x) \in [0, T] \times H$, the functions $\varphi(t, x, \cdot) : U \to H$ is continuous; and iii) For all $(t, x_1, x_2, u) \in [0, T] \times H \times H \times U$,

$$\begin{cases}
|\varphi(t, x_1, u) - \varphi(t, x_2, u)|_H \le C_L |x_1 - x_2|_H, \\
|\varphi(t, 0, u)|_H + |b(t, 0, u)|_H \le C_L.
\end{cases}$$
(6.1)

Consider the following controlled (forward) stochastic evolution equation:

$$\begin{cases} dx = [Ax + a(t, x, u)]dt + b(t, x, u)dW(t) & \text{in } (0, T], \\ x(0) = x_0, \end{cases}$$
(6.2)

where $u \in \mathcal{U}[0,T]$ and $x_0 \in L^{p_0}_{\mathcal{F}_0}(\Omega;H)$ for a given $p_0 \geq 2$. Under the assumption (B1), one can show that the equation (6.2) is well-posed in the sense of mild solution.

Also, we need the following condition:

(B2) Suppose that $g(\cdot, \cdot, \cdot) : [0, T] \times H \times U \to \mathbb{R}$ and $h(\cdot) : H \to \mathbb{R}$ are two functions satisfying: i) For any $(x, u) \in H \times U$, the function $g(\cdot, x, u) : [0, T] \to \mathbb{R}$ is Lebesgue measurable; ii) For any $(t, x) \in [0, T] \times H$, the function $g(t, x, \cdot) : U \to \mathbb{R}$ is continuous; and iii) For all $(t, x_1, x_2, u) \in [0, T] \times H \times H \times U$,

$$\begin{cases}
|g(t, x_1, u) - g(t, x_2, u)|_H + |h(x_1) - h(x_2)|_H \le C_L |x_1 - x_2|_H, \\
|g(t, 0, u)|_H + |h(0)|_H \le C_L.
\end{cases}$$
(6.3)

Define a cost functional $\mathcal{J}(\cdot)$ (for the controlled system (6.2)) as follows:

$$\mathcal{J}(u(\cdot)) \triangleq \mathbb{E}\Big[\int_0^T g(t, x(t), u(t))dt + h(x(T))\Big], \quad \forall u(\cdot) \in \mathcal{U}[0, T], \tag{6.4}$$

where $x(\cdot)$ is the corresponding solution to (6.2).

Let us consider the following optimal control problem for the system (6.2):

Problem (OP) Find a $\bar{u}(\cdot) \in \mathcal{U}[0,T]$ such that

$$\mathcal{J}(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0,T]} \mathcal{J}(u(\cdot)). \tag{6.5}$$

Any $\bar{u}(\cdot)$ satisfying (6.5) is called an optimal control. The corresponding state process $\bar{x}(\cdot)$ is called an optimal state (process), and $(\bar{x}(\cdot), \bar{u}(\cdot))$ is called an optimal pair.

The main aim of this section is to derive the first order necessary optimality condition, i.e., Pontryagin-type maximum principle, for the above Problem (OP).

As we shall see later, the main difficulty to deal with the case of non-convex control domain U is that one needs to study the following $\mathcal{L}(H)$ -valued backward stochastic evolution equation²:

$$\begin{cases}
dP = -(A^* + J^*)Pdt - P(A + J)dt - K^*PKdt \\
-(K^*Q + QK)dt + Fdt + QdW(t) & \text{in } [0, T), \\
P(T) = P_T.
\end{cases}$$
(6.6)

Here and henceforth, $F \in L^1_{\mathbb{F}}(0,T;L^2(\Omega;\mathcal{L}(H))), P_T \in L^2_{\mathcal{F}_T}(\Omega;\mathcal{L}(H)), J \in L^4_{\mathbb{F}}(0,T;L^{\infty}(\Omega;\mathcal{L}(H)))$ and $K \in L^4_{\mathbb{F}}(0,T;L^{\infty}(\Omega;\mathcal{L}(H)))$. For the special case when $H = \mathbb{R}^n$, it is easy to

²Throughout this section, for any operator-valued process (resp. random variable) R, we denote by R^* its pointwise dual operator-valued process (resp. random variable). For example, 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 $|R|_{L^{r_1}_{\mathbb{F}}(0,T;L^{r_2}(\Omega;\mathcal{L}(H)))} = |R^*|_{L^{r_1}_{\mathbb{F}}(0,T;L^{r_2}(\Omega;\mathcal{L}(H)))}$.

see that (6.6) is an $\mathbb{R}^{n\times n}$ (matrix)-backward stochastic differential equation, and therefore, the desired well-posedness follows from that of an \mathbb{R}^{n^2} (vector)-valued backward stochastic differential equation. However, one has to face a real challenge in the study of (6.6) when dim $H = \infty$, without further assumption on the data F and P_T . Indeed, in the infinite dimensional setting, although $\mathcal{L}(H)$ is still a Banach space, it is neither reflexive (needless to say to be a Hilbert space) nor separable even if H itself is separable (See Problem 99 in [4]). As far as we know, in the previous literatures there exists no such a stochastic integration/evolution equation theory in general Banach spaces that can be employed to treat the well-posedness of (6.6). For example, the existing result on stochastic integration/evolution equation in UMD Banach spaces (e.g. [20]) does not fit the present case because, if a Banach space is UMD, then it is reflexive.

To overcome the above-mentioned difficulty, we employ the stochastic transposition method developed in [11]. More precisely, we introduce a concept of relaxed transposition solution to the equation (6.6), and develop a way to study the corresponding well-posedness. Our method can be further modified to treat the second order necessary conditions for stochastic optimal controls and the feedback control design for linear quadratic stochastic optimal control problems in infinite dimensions but all of these topics are beyond the scope of this short course.

6.2 Pontryagin-type maximum principle for convex control domain

In this subsection, we give a necessary condition for optimal controls of Problem (OP) for the case of special control domain U, i.e., U is a convex subset of another separable Hilbert space \widetilde{H} , and the metric of U is introduced by the norm of \widetilde{H} (i.e., $\mathbf{d}(u_1, u_2) = |u_1 - u_2|_{\widetilde{H}}$).

First, we need to study the following H-valued backward stochastic evolution equation:

$$\begin{cases} dy(t) = -A^*y(t)dt + f(t, y(t), Y(t))dt + Y(t)dW(t) & \text{in } [0, T), \\ y(T) = y_T. \end{cases}$$
(6.7)

Here $y_T \in L^2_{\mathcal{F}_T}(\Omega; H)$, $f(\cdot, \cdot, \cdot) : [0, T] \times H \times H \to H$ satisfies

$$\begin{cases}
f(\cdot,0,0) \in L^{1}_{\mathbb{F}}(0,T;L^{2}(\Omega;H)), \\
|f(t,x_{1},y_{1}) - f(t,x_{2},y_{2})|_{H} \leq C_{L}(|x_{1} - x_{2}|_{H} + |y_{1} - y_{2}|_{H}), \\
\text{a.e. } (t,\omega) \in [0,T] \times \Omega, \quad \forall \ x_{1},x_{2},y_{1},y_{2} \in H.
\end{cases}$$
(6.8)

To define the solution to (6.8), we introduce the following (forward) stochastic evolution equation:

$$\begin{cases} dz = (A^*z + v_1)ds + v_2 dW(s) & \text{in } (t, T], \\ z(t) = \eta, \end{cases}$$
(6.9)

where $t \in [0,T]$, $v_1 \in L^1_{\mathbb{F}}(t,T;L^2(\Omega;H))$, $v_2 \in L^2_{\mathbb{F}}(t,T;H)$, $\eta \in L^2_{\mathcal{F}_t}(\Omega;H)$. The equation (6.9) admits a unique mild solution $z \in C_{\mathbb{F}}([t,T];L^2(\Omega;H))$, and

$$|z|_{C_{\mathbb{F}}([t,T];L^{2}(\Omega;H))} \le \mathcal{C}(|\eta|_{L^{2}_{\mathcal{F}_{t}}(\Omega;H)} + |v_{1}|_{L^{1}_{\mathbb{F}}(t,T;L^{2}(\Omega;H))} + |v_{2}|_{L^{2}_{\mathbb{F}}(t,T;H)}). \tag{6.10}$$

We now introduce the following notion.

Definition 6.1 We call $(y(\cdot), Y(\cdot)) \in D_{\mathbb{F}}([0,T]; L^2(\Omega; H)) \times L^2_{\mathbb{F}}(0,T; H)$ a transposition solution to (6.7) if for any $t \in [0,T]$, $v_1(\cdot) \in L^1_{\mathbb{F}}(t,T; L^2(\Omega; H))$, $v_2(\cdot) \in L^2_{\mathbb{F}}(t,T; H)$, $\eta \in L^2_{\mathcal{F}_t}(\Omega; H)$ and the corresponding solution $z \in C_{\mathbb{F}}([t,T]; L^2(\Omega; H))$ to (6.9), it holds that

$$\mathbb{E}\langle z(T), y_T \rangle_H - \mathbb{E} \int_t^T \langle z(s), F(s, y(s), Y(s)) \rangle_H ds$$

$$= \mathbb{E}\langle \eta, y(t) \rangle_H + \mathbb{E} \int_t^T \langle v_1(s), y(s) \rangle_H ds + \mathbb{E} \int_t^T \langle v_2(s), Y(s) \rangle_H ds.$$
(6.11)

We have the following well-posedness result for (6.7) (See [12] for its proof).

Theorem 6.1 For any $y_T \in L^2_{\mathcal{F}_T}(\Omega; H)$ and $f(\cdot, 0, 0) \in L^1_{\mathbb{F}}(0, T; L^2(\Omega; H))$, the equation (6.7) admits a unique transposition solution $(y(\cdot), Y(\cdot)) \in D_{\mathbb{F}}([0, T]; L^2(\Omega; H)) \times L^2_{\mathbb{F}}(0, T; H)$. Furthermore,

$$|(y(\cdot), Y(\cdot))|_{D_{\mathbb{F}}([0,T];L^{p}(\Omega;H)) \times L^{2}_{\mathbb{F}}(0,T;H)} \le \mathcal{C}(|y_{T}|_{L^{p}_{\mathcal{F}_{T}}(\Omega;H)} + |f(\cdot,0,0)|_{L^{1}_{\mathbb{F}}(0,T;L^{2}(\Omega;H))}). \tag{6.12}$$

We introduce the following further assumptions for $a(\cdot,\cdot,\cdot)$, $b(\cdot,\cdot,\cdot)$, $g(\cdot,\cdot,\cdot)$ and $b(\cdot)$.

(B3) The functions a(t, x, u) and b(t, x, u), and the functional g(t, x, u) and h(x) are C^1 with respect to x and u. Moreover, for any $(t, x, u) \in [0, T] \times H \times U$,

$$\begin{cases}
\|a_x(t,x,u)\|_{\mathcal{L}(H)} + \|b_x(t,x,u)\|_{\mathcal{L}(H)} + |g_x(t,x,u)|_H + |h_x(x)|_H \le C_L, \\
\|a_u(t,x,u)\|_{\mathcal{L}(\widetilde{H},H)} + \|b_u(t,x,u)\|_{\mathcal{L}(\widetilde{H},H)} + |g_u(t,x,u)|_{\widetilde{H}} \le C_L.
\end{cases}$$
(6.13)

We have the following result.

Theorem 6.2 Assume that $x_0 \in L^2_{\mathcal{F}_0}(\Omega; H)$. Let the assumptions (B1), (B2) and (B3) hold, and let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be an optimal pair of Problem (OP). Let $(y(\cdot), Y(\cdot))$ be the transposition solution to the equation (6.7) with y_T and $f(\cdot, \cdot, \cdot)$ given by

$$\begin{cases}
y_T = -h_x(\bar{x}(T)), \\
f(t, y_1, y_2) = -a_x(t, \bar{x}(t), \bar{u}(t))^* y_1 - b_x(t, \bar{x}(t), \bar{u}(t))^* y_2 + g_x(t, \bar{x}(t), \bar{u}(t)).
\end{cases} (6.14)$$

Then,

$$\langle a_u(t, \bar{x}(t), \bar{u}(t))^* y(t) + b_u(t, \bar{x}(t), \bar{u}(t))^* Y(t) - g_u(t, \bar{u}(t), \bar{x}(t)), u - \bar{u}(t) \rangle_{\widetilde{H}} \leq 0,$$
a.e. $(t, \omega) \in [0, T] \times \Omega, \quad \forall \ u \in U.$ (6.15)

To prove Theorem 6.2, we need the following result.

Lemma 6.1 If $F(\cdot) \in L^2_{\mathbb{F}}(0,T;\widetilde{H})$ and $\bar{u}(\cdot) \in \mathcal{U}[0,T]$ such that

$$\mathbb{E} \int_0^T \left\langle F(t,\cdot), u(t,\cdot) - \bar{u}(t,\cdot) \right\rangle_{\widetilde{H}} dt \le 0, \tag{6.16}$$

holds for any $u(\cdot) \in \mathcal{U}[0,T]$ satisfying $u(\cdot) - \bar{u}(\cdot) \in L^2_{\mathbb{F}}(0,T;L^2(\Omega;\widetilde{H}))$, then, for any point $u \in U$, the following pointwise inequality holds:

$$\langle F(t,\omega), u - \bar{u}(t,\omega) \rangle_{\widetilde{H}} \le 0, \text{ a.e. } (t,\omega) \in [0,T] \times \Omega.$$
 (6.17)

Proof: We use the contradiction argument. Suppose that the inequality (6.17) did not hold. Then, there would exist a $u_0 \in U$ and an $\varepsilon > 0$ such that

$$\alpha_{\varepsilon} \stackrel{\triangle}{=} \int_{\Omega} \int_{0}^{T} \chi_{\Lambda_{\varepsilon}}(t,\omega) dt d\mathbb{P} > 0,$$

where $\Lambda_{\varepsilon} \triangleq \Big\{ (t, \omega) \in [0, T] \times \Omega : \operatorname{Re} \big\langle F(t, \omega), u_0 - \bar{u}(t, \omega) \big\rangle_{\widetilde{H}} \geq \varepsilon \Big\}$, and $\chi_{\Lambda_{\varepsilon}}$ is the characteristic function of Λ_{ε} . For any $m \in \mathbb{N}$, define $\Lambda_{\varepsilon,m} \stackrel{\triangle}{=} \Lambda_{\varepsilon} \cap \big\{ (t, \omega) \in [0, T] \times \Omega \ \big| \ |\bar{u}(t, \omega)|_{H_1} \leq m \big\}$. It is clear that $\lim_{m \to \infty} \Lambda_{\varepsilon,m} = \Lambda_{\varepsilon}$. Hence, there is an $m_{\varepsilon} \in \mathbb{N}$ such that

$$\int_{\Omega} \int_{0}^{T} \chi_{\Lambda_{\varepsilon,m}}(t,\omega) dt d\mathbb{P} > \frac{\alpha_{\varepsilon}}{2} > 0, \qquad \forall \ m \ge m_{\varepsilon}.$$

Since $\langle F(\cdot), u_0 - \bar{u}(\cdot) \rangle_{\widetilde{H}}$ is **F**-adapted, so is the process $\chi_{\Lambda_{\varepsilon,m}}(\cdot)$. Define

$$\hat{u}_{\varepsilon,m}(t,\omega) = u_0 \chi_{\Lambda_{\varepsilon,m}}(t,\omega) + \bar{u}(t,\omega) \chi_{\Lambda_{\varepsilon,m}^c}(t,\omega), \quad (t,\omega) \in [0,T] \times \Omega.$$

Noting that $|\bar{u}(\cdot)|_{\widetilde{H}} \leq m$ on $\Lambda_{\varepsilon,m}$, we see that $\hat{u}_{\varepsilon,m}(\cdot) \in \mathcal{U}[0,T]$ and satisfies $\hat{u}_{\varepsilon,m}(\cdot) - \bar{u}(\cdot) \in L^2_{\mathbb{F}}(0,T;\widetilde{H})$. Hence, for any $m \geq m_{\varepsilon}$, we obtain that

$$\mathbb{E} \int_{0}^{T} \left\langle F(t), \hat{u}_{\varepsilon,m}(t) - \bar{u}(t) \right\rangle_{\widetilde{H}} dt = \int_{\Omega} \int_{0}^{T} \chi_{\Lambda_{\varepsilon,m}}(t,\omega) \left\langle F(t,\omega), u_{0} - \bar{u}(t,\omega) \right\rangle_{\widetilde{H}} dt d\mathbb{P}$$

$$\geq \varepsilon \int_{\Omega} \int_{0}^{T} \chi_{\Lambda_{\varepsilon,m}}(t,\omega) dt d\mathbb{P} \geq \frac{\varepsilon \alpha_{\varepsilon}}{2} > 0,$$

which contradicts (6.16). This completes the proof of Lemma 6.1.

We are now in a position to prove Theorem 6.2.

Proof of Theorem 6.2: We use the convex perturbation technique and divide the proof into several steps.

Step 1. For the optimal pair $(\bar{x}(\cdot), \bar{u}(\cdot))$, we fix arbitrarily a control $u(\cdot) \in \mathcal{U}[0, T]$ satisfying $u(\cdot) - \bar{u}(\cdot) \in L^2_{\mathbb{R}}(0, T; L^2(\Omega; \widetilde{H}))$. Since U is convex, we see that

$$u^{\varepsilon}(\cdot) = \bar{u}(\cdot) + \varepsilon[u(\cdot) - \bar{u}(\cdot)] = (1 - \varepsilon)\bar{u}(\cdot) + \varepsilon u(\cdot) \in \mathcal{U}[0, T], \quad \forall \ \varepsilon \in [0, 1].$$

Denote by $x^{\varepsilon}(\cdot)$ the state process of (6.2) corresponding to the control $u^{\varepsilon}(\cdot)$. It is easy to show that

$$|x^{\varepsilon}|_{C_{\mathbb{F}}(0,T;L^{2}(\Omega;H))} \le \mathcal{C}\left(1+|x_{0}|_{L^{2}_{\mathcal{F}_{0}}(\Omega;H)}\right), \quad \forall \ \varepsilon \in [0,1].$$

$$(6.18)$$

Write $x_1^{\varepsilon}(\cdot) = \frac{1}{\varepsilon} \left[x^{\varepsilon}(\cdot) - \bar{x}(\cdot) \right]$ and $\delta u(\cdot) = u(\cdot) - \bar{u}(\cdot)$. Since $(\bar{x}(\cdot), \bar{u}(\cdot))$ satisfies (6.2), it is easy to see that $x_1^{\varepsilon}(\cdot)$ satisfies the following stochastic differential equation:

$$\begin{cases}
dx_1^{\varepsilon} = \left(Ax_1^{\varepsilon} + a_1^{\varepsilon}x_1^{\varepsilon} + a_2^{\varepsilon}\delta u\right)dt + \left(b_1^{\varepsilon}x_1^{\varepsilon} + b_2^{\varepsilon}\delta u\right)dW(t) & \text{in } (0, T], \\
x_1^{\varepsilon}(0) = 0,
\end{cases}$$
(6.19)

where for $\varphi = a, b$,

$$\varphi_1^{\varepsilon}(t) = \int_0^1 \varphi_x(t, \bar{x}(t) + \sigma \varepsilon x_1^{\varepsilon}(t), u^{\varepsilon}(t)) d\sigma, \quad \varphi_2^{\varepsilon}(t) = \int_0^1 \varphi_u(t, \bar{x}(t), \bar{u}(t) + \sigma \varepsilon \delta u(t)) d\sigma. \quad (6.20)$$

Consider the following stochastic differential equation:

$$\begin{cases} dx_2 = \left[Ax_2 + a_1(t)x_2 + a_2(t)\delta u \right] dt + \left[b_1(t)x_2 + b_2(t)\delta u \right] dW(t) & \text{in } (0, T], \\ x_2(0) = 0, \end{cases}$$
(6.21)

where for $\varphi = a, b$,

$$\varphi_1(t) = a_x(t, \bar{x}(t), \bar{u}(t)), \quad \varphi_2(t) = a_u(t, \bar{x}(t), \bar{u}(t)).$$
 (6.22)

Step 2. In this step, we shall show that

$$\lim_{\varepsilon \to 0\perp} \left| x_1^{\varepsilon} - x_2 \right|_{L_{\mathbb{F}}^{\infty}(0,T;L^2(\Omega;H))} = 0. \tag{6.23}$$

First, using Burkholder-Davis-Gundy's inequality (See Theorem 2.4) and by the assumption (B1), we find that

$$\mathbb{E}|x_{1}^{\varepsilon}(t)|_{H}^{2} = \mathbb{E}\Big|\int_{0}^{t} S(t-s)a_{1}^{\varepsilon}(s)x_{1}^{\varepsilon}(s)ds + \int_{0}^{t} S(t-s)a_{2}^{\varepsilon}(s)\delta u(s)ds + \int_{0}^{t} S(t-s)b_{1}^{\varepsilon}(s)x_{1}^{\varepsilon}(s)dw(s) + \int_{0}^{t} S(t-s)b_{2}^{\varepsilon}(s)\delta u(s)dw(s)\Big|_{H}^{2}$$

$$\leq C\mathbb{E}\Big[\Big|\int_{0}^{t} S(t-s)a_{1}^{\varepsilon}(s)x_{1}^{\varepsilon}(s)ds\Big|_{H}^{2} + \Big|\int_{0}^{t} S(t-s)b_{1}^{\varepsilon}(s)x_{1}^{\varepsilon}(s)dw(s)\Big|_{H}^{2} + \Big|\int_{0}^{t} S(t-s)b_{2}^{\varepsilon}(s)\delta u(s)dw(s)\Big|_{H}^{2}\Big]$$

$$+\Big|\int_{0}^{t} S(t-s)a_{2}^{\varepsilon}(s)\delta u(s)ds\Big|_{H}^{2} + \Big|\int_{0}^{t} S(t-s)b_{2}^{\varepsilon}(s)\delta u(s)dw(s)\Big|_{H}^{2}\Big]$$

$$\leq C\Big[\int_{0}^{t} \mathbb{E}|x_{1}^{\varepsilon}(s)|_{H}^{2}ds + \int_{0}^{T} \mathbb{E}|\delta u(s)|_{H_{1}}^{2}dt\Big].$$

It follows from (6.24) and Gronwall's inequality that

$$\mathbb{E}|x_1^{\varepsilon}(t)|_H^2 \le C|\bar{u} - u|_{L_{\mathbb{F}}^2(0,T;H_1)}^2, \quad \forall \ t \in [0,T].$$
(6.25)

By a similar computation, we see that

$$\mathbb{E}|x_2(t)|_H^2 \le \mathcal{C}|\bar{u} - u|_{L_{\mathbb{E}}^2(0,T;H_1)}^2, \quad \forall \ t \in [0,T].$$
(6.26)

On the other hand, put $x_3^{\varepsilon} = x_1^{\varepsilon} - x_2$. Then, x_3^{ε} solves the following equation:

$$\begin{cases}
dx_3^{\varepsilon} = \left[Ax_3^{\varepsilon} + a_1^{\varepsilon}(t)x_3^{\varepsilon} + \left(a_1^{\varepsilon}(t) - a_1(t) \right) x_2 + \left(a_2^{\varepsilon}(t) - a_2(t) \right) \delta u \right] dt \\
+ \left[b_1^{\varepsilon}(t)x_3^{\varepsilon} + \left(b_1^{\varepsilon}(t) - b_1(t) \right) x_2 + \left(b_2^{\varepsilon}(t) - b_2(t) \right) \delta u \right] dW(t) & \text{in } (0, T], \\
x_3^{\varepsilon}(0) = 0.
\end{cases}$$
(6.27)

It follows from (6.26)–(6.27) that

$$\begin{split} & \mathbb{E}|x_{3}^{\varepsilon}(t)|_{H}^{2} \\ & = \mathbb{E}\Big|\int_{0}^{t}S(t-s)a_{1}^{\varepsilon}(s)x_{3}^{\varepsilon}(s)ds + \int_{0}^{t}S(t-s)b_{1}^{\varepsilon}(s)x_{3}^{\varepsilon}(s)dW(s) \\ & + \int_{0}^{t}S(t-s)\left[a_{1}^{\varepsilon}(s) - a_{1}(s)\right]x_{2}(s)ds + \int_{0}^{t}S(t-s)\left[b_{1}^{\varepsilon}(s) - b_{1}(s)\right]x_{2}(s)dW(s) \\ & + \int_{0}^{t}S(t-s)\left[a_{2}^{\varepsilon}(s) - a_{2}(s)\right]\delta u(s)ds + \int_{0}^{t}S(t-s)\left[b_{2}^{\varepsilon}(s) - b_{2}(s)\right]\delta u(s)dW(s)\Big|_{H}^{2} \\ & \leq \mathcal{C}\Big[\mathbb{E}\int_{0}^{t}|x_{3}^{\varepsilon}(s)|_{H}^{2}ds + |x_{2}(\cdot)|_{L_{\mathbb{F}}^{\infty}(0,T;L^{2}(\Omega;H))}^{2}\int_{0}^{T}\mathbb{E}\Big(\|a_{1}^{\varepsilon}(s) - a_{1}(s)\|_{\mathcal{L}(H)}^{2} + \|b_{1}^{\varepsilon}(s) - b_{1}(s)\|_{\mathcal{L}(H)}^{2}\Big)dt \\ & + |u - \bar{u}|_{L_{\mathbb{F}}^{2}(0,T;L^{2}(\Omega;H_{1}))}^{2}\int_{0}^{T}\mathbb{E}\Big(\|a_{2}^{\varepsilon}(s) - a_{2}(s)\|_{\mathcal{L}(H_{1},H)}^{2} + \|b_{2}^{\varepsilon}(s) - b_{2}(s)\|_{\mathcal{L}(H_{1},H)}^{2}\Big)dt\Big] \\ & \leq \mathcal{C}(1 + |u - \bar{u}|_{L_{\mathbb{F}}^{2}(0,T;L^{2}(\Omega;H_{1}))}^{2}\Big)\Big\{\mathbb{E}\int_{0}^{t}|x_{3}^{\varepsilon}(s)|_{H}^{2}ds + \int_{0}^{T}\mathbb{E}\Big[\|a_{1}^{\varepsilon}(s) - a_{1}(s)\|_{\mathcal{L}(H)}^{2} + \|b_{1}^{\varepsilon}(s) - b_{1}(s)\|_{\mathcal{L}(H)}^{2} \\ & + \|a_{2}^{\varepsilon}(s) - a_{2}(s)\|_{\mathcal{L}(H_{1},H)}^{2} + \|b_{2}^{\varepsilon}(s) - b_{2}(s)\|_{\mathcal{L}(H_{1},H)}^{2}\Big]dt\Big\}. \end{split}$$

This, together with Gronwall's inequality, implies that

$$\mathbb{E}|x_{3}^{\varepsilon}(t)|_{H}^{2} \leq Ce^{C|u-\bar{u}|_{L_{\mathbb{F}}^{2}(0,T;L^{2}(\Omega;H_{1}))}} \int_{0}^{T} \mathbb{E}\Big[\|a_{1}^{\varepsilon}(s)-a_{1}(s)\|_{\mathcal{L}(H)}^{2} + \|b_{1}^{\varepsilon}(s)-b_{1}(s)\|_{\mathcal{L}(H)}^{2} + \|a_{2}^{\varepsilon}(s)-a_{2}(s)\|_{\mathcal{L}(H_{1},H)}^{2} + \|b_{2}^{\varepsilon}(s)-b_{2}(s)\|_{\mathcal{L}(H_{1},H)}^{2}\Big] ds, \quad \forall t \in [0,T].$$

$$(6.28)$$

Note that (6.25) implies $x^{\varepsilon}(\cdot) \to \bar{x}(\cdot)$ (in H) in probability, as $\varepsilon \to 0$. Hence, by (6.20), (6.22) and the continuity of $a_x(t,\cdot,\cdot)$, $b_x(t,\cdot,\cdot)$, $a_u(t,\cdot,\cdot)$ and $b_u(t,\cdot,\cdot)$, we deduce that

$$\begin{split} \lim_{\varepsilon \to 0} \int_0^T \mathbb{E} \Big[\| a_1^{\varepsilon}(s) - a_1(s) \|_{\mathcal{L}(H)}^2 + \| b_1^{\varepsilon}(s) - b_1(s) \|_{\mathcal{L}(H)}^2 \\ + \| a_2^{\varepsilon}(s) - a_2(s) \|_{\mathcal{L}(H_1, H)}^2 + \| b_2^{\varepsilon}(s) - b_2(s) \|_{\mathcal{L}(H_1, H)}^2 \Big] ds &= 0. \end{split}$$

This, combined with (6.28), gives (6.23).

Step 3. Since $(\bar{x}(\cdot), \bar{u}(\cdot))$ is an optimal pair of Problem (OP), from (6.23), we find that

$$0 \leq \lim_{\varepsilon \to 0} \frac{\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot))}{\varepsilon}$$

$$= \left\{ \mathbb{E} \int_{0}^{T} \left[\left\langle g_{1}(t), x_{2}(t) \right\rangle_{H} + \left\langle g_{2}(t), \delta u(t) \right\rangle_{H_{1}} \right] dt + \mathbb{E} \left\langle h_{x}(\bar{x}(T)), x_{2}(T) \right\rangle_{H} \right\},$$
(6.29)

where

$$g_1(t) = g_x(t, \bar{x}(t), \bar{u}(t)), \quad g_2(t) = g_u(t, \bar{x}(t), \bar{u}(t)).$$

Now, it follows from Itôs formula that

$$-\mathbb{E}\langle h_x(\bar{x}(T)), x_2(T)\rangle_H - \mathbb{E}\int_0^T \langle g_1(t), x_2(t)\rangle_H dt$$

$$= \mathbb{E}\int_0^T \left[\langle a_2(t)\delta u(t), y(t)\rangle_H + \langle b_2(t)\delta u(t), Y(t)\rangle_H\right] dt.$$
(6.30)

Combining (6.29) and (6.30), we find

$$\mathbb{E} \int_{0}^{T} \left\langle a_{2}(t)^{*}y(t) + b_{2}(t)^{*}Y(t) - g_{2}(t), u(t) - \bar{u}(t) \right\rangle_{\widetilde{H}} dt \le 0$$
 (6.31)

holds for any $u(\cdot) \in \mathcal{U}[0,T]$ satisfying $u(\cdot) - \bar{u}(\cdot) \in L^2_{\mathbb{F}}(0,T;L^2(\Omega;\widetilde{H}))$. Hence, by means of Lemma 6.1, we conclude that

$$\langle a_2(t)^* y(t) + b_2(t)^* Y(t) - g_2(t), u - \bar{u}(t) \rangle_{\widetilde{H}} \le 0,$$
 a.e. $[0, T] \times \Omega, \ \forall \ u \in U.$ (6.32)

This completes the proof of Theorem 6.2.

6.3 Pontryagin-type maximum principle for the general case

In this subsection, we give a necessary condition for optimal controls of Problem (OP) for the general case.

6.3.1 Relaxed transposition solution to operator-valued backward stochastic evolution equations

To define the solution to (6.6) in the transposition sense, we need to introduce the following two (forward) stochastic evolution equations:

$$\begin{cases} dx_1 = (A+J)x_1ds + u_1ds + Kx_1dW(s) + v_1dW(s) & \text{in } (t,T], \\ x_1(t) = \xi_1 \end{cases}$$
(6.33)

and

$$\begin{cases} dx_2 = (A+J)x_2ds + u_2ds + Kx_2dW(s) + v_2dW(s) & \text{in } (t,T], \\ x_2(t) = \xi_2. \end{cases}$$
(6.34)

Here

$$\xi_1, \xi_2 \in L^4_{\mathcal{F}_t}(\Omega; H), u_1, u_2 \in L^2_{\mathbb{F}}(t, T; L^4(\Omega; H)), v_1, v_2 \in L^2_{\mathbb{F}}(t, T; L^4(\Omega; H)).$$

Also, we need to introduce the solution space for (6.6). Write

$$C_{\mathbb{F},w}([0,T];L^{2}(\Omega;\mathcal{L}(H))) \stackrel{\triangle}{=} \left\{ P(\cdot,\cdot) \mid P(\cdot,\cdot) \in L_{\mathbb{F}}^{\infty}(0,T;L^{2}(\Omega;\mathcal{L}(H))) \text{ and for every } t \in [0,T] \text{ and } \xi \in L_{\mathcal{F}_{t}}^{4}(\Omega;H), \right.$$

$$\left. P(\cdot,\cdot)\xi \in C_{\mathbb{F}}([t,T];L^{\frac{4}{3}}(\Omega;H)) \text{ and } \left| P(\cdot,\cdot)\xi \right|_{C_{\mathbb{F}}([t,T];L^{\frac{4}{3}}(\Omega;H))} \leq \mathcal{C}|\xi|_{L_{\mathcal{F}_{t}}^{4}(\Omega;H)} \right\}$$

$$(6.35)$$

and

$$\mathcal{Q}[0,T] \stackrel{\triangle}{=} \left\{ \left(Q^{(\cdot)}, \widehat{Q}^{(\cdot)} \right) \mid \text{ For any } t \in [0,T], \text{ both } Q^{(t)} \text{ and } \widehat{Q}^{(t)} \text{ are bounded linear operators} \right. \\
\text{from } L^4_{\mathcal{F}_t}(\Omega; H) \times L^2_{\mathbb{F}}(t,T; L^4(\Omega; H)) \times L^2_{\mathbb{F}}(t,T; L^4(\Omega; H)) \text{ to } L^2_{\mathbb{F}}(t,T; L^{\frac{4}{3}}(\Omega; H)) \\
\text{and } Q^{(t)}(0,0,\cdot)^* = \widehat{Q}^{(t)}(0,0,\cdot) \right\}. \tag{6.36}$$

We now employ the stochastic transposition method, and define the relaxed transposition solution to (6.6) as follows:

Definition 6.2 We call $(P(\cdot), Q^{(\cdot)}, \widehat{Q}^{(\cdot)}) \in C_{\mathbb{F},w}([0,T]; L^2(\Omega; \mathcal{L}(H))) \times \mathcal{Q}[0,T]$ a relaxed transposition solution to the equation (6.6) if for any $t \in [\tau, T]$, $\xi_1, \xi_2 \in L^4_{\mathcal{F}_t}(\Omega; H)$, $u_1(\cdot)$, $u_2(\cdot) \in L^2_{\mathbb{F}}(t, T; L^4(\Omega; H))$ and $v_1(\cdot), v_2(\cdot) \in L^2_{\mathbb{F}}(t, T; L^4(\Omega; H))$, it holds that

$$\mathbb{E}\langle P_{T}x_{1}(T), x_{2}(T)\rangle_{H} - \mathbb{E}\int_{t}^{T} \langle F(s)x_{1}(s), x_{2}(s)\rangle_{H} ds$$

$$= \mathbb{E}\langle P(t)\xi_{1}, \xi_{2}\rangle_{H} + \mathbb{E}\int_{t}^{T} \langle P(s)u_{1}(s), x_{2}(s)\rangle_{H} ds + \mathbb{E}\int_{t}^{T} \langle P(s)x_{1}(s), u_{2}(s)\rangle_{H} ds$$

$$+ \mathbb{E}\int_{t}^{T} \langle P(s)K(s)x_{1}(s), v_{2}(s)\rangle_{H} ds + \mathbb{E}\int_{t}^{T} \langle P(s)v_{1}(s), K(s)x_{2}(s) + v_{2}(s)\rangle_{H} ds$$

$$+ \mathbb{E}\int_{t}^{T} \langle v_{1}(s), \widehat{Q}^{(t)}(\xi_{2}, u_{2}, v_{2})(s)\rangle_{H} ds + \mathbb{E}\int_{t}^{T} \langle Q^{(t)}(\xi_{1}, u_{1}, v_{1})(s), v_{2}(s)\rangle_{H} ds,$$
(6.37)

Here, $x_1(\cdot)$ and $x_2(\cdot)$ solve (6.33) and (6.34), respectively.

We have the following well-posedness result for the equation (6.6) (See [12] for its proof).

Theorem 6.3 Suppose that $L^p_{\mathcal{F}_T}(\Omega)$ $(1 \leq p < \infty)$ is a separable Banach space. Then the equation (6.6) admits one and only one relaxed transposition solution $(P(\cdot), Q^{(\cdot)}, \widehat{Q}^{(\cdot)}) \in C_{\mathbb{F},w}([0,T]; L^2(\Omega; \mathcal{L}(H))) \times \mathcal{Q}[0,T]$. Furthermore,

$$|P|_{C_{\mathbb{F},w}([0,T];L^{2}(\Omega;\mathcal{L}(H)))} + |(Q^{(\cdot)},\widehat{Q}^{(\cdot)})|_{\mathcal{Q}[0,T]} \leq \mathcal{C}(|F|_{L^{1}_{\mathbb{F}}(0,T;L^{2}(\Omega;\mathcal{L}(H)))} + |P_{T}|_{L^{2}_{\mathcal{F}_{T}}(\Omega;\mathcal{L}(H))}).$$

Next, we give a regularity result for the relaxed transposition solution. For this purpose, we first give two preliminary results (See [13] for their proofs).

Lemma 6.2 For each $t \in [0,T]$, if $u_2 = v_2 = 0$ in the equation (6.34), then there exists an operator $U(\cdot,t) \in \mathcal{L}\left(L^4_{\mathcal{F}_t}(\Omega;H); C_{\mathbb{F}}([t,T];L^4(\Omega;H))\right)$ such that the solution to (6.34) can be represented as $x_2(\cdot) = U(\cdot,t)\xi_2$.

Let $\{\Delta_n\}_{n=1}^{\infty}$ be a sequence of partitions of [0,T], that is,

$$\Delta_n \stackrel{\triangle}{=} \left\{ t_i^n \mid i = 0, 1, \dots, n, \text{ and } 0 = t_0^n < t_1^n < \dots < t_n^n = T \right\}$$

such that $\Delta_n \subset \Delta_{n+1}$ and $\delta(\Delta_n) \stackrel{\triangle}{=} \max_{0 \le i \le n-1} (t_{i+1}^n - t_i^n) \to 0$ as $n \to \infty$. We introduce the following subspaces of $L^2_{\mathbb{F}}(0,T;L^4(\Omega;H))$:

$$\mathcal{H}_{n} = \left\{ \sum_{i=0}^{n-1} \chi_{[t_{i}^{n}, t_{i+1}^{n})}(\cdot) U(\cdot, t_{i}^{n}) h_{i} \mid h_{i} \in L_{\mathcal{F}_{t_{i}^{n}}}^{4}(\Omega; H) \right\}.$$
 (6.38)

Here $U(\cdot,\cdot)$ is the operator introduced in Lemma 6.2. We have the following result.

Lemma 6.3 The set $\bigcup_{n=1}^{\infty} \mathcal{H}_n$ is dense in $L^2_{\mathbb{F}}(0,T;L^4(\Omega;H))$.

The regularity result for solutions to (6.6) can be stated as follows (See [13] for its proof).

Lemma 6.4 Suppose that the assumptions in Theorem 6.3 hold and let $(P(\cdot), Q^{(\cdot)}, \widehat{Q}^{(\cdot)})$ be the relaxed transposition solution to the equation (6.6). Then, there exist an $n \in \mathbb{N}$ and two pointwise defined linear operators Q^n and \widehat{Q}^n , both of which are from \mathcal{H}_n to $L^2_{\mathbb{F}}(0,T;L^{\frac{4}{3}}(\Omega;H))$, such that, for any $\xi_1, \xi_2 \in L^4_{\mathcal{F}_0}(\Omega;H)$, $u_1(\cdot), u_2(\cdot) \in L^4_{\mathbb{F}}(\Omega;L^2(0,T;H))$ and $v_1(\cdot), v_2(\cdot) \in \mathcal{H}_n$, it holds that

$$\mathbb{E} \int_{0}^{T} \left\langle v_{1}(s), \widehat{Q}^{(0)}(\xi_{2}, u_{2}, v_{2})(s) \right\rangle_{H} ds + \mathbb{E} \int_{0}^{T} \left\langle Q^{(0)}(\xi_{1}, u_{1}, v_{1})(s), v_{2}(s) \right\rangle_{H} ds
= \mathbb{E} \int_{0}^{T} \left[\left\langle \left(Q^{n} v_{1} \right)(s), x_{2}(s) \right\rangle_{H} + \left\langle x_{1}(s), \left(\widehat{Q}^{n} v_{2} \right)(s) \right\rangle_{H} \right] ds,$$
(6.39)

where, $x_1(\cdot)$ and $x_2(\cdot)$ solve accordingly (6.33) and (6.34) with t = 0. Further, there is a positive constant C(n), depending on n, such that

$$|Q^{n}v_{1}|_{L_{\mathbb{F}}^{2}(0,T;L^{\frac{4}{3}}(\Omega;H))} + |\widehat{Q}^{n}v_{2}|_{L_{\mathbb{F}}^{2}(0,T;L^{\frac{4}{3}}(\Omega;H))} \leq \mathcal{C}(n) (|\widetilde{v}_{1}|_{L_{\mathbb{F}}^{2}(0,T;L^{4}(\Omega;H))} + |\widetilde{v}_{2}|_{L_{\mathbb{F}}^{2}(0,T;L^{4}(\Omega;H))}), \tag{6.40}$$

where

$$\tilde{v}_1 = \sum_{i=0}^{n-1} \chi_{[t_i^n, t_{i+1}^n)}(\cdot) h_i \quad \text{for } v_1 = \sum_{i=0}^{n-1} \chi_{[t_i^n, t_{i+1}^n)}(\cdot) U(\cdot, t_i) h_i$$

and

$$\tilde{v}_2 = \sum_{j=0}^{n-1} \chi_{[t_j^n, t_{j+1}^n)}(\cdot) h_j \quad \text{for } v_2 = \sum_{j=0}^{n-1} \chi_{[t_j^n, t_{j+1}^n)}(\cdot) U(\cdot, t_j) h_j.$$

6.3.2 Statement of the Pontryagin-type maximum principle

We assume the following further conditions for the optimal control problem (OP).

(B4) The function a(t, x, u) and b(t, x, u), and the functional g(t, x, u) and h(x) are C^2 with respect to x, such that for $\psi(t, x, u) = g(t, x, u), h(x)$, it holds that $\varphi_x(t, x, u), \psi_x(t, x, u), \varphi_{xx}(t, x, u)$ and $\psi_{xx}(t, x, u)$ are continuous with respect to u. Moreover, for all $(t, x, u) \in [0, T] \times H \times U$,

$$\begin{cases}
\|a_x(t,x,u)\|_{\mathcal{L}(H)} + \|b_x(t,x,u)\|_{\mathcal{L}(H)} + |\psi_x(t,x,u)|_H \leq C_L, \\
\|a_{xx}(t,x,u)\|_{\mathcal{L}(H\times H,H)} + \|b_{xx}(t,x,u)\|_{\mathcal{L}(H\times H,H)} + \|\psi_{xx}(t,x,u)\|_{\mathcal{L}(H)} \leq C_L.
\end{cases}$$
(6.41)

Let

$$\mathbb{H}(t, x, u, k_1, k_2) \stackrel{\triangle}{=} \langle k_1, a(t, x, u) \rangle_H + \langle k_2, b(t, x, u) \rangle_H - g(t, x, u),$$

$$(t, x, u, k_1, k_2) \in [0, T] \times H \times U \times H \times H.$$

$$(6.42)$$

We have the following result.

Theorem 6.4 Suppose that $L^p_{\mathcal{F}_T}(\Omega)$ $(1 \leq p < \infty)$ is a separable Banach space, U is a separable metric space, and $x_0 \in L^8_{\mathcal{F}_0}(\Omega; H)$. Let the assumptions (B1), (B2) and (B4) hold, and let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be an optimal pair of Problem (OP). Let $(y(\cdot), Y(\cdot))$ be the transposition solution to (6.7) with y_T and $f(\cdot, \cdot, \cdot)$ given by (6.14). Assume that $(P(\cdot), Q^{(\cdot)}, \widehat{Q}^{(\cdot)})$ is the relaxed transposition solution to the equation (6.6) in which P_T , $J(\cdot)$, $K(\cdot)$ and $F(\cdot)$ are given by

$$\begin{cases}
P_T = -h_{xx}(\bar{x}(T)), & J(t) = a_x(t, \bar{x}(t), \bar{u}(t)), \\
K(t) = b_x(t, \bar{x}(t), \bar{u}(t)), & F(t) = -\mathbb{H}_{xx}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)).
\end{cases} (6.43)$$

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

$$\mathbb{H}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t)) - \operatorname{Re} \mathbb{H}(t, \bar{x}(t), u, y(t), Y(t)) \\
-\frac{1}{2} \langle P(t) [b(t, \bar{x}(t), \bar{u}(t)) - b(t, \bar{x}(t), u)], b(t, \bar{x}(t), \bar{u}(t)) - b(t, \bar{x}(t), u) \rangle_{H} \\
\geq 0. \tag{6.44}$$

6.3.3 Proof of the Pontryagin-type stochastic maximum principle

We are now in a position to prove Theorem 6.4.

Proof of Theorem 6.4: We divide the proof into two steps.

Step 1. For each $\varepsilon > 0$, let $E_{\varepsilon} \subset [0,T]$ be a measurable set with measure ε . Put

$$u^{\varepsilon}(t) = \begin{cases} \bar{u}(t), & t \in [0, T] \setminus E_{\varepsilon}, \\ u(t), & t \in E_{\varepsilon}. \end{cases}$$
 (6.45)

where $u(\cdot)$ is an arbitrary given element in $\mathcal{U}[0,T]$.

We introduce some notations which will be used in what follows. For $\varphi = a, b, g$, we let

$$\begin{cases}
\varphi_{1}(t) = \varphi_{x}(t, \bar{x}(t), \bar{u}(t)), & \varphi_{11}(t) = \varphi_{xx}(t, \bar{x}(t), \bar{u}(t)), \\
\tilde{\varphi}_{1}^{\varepsilon}(t) = \int_{0}^{1} \varphi_{x}(t, \bar{x}(t) + \sigma(x^{\varepsilon}(t) - \bar{x}(t)), u^{\varepsilon}(t)) d\sigma, \\
\tilde{\varphi}_{11}^{\varepsilon}(t) = 2 \int_{0}^{1} (1 - \sigma) a_{xx}(t, \bar{x}(t) + \sigma(x^{\varepsilon}(t) - \bar{x}(t)), u^{\varepsilon}(t)) d\sigma,
\end{cases} (6.46)$$

and

$$\begin{cases}
\delta\varphi(t) = \varphi(t, \bar{x}(t), u(t)) - \varphi(t, \bar{x}(t), \bar{u}(t)), \\
\delta\varphi_1(t) = \varphi_x(t, \bar{x}(t), u(t)) - \varphi_x(t, \bar{x}(t), \bar{u}(t)), \\
\delta\varphi_{11}(t) = \varphi_{xx}(t, \bar{x}(t), u(t)) - \varphi_{xx}(t, \bar{x}(t), \bar{u}(t)).
\end{cases} (6.47)$$

Let $x^{\varepsilon}(\cdot)$ be the state process of the system (6.2) corresponding to the control $u^{\varepsilon}(\cdot)$. Then, $x^{\varepsilon}(\cdot)$ solves

$$\begin{cases}
 dx^{\varepsilon} = \left[Ax^{\varepsilon} + a(t, x^{\varepsilon}, u^{\varepsilon}) \right] dt + b(t, x^{\varepsilon}, u^{\varepsilon}) dW(t) & \text{in } (0, T], \\
 x^{\varepsilon}(0) = x_{0}.
\end{cases}$$
(6.48)

It is easy to prove that

$$|x^{\varepsilon}|_{C_{\mathbb{F}}([0,T];L^{8}(\Omega;H))} \le C\left(1+|x_{0}|_{L^{8}_{\mathcal{D}_{0}}(\Omega;H)}\right), \quad \forall \ \varepsilon > 0.$$

$$(6.49)$$

Let $x_1^{\varepsilon}(\cdot) = x^{\varepsilon}(\cdot) - \bar{x}(\cdot)$. Then, by (6.49) and noting that the optimal pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ solves the equation (6.2), we see that $x_1^{\varepsilon}(\cdot)$ satisfies the following stochastic evolution equation:

$$\begin{cases}
dx_1^{\varepsilon} = \left[Ax_1^{\varepsilon} + \tilde{a}_1^{\varepsilon}(t)x_1^{\varepsilon} + \chi_{E_{\varepsilon}}(t)\delta a(t) \right] dt + \left[\tilde{b}_1^{\varepsilon}(t)x_1^{\varepsilon} + \chi_{E_{\varepsilon}}(t)\delta b(t) \right] dW(t) & \text{in } (0, T], \\
x_1^{\varepsilon}(0) = 0.
\end{cases}$$
(6.50)

Consider the following two stochastic differential equations:

$$\begin{cases}
dx_2^{\varepsilon} = \left[Ax_2^{\varepsilon} + a_1(t)x_2^{\varepsilon} \right] dt + \left[b_1(t)x_2^{\varepsilon} + \chi_{E_{\varepsilon}}(t)\delta b(t) \right] dW(t) & \text{in } (0, T], \\
x_2^{\varepsilon}(0) = 0
\end{cases}$$
(6.51)

 and^3

$$\begin{cases}
dx_3^{\varepsilon} = \left[Ax_3^{\varepsilon} + a_1(t)x_3^{\varepsilon} + \chi_{E_{\varepsilon}}(t)\delta a(t) + \frac{1}{2}a_{11}(t)\left(x_2^{\varepsilon}, x_2^{\varepsilon}\right) \right] dt \\
+ \left[b_1(t)x_3^{\varepsilon} + \chi_{E_{\varepsilon}}(t)\delta b_1(t)x_2^{\varepsilon} + \frac{1}{2}b_{11}(t)\left(x_2^{\varepsilon}, x_2^{\varepsilon}\right) \right] dW(t) & \text{in } (0, T], \\
x_3^{\varepsilon}(0) = 0.
\end{cases}$$
(6.52)

Similar to Steps 1-2 in the proof of Theorem 6.2, we can show that

³Recall that, for any C^2 -function $f(\cdot)$ defined on a Banach space X and $x_0 \in X$, $f_{xx}(x_0) \in \mathcal{L}(X,X;X)$. This means that, for any $x_1, x_2 \in X$, $f_{xx}(x_0)(x_1, x_2) \in X$. Hence, by (6.46), $a_{11}(t)(x_2^{\varepsilon}, x_2^{\varepsilon})$ (in (6.52)) stands for $a_{xx}(t, \bar{x}(t), \bar{u}(t))(x_2^{\varepsilon}(t), x_2^{\varepsilon}(t))$. One has a similar meaning for $b_{11}(t)(x_2^{\varepsilon}, x_2^{\varepsilon})$ and so on.

$$|x_1^{\varepsilon}(\cdot)|_{C_{\mathbb{F}}([0,T];L^8(\Omega;H))}^8 \le \mathcal{C}(x_0)\varepsilon^4,\tag{6.53}$$

$$|x_2^{\varepsilon}(\cdot)|_{C_{\mathbb{R}}([0,T];L^8(\Omega;H))}^8 \le \mathcal{C}(x_0)\varepsilon^4,\tag{6.54}$$

$$\max_{t \in [0,T]} \mathbb{E}|x_3^{\varepsilon}(t)|_H^4 \le \mathcal{C}(x_0)\varepsilon^4, \tag{6.55}$$

$$|x_4^{\varepsilon}(\cdot)|_{C_{\mathbb{F}}([0,T];L^2(\Omega;H))} \le \mathcal{C}(x_0)\varepsilon, \tag{6.56}$$

$$|x_5^{\varepsilon}(\cdot)|^2_{C_{\mathbb{F}}([0,T];L^2(\Omega;H))} = o(\varepsilon^2), \quad \text{as } t \to 0,$$
 (6.57)

$$\left| \int_{0}^{1} (1 - \sigma) \left(g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), u^{\varepsilon}(t) \right) - g_{xx} \left(t, \bar{x}(t), u^{\varepsilon}(t) \right) \right) d\sigma \right|_{\mathcal{L}(H)}$$

$$\leq \mathcal{C} \left(\int_{0}^{1} \left| g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), \bar{u}(t) \right) - g_{xx} \left(t, \bar{x}(t), \bar{u}(t) \right) \right|_{\mathcal{L}(H)} d\sigma + \chi_{E_{\varepsilon}}(t) \right),$$

$$(6.58)$$

$$\left| \int_{0}^{1} (1 - \sigma) \left(h_{xx} (\bar{x}(T) + \sigma x_{1}^{\varepsilon}(T)) - h_{xx} (\bar{x}(T)) \right) d\sigma \right|_{\mathcal{L}(H)}$$

$$\leq \mathcal{C} \left(\int_{0}^{1} \left| h_{xx} (\bar{x}(T) + \sigma x_{1}^{\varepsilon}(T)) - h_{xx} (\bar{x}(T)) \right|_{\mathcal{L}(H)} d\sigma + \chi_{E_{\varepsilon}}(t) \right),$$

$$(6.59)$$

and

$$|x_1^{\varepsilon} - x_2^{\varepsilon} - x_3^{\varepsilon}|_{L_{\mathbb{R}}^{\infty}(0,T;L^2(\Omega;H))} = o(\varepsilon), \quad \text{as } \varepsilon \to 0.$$
 (6.60)

Step 2. We need to compute the value of $\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot))$.

$$\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \\
= \mathbb{E} \int_{0}^{T} \left[g(t, x^{\varepsilon}(t), u^{\varepsilon}(t)) - g(t, \bar{x}(t), \bar{u}(t)) \right] dt + \mathbb{E}h(x^{\varepsilon}(T)) - \mathbb{E}h(\bar{x}(T)) \\
= \mathbb{E} \int_{0}^{T} \left\{ \chi_{E_{\varepsilon}}(t) \delta g(t) + \left\langle g_{x}(t, \bar{x}(t), u^{\varepsilon}(t)), x_{1}^{\varepsilon}(t) \right\rangle_{H} \right. \\
\left. + \int_{0}^{1} \left\langle (1 - \sigma) g_{xx}(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), u^{\varepsilon}(t)) x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) \right\rangle_{H} d\sigma \right\} dt \\
+ \mathbb{E} \left\langle h_{x}(\bar{x}(T)), x_{1}^{\varepsilon}(T) \right\rangle_{H} + \mathbb{E} \int_{0}^{1} \left\langle (1 - \sigma) h_{xx}(\bar{x}(T) + \sigma x_{1}^{\varepsilon}(T)) x_{1}^{\varepsilon}(T), x_{1}^{\varepsilon}(T) \right\rangle_{H} d\sigma. \tag{6.61}$$

This, together with the definition of $x_i^{\varepsilon}(\cdot)$ (i=1,2,3,4,5), yields that

$$\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \\
= \mathbb{E} \int_{0}^{T} \left\{ \chi_{E_{\varepsilon}}(t) \delta g(t) + \left\langle \delta g_{1}(t), x_{1}^{\varepsilon}(t) \right\rangle_{H} \chi_{E_{\varepsilon}}(t) + \left\langle g_{1}(t), x_{2}^{\varepsilon}(t) + x_{3}^{\varepsilon}(t) \right\rangle_{H} + \left\langle g_{1}(t), x_{5}^{\varepsilon}(t) \right\rangle_{H} \\
+ \int_{0}^{1} \left\langle (1 - \sigma) \left[g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), u^{\varepsilon}(t) \right) - g_{xx} \left(t, \bar{x}(t), u^{\varepsilon}(t) \right) \right] x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) \right\rangle_{H} d\sigma \\
+ \frac{1}{2} \left\langle \delta g_{11}(t) x_{1}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) \right\rangle_{H} \chi_{E_{\varepsilon}}(t) + \frac{1}{2} \left\langle g_{11}(t) x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right\rangle_{H} + \frac{1}{2} \left\langle g_{11}(t) x_{4}^{\varepsilon}(t), x_{1}^{\varepsilon}(t) + x_{2}^{\varepsilon}(t) \right\rangle_{H} \right\} dt \\
+ \mathbb{E} \left\langle h_{x}(\bar{x}(T)), x_{2}^{\varepsilon}(t) + x_{3}^{\varepsilon}(t) \right\rangle_{H} + \mathbb{E} \left\langle h_{x}(\bar{x}(T)), x_{5}^{\varepsilon}(t) \right\rangle_{H} + \frac{1}{2} \mathbb{E} \left\langle h_{xx}(\bar{x}(T)) x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right\rangle_{H} \\
+ \mathbb{E} \int_{0}^{1} \left\langle (1 - \sigma) \left[h_{xx}(\bar{x}(T) + \sigma x_{1}^{\varepsilon}(T)) - h_{xx}(\bar{x}(T)) \right] x_{1}^{\varepsilon}(T), x_{1}^{\varepsilon}(T) \right\rangle_{H} d\sigma. \tag{6.62}$$

For a.e. $t \in [0, T]$, we find that

$$\left\| \int_{0}^{1} (1 - \sigma) \left[g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), u^{\varepsilon}(t) \right) - g_{xx} \left(t, \bar{x}(t), u^{\varepsilon}(t) \right) \right] d\sigma \right\|_{\mathcal{L}(H \times H, H)}$$

$$= \left\| \int_{0}^{1} (1 - \sigma) \left[g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), \bar{u}(t) \right) - g_{xx} \left(t, \bar{x}(t), \bar{u}(t) \right) \right] d\sigma \right.$$

$$+ \int_{0}^{1} (1 - \sigma) \chi_{E_{\varepsilon}}(t) g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), u(t) \right) d\sigma + \chi_{E_{\varepsilon}}(t) g_{xx} \left(t, \bar{x}(t), u(t) \right) \right\|_{\mathcal{L}(H \times H, H)} d\sigma$$

$$\leq \mathcal{C} \left[\int_{0}^{1} \left\| g_{xx} \left(t, \bar{x}(t) + \sigma x_{1}^{\varepsilon}(t), \bar{u}(t) \right) - g_{xx} \left(t, \bar{x}(t), \bar{u}(t) \right) \right\|_{\mathcal{L}(H \times H, H)} d\sigma + \chi_{E_{\varepsilon}}(t) \right]. \tag{6.63}$$

By (6.62), noting (6.53), (6.54), (6.55), (6.56), (6.57) and (6.58), and using the continuity of both $h_{xx}(x)$ and $g_{xx}(x)$ with respect to x, we end up with

$$\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \\
= \mathbb{E} \int_{0}^{T} \left[\left\langle g_{1}(t), x_{2}^{\varepsilon}(t) + x_{3}^{\varepsilon}(t) \right\rangle_{H} + \frac{1}{2} \left\langle g_{11}(t) x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right\rangle_{H} + \chi_{E_{\varepsilon}}(t) \delta g(t) \right] dt \\
+ \mathbb{E} \left\langle h_{x}(\bar{x}(T)), x_{2}^{\varepsilon}(T) + x_{3}^{\varepsilon}(T) \right\rangle_{H} + \frac{1}{2} \mathbb{E} \left\langle h_{xx}(\bar{x}(T)) x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right\rangle_{H} + o(\varepsilon). \tag{6.64}$$

In the sequel, we shall get rid of $x_2^{\varepsilon}(\cdot)$ and $x_3^{\varepsilon}(\cdot)$ in (6.64) by solutions to the equations (6.7) and (6.6). By the definition of the transposition solution to the equation (6.7) (with y_T and $f(\cdot,\cdot,\cdot)$ given by (6.14)), we obtain that

$$-\mathbb{E}\langle h_x(\bar{x}(T)), x_2^{\varepsilon}(T)\rangle_H - \mathbb{E}\int_0^T \langle g_1(t), x_2^{\varepsilon}(t)\rangle_H dt = \mathbb{E}\int_0^T \langle Y(t), \delta b(t)\rangle_H \chi_{E_{\varepsilon}}(t) dt \quad (6.65)$$

and

$$-\mathbb{E}\langle h_{x}(\bar{x}(T)), x_{3}^{\varepsilon}(T)\rangle_{H} - \mathbb{E}\int_{0}^{T} \langle g_{1}(t), x_{3}^{\varepsilon}(t)\rangle_{H} dt$$

$$= \mathbb{E}\int_{0}^{T} \left\{ \frac{1}{2} \Big[\langle y(t), a_{11}(t) \big(x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \big) \rangle_{H} + \langle Y(t), b_{11}(t) \big(x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \big) \rangle_{H} \Big] + \chi_{E_{\varepsilon}}(t) \Big[\langle y(t), \delta a(t) \rangle_{H} + \langle Y, \delta b_{1}(t) x_{2}^{\varepsilon}(t) \rangle_{H} \Big] \right\} dt.$$

$$(6.66)$$

According to (6.64)–(6.66), we conclude that

$$\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \\
= \frac{1}{2} \mathbb{E} \int_{0}^{T} \left[\left\langle g_{11}(t) x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right\rangle_{H} - \left\langle y(t), a_{11}(t) \left(x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right) \right\rangle_{H} \\
- \left\langle Y, b_{11}(t) \left(x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t) \right) \right\rangle_{H} \right] dt + \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \left[\delta g(t) - \left\langle y(t), \delta a(t) \right\rangle_{H} \\
- \left\langle Y(t), \delta b(t) \right\rangle_{H} \right] dt + \frac{1}{2} \mathbb{E} \left\langle h_{xx}(\bar{x}(T)) x_{2}^{\varepsilon}(T), x_{2}^{\varepsilon}(T) \right\rangle_{H} + o(\varepsilon), \quad \text{as } \varepsilon \to 0. \tag{6.67}$$

By the definition of the relaxed transposition solution to the equation (6.6) (with P_T , $J(\cdot)$, $K(\cdot)$ and $F(\cdot)$ given by (6.43)), we obtain that

$$-\mathbb{E}\langle h_{xx}(\bar{x}(T))x_{2}^{\varepsilon}(T), x_{2}^{\varepsilon}(T)\rangle_{H} + \mathbb{E}\int_{0}^{T}\langle \mathbb{H}_{xx}(t, \bar{x}(t), \bar{u}(t), y(t), Y(t))x_{2}^{\varepsilon}(t), x_{2}^{\varepsilon}(t)\rangle_{H}dt$$

$$= \mathbb{E}\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\langle b_{1}(t)x_{2}^{\varepsilon}(t), P(t)^{*}\delta b(t)\rangle_{H}dt + \mathbb{E}\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\langle P(t)\delta b(t), b_{1}(t)x_{2}^{\varepsilon}(t)\rangle_{H}dt$$

$$+\mathbb{E}\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\langle P(t)\delta b(t), \delta b(t)\rangle_{H}dt + \mathbb{E}\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\langle \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}}\delta b)(t)\rangle_{H}dt$$

$$+\mathbb{E}\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\langle Q^{(0)}(0, 0, \delta b)(t), \delta b(t)\rangle_{H}dt.$$

$$(6.68)$$

Now, we estimate the terms in the right hand side of (6.68). By (6.54), we have

$$\left| \mathbb{E} \int_0^T \chi_{E_{\varepsilon}}(t) \langle b_1(t) x_2^{\varepsilon}(t), P(t)^* \delta b(t) \rangle_H dt + \mathbb{E} \int_0^T \chi_{E_{\varepsilon}}(t) \langle P(t) \delta b(t), b_1(t) x_2^{\varepsilon}(t) \rangle_H dt \right| = o(\varepsilon).$$
(6.69)

In what follows, for any $\tau \in [0,T)$, we choose $E_{\varepsilon} = [\tau, \tau + \varepsilon] \subset [0,T]$. By Lemma 6.3, we can find a sequence $\{\beta_n\}_{n=1}^{\infty}$ such that $\beta_n \in \mathcal{H}_n$ (Recall (6.38) for the definition of \mathcal{H}_n) and $\lim_{n\to\infty} \beta_n = \delta b$ in $L^2_{\mathbb{F}}(0,T;L^4(\Omega;H))$. Hence, for some positive constant $C(x_0)$ (depending on x_0),

$$|\beta_n|_{L^2_{\mathbb{F}}(0,T;L^4(\Omega;H))} \le C(x_0) < \infty, \qquad \forall \ n \in \mathbb{N}, \tag{6.70}$$

and there is a subsequence $\{n_k\}_{k=1}^{\infty} \subset \{n\}_{n=1}^{\infty}$ such that

$$\lim_{k \to \infty} |\beta_{n_k}(t) - \delta b(t)|_{L^4_{\mathcal{F}_t}(\Omega; H)} = 0 \quad \text{for a.e. } t \in [0, T].$$

$$(6.71)$$

Denote by Q^{n_k} and \widehat{Q}^{n_k} the corresponding pointwise defined linear operators from \mathcal{H}_{n_k} to $L^2_{\mathbb{F}}(0,T;L^{\frac{4}{3}}(\Omega;H))$, given in Lemma 6.4.

Consider the following equation:

$$\begin{cases}
 dx_{2,n_k}^{\varepsilon} = \left[Ax_{2,n_k}^{\varepsilon} + a_1(t)x_{2,n_k}^{\varepsilon} \right] dt + \left[b_1(t)x_{2,n_k}^{\varepsilon} + \chi_{E_{\varepsilon}}(t)\beta_{n_k}(t) \right] dW(t) & \text{in } (0,T], \\
 x_{2,n_k}^{\varepsilon}(0) = 0.
\end{cases}$$
(6.72)

We have

$$\mathbb{E}|x_{2,n_k}^{\varepsilon}(t)|_H^4$$

$$= \mathbb{E} \Big| \int_{0}^{t} S(t-s)a_{1}(s)x_{2,n_{k}}^{\varepsilon}(s)ds + \int_{0}^{t} S(t-s)b_{1}(s)x_{2,n_{k}}^{\varepsilon}(s)dW(s) + \int_{0}^{t} S(t-s)\chi_{E_{\varepsilon}}(s)\beta_{n_{k}}(s)dW(s) \Big|_{H}^{4}$$

$$\leq C \Big[\mathbb{E} \Big| \int_{0}^{t} S(t-s)a_{1}(s)x_{2,n_{k}}^{\varepsilon}(s)ds \Big|_{H}^{4} + \mathbb{E} \Big| \int_{0}^{t} S(t-s)b_{1}(s)x_{2,n_{k}}^{\varepsilon}(s)dW(s) \Big|_{H}^{4}$$

$$+ \mathbb{E} \Big| \int_{0}^{t} S(t-s)\chi_{E_{\varepsilon}}(s)\beta_{n_{k}}(s)dW(s) \Big|_{H}^{4} \Big]$$

$$\leq C \Big[\int_{0}^{t} \mathbb{E} |x_{2,n_{k}}^{\varepsilon}(s)|_{H}^{4}ds + \varepsilon \int_{E_{\varepsilon}} \mathbb{E} |\beta_{n_{k}}(s)|_{H}^{4}ds \Big].$$

$$(6.73)$$

By (6.70) and thanks to Gronwall's inequality, (6.73) leads to

$$|x_{2,n_k}^{\varepsilon}(\cdot)|_{L_{\mathbb{F}}^{\infty}(0,T;L^4(\Omega;H))}^4 \le C(x_0,k)\varepsilon^2.$$
(6.74)

Here and henceforth, $C(x_0, k)$ is a generic constant (depending on x_0, k, T, A and C_L), which may be different from line to line. For any fixed $k \in \mathbb{N}$, since $Q^{n_k}\beta_{n_k} \in L^2_{\mathbb{F}}(0, T; L^{\frac{4}{3}}(\Omega; H))$, by (6.74), we find that

$$\left| \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \left\langle \left(Q^{n_{k}} \beta_{n_{k}} \right)(t), x_{2,n_{k}}^{\varepsilon}(t) \right\rangle_{H} dt \right|$$

$$\leq |x_{2,n_{k}}^{\varepsilon}(\cdot)|_{L_{\mathbb{F}}^{\infty}(0,T;L^{4}(\Omega;H))} \int_{E_{\varepsilon}} \left| \left(Q^{n_{k}} \beta_{n_{k}} \right)(t) \right|_{L_{\mathcal{F}_{t}}^{\frac{4}{3}}(\Omega;H)} dt \qquad (6.75)$$

$$\leq C(x_{0},k) \sqrt{\varepsilon} \int_{E_{\varepsilon}} \left| \left(Q^{n_{k}} \beta_{n_{k}} \right)(t) \right|_{L_{\mathcal{F}_{t}}^{\frac{4}{3}}(\Omega;H)} dt = o(\varepsilon), \quad \text{as } \varepsilon \to 0.$$

Similarly,

$$\left| \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \left\langle x_{2,n_{k}}^{\varepsilon}(t), \left(\widehat{Q}^{n_{k}} \beta_{n_{k}} \right)(t) \right\rangle_{H} dt \right| = o(\varepsilon), \quad \text{as } \varepsilon \to 0.$$
 (6.76)

From (6.39) in Theorem 6.4, and noting that both Q^{n_k} and \widehat{Q}^{n_k} are pointwise defined, we arrive at the following equality:

$$\mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t)\beta_{n_{k}}(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}}\beta_{n_{k}})(t) \right\rangle_{H} dt + \mathbb{E} \int_{0}^{T} \left\langle Q^{(0)}(0, 0, \chi_{E_{\varepsilon}}\beta_{n_{k}})(t), \chi_{E_{\varepsilon}}\beta_{n_{k}}(t) \right\rangle_{H} dt$$

$$= \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}} \left[\left\langle \left(Q^{n_{k}}\beta_{n_{k}} \right)(t), x_{2,n_{k}}^{\varepsilon}(t) \right\rangle_{H} + \left\langle x_{2,n_{k}}^{\varepsilon}(t), \left(\widehat{Q}^{n_{k}}\beta_{n_{k}} \right)(t) \right\rangle_{H} \right] dt. \tag{6.77}$$

Hence,

$$\mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t)\delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}}\delta b)(t) \right\rangle_{H} dt + \mathbb{E} \int_{0}^{T} \left\langle Q^{(0)}(0, 0, \chi_{E_{\varepsilon}}\delta b)(t), \chi_{E_{\varepsilon}}(t)\delta b(t) \right\rangle_{H} dt \\
-\mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \left[\left\langle \left(Q^{n_{k}} \beta_{n_{k}} \right)(t), x_{2,n_{k}}^{\varepsilon}(t) \right\rangle_{H} + \left\langle x_{2,n_{k}}^{\varepsilon}(t), \left(\widehat{Q}^{n_{k}} \beta_{n_{k}} \right)(t) \right\rangle_{H} \right] dt \\
= \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t)\delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}}\delta b)(t) \right\rangle_{H} dt + \mathbb{E} \int_{0}^{T} \left\langle Q^{(0)}(0, 0, \chi_{E_{\varepsilon}}\delta b)(t), \chi_{E_{\varepsilon}}(t)\delta b(t) \right\rangle_{H} dt \\
-\mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t)\beta_{n_{k}}(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}}\beta_{n_{k}})(t) \right\rangle_{H} dt - \mathbb{E} \int_{0}^{T} \left\langle Q^{(0)}(0, 0, \chi_{E_{\varepsilon}}\beta_{n_{k}})(t), \chi_{E_{\varepsilon}}(t)\beta_{n_{k}}(t) \right\rangle_{H} dt. \tag{6.78}$$

It is easy to see that

$$\begin{split} & \left| \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \delta b)(t) \right\rangle_{H} dt - \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \beta_{n_{k}}(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_{k}})(t) \right\rangle_{H} dt \right| \\ & \leq \left| \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \delta b)(t) \right\rangle_{H} dt - \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_{k}})(t) \right\rangle_{H} dt \right| \\ & + \left| \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_{k}})(t) \right\rangle_{H} dt - \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \beta_{n_{k}}(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_{k}})(t) \right\rangle_{H} dt \right|. \end{split}$$

From (6.71) and the density of the Lebesgue points, we find that for a.e. $\tau \in [0, T)$, it holds that

$$\lim_{k \to \infty} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big| \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \delta b)(t) \right\rangle_{H} dt$$

$$- \mathbb{E} \int_{0}^{T} \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_{k}})(t) \right\rangle_{H} dt \Big|$$

$$\leq \lim_{k \to \infty} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big[\int_{0}^{T} \chi_{E_{\varepsilon}}(t) \Big(\mathbb{E} |\delta b(t)|_{H}^{4} \Big)^{\frac{1}{2}} dt \Big]^{\frac{1}{2}} |\widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}}(\delta b - \beta_{n_{k}}))|_{L_{\mathbb{F}}^{2}(0,T;L^{\frac{4}{3}}(\Omega;H))}$$

$$\leq C \lim_{k \to \infty} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big[\int_{0}^{T} \chi_{E_{\varepsilon}}(t) \Big(\mathbb{E} |\delta b(t)|_{H}^{4} \Big)^{\frac{1}{2}} dt \Big]^{\frac{1}{2}} |\chi_{E_{\varepsilon}}(\delta b - \beta_{n_{k}})|_{L_{\mathbb{F}}^{2}(0,T;L^{4}(\Omega;H))}$$

$$\leq C \lim_{k \to \infty} \lim_{\varepsilon \to 0} \frac{|\delta b(\tau)|_{L_{\mathcal{F}_{\tau}}^{4}(\Omega;H)} \Big[\int_{0}^{T} \chi_{E_{\varepsilon}}(t) \Big(\mathbb{E} |\delta b(t) - \beta_{n_{k}}(t)|_{H}^{4} \Big)^{\frac{1}{2}} dt \Big]^{\frac{1}{2}}$$

$$= C \lim_{k \to \infty} \lim_{\varepsilon \to 0} |\delta b(\tau)|_{L_{\mathcal{F}_{\tau}}^{4}(\Omega;H)} \Big[\frac{1}{\varepsilon} \int_{\tau}^{\tau + \varepsilon} |\delta b(t) - \beta_{n_{k}}(t)|_{L_{\mathcal{F}_{\tau}}^{4}(\Omega;H)}^{2} dt \Big]^{\frac{1}{2}}$$

$$= C \lim_{k \to \infty} |\delta b(\tau)|_{L_{\mathcal{F}_{\tau}}^{4}(\Omega;H)} |\delta b(\tau) - \beta_{n_{k}}(\tau)|_{L_{\mathcal{F}_{\tau}}^{4}(\Omega;H)}$$

$$= 0.$$

Similarly,

$$\lim_{k\to\infty}\lim_{\varepsilon\to0}\frac{1}{\varepsilon}\Big|\mathbb{E}\int_{0}^{T}\left\langle\chi_{E_{\varepsilon}}(t)\delta b(t),\widehat{Q}^{(0)}(0,0,\chi_{E_{\varepsilon}}\beta_{n_{k}})(t)\right\rangle_{H}dt$$

$$-\mathbb{E}\int_{0}^{T}\left\langle\chi_{E_{\varepsilon}}(t)\beta_{n_{k}}(t),\widehat{Q}^{(0)}(0,0,\chi_{E_{\varepsilon}}\beta_{n_{k}})(t)\right\rangle_{H}dt\Big|$$

$$\leq \lim_{k\to\infty}\lim_{\varepsilon\to0}\frac{1}{\varepsilon}\Big|\widehat{Q}^{(0)}(0,0,\chi_{E_{\varepsilon}}\beta_{n_{k}})\Big|_{L_{\varepsilon}^{2}(0,T;L^{\frac{4}{3}}(\Omega;H))}\Big[\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\Big(\mathbb{E}|\delta b(t)-\beta_{n_{k}}(t)|_{H}^{4}\Big)^{\frac{1}{2}}dt\Big]^{\frac{1}{2}}$$

$$\leq C\lim_{k\to\infty}\lim_{\varepsilon\to0}\frac{1}{\varepsilon}\Big|\chi_{E_{\varepsilon}}\beta_{n_{k}}\Big|_{L_{\varepsilon}^{2}(0,T;L^{4}(\Omega;H))}\Big[\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\Big(\mathbb{E}|\delta b(t)-\beta_{n_{k}}(t)|_{H}^{4}\Big)^{\frac{1}{2}}dt\Big]^{\frac{1}{2}}$$

$$\leq C\lim_{k\to\infty}\lim_{\varepsilon\to0}\frac{1}{\varepsilon}\Big\{\Big|\chi_{E_{\varepsilon}}\delta b\Big|_{L_{\varepsilon}^{2}(0,T;L^{4}(\Omega;H))}\Big[\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\Big(\mathbb{E}|\delta b(t)-\beta_{n_{k}}(t)|_{H}^{4}\Big)^{\frac{1}{2}}dt\Big]^{\frac{1}{2}}$$

$$+\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\Big(\mathbb{E}|\delta b(t)-\beta_{n_{k}}(t)|_{H}^{4}\Big)^{\frac{1}{2}}dt\Big\}$$

$$\leq C\lim_{k\to\infty}\lim_{\varepsilon\to0}\Big\{\frac{|\delta b(\tau)|_{L_{\varepsilon}^{4}}(\Omega;H)}{\sqrt{\varepsilon}}\Big[\int_{0}^{T}\chi_{E_{\varepsilon}}(t)\Big(\mathbb{E}|\delta b(t)-\beta_{n_{k}}(t)|_{H}^{4}\Big)^{\frac{1}{2}}dt\Big\}$$

$$=C\lim_{k\to\infty}\lim_{\varepsilon\to0}\Big\{|\delta b(\tau)|_{L_{\varepsilon}^{4}}(\Omega;H)\Big[\frac{1}{\varepsilon}\int_{\tau}^{\tau+\varepsilon}|\delta b(t)-\beta_{n_{k}}(t)|_{H}^{2}\Big]^{\frac{1}{2}}dt\Big\}$$

$$=C\lim_{k\to\infty}\lim_{\varepsilon\to0}\Big\{|\delta b(\tau)|_{L_{\varepsilon}^{4}}(\Omega;H)\Big[\delta b(\tau)-\beta_{n_{k}}(t)|_{L_{\varepsilon}^{4}}(\Omega;H)dt\Big\}$$

$$=C\lim_{k\to\infty}\Big[|\delta b(\tau)|_{L_{\varepsilon}^{4}}(\Omega;H)|\delta b(\tau)-\beta_{n_{k}}(\tau)|_{L_{\varepsilon}^{4}}(\Omega;H)+|\delta b(\tau)-\beta_{n_{k}}(\tau)|_{L_{\varepsilon}^{4}}^{2}(\Omega;H)\Big]$$

$$=0.$$
(6.81)

From (6.79)–(6.81), we find that

$$\lim_{k \to \infty} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big| \mathbb{E} \int_0^T \left\langle \chi_{E_{\varepsilon}}(t) \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \delta b)(t) \right\rangle_H dt$$

$$- \mathbb{E} \int_0^T \left\langle \chi_{E_{\varepsilon}}(t) \beta_{n_k}(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_k})(t) \right\rangle_H dt \Big| = 0.$$
(6.82)

By a similar argument, we obtain that

$$\lim_{k \to \infty} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big| \mathbb{E} \int_0^T \left\langle Q^{(0)}(0, 0, \chi_{E_{\varepsilon}} \delta b)(t), \chi_{E_{\varepsilon}}(t) \delta b(t) \right\rangle_H dt$$

$$- \mathbb{E} \int_0^T \left\langle Q^{(0)}(0, 0, \chi_{E_{\varepsilon}} \beta_{n_k})(t), \chi_{E_{\varepsilon}}(t) \beta_{n_k}(t) \right\rangle_H dt \Big| = 0.$$

$$(6.83)$$

From (6.75)–(6.78) and (6.82)–(6.83), we obtain that

$$\left| \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \langle \delta b(t), \widehat{Q}^{(0)}(0, 0, \chi_{E_{\varepsilon}} \delta b)(t) \rangle_{H} dt + \mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \langle Q^{(0)}(0, 0, \delta b)(t), \delta b(t) \rangle_{H} dt \right|$$

$$= o(\varepsilon), \quad \text{as } \varepsilon \to 0.$$
(6.84)

Combining (6.67), (6.68), (6.69) and (6.84), we end up with

$$\begin{split} &\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \\ &= \mathbb{E} \int_{0}^{T} \left[\delta g(t) - \left\langle y(t), \delta a(t) \right\rangle_{H} - \left\langle Y(t), \delta b(t) \right\rangle_{H} - \frac{1}{2} \left\langle P(t) \delta b(t), \delta b(t) \right\rangle_{H} \right] \chi_{E_{\varepsilon}}(t) dt + o(\varepsilon). \end{split}$$

Since $\bar{u}(\cdot)$ is the optimal control, $\mathcal{J}(u^{\varepsilon}(\cdot)) - \mathcal{J}(\bar{u}(\cdot)) \geq 0$. Thus,

$$\mathbb{E} \int_{0}^{T} \chi_{E_{\varepsilon}}(t) \left[\left\langle y(t), \delta a(t) \right\rangle_{H} + \left\langle Y(t), \delta b(t) \right\rangle_{H} - \delta g(t) + \frac{1}{2} \left\langle P(t) \delta b(t), \delta b(t) \right\rangle_{H} \right] dt \leq o(\varepsilon), \quad (6.85)$$

as $\varepsilon \to 0$.

Finally, by
$$(6.85)$$
, we obtain (6.44) . This completes the proof of Theorem 6.4.

It is worth to mention that, the stochastic transposition method has some other applications, say it can be used to establish the equivalence between the existence of optimal feedback operator for infinite dimensional stochastic linear quadratic control problems with random coefficients and the solvability of the corresponding operator-valued, backward stochastic Riccati equations (See [16] for more details).

References

- [1] G. Da Prato and J. Zabczyk. Stochastic equations in infinite dimensions. Cambridge University Press, Cambridge, 1992.
- [2] H. Frankowska, H. Zhang and X. Zhang. First and second order necessary conditions for stochastic optimal controls. ArXiv: 1603.08274, 2016.
- [3] A. V. Fursikov and O. Yu. Imanuvilov. Controllability of evolution equations. Lecture Notes Series 34, Research Institute of Mathematics, Seoul National University, Seoul, Korea, 1994.
- [4] P. R. Halmos. Measure theory. D. Van Nostrand Company, Inc., New York, 1950.
- [5] G. Lebeau and L. Robbiano. Contrôle exact de l'équation de la chaleur. Comm. Partial Differential Equations. 20 (1995), 335–356.
- [6] G. Lebeau and E. Zuazua. Null controllability of a system of linear thermoelasticity. Arch. Rational Mech. Anal. 141 (1998), 297–329.
- [7] J. L. Lions. Optimal control of systems governed by partial differential equations. Springer-Verlag, Berlin, Heidelberg, New York, 1971.
- [8] Q. L\u00fc. Some results on the controllability of forward stochastic parabolic equations with control on the drift. J. Funct. Anal. 260 (2011), 832–851.
- [9] Q. Lü. A lower bound on local energy of partial sum of eigenfunctions for Laplace-Beltrami operators. ESAIM Control Optim. Calc. Var. 19 (2013), 255–273.

- [10] Q. Lü, J. Yong and X. Zhang. Representation of Itô integrals by Lebesgue/Bochner integrals. J. Eur. Math. Soc. 14 (2012), 1795–1823.
- [11] Q. Lü and X. Zhang. Well-posedness of backward stochastic differential equations with general filtration. J. Differential Equations. 254 (2013), 3200–3227.
- [12] Q. Lü and X. Zhang. General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions. Springer Briefs in Mathematics, Springer, New York, 2014.
- [13] Q. Lü and X. Zhang. Transposition method for backward stochastic evolution equations revisited, and its application. Math. Control Relat. Fields. 5 (2015), 529–555.
- [14] Q. Lü and X. Zhang. Mathematical theory for stochastic distributed parameter control systems. A book in preparation.
- [15] Q. Lü and X. Zhang. Control theory for stochastic distributed parameter systems: Recent progresses and open problems. A survey paper in preparation.
- [16] Q. Lü and X. Zhang. Optimal feedback for stochastic linear quadratic control and backward stochastic Riccati equations in infinite dimensions. Preprint.
- [17] N. I. Mahmudov and M. A. McKibben. On backward stochastic evolution equations in Hilbert spaces and optimal control. Nonlinear Anal. 67 (2007), 1260–1274.
- [18] S. Peng. A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim. 28 (1990), 966–979.
- [19] S. Peng. Backward stochastic differential equation and exact controllability of stochastic control systems. Progr. Natur. Sci. (English Ed.). 4 (1994), 274–284.
- [20] J. M. A. M. van Neerven, M. C. Veraar and L. W. Weis. Stochastic evolution equations in UMD Banach spaces. J. Funct. Anal. 255 (2008), 940–993.
- [21] L. C. G. Rogers and D. Williams. Diffusions, Markov processes, and martingales. Vol. 2. Itô calculus. John Wiley & Sons, Inc., New York, 1987.
- [22] S. Tang and X. Zhang, Null controllability for forward and backward stochastic parabolic equations. SIAM J. Control Optim. 48 (2009), 2191–2216.
- [23] J. Yong and X. Y. Zhou. Stochastic controls: Hamiltonian systems and HJB equations. Springer-Verlag, New York, 1999.
- [24] 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.
- [25] H. Zhang and X. Zhang. Pointwise second-order necessary conditions for stochastic optimal controls, Part II: The general case. ArXiv:1509.07995, 2015.