



HAL
open science

Optimality Conditions for Parabolic Stochastic Optimal Control Problems with Boundary Controls

Piero Visconti

► **To cite this version:**

Piero Visconti. Optimality Conditions for Parabolic Stochastic Optimal Control Problems with Boundary Controls. 2024. hal-04360452v2

HAL Id: hal-04360452

<https://hal.science/hal-04360452v2>

Preprint submitted on 14 Mar 2024

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

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

Optimality Conditions for Parabolic Stochastic Optimal Control Problems with Boundary Controls

Piero Visconti*

March 14, 2024

Abstract

Optimality conditions are provided for a class of control problems driven by a cylindrical Wiener process, which amount to a stochastic maximum principle in differential form. The control is considered to act on the drift and the volatility, both of which may be unbounded operators, which allows us to consider SPDEs with control and/or noise on the boundary. By the factorization method, a regularizing property is established for the state equation which is then employed to prove, by duality, a similar result for the backward time costate equation. The costate equation is understood in the sense of transposition. Finally, the cost is shown to be Gateaux differentiable and its derivative is represented in terms of the costate, the optimality condition is deduced using results of set valued analysis.

Keywords: Stochastic optimal control, Boundary control, Necessary optimality conditions, Backward stochastic evolution equations.

1 Introduction

The aim of this paper is to give optimality conditions in the form of a maximum principle for optimal control problems involving stochastic parabolic evolution equations. Specifically, we consider the situation where all functions defining the dynamics are allowed to take values outside the state space, which enables us to consider a broader class of boundary control problems for stochastic partial differential equations (SPDEs) than is covered in the literature.

Let \mathbb{X} and \mathbb{H} be separable Hilbert spaces, let \mathbb{U} be a Banach space, $U^{ad} \subset \mathbb{U}$ be a closed subset, and let Ω be a probability space. Suppose W is a cylindrical Wiener process on \mathbb{H} with underlying probability space Ω . Let A be the generator of an analytic semigroup on \mathbb{X} .

Consider the stochastic evolution equation (SEE)

$$dx(t) = [Ax(t) + b(t, x(t), u(t))] dt + \sigma(t, x(t), u(t))dW(t), \quad x(0) = x_0. \quad (1.1)$$

Consider as well the cost function

$$\mathcal{J}(x, u) = \mathbb{E} \int_0^T f(t, x(t), u(t))dt + \mathbb{E}h(x(T)). \quad (1.2)$$

A pair of predictable stochastic processes (x, u) with values in $\mathbb{X} \times U^{ad}$ is said to be an admissible pair if it satisfies equation (1.1), and it is an optimal pair if $\mathcal{J}(x, u)$ is minimal among admissible pairs.

*INSA Rouen Normandie, Normandie Univ, LMI UR 3226, F-76000 Rouen, France. E-mail: piero.visconti@insa-rouen.fr

If we set $\mathbb{X} = \mathbb{R}^n$ and omit σ and W , we are left with a control system governed by an ordinary differential equation, and the expectation can be omitted from the cost functional. In the influential work [5], the costate equation

$$\dot{\bar{y}}(t) = -[A^*y + b_x(t, \bar{x}(t), \bar{u}(t))^* \bar{y}(t) - f_x(t, \bar{x}(t), \bar{u}(t))], \quad y(T) = -h_x(\bar{x}(T)), \quad (1.3)$$

was introduced, as well as the Hamiltonian

$$H(t, x, u, y) = \langle y, b(t, x, u) \rangle_{\mathbb{X}^*, \mathbb{X}} - f(t, x, u), \quad (1.4)$$

and it was established that under mild assumptions (most notably differentiability with respect to (w.r.t.) x and continuity w.r.t. u of the functions b , f and h), for a given optimal pair (\bar{x}, \bar{u}) , and the respective solution \bar{y} of (1.3), the function

$$U^{ad} \ni u \mapsto H(t, \bar{x}(t), u, \bar{y}(t))$$

is maximized at $\bar{u}(t)$ a.e. in $[0, T]$. The power of this results is that it reduces an optimization problem in infinite dimensions (the space of controls) to a family of optimization problems in U^{ad} . We refer to [32, 7].

For optimal control in infinite dimensions, we refer to the monographs [22, 31]. More specifically, for semilinear parabolic equations with boundary control, we refer to [6, 28, 29], it is worth noting that the adjoint equation is also parabolic and typically the same regularity holds for its solutions as for those of the state equation, this is of particular importance when considering boundary controls, as the optimality conditions often involve a trace of the costate.

We briefly consider again $\mathbb{X} = \mathbb{R}^n$ and now let $\mathbb{H} = \mathbb{R}^d$ and consider σ in the control system again, the dynamics now involve two expressions, as a consequence, the costate includes an additional process q , taking values in $R^{n \times d}$. The costate equation is given by

$$\begin{aligned} dy(t) &= -[A^*y + b_x(t, x(t), u(t))^* y(t) + \sigma_x(t, x(t), u(t))^* q(t) - f_x(t, x(t), u(t))]dt + q(t)dW(t), \\ y(T) &= -h_x(x(T)). \end{aligned} \quad (1.5)$$

In this case, the process q is part of the solution, and the costate corresponds to the pair of processes (y, q) .

We define the Hamiltonian

$$H(t, x, u, y, q) = \langle y, b(t, x, u) \rangle_{\mathbb{X}^*, \mathbb{X}} + tr [q^* \sigma(t, x, u)] - f(t, x, u), \quad (1.6)$$

where tr denotes the trace of an matrix or operator (recall that $R^{n \times d}$ is a Hilbert space with Frobenius product).

The naive extension of the maximum principle would say that for any optimal pair (\bar{x}, \bar{u}) , there are solutions (\bar{y}, \bar{q}) of equation (1.5) such that the stochastic Hamiltonian (1.6) is maximized at $\bar{u}(t)$ a.e. in $[0, T] \times \Omega$. However, even in finite dimensions, the Hamiltonian is not necessarily maximized at an optimal control (see [34, Chapter 3, Example 3.1] for a simple counterexample with $\mathbb{X} = \mathbb{U} = \mathbb{H} = \mathbb{R}$). A maximum principle can be established under additional assumptions, one may point as very early examples to [20, 19] (σ independent of the control) or [4] (convexity assumptions).

In [27], the author proves the maximum principle for general stochastic differential equations, in such generality, it is necessary to impose more regularity for b, σ, f and h (up to second derivatives

w.r.t. x) and introduce a second order adjoint equation with solutions (Y, Q) taking values in $\mathcal{L}(\mathbb{X}) \times \mathcal{L}(\mathbb{H}, \mathcal{L}(\mathbb{X}))$ and add to the Hamiltonian the quadratic correction term

$$\frac{1}{2} \text{tr}[(\sigma(t, \bar{x}(t), u) - \sigma(t, \bar{x}(t), \bar{u}(t)))^* Y(t) (\sigma(t, \bar{x}(t), u) - \sigma(t, \bar{x}(t), \bar{u}(t)))). \quad (1.7)$$

We refer the reader to [34] and the references therein for a revision of what is by now the classical theory of stochastic optimal control in finite dimension. We mention [14], where the authors are able to obtain a first order optimality condition involving only first derivatives of b, σ, f and h without the assumption that the admissible control set U^{ad} is convex, the present paper takes a similar approach.

We revert back now to our original setting in which \mathbb{X} and \mathbb{H} are separable Hilbert spaces and \mathbb{U} is a Banach space. We refer to [2] for a pioneering work, where it is shown that, while a maximum principle doesn't necessarily hold, assuming differentiability w.r.t. u and imposing that the admissible control set is a convex subset of a Hilbert space \mathbb{U} , the derivative of the Hamiltonian at an optimal control is not positive along feasible directions.

See [25] and the rich references therein for a recent review on the state of control theory for infinite dimensional stochastic control systems. We refer in particular to [12, 15, 24], all of which give meaning (in different ways) to the correction term (1.7) in the infinite dimensional case, this is a delicate technical question and in the case that σ is unbounded (i.e. does not belong to $\mathcal{L}(\mathbb{H}, \mathbb{X})$), it would not be enough to have Y take values in $\mathcal{L}(\mathbb{X})$, one should prove that it is in some sense regularizing, this kind of well-posedness for the second order adjoint is not available at the moment, we therefore aim for an optimality condition in the spirit of [2, 14], making use of only one adjoint equation.

Unlike in the deterministic setting, backward time stochastic equations are not equivalent to their forward time counterparts through a simple change of variable as the adaptedness of solutions is lost in the process, however, much like in the deterministic case, it is possible to establish additional regularity in the parabolic case. This regularity for the costate is relevant because, a priori, the Hamiltonian as found in equation (1.6) is not well defined when b and σ are unbounded. For a maximum principle involving stochastic boundary control systems we refer to [18, 35] where additional regularity of the costate y is shown to hold and optimality conditions are established. In order to treat the case where σ depends on the control, the properties of the process q become relevant, even when σ is assumed to take values in $\mathcal{L}(\mathbb{H}, \mathbb{X})$ it could be necessary to establish additional regularity for q , for instance, when the noise is infinite dimensional (see [16], where the authors need to prove an integrability condition for the trace of q). As far as we know, there is no work establishing optimality conditions for a controlled SEE with σ being simultaneously unbounded and depending on the control. The main novelty of the present paper is to consider unbounded b and σ , both possibly depending on u .

The rest of the paper is organized as follows: In section 2 we introduce the notation and some preliminary concepts. In section 3, we provide a mathematical setting and define the optimal control problem considered. In section 4, we provide well posedness results for the state equation, in doing so, we refine the known results for the regularity of solutions of parabolic stochastic evolution equations, which may be of independent interest. In section 5 we prove the well-posedness and regularity of the solutions of the costate equation. Section 6 concerns the optimality conditions of the solutions of our optimal control problem, finally, we give two worked out examples for which we write our costate equation and optimality conditions.

2 Notation and preliminaries

2.1 General notation

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. We denote by $\mathcal{L}(X, Y)$ the space of bounded linear operators from X to Y , which we endow with the operator norm:

$$\|B\|_{\mathcal{L}(X, Y)} = \sup\{\|Bx\|_Y : \|x\|_X \leq 1\}.$$

Denote $\mathcal{L}(X) := \mathcal{L}(X, X)$, we denote $Id_X \in \mathcal{L}(X)$ the identity operator $Id_X x = x$, when it is clear from context, we omit the subscript. We define the dual space $X^* := \mathcal{L}(X, \mathbb{R})$. For $x^* \in X^*$ we denote $x^*(x) = \langle x, x^* \rangle_{X, X^*}$.

For $B \in \mathcal{L}(X, Y)$, define the adjoint operator $B^* \in \mathcal{L}(Y^*, X^*)$, given by

$$\langle x, B^* y^* \rangle_{X, X^*} := \langle Bx, y^* \rangle_{Y, Y^*}.$$

2.2 Analytic semigroups and interpolation spaces

Let A be the infinitesimal generator of an analytic semigroup S . For large enough $\beta \in \rho(A) \cap \mathbb{R}$, we have

$$(\beta Id - A)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-\beta t} S(t) dt \quad (2.1)$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$.

We denote $X_\alpha = D((\beta Id - A)^\alpha)$ with the norm $\|x\|_{X_\alpha} = \|(\beta - A)^\alpha x\|_X$ and $X_{-\alpha}$ the completion of X with the norm $\|(\beta Id - A)^{-\alpha} x\|_X$.

Analytic semigroups have the following regularizing property:

Proposition 2.1 ([26, Chapter 2, Theorem 6.13]). *Let S be an analytic semigroup and $\{X_\alpha\}_\alpha$ defined as before. Then for all $t > 0$ and $\alpha \geq 0$ we have $S(t) \in \mathcal{L}(X, X_\alpha)$. Moreover, there exists M_α such that for all $t > 0$,*

$$\|S(t)\|_{\mathcal{L}(X, X_\alpha)} \leq M_\alpha t^{-\alpha} e^{\beta t}.$$

2.3 Stochastic processes

Let $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ be a probability space and let the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ be an increasing collection of σ -algebras contained in \mathcal{F}_∞ .

We define \mathcal{F}_{pr} to be the smallest σ -algebra on $\mathbb{R}_+ \times \Omega$ containing all sets of the form $(s, t] \times B_s$ with $B_s \in \mathcal{F}_s$ and $s < t$, and of the form $\{0\} \times B_0$ for $B_0 \in \mathcal{F}_0$.

A function $f : \mathbb{R}_+ \times \Omega \rightarrow M$ where (M, \mathcal{F}_M) is any measurable space is said to be \mathbb{F} -predictable if it's measurable w.r.t. the predictable σ -algebra \mathcal{F}_{pr} , i.e. if $f^{-1}(B) \in \mathcal{F}_{pr}$ for all $B \in \mathcal{F}_M$. We write $M_{\mathbb{F}}$ for the set of M -valued predictable functions. When M is a topological space (in particular, for a Banach space), we consider on it the Borel σ -algebra.

For $T \in (0, \infty]$, $1 \leq p, q \leq \infty$ and a Banach space X we will use the following notations:

- $L_{\mathbb{F}}^p(\Omega; L^q(0, T; X)) = \left\{ x \in X_{\mathbb{F}} : \mathbb{E} \left(\int_0^T \|x_t\|_X^q dt \right)^{p/q} < \infty \right\}$ with norm $\left[\mathbb{E} \left(\int_0^T \|(\cdot)_t\|_X^q dt \right)^{p/q} \right]^{1/p}$.
- $L_{\mathbb{F}}^q(0, T; L^p(\Omega; X)) = \left\{ x \in X_{\mathbb{F}} : \int_0^T (\mathbb{E} \|x_t\|_X^p)^{q/p} dt < \infty \right\}$ with norm $\left[\int_0^T (\mathbb{E} \|(\cdot)_t\|_X^p)^{q/p} dt \right]^{1/q}$.

Both defined up to equality a.e. in $[0, T] \times \Omega$ with the obvious adjustment when p or q are ∞ .

- $L_{\mathbb{F}}^p(\Omega; C([0, T], X))$ is the subspace of $L_{\mathbb{F}}^p(\Omega; L^\infty(0, T; X))$ consisting of processes x which almost surely have continuous paths, i.e. for almost every $\omega \in \Omega$, $t \mapsto x_t(\omega)$ is continuous from $[0, T]$ to X .
- $D_{\mathbb{F}}([0, T], L^p(\Omega; X))$ is the subspace of $L_{\mathbb{F}}^\infty(0, T; L^p(\Omega; X))$ consisting of processes x which are càdlàg (right continuous with left limits) from $[0, T]$ into $L_{\mathcal{F}_T}^p(\Omega; X)$, i.e. for every $t \in [0, T]$,

$$\lim_{s \rightarrow t^+} \mathbb{E} \|x_t - x_s\|^p = 0$$

and the limit $\lim_{s \rightarrow t^-} x_s$ in $L_{\mathcal{F}_t}^p(X)$ exists.

- $C_{\mathbb{F}}([0, T], L^p(\Omega; X))$ is the subspace of $L_{\mathbb{F}}^\infty(0, T; L^p(\Omega; X))$ consisting of processes x which are continuous from $[0, T]$ into $L_{\mathcal{F}_T}^p(\Omega; X)$, i.e. for every $t \in [0, T]$,

$$\lim_{s \rightarrow t} \mathbb{E} \|x_t - x_s\|^p = 0.$$

2.4 Hilbert-Schmidt space & stochastic integration

Let Z and K be Hilbert spaces. By the Riesz representation theorem, it is possible to identify Z with Z^* and K with K^* , and thus the adjoint P^* of a bounded operator $P \in \mathcal{L}(Z, K)$ may be thought of as belonging to $\mathcal{L}(K, Z)$, and it is defined by the relation $\langle z, P^*k \rangle_Z = \langle Pz, k \rangle_K$.

Let I be an arbitrary set of indices, we say that $\{e_i\}_{i \in I} \subset Z$ is a Hilbert basis of Z if it's an orthonormal set, i.e.

$$\langle e_i, e_j \rangle_Z = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

and $\{e_i\}_{i \in I}$ spans a dense subset of Z .

Fix a Hilbert basis $\{e_i\}_{i \in I}$. For any positive bounded operator $T_+ \in \mathcal{L}(Z)$, i.e. one satisfying $\langle T_+z, z \rangle_Z \geq 0$ for all $z \in Z$, we define

$$\text{tr}(T_+) = \sum_{i \in I} \langle T_+e_i, e_i \rangle_Z.$$

It can be checked that $\text{tr}(T_+)$ doesn't depend on the choice of Hilbert basis $\{e_i\}_{i \in I}$. Define the space $\mathcal{L}_1(Z)$ of trace-class operators on Z as

$$\mathcal{L}_1(Z) := \{T \in \mathcal{L}(Z) : \text{tr}(|T|) < \infty\},$$

where $|T| = (T^*T)^{1/2}$, then $\text{tr}(T)$ is also well defined by the same formula for $T \in \mathcal{L}_1(Z)$.

We denote $\mathcal{L}_2(Z, K)$ the space of Hilbert-Schmidt operators $P : Z \rightarrow K$, defined as

$$\mathcal{L}_2(Z, K) := \{P \in \mathcal{L}(Z, K) : P^*P \in \mathcal{L}_1(Z)\} = \{P \in \mathcal{L}(Z, K) : \text{tr}(P^*P) < \infty\}.$$

We list below some facts about Hilbert-Schmidt and trace-class operators (see for instance [30, Section VI.6]):

Proposition 2.2. *Let Z and K be Hilbert spaces. The following statements hold.*

- *Every Hilbert-Schmidt operator is compact.*
- *Every trace-class operator is Hilbert-Schmidt: $\mathcal{L}_1(Z) \subset \mathcal{L}_2(Z)$. In particular, every trace-class operator is compact.*

- Let $T \in \mathcal{L}_1(Z)$. There exists (at least) a Hilbert basis $\{e_i\}_{i \in I}$ of Z such that

$$\sum_{i \in I} \|Te_i\|_Z \leq \infty.$$

One such basis is the one consisting of eigenvectors of $|T|$ (whose existence follows from Hilbert's spectral decomposition theorem).

- The composition QP of two Hilbert-Schmidt operators $P \in \mathcal{L}_2(Z, K)$ and $Q \in \mathcal{L}_2(K, Z)$ is a trace-class operator on Z .
- $\mathcal{L}_2(Z, K)$ is itself a Hilbert space with the inner product $(P, Q) = \text{tr}(Q^*P)$, which is again well defined by

$$\text{tr}(Q^*P) = \sum_{i \in I} \langle Q^*Pe_i, e_i \rangle_Z = \sum_{i \in I} \langle Pe_i, Qe_i \rangle_K$$

and is independent of the choice of basis. If Z, K are separable, $\mathcal{L}_2(Z, K)$ is again separable.

We have the following identification:

Proposition 2.3. *Assume Z, K are Hilbert spaces. Then $(\mathcal{L}_2(Z, K))^* = \mathcal{L}_2(Z, K^*)$ with the duality*

$$\langle P, Q \rangle_{\mathcal{L}_2(Z, K), \mathcal{L}_2(Z, K^*)} = \sum_{i \in I} \langle Pe_i, Qe_i \rangle_{K, K^*},$$

for an orthonormal basis $\{e_i\}_{i \in I}$ of Z . Moreover, if H is another Hilbert space densely and continuously included in K , then $\mathcal{L}_2(Z, H) \subset \mathcal{L}_2(Z, K)$ and $\mathcal{L}_2(Z, K^*) \subset \mathcal{L}_2(Z, H^*)$ with dense and continuous inclusions and

$$\langle P, Q \rangle_{\mathcal{L}_2(Z, K), \mathcal{L}_2(Z, K^*)} = \langle P, Q \rangle_{\mathcal{L}_2(Z, H), \mathcal{L}_2(Z, H^*)}$$

when $P \in \mathcal{L}_2(Z, H)$ and $Q \in \mathcal{L}_2(Z, K^*)$.

See appendix A for a proof.

We now define a cylindrical Wiener process. Let \mathbb{H} be a separable Hilbert space densely embedded in another separable Hilbert space \mathbb{H}_{-1} such that the inclusion $\mathcal{I} : \mathbb{H} \rightarrow \mathbb{H}_{-1}$ is Hilbert-Schmidt. Let $W : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{H}_{-1}$ be a measurable function, we write $W(\omega, t) = W_t(\omega)$ and often omit dependence on ω . We say that W is a cylindrical Wiener process on \mathbb{H} if for all $h, k \in \mathbb{H}_{-1}$, $t, s \in \mathbb{R}_+$, $\langle W_t, h \rangle_{\mathbb{H}_{-1}}$ is a centered Gaussian random variable, $W(\omega, \cdot) \in C(\mathbb{R}_+, \mathbb{H}_{-1})$ a.s. and

$$\mathbb{E}_{\mathbb{P}} [\langle W_t, h \rangle_{\mathbb{H}_{-1}} \langle W_s, k \rangle_{\mathbb{H}_{-1}}] = (t \wedge s) \langle \mathcal{I}^*h, \mathcal{I}^*k \rangle_{\mathbb{H}}.$$

We say the filtration \mathbb{F} is admissible for W if for all $t \geq 0$, W_t is \mathcal{F}_t -measurable and $W_s - W_t$ is independent of \mathcal{F}_t for any $s > t$. For a given process W , an admissible filtration can always be constructed, the smallest such filtration is called the natural filtration of W and it consists of $\{\mathcal{F}_t^W\}_{t \in \mathbb{R}_+}$, where \mathcal{F}_t^W is the smallest σ -algebra for which W_s is measurable for all $s \in [0, t]$.

In the rest of the paper, we assume the filtration \mathbb{F} is admissible for W .

For a Hilbert space K , the Itô stochastic integral is then defined for predictable $\mathcal{L}_2(\mathbb{H}, K)$ -valued processes which belong to $L_{loc}^2(0, \infty; \mathcal{L}_2(\mathbb{H}, K))$ a.s. (see for instance [10, Part I, Chapter 4]).

3 Setting of the problem

In this section we introduce the functional setting for our control problem, we then give conditions under which the state equation is well-posed and the cost is well defined, we also provide some estimates that will be useful in the rest of the paper.

Let \mathbb{X} be a separable Hilbert space, $A : D(A) \subset \mathbb{X} \rightarrow \mathbb{X}$ is the infinitesimal generator of an analytic semigroup. For $s \in \mathbb{R}$, write $\mathbb{X}_s = D((\alpha Id - A)^s)$ and $\mathbb{X}_s^* = D((\alpha Id - A^*)^s)$ for $\alpha \in \mathbb{R}$ large enough, so that $(\mathbb{X}_s)^* = \mathbb{X}_{-s}^*$ (we don't identify \mathbb{X} with \mathbb{X}^*).

Let \mathbb{H} be a separable Hilbert space and W a cylindrical Wiener process on \mathbb{H} . Let $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ a probability space endowed with a right-continuous filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ (i.e. $\mathcal{F}_t = \cap_{s>t} \mathcal{F}_s$ for every $t \in \mathbb{R}_+$) which is admissible for W .

The control space \mathbb{U} is assumed to be a reflexive and separable Banach space.

Consider the state equation

$$dx(t) = [Ax(t) + b(t, x(t), u(t))] dt + \sigma(t, x(t), u(t)) dW(t), \quad x(0) = x_0 \in L^p_{\mathcal{F}_0}(\Omega; \mathbb{X}). \quad (3.1)$$

Hypothesis 3.1. Assume $p > 2$ and $\eta \in \mathbb{R}$. Let $b : [0, T] \times \mathbb{X} \times \mathbb{U} \times \Omega \rightarrow \mathbb{X}_{-1/2}$ and $\sigma : [0, T] \times \mathbb{X} \times \mathbb{U} \times \Omega \rightarrow \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})$. Assume the following conditions hold:

1.

$$\eta < \frac{1}{2} - \frac{1}{p} \quad (3.2)$$

2. $b(\cdot, x, u), \sigma(\cdot, x, u)$ are predictable for any $x \in \mathbb{X}, u \in \mathbb{U}$.

3. b, σ are Lipschitz continuous w.r.t. x, u , i.e. for some $C_b, C_\sigma \geq 0$

$$\|b(t, x_1, u_1) - b(t, x_2, u_2)\|_{\mathbb{X}_{-1/2}} \leq C_b (\|x_1 - x_2\|_{\mathbb{X}} + \|u_1 - u_2\|_{\mathbb{U}}), \quad (3.3)$$

$$\|\sigma(t, x_1, u_1) - \sigma(t, x_2, u_2)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} \leq C_\sigma (\|x_1 - x_2\|_{\mathbb{X}} + \|u_1 - u_2\|_{\mathbb{U}}). \quad (3.4)$$

4. $b(\cdot, 0, 0) \in L^2_{\mathbb{F}}(0, T; L^p(\Omega; \mathbb{X}_{-1/2}))$, $\sigma(\cdot, 0, 0) \in L^p_{\mathbb{F}}(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})))$.

Definition 3.2. Let $x_0 \in \mathbb{X}$ a.s. A predictable \mathbb{X} -valued process x is said to be a **mild solution** of (3.1) if for every $t \geq 0$ we have $S(t-s)b(s, x(s), u(s)) \in L^1(0, t; \mathbb{X})$ a.s., $S(t-s)\sigma(s, x(s), u(s)) \in L^2(0, t; \mathcal{L}_2(\mathbb{H}, \mathbb{X}))$ a.s. and

$$x(t) = S(t)x_0 + \int_0^t S(t-s)b(s, x(s), u(s))ds + \int_0^t S(t-s)\sigma(s, x(s), u(s))dW(s) \quad a.s. \quad (3.5)$$

Definition 3.2 is similar to the notion of mild solution given in [10, Chapter 7], the difference is that we have relaxed the assumption that $b(s, x(s), u(s))$ takes values in \mathbb{X} , and that $\sigma(s, x(s), u(s))$ takes values in $\mathcal{L}_2(\mathbb{H}, \mathbb{X})$, we assume instead that they belong to the larger spaces $\mathbb{X}_{-1/2}$ and $\mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})$. We don't assume that $\sigma(s, x(s), u(s))$ takes values in $\mathcal{L}(\mathbb{H}, \mathbb{X})$ either.

Example 3.3. Let W_{st} be a cylindrical Wiener process on $L^2(0, 1)$ (from now on $L^p := L^p(0, 1)$), so that \dot{W}_{st} is (formally) a space-time white noise. Let w_0 and w_1 be 1-D Wiener processes, independent of each other and of W_{st} .

Consider the equation

$$\begin{aligned} \partial_t x &= \partial_{\xi\xi} x + b^{st}(t, \xi, x, u_{st}) + \sigma^{st}(t, \xi, u_{st}) \dot{W}_{st}(t, \xi) & t \in (0, T), \xi \in (0, 1) \\ \partial_n x(t, i) &= b^i(t, u_i) + \sigma^i(t, u_i) \dot{w}_i(t) & t \in (0, T), i = 0, 1 \\ x(0, \xi) &= x_0(\xi) & \xi \in (0, 1) \end{aligned} \quad (3.6)$$

where $x = x(t, \xi)$, $u_{st} = u_{st}(t, \xi)$ and $u_i = u_i(t)$, $i = 0, 1$.

1. Let $b^{st} : [0, T] \times \Omega \times (0, 1) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuously differentiable and uniformly Lipschitz continuous w.r.t. x, u a.e. in $[0, T] \times \Omega \times (0, 1)$. Assume $b^{st}(\cdot, \xi, x, u_{st})$ is predictable for fixed values of (ξ, x, u_{st}) and $b^{st}(\cdot, \xi, 0, 0) \in L^p([0, T] \times \Omega \times (0, 1))$.
2. Let $b^0 : [0, T] \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuously differentiable and uniformly Lipschitz continuous w.r.t. u a.e. in $[0, T] \times \Omega$. Assume $b^0(\cdot, u_0)$ is predictable for fixed values of u_0 and $b^0(\cdot, 0) \in L^p(\Omega L^2(0, T))$.
3. Let $b^1 : [0, T] \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuously differentiable and uniformly Lipschitz continuous w.r.t. u a.e. in $[0, T] \times \Omega$. Assume $b^1(\cdot, u_1)$ is predictable for fixed values of u_1 and $b^1(\cdot, 0) \in L^p(\Omega L^2(0, T))$.
4. Let $\sigma^{st} : [0, T] \times \Omega \times (0, 1) \times \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuously differentiable and uniformly Lipschitz continuous w.r.t. u a.e. in $[0, T] \times \Omega \times (0, 1)$. Assume $\sigma^{st}(\cdot, \xi, u_{st})$ is predictable for fixed values of (ξ, u_{st}) and $\sigma^{st}(\cdot, \xi, 0) \in L^p([0, T] \times \Omega \times (0, 1))$.
5. Let $\sigma^0 : [0, T] \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuously differentiable and uniformly Lipschitz continuous w.r.t. u a.e. in $[0, T] \times \Omega$. Assume $\sigma^0(\cdot, u_0)$ is predictable for fixed values of u_0 and $\sigma^0(\cdot, 0) \in L^p([0, T] \times \Omega)$.
6. Let $\sigma^1 : [0, T] \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuously differentiable and uniformly Lipschitz continuous w.r.t. u a.e. in $[0, T] \times \Omega$. Assume $\sigma^1(\cdot, u_1)$ is predictable for fixed values of u_1 and $\sigma^1(\cdot, 0) \in L^p([0, T] \times \Omega)$.

We consider $A = \partial_{\xi\xi}$ and $D(A) = \{x \in H^2(0, 1) : \partial_\xi x(0) = \partial_\xi x(1) = 0\}$, so that A generates an analytic semigroup on L^2 . Define $\mathcal{H}_{2s} = D((\alpha Id - A)^s)$ for a fixed $\alpha > 0$. With this, $\mathcal{H}_s = H^s(0, 1)$ for $s \in (-\frac{1}{2}, \frac{3}{2})$ (see [21, Theorem 3A.1]). Let $r = \frac{2}{p-2}$, notice that due to [11, Theorem 6.7], we have the inclusions $\mathcal{H}_{1/p} \subset L^r \subset L^2 \subset L^{r^*} \subset \mathcal{H}_{-1/p}$, where $\frac{1}{r} + \frac{1}{r^*} = 1$.

The function $b^{st}(t, \xi, \cdot, \cdot)$ is Gateaux differentiable w.r.t. (x, u) from $L^2 \times L^p$ to L^2 . Its derivatives at x, u are the multiplications by $b_x^{st}(t, \xi, x(\xi), u(\xi))$ and $b_u^{st}(t, \xi, x(\xi), u(\xi))$.

Notice that for $\Sigma^{st} \in L^p$, the product $h \mapsto \Sigma^{st}h$ maps L^2 linearly and continuously into L^{r^*} , and therefore, into $\mathcal{H}_{-1/p}$.

It follows from the spectral decomposition of A that when $\eta > \frac{1}{4} + \frac{1}{2p}$, the inclusion from $\mathcal{H}_{-1/p}$ into $\mathcal{H}_{-2\eta}$ is Hilbert-Schmidt.

The function $\sigma^{st}(t, \xi, \cdot)$ is Gateaux differentiable w.r.t. u from L^p to L^p , and its derivative at u is the multiplication by $\sigma_u^{st}(t, \xi, u(\xi))$. Therefore, the mapping $u \mapsto [h \mapsto \sigma^{st}(t, \cdot, u(\cdot))h]$ is Gateaux differentiable w.r.t. u from L^p to $\mathcal{L}_2(L^2, \mathcal{H}_{-2\eta})$.

For $u \in \mathbb{R}^2$, define $b^\partial(t, u) = (\alpha Id - A)z$ where

$$\alpha z - \partial_{\xi\xi} z = 0, \quad \partial_n z(i) = b^i(t, u_i) \quad i = 0, 1,$$

so that $b^\partial(t, \cdot)$ is Gateaux differentiable from \mathbb{R}^2 into $\mathcal{H}_{-\beta}$ for $\beta > 1/2$ (see for instance [21, Section 3.3]). Notice that, integrating by parts, the adjoint $(b_u^\partial(t, u))^*$ is then the following Dirichlet trace operator

$$(b_u^\partial(t, u))^* y = (b_u^0(t, u)y(0), b_u^1(t, u)y(1)).$$

Define as well $[\sigma^\partial(t)u]w = (\alpha Id - A)z$ where

$$\alpha z - \partial_{\xi\xi} z = 0, \quad \partial_n z(i) = \sigma^i(t, u_i)w_i \quad i = 0, 1,$$

we have that $\sigma^\partial(t, \cdot)$ is Gateaux differentiable from \mathbb{R}^2 into $\mathcal{L}(\mathbb{R}^2, \mathcal{H}_{-\beta}) \simeq \mathcal{L}_2(\mathbb{R}^2, \mathcal{H}_{-\beta})$ for $\beta > 1/2$. Notice that, integrating by parts, the adjoint $(\sigma_u^\partial(t, u))^*$ is then the following Dirichlet trace operator

$$(\sigma_u^\partial(t, u))^*(q_0, q_1) = (\sigma_u^0(t, u)q_0(0), \sigma_u^1(t, u)q_1(1)).$$

Define $\mathbb{H} = L^2 \oplus \mathbb{R}^2$ so that $W = (W_{st}, w_0, w_1)$ is a cylindrical Wiener process on \mathbb{H} .

In the notation of 6.2, denote $\mathbb{X}_s = \mathcal{H}_{2s}$.

Denote also $U^{\text{ad}} := \mathbb{U} = L^p \oplus \mathbb{R}^2$, we write $u = (u_{st}, u_0, u_1)$, and define

$$b(t, x, u) = b^{st}(t, \xi, x, u_{st}) + b^\partial(t, u_0, u_1) \in \mathcal{H}_{-1}.$$

Similarly, we define

$$\sigma(t, x, u) = (\sigma^{st}(t, \xi, u_{st}) \quad \sigma^\partial(t, u_0, u_1)) \in \mathcal{L}_2(\mathbb{H}, \mathcal{H}_{-2\eta}).$$

We have hypothesis 6.2 for $\eta \in (\frac{1}{4}, \frac{1}{2})$ and $p > \max\{\frac{2}{1-2\eta}, \frac{2}{4\eta-1}\}$ (this imposes $p > 6$, and p may be chosen arbitrarily close to 6 by choosing $\eta = 1/3$).

This SPDE is similar to some of those considered [9], with the difference that we consider a variable volatility term on the boundary condition which is controlled.

It is worth mentioning that in [16], the authors take hypotheses similar to 3.1, however, neither setting completely encompasses the other. In the particular case that A is analytic, [16, Assumption 3.3] implies that the inclusion from \mathbb{X} into $\mathbb{X}_{-\eta}$ is Hilbert-Schmidt for some $\eta \in (0, \frac{1}{2})$, and therefore [16, Assumption 2.1] implies hypothesis 3.1.

The advantage of the present approach is twofold: In removing the assumptions that b and σ take values in \mathbb{X} and $\mathcal{L}(\mathbb{H}, \mathbb{X})$, this setting gives flexibility in choosing a functional setting for a concrete equation, for instance, in example 3.3 we are able to consider the control space to be $L^p(0, 1)$ (which is reflexive) instead of $L^\infty(0, 1)$. More importantly, our approach yields improved regularity for the costate, which allows us to define the Hamiltonian and obtain optimality conditions even when $b(s, x(s), u(s)) \notin \mathbb{X}$ and $\sigma(s, x(s), u(s)) \notin \mathcal{L}(\mathbb{H}, \mathbb{X})$, both of which happen, for instance, in problems with a control on the boundary whose action is perturbed multiplicatively by a noise.

For $x \in L^p_{\mathbb{R}}(\Omega; C([0, T], \mathbb{X}))$ and $u \in L^p_{\mathbb{R}}(\Omega; L^p(0, T; \mathbb{U}))$ (chosen freely, not solving any particular equation) define

$$\mathcal{J}(x, u) = \mathbb{E} \int_0^T f(t, x(t), u(t)) dt + \mathbb{E} h(x(T)). \quad (3.7)$$

The following will be our working hypothesis in order to have a well defined cost function.

Hypothesis 3.4. *Let $p > 2$. The mappings $f : [0, T] \times \mathbb{X} \times \mathbb{U} \times \Omega \rightarrow \mathbb{R}$ and $h : \mathbb{X} \times \Omega \rightarrow \mathbb{R}$ satisfy the following:*

1. $f(\cdot, x, u)$ is predictable and $h(x)$ is \mathcal{F}_T -measurable for any $x \in \mathbb{X}$, $u \in \mathbb{U}$.
2. For some $C_0 \geq 0$:

$$|f(t, x, u)| \leq C_0(1 + \|u\|_{\mathbb{U}}^p + \|x\|_{\mathbb{X}}^p).$$

$$|h(x)| \leq C_0(1 + \|x\|_{\mathbb{X}}^p).$$

3. For some $C_1 \geq 0$:

$$|f(t, x_1, u_1) - f(t, x_2, u_2)| \leq C_1 (1 + \|u_1\|_{\mathbb{U}} + \|u_2\|_{\mathbb{U}} + \|x_1\|_{\mathbb{X}} + \|x_2\|_{\mathbb{X}})^{p-1} (\|x_1 - x_2\|_{\mathbb{X}} + \|u_1 - u_2\|_{\mathbb{U}}).$$

$$|h(x_1) - h(x_2)| \leq C_1 (1 + \|x_1\|_{\mathbb{X}} + \|x_2\|_{\mathbb{X}})^{p-1} \|x_1 - x_2\|_{\mathbb{X}}$$

We now define our optimal control problem. Let $U^{ad} \subset \mathbb{U}$ be a closed subset. Define

$$\mathcal{U}^{ad} := \left\{ u \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U})) : u \in U^{ad} \text{ a.e. } [0, T] \times \Omega \right\} \quad (3.8)$$

Problem \mathcal{P} . Let $x_0 \in L_{\mathbb{F}}^p(\Omega; \mathbb{X})$. Find a pair $(\tilde{x}, \tilde{u}) \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X})) \times \mathcal{U}^{ad}$ which solve equation (3.1), such that for all $(x, u) \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X})) \times \mathcal{U}^{ad}$ also solving equation (3.1) we have

$$\mathcal{J}(\tilde{x}, \tilde{u}) \leq \mathcal{J}(x, u).$$

We call such (\tilde{x}, \tilde{u}) an optimal pair, \tilde{x} an optimal trajectory and \tilde{u} an optimal control for the problem.

4 State equation

We will now establish the basic existence and uniqueness results for our state equation. In order to do so, we first refine some results on the existing literature for stochastic evolution equations. Specifically, regarding the regularizing properties of the stochastic convolution which defines the mild solution of equation (3.1).

4.1 Well-posedness for parabolic stochastic evolution equations

The aim of this subsection is to extend the regularity results in [10, Theorem 5.15] to the case where the stochastic convolution has a source that isn't constant, in order to be able to incorporate the action of the control. Proposition 4.2 states the regularity for the stochastic convolution, and mixes techniques from the proof of [10, Theorem 5.15] and that of [10, Proposition 7.3]. We then show the regularity of mild solutions for linear and semilinear equations, in propositions 4.3 and 4.4 respectively.

We first establish a generalized version of the factorization formula [10, Theorem 5.11], here we allow the integrand to be a stochastic process, and we allow its values to be unbounded operators. The proof is essentially contained in that of [17, Proposition 3.2], but in said proposition the authors consider $g : [0, T] \times \Omega \rightarrow \mathcal{L}(\mathbb{H}, \mathbb{X})$.

Lemma 4.1. Suppose A is the infinitesimal generator of a C^0 -semigroup S , write $\mathbb{X}_1 = D(A)$ and $\mathbb{X}_{-1} = (D(A^*))^*$, so that $\mathbb{X}_1 \subset \mathbb{X} \subset \mathbb{X}_{-1}$. Assume for some $\alpha \in (0, \frac{1}{2})$, $p > \frac{1}{\alpha}$, $g : [0, T] \times \Omega \rightarrow \mathcal{L}(\mathbb{H}, \mathbb{X}_{-1})$ is predictable, $S(t-s)g(s) \in \mathcal{L}_2(\mathbb{H}, \mathbb{X})$ a.e. on $\{0 < s < t < T\} \times \Omega$ and

$$\int_0^T \left(\mathbb{E} \left[\int_0^t (t-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{p/2} dt < \infty. \quad (4.1)$$

Then $x(t) = \int_0^t S(t-s)g(s)dW(s)$ has a.s. continuous paths in \mathbb{X} . Moreover, $x \in L^p(\Omega; C([0, T], \mathbb{X}))$ and

$$\|x\|_{L^p_{\mathbb{F}}(\Omega; C([0, T], \mathbb{X}))} \leq C \left(\int_0^T \left(\mathbb{E} \left[\int_0^t (t-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{p/2} dt \right)^{1/p}$$

with C only depending on A, α, p, T (and not on g), moreover, $C \rightarrow 0$ as $T \rightarrow 0$.

Proof. It is clear that $y(t) = \int_0^t (t-s)^{-\alpha} S(t-s)g(s)dW(s)$ is well defined for almost all $t \in [0, T]$ and by the Burkholder-Davis-Gundy inequality ([10, theorem 4.36])

$$(\mathbb{E}[\|y(t)\|_{\mathbb{X}}^p])^{1/p} \leq C_p \left(\mathbb{E} \left[\int_0^t (t-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{1/2},$$

for any $p > \frac{1}{\alpha}$, with which

$$\begin{aligned} & \left(\mathbb{E} \left[\int_0^T \|y(t)\|_{\mathbb{X}}^p dt \right] \right)^{1/p} = \left(\int_0^T \mathbb{E} [\|y(t)\|_{\mathbb{X}}^p] dt \right)^{1/p} \\ & \leq C_p \left(\int_0^T \mathbb{E} \left[\int_0^t (t-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right]^{p/2} dt \right)^{1/p} < \infty \end{aligned}$$

in light of (4.1), so that $y \in L^p_{\mathbb{F}}(\Omega; L^p(0, T; \mathbb{X}))$.

We now want to rewrite x as a deterministic integral of a process involving y , to this end we apply the stochastic Fubini theorem ([10, Theorem 4.33]) to the function defined for fixed $t \in [0, T]$ by

$$(r, s) \mapsto (t-r)^{\alpha-1}(r-s)^{\alpha} S(t-s) \mathbb{I}_{0 < s < r < t} g(s) \in L^1([0, t], L^2(0, t; \mathcal{L}_2(\mathbb{H}, \mathbb{X}))).$$

We have for $p > \frac{1}{\alpha}$, $p^* := \frac{p}{p-1} < \frac{1}{1-\alpha}$, so, letting $M_T = \sup_{t \in [0, T]} \|S(t)\|$,

$$\begin{aligned} & \int_0^t (t-r)^{\alpha-1} \left(\mathbb{E} \left[\int_0^r (r-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{1/2} dr \\ & \leq \left(\int_0^t (t-r)^{(\alpha-1)p^*} dr \right)^{1/p^*} \left(\int_0^t \left(\mathbb{E} \left[\int_0^r (r-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{p/2} dt \right)^{1/p} \\ & \leq \left(\int_0^t (t-r)^{(\alpha-1)p^*} dr \right)^{1/p^*} \left(\int_0^t \left(\mathbb{E} \left[\int_0^r (r-s)^{-2\alpha} \|S(t-r)\| \|S(r-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{p/2} dt \right)^{1/p} \\ & \leq M_T \left(\int_0^t (t-r)^{(\alpha-1)p^*} dr \right)^{1/p^*} \left(\int_0^T \left(\mathbb{E} \left[\int_0^r (r-s)^{-2\alpha} \|S(r-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] \right)^{p/2} dt \right)^{1/p} < \infty \end{aligned}$$

for all $t \in [0, T]$, so we may indeed interchange the stochastic and deterministic integrals, notice that $\int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} dr = c$ doesn't depend on s, t , with this

$$\begin{aligned} x(t) &= \int_0^t S(t-s)g(s)dW(s) = c^{-1} \int_0^t \int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} dr S(t-s)g(s)dW(s) \\ &= c^{-1} \int_0^t \int_0^r (t-r)^{\alpha-1}(r-s)^{-\alpha} S(t-s)g(s)dW(s) dr \end{aligned}$$

$$\begin{aligned}
&= c^{-1} \int_0^t (t-r)^{\alpha-1} \int_0^r (r-s)^{-\alpha} S(t-r)S(r-s)g(s)dW(s)dr \\
&= c^{-1} \int_0^t (t-r)^{\alpha-1} S(t-r) \int_0^r (r-s)^{-\alpha} S(r-s)g(s)dW(s)dr = c^{-1} \int_0^t (t-r)^{\alpha-1} S(t-r)y(r)dr.
\end{aligned}$$

We now have by [10, Proposition 5.9] that $y \mapsto x(t) = c^{-1} \int_0^t (t-r)^{\alpha-1} S(t-r)y(r)dr$ maps $L^q(0, T; \mathbb{X})$ into $C([0, T]; \mathbb{X})$ for q large enough, namely, when $q > \frac{1}{\alpha}$. Applying this result for $q = p$, we find that $c^{-1} \int_0^t (t-r)^{\alpha-1} S(t-r)y(r)dr \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$ so that $x \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$, redefining x on a subset of $\Omega \times [0, T]$ of null measure if necessary.

Finally,

$$\begin{aligned}
\|x\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} &\leq \left\| \sup_{t \in [0, T]} \left[\sup_{s \in [0, t]} \|S(s)\| \|s^{\alpha-1}\|_{L^{p^*}(0, t)} \|y\|_{L^p(0, t; \mathbb{X})} \right] \right\|_{L^p(\Omega)} \\
&\leq \sup_{t \in [0, T]} \|S(t)\| \|t^{\alpha-1}\|_{L^{p^*}(0, T)} \left\| \|y\|_{L^p(0, T; \mathbb{X})} \right\|_{L^p(\Omega)} = \sup_{t \in [0, T]} \|S(t)\| \|t^{\alpha-1}\|_{L^{p^*}(0, T)} \|y\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{X}))} \\
&\leq M_T \|t^{\alpha-1}\|_{L^{p^*}(0, T)} \|y\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{X}))} = M_T \|t^{\alpha-1}\|_{L^{p^*}(0, T)} \left(\mathbb{E} \left[\int_0^T \|y(t)\|_{\mathbb{X}}^p dt \right] \right)^{1/p} \\
&\leq M_T \|t^{\alpha-1}\|_{L^{p^*}(0, T)} C_p \left(\int_0^T \mathbb{E} \left[\int_0^t (t-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X})}^2 ds \right] dt \right)^{1/p}.
\end{aligned}$$

When $T \rightarrow 0$, we have $\|t^{\alpha-1}\|_{L^{p^*}(0, T)} \rightarrow 0$ so that $C \rightarrow 0$. \square

In lemma 4.1, there is no additional requirement made directly on A beyond generating a C^0 -semigroup (hence the explicit introduction of \mathbb{X}_{-1} , as the construction of \mathbb{X}_s made in the previous section for analytic semigroups doesn't apply). In the sequel, we will establish the integrability assumption (4.1) by imposing that the semigroup generated by A is analytic.

Proposition 4.2. *Let A be the infinitesimal generator of the analytic semigroup S on \mathbb{X} , for some $\beta \in \rho(A)$ fixed, denote $\mathbb{X}_s = D((\beta Id - A)^s)$ the respective interpolation spaces for $s \in \mathbb{R}$ (see section 2). If there exist $\eta \geq 0$ and $p > 2$ such that $g \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)))$, then $x(t) = \int_0^t S(t-s)g(s)dW(s)$ has continuous paths in \mathbb{X}_γ a.s. for $\gamma < \eta + \frac{1}{2} - \frac{1}{p}$. Moreover, we have $x \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}_\gamma))$ and*

$$\|x\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}_\gamma))} \leq C \|g\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))}$$

with C only depending on A, η, γ, p, T (and not on g).

Proof. Fix $\gamma < \eta + \frac{1}{2} - \frac{1}{p}$ and choose $\alpha \in (\frac{1}{p}, \eta + \frac{1}{2} - \gamma)$ (in particular, $p > \frac{1}{\alpha}$), so that

$$\|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\gamma)} \leq \|S(t-s)\|_{\mathcal{L}(\mathbb{X}_\eta, \mathbb{X}_\gamma)} \|g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)} \leq C(t-s)^{\eta-\gamma} \|g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)},$$

due to proposition 2.1, from which it follows that

$$\begin{aligned}
&\int_0^T \mathbb{E} \left[\int_0^t (t-s)^{-2\alpha} \|S(t-s)g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\gamma)}^2 ds \right]^{p/2} dt \\
&\leq C \int_0^T \mathbb{E} \left[\int_0^t (t-s)^{-2(\alpha-\eta+\gamma)} \|g(s)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)}^2 ds \right]^{p/2} dt
\end{aligned}$$

$$\leq C \left(\frac{T^{1-2(\alpha-\eta+\gamma)}}{1-2(\alpha-\eta+\gamma)} \right)^{p/2} \|g\|_{L_{\mathbb{F}}^p(\Omega; L^p(0,T; \mathcal{L}(\mathbb{H}, \mathbb{X}_\eta))}^p < \infty$$

by Young's convolution inequality with exponents $1, \frac{p}{2}$. The result follows by lemma 4.1. \square

Proposition 4.3. *Let A be the infinitesimal generator of the analytic semigroup S on \mathbb{X} , for some $\beta \in \rho(A)$ fixed, denote $\mathbb{X}_s = D((\beta Id - A)^s)$ the respective interpolation spaces for $s \in \mathbb{R}$ (see section 2). Assume $f \in L_{\mathbb{F}}^2(0, T; X_{\gamma-\frac{1}{2}})$ and there exists $\eta \geq 0$ such that $g \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)))$ and we assume also $x_0 \in \mathbb{X}_\gamma$ a.s. is \mathcal{F}_0 -measurable with $\gamma < \eta + \frac{1}{2} - \frac{1}{p}$, then the mild solution (3.5) of*

$$\begin{aligned} dx(t) &= [Ax(t) + f(t)]dt + g(t)dW(t) \\ x(0) &= x_0 \end{aligned}$$

is well defined and has continuous paths in \mathbb{X}_γ a.s. Moreover, there exists $C = C(A, \gamma, \eta, p, T)$, such that if $\mathbb{E}[\|f\|_{L^2(0,T; X_{\gamma-\frac{1}{2}})}^p + \|x_0\|_{\mathbb{X}_\gamma}^p] < \infty$, then

$$\|x\|_{L_{\mathbb{F}}^p(\Omega; C([0,T], \mathbb{X}_\gamma))} \leq C \left(\|f\|_{L_{\mathbb{F}}^p(\Omega; L^2(0,T; X_{\gamma-\frac{1}{2}}))} + \|g\|_{L_{\mathbb{F}}^p(\Omega; L^p(0,T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))} + \|x_0\|_{L_{\mathcal{F}_0}^p(\Omega; \mathbb{X}_\gamma)} \right).$$

Proof. We decompose $x = x^{det} + x^{stoch}$ with

$$x^{det}(t) = S(t)x_0 + \int_0^t S(t-s)f(s)ds, \quad x^{stoch}(t) = \int_0^t S(t-s)g(s)dW(s).$$

The result follows from [3, Chapter II.1 Theorem 3.1 & Remark 4.2] for x^{det} and from proposition 4.2 for x^{stoch} . \square

We now will consider the following equation:

$$\begin{aligned} dx(t) &= [Ax(t) + b(t, x(t))]dt + \sigma(t, x(t))dW(t) \\ x(0) &= x_0 \end{aligned} \tag{4.2}$$

Proposition 4.4. *Let $\gamma < \eta + \frac{1}{2} - \frac{1}{p}$. Suppose A is the infinitesimal generator of the analytic semigroup S on \mathbb{X} , denote $\mathbb{X}_s = D((\beta Id - A)^s)$ the respective interpolation spaces for $s \in \mathbb{R}$ (see section 2). Assume additionally that $b : [0, T] \times \mathbb{X}_\gamma \times \Omega \rightarrow \mathbb{X}_{\gamma-\frac{1}{2}}$ and $\sigma : [0, T] \times \mathbb{X}_\gamma \times \Omega \rightarrow \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)$ satisfy*

$$\|b(t, x_1) - b(t, x_2)\|_{\mathbb{X}_{\gamma-\frac{1}{2}}} \leq L_b(t) \|x_1 - x_2\|_{X_\gamma}, \quad \|\sigma(t, x_1) - \sigma(t, x_2)\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta)} \leq L_\sigma(t) \|x_1 - x_2\|_{X_\gamma} \tag{4.3}$$

for some $L_b \in L^2(0, T)$ and $L_\sigma \in L^p(0, T)$. Assume also that

$$x_0 \in L_{\mathcal{F}_0}^p(\Omega; \mathbb{X}_\gamma), \quad b(\cdot, 0) \in L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{\gamma-\frac{1}{2}})), \quad \sigma(\cdot, 0) \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))).$$

Then problem (4.2) has a unique mild solution $x \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}_\gamma))$ and there exists $C = C(A, \gamma, \eta, p, T, L_b, L_\sigma)$ such that

$$\|x\|_{L_{\mathbb{F}}^p(\Omega; C([0,T], \mathbb{X}_\gamma))} \leq C \left(\|x_0\|_{L_{\mathcal{F}_0}^p(\Omega; \mathbb{X}_\gamma)} + \|b(\cdot, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^2(0,T; \mathbb{X}_{\gamma-\frac{1}{2}}))} + \|\sigma(\cdot, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^p(0,T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))} \right) \tag{4.4}$$

Proof. Let $t_1 < t_2 \leq T$. We define for $z \in L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma))$ and $x_{t_1} \in L_{\mathcal{F}_{t_1}}^p(\Omega; \mathbb{X}_\gamma)$ the application $F : z \mapsto F(z)$ given by

$$F(z)(t) = S(t - t_1)x_{t_1} + \int_{t_1}^t S(t - s)b(s, y(s))ds + \int_{t_1}^t S(t - s)\sigma(s, y(s))dW(s).$$

Since

$$\|b(\cdot, y(\cdot))\|_{L_{\mathbb{F}}^p(\Omega; L^2(t_1, t_2; \mathbb{X}_{\gamma - \frac{1}{2}}))} \leq \|b(\cdot, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^2(t_1, t_2; \mathbb{X}_{\gamma - \frac{1}{2}}))} + \|L_b\|_{L^2(t_1, t_2)} \|y\|_{L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma))} < \infty$$

and

$$\|\sigma(\cdot, y(\cdot))\|_{L_{\mathbb{F}}^p(\Omega; L^p(t_1, t_2; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))} \leq \|\sigma(\cdot, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^p(t_1, t_2; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))} + \|L_\sigma\|_{L^p(0, T)} \|y\|_{L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma))} < \infty$$

it follows by proposition 4.3 that $F : L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma)) \rightarrow L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma))$ is well defined. Clearly, x is a mild solution of equation (3.1) (on $[T_l, T_u]$) if and only if it is a fixed point of F .

Notice that

$$\|F(z_1) - F(z_2)\|_{L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma))} \leq C(\|L_b\|_{L^2(t_1, t_2)} + \|L_\sigma\|_{L^p(t_1, t_2)}) \|z_1 - z_2\|_{L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}_\gamma))}$$

it follows that for t_1, t_2 sufficiently small, F is Lipchitz with constant smaller than one, so it has a unique fixed point. Existence on $[0, T]$ follows by subdividing the interval and finding fixed points successively, uniqueness follows from the uniqueness on each subinterval. The estimate follows from the previous inequalities. \square

Remark 4.5. For $x_1, x_2 \in L_{\mathbb{F}}^p(\Omega; C([T_l, T_u], \mathbb{X}_\gamma)) \subset C_{\mathbb{F}}([T_l, T_u], L^p(\Omega; \mathbb{X}_\gamma))$ we have

$$b(\cdot, x_1(\cdot)) - b(\cdot, x_2(\cdot)) \in L_{\mathbb{F}}^2(T_l, T_u; L^p(\Omega; \mathbb{X}_{\gamma - \frac{1}{2}})).$$

In particular, if we assume that $b(\cdot, 0) \in L_{\mathbb{F}}^2(T_l, T_u; L^p(\Omega; \mathbb{X}_{\gamma - \frac{1}{2}}))$, we find $b(\cdot, x(\cdot)) \in L_{\mathbb{F}}^2(T_l, T_u; L^p(\Omega; \mathbb{X}_{\gamma - \frac{1}{2}}))$.

4.2 Controlled state equation

Proposition 4.6. Under hypothesis 3.1, for any $x_0 \in L_{\mathbb{F}}^p(\Omega; \mathbb{X})$ and $u \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$, equation (3.1) admits a unique mild solution $x^u \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$. Moreover, there exists $C = C(A, \eta, p, T, C_b, C_\sigma)$ such that

$$\begin{aligned} & \|x^u\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} \\ & \leq C(\|b(\cdot, 0, 0)\|_{L_{\mathbb{F}}^2(0, T; L^p(\Omega; \mathbb{X}_{-1/2}))} + \|\sigma(\cdot, 0, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta}))} + \|x_0\|_{L_{\mathbb{F}}^p(\Omega; \mathbb{X})} + \|u\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))}). \end{aligned} \quad (4.5)$$

Proof. Follows from proposition 4.4 (for $\gamma = 0$) and the fact that

$$\begin{aligned} \|b(\cdot, 0, u(\cdot))\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-1/2}))} & \leq \|b(\cdot, 0, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-1/2}))} + C_b \|u\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{U}))} \\ & \leq \|b(\cdot, 0, 0)\|_{L_{\mathbb{F}}^2(0, T; L^p(\Omega; \mathbb{X}_{-1/2}))} + C_b \|u\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{U}))} \end{aligned}$$

and

$$\|\sigma(\cdot, 0, u(\cdot))\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta}))} \leq \|\sigma(\cdot, 0, 0)\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta}))} + C_\sigma \|u\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))}.$$

\square

Fix $x_0 \in L_{\mathbb{F}}^p(\Omega; \mathbb{X})$. Denote x^u the solution of equation (3.1) for a given $u \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$.

Proposition 4.7. *Assume hypothesis 3.1. Then the mapping*

$$u \mapsto x^u : L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U})) \rightarrow L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$$

is Lipschitz continuous.

Proof. Let u_1 and u_2 be in $L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$. We have

$$\begin{aligned} d(x^{u_1} - x^{u_2})(t) &= [A(x^{u_1} - x^{u_2})(t) + b(t, x^{u_1}(t), u_1(t)) - b(t, x^{u_2}(t), u_2(t))] dt \\ &\quad + [\sigma(t, x^{u_1}(t), u_1(t)) - \sigma(t, x^{u_2}(t), u_2(t))] dW(t), \\ x^{u_1}(0) - x^{u_2}(0) &= 0 \in L_{\mathcal{F}_0}^p(\Omega; \mathbb{X}). \end{aligned}$$

The estimates of proposition 4.4 provide $C \geq 0$ such that for any $t \in [0, T]$

$$\begin{aligned} \|x^{u_1} - x^{u_2}\|_{L_{\mathbb{F}}^p(\Omega; C([0, t], \mathbb{X}))}^p &\leq C \left(\|b(\cdot, x^{u_1}(\cdot), u_1(\cdot)) - b(\cdot, x^{u_2}(\cdot), u_2(\cdot))\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, t; \mathbb{X}_{-1/2}))}^p \right. \\ &\quad \left. + \|\sigma(\cdot, x^{u_1}(\cdot), u_1(\cdot)) - \sigma(\cdot, x^{u_2}(\cdot), u_2(\cdot))\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, t; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))}^p \right) \\ &\leq C(C_b^p + C_\sigma^p) \mathbb{E} \left[\left(\int_0^t [\|x^{u_1}(s) - x^{u_2}(s)\|_{\mathbb{X}} + \|u_1(s) - u_2(s)\|_{\mathbb{U}}]^2 ds \right)^{p/2} \right. \\ &\quad \left. + \int_0^t [\|x^{u_1}(s) - x^{u_2}(s)\|_{\mathbb{X}} + \|u_1(s) - u_2(s)\|_{\mathbb{U}}]^p ds \right] \\ &\leq \bar{C} \int_0^t \|x^{u_1} - x^{u_2}\|_{L_{\mathbb{F}}^p(\Omega; C([0, s], \mathbb{X}))}^p ds + \|u_1 - u_2\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, t; \mathbb{U}))}^p, \end{aligned}$$

and the result follows from the Grönwall inequality. \square

5 Adjoint equations

Throughout this section we assume that hypotheses 3.1 and 3.4 hold. We wish to quantify variations in the cost in terms of changes in the trajectories. For this we introduce the adjoint or costate equation at the pair (x^u, u) :

Define, for any $u \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$, x^u as the solution of

$$dx^u(t) = [Ax^u(t) + b(t, x^u(t), u(t))]dt + \sigma(t, x^u(t), u(t))dW(t), \quad x^u(0) = x_0.$$

Define the costate (y^u, q^u) as the solution of the following backward stochastic evolution equation (BSEE):

$$\begin{aligned} dy^u(t) &= -[A^*y^u(t) + b_x(t, x^u(t), u(t))^*y^u(t) + \sigma_x(t, x^u(t), u(t))^*q^u(t) - f_x(t, x^u(t), u(t))]dt + q^u(t)dW(t), \\ y^u(T) &= -h_x(x^u(T)). \end{aligned} \tag{5.1}$$

Equation (5.1) may be written more compactly as

$$\begin{aligned} dy^u(t) &= -[A^*y^u(t) + H_x(t, x^u(t), u(t), y^u(t), q^u(t))]dt + q^u(t)dW(t), \\ y^u(T) &= -h_x(x^u(T)). \end{aligned} \tag{5.2}$$

where

$$H(t, x, u, y, q) = \langle y, b(t, x, u) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} + \langle q, \sigma(t, x, u) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} - f(t, x, u).$$

The aim of this section is to give a framework for equation (5.1) and establish existence, uniqueness and regularity results.

Hypothesis 5.1. Assume hypothesis 3.1 holds and b, σ are differentiable w.r.t. x a.e. on $[0, T] \times \Omega$.

Hypothesis 5.2. Assume hypothesis 3.4 holds and f is differentiable w.r.t. x a.e. on $[0, T] \times \Omega$ and h is differentiable w.r.t. x a.s.

Remark 5.3. Assuming hypotheses 5.1 and 5.2, the Hamiltonian

$$H(t, x, u, y, q) = \langle y, b(t, x, u) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} + \langle q, \sigma(t, x, u) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} - f(t, x, u),$$

is differentiable w.r.t. x and its derivative H_x is given by $b_x(t, x, u)^* y + \sigma_x(t, x, u)^* q - f_x(t, x, u)$. With this, equation (5.2) is not only a formalism, but actually the same as equation (5.1).

5.1 Regularity for solutions of parabolic BSEEs

Notice that in the BSEE (5.1) equation both y, q are unknown. We now consider the general BSEE:

$$dy(t) = -[A^* y(t) + g(t, y(t), q(t))] dt + q dW(t), \quad y(T) = y_T \in L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*). \quad (5.3)$$

Definition 5.4. Let A be the generator of an analytic semigroup and

$$g : \Omega \times [0, T] \times \mathbb{X}_{1/2}^* \times \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*) \rightarrow \mathbb{X}^*$$

be a function. Let $p > 2$ and write $p^* = \frac{p}{p-1}$. We say the pair

$$(y, q) \in \left[D_{\mathbb{F}}([0, T], L^{p^*}(\Omega; \mathbb{X})) \cap L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \right] \times L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))$$

is a **transposition solution** of equation (5.3) if $g(\cdot, y(\cdot), q(\cdot)) \in L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))$ and for every $t \in [0, T]$, $x_t \in L_{\mathcal{F}_t}^p(\Omega; \mathbb{X})$, $B \in L_{\mathbb{F}}^2(t, T; L^p(\Omega; \mathbb{X}_{-1/2}))$ and $\Sigma \in L_{\mathbb{F}}^p(t, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})))$ we have

$$\begin{aligned} & \mathbb{E} \langle y_T, x(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_t^T \langle g(s, y(s), q(s)), x(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ &= \mathbb{E} \langle y(t), x_t \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_t^T \langle y(s), B(s) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} ds + \mathbb{E} \int_t^T \langle q(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} ds, \end{aligned} \quad (5.4)$$

where

$$dx(t) = [Ax(t) + B(t)] dt + \Sigma(t) dW(t), \quad x(t) = x_t.$$

The preceding definition differs from [25, Definition 4.13] in the fact that we assume the weaker condition

$$g : \Omega \times [0, T] \times \mathbb{X}_{1/2}^* \times \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*) \rightarrow \mathbb{X}^*$$

instead of

$$g : \Omega \times [0, T] \times \mathbb{X}^* \times \mathcal{L}_2(\mathbb{H}, \mathbb{X}^*) \rightarrow \mathbb{X}^*,$$

it is for this reason that we impose

$$(y, q) \in L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \times L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*))).$$

We begin by considering the case where g doesn't depend on y, q . Consider

$$dy(t) = -[A^* y(t) + G(t)] dt + q dW(t), \quad y(T) = y_T \in L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*). \quad (5.5)$$

Proposition 5.5. *If $G \in L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))$ and $y_T \in L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X})$, then equation (5.5) has a unique transposition solution (y, q) . Moreover*

$$\begin{aligned} & \|y\|_{D_{\mathbb{F}}([0, T], L^{p^*}(\Omega; \mathbb{X}))} + \|y\|_{L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|q\|_{L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*))} \\ & \leq C \left(\|y_T\|_{L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X})} + \|G\|_{L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))} \right). \end{aligned}$$

Before we prove proposition 5.5, we state the following lemma.

Lemma 5.6 ([25, Theorem 4.16]). *If $G \in L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))$ and $y_T \in L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X})$, then there exists a unique pair of processes*

$$(y, q) \in D_{\mathbb{F}}([0, T], L^p(\Omega; \mathbb{X})) \times L_{\mathbb{F}}^2(0, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*)))$$

such that for every $t \in [0, T]$, $x_t \in L_{\mathcal{F}_t}^p(\Omega; \mathbb{X})$, $B \in L_{\mathbb{F}}^1(t, T; L^p(\Omega; \mathbb{X}))$ and $\Sigma \in L_{\mathbb{F}}^2(t, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X})))$ we have

$$\begin{aligned} & \mathbb{E}\langle y_T, x(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_t^T \langle G(s), x(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ & = \mathbb{E}\langle y(t), x_t \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_t^T \langle y(s), B(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds + \mathbb{E} \int_t^T \langle q(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X})} ds. \end{aligned} \tag{5.6}$$

where

$$dx(t) = [Ax(t) + B(t)] dt + \Sigma(t)dW(t), \quad x(t) = x_t.$$

Moreover

$$\begin{aligned} & \|y\|_{D_{\mathbb{F}}([0, T], L^{p^*}(\Omega; \mathbb{X}))} + \|q\|_{L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}^*))} \\ & \leq C \left(\|y_T\|_{L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X})} + \|G\|_{L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))} \right). \end{aligned}$$

Proof of proposition 5.5. First, by proposition 4.3, $(B, \Sigma) \mapsto \Gamma(B, \Sigma) = (x(T), x)$, where

$$dx(t) = [Ax(t) + B(t)] dt + \Sigma(t)dW(t), \quad x(0) = 0.$$

defines a bounded linear operator from $L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-1/2})) \times L_{\mathbb{F}}^p(0, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})))$ to $L_{\mathcal{F}_T}^p(\Omega; \mathbb{X}) \times L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$. Notice that

$$G \in L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*)) \subset [C_{\mathbb{F}}(0, T; L^p(\Omega; \mathbb{X}))]^* \subset [L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))]^*.$$

Define $(\tilde{y}, \tilde{q}) = \Gamma^*(y_T, g)$ so that

$$\begin{aligned} & \mathbb{E}\langle y_T, x(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_0^T \langle G(s), x(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ & = \langle \tilde{y}, B \rangle_{[L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-1/2}))]^*, L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-1/2}))} + \mathbb{E} \int_0^T \langle \tilde{q}(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} ds. \end{aligned}$$

Due to [23, corollary 2.3, remark 2.4] we have

$$[L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-1/2}))]^* \subset [L_{\mathbb{F}}^2(0, T; L^p(\Omega; \mathbb{X}_{-1/2}))]^* = L_{\mathbb{F}}^2(0, T; L^p(\Omega; \mathbb{X}_{1/2}^*)),$$

so that

$$\|\tilde{y}\|_{L_{\mathbb{F}}^2(0,T;L^{p^*}(\Omega;\mathbb{X}_{1/2}^*))} + \|\tilde{q}\|_{L_{\mathbb{F}}^{p^*}(0,T;L^{p^*}(\Omega;\mathcal{L}_2(\mathbb{H},\mathbb{X}_{\eta}^*)))} \leq \|\Gamma\| \left(\|y_T\|_{L_{\mathcal{F}_T}^{p^*}(\Omega;\mathbb{X})} + \|G\|_{L_{\mathbb{F}}^1(0,T;L^{p^*}(\Omega;\mathbb{X}^*))} \right).$$

When $B \in L_{\mathbb{F}}^2(0,T;L^p(\Omega;\mathbb{X}_{-1/2}))$, the term $\langle \tilde{y}, B \rangle_{[L_{\mathbb{F}}^p(\Omega;L^2(0,T;\mathbb{X}_{-1/2}))]^*, L_{\mathbb{F}}^p(\Omega;L^2(0,T;\mathbb{X}_{-1/2}))}$ can be replaced by $\mathbb{E} \int_0^T \langle \tilde{y}(s), B(s) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} ds$, meaning

$$\begin{aligned} & \mathbb{E} \langle y_T, x(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_0^T \langle G(s), x(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ &= \mathbb{E} \int_0^T \langle \tilde{y}(s), B(s) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} ds + \mathbb{E} \int_0^T \langle \tilde{q}(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} ds, \end{aligned} \quad (5.7)$$

Now, let

$$(y, q) \in D_{\mathbb{F}}([0, T], L^{p^*}(\Omega; \mathbb{X})) \times L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*)))$$

be the processes given by 5.6.

Let (B, Σ) belong to

$$L_{\mathbb{F}}^2(0, T; L^p(\Omega; \mathbb{X})) \times L_{\mathbb{F}}^p(0, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}))),$$

then

$$\begin{aligned} & \mathbb{E} \int_0^T \langle y(s), B(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds + \mathbb{E} \int_0^T \langle q(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X})} ds \\ &= \mathbb{E} \langle y_T, x(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_0^T \langle G(s), x(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ &= \mathbb{E} \int_0^T \langle \tilde{y}(s), B(s) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} ds + \mathbb{E} \int_0^T \langle \tilde{q}(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} ds \\ &= \mathbb{E} \int_0^T \langle \tilde{y}(s), B(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds + \mathbb{E} \int_0^T \langle \tilde{q}(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X})} ds. \end{aligned} \quad (5.8)$$

It follows that $(y, q) = (\tilde{y}, \tilde{q})$. Let $t \in [0, T)$, $x_t \in L_{\mathcal{F}_t}^p(\Omega; \mathbb{X})$ and (B, Σ) be supported on $[t, T]$, then (5.8) together with (5.6) imply (5.4).

Finally, the result follows because

$$L_{\mathbb{F}}^2(t, T; L^p(\Omega; \mathbb{X})) \times L_{\mathbb{F}}^p(t, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X})))$$

is dense in

$$L_{\mathbb{F}}^2(t, T; L^p(\Omega; \mathbb{X}_{-1/2})) \times L_{\mathbb{F}}^p(t, T; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta}))).$$

□

Our next step is to use a fixed point argument to deduce well-posedness for equation (5.3) from proposition 5.5.

Theorem 5.7. *Let $y_T \in L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)$. Let*

$$g : \Omega \times [0, T] \times \mathbb{X}_{1/2}^* \times \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*) \rightarrow \mathbb{X}^*$$

be a function such that $g(\cdot, y, q)$ is predictable for every fixed $(y, q) \in \mathbb{X}_{1/2}^* \times \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)$ and

$$g(\cdot, 0, 0) \in L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*)).$$

Suppose additionally that

$$\|g(t, y_1, q_1) - g(t, y_2, q_2)\|_{\mathbb{X}^*} \leq L \left(\|y_1 - y_2\|_{\mathbb{X}_{1/2}^*} + \|q_1 - q_2\|_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)} \right).$$

Then equation (5.3) has a unique transposition solution (y, q) . Moreover, there exists C depending only on A, p, η, T, L such that

$$\begin{aligned} & \|y\|_{D_{\mathbb{F}}([0, T], L^{p^*}(\Omega; \mathbb{X}))} + \|y\|_{L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|q\|_{L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \\ & \leq C \left(\|y_T\|_{L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)} + \|g(\cdot, 0, 0)\|_{L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))} \right). \end{aligned}$$

Proof. It follows from proposition 4.3 that for $0 \leq t_1 < t_2 \leq T$ there exists a bounded operator from

$$\Gamma_{t_1}^{t_2} : L_{\mathcal{F}_{t_1}}^p(\Omega; \mathbb{X}) \times L_{\mathbb{F}}^2(t_1, t_2; L^p(\Omega; \mathbb{X}_{-1/2})) \times L_{\mathbb{F}}^p(t_1, t_2; L^p(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta}))) \rightarrow L_{\mathcal{F}_{t_2}}^p(\Omega; \mathbb{X}) \times L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}))$$

given by $\Gamma_{t_1}^{t_2}(x_{t_1}, B, \Sigma) = (x(t_2), x)$ where x solves

$$dx = Bdt + \Sigma dW, \quad x(t_1) = x_{t_1}.$$

Moreover, there exists some uniform constant Λ such that $\|\Gamma_{t_1}^{t_2}\| \leq \Lambda$ (we may take $\Lambda = \|\Gamma_0^T\|$).

Notice that

$$L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}^*)) \subset [C_{\mathbb{F}}(t_1, t_2; L^p(\Omega; \mathbb{X}))]^* \subset [L_{\mathbb{F}}^p(\Omega; C([t_1, t_2], \mathbb{X}))]^*$$

and due to [23, corollary 2.3 & remark 2.4] we have

$$[L_{\mathbb{F}}^p(\Omega; L^2(t_1, t_2; \mathbb{X}_{-1/2}))]^* \subset [L_{\mathbb{F}}^2(t_1, t_2; L^p(\Omega; \mathbb{X}_{-1/2}))]^* = L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)).$$

For $(Y, Q) \in L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \times L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))$ we have

$$\begin{aligned} & \|g(\cdot, Y(\cdot), Q(\cdot))\|_{L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}^*))} \\ & \leq \|g(\cdot, 0, 0)\|_{L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}^*))} + L \left(\|Y\|_{L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|Q\|_{L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \right) \\ & \leq \|g(\cdot, 0, 0)\|_{L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}^*))} + L \left((t_2 - t_1)^{1/2} \|Y\|_{L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + (t_2 - t_1)^{1/p^*} \|Q\|_{L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \right) \\ & < \infty \end{aligned}$$

which makes the mapping

$$\begin{aligned} F & : L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \times L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*))) \\ & \rightarrow L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \times L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*))) \end{aligned}$$

given by

$$F(Y, Q) = \pi_{2,3}(\Gamma_{t_1}^{t_2})^*(y_{t_2}, g(\cdot, Y(\cdot), Q(\cdot))),$$

where $\pi_{2,3}$ projects $L_{\mathcal{F}_{t_1}}^{p^*}(\Omega; \mathbb{X}^*) \times L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \times L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))$ onto the subspace $L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*)) \times L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))$, well defined for $y_{t_2} \in L_{\mathcal{F}_{t_2}}^{p^*}(\Omega; \mathbb{X}^*)$. Moreover, we have

$$\begin{aligned} & \|\pi_2 F(Y_1, Q_1) - \pi_2 F(Y_2, Q_2)\|_{L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|\pi_3 F(Y_1, Q_1) - \pi_3 F(Y_2, Q_2)\|_{L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \\ & \leq \Lambda \|g(\cdot, Y_1(\cdot), Q_1(\cdot)) - g(\cdot, Y_2(\cdot), Q_2(\cdot))\|_{L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}^*))} \\ & \leq \Lambda L \left((t_2 - t_1)^{1/2} \|Y_1 - Y_2\|_{L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + (t_2 - t_1)^{1/p^*} \|Q_1 - Q_2\|_{L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \right) \\ & \leq \Lambda L \left((t_2 - t_1)^{1/2} + (t_2 - t_1)^{1/p^*} \right) \left(\|Y_1 - Y_2\|_{L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|Q_1 - Q_2\|_{L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \right), \end{aligned}$$

so that F is a contraction for small enough $t_2 - t_1$. In such case, the Banach fixed point theorem yields the existence of a fixed point (y, q) , which moreover satisfies

$$\begin{aligned} & \|y\|_{L_{\mathbb{F}}^2(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|q\|_{L_{\mathbb{F}}^{p^*}(t_1, t_2; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*)))} \\ & \leq \frac{\Lambda \left(\|y_{t_2}\|_{L_{\mathcal{F}_{t_2}}^{p^*}(\Omega; \mathbb{X})} + \|g(\cdot, 0, 0)\|_{L_{\mathbb{F}}^1(t_1, t_2; L^{p^*}(\Omega; \mathbb{X}^*))} \right)}{1 - \Lambda L \left((t_2 - t_1)^{1/2} + (t_2 - t_1)^{1/p^*} \right)}. \end{aligned}$$

We partition $0 = T_0 < T_1 < \dots < T_n = T$ such that F is a contraction on each subinterval $[T_{i-1}, T_i]$. Define (y, q) as the fixed point of F on $[T_{n-1}, T_n]$ with $y_{T_n} = y_T$. Define inductively $y_{T_{i-1}} = \pi_1(\Gamma_{T_{i-1}}^{T_i})^*(y_{T_i}, g(\cdot, y(\cdot), q(\cdot)))$ and (y, q) on $[T_{i-2}, T_{i-1}]$ as the fixed point of F on that subinterval. Fix i , we clearly have

$$\begin{aligned} & \mathbb{E}\langle y_{T_i}, x(T_i) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_{T_{i-1}}^{T_i} \langle g(s, y(s), q(s)), x(s) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ & = \mathbb{E}\langle y_{T_{i-1}}, x_{T_{i-1}} \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_{T_{i-1}}^{T_i} \langle y(s), B(s) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} ds + \mathbb{E} \int_{T_{i-1}}^{T_i} \langle q(s), \Sigma(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} ds, \end{aligned}$$

where x solves $dx = Bdt + \Sigma dW$ and $x(T_{i-1}) = x_{T_{i-1}}$. To conclude, letting x solve

$$dx = Bdt + \Sigma dW$$

with $x_{T_0} = x_0 = 0$, setting $x_{T_{i-1}} = x(T_{i-1})$ and summing over i , we find that (y, q) is a transposition solution. The uniqueness follows from uniqueness on each subinterval.

The bound for the norm on $[0, T]$ follows from the bound on each subinterval. \square

5.2 Adjoint equations for the optimal control problem

Notice that the Lipschitz conditions for b and σ given in hypothesis 3.1 guarantee that the derivatives given by hypothesis 5.1 are uniformly bounded. On the other hand, the local Lipschitz condition for f given in hypothesis 3.4, means that the derivative $f_x(t, x^u(t), u(t))$ given by hypothesis 5.2 satisfies

$$\|f_x(t, x^u(t), u(t))\|_{\mathbb{X}^*} \leq C_1 (1 + 2 \|u(t)\|_{\mathbb{U}} + 2 \|x^u(t)\|_{\mathbb{X}})^{p-1}$$

and therefore belongs to $L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathbb{X}^*)) \subset L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))$, and similarly

$$\|h_x(x^u(T))\|_{\mathbb{X}^*} \leq C_1 (1 + 2 \|x^u(T)\|_{\mathbb{X}})^{p-1},$$

so that it belongs to $L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)$. These facts combined mean that the function defined by

$$g(t, y, q) = H_x(t, x^u(t), u(t), y, q) = b_x(t, x^u(t), u(t))^* y + \sigma_x(t, x^u(t), u(t))^* q - f_x(t, x^u(t), u(t))$$

satisfies the hypotheses of theorem 5.7. As a result, we have:

Proposition 5.8. *Let 5.1 and 5.2 hold. Let $x_0 \in L_{\mathbb{F}}^p(\Omega; \mathbb{X})$, $u \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$ be given, let $x^u \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$ be the respective solution of equation (3.1).*

Then equation (5.1) admits a unique transposition solution (y^u, q^u) . Moreover

$$\begin{aligned} & \|y^u\|_{L_{\mathbb{F}}^2(0, T; L^{p^*}(\Omega; \mathbb{X}_{1/2}^*))} + \|q^u\|_{L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{\eta}^*))} \\ & \leq C \left(\|h_x(x(T))\|_{L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)} + \|f_x(\cdot, x^u(\cdot), u(\cdot))\|_{L_{\mathbb{F}}^1(0, T; L^{p^*}(\Omega; \mathbb{X}^*))} \right) \\ & \leq \tilde{C} \left(1 + \|x_0\|_{L_{\mathcal{F}_0}^p(\Omega; \mathbb{X})} + \|u\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} \right)^{p-1}. \end{aligned}$$

6 Optimality conditions

In this section, our objective is to give necessary optimality conditions for solutions of problem \mathcal{P} .

Our first step is to parametrize the problem in terms of the control. As we have seen, for fixed $u \in \mathcal{U}^{ad}$, under hypothesis 3.1, we have existence and uniqueness of a trajectory $x \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))$ solving (3.1). This means that the set of pairs $(x, u) \in L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X})) \times \mathcal{U}^{ad}$ solving (3.1) is parametrized by $u \in \mathcal{U}^{ad}$, and this parametrization is given by $u \mapsto (x^u, u)$. With this, we may interpret problem \mathcal{P} as

$$\min_{u \in \mathcal{U}^{ad}} J(u) \tag{6.1}$$

where

$$J(u) = \mathcal{J}(x^u, u). \tag{6.2}$$

Under hypotheses 3.1 and 3.4, J is well defined and finite. Furthermore we have the following local Lipschitz property.

Proposition 6.1. *Assuming hypotheses 3.1 and 3.4, the cost function $J : L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U})) \rightarrow \mathbb{R}$ is Lipschitz continuous on bounded subsets of $L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$.*

Proof. It follows from proposition 4.7 and hypothesis 3.4, since

$$\begin{aligned} |J(u_1) - J(u_2)| & \leq \|f(\cdot, x^{u_1}(\cdot), u_1(\cdot)) - f(\cdot, x^{u_2}(\cdot), u_2(\cdot))\|_{L_{\mathbb{F}}^1(0, T; L^1(\Omega))} + \|h(x^{u_1}(T)) - h(x^{u_2}(T))\|_{L_{\mathcal{F}_T}^1(\Omega)} \\ & \leq C_0 \left(T^{1/p} + \|u_1\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} + \|u_2\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} + T^{1/p} \|x^{u_1}\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} + T^{1/p} \|x^{u_2}\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} \right)^{p-1} \\ & \quad \times \left(\|x^{u_1} - x^{u_2}\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} + \|u_1 - u_2\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} \right) \\ & \quad + C_0 \left(1 + \|x^{u_1}(T)\|_{L_{\mathcal{F}_T}^p(\Omega; \mathbb{X})} + \|x^{u_2}(T)\|_{L_{\mathcal{F}_T}^p(\Omega; \mathbb{X})} \right)^{p-1} \|x^{u_1}(T) - x^{u_2}(T)\|_{L_{\mathcal{F}_T}^p(\Omega; \mathbb{X})} \\ & \leq C \left(1 + \|u_1\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} + \|u_2\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} \right)^{p-1} \|u_1 - u_2\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))}. \end{aligned}$$

□

Our next step is to quantify the variations in the cost in terms of variations in the Hamiltonian.

Hypothesis 6.2. Assume hypothesis 5.1 holds and b, σ are differentiable w.r.t. u . Assume also that the derivatives $b_x, b_u, \sigma_x, \sigma_u$ are strongly continuous w.r.t. x, u a.e. on $[0, T] \times \Omega$.

Hypothesis 6.3. Assume hypothesis 5.2 holds and f is differentiable w.r.t. u . Assume also that the derivatives f_x, f_u are strongly continuous w.r.t. x, u a.e. on $[0, T] \times \Omega$ and h_x is strongly continuous w.r.t. x a.s.

Notice that hypotheses 6.2 and 6.3 are similar to hypotheses 3.1 and 3.4. In both cases, the only additional assumption is differentiability w.r.t. u and strong continuity of all derivatives. Because for a Banach space X strong continuity into $\mathcal{L}(X, \mathbb{C})$ is equivalent to weak* continuity into X^* , we have that f_x, f_u, h_x are weak* continuous, and therefore, weak continuous, by the reflexivity of \mathbb{X} and \mathbb{U} .

Remark 6.4. Assuming hypotheses 6.2 and 6.3, the Hamiltonian

$$H(t, x, u, y, q) = \langle y, b(t, x, u) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} + \langle q, \sigma(t, x, u) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} - f(t, x, u),$$

is differentiable w.r.t. u and its derivative H_u is given by $b_u(t, x, u)^* y + \sigma_u(t, x, u)^* q - f_u(t, x, u)$.

Proposition 6.5. Fix $u, v \in L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$, and write $u^\varepsilon = u + \varepsilon v$. Assume hypotheses 6.2 and 6.3. Then the costate equation 5.1 for u admits a unique transposition solution (y^u, q^u) , we have $H_u(\cdot, x^u(\cdot), u(\cdot), y^u(\cdot), q^u(\cdot)) \in L_{\mathbb{F}}^{p^*}(\Omega; L^{p^*}(0, T; \mathbb{U}^*))$ and we have

$$\left| J(u^\varepsilon) - J(u) + \varepsilon \mathbb{E} \int_0^T \langle H_u(t, x^u(t), u(t), y^u(t), q^u(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \right| = o(\varepsilon).$$

The proof is in appendix B.

Remark 6.6.

Let X be a Banach space, let $K \subset X$ be a subset. We define $T_K^b(x)$, the adjacent cone to K at $x \in X$ (see, for instance, [1, section 4.1]) as

$$\{d \in X : \forall [\{\varepsilon_k\}_{k \in \mathbb{N}} \subset (0, \infty) : \varepsilon_k \rightarrow 0 \text{ as } k \rightarrow \infty] \exists [\{d_k\}_{k \in \mathbb{N}} \subset X : d_k \rightarrow d \text{ as } k \rightarrow \infty \text{ and } x + \varepsilon_k d_k \in K \quad \forall k]\}.$$

Proposition 6.7. Assume hypotheses 6.2 and 6.3 hold. Let \bar{u} be a solution of problem \mathcal{P} . Then the costate equation 5.1 for $u = \bar{u}$ admits a unique transposition solution $(y^{\bar{u}}, q^{\bar{u}})$ and we have

$$\mathbb{E} \int_0^T \langle H_u(t, x^{\bar{u}}(t), \bar{u}(t), y^{\bar{u}}(t), q^{\bar{u}}(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \leq 0 \quad \forall v \in T_{\mathcal{U}^{ad}}^b(\bar{u}). \quad (6.3)$$

Proof. It follows from propositions 6.1 and 6.5. Let $v \in T_{\mathcal{U}^{ad}}^b(\bar{u})$, $\{\varepsilon_k\}_{k \in \mathbb{N}} \subset (0, \infty)$, and $\{v_k\}_{k \in \mathbb{N}} \subset L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))$ satisfying $(\varepsilon_k, v_k) \rightarrow (0, v)$ such that $\bar{u} + \varepsilon_k v_k \in \mathcal{U}^{ad}$. Denote still $\{\varepsilon_k\}_{k \in \mathbb{N}}$ the subsequence given by proposition 6.5. Let L be the Lipschitz constant of J in a ball to which $\bar{u} + \varepsilon_k v_k, \bar{u} + \varepsilon_k v$ are confined. We have

$$\begin{aligned} 0 &\leq J(\bar{u} + \varepsilon_k v_k) - J(\bar{u}) = J(\bar{u} + \varepsilon_k v_k) - J(\bar{u} + \varepsilon_k v) + J(\bar{u} + \varepsilon_k v) - J(\bar{u}) \\ &\leq L \varepsilon_k \|v - v_k\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))} - \varepsilon_k \mathbb{E} \int_0^T \langle H_u(t, x^{\bar{u}}(t), \bar{u}(t), y^{\bar{u}}(t), q^{\bar{u}}(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt + o(\varepsilon_k) \\ &= -\varepsilon_k \mathbb{E} \int_0^T \langle H_u(t, x^{\bar{u}}(t), \bar{u}(t), y^{\bar{u}}(t), q^{\bar{u}}(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt + o(\varepsilon_k). \end{aligned}$$

□

Theorem 6.8. *Assume hypotheses 6.2 and 6.3 hold. Let \bar{u} be a solution of problem \mathcal{P} . Then the costate equation 5.1 for $u = \bar{u}$ admits a unique transposition solution $(y^{\bar{u}}, q^{\bar{u}})$ and we have*

$$\left[\langle H_u(t, x^{\bar{u}}(t), \bar{u}(t), y^{\bar{u}}(t), q^{\bar{u}}(t)), v \rangle_{\mathbb{U}^*, \mathbb{U}} \leq 0 \quad \forall v \in T_{U^{ad}}^b(\bar{u}(t)) \right] \text{ a.e. in } [0, T] \times \Omega. \quad (6.4)$$

We will need the following result, which is proven in [33, Lemma 4.6] for $p = 2$, and X finite dimensional; and in [13, Lemma 3.2.], also for $p = 2$, with the Clarke tangent cone in place of the adjacent cone and with the assumption that X is a separable Hilbert space.

Lemma 6.9 ([33, Lemma 4.6],[13, Lemma 3.2.]). *Assume X is separable, (S, Σ, μ) is a σ -finite measure space, $p \in [1, \infty)$ and $p^* = \frac{p}{p-1}$ ($1^* = \infty$). Denote*

$$\mathcal{K} = \{u \in L^p(S, \mu; X) : u(\cdot) \in K \quad \mu - \text{a.e.}\}.$$

Suppose also that $F \in L^{p^}(S, \mu; X^*)$ and for every $v \in T_{\mathcal{K}}^b(u)$,*

$$\int_S \langle F(s), v(s) \rangle_{X^*, X} d\mu(s) \leq 0.$$

Then $\langle F(s), v \rangle_{X^, X} \leq 0 \quad \forall v \in T_{\mathcal{K}}^b(u(s))$ for μ -a.e. $s \in S$.*

The proof is omitted as it is the same as that of [13, Lemma 3.2.], replacing 2 with p and the Clarke cone with the adjacent cone, this can be done because the latter is closed as well, and therefore $T_{\mathcal{K}}^b(u(s))$ is a measurable set-valued mapping with closed values, so a suitable measurable selection can be constructed (see [1, theorem 8.5.1] for the measurability of $T_{\mathcal{K}}^b(u(s))$). In the proof of [13, lemma 3.2.], X being Banach and separable are the sole relevant properties of the space X .

Proof of theorem 6.8. Follows from proposition 6.7 and lemma 6.9. □

Example 6.10. *Recall the setting of example 3.3.*

For

$$J(u) = \frac{1}{p} \mathbb{E} \int_0^T [\alpha_{st} \|u_{st}(t)\|_{L^p}^p + \alpha_0 |u_0(t)|^p + \alpha_1 |u_1(t)|^p] dt + \frac{\alpha_x}{2} \mathbb{E} \int_0^T \|x(t)\|_{L^2(0,1)}^2 dt + \frac{\alpha_T}{2} \mathbb{E} \|x(T)\|_{L^2(0,1)}^2$$

we have hypothesis 6.3. Moreover, if the functions b and σ are linear, existence of a solution (\tilde{u}, \tilde{x}) of the optimal control problem is guaranteed by the Weierstrass minimization theorem, as the equation is linear, the cost convex and coercive, and the control space is a reflexive Banach space.

Formally, the abstract costate equation 5.1 may be regarded as

$$\begin{aligned} \partial_t y(t, \xi) &= -\partial_{\xi\xi} y(t, \xi) - b_x^{st}(s, \xi, x, u) y(t, \xi) + \alpha_x x(s, \xi) \\ &\quad + q_{st}(t, \xi) \dot{W}_{st}(t, \xi) + q_0(t, \xi) \dot{w}_0(t) + q_1(t, \xi) \dot{w}_1(t) \quad t \in (0, T), \xi \in (0, 1) \\ \partial_n y(t, i) &= 0 \quad t \in (0, T), i = 0, 1 \\ y(T, \xi) &= -\alpha_3 x(T, \xi) \quad \xi \in (0, 1) \end{aligned}$$

where

$$(q_{st} \quad q_0 \quad q_1) = q \in \mathcal{L}_2(\mathbb{H}, \mathcal{H}_{2\eta}) \simeq \mathcal{L}_2(L^2, \mathcal{H}_{2\eta}) \oplus \mathcal{L}_2(\mathbb{R}, \mathcal{H}_{2\eta}) \oplus \mathcal{L}_2(\mathbb{R}, \mathcal{H}_{2\eta}) \simeq \mathcal{L}_2(L^2, \mathcal{H}_{2\eta}) \oplus \mathcal{H}_{2\eta} \oplus \mathcal{H}_{2\eta}.$$

Notice that because $\eta > \frac{1}{4} + \frac{1}{2p} > \frac{1}{4}$, the inclusions from $\mathcal{H}_{2\eta}$ into $\mathcal{H}_{1/p}$, and hence into L^2 , are Hilbert-Schmidt, therefore, $\mathcal{L}_2(L^2, \mathcal{H}_{2\eta}) \subset \mathcal{L}_1(L^2)$. Fix $q^{st} \in \mathcal{L}_2(L^2, \mathcal{H}_{2\eta})$, if we choose a Hilbert basis $\{\varphi_i\}_{i \in \mathbb{N}}$ of L^2 which consists of eigenvectors of $|q^{st}|$, then

$$\begin{aligned} \int_0^1 \sum_{i \in \mathbb{N}} \|[q^{st}\varphi_i](\xi)\varphi_i(\xi)|d\xi &= \sum_{i \in \mathbb{N}} \int_0^1 \|[q^{st}\varphi_i](\xi)\varphi_i(\xi)|d\xi \leq \sum_{i \in \mathbb{N}} \|q^{st}\varphi_i\|_{L^2} \|\varphi_i\|_{L^2} \\ &= \sum_{i \in \mathbb{N}} \|q^{st}\varphi_i\|_{L^2} = \sum_{i \in \mathbb{N}} \||q^{st}|\varphi_i\|_{L^2} = \sum_{i \in \mathbb{N}} \langle |q^{st}|\varphi_i, \varphi_i \rangle_{L^2} < \infty, \end{aligned}$$

it follows that the function $[Tr(q^{st})](\xi) := \sum_{i \in \mathbb{N}} [q^{st}\varphi_i](\xi)\varphi_i(\xi)$ is well defined for a.e. $\xi \in (0, 1)$ and belongs to L^1 , moreover, for $\Sigma^{st} \in L^\infty \subset L^p$, and $[\mathcal{B}\Sigma^{st}]h = \Sigma^{st}h$, we have

$$\begin{aligned} \int_0^1 \Sigma^{st}[\mathcal{B}^*q^{st}](\xi)d\xi &= \langle \mathcal{B}\Sigma^{st}, q^{st} \rangle_{\mathcal{L}_2(L^2, \mathcal{H}_{-2\eta}), \mathcal{L}_2(L^2, \mathcal{H}_{2\eta})} = \sum_{i \in \mathbb{N}} \langle \Sigma^{st}\varphi_i, q^{st}\varphi_i \rangle_{\mathcal{H}_{-2\eta}, \mathcal{H}_{2\eta}} \\ &= \sum_{i \in \mathbb{N}} \langle \Sigma^{st}\varphi_i, q^{st}\varphi_i \rangle_{L^{r^*}, L^r} = \int_0^1 \Sigma^{st}[Tr(q^{st})](\xi)d\xi. \end{aligned}$$

It follows that $Tr(q^{st}) = \mathcal{B}^*q^{st}$ and therefore belongs to L^{p^*} and depends linearly and continuously on q^{st} (the linearity is not obvious from the definition because the choice of $\{\varphi_i\}_{i \in \mathbb{N}}$ depends on q^{st}).

Recalling that $(b_u^\partial)^*$, $(\sigma_u^\partial)^*$ correspond to the Dirichlet traces we have, denoting $\tilde{b}_u = b_u(t, x, u)$ and similarly for $\tilde{\sigma}_u$ and \tilde{f}_u ,

$$\tilde{b}_u^* y^u(t) + \tilde{\sigma}_u^* q^u(t) - \tilde{f}_u = \begin{pmatrix} b_u^{st}(t, \cdot, x, u_{st})y^{\bar{u}}(t, \cdot) + \sigma_u^{st}(t, \cdot, u_{st})[Tr(q_{st}^{\bar{u}}(t))](\cdot) - \alpha_{st}|\bar{u}_{st}|^{p-2}\bar{u}_{st}(t, \cdot) \\ b_u^0(t, u_0)y^u(t, 0) + \sigma_u^0(t, u_0)q_0^u(t, 0) - \alpha_0|u_0|^{p-2}u_0(t) \\ b_u^1(t, u_1)y^u(t, 1) + \sigma_u^1(t, u_1)q_1^u(t, 1) - \alpha_1|u_1|^{p-2}u_1(t) \end{pmatrix},$$

It follows from theorem 6.8, because $T_{Uad}(\cdot) \equiv \mathbb{U}$, that

$$b_u^{st}(t, \xi, x, u_{st})y^{\bar{u}}(t, \cdot) + \sigma_u^{st}(t, \xi, u_{st})[Tr(q_{st}^{\bar{u}}(t))](\cdot) = \alpha_{st}|\bar{u}_{st}|^{p-2}\bar{u}_{st}(t, \xi) \quad (6.5)$$

and

$$b_u^i(t, \bar{u}_i)y^{\bar{u}}(t, i) + \sigma_u^i(t, \bar{u}_i)q_i^{\bar{u}}(t, i) = \alpha_i|\bar{u}|^{p-2}\bar{u}(t, i), \quad i = 0, 1. \quad (6.6)$$

Example 6.11. Let $\mathcal{D} \subset \mathbb{R}^d$. Let W_{st} be a cylindrical Wiener process on $L^2(\mathcal{D})$ (from now on $L^p := L^p(\mathcal{D})$).

$$\begin{aligned} \partial_t x(t, \xi) &= \Delta x(t, \xi) + (-\Delta)^{-\gamma} u(t, \xi) \dot{W}_{st}(t, \xi) & t \in (0, T), \xi \in \mathcal{D} \\ x(t, \xi) &= 0 & t \in (0, T), \xi \in \partial\mathcal{D} \\ x(0, \xi) &= x_0(\xi) & \xi \in \mathcal{D} \end{aligned} \quad (6.7)$$

We consider $A = \Delta$ and $D(A) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$, so that A generates an analytic semigroup on L^2 . Define $\mathcal{H}_{2s} = D((-A)^s)$. Notice that $\mathcal{H}_s = H^s(\mathcal{D})$ for $s \in (-\frac{1}{2}, \frac{1}{2})$.

Let $p > 2d$, $r = \frac{2}{p-2}$, notice that due to [11, Theorem 6.7], we have the inclusions $\mathcal{H}_{d/p} \subset L^r \subset L^2 \subset L^{r^*} \subset \mathcal{H}_{-d/p}$, where $\frac{1}{r} + \frac{1}{r^*} = 1$.

For $u \in L^p(\mathcal{D})$ we define $\Sigma(u)x = ux$, we have $\frac{1}{p} + \frac{1}{r} = \frac{1}{2}$, with this, $\Sigma(u) \in \mathcal{L}(L^2, L^{r^*})$, and

$$\|\Sigma(u)\|_{\mathcal{L}(L^2, L^{r^*})} = \|u\|_{L^p(\mathcal{D})}.$$

It follows from the spectral decomposition of A that when $\eta > \frac{d}{4} + \frac{d}{2p} - \gamma$,

$$(-\Delta)^{-\gamma}\Sigma(u) \in \mathcal{L}_2(L^2, \mathcal{H}_{-2\eta}) \quad \text{and} \quad \|(-\Delta)^{-\gamma}\Sigma(u)\|_{\mathcal{L}_2(L^2, \mathcal{H}_{-2\eta})} \leq C \|u\|_{L^p}.$$

W_{st} is a cylindrical Wiener process on L^2 . Denote $\mathbb{X}_s = \mathcal{H}_{2s}$.

Denote $\mathbb{U} := L^p$, and define $\sigma(u) = (-\Delta)^{-\gamma}\Sigma(u)$ we have hypothesis 6.2 for $\eta \in (\frac{1}{4}, \frac{1}{2})$ and

$$p > \max \left\{ \frac{2}{1-2\eta}, \frac{2d}{4\eta + 4\gamma - d} \right\},$$

this imposes $\gamma > \frac{d-2}{4}$.

For

$$J(u) = \frac{1}{2} \mathbb{E} \int_0^T \alpha_x \|x(t)\|_{L^2(0,1)}^2 dt + \frac{1}{p} \mathbb{E} \int_0^T \alpha_u \|u(t)\|_{\mathbb{U}}^p dt + \frac{\alpha_T}{2} \mathbb{E} \|x(T)\|_{L^2(0,1)}^2$$

we have hypothesis 6.3.

Formally, the abstract costate equation 5.1 may be regarded as

$$\begin{aligned} \partial_t y(t, \xi) &= -\partial_{\xi\xi} y(t, \xi) + \alpha_x x(s, \xi) + q(t, \xi) \dot{W}_{st}(t, \xi) & t \in (0, T), \xi \in \mathcal{D} \\ y(t, i) &= 0 & t \in (0, T), i = 0, 1 \\ y(T, \xi) &= -\alpha_T x(T, \xi) & \xi \in \mathcal{D}. \end{aligned}$$

Notice that because $\eta > \frac{d}{4} + \frac{d}{2p} - \gamma > \frac{d}{4} - \gamma$, the inclusions from $\mathcal{H}_{2\eta}$ into $\mathcal{H}_{d/p-2\gamma}$, and hence into $\mathcal{H}_{-2\gamma}$, are Hilbert-Schmidt, therefore, for $q \in \mathcal{L}_2(L^2, \mathcal{H}_{2\eta})$, we have $(-\Delta)^{-\gamma}q \in \mathcal{L}_1(L^2)$. Fix $q \in \mathcal{L}_2(L^2, \mathcal{H}_{2\eta})$, if we choose a Hilbert basis $\{\varphi_i\}_{i \in \mathbb{N}}$ of L^2 which consists of eigenvectors of $|(-\Delta)^{-\gamma}q|$, then

$$\begin{aligned} \int_{\mathcal{D}} \sum_{i \in \mathbb{N}} | [(-\Delta)^{-\gamma}q\varphi_i](\xi)\varphi_i(\xi) | d\xi &= \sum_{i \in \mathbb{N}} \int_{\mathcal{D}} | [(-\Delta)^{-\gamma}q\varphi_i](\xi)\varphi_i(\xi) | d\xi \leq \sum_{i \in \mathbb{N}} \|(-\Delta)^{-\gamma}q\varphi_i\|_{L^2} \|\varphi_i\|_{L^2} \\ &= \sum_{i \in \mathbb{N}} \|(-\Delta)^{-\gamma}q\varphi_i\|_{L^2} = \sum_{i \in \mathbb{N}} \| |(-\Delta)^{-\gamma}q| \varphi_i \|_{L^2} = \sum_{i \in \mathbb{N}} \langle |(-\Delta)^{-\gamma}q| \varphi_i, \varphi_i \rangle_{L^2} < \infty, \end{aligned}$$

it follows that the function $[Tr((-\Delta)^{-\gamma}q)](\xi) := \sum_{i \in \mathbb{N}} [(-\Delta)^{-\gamma}q\varphi_i](\xi)\varphi_i(\xi)$ is well defined for a.e. $\xi \in \mathcal{D}$ and belongs to L^1 , moreover, for $u \in L^\infty \subset L^p$, we have

$$\begin{aligned} \int_{\mathcal{D}} u(\xi) [\sigma_u^* q](\xi) d\xi &= \langle \sigma_u u, q \rangle_{\mathcal{L}_2(L^2, \mathcal{H}_{-2\eta}), \mathcal{L}_2(L^2, \mathcal{H}_{2\eta})} = \sum_{i \in \mathbb{N}} \langle (-\Delta)^{-\gamma}u\varphi_i, q\varphi_i \rangle_{\mathcal{H}_{-2\eta}, \mathcal{H}_{2\eta}} \\ &= \sum_{i \in \mathbb{N}} \langle u\varphi_i, (-\Delta)^{-\gamma}q\varphi_i \rangle_{\mathcal{H}_{-2\eta-2\gamma}, \mathcal{H}_{2\eta+2\gamma}} = \sum_{i \in \mathbb{N}} \langle u\varphi_i, (-\Delta)^{-\gamma}q\varphi_i \rangle_{L^{r^*}, L^r} = \int_{\mathcal{D}} u [Tr((-\Delta)^{-\gamma}q)](\xi) d\xi. \end{aligned}$$

It follows that $Tr((-\Delta)^{-\gamma}q) = \sigma_u^* q$ and therefore belongs to L^{p^*} and depends linearly and continuously on q (the linearity is not obvious from the definition because the choice of $\{\varphi_i\}_{i \in \mathbb{N}}$ depends on q).

It follows from theorem 6.8, because $T_{Uad}(\cdot) \equiv L^p$ that

$$b_u(t, \bar{x}, \bar{u})^* y^{\bar{u}}(t) + \sigma_u(t, \bar{x}, \bar{u})^* q^{\bar{u}}(t) - f_u(t, \bar{x}, \bar{u}) = Tr[(-\Delta)^\gamma q^{\bar{u}}] - \alpha_u |\bar{u}|^{p-2} \bar{u}(t) = 0.$$

Acknowledgments & Funding

The author expresses gratitude to Ioana Ciotir and Hasnaa Zidani from LMI - INSA Rouen Normandie, for fruitful discussions and invaluable support in the preparation of this manuscript.

This research work has been partially supported by Normandy Region and the European Union through ERDF research and innovation program, under the grant for "chaire d'excellence COPTI". The author also acknowledges support from Agence Nationale de la Recherche (ANR), project ANR-22-CE40-0010 COSS, and support from FMJH Program Gaspard Monge in Optimization and operations research including support to this program by EDF.

A Dual of Hilbert-Schmidt space

Proof of proposition 2.3. Recall that $\|P\|_{\mathcal{L}_2(Z,K)}^2 = \sum_{i \in I} \|Pe_i\|_K^2$ and $\|Q\|_{\mathcal{L}_2(Z,K^*)}^2 = \sum_{i \in I} \|Qe_i\|_{K^*}^2$. So that, by the Cauchy-Schwarz inequality in $\ell^2(I)$,

$$\langle P, Q \rangle_{\mathcal{L}_2(Z,K), \mathcal{L}_2(Z,K^*)} \leq \|P\|_{\mathcal{L}_2(Z,K)} \|Q\|_{\mathcal{L}_2(Z,K^*)}.$$

It follows that the duality defines a mapping $\mathcal{I} : \mathcal{L}_2(Z, K^*) \rightarrow (\mathcal{L}_2(Z, K))^*$ with at most unit norm. Fix $Q \in \mathcal{L}_2(Z, K^*)$, for each $i \in I$, define $P_i \in K$ as the Riesz representative of Qe_i , so that

$$\langle k, Qe_i \rangle_{K, K^*} = \langle k, P_i \rangle_K,$$

meaning

$$\langle P_i, Qe_i \rangle_{K, K^*} = \langle P_i, P_i \rangle_K = \|P_i\|_K^2 = \|Qe_i\|_{K^*}^2,$$

since $\|P_i\|_K \in \ell^2(I)$ it follows that P^Q defined as $P^Q e_i = P_i$ belongs to $\mathcal{L}_2(Z, K)$ and $\|P^Q\|_{\mathcal{L}_2(Z,K)} = \|Q\|_{\mathcal{L}_2(Z,K^*)}$, with this,

$$\|P^Q\|_{\mathcal{L}_2(Z,K)} \|Q\|_{\mathcal{L}_2(Z,K^*)} = \|Q\|_{\mathcal{L}_2(Z,K^*)}^2 = \sum_{i \in \mathbb{N}} \|Qe_i\|_{K^*}^2 = \sum_{i \in \mathbb{N}} \langle P_i, Qe_i \rangle_{K, K^*} = \langle P^Q, \mathcal{I}Q \rangle_{\mathcal{L}_2(Z,K), (\mathcal{L}_2(Z,K))^*},$$

so that \mathcal{I} is in fact an isometry from $\mathcal{L}_2(Z, K^*)$ into $(\mathcal{L}_2(Z, K))^*$.

We have left to prove that \mathcal{I} is onto. Let $F \in (\mathcal{L}_2(Z, K))^*$. Begin by noticing that $\mathcal{L}_2(Z, K)$ is a Hilbert space so that, again by the Riesz representation theorem, there must exist $P^F \in \mathcal{L}_2(Z, K)$ such that

$$\langle P, F \rangle_{\mathcal{L}_2(Z,K), (\mathcal{L}_2(Z,K))^*} = \langle P, P^F \rangle_{\mathcal{L}_2(Z,K)} = \sum_{i \in I} \langle Pe_i, P^F e_i \rangle_K,$$

now define $Q_i \in K^*$ given by

$$\langle Q_i, k \rangle_{K, K^*} = \langle P^F e_i, k \rangle_K$$

with which $\|Q_i\|_{K^*} = \|P^F e_i\|_K$ so that Q^F defined by $Q^F e_i = Q_i$ belongs to $\mathcal{L}_2(Z, K^*)$ and

$$\begin{aligned} \langle P, Q^F \rangle_{\mathcal{L}_2(Z,K), \mathcal{L}_2(Z,K^*)} &= \sum_{i \in I} \langle Pe_i, Qe_i \rangle_{K, K^*} = \sum_{i \in I} \langle Pe_i, P^F e_i \rangle_K \\ &= \sum_{i \in I} \langle Pe_i, P^F e_i \rangle_K = \langle P, P^F \rangle_{\mathcal{L}_2(Z,K), (\mathcal{L}_2(Z,K))^*}, \end{aligned}$$

so that $\mathcal{I}Q^F = F$.

The continuity of the inclusions $\mathcal{L}_2(Z, H) \subset \mathcal{L}_2(Z, K)$ and $\mathcal{L}_2(Z, K^*) \subset \mathcal{L}_2(Z, H^*)$ is direct, the density follows from the density of finite rank operators in the respective Hilbert-Schmidt spaces and the density of the inclusions $H \subset K$ and $K^* \subset H^*$.

The identity

$$\langle P, Q \rangle_{\mathcal{L}_2(Z, K), \mathcal{L}_2(Z, K^*)} = \langle P, Q \rangle_{\mathcal{L}_2(Z, H), \mathcal{L}_2(Z, H^*)}$$

follows from the fact that $\langle Pe_i, Qe_i \rangle_{K, K^*} = \langle Pe_i, Qe_i \rangle_{H, H^*}$ when $Pe_i \in H$ and $Qe_i \in K^*$. \square

B Proof of proposition 6.5

Proof of proposition 6.5. For $x_1, x_2 \in \mathbb{X}$ and $u_1, u_2 \in \mathbb{U}$, define

$$\begin{aligned} \tilde{b}_x(\cdot, x_1, x_2, u_1, u_2) &= \int_0^1 b_x(\cdot, \theta x_1 + (1 - \theta)x_2, \theta u_1 + (1 - \theta)u_2) d\theta, \\ \tilde{b}_u(\cdot, x_1, x_2, u_1, u_2) &= \int_0^1 b_u(\cdot, \theta x_1 + (1 - \theta)x_2, \theta u_1 + (1 - \theta)u_2) d\theta, \\ \tilde{\sigma}_x(\cdot, x_1, x_2, u_1, u_2) &= \int_0^1 \sigma_x(\cdot, \theta x_1 + (1 - \theta)x_2, \theta u_1 + (1 - \theta)u_2) d\theta, \\ \tilde{\sigma}_u(\cdot, x_1, x_2, u_1, u_2) &= \int_0^1 \sigma_u(\cdot, \theta x_1 + (1 - \theta)x_2, \theta u_1 + (1 - \theta)u_2) d\theta, \\ \tilde{f}_x(\cdot, x_1, x_2, u_1, u_2) &= \int_0^1 f_x(\cdot, \theta x_1 + (1 - \theta)x_2, \theta u_1 + (1 - \theta)u_2) d\theta, \\ \tilde{f}_u(\cdot, x_1, x_2, u_1, u_2) &= \int_0^1 f_u(\cdot, \theta x_1 + (1 - \theta)x_2, \theta u_1 + (1 - \theta)u_2) d\theta, \end{aligned}$$

and

$$\tilde{h}_x(x_1, x_2) = \int_0^1 h_x(\theta x_1 + (1 - \theta)x_2) d\theta$$

Write

So that, for $\varphi = b, \sigma, h, f$, we have

$$\varphi(\cdot, x_2, u_2) - \varphi(\cdot, x_1, u_1) = \tilde{\varphi}_x(\cdot, x_1, x_2, u_1, u_2)(x_2 - x_1) + \tilde{\varphi}_u(\cdot, x_1, x_2, u_1, u_2)(u_2 - u_1).$$

Due to hypothesis 6.2 and 6.3,

$$(\tilde{\varphi}_x(\cdot, x_1, x_2, u_1, u_2), \tilde{\varphi}_u(\cdot, x_1, x_2, u_1, u_2)) \rightarrow (\varphi_x(\cdot, x_1, u_1), \varphi_u(\cdot, x_1, u_1))$$

strongly when $x_2 \rightarrow x_1$ and $u_2 \rightarrow u_1$.

We introduce the notations $\tilde{\varphi}_x^\varepsilon(\cdot) = \tilde{\varphi}_x(\cdot, x^u, x^{u^\varepsilon}, u, u^\varepsilon)$ and $\tilde{\varphi}_u^\varepsilon(\cdot) = \tilde{\varphi}_u(\cdot, x^u, x^{u^\varepsilon}, u, u^\varepsilon)$, with which

$$\varphi(\cdot, x^{u^\varepsilon}, u^\varepsilon) - \varphi(\cdot, x^u, u) = \tilde{\varphi}_x^\varepsilon(\cdot)(x^{u^\varepsilon} - x^u) + \varepsilon \tilde{\varphi}_u^\varepsilon(\cdot)v,$$

as well as $\hat{\varphi}_x(t) = \varphi_x(t, x^u(t), u(t))$ and $\hat{\varphi}_u(t) = \varphi_u(t, x^u(t), u(t))$.

Consider the equation

$$\begin{aligned} d\delta_v^u(t) &= [A\delta_v^u(t) + \hat{b}_x(t)\delta_v^u(t) + \hat{b}_u(t)v(t)]dt + [\hat{\sigma}_x(t)\delta_v^u(t) + \hat{\sigma}_u(t)v(t)]dW(t) \\ \delta_v^u(0) &= 0 \end{aligned} \tag{B.1}$$

Because of theorem 4.4, there exists a unique $\delta_v^u \in L_{\mathcal{F}}^p(\Omega; C([0, T], \mathbb{X}))$ solving equation (B.1), moreover $\|\delta_v^u\|_{L_{\mathcal{F}}^p(\Omega; C([0, T], \mathbb{X}))} \leq C \|v\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathbb{U}))}$.

We now define $r_v^\varepsilon = x^{u^\varepsilon} - x^u - \varepsilon \delta_v^u$, which the is a solution of

$$\begin{aligned} dr_v^\varepsilon(t) &= [Ar_v^\varepsilon(t) + \tilde{b}_x^\varepsilon r_v^\varepsilon(t) + \varepsilon[(\tilde{b}_x^\varepsilon(t) - \hat{b}_x(t))\delta_v^u(t) + (\tilde{b}_u^\varepsilon(t) - \hat{b}_u(t))v(t)]]dt \\ &\quad + [\tilde{\sigma}_x^\varepsilon r_v^\varepsilon(t) + \varepsilon[(\tilde{\sigma}_x^\varepsilon(t) - \hat{\sigma}_x(t))\delta_v^u(t) + (\tilde{\sigma}_u^\varepsilon(t) - \hat{\sigma}_u(t))v(t)]]dW(t). \\ r_v^\varepsilon(0) &= 0 \end{aligned} \quad (\text{B.2})$$

It follows again from theorem 4.4 that

$$\begin{aligned} \|r_v^\varepsilon\|_{L_{\mathcal{F}}^p(\Omega; C([0, T], \mathbb{X}))} &\leq \varepsilon \left\| (\tilde{b}_x^\varepsilon(t) - \hat{b}_x(t))\delta_v^u(t) + (\tilde{b}_u^\varepsilon(t) - \hat{b}_u(t))v(t) \right\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-\frac{1}{2}}))} \\ &\quad + \varepsilon \left\| (\tilde{\sigma}_x^\varepsilon(t) - \hat{\sigma}_x(t))\delta_v^u(t) + (\tilde{\sigma}_u^\varepsilon(t) - \hat{\sigma}_u(t))v(t) \right\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))}. \end{aligned}$$

Let $\{\varepsilon_k\}_{k \in \mathbb{N}} \subset (0, \infty)$ be a sequence such that $\varepsilon_k \rightarrow 0$.

Due to proposition 4.7, $x^{u^{\varepsilon_k}} \rightarrow x^u$ in $L_{\mathcal{F}}^p(\Omega; C([0, T], \mathbb{X}))$ as $k \rightarrow \infty$, with this, there is a subsequence $\{\varepsilon_{k_j}\}_{k_j \in \mathbb{N}}$ for which $x^{u^{\varepsilon_{k_j}}} \rightarrow x^u$ uniformly on $[0, T]$ a.s. as $k \rightarrow \infty$. Because of the strong continuity of the derivatives and the dominated convergence theorem, we have

$$\left\| (\tilde{b}_x^{\varepsilon_{k_j}} - \hat{b}_x)\delta_v^u(t) + (\tilde{b}_u^{\varepsilon_{k_j}} - \hat{b}_u)v(t) \right\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-\frac{1}{2}}))} \rightarrow 0$$

and

$$\left\| (\tilde{\sigma}_x^{\varepsilon_{k_j}} - \hat{\sigma}_x)\delta_v^u(t) + (\tilde{\sigma}_u^{\varepsilon_{k_j}} - \hat{\sigma}_u)v(t) \right\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))} \rightarrow 0,$$

because the choice of $\{\varepsilon_k\}$ is arbitrary, it follows that

$$\left\| (\tilde{b}_x^\varepsilon - \hat{b}_x)\delta_v^u(t) + (\tilde{b}_u^\varepsilon - \hat{b}_u)v(t) \right\|_{L_{\mathbb{F}}^p(\Omega; L^2(0, T; \mathbb{X}_{-\frac{1}{2}}))} \rightarrow 0$$

and

$$\left\| (\tilde{\sigma}_x^\varepsilon - \hat{\sigma}_x)\delta_v^u(t) + (\tilde{\sigma}_u^\varepsilon - \hat{\sigma}_u)v(t) \right\|_{L_{\mathbb{F}}^p(\Omega; L^p(0, T; \mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta))} \rightarrow 0,$$

and therefore

$$\|r_v^\varepsilon\|_{L_{\mathcal{F}}^p(\Omega; C([0, T], \mathbb{X}))} = o(\varepsilon).$$

It follows from the definition of the transposition solution (y^u, q^u) of equation (5.1), that

$$\begin{aligned} &\mathbb{E}\langle -h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \\ &+ \mathbb{E} \int_0^T \langle b_x(s, x^u(s), u(s))^* y^u(t) + \sigma_x(s, x^u(s), u(s))^* q^u(t) - f_x(t, x^u(t), u(t)), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ &= + \mathbb{E} \int_0^T \langle y^u(s), b_x(s, x^u(s), u(s))\delta_v^u(s) + b_u(s, x^u(s), u(s))v(s) \rangle_{\mathbb{X}_{1/2}^*, \mathbb{X}_{-1/2}} ds \\ &+ \mathbb{E} \int_0^T \langle q^u(s), \sigma_x(s, x^u(s), u(s))\delta_v^u(s) + \sigma_u(s, x^u(s), u(s))v(s) \rangle_{\mathcal{L}_2(\mathbb{H}, \mathbb{X}_\eta^*), \mathcal{L}_2(\mathbb{H}, \mathbb{X}_{-\eta})} ds, \end{aligned} \quad (\text{B.3})$$

it follows that

$$\begin{aligned} &\mathbb{E}\langle -h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \mathbb{E} \int_0^T \langle -f_x(t, x^u(t), u(t)), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} ds \\ &= \mathbb{E} \int_0^T \langle b_u(t, x^u(t), u(t))^* y^u(t) + \sigma_u(t, x^u(t), u(t))^* q^u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt. \end{aligned} \quad (\text{B.4})$$

The local Lipschitz condition for f given in hypothesis 3.4, means that the derivative $f_x(t, x^u(t), u(t))$ given by hypothesis 5.2 satisfies

$$\|f_x(t, x^u(t), u(t))\|_{\mathbb{X}^*} \leq C_1 (1 + 2 \|u(t)\|_{\mathbb{U}} + 2 \|x^u(t)\|_{\mathbb{X}})^{p-1}$$

with which

$$\|f_u(\cdot, x^u(\cdot), u(\cdot))\|_{L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathbb{X}^*))} \leq C_1 \left(T^{1/p} + 2 \|u(t)\|_{\mathcal{L}_{\mathbb{F}}^p(0, T; L^p(\Omega; \mathbb{U}))} + 2T^{1/p} \|x^u(t)\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} \right)^{p-1},$$

the derivative $f_u(t, x^u(t), u(t))$ given by hypothesis 6.3 satisfies

$$\|f_u(t, x^u(t), u(t))\|_{\mathbb{U}^*} \leq C_1 (1 + 2 \|u(t)\|_{\mathbb{U}} + 2 \|x^u(t)\|_{\mathbb{X}})^{p-1}$$

with which

$$\|f_u(\cdot, x^u(\cdot), u(\cdot))\|_{L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathbb{U}^*))} \leq C_1 \left(T^{1/p} + 2 \|u(t)\|_{\mathcal{L}_{\mathbb{F}}^p(0, T; L^p(\Omega; \mathbb{U}))} + 2T^{1/p} \|x^u(t)\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} \right)^{p-1}.$$

Similarly the local Lipschitz condition for h given in hypothesis 3.4, means that the derivative $h_x(x^u(T))$ given by hypothesis 5.2 satisfies

$$\|h_x(x^u(T))\|_{\mathbb{X}^*} \leq C_1 (1 + 2 \|x^u(T)\|_{\mathbb{X}})^{p-1}$$

with which

$$\|h_x(x^u(T))\|_{L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)} \leq C_1 \left(1 + 2 \|x^u(T)\|_{L_{\mathcal{F}_T}^p(\Omega; \mathbb{X})} \right)^{p-1}.$$

We may show in a similar way that \tilde{f}_x^ε , \tilde{f}_u^ε and \tilde{h}_x^ε are uniformly bounded for $\varepsilon \in (0, 1)$ in $L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathbb{X}^*))$, $L_{\mathbb{F}}^{p^*}(0, T; L^{p^*}(\Omega; \mathbb{U}^*))$ and $L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)$ respectively.

We now have:

$$\begin{aligned} & J(u^\varepsilon) - J(u) + \varepsilon \mathbb{E} \int_0^T \langle H_u(t, x^u(t), u(t), y^u(t), q^u(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \\ &= \mathbb{E} \int_0^T f(t, x^{u^\varepsilon}(t), u^\varepsilon(t)) dt + \mathbb{E} h(x^{u^\varepsilon}(T)) - \mathbb{E} \int_0^T f(t, x^u(t), u(t)) dt - \mathbb{E} h(x^u(T)) \\ &+ \varepsilon \mathbb{E} \int_0^T \langle H_u(t, x^u(t), u(t), y^u(t), q^u(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \\ &= \mathbb{E} \int_0^T \left[\langle \tilde{f}_x^\varepsilon(t), x^{u^\varepsilon}(t) - x^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \varepsilon \langle \tilde{f}_u^\varepsilon(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} \right] dt + \mathbb{E} \langle \tilde{h}_x^\varepsilon, x^{u^\varepsilon}(T) - x^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \\ &+ \varepsilon \mathbb{E} \int_0^T \langle b_u(t, x^u(t), u(t))^* y^u(t) + \sigma_u(t, x^u(t), u(t))^* q^u(t) - f_u(t, x^u(t), u(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \\ &= \mathbb{E} \int_0^T \left[\langle \tilde{f}_x^\varepsilon(t), r_v^\varepsilon(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \varepsilon (\langle \tilde{f}_x^\varepsilon(t), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \langle \tilde{f}_u^\varepsilon(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}}) \right] dt + \mathbb{E} \left[\langle \tilde{h}_x^\varepsilon r_v^\varepsilon(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \varepsilon \langle \tilde{h}_x^\varepsilon, \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \right] \\ &+ \varepsilon \mathbb{E} \langle -h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \varepsilon \mathbb{E} \int_0^T \langle -f_x(t, x^u(t), u(t)), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} dt + \varepsilon \mathbb{E} \int_0^T \langle -f_u(t, x^u(t), u(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \\ &= \mathbb{E} \int_0^T \left[\langle \tilde{f}_x^\varepsilon(t), r_v^\varepsilon(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \varepsilon (\langle \tilde{f}_x^\varepsilon(t) - \hat{f}_x(t), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \langle \tilde{f}_u^\varepsilon(t) - \hat{f}_u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}}) \right] dt \\ &+ \mathbb{E} \left[\langle \tilde{h}_x^\varepsilon r_v^\varepsilon(T) \rangle_{\mathbb{X}^*, \mathbb{X}} + \varepsilon \langle \tilde{h}_x^\varepsilon - h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \right] \end{aligned}$$

it follows that

$$\begin{aligned}
& \left| J(u^\varepsilon) - J(u) + \varepsilon \mathbb{E} \int_0^T \langle H_u(t, x^u(t), u(t), y^u(t), q^u(t)), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \right| \\
& \leq \left\| \tilde{f}_x^\varepsilon \right\|_{L_{\mathbb{F}}^{p^*}(\Omega; L^1(0, T; X^*))} \|r_v^\varepsilon\|_{L_{\mathbb{F}}^p(\Omega; C([0, T], \mathbb{X}))} + \left\| \tilde{h}_x^\varepsilon \right\|_{L_{\mathcal{F}_T}^{p^*}(\Omega; X^*)} \|r_v^\varepsilon(T)\|_{L_{\mathcal{F}_T}^p(\Omega; \mathbb{X})} \\
& + \varepsilon \left| \mathbb{E} \int_0^T \left[\langle \tilde{f}_x^\varepsilon(t) - \hat{f}_x(t), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \langle \tilde{f}_u^\varepsilon(t) - \hat{f}_u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} \right] dt + \mathbb{E} \langle \tilde{h}_x^\varepsilon - h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \right| \\
& = \varepsilon \left| \mathbb{E} \int_0^T \left[\langle \tilde{f}_x^\varepsilon(t) - \hat{f}_x(t), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \langle \tilde{f}_u^\varepsilon(t) - \hat{f}_u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} \right] dt + \mathbb{E} \langle \tilde{h}_x^\varepsilon - h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \right| + o(\varepsilon).
\end{aligned}$$

As before, let $\{\varepsilon_k\}_{k \in \mathbb{N}} \subset (0, \infty)$ be a sequence such that $\varepsilon_k \rightarrow 0$.

Due to proposition 4.7, $x^{u^{\varepsilon_{k_j}}} \rightarrow x^u$ in $L_{\mathcal{F}}^p(\Omega; C([0, T], \mathbb{X}))$ as $j \rightarrow \infty$, with this, there is a subsequence, which we still denote $\{\varepsilon_{k_j}\}_{j \in \mathbb{N}}$, for which $x^{u^{\varepsilon_{k_j}}} \rightarrow x^u$ uniformly on $[0, T]$ a.s. as $j \rightarrow \infty$.

Because of the strong continuity of the derivatives, it follows that $\tilde{f}_x^{\varepsilon_{k_j}}(t) \rightarrow \hat{f}_x(t)$ in \mathbb{X}^* a.e. on $[0, T] \times \Omega$ and $\tilde{h}_x^{\varepsilon_{k_j}} \rightarrow h_x(x^u(T))$ in \mathbb{X}^* a.s. on Ω (this because for a Banach space X strong convergence in $\mathcal{L}(X, \mathbb{C})$ is equivalent to weak* convergence in X^*).

Because $L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathbb{X}^*))$, $L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathbb{U}^*))$, and $L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*)$ are reflexive, there is a subsequence, which we still denote $\{\varepsilon_{k_j}\}_{j \in \mathbb{N}}$, and elements

$$\Phi_x \in L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathbb{X}^*)), \quad \Phi_u \in L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathbb{U}^*)) \quad \text{and} \quad \Phi_T \in L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*),$$

for which $\tilde{f}_x^{\varepsilon_{k_j}} \rightarrow \Phi_x$, $\tilde{f}_u^{\varepsilon_{k_j}} \rightarrow \Phi_u$ and $\tilde{h}_x^{\varepsilon_{k_j}} \rightarrow \Phi_T$ as $j \rightarrow \infty$.

Because of the Mazur lemma there exists a sequence of weights $a_{j,i}$ such that $a_{j,\cdot}$ is supported on a finite subset of $\{i \in \mathbb{N} : i \geq j\}$, $0 \leq a_{j,i}$ and $\sum_{i \geq j} a_{j,i} = 1$, satisfying

$$\Phi_x^j := \sum_{i \geq j} a_{j,i} \tilde{f}_x^{\varepsilon_{k_i}} \rightarrow \Phi_x \text{ in } L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathbb{X}^*)),$$

$$\Phi_u^j := \sum_{i \geq j} a_{j,i} \tilde{f}_u^{\varepsilon_{k_i}} \rightarrow \Phi_u \text{ in } L_{\mathbb{F}}^{p^*}(0, T; L^p(\Omega; \mathbb{U}^*))$$

and

$$\Phi_T^j := \sum_{i \geq j} a_{j,i} \tilde{h}_x^{\varepsilon_{k_i}} \rightarrow \Phi_T \text{ in } L_{\mathcal{F}_T}^{p^*}(\Omega; \mathbb{X}^*).$$

With this, there is a subsequence $\{j_\ell\}_{\ell \in \mathbb{N}}$ such that $\Phi_x^{j_\ell} \rightarrow \Phi_x$ and $\Phi_u^{j_\ell} \rightarrow \Phi_u$ a.e. on $[0, T] \times \Omega$ and $\Phi_T^{j_\ell} \rightarrow \Phi_T$ a.s. on Ω . This implies that $\Phi_x = \hat{f}_x$, $\Phi_u = \hat{f}_u$ and $h_x(x^u(T)) = \Phi_T$. The weak convergence means that

$$\mathbb{E} \int_0^T \left[\langle \tilde{f}_x^{\varepsilon_{k_j}}(t) - \hat{f}_x(t), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \langle \tilde{f}_u^{\varepsilon_{k_j}}(t) - \hat{f}_u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} \right] dt + \mathbb{E} \langle \tilde{h}_x^{\varepsilon_{k_j}} - h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \rightarrow 0,$$

because the choice of $\{\varepsilon_k\}$ is arbitrary, it follows that

$$\mathbb{E} \int_0^T \left[\langle \tilde{f}_x^\varepsilon(t) - \hat{f}_x(t), \delta_v^u(t) \rangle_{\mathbb{X}^*, \mathbb{X}} + \langle \tilde{f}_u^\varepsilon(t) - \hat{f}_u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} \right] dt + \mathbb{E} \langle \tilde{h}_x^\varepsilon - h_x(x^u(T)), \delta_v^u(T) \rangle_{\mathbb{X}^*, \mathbb{X}} \rightarrow 0.$$

□

Remark B.1. *The assumption that \mathbb{U} is reflexive can be relaxed if we assume that f is Lipschitz continuous w.r.t. u , as in that case we can treat the term $\mathbb{E} \int_0^T \langle \tilde{f}_u^\varepsilon(t) - \hat{f}_u(t), v(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt$ using the dominated convergence theorem.*

References

- [1] Jean-Pierre Aubin and Hélène Frankowska. *Set-valued analysis*. Springer Science & Business Media, 2009.
- [2] Alain Bensoussan. Stochastic maximum principle for distributed parameter systems. *Journal of the Franklin Institute*, 315(5-6):387–406, 1983.
- [3] Alain Bensoussan, Giuseppe Da Prato, Michel C. Delfour, and Sanjoy K. Mitter. *Representation and control of infinite dimensional systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, second edition, 2007.
- [4] Jean-Michel Bismut. *Convex analysis and probability*. PhD thesis, Université Paris VI, 1973.
- [5] Vladimir Grigorjewitsch Boltyanskiy, Revaz Valer’yanovich Gamkrelidze, Evgenii Frolovich Mishchenko, and Lev Semenovich Pontryagin. *Mathematical theory of optimal processes*. 1962.
- [6] Eduardo Casas. Pontryagin’s principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM Journal on Control and Optimization*, 35(4):1297–1327, 1997.
- [7] Francis Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264. Springer, 2013.
- [8] Giuseppe Da Prato, S Kwapien, and Jerzy Zabczyk. Regularity of solutions of linear stochastic equations in hilbert spaces. *Stochastics: An International Journal of Probability and Stochastic Processes*, 23(1):1–23, 1988.
- [9] Giuseppe Da Prato and Jerzy Zabczyk. Evolution equations with white-noise boundary conditions. *Stochastics and Stochastic Reports*, 42(3-4):167–182, 1993.
- [10] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.
- [11] Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker’s guide to the fractional sobolev spaces. *Bulletin des sciences mathématiques*, 136(5):521–573, 2012.
- [12] Kai Du and Qingxin Meng. A maximum principle for optimal control of stochastic evolution equations. *SIAM Journal on Control and Optimization*, 51(6):4343–4362, 2013.
- [13] Hélène Frankowska and Qi Lü. First and second order necessary optimality conditions for controlled stochastic evolution equations with control and state constraints. *Journal of Differential Equations*, 268(6):2949–3015, 2020.
- [14] Hélène Frankowska, Haisen Zhang, and Xu Zhang. First and second order necessary conditions for stochastic optimal controls. *Journal of Differential Equations*, 262(6):3689–3736, 2017.
- [15] Marco Fuhrman, Ying Hu, and Gianmario Tessitore. Stochastic maximum principle for optimal control of spdes. *Applied Mathematics & Optimization*, 68:181–217, 2013.
- [16] Marco Fuhrman, Ying Hu, and Gianmario Tessitore. Stochastic maximum principle for optimal control of partial differential equations driven by white noise. *Stochastics and Partial Differential Equations: Analysis and Computations*, 6:255–285, 2018.

- [17] Marco Fuhrman and Gianmario Tessitore. Nonlinear kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. *The Annals of Probability*, 30(3):1397–1465, 2002.
- [18] Giuseppina Guatteri. Stochastic maximum principle for spdes with noise and control on the boundary. *Systems & control letters*, 60(3):198–204, 2011.
- [19] Harold Kushner. Necessary conditions for continuous parameter stochastic optimization problems. *SIAM Journal on Control*, 10(3):550–565, 1972.
- [20] Harold Kushner and Fred Schweppe. A maximum principle for stochastic control systems. *Journal of Mathematical Analysis and Applications*, 8(2):287–302, 1964.
- [21] Irena Lasiecka and Roberto Triggiani. *Control theory for partial differential equations: Volume 1, Abstract parabolic systems: Continuous and approximation theories*, volume 1. Cambridge University Press, 2000.
- [22] Xungjing Li and Jiongmin Yong. *Optimal control theory for infinite dimensional systems*. Birkhäuser Boston, 1995.
- [23] Qi Lü, Jiongmin Yong, and Xu Zhang. Representation of itô integrals by lebesgue/bochner integrals. *Journal of the European Mathematical Society*, 14(6):1795–1823, 2012.
- [24] Qi Lü and Xu Zhang. *General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions*. Springer, 2014.
- [25] Qi Lü and Xu Zhang. Mathematical control theory for stochastic partial differential equations. *Probability Theory and Stochastic Modelling*, 101, 2021.
- [26] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44. Springer-Verlag, 1983.
- [27] Shige Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization*, 28(4):966–979, 1990.
- [28] Jean-Pierre Raymond and Hasnaa Zidani. Pontryagin’s principle for state-constrained control problems governed by parabolic equations with unbounded controls. *SIAM Journal on Control and Optimization*, 36(6):1853–1879, 1998.
- [29] Jean-Pierre Raymond and Hasnaa Zidani. Hamiltonian pontryagin’s principles for control problems governed by semilinear parabolic equations. *Applied Mathematics and Optimization*, 39(2):143–177, 1999.
- [30] Michael Reed and Barry Simon. *Functional analysis*, volume I. Academic press, 1981.
- [31] Fredi Tröltzsch. *Optimal control of partial differential equations: theory, methods, and applications*, volume 112. American Mathematical Soc., 2010.
- [32] Richard B. Vinter. *Optimal control*, volume 2. Springer, 2010.
- [33] Tianxiao Wang and Haisen Zhang. Optimal control problems of forward-backward stochastic volterra integral equations with closed control regions. *SIAM Journal on Control and Optimization*, 55(4):2574–2602, 2017.

- [34] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.
- [35] Huaiqiang Yu and Bin Liu. Optimality conditions for stochastic boundary control problems governed by semilinear parabolic equations. *Journal of Mathematical Analysis and Applications*, 395(2):654–672, 2012.