Article:

Geldhauser, C. and Valdinoci, E. (2018) Optimizing the fractional power in a model with stochastic PDE constraints. Advanced Nonlinear Studies, 18 (4). pp. 649-669. ISSN: 1536-1365

https://doi.org/10.1515/ans-2018-2031

Research Article

Carina Geldhauser and Enrico Valdinoci*

Optimizing the Fractional Power in a Model with Stochastic PDE Constraints

https://doi.org/10.1515/ans-2018-2031
Received April 23, 2018; revised August 20, 2018; accepted August 24, 2018

Abstract: We study an optimization problem with SPDE constraints, which has the peculiarity that the control parameter s is the s-th power of the diffusion operator in the state equation. Well-posedness of the state equation and differentiability properties with respect to the fractional parameter s are established. We show that under certain conditions on the noise, optimality conditions for the control problem can be established.

Keywords: Stochastic Heat Equation, Optimization, Optimal Control, Fractional Parameter

MSC 2010: 65K10, 35R60, 35R11

Communicated by: Laurent Veron

1 Introduction

Generally speaking, optimal control problems with constraints are formulated as

$$\min_{y\in Y, u\in U} \mathcal{J}(y, u) \quad \text{subject to } \operatorname{Constr}(y, u) = 0,$$

where $\mathcal{J}$ is a cost functional, y the state variable, u the control variable and Constr is a constraint, usually in the form of an equation for y, called the "state equation". An important subcategory arises when the Constr is a partial differential equation, so that the task is the identification of coefficient functions or right-hand sides in the PDE: these are often called "identification problems" in the literature.

The purpose of this work is to study an identification problem with two peculiarities: (i) the control variable appears as the (fractional) exponent of a diffusion operator, and (ii) the constraint will be a stochastic PDE in the sense of a PDE driven by a Wiener Process. The optimal control parameter is therefore the answer to the question "*what is the optimal (fractional) diffusion pattern?*", which appears in applications, for instance, in the following way: In biology, y represents the density of a biological species exhibiting anomalous diffusion driven by a fractional operator and combined with a random perturbation. Such fractional diffusion processes are supposed to model very well the forage behavior or certain species; see, e.g., [12].

In engineering and economics, the problem of optimization under uncertainties is also wide-spread: Our model could serve as a very first toy problem to mathematically investigate the maximization of the probabilistic incremental Net Present Value for selecting the location of injection and production wells in petroleum engineering. The geometry and extension of such wells are crucial to the success of oil extraction in mature oil fields, where the diffusion of injected polymers within the oil field is studied; see, e.g., [23, 26].

Carina Geldhauser, Chebyshev Laboratory, St. Petersburg State University, 14th Line V.O., 29B, Saint Petersburg 199178 Russia, e-mail: k.geldhauser@spbu.ru. http://orcid.org/0000-0002-9997-6710
***Corresponding author: Enrico Valdinoci,** Department of Mathematics and Statistics, University of Western Australia, 35 Stirling Highway, Crawley, Perth, Western Australia 6009, Australia; and Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini 50, 20133 Milan; and Istituto di Matematica Applicata e Tecnologie Informatiche, Consiglio Nazionale delle Ricerche, Via Ferrata 1, 27100 Pavia, Italy, e-mail: enrico.valdinoci@uwa.edu.au. http://orcid.org/0000-0001-6222-2272

We stress that in the available literature the term “stochastic PDE constraints” usually refers to deterministic PDEs with “random input” in the sense of random coefficients of the PDE or of the force term; see [5, 7, 10, 11, 13, 27]. These problems, where the cost functional has deterministic output (due to the usage of $(L^2(D) \otimes L^2(\Omega))$-norms see, e.g., [11, 21]), are interesting due to their challenges for numerical approximation, in particular avoiding the “curse of dimensionality”. A different approach is to study expectation and variance of random cost functionals under stochastic constraints. These arise in economics, for example when optimizing a portfolio with finitely many assets; see, e.g., [9, 18]. The scope here is to find efficient portfolios, namely those minimizing the risk (i.e. the uncertainty of the return) or maximize the mean return for a given risk value using stochastic dominance constraints; see also [17] for an overview on the (finite-dimensional) mean variance analysis.

Motivated by such applications, this work derives optimality conditions and the existence of optimal controls for a random cost functional, a task which was, to the authors’ best knowledge, not considered up to now.

The second peculiarity in our approach is that the control variable of our problem is the fractional power of the differential operator. Our work is therefore the prototypical stochastic extension of the work [24], where this class of identification problems was introduced for the first time in a deterministic setting. This new type of problem poses several interesting mathematical challenges, among which we mention the need for a compactness theorem adapted for variable Banach spaces and the need for pathwise existence of the stochastic convolution, which is crucial for the derivation of the optimal random cost functional.

The control theory of fractional operators of diffusion type is a very new topic. Available results include the recent papers [2–4, 6]. In these works, however, the fractional operator was fixed a priori. In our case, the type of fractional-order operator itself is to be determined. From the point of view of applications, it is natural to optimize over the fractional power s: As a possible application, we can interpret the model as optimizing the mean radius of search for qualified workforce around a given location (normally the company’s production site). Uncertainty enters into these questions when considering non-negligible fluctuations in the mobility of the workforce, for example due to personal constraints. We note that the use of mathematical models to deal with problems in the job market is an important topic of contemporary research; see, e.g., [19, 20, 25] and the references therein.

Problem statement. Let $D \subset \mathbb{R}$ be a given bounded, open domain, and denote by $D_T := D \times (0, T)$ the space-time cylinder. In D_T, we consider the evolution of a fractional diffusion process governed by the s-th power of a positive definite operator $\mathcal{L}$, which has a discrete spectrum. Note that the fractional parameter $s > 0$ can also be greater than one. The prototypical example of $\mathcal{L}$ which we have in mind is (minus) the Laplacian endowed with Dirichlet boundary conditions with domain $H^2(D) \cap H_0^1(D)$.

For a given target function $y_{D_T}(x, t) \in L^2(D_T)$, and a non-negative smooth penalty function $\Phi(s)$, for each $\omega \in \Omega$ we want to prove the existence of a pathwise optimal cost $\mathcal{J}(\omega)$, defined as a minimizer in s and y of the cost functional

$$\mathcal{J}(y, s, \omega) = \int_0^T \int_D |y(s, x, t, \omega) - y_{D_T}(x, t)|^2 \, dx \, dt + \Phi(s) \tag{1.1}$$

subject to the state equation

$$\begin{cases} dy(t) + \mathcal{L}^s y(t) \, dt = dW(t) & \text{in } D \times [0, T] \\ y(\cdot, 0) = y_0 & \text{in } D, \end{cases} \tag{1.2}$$

where $y_0 \in L^2(D)$ is a given initial condition and $W = W(x, t)$ an L^2-Wiener process. The minimizer $\mathcal{J}(\omega)$ of (1.1) subject to (1.2) is called the *solution to the identification problem* (IP).

The penalty function $\Phi(s)$ is given a priori, and, from a technical point of view, it has to be chosen such that the problem has sufficient compactness properties in s. To this end, to simplify technicalities, we assume that $\Phi \in C^2(0, L)$ for some $L \in (0, +\infty]$, that Φ is non-negative and that it satisfies

$$\lim_{s \to 0^+} \Phi(s) = +\infty = \lim_{s \to L^-} \Phi(s). \tag{1.3}$$

In [24], as typical examples of functions satisfying these assumptions, the cases

$$\Phi(s) = \frac{1}{s(L-s)} \quad \text{when } L \in (0, +\infty),$$

and

$$\Phi(s) = \frac{e^s}{s} \quad \text{when } L = +\infty$$

were explicitly taken into consideration.

Note that the operator $\mathcal{L}^s$ is defined as the s-th power of $\mathcal{L}$, and this definition does not correspond to the usual definition of a fractional Laplacian operator via a singular integral. We refer to, e.g., [1, 22], and to Section 2.1 here for details about this point.

Optimizing the fractional exponent s is challenging already in the deterministic case since, when the fractional parameter s changes, so does the domain of definition of the operator $\mathcal{L}$, and with it the underlying space of functions of the fractional operator. This causes difficulties, e.g., when proving the existence of optimal controls, as the usual compactness arguments are not directly applicable. Similar to the deterministic framework of [24], we tackle this issue by a hand-tailored compactness argument.

Outline of this work. The structure of this work is as follows: In Section 3, we establish existence results of solutions to (1.2) and identify the set of admissible controls. In Section 4, we derive the differentiability properties of the control-to-state mapping $s \mapsto u(s)$, and then use them to identify necessary and sufficient optimality conditions for the control problem (IP), which means optimizing (1.1) subject to (1.2). Then in Section 5 we prove the existence of optimal controls, namely Theorem 5.2, and, more specifically, we show that $J(s, \omega)$ attains a minimum if ω is fixed, and s is in the set of admissible controls. The paper ends with Appendix A stating an ancillary result of Borel–Cantelli type, which is used in the main proofs.

We remark that the optimal fractional parameter $\bar{s} = \bar{s}(\omega)$ depends on ω since it is obtained by the optimizing problem in (IP) for a fixed ω (but we often write simply $\bar{s}$ instead of $\bar{s}(\omega)$ for typographical convenience).

The main results of this work are Theorem 4.4 on the optimality conditions, and Theorem 5.2 on the existence of optimal controls. Before diving into technicalities, we state a "toy version" of our main results as follows.

Theorem 1.1. *Under suitable assumptions on the regularity of the noise and on the initial data, the control problem* (IP) *has a solution, that is, for almost every fixed $\omega \in \Omega$, the cost functional $\mathcal{J}(\omega)$ attains a minimum in the set of admissible controls.*

Moreover, the following optimality conditions hold for a fixed realization $\omega \in \Omega$:

(i) Necessary condition: *If $\bar{s} = \bar{s}(\omega)$ is an optimal parameter for* (IP)*, and $y(\bar{s})$ is the associated unique solution to the state system* (1.2)*, then for almost every $\omega \in \Omega$,*

$$\int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s})\, dx\, dt + \Phi'(\bar{s}) = 0. \tag{1.4}$$

(ii) Sufficient condition: *If $\bar{s} = \bar{s}(\omega) \in (0, L)$ satisfies the necessary condition* (1.4)*, and if in addition*

$$\int_0^T \int_D (\partial_s y(\bar{s}))^2 + (y(\bar{s}) - y_D)\partial_{ss}^2 y(\bar{s})\, dx\, dt + \Phi''(\bar{s}) > 0$$

for almost every $\omega \in \Omega$, then $\bar{s}$ is optimal for (IP).

The "suitable assumptions" are made precise in Assumptions 3.1, 3.2 and 3.3, that are all stated at the beginning of Section 3. Roughly speaking, the assumptions are that the covariance operator Q and the linear operator $\mathcal{L}$ can be diagonalized in the same basis of eigenfunctions (this is the content of Assumption 3.1), that the eigenvalues of the diffusive operator are positive and diverging (this is the content of Assumption 3.2), in fact they diverge sufficiently fast to make a fractional series summable, and the size of the eigenvalues of Q

is controlled by the decay of the eigenvalues of the diffusive operator in a suitable duality sense (a precise statement of this is given in Assumption 3.3).

We would like to point out that the results we obtained are for a fixed realization of the solution y, they do not imply that the solution $\mathcal{J}(\omega, \bar{s})$ to the identification problem (IP) is a random variable since, due to the minimizing procedure, measurability properties may be lost. Therefore, the study of expectation and variance of our random cost functional, as in the finite-dimensional case [9, 17, 18], remains an open problem.

2 Notation and Setup

2.1 The Functional Analytic Setting

We denote by $D \subset \mathbb{R}$ a bounded domain and by $x \in D$ the space variable. We will work in the space $L^2(D)$ of square-integrable functions over D, and denote by $\langle \cdot , \cdot \rangle$ the scalar product in $L^2(D)$.

Let $\mathcal{L} : D(\mathcal{L}) \subset L^2(D) \to L^2(D)$ be a densely defined, linear, self-adjoint, positive operator, which is not necessarily bounded but with compact inverse. Hence there exist an orthonormal basis $\{e_j\}_{j\in\mathbb{N}}$ of $L^2(D)$ made of eigenfunctions of $\mathcal{L}$ and a sequence of real numbers λ_j such that $\mathcal{L}e_j = \lambda_j e_j$ and $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_j \to +\infty$ as $j \to +\infty$, the corresponding eigenvalues of $\mathcal{L}$.

In this setting, we can write every function $v \in L^2(D)$ in the form

$$v = \sum_{j=1}^{+\infty} \langle v, e_j \rangle e_j$$

and denote

$$v_j := \langle v, e_j \rangle,$$

so that

$$v = \sum_{j=1}^{+\infty} v_j e_j.$$

The domain of $\mathcal{L}$ is characterized by

$$D(\mathcal{L}) = \Big\{ v \in H : \sum_{j=1}^{+\infty} \lambda_j^2 \langle v, e_j \rangle^2 < +\infty \Big\}.$$

Thus, $-\mathcal{L}$ is the generator of an analytic semigroup of contractions which has the well-known structure

$$S(t)v = \sum_{j=1}^{+\infty} e^{-\lambda_j t} v_j e_j. \tag{2.1}$$

In our framework, the semigroup structure will be a crucial property when we study the features of the trajectories of solutions to (1.2); see the forthcoming Lemma 3.13. In analogy to [24], we use for v in the domain of $\mathcal{L}$ the notation

$$v \in \mathcal{H}^1 := \{\phi \in L^2(D) : \{\lambda_j \langle \phi, e_j \rangle\}_{j\in\mathbb{N}} \in l^2\}.$$

In this way we can write

$$\mathcal{L}v = \sum_{j=1}^{+\infty} \lambda_j \langle v, e_j \rangle e_j.$$

Similarly, given $s > 0$, we can define the (spectral) s-th power of $\mathcal{L}$ via

$$\mathcal{L}^s v = \sum_{j=1}^{+\infty} \lambda_j^s \langle v, e_j \rangle e_j, \tag{2.2}$$

and describe the domain of $\mathcal{L}^s$ as

$$D(\mathcal{L}^s) = \Big\{ v = \sum_{j=1}^{+\infty} v_j e_j : v_j \in \mathbb{R} \text{ with } \|v\|_s^2 := \|\mathcal{L}^s v\|^2 = \sum_{j=1}^{+\infty} \lambda_j^{2s} v_j^2 < +\infty \Big\}.$$

It is a classical approach in SPDE to define fractional powers of linear operators in this way; see, e.g., [8, 16] or also the more recent work [14]. Next, we define the space

$$\mathcal{H}^s := \{v \in L^2(D) : \|v\|_{\mathcal{H}^s} < +\infty\}$$

with the norm

$$\|v\|_{\mathcal{H}^s} := \Big(\sum_{j=1}^{+\infty} \lambda_j^{2s} |\langle v, e_j\rangle|^2\Big)^{1/2}. \tag{2.3}$$

2.2 The Probabilistic Setting

The above functional analytic setting was dwelling on properties of linear operators in Hilbert spaces. The classical theory of SPDEs builds upon the very same framework, giving conditions to make sense of solutions to PDEs with infinite-dimensional noise terms such as $dW(x, t)$. First of all, we recall that stochastic differential equations have a rigorous meaning only in their integral form, which for (1.2) reads

$$y(s, x, t) = y(s, x, 0) + \int_0^t \mathcal{L}^s y(s, x, \tau)\, d\tau + W(x, t). \tag{2.4}$$

If (2.4) holds $\mathbb{P}$-almost surely, then we call these *strong solutions* to SPDEs; see (3.9). We will use a slightly different notion of solutions here (see Definition 3.6), which is based on the possibility to write y as an infinite sum along the orthonormal basis of the Hilbert space; see (3.3). We discuss the two solutions concepts in Remark 3.7, after having laid out the necessary framework and properties.

Now we introduce the necessary notation and standard assumptions in order to write the noise $W(x, t)$ as an infinite sum of independent and identically distributed Brownian Motions. By $W : \Omega \times [0, T] \to L^2(D)$ we denote a Q-Wiener process with values in $L^2(D)$. The underlying probability space is $(\Omega, \mathcal{F}, \mathbb{P})$, and we assume that the Wiener process is adapted to a normal filtration $\mathcal{F}_t \in \mathcal{F}$. We assume that the covariance operator Q of W is linear, bounded, self-adjoint, positive semidefinite, and that its trace is finite, namely

$$\operatorname{Tr} Q < +\infty. \tag{2.5}$$

This implies that the sum of the eigenvalues μ_j of Q is bounded. Note that a Q-Wiener process in $L^2(D)$ can be approximated in $L^2(\Omega, C([0, T], L^2(D)))$ by a sequence of i.i.d. Brownian motions $\{B_j\}_{j\in\mathbb{N}}$:

$$W(x, t) = \sum_{j=1}^{+\infty} \sqrt{\mu_j}\, e_j(x) B_j(t), \tag{2.6}$$

and by means of an exponential inequality and the Borel–Cantelli lemma, the convergence can be obtained uniformly with probability one. Thus, the sample paths of $W(t)$ belong to $C([0, T], L^2(D))$ almost surely, and we may therefore choose a continuous version.

Note that without the trace-class assumption in (2.5), the sum in (2.6) would not converge in L^2, but only in a larger space. In fact, due to the lack of space regularity of the noise, even the meaning of a simple SPDE such as (1.2) is unclear. Here we will not dwell on weaker notions of solutions, as have been developed in recent years, because we need quite some regularity of solutions to ensure that the optimality conditions can be formulated in a meaningful way.

Therefore, in our discussion we will a priori restrict ourselves to trace-class noises and elliptic operators generating analytic semigroups as in (2.1), which are sufficiently regularizing in order to compete with the roughness of the noise; see Assumptions 3.2 and 3.3 below.

Note that the smaller s, the less $\mathcal{L}$ is regularizing our solution, and the stricter assumptions we need to impose on Q. The exact conditions for the regularity of the solutions depend therefore on the interplay between Q and s, and they are stated in Assumption 3.3.

In the forthcoming analysis, especially Section 3.1, we will make precise statements on the conditions on Q which are necessary to ensure that our solution takes values in the space $\mathcal{H}^s$, which was defined in (2.3). We will also see that due to the influence of the noise, the set of admissible controls differs from the deterministic case.

3 Construction of Solutions

Up to now, the discussion of the linear operator $\mathcal{L}^s$ and the covariance operator Q have been somehow informal since they aimed at expressing the main ideas without going into too technical statements. We now make the above-mentioned approach more precise. For this, as we aim to expand a solution of (1.2) along an orthonormal basis of a Hilbert space, it is convenient to assume that Q has a common set of eigenfunctions with $\mathcal{L}^s$ (and so with $\mathcal{L}$), to which we have already hinted on by using the same notation for the eigenfunctions in (2.6). More explicitly, we suppose the following.

Assumption 3.1. For any $j \in \mathbb{N}$, we have that $Qe_j = \mu_j e_j$ and $\mathcal{L}e_j = \lambda_j e_j$ (and thus $\mathcal{L}^s e_j = \lambda_j^s e_j$). In addition,

$$\sum_{j=1}^{+\infty} \mu_j < +\infty. \tag{3.1}$$

We remark that (3.1) is just a restatement of (2.5). For our purposes, it will also be technically advantageous to avoid operators with zero or negative eigenvalues, in view of compactness and regularity theory. Precisely, from now on we will assume the following.

Assumption 3.2. The eigenvalues of $\mathcal{L}$ satisfy

$$\alpha < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_j \to +\infty \tag{3.2}$$

for some $\alpha > 0$.

The latter is a standard assumption in SPDEs, and it is satisfied for example by the operators $\mathcal{L} = (-\Delta + \alpha)$ with either Neumann or Dirichlet conditions, or $\mathcal{L} = -\Delta$ with Dirichlet conditions.

In the spirit of the spectral definition of the fractional Laplacian in (2.2), we want to find solutions of the state equation (1.2) by approximation with real-valued stochastic processes $y_j(t, s) := \langle y(\cdot, t), e_j\rangle$, where $e_j(x) \in H_0^1(D)$ is an orthonormal basis of $L^2(D)$ built out of eigenfunctions of $\mathcal{L}$. In other words, for fixed s we define the solution of (1.2) as the infinite series

$$y(s)(x, t) = \sum_{j=1}^{+\infty} \langle y(s)(x, t), e_j(x)\rangle e_j(x) = \sum_{j=1}^{+\infty} y_j(s, t) e_j(x). \tag{3.3}$$

Lemma 3.8 will show that this sum is convergent, i.e. (3.3) is well-defined, however, this is not enough to define a solution to (1.2) which exists pathwise, and this property is in turn necessary to show the existence of optimal controls. For this, we need another assumption.

Assumption 3.3. We assume that s is such that

$$\sum_{j=1}^{+\infty} \lambda_j^{-s} < +\infty, \tag{3.4}$$

$$\sum_{j=1}^{+\infty} \mu_j \lambda_j^{s} < +\infty. \tag{3.5}$$

In a sense, Assumption 3.3 is a strengthening of Assumption 3.2 (which is always assumed in the following without further mentioning it).

Definition 3.4. We say that a control $s \in (0, L)$ is *admissible* if it satisfies (3.4) and (3.5). We collect all such s in the set $\mathscr{S}$ and call it the *set of admissible controls*. Moreover, we denote the interior of $\mathscr{S}$ by $\mathscr{S}^\circ$.

Example 3.5. Set $\mathcal{L} = \Delta$ on $(0, \pi)$ with Dirichlet boundary conditions. Then the eigenfunctions read

$$e_j(x) := c_j \sin(jx),$$

and the corresponding eigenvalues are $\lambda_j = j^2$. For (3.4) we get

$$\sum_{j=1}^{+\infty} \lambda_j^{-s} = \sum_{j=1}^{+\infty} j^{-2s},$$

which is convergent for $s > \frac{1}{2}$. As the penalty function Φ is defined for $s \in (0, L)$ (see the problem statement, especially the paragraph around (1.3) for details), we conclude that we can take the set of admissible controls as the interval $\mathscr{S} = (\frac{1}{2}, L)$.

With these preparatory tools at hand, we can finally state our solution concept, and prove the existence of such solutions.

Definition 3.6. We say that $y(s) : \Omega \times D \times [0, T] \to \mathbb{R}$ is an admissible solution to the state equation in (1.2) with initial condition $y_0 \in \mathcal{H}^{s/2}$ if and only if the following conditions are satisfied:

(i) The random variable
$$\omega \mapsto \|y(s, \omega)\|_{L^2([0,T];\mathcal{H}^s)}$$
is almost surely finite for a fixed $s \in \mathscr{S}$.

(ii) For fixed $s \in \mathscr{S}$, we have that
$$y(x, t) = \sum_{j=1}^{+\infty} y_j(t, s) e_j(x),$$
and the stochastic processes $y_j(t, s)$ solve the Itô diffusion equation
$$dy_j(t) = -\lambda_j^s y_j(t)\,dt + \sqrt{\mu_j} dB_j(t), \quad j \in \mathbb{N}, \tag{3.6}$$
or, in integral form,
$$y_j(t) = y_j(0) - \lambda_j^s \int_0^t y_j(\tau)\,d\tau + \sqrt{\mu_j} B_j(t), \quad j \in \mathbb{N}$$

for every $t \in (0, T)$.

Notice that, as $\sqrt{\mu_j}$ and λ_j^s are constant for fixed j and $-\lambda_j^s y_j(t)$ is Lipschitz continuous, for fixed s and for every j the Itô equation (3.6) has a unique strong solution which depends continuously on the initial data, as proved for example on [15, p. 212]. We can explicitly solve (3.6) by applying Itô's formula to $e^{\lambda_j^s t} y_j(t)$ and obtain
$$y_j(t) = y_{j,0} e^{-\lambda_j^s t} + \sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)}\,dB_j(\tau). \tag{3.7}$$

Notation. The stochastic process in (3.7) consists of two parts, a deterministic mean and a random perturbation, which is a stochastic integral. To ease the forthcoming estimates, we will abbreviate the mean part, which is a function of time depending on the parameter s, by $m_j(t, s)$. Moreover, the one-dimensional stochastic integral appears often as a summand in our calculations, and we abbreviate it by $W^j_{\mathcal{L},s}(t)$, indicating that it is a stochastic process involving the (semigroup) of the operator $\mathcal{L}^s$. In formula, we set
$$m_j(t, s) := y_{j,0} e^{-\lambda_j^s t}, \quad W^j_{\mathcal{L},s}(t) := \sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)}\,dB_j(\tau). \tag{3.8}$$

The main part of this section will be dedicated to show that solutions in the sense of Definition 3.6 exist. This is the content of Theorem 3.11, whose proof needs some preparation.

Remark 3.7 (Comparison to Strong Solutions). Our notion of solutions, namely the one in Definition 3.6, resembles the definition of a *strong solution* to SPDEs (see [8]) which reads for (1.2) as
$$y(t) = y_0 + \int_0^t \mathcal{L}^s y(\tau) d\tau + W(t) \quad \mathbb{P}\text{-a.s.} \tag{3.9}$$

Strong solutions are required to be in the domain of the differential operator, which requires
$$y(\cdot\,, t, s) \in L^2(\Omega, \mathscr{H}^s)$$
for any $t \in (0, T]$ and $s \in \mathscr{S}$, which will be shown in Proposition 3.9. Moreover, for strong solutions it is

required that

$$\int_0^T \mathcal{L}^s y(\tau)d\tau < \infty \quad \mathbb{P}\text{-almost surely},$$

which is the statement of condition (i) in Definition 3.6. However, we decided to propose Definition 3.6 as our solution concept as we need the very explicit description of the solution as an infinite series in order to be able to derive concrete optimality conditions and the existence of optimal controls. In this setting, we observe that our solutions in the sense of Definition 3.6 are also strong solutions.

Lemma 3.8. *Let $\mathcal{L}$ satisfy Assumption 3.2 and let Q satisfy* (3.1). *Let the initial data $y_0 \in L^2(D)$ be deterministic. Then the sum appearing in* (3.3) *is convergent in $L^2(\Omega, L^2(D))$, and almost surely in $L^2(D)$, and its limit $y(s)(x,t)$ is an $L^2(D)$-valued adapted stochastic process.*

Proof. We show first that for fixed t the series in (3.3) converges in $L^2(\Omega, L^2(D))$. The sum (3.3) reads formally

$$y(s)(x,t) = \sum_{j=1}^{+\infty} e_j(x) y_{j,0} e^{-\lambda_j^s t} + \sum_{j=1}^{+\infty} e_j(x)\sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau). \tag{3.10}$$

Since $e^{-\lambda_j^s t} \le 1$ for all $s, t > 0$, we get for the first term in (3.10) that

$$\Big\|\sum_{j=1}^{+\infty} e_j y_{j,0} e^{-\lambda_j^s t}\Big\|^2_{L^2(D)} = \sum_{j=1}^{+\infty} |y_{j,0}|^2 e^{-\lambda_j^s t} \le \|y_0\|^2_{L^2(D)}, \tag{3.11}$$

which is finite by assumption.

To show the convergence of the second term in (3.10), we recall from (3.8) the notation

$$W^j_{\mathcal{L},s}(t) := \sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau),$$

and start by looking at the partial sum $\sum_{j=1}^n e_j(x) W^j_{\mathcal{L},s}(t)$. As this partial sum has finitely many summands, we can exchange expectation and summation, use the one-dimensional Itô's Isometry and the lower bound assumption on the eigenvalues (3.2) to obtain

$$\begin{aligned}
\mathbb{E}\Big[\Big\|\sum_{j=1}^n e_j(x) W^j_{\mathcal{L},s}(t)\Big\|^2_{L^2(D)}\Big] &= \sum_{j=1}^n \mathbb{E}\Big[\Big|\sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\Big]\\
&= \sum_{j=1}^n \mu_j \int_0^t e^{-2\lambda_j^s(t-\tau)}\,d\tau\\
&= \sum_{j=1}^n \frac{\mu_j}{2\lambda_j^s}(1 - e^{-2\lambda_j^s t})\\
&\le \frac{1}{2}\sum_{j=1}^n \frac{\mu_j}{\lambda_j^s}\\
&\le \frac{1}{2\alpha^s}\sum_{j=1}^n \mu_j,
\end{aligned} \tag{3.12}$$

which is finite due to (3.1). Similarly, we can calculate for $m > n$,

$$\begin{aligned}
\mathbb{E}\Big[\Big\|\sum_{j=1}^m e_j(x) W^j_{\mathcal{L},s}(t) - \sum_{j=1}^n e_j(x) W^j_{\mathcal{L},s}(t)\Big\|^2_{L^2(D)}\Big] &= \mathbb{E}\Big[\Big\|\sum_{l=n+1}^m e_l(x) W^l_{\mathcal{L},s}(t)\Big\|^2_{L^2(D)}\Big]\\
&= \sum_{l=n+1}^m \frac{\mu_l}{2\lambda_l^s}(1 - e^{-2\lambda_l^s t})\\
&\le \frac{1}{2\alpha^s}\sum_{l=n+1}^m \mu_l,
\end{aligned} \tag{3.13}$$

and it follows that the sequence of partial sums $\sum_{j=1}^{n} e_j(x) W^j_{\mathcal{L},s}(t)$ is a Cauchy sequence for fixed control s and time t.

By (3.10)–(3.13) it follows that the series in (3.3) is convergent in $L^2(\Omega, L^2(D))$. This, by recalling Lemma A.1, gives also that this series is almost surely finite in $L^2(D)$. In addition, from (3.1) and (3.12) we deduce the following boundedness in time:

$$\sup_{t\leq T} \mathbb{E}\Big[\Big\| \sum_{j=n+1}^{+\infty} e_j(x) W^j_{\mathcal{L},s}(t) \Big\|^2_{L^2(D)}\Big] \leq \frac{1}{2\alpha^s} \sum_{l=n+1}^{+\infty} \mu_l \to 0$$

as $n \to +\infty$, and so $y(s)(\cdot, t)$ is an $\mathcal{F}_t$-adapted $L^2(D)$-valued process. □

3.1 Properties of the Solution Which Need Assumption 3.3

As already announced in the beginning of this section, differently from the deterministic case described in [24], an additional assumption is necessary in the stochastic case to ensure the appropriate spatial regularity of the solution. The following proposition will now make transparent why Assumption 3.3 is the appropriate condition for our purposes.

Proposition 3.9. *Let the initial data $y_0 \in L^2(D)$ be deterministic, and let Assumptions 3.1 and 3.3 be satisfied. Then the solution to the state equation* (1.2) *satisfies*

$$y(s, t, \cdot) \in L^2(\Omega, \mathcal{H}^s) \tag{3.14}$$

for any fixed $s \in \mathscr{S}$ and $t \in [0, T]$.

Proof. Recalling (2.3) and (3.8), for fixed s we define

$$\kappa(t) := \sup_{r>0}(r^2 e^{-rt}).$$

Then we have that

$$\Big\| \sum_{j=1}^{+\infty} e_j y_{j,0} e^{-\lambda_j^s t} \Big\|^2_{\mathcal{H}^s} = \sum_{j=1}^{+\infty} \lambda_j^{2s} |y_{j,0}|^2 e^{-2\lambda_j^s t} \leq \kappa(t) \sum_{j=1}^{+\infty} |y_{j,0}|^2 = \kappa(t) \|y_0\|^2_{L^2(D)}. \tag{3.15}$$

Also, we exchange expectation and summation, and we apply Itô's Isometry to get

$$\begin{aligned}
\mathbb{E}\Big[\Big\| \sum_{j=1}^{N} e_j W^j_{\mathcal{L},s}(t) \Big\|^2_{\mathcal{H}^s}\Big] &= \mathbb{E}\Big[\sum_{j=1}^{N} \lambda_j^{2s} |W^j_{\mathcal{L},s}(t)|^2\Big] \\
&= \sum_{j=1}^{N} \lambda_j^{2s} \mu_j \mathbb{E}\Big[\Big(\int_0^t e^{-\lambda_j^s(t-\tau)}\, dB_j(\tau)\Big)^2\Big] \\
&= \sum_{j=1}^{N} \lambda_j^{2s} \mu_j \int_0^t e^{-2\lambda_j^s(t-\tau)}\, d\tau \\
&= \frac{1}{2} \sum_{j=1}^{N} \lambda_j^s \mu_j (1 - e^{-2\lambda_j^s t}) \\
&\leq \frac{1}{2} \sum_{j=1}^{N} \lambda_j^s \mu_j,
\end{aligned} \tag{3.16}$$

which is finite, in light of (3.5) from Assumption 3.3. From (3.15) and (3.16) we obtain (3.14), as desired. □

The next proposition deals with the almost sure finiteness of the integral $\int_0^t \mathcal{L}^s y(s, \tau)\, d\tau$.

Proposition 3.10. *Let Assumptions 3.1 and 3.3 be satisfied and let the initial data y_0 be deterministic, with $y_0 \in \mathcal{H}^{s/2}$. Then the solution to the state equation* (1.2) *satisfies*

$$\|y(s)\|_{L^2(\Omega\times[0,T];\mathcal{H}^s)} \le C \tag{3.17}$$

for some $C > 0$. Moreover, for a fixed $s \in \mathscr{S}$, the random variable

$$\omega \mapsto \|y(s,\omega)\|_{L^2([0,T];\mathcal{H}^s)}$$

is almost surely finite.

Proof. Using (3.7), we have that

$$\|y(s)\|^2_{\mathcal{H}^s} = \sum_{j=1}^{+\infty} \lambda_j^{2s}|y_j(s)|^2 = \sum_{j=1}^{+\infty} \lambda_j^{2s}\Big|y_{j,0}e^{-\lambda_j^s t} + \sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2.$$

Hence, using the standard estimate $(a+b)^2 \le 2(a^2+b^2)$, we conclude that

$$\frac{1}{2}\|y(s)\|^2_{\mathcal{H}^s} \le \sum_{j=1}^{+\infty} \lambda_j^{2s}|y_{j,0}e^{-\lambda_j^s t}|^2 + \sum_{j=1}^{+\infty} \lambda_j^{2s}\mu_j\Big|\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2,$$

and therefore

$$\begin{aligned}\frac{1}{2}\|y(s)\|^2_{L^2(\Omega\times[0,T];\mathcal{H}^s)} &= \frac{1}{2}\mathbb{E}\Big[\int_0^T \|y(s)\|^2_{\mathcal{H}^s}\,dt\Big]\\ &\le \int_0^T \sum_{j=1}^{+\infty} \lambda_j^{2s}|y_{j,0}e^{-\lambda_j^s t}|^2\,dt + \mathbb{E}\Big[\int_0^T \sum_{j=1}^{+\infty} \lambda_j^{2s}\mu_j\Big|\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\,dt\Big].\end{aligned} \tag{3.18}$$

Now we analyze the right-hand side of (3.18) term by term. First of all,

$$\begin{aligned}\int_0^T \sum_{j=1}^{+\infty} \lambda_j^{2s}|y_{j,0}e^{-\lambda_j^s t}|^2\,dt &= \int_0^T \sum_{j=1}^{+\infty} \lambda_j^{2s}|y_{j,0}|^2 e^{-2\lambda_j^s t}\,dt\\ &= \sum_{j=1}^{+\infty} \lambda_j^{2s}|y_{j,0}|^2 \frac{1-e^{-2\lambda_j^s T}}{2\lambda_j^s}\\ &\le \frac{1}{2}\sum_{j=1}^{+\infty} \lambda_j^{s}|y_{j,0}|^2\\ &= \frac{1}{2}\|y_0\|^2_{\mathcal{H}^{s/2}}.\end{aligned} \tag{3.19}$$

Furthermore, by Itô's Isometry,

$$\mathbb{E}\Big[\Big|\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\Big] = \mathbb{E}\Big[\int_0^t e^{-2\lambda_j^s(t-\tau)}\,d\tau\Big] = \frac{1-e^{-2\lambda_j^s t}}{2\lambda_j^s} \le \frac{1}{2\lambda_j^s},$$

and therefore, for any fixed $N \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}\Big[\int_0^T \sum_{j=1}^{N} \lambda_j^{2s}\mu_j\Big|\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\,dt\Big] &= \int_0^T \sum_{j=1}^{N} \lambda_j^{2s}\mu_j\mathbb{E}\Big[\Big|\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\Big]\,dt\\ &\le \frac{1}{2}\int_0^T \sum_{j=1}^{N} \lambda_j^{s}\mu_j\,dt\\ &= \frac{T}{2}\sum_{j=1}^{N} \lambda_j^{s}\mu_j\\ &\le c(T,s)\end{aligned}$$

for some $c(T,s) > 0$, thanks to Assumption 3.3. This and Fatou's Lemma imply that

$$\mathbb{E}\Big[\int_0^T \sum_{j=1}^{+\infty} \lambda_j^{2s}\mu_j \Big|\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\,dt\Big] \le c(T,s). \tag{3.20}$$

Then, combining (3.18)–(3.20), we complete the proof of (3.17).

Finally, the almost sure statement in Proposition 3.10 follows from (3.17) and Lemma A.1. □

Theorem 3.11. *Let Assumptions 3.1 and 3.3 be satisfied. Let the initial data y_0 be deterministic, with $y_0 \in \mathcal{H}^{s/2}$. Then, for every $s \in \mathscr{S}$, there exists a unique solution $y = y(s)$ to the state system* (1.2) *in the sense of Definition 3.6.*

Proof. Condition (i) in Definition 3.6 was verified in Proposition 3.10. To fulfill condition (ii) of Definition 3.6, we choose a deterministic initial condition $y_{j,0} = \langle y_0(x), e_j(x)\rangle \in \mathbb{R}$, and employ the series approximation of the Q-Wiener process (2.6) to get the infinite system of Itô equations

$$dy_j(t) = -\lambda_j^s y_j(t)\,dt + \sqrt{\mu_j}\,dB_j(t) \quad \text{for } j \in \mathbb{N}. \tag{3.21}$$

As $\sqrt{\mu_j}$ and λ_j^s are constant for fixed j, and $-\lambda_j^s y_j(t)$ is Lipschitz continuous for fixed s, each Itô equation in (3.21) has a unique strong solution $y_j(t,s)$, which depends continuously on the initial data, as was proved for example on [15, p. 212].

Lemma 3.8 shows that the sum $y(x,t) = \sum_{j=1}^{+\infty} y_j(t,s)e_j(x)$ is convergent, and its limit $y(s)(x,t)$ is an $L^2(D)$-valued adapted stochastic process, which concludes the proof. □

3.2 Further Space-Time Regularity and Hölder Continuity

In this section, we prove further properties of solutions to the state system (1.2), which we will need in Section 5.

Proposition 3.12. *Let $\mathcal{L}$ satisfy Assumption 3.2 and let Q satisfy* (3.1). *Let $y_0 \in L^2(D)$ be deterministic. Then any solution $y = y(s)$ to the state equation* (1.2) *satisfies the estimate*

$$\|y(s)\|_{L^2(\Omega,L^2(D\times[0,T]))} \le C. \tag{3.22}$$

Moreover, for a fixed $s \in (0,+\infty)$, the random variable

$$\omega \mapsto \|y(s,\omega)\|_{L^2(D\times[0,T])}$$

is almost surely finite.

Proof. By (3.3) and (3.7), we know that

$$\begin{aligned}
\|y(s)\|^2_{L^2(D)} &= \sum_{j=1}^{+\infty} |y_j(s)|^2 \\
&= \sum_{j=1}^{+\infty} \Big| y_{j,0}(s)e^{-\lambda_j^s t} + \sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2 \\
&\le 2\Big[\sum_{j=1}^{+\infty} |y_{j,0}(s)e^{-\lambda_j^s t}|^2 + \sum_{j=1}^{+\infty} \Big|\sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\Big],
\end{aligned}$$

and therefore

$$\begin{aligned}
\|y(s)\|^2_{L^2(\Omega,L^2(D\times[0,T]))} &= \mathbb{E}\Big[\int_0^T \|y(s)\|^2_{L^2(D)}\,dt\Big] \\
&\le 2\Big\{\int_0^T \sum_{j=1}^{+\infty} |y_{0,j}(s)e^{-\lambda_j^s t}|^2 + \mathbb{E}\Big[\int_0^T \sum_{j=1}^{+\infty} \Big|\sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}\,dB_j(\tau)\Big|^2\,dt\Big]\Big\}.
\end{aligned} \tag{3.23}$$

For the first term in the right-hand side of (3.23) we calculate, using Assumption 3.2, that

$$\int_0^T \sum_{j=1}^{+\infty} |y_{0,j}(s) e^{-\lambda_j^s t}|^2 \, dt = \frac{1}{2} \sum_{j=1}^{+\infty} |y_{0,j}(s)|^2 \frac{1}{2\lambda_j^s} (1 - e^{-2\lambda_j^s T}) \le \frac{1}{2\alpha^s} \|y_0\|_{L^2}^2. \tag{3.24}$$

For the second term in the right-hand side of (3.23) we use Itô's Isometry to get

$$\mathbb{E}\Big[\Big(\int_0^t e^{-\lambda_j^s (t-\tau)} \, dB_j(\tau)\Big)^2\Big] = \mathbb{E}\Big[\int_0^t e^{-2\lambda_j^s (t-\tau)} \, d\tau\Big] = \frac{1 - e^{-2\lambda_j^s t}}{2\lambda_j^s}.$$

Hence, we consider the partial sum $j \le N$, due to which we can exchange expectation and summation, and conclude that

$$\begin{aligned}
\mathbb{E}\Big[\int_0^T \sum_{j=1}^{N} \Big|\sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)} \, dB_j(\tau)\Big|^2 \, dt\Big] &= \int_0^T \sum_{j=1}^{N} \mu_j \mathbb{E}\Big[\Big(\int_0^t e^{-\lambda_j^s (t-\tau)} \, dB_j(\tau)\Big)^2\Big] \, dt \\
&= \int_0^T \sum_{j=1}^{N} \frac{\mu_j}{2\lambda_j^s} (1 - e^{-2\lambda_j^s t}) \, dt \\
&\le \int_0^T \sum_{j=1}^{N} \frac{\mu_j}{2\lambda_j^s} \, dt \\
&\le c(T, s)
\end{aligned} \tag{3.25}$$

for some $c(T, s) > 0$, thanks to Assumption 3.2 and the trace-class property of the noise in (3.1). Then, combining (3.24) and (3.25) and applying Fatou's Lemma, we obtain (3.22), as desired.

Then from (3.22) and Lemma A.1 we also obtain that

$$\omega \mapsto \|y(s, \omega)\|_{L^2(D \times [0,T]))}$$

is almost surely finite. □

Note that for Proposition 3.12 $y(s)(x, t)$ is only required to be an $L^2(D)$-valued adapted stochastic process, as proved in Lemma 3.8. The proof used only $L^2(\Omega, L^2(D \times [0, T]))$-norms, no additional $\mathcal{H}^s$-regularity is needed. Therefore, Assumption 3.3 is not needed in Proposition 3.12.

To ensure sufficient compactness properties needed to prove the existence of optimal controls in Section 5, we need to quantify the Hölder continuity in time of solutions to equation (1.2) in dependence of s. This is proved via the factorization method (see [8, Chapter II.5.3]), which uses the semigroup generated by $\mathcal{L}^s$ and works with interpolation spaces.

Lemma 3.13. *Let the initial data $y_0 \in L^2(D)$ be deterministic, and let Assumptions 3.1 and 3.3 be satisfied. Then the sample paths of the process $y(s)(x, t)$ are in $C^\delta([0, T], L^2(D))$ for arbitrary $\delta \in (0, \frac{1}{2})$.*

Proof. It suffices to verify that the trajectories of the stochastic convolution are δ-Hölder continuous. According to [8, Theorem 5.15], this holds with $\delta \in (0, \beta - \epsilon)$ if the following condition on the Hilbert–Schmidt norm of $S(t)Q$, where $S(t)$ is the semigroup generated by $\mathcal{L}^s$ (see (2.1)), is satisfied:

$$\int_0^T t^{-2\beta} \|S(t)Q\|_{HS}^2 \, dt < +\infty, \tag{3.26}$$

where

$$\|P\|_{HS}^2 := \sqrt{\sum_{j=1}^{+\infty} \|Pe_j\|_{L^2(D)}^2}.$$

We observe that, in light of Assumption 3.1 and (2.1),

$$\|S(t)Q\|_{HS}^2=\sum_{j=1}^{+\infty}\|S(t)Qe_j\|_{L^2(D)}^2=\sum_{j=1}^{+\infty}\mu_j^2\|S(t)e_j\|_{L^2(D)}^2=\sum_{j=1}^{+\infty}\mu_j^2\|e^{-\lambda_j t}e_j\|_{L^2(D)}^2=\sum_{j=1}^{+\infty}\mu_j^2e^{-2\lambda_j t}.$$

Notice also that μ_j is a bounded sequence due to (3.1). Therefore,

$$\|S(t)Q\|_{HS}^2\le\sum_{j=1}^{+\infty}\mu_j^2\le\sup_{j\in\mathbb{N}}\mu_j\sum_{j=1}^{+\infty}\mu_j,$$

which is finite, again by virtue of (3.1).

This gives that (3.26) is verified when $\beta<\frac{1}{2}$, which proves the desired result. □

4 Differentiability of the Control-to-State Operator

In this section, we prepare the way to formulate the optimality conditions. For this it is necessary to look at the partial derivative of solutions to (1.2) in the control variable s. Due to the need for ω-wise (or "pathwise") definitions of such objects, we first prove a preliminary result on Wiener Integrals.

4.1 A Property of the Wiener Integral

In general, the stochastic integral is a random variable, which does not a priori make sense pathwise: the integrator, in our case Brownian Motion, is not of bounded variation, and therefore the stochastic integral enjoys much weaker properties than a Riemann–Stieltjes-Integral.

In this section, we take a look at the stochastic integrals appearing in our analysis. First of all, note that the integrands are deterministic, thus providing a special case that is called "Wiener Integrals". Moreover, due to the regularity of the integrands, which are of the form $\exp(-\lambda_j^s\tau)$, we are able to give conditions on when the operator $\frac{d}{ds}$ applied to the Wiener Integral is well-defined as an s-dependent random variable.

Lemma 4.1. *For any $j\in\mathbb{N}$, let $g_j:\mathscr{S}\times[0,T]$. Assume that, for any $t\in[0,T]$, the map $\mathscr{S}\ni s\mapsto g_j(t,s)$ is C^2. Let*

$$G_j(t,s):=\int_0^t g_j(\tau,s)\,dB_j(\tau)$$

and

$$H_j(t,s):=\int_0^t \partial_s g_j(\tau,s)\,dB_j(\tau).$$

Assume that, for any $j\in\mathbb{N}$ and $s\in\mathscr{S}$,

$$|g_j(t,s)|+|\partial_s g_j(t,s)|+|\partial_s^2 g_j(t,s)|\le C(s)\sqrt{\mu_j}\,\Gamma(t),\tag{4.1}$$

with $C(s)\in(0,+\infty)$ for any fixed s,

$$M(s):=\sup_{\sigma\in(s/2,2s)}C(\sigma)<+\infty\quad\text{for any fixed } s\in\mathscr{S},\tag{4.2}$$

and

$$\Gamma\in L^2([0,T],[0,+\infty)).\tag{4.3}$$

Then

$$\partial_s\sum_{j=1}^{+\infty}G_j(t,s)e_j(x)=\sum_{j=1}^{+\infty}\partial_s G_j(t,s)e_j(x)=\sum_{j=1}^{+\infty}H_j(t,s)e_j(x)\tag{4.4}$$

as functions in $L^2(\Omega,L^2(D\times[0,T]))$.

Proof. First of all, we check that

$$\sum_{j=1}^{+\infty} G_j(t,s)e_j(x) \in L^2(\Omega, L^2(D \times [0,T])). \tag{4.5}$$

To this end, we exploit (4.1) and (4.2) to see that

$$|G_j(t,s)| = \Big|\int_0^t g_j(\tau,s)\,dB_j(\tau)\Big| \le C(s)\sqrt{\mu_j}\int_0^t \Gamma(\tau)\,dB_j(\tau).$$

Then, making use of (4.3) together with Itô's Isometry, we see that, for every $M \ge N \in \mathbb{N}$,

$$\begin{aligned}
\Big\|\sum_{j=N}^{M} G_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,L^2(D\times[0,T]))} &= \mathbb{E}\Big[\int_0^T \Big\|\sum_{j=N}^{M} G_j(t,s)e_j(x)\Big\|^2_{L^2(D)}\,dt\Big]\\
&= \mathbb{E}\Big[\int_0^T \sum_{j=N}^{M} |G_j(t,s)|^2\,dt\Big]\\
&\le C^2(s)\sum_{j=N}^{M}\mu_j \int_0^T \mathbb{E}\Big[\Big|\int_0^t \Gamma(\tau)\,dB_j(\tau)\Big|^2\Big]\,dt\\
&\le C^2(s)\sum_{j=N}^{M}\mu_j \int_0^T \mathbb{E}\Big[\int_0^t \Gamma^2(\tau)\,d\tau\Big]\,dt\\
&\le C(s,T)\sum_{j=N}^{M}\mu_j.
\end{aligned}$$

Notice that the latter quantity is infinitesimal for large N and M thanks to (3.1), and therefore the series in (4.5) produces a Cauchy sequence, thus proving (4.5).

Fixed $j \in \mathbb{N}$, $t \in [0,T]$, $s \in \mathscr{S}$ and $h \in (-1,1)$ (to be taken sufficiently small), we notice that

$$\begin{aligned}
|G_j(t,s+h) - G_j(t,s) - hH_j(t,s)| &= \Big|\int_0^t (g_j(\tau,s+h) - g_j(\tau,s) - h\partial_s g_j(\tau,s))\,dB_j(\tau)\Big|\\
&= \Big|\int_0^t \Big(\int_0^h \partial_s g_j(\tau,s+\sigma)\,d\sigma - h\partial_s g_j(\tau,s)\Big)\,dB_j(\tau)\Big|\\
&= \Big|\int_0^t \Big(\int_0^h (\partial_s g_j(\tau,s+\sigma) - \partial_s g_j(\tau,s))\,d\sigma\Big)\,dB_j(\tau)\Big|\\
&= \Big|\int_0^t \Big(\int_0^h \Big(\int_0^\sigma \partial_s^2 g_j(\tau,s+\rho)\,d\rho\Big)\,d\sigma\Big)\,dB_j(\tau)\Big|\\
&\le \sqrt{\mu_j}\int_0^t \Big(\int_0^h \Big(\int_0^\sigma C(s+\rho)\Gamma(\tau)\,d\rho\Big)\,d\sigma\Big)\,dB_j(\tau)\\
&\le M(s)\sqrt{\mu_j}\int_0^t \Big(\int_0^h \Big(\int_0^\sigma \Gamma(\tau)\,d\rho\Big)\,d\sigma\Big)\,dB_j(\tau)\\
&= \frac{M(s)\sqrt{\mu_j}\,h^2}{2}\int_0^t \Gamma(\tau)\,dB_j(\tau)
\end{aligned}$$

as long as h is small enough, thanks to (4.1) and (4.2). As a consequence,

$$\Big|\frac{G_j(t,s+h)-G_j(t,s)}{h}-H_j(t,s)\Big|^2 \le \frac{M^2(s)\mu_j h^2}{4}\Big(\int_0^t \Gamma(\tau)\,dB_j(\tau)\Big)^2.$$

This, Itô's Isometry and (4.3) lead to

$$\begin{aligned}\mathbb{E}\Big[\Big|\frac{G_j(t,s+h)-G_j(t,s)}{h}-H_j(t,s)\Big|^2\Big] &\le \frac{M^2(s)\mu_j h^2}{4}\mathbb{E}\Big[\Big(\int_0^t \Gamma(\tau)\,dB_j(\tau)\Big)^2\Big]\\ &= \frac{M^2(s)\mu_j h^2}{4}\mathbb{E}\Big[\int_0^t \Gamma^2(\tau)\,d\tau\Big]\\ &\le C(s,T)\mu_j h^2\end{aligned}$$

for some $C(s,T)>0$. This estimate and Fatou's Lemma give that

$$\mathbb{E}[|\partial_s G_j(t,s)-H_j(t,s)|^2] \le \lim_{h\to 0}\mathbb{E}\Big[\Big|\frac{G_j(t,s+h)-G_j(t,s)}{h}-H_j(t,s)\Big|^2\Big] \le 0.$$

Consequently, for any $N\in\mathbb{N}$,

$$\mathbb{E}\Big[\int_0^T \sum_{j=1}^N |\partial_s G_j(t,s)-H_j(t,s)|^2\,dt\Big]=0.$$

Thus, using again Fatou's Lemma, we obtain

$$\begin{aligned}0 &= \lim_{N\to+\infty}\mathbb{E}\Big[\int_0^T \sum_{j=1}^N |\partial_s G_j(t,s)-H_j(t,s)|^2\,dt\Big]\\ &\ge \mathbb{E}\Big[\int_0^T \sum_{j=1}^{+\infty} |\partial_s G_j(t,s)-H_j(t,s)|^2\,dt\Big]\\ &= \mathbb{E}\Big[\int_0^T \Big\|\sum_{j=1}^{+\infty}(\partial_s G_j(t,s)-H_j(t,s))e_j(x)\Big\|^2_{L^2(D)}\,dt\Big]\\ &= \Big\|\sum_{j=1}^{+\infty}(\partial_s G_j(t,s)-H_j(t,s))e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])},\end{aligned}$$

and accordingly

$$\sum_{j=1}^{+\infty}\partial_s G_j(t,s)e_j(x)=\sum_{j=1}^{+\infty}H_j(t,s)e_j(x) \tag{4.6}$$

in $L^2(\Omega, D\times[0,T])$.

Now we observe that, for any $N\le M\in\mathbb{N}$,

$$\begin{aligned}&\Big\|\sum_{j=N}^M \frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)\Big\|^2_{L^2(D)}+\Big\|\sum_{j=N}^M H_j(t,s)e_j(x)\Big\|^2_{L^2(D)}\\ &\quad= \sum_{j=N}^M \frac{|G_j(t,s+h)-G_j(t,s)|^2}{h^2}+\sum_{j=N}^M |H_j(t,s)|^2\\ &\quad= \sum_{j=N}^M \frac{1}{h^2}\Big|\int_0^t (g_j(\tau,s+h)-g_j(\tau,s))\,dB_j(\tau)\Big|^2+\sum_{j=N}^M\Big|\int_0^t \partial_s g_j(\tau,s)\,dB_j(\tau)\Big|^2\\ &\quad\le M^2(s)\sum_{j=N}^M \mu_j\Big|\int_0^t \Gamma(\tau)\,dB_j(\tau)\Big|^2+C^2(s)\sum_{j=N}^M \mu_j\Big|\int_0^t \Gamma(\tau)\,dB_j(\tau)\Big|^2,\end{aligned}$$

in view of (4.1) and (4.2). Using Itô's Isometry and (4.3), we thereby find that

$$\begin{aligned}
&\mathbb{E}\Big[\Big\|\sum_{j=N}^{M}\frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)\Big\|^2_{L^2(D)}+\Big\|\sum_{j=N}^{M}H_j(t,s)e_j(x)\Big\|^2_{L^2(D)}\Big]\\
&\quad\le M^2(s)\Big\{\mathbb{E}\Big[\sum_{j=N}^{M}\mu_j\Big|\int_0^t\Gamma(\tau)\,dB_j(\tau)\Big|^2\Big]+\mathbb{E}\Big[\sum_{j=N}^{M}\mu_j\Big|\int_0^t\Gamma(\tau)\,dB_j(\tau)\Big|^2\Big]\Big\}\\
&\quad= M^2(s)\Big\{\sum_{j=N}^{M}\mu_j\mathbb{E}\Big[\int_0^t\Gamma^2(\tau)\,d\tau\Big]+\sum_{j=N}^{M}\mu_j\mathbb{E}\Big[\int_0^t\Gamma^2(\tau)\,d\tau\Big]\Big\}\\
&\quad\le C(s,T)\sum_{j=N}^{M}\mu_j
\end{aligned}$$

for some $C(s,T)>0$. Therefore, recalling (3.1), we have that for any $\epsilon>0$ there exists $N_{\epsilon,s,T}\in\mathbb{N}$ such that if $M\ge N\ge N_{\epsilon,s,T}$, we have that

$$\Big\|\sum_{j=N}^{M}\frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}+\Big\|\sum_{j=N}^{M}H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}\le\epsilon,$$

and so, by Fatou's Lemma,

$$\Big\|\sum_{j=N_{\epsilon,s,T}+1}^{+\infty}\frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}+\Big\|\sum_{j=N_{\epsilon,s,T}+1}^{+\infty}H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}\le\epsilon.$$

As a consequence, using Itô's Isometry once again, we obtain

$$\begin{aligned}
&\frac12\Big\|\sum_{j=1}^{+\infty}\frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)-\sum_{j=1}^{+\infty}H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}\\
&\quad\le\Big\|\sum_{j=1}^{N_{\epsilon,s,T}}\frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)-\sum_{j=1}^{N_{\epsilon,s,T}}H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}+\epsilon\\
&\quad=\Big\|\sum_{j=1}^{N_{\epsilon,s,T}}\int_0^t\Big(\frac{g_j(\tau,s+h)-g_j(\tau,s)}{h}-\partial_s g_j(\tau,s)\Big)\,dB_j(\tau)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}+\epsilon\\
&\quad=\mathbb{E}\Big[\int_0^T\sum_{j=1}^{N_{\epsilon,s,T}}\Big|\int_0^t\Big(\frac{g_j(\tau,s+h)-g_j(\tau,s)}{h}-\partial_s g_j(\tau,s)\Big)\,dB_j(\tau)\Big|^2\,dt\Big]+\epsilon\\
&\quad=\sum_{j=1}^{N_{\epsilon,s,T}}\int_0^T\mathbb{E}\Big[\int_0^t\Big(\frac{g_j(\tau,s+h)-g_j(\tau,s)}{h}-\partial_s g_j(\tau,s)\Big)^2\,d\tau\Big]\,dt+\epsilon.
\end{aligned}$$

Hence, by (4.1)–(4.3),

$$\begin{aligned}
&\frac12\Big\|\sum_{j=1}^{+\infty}\frac{G_j(t,s+h)-G_j(t,s)}{h}e_j(x)-\sum_{j=1}^{+\infty}H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}\\
&\quad\le M^2(s)h^2\sum_{j=1}^{N_{\epsilon,s,T}}\mu_j\int_0^T\mathbb{E}\Big[\int_0^t\Gamma^2(\tau)\,d\tau\Big]\,dt+\epsilon\\
&\quad\le C(s,T)h^2\sum_{j=1}^{N_{\epsilon,s,T}}\mu_j+\epsilon.
\end{aligned}$$

This, (4.5) and Fatou's Lemma yield that

$$\begin{aligned}\epsilon &= \lim_{h\to 0} C(s,T)h^2 \sum_{j=1}^{N_{\epsilon,s,T}} \mu_j + \epsilon \\ &\geq \frac{1}{2}\lim_{h\to 0}\Big\| \sum_{j=1}^{+\infty} \frac{G_j(t,s+h)-G_j(t,s)}{h} e_j(x) - \sum_{j=1}^{+\infty} H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])} \\ &= \frac{1}{2}\lim_{h\to 0}\Big\| \frac{1}{h}\Big(\sum_{j=1}^{+\infty} G_j(t,s+h)e_j(x) - \sum_{j=1}^{+\infty} G_j(t,s)e_j(x)\Big) - \sum_{j=1}^{+\infty} H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])} \\ &\geq \frac{1}{2}\Big\| \lim_{h\to 0}\frac{1}{h}\Big(\sum_{j=1}^{+\infty} G_j(t,s+h)e_j(x) - \sum_{j=1}^{+\infty} G_j(t,s)e_j(x)\Big) - \sum_{j=1}^{+\infty} H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])} \\ &= \frac{1}{2}\Big\| \partial_s \sum_{j=1}^{+\infty} G_j(t,s)e_j(x) - \sum_{j=1}^{+\infty} H_j(t,s)e_j(x)\Big\|^2_{L^2(\Omega,D\times[0,T])}.\end{aligned}$$

Since ϵ can be taken arbitrarily small, we thereby conclude that

$$\partial_s \sum_{j=1}^{+\infty} G_j(t,s)e_j(x) = \sum_{j=1}^{+\infty} H_j(t,s)e_j(x)$$

in $L^2(\Omega, D\times[0,T])$. This and (4.6) give (4.4), as desired. □

Next, we recall [24, Lemma 2.2], which is an auxiliary result on the derivatives of a function of exponential type.

Lemma 4.2. *Define for fixed $\lambda > 0$ and $t > 0$ the real-valued function*

$$E_{\lambda,t}(s) := e^{-\lambda^s t} \quad \text{for } s > 0. \tag{4.7}$$

Then there exist constants $C_i > 0$ such that, for all $\lambda > 0$, $t \in (0,T]$ and $s > 0$, we have that

$$|E_{\lambda,t}(s)| \leq C_0$$

and

$$\Big|\frac{d^k}{ds^k} E_{\lambda,t}(s)\Big| \leq \frac{C_k}{s^k}(1+|\ln(t)|^k), \quad \text{for all } 1 \leq k \leq 4. \tag{4.8}$$

Using Lemmata 4.1 and 4.2, we can now take into account the first and second derivatives of the solutions with respect to the fractional parameter s, according to the following result.

Proposition 4.3. *Let $\mathcal{L}$ satisfy Assumption 3.2 and let Q satisfy* (3.1). *Let the initial data $y_0 \in L^2(D)$ be deterministic. Then*

$$\partial_s y(s) = \sum_{j=1}^{+\infty} \partial_s y_j(\cdot, s)e_j \quad \text{and} \quad \partial^2_{ss} y(s) = \sum_{j=1}^{+\infty} \partial_{ss} y_j(\cdot, s)e_j \tag{4.9}$$

are functions in $L^2(\Omega, L^2(D\times[0,T]))$.

Moreover, for a fixed $s \in (0,+\infty)$, the random variables

$$\omega \mapsto \|\partial_s y(s,\omega)\|_{L^2(D\times[0,T])} \quad \text{and} \quad \omega \mapsto \|\partial_{ss} y(s,\omega)\|_{L^2(D\times[0,T])}$$

are almost surely finite.

Proof. From (3.7) and (4.7) we know that

$$y_j(t,s) = y_{j,0}E_{\lambda_j,t} + \sqrt{\mu_j}\int_0^t E_{\lambda_j,t-\tau}\, dB_j(\tau). \tag{4.10}$$

Now we exploit Lemma 4.1, used here with $g_j := \sqrt{\mu_j}E_{\lambda_j,t}$, in the case of the first derivative, and

$$g_j := \sqrt{\mu_j}\frac{dE_{\lambda_j,t}}{ds}$$

in the case of the second derivative: in this setting, in light of (4.8) we can take $C(s) := C(\frac{1}{s} + \frac{1}{s^4})$, with $C > 0$ and $\Gamma(t) := 1 + |\ln t|^4$ in Lemma 4.1, and then assumptions (4.1)–(4.3) are satisfied.

Accordingly, from (4.10) and Lemma 4.1 we obtain (4.9), as desired.

Then, by Lemma A.1, we conclude that the first and second derivatives of the solution with respect to s are almost surely finite in $L^2(D \times [0, T])$. □

Note that for Proposition 4.3 the function $y(s)(x, t)$ is only required to be an $L^2(D)$-valued adapted stochastic process, as proved in Lemma 3.8. The proof used only $L^2(\Omega, L^2(D \times [0, T]))$-norms, no additional $\mathcal{H}^s$-regularity is needed. Therefore, Assumption 3.3 is not needed in Proposition 4.3.

4.2 Optimality Conditions

In this section, we establish first-order necessary conditions and sufficient optimality conditions of optimal controls.

Theorem 4.4. *Let $y_0 \in L^2(D)$ be deterministic, and let $y = y(s)$ be a solution to the state equation* (1.2) *in the sense of the $L^2(D)$-valued stochastic process $y(s) : \Omega \times [0, T] \to L^2(D)$ of Lemma 3.8. Then the following holds true for a fixed realization $\omega \in \Omega$:*

(i) Necessary condition: *If $\bar{s} = \bar{s}(\omega)$ is an optimal parameter for* (IP) *and $y(\bar{s})$ is the associated unique solution to the state system* (1.2), *then for almost every $\omega \in \Omega$,*

$$\int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s})\,dx\,dt + \Phi'(\bar{s}) = 0. \tag{4.11}$$

(ii) Sufficient condition: *If $\bar{s} = \bar{s}(\omega) \in (0, L)$ satisfies the necessary condition* (4.11), *and if in addition*

$$\int_0^T \int_D (\partial_s y(\bar{s}))^2 + (y(\bar{s}) - y_D)\partial^2_{ss} y(\bar{s})\,dx\,dt + \Phi''(\bar{s}) > 0$$

for almost every $\omega \in \Omega$, then $\bar{s}$ is optimal for (IP).

Proof. By Proposition 4.3, the map

$$s \mapsto \mathcal{J}(s) := \mathcal{J}(y(s), s)$$

is twice differentiable on $(0, +\infty)$. By the chain rule,

$$\begin{aligned}\mathcal{J}'(\bar{s}) &= \frac{d}{ds}\mathcal{J}(y(\bar{s}), \bar{s})\\ &= \partial_y \mathcal{J}(y(\bar{s}), \bar{s}) \circ \partial_s y(\bar{s}) + \partial_s \mathcal{J}(y(\bar{s}), \bar{s})\\ &= \int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s})\,dx\,dt + \Phi'(\bar{s}),\end{aligned}$$

and assertion (i) follows. Also, assertion (ii) is a consequence of the following computation:

$$\begin{aligned}\mathcal{J}''(\bar{s}) &= \frac{d}{ds}\mathcal{J}(y(\bar{s}), \bar{s})\\ &= \partial_y \mathcal{J}(y(\bar{s}), \bar{s}) \circ \partial_s y(\bar{s}) + \partial_s \mathcal{J}(y(\bar{s}), \bar{s})\\ &= \int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s})\,dx\,dt + \Phi'(\bar{s}).\end{aligned}$$

The proof of Theorem 4.4 is thus complete. □

5 Existence of Optimal Controls

The existence of pathwise optimal controls is shown by checking that, for fixed $\omega \in \Omega$, there exists a subsequence $y(s_k)$ which strongly converges to the optimal y in $L^2(D \times [0, T])$.

To show the strong convergence, we use a compactness result, which proves that under certain assumptions solutions enjoy a suitable Hölder regularity in time which is independent of the fractional exponent.

Lemma 5.1 (Compactness Lemma). *Let the initial data y_0 be deterministic, with $y_0 \in \mathcal{H}^{s/2}$. Let Assumptions 3.1 and 3.3 be satisfied.*

Then, for a fixed realization $\omega \in \Omega$, the sequence $\{y_{s_k}(\omega)\}_{k\in\mathbb{N}}$ of solutions to the state equation (1.2) *with initial datum y_0 contains a subsequence that converges strongly in $L^2(D \times [0, T])$.*

Proof. Recall that for solutions of (1.2) in the sense of Definition 3.6 we know the following:

(i) By Proposition 3.10, for all $s_k \in \mathscr{S}$ and almost every $\omega \in \Omega$,

$$\sup_k (\|y_{s_k}(\omega)\|_{L^2([0,T],\mathcal{H}^{s_k})}) < +\infty.$$

(ii) By Proposition 3.12, for all $s_k \in (0, L)$ and for almost every $\omega \in \Omega$,

$$\sup_k (\|y_{s_k}(\omega)\|_{L^2(D\times[0,T])}) < +\infty.$$

(iii) By Lemma 3.13, the trajectories of the family of stochastic processes $y_{s_k}(t)$ are in $C^{\delta_k}([0, T], L^2(D))$ for every k and $\delta_k \geq \delta_* \geq \delta_0 > 0$.

Therefore, we know that y_{s_k} is a sequence (in k) of $L^2(D)$-valued stochastic processes (in (x, t)) with δ-Hölder continuous sample paths and $y_{s_k}(\omega) \in L^2([0, T], \mathcal{H}^{s_k})$ for fixed $\omega \in \Omega$. Notice that, by (iii),

$$C \geq \|y_{s_k}(t)\|^2_{L^2(D)} = \sum_{i=1}^{+\infty} |y_{s_k,i}(t)|^2$$

and

$$C|t - t'|^{\delta_k} \geq \|y_{s_k}(t) - y_{s_k}(t')\|^2_{L^2(D)} = \sum_{i=1}^{+\infty} |y_{s_k,i}(t) - y_{s_k,i}(t')|^2,$$

and so the infinite string $(\{y_{s_k,1}\}_{k\in\mathbb{N}}, \{y_{s_k,2}\}_{k\in\mathbb{N}}, \dots)$ lies in the space

$$C^{\delta_0}([0, T]) \times C^{\delta_0}([0, T]) \times \cdots .$$

Hence, there exists a subsequence denoted by $(s_k)_m$ which converges in this product space to an infinite string of the form $((y^*_s)_1, (y^*_s)_2, \dots)$, and every $(y^*_s)_j$ is in $C^{\delta_0}([0, T])$. We define

$$y^*(x, t) = \sum_{j\in\mathbb{N}} y^*_j e_j(x).$$

The convergence of $y_{(s_k)_m} \to y^*$ follows exactly as in the compactness lemma in the deterministic case, which is [24, Lemma 6.1], by using also (i) and (ii). The details are therefore omitted. □

Theorem 5.2. *Let the initial data y_0 be deterministic, and let Assumptions 3.1 and 3.3 be satisfied. Moreover, let the initial data satisfy*

$$\sup_{s\in\mathscr{S}} \|y_0\|_{\mathcal{H}^s} < +\infty. \tag{5.1}$$

Then, for almost every fixed $\omega \in \Omega$, the functional $\mathcal{J}(\omega)$ attains a minimum in $\mathscr{S}^\circ$, and moreover

$$\inf_{s\in\mathscr{S}} \mathcal{J}(\omega) < +\infty.$$

Proof. Note first that, by our assumptions on $\Phi(s)$, we can find $s^* \in \mathscr{S}^\circ$ such that $\mathcal{J}(s^*, \omega) < +\infty$, and, in view of (1.3), we infer that

$$0 < \inf_{s\in\mathscr{S}^\circ} \mathcal{J}(s, \omega) < +\infty \quad \text{for any fixed } \omega \in \Omega.$$

We pick a minimizing sequence $\{s_k\}_{k\in\mathbb{N}} \subset \mathscr{S}^\circ$, and consider for every $k \in \mathbb{N}$ the unique solution $y_k = y(s_k)$ to the state system (1.2) with initial datum y_0. Without loss of generality, we can assume

$$\mathcal{J}(s_k) \leq 1 + \mathcal{J}(s^*) \quad \text{for all } k \in \mathbb{N} \text{ for fixed } \omega \in \Omega.$$

This and (1.1) give the almost sure finiteness of $\|y_k(\omega)\|_{L^2(D\times[0,T])}$.

In view of (1.3), the minimizing sequence s_k is bounded and we may assume without loss of generality that $s_k \to \bar{s}$ for some $\bar{s} \in \mathscr{S}^\circ$.

Recalling (5.1) and Proposition 3.12, we can apply the compactness result in Lemma 5.1, with $\delta_0 = \frac{1}{4}$, and select a (not relabeled) subsequence such that $\{y_k\}_{k\in\mathbb{N}}$ converges strongly in $L^2(D \times [0,T])$ for fixed ω to a limit $\bar{y}$. Then, thanks to[1] the uniqueness of solutions to the deterministic optimization problem, which is [24, Theorem 4.2], the identification $\bar{y}(\omega) = y(\bar{s}, \omega)$ is meaningful at the level of fixed ω. □

A An Auxiliary Result of Borel–Cantelli Type

We state here a simple consequence of the Borel–Cantelli Lemma, which is used several times in the proofs of the main results.

Lemma A.1. *Let Z be a Banach space, with norm $\|\cdot\|_Z$, and $z : \Omega \to Z$. Assume that*

$$\|z\|_{L^2(\Omega,Z)} < +\infty. \tag{A.1}$$

Then the random variable

$$\Omega \ni \omega \mapsto \|z(\omega)\|_Z$$

is almost surely finite.

Proof. For any $m \in \mathbb{N} \cup \{+\infty\}$, we define

$$A_m := \{\omega \in \Omega : \|z(\omega)\|_Z^2 \geq 2^m\}.$$

From (A.1) and the Chebychev inequality we see that

$$\mathbb{P}(A_m) \leq \frac{1}{2^m}\mathbb{E}[\|z\|_Z^2] = \frac{1}{2^m}\|z\|^2_{L^2(\Omega,Z)},$$

and therefore

$$\sum_{m=0}^{+\infty} \mathbb{P}(A_m) < +\infty.$$

From this and the Borel–Cantelli lemma we conclude that

$$0 = \mathbb{P}(A_\infty) = \mathbb{P}(\{\omega \in \Omega : \|z(\omega)\|_Z^2 = +\infty\}),$$

which leads to the desired result. □

Acknowledgment: We are very much indebted to the Referee for her or his very fruitful comments. The first author thanks Nina Gantert and Mauro Rosestolato for enlightening discussions on pathwise stochastic integration.

Funding: Research of Sections 1–3 was supported by the Russian Science Foundation Grant 14-21-00035. The second author was supported by the Australian Research Council Discovery Project DP170104880 NEW Nonlocal Equations at Work.

1 As a side remark, we note that the ω-wise identification $\bar{y}(\omega) = y(\bar{s}, \omega)$ is not enough to ensure that the optimal $y(\bar{s})$ found in Theorem 5.2 is an adapted stochastic process, since $\mathbb{P}$-measurability may be lost when passing to the limit. Therefore, it remains an open problem to show that $y(\bar{s})$ as a function of (ω, x, t) is a solution to (1.2) in the sense of Definition 3.6.

References

[1] N. Abatangelo and E. Valdinoci, *Getting Acquainted with the Fractional Laplacian*, Springer INdAM Ser., Springer, Cham, 2019.
[2] H. Antil and E. Otárola, A FEM for an optimal control problem of fractional powers of elliptic operators, *SIAM J. Control Optim.* **53** (2015), no. 6, 3432–3456.
[3] H. Antil, E. Otárola and A. J. Salgado, A space-time fractional optimal control problem: Analysis and discretization, *SIAM J. Control Optim.* **54** (2016), no. 3, 1295–1328.
[4] H. Antil, E. Otárola and A. J. Salgado, Some applications of weighted norm inequalities to the error analysis of PDE-constrained optimization problems, *IMA J. Numer. Anal.* **38** (2018), no. 2, 852–883.
[5] P. Benner, A. Onwunta and M. Stoll, Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs, *SIAM J. Matrix Anal. Appl.* **37** (2016), no. 2, 491–518.
[6] D. Bors, Optimal control of nonlinear systems governed by dirichlet fractional Laplacian in the minimax framework, preprint (2015), http://arxiv.org/abs/1509.01283.
[7] P. Chen, A. Quarteroni and G. Rozza, Stochastic optimal Robin boundary control problems of advection-dominated elliptic equations, *SIAM J. Numer. Anal.* **51** (2013), no. 5, 2700–2722.
[8] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, 2nd ed., Encyclopedia Math. Appl. 152, Cambridge University Press, Cambridge, 2014.
[9] D. Dentcheva and A. Ruszczyński, Portfolio optimization with stochastic dominance constraints, *J. Banking Finance* **30** (2006), no. 2, 433–451.
[10] M. D. Gunzburger, H.-C. Lee and J. Lee, Error estimates of stochastic optimal Neumann boundary control problems, *SIAM J. Numer. Anal.* **49** (2011), no. 4, 1532–1552.
[11] L. S. Hou, J. Lee and H. Manouzi, Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs, *J. Math. Anal. Appl.* **384** (2011), no. 1, 87–103.
[12] N. E. Humphries, H. Weimerskirch, N. Queiroz, E. J. Southall and D. W. Sims, Foraging success of biological lévy flights recorded in situ, *Proc. Natl. Acad. Sci. USA* **109** (2012), no. 19, 7169–7174.
[13] D. P. Kouri, M. Heinkenschloss, D. Ridzal and B. G. van Bloemen Waanders, A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty, *SIAM J. Sci. Comput.* **35** (2013), no. 4, A1847–A1879.
[14] R. Kruse and S. Larsson, Optimal regularity for semilinear stochastic partial differential equations with multiplicative noise, *Electron. J. Probab.* **17** (2012), Article ID 65.
[15] J.-F. Le Gall, *Brownian Motion, Martingales, and Stochastic Calculus*, Grad. Texts in Math. 274, Springer, Cham, 2016.
[16] G. J. Lord, C. E. Powell and T. Shardlow, *An Introduction to Computational Stochastic PDEs*, Cambridge Texts Appl. Math., Cambridge University Press, New York, 2014.
[17] H. M. Markowitz, *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Basil Blackwell, Oxford, 1987.
[18] B. Pasik-Duncan, Asymptotic distribution of some quadratic functionals of linear stochastic evolution systems, *J. Optim. Theory Appl.* **75** (1992), no. 2, 389–400.
[19] B. Perthame, E. Ribesy and D. Salort, Career plans and wage structures: a mean field game approach, *Math. Eng.* **1** (2019), no. 1, DOI 10.3934/Mine.2018.1.47.
[20] W. Price, A. Martel and K. Lewis, A review of mathematical models in human resource planning, *Omega* **8** (1980), no. 6, 639–645.
[21] E. Rosseel and G. N. Wells, Optimal control with stochastic PDE constraints and uncertain controls, *Comput. Methods Appl. Mech. Engrg.* **213/216** (2012), 152–167.
[22] R. Servadei and E. Valdinoci, On the spectrum of two different fractional operators, *Proc. Roy. Soc. Edinburgh Sect. A* **144** (2014), no. 4, 831–855.
[23] M. Sieberer, T. Clemens, J. Peisker and S. Ofori, Polymer flood field implementation - pattern configuration and horizontal versus vertical wells, in: *SPE Improved Oil Recovery Conference Paper* (Tulsa 2018), Society of Petroleum Engineers, (2018), DOI 10.2118/190233-MS.
[24] J. Sprekels and E. Valdinoci, A new type of identification problems: Optimizing the fractional order in a nonlocal evolution equation, *SIAM J. Control Optim.* **55** (2017), no. 1, 70–93.
[25] B. Stewart, D. Webster, S. Ahmad and J. Matson, Mathematical models for developing a flexible workforce, *Int. J. Production Econ.* **36** (1994), no. 3, 243–254.
[26] S. Taware, A. H. Alhuthali, M. Sharma and A. Datta-Gupta, Optimal rate control under geologic uncertainty: Water flood and EOR processes, *Optim. Eng.* **18** (2017), no. 1, 63–86.
[27] H. Tiesler, R. M. Kirby, D. Xiu and T. Preusser, Stochastic collocation for optimal control problems with stochastic PDE constraints, *SIAM J. Control Optim.* **50** (2012), no. 5, 2659–2682.