OPTIMIZING THE FRACTIONAL POWER IN A MODEL WITH STOCHASTIC PDE CONSTRAINTS

CARINA GELDHAUSER AND ENRICO VALDINOCI

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.

1. Introduction

Generally speaking, optimal control problems with constraints are formulated as

$$\min_{y\in Y, u\in U} \mathcal{J}(y,u) \quad \text{subject to } Constr(y,u)=0 \tag{1.1}$$

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: (1) the control variable appears as the (fractional) exponent of a diffusion operator, and (2) the constraint will be a stochastic PDE in the sense of a PDE driven by a Wiener Process. The model we present possesses a biological interpretation, in which 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 behaviour or certain species, see e.g. [11].

The problem of optimization under uncertainties is wide-spread in engineering and economics. Our model could serve as a very first toy problem to understand 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. [21, 24].

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 [4, 6, 9, 10, 12] and [25]. 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. [10], [19]), are

Chebyshev Laboratory, St. Petersburg State University, 14th Line V.O., 29B, Saint Petersburg 199178 Russia.

Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini 50, 20133 Milan, Italy and Istituto di Matematica Applicata e Tecnologie Informatiche, Consiglio Nazionale delle Ricerche, Via Ferrata 1, 27100 Pavia, Italy, and School of Mathematics and Statistics, University of Western Australia, 35 Stirling Highway, Crawley, Perth, Western Australia 6009, Australia

E-mail addresses: k.geldhauser@spbu.ru, enrico@mat.uniroma3.it.

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. [8] or [17]. 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 [16] 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 author's 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 [22], 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], [3], [1], and [5]. 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.[18], [23] 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 be also 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 $\Phi(s)$, we want to prove the existence of a random variable $\mathcal{J}(\omega)$, defined as a minimizer in s and y of the cost functional

$$\mathcal{J}(y, s, \omega) = \int_0^T \int_D |y(s, x, t, \omega) - y_{D_T}(x, t)|^2 dx dt + \Phi(s) \tag{1.2}$$

subject to the state equation

$$\begin{aligned} dy(t) + \mathcal{L}^s y(t) dt &= dW(t) \qquad \text{in } D \times [0, T] \\ y(., 0) &= y_0 \qquad \text{in } D \end{aligned} \tag{1.3}$$

where $f(x,t) \in L^2(D_T)$ and W is an L^2-Wiener process. The minimizing random variable $\mathcal{J}(\omega)$ of (1.2) subject to (1.3) is called the *solution to the identification problem (IP)*.

The penalty function $\Phi(s)$ is given a priori, to simplify technicalities we assume that $\Phi \in C^2(0, L)$ (for some $L \in (0, +\infty]$) is non-negative and satisfies

$$\lim_{s \to 0} \Phi(s) \;=\; +\infty \;=\; \lim_{s \to L} \Phi(s). \tag{1.4}$$

Moreover, from a technical point of view, $\Phi(s)$ has to be chosen such that the problem has sufficient compactness properties in s. A possible choice used in [22] is $\Phi(s) = \frac{1}{s(L-s)}$, when $L \in (0, +\infty)$, or $\Phi(s) = \frac{e^s}{s}$ when $L = +\infty$.

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 e.g. to [20] and to Section 2.1 here for details about this point.

To optimize the fractional exponent s is challenging already arise in the deterministic case, as, 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 [22], 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 of solutions to (1.3) 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 (1.2) subject to (1.3). Finally, in section 5 we prove the existence of optimal controls, Theorem 5.3, 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 main results of this work are Theorem 5.3 and Theorem 4.5 on the optimality conditions, which we state here in a non-technical form.

Theorem 1.1. *Under natural 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 condition hold for a fixed realisation $\omega \in \Omega$:
(i) necessary condition: *If $\bar{s}$ is an optimal parameter for (IP) and $y(\bar{s})$ the associated unique solution to the state system* (1.3)*, then for almost every $\omega \in \Omega$*

$$\int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s})\, dxdt \; + \; \Phi'(\bar{s}) \; = \; 0. \tag{1.5}$$

(ii) sufficient condition: *If $\bar{s} \in (0, L)$ satisfies the necessary condition* (4.29) *and if in addition*

$$\int_0^T \int_D (\partial_s y(\bar{s}))^2 + (y(\bar{s}) - y_D)\partial^2_{ss} y(\bar{s})\, dxdt \; + \; \Phi''(\bar{s}) \; > \; 0 \tag{1.6}$$

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

2. Notation and setup

2.1. **The functional analytic setting.** We denote by $D \subset \mathbb{R}$ a bounded domain and $x \in D$ the space variable. We will work in the space $L^2(D)$ of square-integrable functions over D, and denote by $\langle ., . \rangle$ the scalar product in $L^2(D)$. We can write every $v \in L^2(D)$ in the form $v = \sum_{j=1}^{+\infty} \langle v, e_j \rangle e_j$, and denote

$$v_j := \langle v, e_j \rangle \tag{2.1}$$

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

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 $0 < \lambda_1 \leq \lambda_2 \leq \ldots \lambda_j \to +\infty$ as $j \to +\infty$ the corresponding eigenvalues of $\mathcal{L}$.

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

$$D(\mathcal{L}) = \left\{ v \in H : \sum_{j\in\mathbb{N}} \lambda_j^2 \langle v, e_j\rangle^2 < +\infty \right\}. \tag{2.2}$$

Thus, $-\mathcal{L}$ is the generator of an analytic semigroup of contractions which has the well-known structures $S(t) = \sum_{j=1}^{+\infty} e^{-\lambda_j t} v_j(x) v_j(y)$.

In analogy to [22], 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

$$\mathcal{L}v = \sum_{j\in\mathbb{N}} \lambda_j \langle v, e_j\rangle e_j. \tag{2.3}$$

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

$$\mathcal{L}^s v = \sum_{j\in\mathbb{N}} \lambda_j^s \langle v, e_j\rangle e_j \tag{2.4}$$

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

$$D(\mathcal{L}^s) \;=\; \left\{ v = \sum_{j=1}^{+\infty} v_j e_j : v_j \in \mathbb{R}, \text{ with } \|v\|_s^2 := \|\mathcal{L}^s v\|^2 = \sum_{j\in\mathbb{N}} \lambda_j^{2s} v_j^2 < +\infty \right\}. \tag{2.5}$$

To define fractional powers of linear operators in this way is a classical approach in SPDE, see [7], [15] or also the more recent work [13]. Next, we define the space $\mathcal{H}^s := \left\{ v \in L^2(D) : \|v\|_{\mathcal{H}^s} < +\infty \right\}$ with the norm

$$\|v\|_{\mathcal{H}^s} := \left(\sum_{j\in\mathbb{N}} \lambda_j^{2s} |\langle v, e_j\rangle|^2 \right)^{1/2}. \tag{2.6}$$

Additional assumption. It will be technically advantageous (see e.g. in (3.7)) to assume that the eigenvalues of $\mathcal{L}$ are bounded away from zero, in formula

$$\alpha < \lambda_1 \leq \lambda_2 \leq \ldots \lambda_j \longrightarrow +\infty. \tag{2.7}$$

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

2.2. **The probabilistic setting.** We denote by $W : \Omega \times [0, T] \to L^2(D)$ 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 is linear, bounded, self-adjoint, positive semidefinite. Moreover, it is convenient to assume that Q has a common set of eigenfunctions with $\mathcal{L}^s$ (and so with $\mathcal{L}$). We fix notation as $Qe_k = \mu_k e_k$. Finally, we assume that $\operatorname{Tr} Q < +\infty$, which implies that the sum of the eigenvalues of Q is bounded.

Note that the 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 $\{B_j\}_{j\in\mathbb{N}}$

$$W(x, t) = \sum_{j=1}^{+\infty} \sqrt{\mu_j} e_j(x) B_j(t), \tag{2.8}$$

and means of an exponential inequality and 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.

3. Construction of solutions

3.1. **Itô equations and notion of solutions.** In the spirit of the definition (2.4), we want to find solutions of the state equation (1.3) by approximation with real-valued stochastic processes $y_j(t,s) := \langle y(.,t), e_j\rangle$, 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 (1.3) as the infinite series

$$y(s)(x,t) = \sum_{j=1}^{+\infty} \langle y(s)(x,t), e_j(x)\rangle e_j(x) = \sum_{j=1}^{+\infty} y_j(s,t) e_j(x). \tag{3.1}$$

Choosing a deterministic initial condition $y_{j,0} = \langle y_0(x), e_j(x)\rangle \in \mathbb{R}$, and employing the series approximation of the Q-Wiener process (2.8), we get the infinite system of Itô equations

$$dy_j(t) = -\lambda_j^s y_j(t)dt + \sqrt{\mu_j}\, dB_j(t). \tag{3.2}$$

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 (3.2) has a unique strong solution which depends continuously on the initial data, as proved for example on page 212 in [14].

We can explicitly solve (3.2) by applying Itô's formula to $e^{\lambda_j^s t} y_j(t)$ and get the independent *fractional Ornstein-Uhlenbeck processes*

$$y_j(t) = y_{j,0} e^{-\lambda_j^s t} + \sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)} dB_j(\tau). \tag{3.3}$$

Notation: We will often write

$$m_j(t,s) := y_{j,0} e^{-\lambda_j^s t}, \qquad W_{\mathcal{L},s}^j(t) := \sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)} dB_j(\tau). \tag{3.4}$$

Lemma 3.1. *Let the initial data $y_0 \in L^2(D)$ be deterministic. Then the sum appearing in* (3.1) *is convergent and its limit $y(s)(x,t)$ is a $L^2(D)$-valued adapted stochastic process.*

Proof. We show first that for fixed t the series (3.1) converges in $L^2(\Omega, L^2(D))$. For this, recall that we have identified above the summands for fixed s as fractional Ornstein-Uhlenbeck processes, so the sum (3.1) reads formally

$$y(s)(x,t) \;=\; \sum_{j=1}^{+\infty} e_j(x) y_{j,0} e^{-\lambda_j^s t} \;+\; \sum_{j=1}^{+\infty} e_j(x) \sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)} dB_j(\tau). \tag{3.5}$$

As $e^{-\lambda_j^s t} \leq 1$ for all $s, t > 0$, we get for the first term

$$\sum_{j=1}^{+\infty} \left| \langle y_0, e_j(x)\rangle e^{-\lambda_j^s t} \right|^2 \leq \sum_{j=1}^{+\infty} |\langle y_0, e_j(x)\rangle|^2 =: \|y_0\|_{L^2(D)}^2 \tag{3.6}$$

which is finite by assumption.

To show the convergence of the second term, the "random perturbation" part, we denote the partial sum by $W_{\mathcal{L},s}^n(t) = \sum_{j=1}^n e_j(x) W_{\mathcal{L},s}^j(t)$. As the sum is finite, we can exchange expectation and summation, use the one-dimensional Itô Isometry and the lower bound assumption on the eigenvalues (2.7) to obtain

$$
\begin{aligned}
\mathbb{E}\left\|W^n_{\mathcal{L},s}(t)\right\|^2_{L^2(D)} &= \sum_{j=1}^{n}\mathbb{E}\left|\sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}dB_j(\tau)\right|^2 \\
&= \sum_{j=1}^{n}\mu_j\int_0^t e^{-2\lambda_l^s(t-\tau)}d\tau \\
&= \sum_{j=1}^{n}\frac{\mu_j}{2\lambda_j^s}\underbrace{\left(1-e^{-2\lambda_j^s t}\right)}_{\leq 1} \overset{(2.7)}{\leq} \frac{1}{2\alpha^s}\sum_{j=1}^{n}\mu_j
\end{aligned} \tag{3.7}
$$

which is finite as μ_j are summable. Similarly, we can calculate for $m > n$

$$
\begin{aligned}
\mathbb{E}\left\|W^m_{\mathcal{L},s}(t)-W^n_{\mathcal{L},s}(t)\right\|^2_{L^2(D)} &= \mathbb{E}\left\|\sum_{l=n+1}^{m}e_l(x)W^l_{\mathcal{L},s}(t)\right\|^2_{L^2(D)} \\
&= \sum_{l=n+1}^{m}\frac{\mu_l}{2\lambda_l^s}\left(1-e^{-2\lambda_l^s t}\right) \\
&\leq \frac{1}{2\alpha^s}\sum_{l=n+1}^{m}\mu_l
\end{aligned} \tag{3.8}
$$

and follow that $W^n_{\mathcal{L},s}(t)$ is a Cauchy sequence for fixed control s and time t.

Note that from (3.7) we can already deduce boundedness in time

$$
\sup_{t\leq T}\mathbb{E}\left\|W^n_{\mathcal{L},s}(t)\right\|^2_{L^2(D)} \overset{(2.7)}{\leq} \frac{1}{2\alpha^s}\sum_{j=1}^{n}\mu_j \leq \frac{1}{2\alpha^s}TrQ \tag{3.9}
$$

from which we conclude that for $n \to +\infty$

$$
\sup_{t\leq T}\mathbb{E}\left\|W_{\mathcal{L},s}(t)-W^n_{\mathcal{L},s}(t)\right\|^2_{L^2(D)} = \frac{1}{2\alpha^s}\sum_{l=n+1}^{+\infty}\mu_l \longrightarrow 0 \tag{3.10}
$$

and so $y(s)(.,t)$ is a $\mathcal{F}_t$-adapted $L^2(D)$-valued process. □

3.2. **Properties of the solution which need faster decay properties.** In contrary to the deterministic case described in [22], an additional assumption is necessary in the stochastic case:

Assumption 3.2. *We assume that* $\mu_j \sim \lambda_j^{-2s-\epsilon}$ *and* s *is such that*

$$
\sum_{j=1}^{+\infty}\lambda_j^{-s} < +\infty \tag{3.11}
$$

Notation From now on, we call the set of all s satisfying Assumption 3.2 the *set of admissible controls* and denote it by $\mathscr{S}$, and denote its interior by $\mathscr{S}^\circ$.

Example 3.3. *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* $\lambda_j = j^2$*. We get for* (3.11)

$$
\sum_{j=1}^{+\infty}\lambda_j^{-s-\varepsilon} = \sum_{j=1}^{+\infty}j^{-2s-2\varepsilon}. \tag{3.12}
$$

From this we conclude that $\mathscr{S} = \left(\frac{1}{2}, L\right)$.

Proposition 3.4. *Let $\mathcal{L}$, s and Q be such that Assumption 3.2 holds. Let the initial data $y_0 \in L^2(D)$ be deterministic. Then the solution to the state equation* (1.3) *satisfies*

$$y(s,t,.) \in L^2(\Omega, \mathcal{H}^s(D)) \tag{3.13}$$

Proof. Recalling (2.6) and (3.4), we calculate for fixed s, using that the constant $\kappa(t) := \sup_r (re^{-rt})$ is finite for $t \in (0,T]$,

$$\lambda_j^s |m_j(t,s)| = |\langle y_0, e_j\rangle| \lambda_j^s e^{-\lambda_j^s t} \leq \kappa(t)|\langle y_0, e_j\rangle| \tag{3.14}$$

and therefore, as $y_0 \in L^2(D)$, the sequence $\{\lambda_j^s m_j(t,s)\} \in \ell^2$ for any $t \in (0,T]$.

As for the random perturbation term, we denote the partial sum $\mathbb{X}_N^2 := \sum_{j=1}^N \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2$. As we have just a finite sum, we can exchange expectation and summation and apply Itô's Isometry to get

$$\begin{aligned} \mathbb{E}\left[\mathbb{X}_N^2\right] &= \sum_{j=1}^N \lambda_j^{2s}\mu_j \mathbb{E}\left[\left(\int_0^t e^{-\lambda_j^s(t-\tau)} dB_j(\tau)\right)^2\right] \\ &= \sum_{j=1}^N \mu_j \lambda_j^{2s} \int_0^t e^{-2\lambda_j^s(t-\tau)} d\tau \\ &= \sum_{j=1}^N \mu_j \lambda_j^s \left(1 - e^{-2\lambda_j^s t}\right) \\ &\leq \sum_{j=1}^N \mu_j \lambda_j^s \sim \sum_{j=1}^N \lambda_j^{-s-\varepsilon}. \end{aligned} \tag{3.15}$$

Therefore,

$$\begin{aligned} \mathbb{E}\Big[\mathbb{X}_N^2 - \mathbb{X}_M^2\Big] &= \mathbb{E}\left[\sum_{j=1}^N \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2 - \sum_{j=1}^M \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2\right] = \mathbb{E}\left[\sum_{l=N+1}^M \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2\right] \\ &\leq \sum_{l=N+1}^M \lambda_j^{-s-\varepsilon} \end{aligned} \tag{3.16}$$

and this sum goes to zero due to the assumption (3.11) on the eigenvalues. Therefore, $\mathbb{X}_N^2 := \sum_{j=1}^N \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2$ is a Cauchy sequence for fixed control s and time t, which concludes the proof. □

3.3. Solution concept and existence of solutions.

Definition 3.5. *We say that $y : \Omega \times D \times [0,T] \to \mathbb{R}$ is an admissible solution to the state equation* (1.3) *if and only if the following conditions are satisfied:*

1. *$y(.t) \in L^2(\Omega, \mathscr{H}^s(D))$ for any $t \in (0,T]$*
2. *For fixed s, $y(x,t) = \sum_{j=1}^{+\infty} y_j(t,s)e_j(x)$ and the stochastic processes $y_j(t,s)$ solve the Itô diffusion equation*

$$dy_j(t) = -\lambda_j^s y_j(t)dt + \sqrt{\mu_j} dB_j(t) \tag{3.17}$$

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

Theorem 3.6. *Let $\mathcal{L}$, and Q be such that Assumption 3.2 holds. Let the initial data $y_0 \in L^2(D)$ be deterministic. Then there exists for every $s \in \mathscr{S}$ a unique solution $y = y(s)$ to the state system* (1.3) *in the sense of Definition 3.5.*

Proof. Condition (1) in Definition 3.5 was shown to hold in Proposition 3.4, condition (2) holds by construction.

As $\sum_{j=1}^{+\infty} \sqrt{\mu_j} e_j(x) B_j(t) \to W(x,t)$ uniformly with probability one for a Q-Wiener process (see (2.8)), the sum $y(x,t) = \sum_{j=1}^{+\infty} y_j(t,s) e_j(x)$ is defined for almost every t. □

Remark 3.7 (Comparison to strong solutions)**.** Our definition above resembles the definition of a *strong solution* to SPDEs, see [7]. Strong solutions are required to be in the domain of the differential operator, which is also the case for Definition 3.5. However, the solution formulation for a strong solution,

$$y(t) = y_0 + \int_0^T \mathcal{L}^s y(\tau) d\tau + W(t) \qquad \mathbb{P} - a.s., \tag{3.18}$$

is disadvantageous for our analysis, as we need the very explicit description of the solution of Definition 3.5 in order to be able to derive concrete optimality conditions and the existence of optimal controls. In our analysis, the almost sure finiteness of $\int_0^T \mathcal{L}^s y(\tau) d\tau$ as required in (3.18) is not needed, it is sufficient to have finiteness in the L^2-sense, which is Proposition 3.8 below.

3.4. **Combined space-time regularity.** In this section we prove space-time regularity results on solutions to the state system (1.3), which we will need in sections 4 and 5.

Proposition 3.8. *Let $\mathcal{L}$, and Q be such that Assumption 3.2 holds. Let the initial data y_0 be deterministic and moreover $y_0 \in \mathcal{H}^{s/2}(D)$. Then the solution to the state equation* (1.3) *satisfies*

$$\|y(s)\|_{L^2(\Omega \times [0,T]; \mathcal{H}^s)} \leq C \tag{3.19}$$

for some $C > 0$.

Proof. We need to show

$$\mathbb{E}\left[\int_0^T \left(\sum_{j=1}^{+\infty} \lambda_j^{2s} \left| y_{j,0} e^{-\lambda_j^s t} + \sqrt{\mu_j} \int_0^t e^{-\lambda_j^s (t-\tau)} dB_j(\tau) \right|^2 \right)^{1/2} dt \right] < +\infty. \tag{3.20}$$

We use the algebraic estimate $(a+b)^2 = a^2 + b^2 + 2ab \leq 3(a^2+b^2)$ to expand

$$
\begin{aligned}
(3.19) &= \int_0^T \mathbb{E}\left[\|y(s)\|^2_{\mathcal{H}^s}\right] dt \\
&= \int_0^T \mathbb{E}\left(\sum_{j=1}^{+\infty} \lambda_j^{2s}\left(m_j(s,t) + W^j_{\mathcal{L},s}(s,t)\right)^2\right)^{1/2} dt \\
&\leq 3\int_0^T \mathbb{E}\left(\sum_{j=1}^{+\infty} \lambda_j^{2s} m_j^2(s,t) \ + \ \sum_{j=1}^{+\infty} \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2\right)^{1/2} dt \\
&\leq 3\cdot\sqrt{2}\int_0^T \mathbb{E}\left[\left(\sum_{j=1}^{+\infty} \lambda_j^{2s} m_j^2(s,t)\right)^{1/2} + \left(\sum_{j=1}^{+\infty} \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2\right)^{1/2}\right] dt \\
&= 3\cdot\sqrt{2}\left\{\int_0^T \mathbb{E}\left(\sum_{j=1}^{+\infty} \lambda_j^{2s} m_j^2(s,t)\right)^{1/2} dt \ + \ \int_0^T \mathbb{E}\left(\sum_{j=1}^{+\infty} \lambda_j^{2s}(W^j_{\mathcal{L},s}(s,t))^2\right)^{1/2} dt\right\}
\end{aligned}
\tag{3.21}
$$

where we used in the second last line that $a \leq b \Rightarrow (a+b)^{1/2} \leq (2b)^{1/2} \leq \sqrt{2}(a^{1/2}+b^{1/2})$.

We analyze (3.21) termwise: For the first term, which is deterministic, we take a finite sum and exchange summation and integration to calculate

$$
\begin{aligned}
\int_0^T \sum_{j=1}^{N} \lambda_j^{2s} m_j^2(s,t) dt &= \int_0^T \sum_{j=1}^{N} y_{0,j}^2 \lambda_j^{2s} e^{-2\lambda_j^s t} dt \\
&= \sum_{j=1}^{N} y_{0,j}^2 \frac{\lambda_j^s}{-2} \underbrace{\left(e^{-2\lambda_j^s T} - 1\right)}_{-1\leq * \leq 0} \\
&= \sum_{j=1}^{N} y_{0,j}^2 \frac{\lambda_j^s}{2} \underbrace{\left(1 - e^{-2\lambda_j^s T}\right)}_{0\leq * \leq 1} \\
&\leq \frac{1}{2}\|y_0\|^2_{\mathcal{H}^{s/2}}
\end{aligned}
\tag{3.22}
$$

and therefore

$$
\begin{aligned}
\int_0^T \sum_{j=1}^{N} \lambda_j^{2s} m_j^2(s,t) - \sum_{j=1}^{M} \lambda_j^{2s} m_j^2(s,t) dt &= \int_0^T \sum_{j=N+1}^{M} y_{0,j}^2 \lambda_j^{2s} e^{-2\lambda_j^s t} dt \\
&\leq \frac{1}{2} \sum_{j=N+1}^{M} y_{0,j}^2 \lambda_j^s \longrightarrow 0
\end{aligned}
\tag{3.23}
$$

as $\|y_0\|^2_{\mathcal{H}^{s/2}}$ is finite. For the second term we look at a partial sum and employ Itô's isometry to get

$$\begin{aligned}\int_0^T \sum_{j=1}^N \lambda_j^{2s}\mu_j \mathbb{E}\left[|\int_0^t e^{-\lambda_j^s(t-\tau)}dB_j(\tau)|^2\right] dt &\leq \int_0^T \sum_{j=1}^N \lambda_j^s \mu_j \left(1-e^{-2\lambda_j^s t}\right) dt \\ &\leq \sum_{j=1}^N \left(\lambda_j^s\mu_j T - \frac{\mu_j}{2}e^{-2\lambda_j^s T} + \frac{\mu_j}{2}\right) \\ &\leq \sum_{j=1}^N \left(\lambda_j^s T + \frac{1}{2}\right)\mu_j\end{aligned} \tag{3.24}$$

which is finite due to Assumption (3.2). Now we look at the difference between two partial sums, and conclude

$$\int_0^T \mathbb{E}\left[|\sum_{j=1}^N \lambda_j^{2s} W^j_{\mathcal{L},s}(t) - \sum_{j=1}^M \lambda_j^{2s} W^j_{\mathcal{L},s}(t)|^2\right] dt \leq \sum_{j=N}^M \left(\lambda_j^s\mu_j T + \frac{\mu_j}{2}\right) \longrightarrow 0 \tag{3.25}$$

due to Assumption (3.2). □

In order for the solution y to qualify as a minimizer of the functional J, we need more regularity properties of y, which we prove in the next proposition. Recall that $D_T := D \times (0,T)$ denotes the space-time cylinder.

Proposition 3.9. *Let $y_0 \in L^2(D)$ be deterministic. Then any solution $y = y(s)$ to the state equation* (1.3) *satisfies the a priori estimate*

$$\|y(s)\|_{L^2(\Omega, L^2(D_T))} \leq C. \tag{3.26}$$

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.

Note that we did not have to require $s \in \mathscr{S}$ here, as we need only space regularity $L^2(D)$.

Proof. By definition,

$$\begin{aligned}(3.26) &= \mathbb{E}\left[\int_0^T \|y(s)\|^2_{L^2(D)} dt\right] \\ &= \int_0^T \sum_{j=1}^{+\infty} |y_{0,j}(s)e^{-\lambda_j^s t}|^2 \;+\; \mathbb{E}\left[\int_0^T |\sum_{j=1}^{+\infty}\sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}dB_j(\tau)|^2 dt\right].\end{aligned} \tag{3.27}$$

For the first term we calculate

$$\int_0^T \sum_{j=1}^{+\infty} |y_{0,j}(s)e^{-\lambda_j^s t}|^2 dt = \frac{1}{2}\sum_{j=1}^{+\infty} |y_{0,j}(s)|^2 \frac{1}{2\lambda_j^s} \underbrace{\left|e^{-2\lambda_j^s T} - 1\right|}_{\leq 1} \leq \frac{1}{2\alpha}\|y_0\|^2_{L^2}. \tag{3.28}$$

For the second term, we consider the partial sum $j \leq N$, due to which we can exchange expectation and summation, and use Itô's Isometry to get

$$
\begin{aligned}
\mathbb{E}\left[\int_0^T |\sum_{j=1}^N W^j_{\mathcal{L},s}(t)|^2 dt\right] &= \mathbb{E}\left[\int_0^T \left|\sum_{j=1}^N \sqrt{\mu_j}\int_0^t e^{-\lambda_j^s(t-\tau)}dB_j(\tau)\right|^2 dt\right] \\
&\leq \int_0^T \sum_{j=1}^N \mu_j \mathbb{E}\left[\left(\int_0^t e^{-\lambda_j^s(t-\tau)}dB_j(\tau)\right)^2\right] dt \\
&\leq \int_0^T \sum_{j=1}^N \frac{\mu_j}{2\lambda_j^s}\left(1-e^{-2\lambda_j^s t}\right) dt \qquad (3.29)\\
&\leq \sum_{j=1}^N \frac{\mu_j}{2\lambda_j^s}T + \frac{\mu_j}{4\lambda_j^{2s}}\underbrace{\left(e^{-2\lambda_j^s T}-1\right)}_{\in(-1,0)} \\
&\leq \left(\frac{T}{2\alpha^s}-\frac{1}{4\alpha^{2s}}\right)\sum_{j=1}^N \mu_j \ \leq c(T,s).
\end{aligned}
$$

Therefore,

$$
\mathbb{E}\left[\int_0^T |\sum_{l=N+1}^M W^j_{\mathcal{L},s}(t)|^2 dt\right] \leq \left(\frac{T}{2\alpha}+\frac{1}{4\alpha^2}\right)\sum_{l=N+1}^M \mu_l \ \longrightarrow 0 \qquad (3.30)
$$

which proves convergence of $y(s)$ in $L^2(\Omega, L^2(D_T))$.

To prove the almost sure statement, we define the random variable $\mathbb{W}_T^N := \int_0^T \sum_{j=1}^N W^j_{\mathcal{L},s}(t)dt$. We apply Chebychev's inequality on (3.29) to see

$$
\begin{aligned}
\mathbb{P}\left(\mathbb{W}_T^N > N\right) \ &\leq \ \frac{1}{N^2}\mathbb{E}\left[\left(\mathbb{W}_T^N\right)^2\right] = \frac{1}{N^2}\mathbb{E}\left[\left(\int_0^T \sum_{j=1}^N W^j_{\mathcal{L},s}(t)dt\right)^2\right] \qquad (3.31)\\
&\leq \ \frac{1}{N^2}\mathbb{E}\left[\int_0^T |\sum_{j=1}^N W^j_{\mathcal{L},s}(t)|^2 dt\right] \ \leq c(T,s)\frac{1}{N^2}.
\end{aligned}
$$

From (3.31) we see

$$
\sum_{N=1}^{+\infty} \mathbb{P}\left(\mathbb{W}_T^N > N\right) \ \leq c(T,s)\sum_{N=1}^{+\infty}\frac{1}{N^2} \qquad (3.32)
$$

and conclude by the Borel-Cantelli Lemma that the random variable $\|y(s)\|_{L^2(D\times[0,T])}$ is almost surely finite. □

3.5. **Hölder continuity.** 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 (1.3) in dependence of s. This is proved via the factorization method (see [7], Chapter II.5.3), which works with interpolation spaces.

Lemma 3.10. *Let $\mathcal{L}$, and Q be such that Assumption 3.2 holds. Let y_0 be deterministic. 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 of the stochastic convolution are δ-Hölder continuous. According to [7], 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 $\mathcal{L}^s$, is satisfied:

$$\int_0^T t^{-2\beta} \|S(t)Q\|_{HS}^2 dt < +\infty. \tag{3.33}$$

Note that $S(t)$ is self-adjoint, and we calculate, using Assumption 3.2 and Hölder's inequality

$$\begin{aligned}
\|S(t)Q\|_{HS}^2 = Tr\left(S(t)QS(t)\right)) &= \sum_{j=1}^{+\infty} \langle (S(t)QS(t))e_j, e_j \rangle \\
&= \sum_{j=1}^{+\infty} \langle S(t)e_j, e_j \rangle \langle Qe_j, e_j \rangle \langle S^* t)e_j, e_j \rangle \\
&= \sum_{j=1}^{+\infty} \mu_j e^{-2\lambda_j^s t} \\
&= \sum_{j=1}^{+\infty} \lambda_j^{-2s-\varepsilon} e^{-2\lambda_j^s t} \\
&\leq \left(\sum_{j=1}^{+\infty} \left(\lambda_j^{-2s-\varepsilon}\right)^2 \right)^{1/2} \left(\sum_{j=1}^{+\infty} e^{-4\lambda_j^s t} \right)^{1/2}
\end{aligned} \tag{3.34}$$

which is finite due to (3.11) and we conclude that (3.33) is verified when $\beta < \frac{1}{2}$, which proves the Lemma. □

4. Differentiability of the control-to-state operator

As we are optimizing with respect to the exponent s of $\mathcal{L}^s$, we make the dependence on the control parameter explicit by defining the *control-to-state operator* $\mathcal{S} : s \to y(s)$ via

$$\begin{aligned}
\mathcal{S}(\omega) : (0, +\infty) &\to \mathcal{L}\left((0, +\infty), C([0,T], L^2(D))\right) \\
s &\mapsto \mathcal{S}(s)(x,t) = y(s)(x,t)
\end{aligned} \tag{4.1}$$

This means that we need our solution to exist pathwise, which is ensured by Lemma 3.10. We start our derivation of necessary and sufficient optimality conditions with a basic property of the Wiener Integral, which we prove here for the reader's convenience.

4.1. A property of the Wiener Integral.

Lemma 4.1. *Fix $s \in \mathscr{S}$ and let $\tau \in [0,T]$. Let $g(s, \cdot) \in C([0,T]) \cap C^2(\mathscr{S})$ and B one-dimensional Brownian Motion. Then*

$$\frac{d}{ds} \int_0^T g(s,\tau) dB(\tau) = \int_0^T \partial_s g(s,\tau) dB(\tau) \tag{4.2}$$

and the random variable $\int_0^T \partial_s g(s,\tau) dB(\tau)$ belongs to $L^2(\Omega)$.

Proof. Let Σ be the set of all decompositions $\{0 = t_0 < t_1 < \ldots < t_n = T\}$ of $[0,T]$, $\sigma = \{t_0, t_1, \ldots, t_n\}$ an element of Σ with $|\sigma| = \max_{1 \leq i \leq n} |t_i - t_{i-1}|$ and define the Riemann sums

$$I_\sigma = \sum_{j=1}^{n} g(s, t_{j-1})(B(t_j) - B(t_{j-1})). \tag{4.3}$$

Then, as we have a finite sum,

$$\begin{aligned} \frac{d}{ds} I_\sigma &= \frac{d}{ds} \sum_{j=1}^{n} g(s, t_{j-1}) \left(B(t_j) - B(t_{j-1}) \right) \\ &= \sum_{j=1}^{n} \partial_s g(s, t_{j-1}) \left(B(t_j) - B(t_{j-1}) \right). \end{aligned} \tag{4.4}$$

We show now that

(4.5)

$$\lim_{|\sigma|\to 0} \frac{d}{ds} I_\sigma = \lim_{|\sigma\to 0|} \sum_{j=1}^{n} \partial_s g(s, t_{j-1}) \left(B(t_j) - B(t_{j-1}) \right) := \lim_{|\sigma|\to 0} \tilde{I}_\sigma := \int_0^T \partial_s g(s, \tau) dB(\tau).$$

For this we pick a second partition $\tilde{\sigma} = \left\{ \widetilde{t_0}, \widetilde{t_1}, \ldots, \widetilde{t_n} \right\} \in \Sigma$ and we define $\delta = \delta(\epsilon)$ such that $|\sigma| < \delta$ and $|\tilde{\sigma}| < \delta$. Take now the union of the two partitions $\eta = \sigma \cup \tilde{\sigma}$ by defining

$$\eta = \{r_0, r_1, \ldots, r_n\} = \{t_0, t_1, \ldots, t_n\} \cup \left\{ \widetilde{t_0}, \widetilde{t_1}, \ldots, \widetilde{t_n} \right\}$$

As the fractional parameter s is fixed, we define for simplicity $f(\tau) = \partial_s g(s, \tau)$. We show now that the limit $\lim_{|\sigma|\to 0} \tilde{I}_\sigma$ of the right hand side of (4.5) exists by showing that

$$\mathbb{E} \left[|\tilde{I}_\sigma(f) - \tilde{I}_{\tilde{\sigma}}(f)|^2 \right] \to 0. \tag{4.6}$$

For this, note that

$$\tilde{I}_\sigma(f) - \tilde{I}_{\tilde{\sigma}}(f) = \sum_{j=1}^{n} f_1(r_{j-1}) - f_2(r_{j-1}) \left(B(r_j) - B(r_{j-1}) \right) \tag{4.7}$$

where f_1 is defined in a stepwise manner on the whole interval $[0, T]$ with values $f(t_l)$ taken from the partition σ and f_2 on $\tilde{\sigma}$ in the same way.

By continuity of f, we deduce from $r_j - r_{j-1} \leq \delta$ (which holds as $|\sigma| < \delta$ and $|\tilde{\sigma}| < \delta$) that $|f(r_{j-1}) - f(r_{j-1})|^2 < \epsilon$ and therefore also

$$\mathbb{E} \left[|\tilde{I}_\sigma(f) - \tilde{I}_{\tilde{\sigma}}(f)|^2 \right] = \sum_{j=1}^{n} \left(f_1(r_{j-1}) - f_2(r_{j-1}) \right)^2 (r_j - r_{j-1}) \leq \epsilon T \tag{4.8}$$

where we employed the fact that $\mathbb{E}[\tilde{I}_\sigma^2] = \sum_{j=1}^{n} f(t_{j-1})^2 (t_j - t_{j-1})$. This proves (4.6) and we obtain the statement of the Lemma. □

4.2. **The differential of the control-to-state operator.** We first recall Lemma 2.2 of [22], which is an auxiliary result on the derivatives of a function of exponential type. For this and for the following computations, it is convenient to introduce two $L^2([0, T])$-functions as in the proof of Theorem 2.3 in [22], namely $\phi_k(t) = 1 + |\ln(t)|^k$ and $\psi_k(t) = \int_0^t \left(1 + |\ln(t - \tau)|^k\right) d\tau$.

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

$$E_{\lambda,t}(s) := e^{-\lambda^s t} \qquad \text{for } s > 0. \tag{4.9}$$

Then there exists constants C_i such that for all $\lambda > 0$, $t \in (0, T]$ and $s > 0$ the function (4.9) satisfies $|E_{\lambda,t}(s)| \leq C_0$ and for $1 \leq k \leq 4$ holds

$$\left| \frac{d^k}{ds^k} E_{\lambda,t}(s) \right| \leq C_k \frac{1}{s^k} \left(1 + |\ln(t)|^k\right) = \frac{C_k}{s^k} \phi(t)^k \tag{4.10}$$

where we defined $L^2([0, T]) \ni \phi_k(t) := (1 + |\ln(t)|)^k$.

Proposition 4.3. *Let $y(s)(x,t)$ as in (3.1) and $y_j(t)$ as in (3.3). Let the initial condition $y_0 \in L^2(D)$ be deterministic. Then the functions*

$$\partial_s y(\bar{s}) := \sum_{j=1}^{+\infty} \partial_s y_j(\cdot,\bar{s}) e_j \quad \text{and} \quad \partial^2_{ss} y(\bar{s}) := \sum_{j=1}^{+\infty} \partial^2_{ss} y_j(\cdot,\bar{s}) e_j \tag{4.11}$$

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

Moreover, for a fixed $s \in (0,+\infty)$, the random variables $\omega \mapsto \|\partial_s y(s,\omega)\|_{L^2(D\times[0,T])}$ and $\omega \mapsto \|\partial_{ss} y(s,\omega)\|_{L^2(D\times[0,T])}$ are almost surely finite.

Proof. We estimate the deterministic part and the stochastic part separately.

Step 1: derivatives of $m_j(t,s)$. The functions $m_j(t,s)$ are deterministic, and we can argue as in [22] to get

$$\left|\frac{\partial^k}{\partial s^k} m_j(t,\bar{s})\right| \le |\langle y_0, e_j\rangle| \left|\frac{d^k}{d^k} E_{\lambda_j,t}(\bar{s})\right| \le \frac{C_s}{\bar{s}^k}(1+|\ln(t)|^k)|\langle y_0, e_j\rangle| \tag{4.12}$$

which is finite as $\phi_k(t) = 1+|\ln(t)|^k \in L^2(0,T)$. We then follow for $1\le k\le 2$

$$\left\|\sum_{j=N}^{N+M} \frac{\partial^k}{\partial s^k} m_j(t,\bar{s}) e_j\right\|^2_{L^2(D\times[0,T])} \le C(k,T)\bar{s}^{-2k}\int_0^T \phi_k^2(t)dt \sum_{j=N}^{N+M} |\langle y_0, e_j\rangle|^2 \longrightarrow 0. \tag{4.13}$$

Step 2: derivatives in s for the stochastic integral. As justified in Lemma 4.1, we can exchange differentiation with respect to s and the stochastic integration and derive therefore the integrand $g(t-\tau) := e^{-\lambda_j^s(t-\tau)}$ with respect to s first. Performing the stochastic integration with the majorant (4.10), we get

$$\begin{aligned}\mathbb{E}(\int_0^t \partial_s g(t-\tau) dB_j(\tau))^2 &= \int_0^t (\partial_s g(t-\tau))^2 d\tau \\ &\le \frac{C}{s^2}\int_0^t (1+|\ln(t-\tau)|)^2\, d\tau \\ &\le s^{-2} C(k,T)\end{aligned} \tag{4.14}$$

where we employed that

$$(1+|\ln(t-\tau)|)^k \le 2^k + 2^k|\ln(t-\tau)|^k = 2^k\left(1+|\ln(t-\tau)|^k\right) = 2^k\psi_k(t). \tag{4.15}$$

Similarly,

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^t \partial_{ss} g(t-\tau) dB_j(\tau)\right)^2\right] &= \int_0^t (\partial_{ss} g(t-\tau))^2 d\tau \\ &\le \frac{C}{s^4}\int_0^t (1+|\ln(t-\tau)|)^4\, d\tau \\ &\le s^{-4} C(k,T).\end{aligned} \tag{4.16}$$

We then follow for $1\le k\le 2$

$$\begin{aligned}\left\|\sum_{j=N}^{N+M} \mathbb{E}\left[\frac{d^k}{ds^k} W^j_{\mathcal{L},s}(t,s)\right]\right\|^2_{L^2(D\times[0,T])} &= \left\|\sum_{j=N}^{N+M} \int_0^t (\partial_s^k g(t-\tau))^2 d\tau\right\|^2_{L^2(D\times[0,T])} \\ &\le \sum_{j=N}^{N+M} C(k,T)\bar{s}^{-2k}\int_0^T \psi_k^2(t)dt \longrightarrow 0\end{aligned} \tag{4.17}$$

which shows the finiteness of (4.11) in $L^2\left(\Omega, L^2(D\times[0,T])\right)$.

Step 3: almost sure statement. Define the random variable $\mathbb{W}_{T,k}^N := \int_0^T \sum_{j=1}^N \frac{d^k}{ds^k} W_{\mathcal{L},s}^j(t)dt$ with $k=1,2$. By (4.14), (4.16) and Chebychev's inequality, we infer that

(4.18)
$$\mathbb{P}\left(\mathbb{W}_{T,k}^N > N\right) \leq \frac{1}{N^2}\mathbb{E}\left[\left(\mathbb{W}_{T,k}^N\right)^2\right] \leq \frac{1}{N^2}\mathbb{E}\left[\int_0^T \frac{d^k}{ds^k}\left(\sum_{j=1}^N W_{\mathcal{L},s}^j(t)\right)^2 dt\right] \leq c(T,s)\frac{1}{N^2}$$

and so

$$\sum_{N=1}^{+\infty} \mathbb{P}\left(\mathbb{W}_{T,k}^N > N\right) \leq c(T,s)\sum_{N=1}^{+\infty}\frac{1}{N^2}. \tag{4.19}$$

Consequently, the almost sure finiteness of $\|\partial_k y(s)\|_{L^2(D\times[0,T])}$ and $\|\partial_{kk} y(s)\|_{L^2(D\times[0,T])}$ follows by Borel-Cantelli's Lemma. □

In the next theorem, we characterize the differential operator $\mathcal{D}_s\mathcal{S}$ for fixed ω in terms of explicitly know quantities, which is crucial for the upcoming derivation of the optimality conditions.

Theorem 4.4. *Let the initial condition $y_0 \in L^2(D)$ be deterministic. Then, for almost every realisation $\mathcal{S}(\omega)$ of $\mathcal{S}$ with $\omega \in \Omega$, the control-to-state operator $\mathcal{S}$ as defined in* (4.1) *is twice differentiable and for every $\bar{s} \in (0,+\infty)$ the first and second derivatives can be identified with the functions $\partial_s y(\bar{s},\omega)$ and $\partial_{ss}^2 y(\bar{s},\omega)$ as*

$$\mathcal{D}_s\mathcal{S}(\bar{s})(h) = h\partial_s y(\bar{s}) \quad \text{and} \quad \mathcal{D}_{ss}^2\mathcal{S}(\bar{s})(h)(\tilde{h}) = h\tilde{h}\partial_{ss}^2 y(\bar{s}) \tag{4.20}$$

Proof. It is sufficient to prove (4.20) in the L^2-sense, as then the statement follows from the almost sure existence of $\partial_s y(\bar{s})$ and $\partial_{ss}^2 y(\bar{s})$, see Proposition 4.3.

Step 1: the initial condition. We apply Taylor's theorem on the function $E_{\lambda,t}(s)$ defined in (4.9) and apply the estimates from Lemma 4.2 to get for a point $\xi_h \in (\bar{s}-|h|, \bar{s}+|h|)$

$$\begin{aligned}\left|E_{\lambda_j,t}(\bar{s}+h) - E_{\lambda_j,t}(\bar{s}) - hE'_{\lambda_j,t}(\bar{s})\right| &= \frac{h^2}{2}\left|E''_{\lambda_j,t}(\xi_h)\right| \\ &\leq \frac{h^2}{2}c\cdot\xi_h^2\phi_2(t) \leq ch^2\cdot\bar{s}^{-2}\phi_2(t)\end{aligned} \tag{4.21}$$

where we recall that $\phi_2(t) := (1+|\ln(t)|)^2 \in L^2([0,T])$. Therefore,

$$\int_0^t \left|E_{\lambda_j,\tau-t}(\bar{s}+h) - E_{\lambda_j,\tau-t}(\bar{s}) - hE'_{\lambda_j,\tau-t}(\bar{s})\right| d\tau \leq c\bar{s}^{-2}\cdot h^2\int_0^t \phi_2(t-\tau)d\tau. \tag{4.22}$$

Analogously,

$$\int_0^t \left|E'_{\lambda_j,\tau-t}(\bar{s}+h) - E'_{\lambda_j,\tau-t}(\bar{s}) - hE''_{\lambda_j,\tau-t}(\bar{s})\right| d\tau \leq c\bar{s}^{-3}\cdot h^2\int_0^t \phi_3(t-\tau)d\tau. \tag{4.23}$$

Consequently, we can estimate

$$|m_j(t,\bar{s}+h) - m_j(t,\bar{s}) - h\partial_s m_j(t,\bar{s})|^2 \leq K(\bar{s})\cdot h^4\phi_2^2(t)|\langle y_0, e_j\rangle|^2 \tag{4.24}$$

and

$$\left|\partial_s m_j(t,\bar{s}+h) - \partial_s m_j(t,\bar{s}) - h\partial_{ss}^2 m_j(t,\bar{s})\right|^2 \leq K(\bar{s})\cdot h^4\phi_3^2(t)|\langle y_0, e_j\rangle|^2. \tag{4.25}$$

Step 2: the stochastic convolution. Using Lemma 4.1 and the additivity of the Wiener Integral, we get

$$\begin{aligned}
&\mathbb{E}\Big|W^j_{\mathcal{L},s}(t,\bar{s}+h) - W^j_{\mathcal{L},s}(t,\bar{s}) - h\frac{d}{ds}W^j_{\mathcal{L},s}(t,\bar{s})\Big|^2 \\
&\le \mu_j \mathbb{E}\left|\int_0^t e^{-\lambda_j^{\bar{s}+h}(t-\tau)}dB_j(\tau) - \int_0^t e^{-\lambda_j^{\bar{s}}(t-\tau)}dB_j(\tau) - h\frac{d}{ds}\int_0^t e^{-\lambda_j^{\bar{s}}(t-\tau)}dB_j(\tau)\right|^2 \\
&\le \mu_j \mathbb{E}\left|\int_0^t e^{-\lambda_j^{\bar{s}+h}(t-\tau)} - e^{-\lambda_j^{\bar{s}}(t-\tau)} - h\partial_s e^{-\lambda_j^{\bar{s}}(t-\tau)}\, dB_j(\tau)\right|^2 \\
&\le \mu_j \int_0^t \left|e^{-\lambda_j^{\bar{s}+h}(t-\tau)} - e^{-\lambda_j^{\bar{s}}(t-\tau)} - h\partial_s e^{-\lambda_j^{\bar{s}}(t-\tau)}\right|^2 d\tau \\
&\le \mu_j \left(ch^2\mu_j\bar{s}^{-2}\int_0^t \phi_2(t-\tau)d\tau\right)^2 \\
&\le K(\bar{s})\mu_j \cdot h^4\phi_2^2(t)
\end{aligned} \tag{4.26}$$

where we employed (4.22) in the second-last inequality. Analogously, employing Lemma 4.1 and (4.23), we get

$$\begin{aligned}
\mathbb{E}\Big|\frac{d}{ds}W^j_{\mathcal{L},s}(t,\bar{s}+h) - \frac{d}{ds}W^j_{\mathcal{L},s}(t,\bar{s}) - h\frac{d^2}{ds^2}W^j_{\mathcal{L},s}(t,\bar{s})\Big|^2 &\le \mu_j\left(ch^2\bar{s}^{-3}\int_0^t \phi_3(t-\tau)d\tau\right)^2 \\
&\le \mu_j K(\bar{s})\cdot h^4\phi_3^2(t).
\end{aligned} \tag{4.27}$$

Therefore, we get

$$\begin{aligned}
&\mathbb{E}\Big[\Big\|y(\bar{s}+h) - y(\bar{s}) - h\sum_{j=1}^{+\infty}\partial_s y_j(\cdot,\bar{s})e_j(x)\Big\|^2_{L^2(D\times[0,T])}\Big] \\
&\qquad\le \lim_{N\to+\infty}\sum_{j=1}^{N}\mathbb{E}\Big[\int_0^T |y_j(t,\bar{s}+h) - y_j(t,\bar{s}) - h\partial_s y_j(t,\bar{s})e_j(x)|^2\, dt\Big] \\
&\qquad\le C(\bar{s})\cdot h^4
\end{aligned} \tag{4.28}$$

which means that for any fixed $s \in \mathrm{S}$, the $L^2(D\times[0,T])$-valued random variable $\mathcal{D}_s\mathcal{S}(\bar{s})$ can be identified in the $L^2(\Omega)$ sense with the random variable $\partial_s y(\bar{s}) \in L^2\left(\Omega, L^2(D\times[0,T])\right)$, which proves (4.20) in the $L^2(\Omega)$ sense. The statement then follows from the almost sure existence of $\partial_s y(\bar{s})$ and $\partial^2_{ss}y(\bar{s})$, see Proposition 4.3. □

4.3. **Optimality conditions.** In this section, we establish first-order necessary conditions and sufficient optimality conditions of optimal controls. Usually, in optimal control theory the first-order conditions are formulated in terms of a variational inequality, which encodes possible control constraints, and an adjoint state equation.

As we have explicit formulas, i.e. for the representation of the solution y and its derivatives in s, we can avoid these abstract concepts.

Theorem 4.5. *Let the assumptions of Theorem 4.4 be satisfied. Moreover, let $y_0 \in L^2(D)$ be deterministic. Then the following holds true for a fixed realisation $\omega \in \Omega$:*

(i) necessary condition: *If $\bar{s}$ is an optimal parameter for (IP) and $y(\bar{s})$ the associated unique solution to the state system* (1.3)*, then for almost every $\omega \in \Omega$*

$$\int_0^T\int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s})\, dxdt \;+\; \Phi'(\bar{s}) \;=\; 0. \tag{4.29}$$

(ii) sufficient condition: *If $\bar{s} \in (0, L)$ satisfies the necessary condition (4.29) and if in addition*

$$\int_0^T \int_D (\partial_s y(\bar{s}))^2 + (y(\bar{s}) - y_D)\partial^2_{ss} y(\bar{s})\, dxdt \; + \; \Phi^{''}(\bar{s}) \; > \; 0 \tag{4.30}$$

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

Proof. We focus on the reduced cost functional $\mathcal{J}(\omega, \mathcal{S}(s), s)$ in dependence on s: By Theorem 4.4, $s \mapsto \mathcal{J}(s) := \mathcal{J}(y(s), s)$ is twice differentiable on $(0, +\infty)$. By the chain rule,

$$\begin{aligned} \mathcal{J}'(\bar{s}) &= \frac{d}{ds}\mathcal{J}(y(\bar{s}), \bar{s}) = \partial_y \mathcal{J}(y(\bar{s}), \bar{s}) \circ D_s\mathcal{S}(\bar{s}) \; + \; \partial_s \mathcal{J}(y(\bar{s}), \bar{s}) \\ &= \int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s}) dxdt \; + \; \Phi^{'}(\bar{s}) \end{aligned} \tag{4.31}$$

and assertion (i) follows. Assertion (ii) follows from

$$\begin{aligned} \mathcal{J}^{''}(\bar{s}) &= \frac{d}{ds}\mathcal{J}(y(\bar{s}), \bar{s}) = \partial_y \mathcal{J}(y(\bar{s}), \bar{s}) \circ D_s\mathcal{S}(\bar{s}) \; + \; \partial_s \mathcal{J}(y(\bar{s}), \bar{s}) \\ &= \int_0^T \int_D (y(\bar{s}) - y_D)\partial_s y(\bar{s}) dxdt \; + \; \Phi^{'}(\bar{s}). \end{aligned} \tag{4.32}$$

□

5. Existence of optimal controls

The existence of optimal controls is shown by showing that 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 to find such a strongly converging subsequence. The compactness result proves that under certain assumptions there exists a minimum Hölder regularity in time which is independent of the fractional exponent.

Assumption 5.1. (1) *The sequence of eigenvalues $\lambda_k \to +\infty$ as $k \to +\infty$,*

(2) *For almost every $\omega \in \Omega$,*

$$\sup_k \left(\|y_k(\omega)\|_{L^2([0,T], \mathcal{H}^{s_k}(D))} \right) < +\infty, \tag{5.1}$$

(3) *For almost every $\omega \in \Omega$,*

$$\sup_k \left(\|y_k(\omega)\|_{L^2([0,T] \times D)} \right) < +\infty, \tag{5.2}$$

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

Lemma 5.2 (Compactness lemma)**.** *Given a sequence (in k) of $L^2(D)$-valued stochastic processes (in (x, t)) with δ-Hölder continuous sample paths and for which $y_k(\omega) \in L^2([0, T], \mathcal{H}^{s_k}(D))$ for fixed $\omega \in \Omega$. Let Assumptions 5.1 hold for y_k.*

Then, for a fixed realisation $\omega \in \Omega$, the sequence $\{y_k(\omega)\}_{k \in \mathbb{N}}$ contains a subsequence that converges strongly in $L^2(D \times [0, T])$.

Proof. Assumptions 5.1 ensure that the infinite string $\left(\{y_{k,1}\}_{k\in\mathbb{N}}, \{y_{k,2}\}_{k\in\mathbb{N}}, \ldots\right)$ lies in the space

$$C^{\delta_0}([0, T]) \times C^{\delta_0}([0, T]) \times \ldots \tag{5.3}$$

Hence, there is a subsequence denoted by k_m which converges in this product space to an infinite string of the form $(y_1^*, y_2^*, \ldots)$, and every $y_j^* \in C^{\delta_0}([0, T])$. We define

$$y^*(x, t) = \sum_{j \in \mathbb{N}} y_j^* e_j(x). \tag{5.4}$$

The convergence of $y_{k_m} \longrightarrow y^*$ follows exactly as in the compactness lemma in the deterministic case, which is Lemma 6.1. of [22], and is therefore omitted. □

Theorem 5.3. *Suppose that Assumption 3.2 and Assumption 5.1 are satisfied. Moreover, let the initial data satisfy* $\sup_{s\in\mathscr{S}} \|y_0\|_{\mathcal{H}^s} < +\infty$.

Then the control problem (IP) has a solution, that is, for almost every fixed $\omega \in \Omega$, $\mathcal{J}(\omega)$ *attains a minimum in* $\mathscr{S}^\circ$ *and moreover*

$$\inf_{s\in\mathscr{S}} \mathcal{J}(\omega) < +\infty. \tag{5.5}$$

Proof. Note first that, by assumptions on $\Phi(s)$, we can find $s^* \in \mathscr{S}^\circ$ such that $\mathcal{J}(s^*,\omega) < +\infty$ and due to (1.4), we infer

$$0 < \inf_{s\in\mathscr{S}^\circ} \mathcal{J}(s,\omega) < +\infty \qquad \text{for fixed } \omega \in \Omega. \tag{5.6}$$

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 $y_k = \mathcal{S}(s_k)$ to the state system (1.3). Without loss of generality, we can assume

$$\mathcal{J}(s_k) \leq 1 + \mathcal{J}(s^*) \qquad \forall k \in \mathbb{N} \text{ for fixed } \omega \in \Omega. \tag{5.7}$$

This gives us, first of all, employing the form of $\mathcal{J}$, the almost sure finiteness of $\|y_k(\omega)\|_{L^2(D\times[0,T])}$. Due to the form of the penalty function Φ, the minimizing sequence s_k is bounded and we may assume without loss of generality that $s_k \to \bar{s}$ for some $\bar{s} \in \mathscr{S}^\circ$. As the initial data satisfies $\sup_{s\in\mathscr{S}^\circ} \|y(0)\|_{\mathcal{H}^s} < +\infty$ and recalling the a-priori estimates, we can apply the compactness result in Lemma 5.2, with $\delta_0 = \frac{1}{4}$, and select a subsequence, which we again index by k, such that $\{y_k\}_{k\in\mathbb{N}}$ converges strongly in $L^2(D \times [0,T])$ for fixed ω to a random variable $\bar{y}$. The identification $\bar{y} = y(\bar{s})$, follows directly from the uniqueness of solutions to (1.3) in the sense of Definition 3.5. □

Acknowledgements

The first author is supported by the Russian Science Foundation grant Nr. 14-21-00035 and thanks Nina Gantert and Mauro Rosestolato for helpful suggestions leading to Lemma 4.1. The second author is supported by the Australian Research Council Discovery Project DP170104880 NEW Nonlocal Equations at Work.

References

[1] H. Antil, E. Otarola, and A. J. Salgado. Some applications of weighted norm inequalities to the error analysis of pde constrained optimization problems. Preprint, http://arxiv.org/abs/1505.03919, 2015.
[2] H. Antil and E. Otárola. A fem for an optimal control problem of fractional powers of elliptic operators. *SIAM Journal on Control and Optimization*, 53(6):3432–3456, 2015.
[3] H. Antil, E. Otárola, and A. J. Salgado. A space-time fractional optimal control problem: Analysis and discretization. *SIAM Journal on Control and Optimization*, 54(3):1295–1328, 2016.
[4] P. Benner, A. Onwunta, and M. Stoll. Block-diagonal preconditioning for optimal control problems constrained by pdes with uncertain inputs. *SIAM Journal on Matrix Analysis and Applications*, 37(2):491–518, 2016.
[5] D. Bors. Optimal control of nonlinear systems governed by dirichlet fractional laplacian in the minimax framework. Preprint, http://arxiv.org/abs/1509.01283, 2015.
[6] P. Chen, A. Quarteroni, and G. Rozza. Stochastic optimal robin boundary control problems of advection-dominated elliptic equations. *SIAM Journal on Numerical Analysis*, 51(5):2700–2722, 2013.
[7] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.
[8] D. Dentcheva and A. Ruszczyński. Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, 30(2):433–451, 2006.
[9] M. D. Gunzburger, H.-C. Lee, and J. Lee. Error estimates of stochastic optimal Neumann boundary control problems. *SIAM J. Numer. Anal.*, 49(4):1532–1552, 2011.

[10] 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(1):87–103, 2011.

[11] N. E. Humphries, H. Weimerskirch, N. Queiroz, E. J. Southall, and D. W. Sims. Foraging success of biological lévy flights recorded in situ. *Proceedings of the National Academy of Sciences*, 109(19):7169–7174, 2012.

[12] 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(4):A1847–A1879, 2013.

[13] R. Kruse and S. Larsson. Optimal regularity for semilinear stochastic partial differential equations with multiplicative noise. *Electron. J. Probab.*, 17:19 pp., 2012.

[14] J.-F. Le Gall. *Brownian motion, martingales, and stochastic calculus*, volume 274 of *Graduate Texts in Mathematics*. Springer, french edition, 2016.

[15] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs*. Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, 2014.

[16] H. M. Markowitz and G. P. Todd. *Mean-variance analysis in portfolio choice and capital markets*, volume 66. John Wiley & Sons, 2000.

[17] B. Pasik-Duncan. Asymptotic distribution of some quadratic functionals of linear stochastic evolution systems. *Journal of Optimization Theory and Applications*, 75(2):389–400, 1992.

[18] W. Price, A. Martel, and K. Lewis. A review of mathematical models in human resource planning. *Omega*, 8(6):639 – 645, 1980.

[19] E. Rosseel and G. N. Wells. Optimal control with stochastic pde constraints and uncertain controls. *Computer Methods in Applied Mechanics and Engineering*, 213:152–167, 2012.

[20] R. Servadei and E. Valdinoci. On the spectrum of two different fractional operators. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 144(4):831–855, 2014.

[21] M. Sieberer, T. Clemens, J. Peisker, and S. Ofori. Polymer flood field implementation - pattern configuration and horizontal versus vertical wells. *SPE Improved Oil Recovery Conference Paper*, 01 2018.

[22] J. Sprekels and E. Valdinoci. A new type of identification problems: Optimizing the fractional order in a nonlocal evolution equation. *SIAM Journal on Control and Optimization*, 55(1):70–93, 2017.

[23] B. Stewart, D. Webster, S. Ahmad, and J. Matson. Mathematical models for developing a flexible workforce. *International Journal of Production Economics*, 36(3):243 – 254, 1994.

[24] S. Taware, A. H. Alhuthali, M. Sharma, and A. Datta-Gupta. Optimal rate control under geologic uncertainty: water flood and eor processes. *Optimization and Engineering*, 18(1):63–86, Mar 2017.

[25] 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(5):2659–2682, 2012.