A NEW TYPE OF IDENTIFICATION PROBLEMS: OPTIMIZING THE FRACTIONAL ORDER IN A NONLOCAL EVOLUTION EQUATION

JÜRGEN SPREKELS* AND ENRICO VALDINOCI†

Abstract. In this paper, we consider a rather general linear evolution equation of fractional type, namely a diffusion type problem in which the diffusion operator is the sth power of a positive definite operator having a discrete spectrum in $\mathbb{R}^+$. We prove existence, uniqueness and differentiability properties with respect to the fractional parameter s. These results are then employed to derive existence as well as first-order necessary and second-order sufficient optimality conditions for a minimization problem, which is inspired by considerations in mathematical biology.

In this problem, the fractional parameter s serves as the "control parameter" that needs to be chosen in such a way as to minimize a given cost functional. This problem constitutes a new class of identification problems: while usually in identification problems the type of the differential operator is prescribed and one or several of its coefficient functions need to be identified, in the present case one has to determine the type of the differential operator itself.

This problem exhibits the inherent analytical difficulty that with changing fractional parameter s also the domain of definition, and thus the underlying function space, of the fractional operator changes.

Key words. Fractional operators, identification problems, first-oder necessary and second-order sufficient optimality conditions, existence, uniqueness, regularity.

AMS subject classifications. 49K21, 35S11, 49R05, 47A60.

1. Introduction.

Let $\Omega \subset \mathbb{R}^n$ be a given open domain and, with a given $T > 0$, $Q := \Omega \times (0, T)$. We consider in Ω the evolution of a fractional diffusion process governed by the $s-$power of a positive definite operator $\mathcal{L}$. In this paper, we study, for a given $L \in (0, +\infty) \cup \{+\infty\}$, the following identification problem for fractional evolutionary systems:

(IP) Minimize the cost function

$$J(y, s) := \frac{1}{2} \int_0^T \int_\Omega \left| y(x,t) - y_Q(x,t) \right|^2 dx\, dt \, + \, \varphi(s) \tag{1.1}$$

with s in the interval $(0, L)$, subject to the fractional evolution problem

$$\partial_t y \, + \, \mathcal{L}^s y \, = \, f \quad \text{in } Q, \tag{1.2}$$

$$y(\cdot, 0) \, = \, y_0 \quad \text{in } \Omega. \tag{1.3}$$

In this connection, $y_Q \in L^2(Q)$ is a given target function, and $\varphi \in C^2(0, L)$ is a nonnegative penalty function satisfying

$$\lim_{s \searrow 0} \varphi(s) = +\infty = \lim_{s \nearrow L} \varphi(s). \tag{1.4}$$

*Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, and Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany.

†Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany, and Dipartimento di Matematica Federigo Enriques, Università degli Studi di Milano, Via Cesare Saldini 50, 20133 Milano, Italy.

Examples of penalty functions which fulfill (1.4) are

$$\begin{aligned} \varphi(s) &= \frac{1}{s\,(L-s)} \qquad \text{for } s \in (0,L), \text{ if } L \neq +\infty \\ \text{and } \varphi(s) &= \frac{e^s}{s} \qquad \text{for } s \in (0,L), \text{ if } L = +\infty. \end{aligned} \tag{1.5}$$

The properties of the right-hand side f and of the initial datum y_0 will be specified later.

Problem **(IP)** defines a class of identification problems which, to the authors' best knowledge, has never been studied before. Indeed, while there exists a vast literature on the identification of coefficient functions or of right-hand sides in parabolic and hyperbolic evolution equations (which cannot be cited here), there are only but a few contributions to the control theory of fractional operators of diffusion type. In this connection, we refer the reader to the recent papers [1], [2], [3] and [4]. However, in these works the fractional operator was fixed and given a priori. In contrast to these papers, in our case the type of the fractional order operator itself, which is defined by the parameter s, is to be determined.

The fact that the fractional order parameter s is the "control variable" in our problem entails a mathematical difficulty, namely, that with changing s also the domain of $\mathcal{L}^s$ changes. As a consequence, in the functional analytic framework also the underlying solution space changes with s. From this, mathematical difficulties have to be expected. For instance, simple compactness arguments are likely not to work if existence is to be proved. In order to overcome this difficulty, we present in Section 4 (see the compactness result of Lemma 6) an argument which is based on Tikhonov's compactness theorem.

Another feature of the problem **(IP)** is the following: if we want to establish necessary and sufficient optimality conditions, then we have to derive differentiability properties of the control-to-state ($s \mapsto y$) mapping. A major part of this work is devoted to this analysis.

In this paper, the fractional power of the diffusive operator is seen as an "optimization parameter". This type of problems has natural applications. For instance, a biological motivation is the following: in the study of the diffusion of biological species (see, e.g., [6, 8, 12, 11] and the references therein) there is experimental evidence (see [18, 9]) that many predatory species follow "fractional" diffusion patterns instead of classical ones: roughly speaking, for instance, suitably long excursions may lead to a more successful hunting strategy. In this framework, optimizing over the fractional parameter s reflects into optimizing over the "average excursion" in the hunting procedure, which plays a crucial role for the survival and the evolution of a biological population (and, indeed, different species in nature adopt different fractional diffusive behaviors).

In this connection[1], the solution y to the state system (1.2), (1.3) can be thought

[1] As a technical remark, we point out that, strictly speaking, in view of their probabilistic and statistical interpretations, many of the experiments available in the literature are often more closely related to fractional operators of integrodifferential type rather than to fractional operators of spectral type, and these two notions are, in general, not the same (see e.g. [15]), although they coincide, for instance, on the torus, and are under reasonable assumptions asymptotic to each other in large domains (see e.g. Theorem 1 in [14] for precise estimates). Of course, the problem considered in this paper does not aim to be exhaustive, and other types of operators and cost functions may be studied as well, and, in fact, in concrete situations different "case by case" analytic and phenomenological considerations may be needed to produce detailed models which are as accurate as possible for "real life" applications.

of as the spatial density of the predators (where the birth and death rates of the population are not taken into account here, but rather its capability of adapting to the environmental situation). In this sense, the minimization of J is related to finding the "optimal" distribution for the population (for instance, in terms of the availability of resources, possibility of using favorable environments, distributions of possible preys, favorable conditions for reproduction, etc.). Differently from the existing literature, this optimization is obtained here by changing the nonlocal diffusion parameter s, where, roughly speaking, a small s corresponds to a not very dynamic population and a large s to a rather mobile one.

The growth condition (1.4) has to be understood against this biological background: in nature, neither a complete immobility of the individuals (i. e., the choice $s = 0$) nor an extremely fast diffusion (observe that even the extreme case $s = L = +\infty$ is allowed in our setting) are likely to guarantee the survival of the species. In this connection, we may interpret the target function y_Q as, e. g., the spatial distribution of the prey. To adapt their strategy, the predators must know these seasonal distributions a priori; however, this is often the case from long standing experience. We also remark that in nature the prey species in turn adapt their behavior to the strategy of the predators; it would thus be more realistic to consider a predator-prey system with two (possibly different) values of s. Such an analysis, however, goes beyond the scope of this work in which we confine ourselves to the simplest possible situation.

The remainder of the paper is organized as follows: in the following section, we formulate the functional analytic framework of our problem and prove the basic well-posedness results for the state system (1.2), (1.3), as well as its differentiability properties with respect to the parameter s. Afterwards, in Section 3, we study the problem **(IP)** and establish the first-order necessary and the second-order sufficient conditions of optimality. Some elementary explicit examples are also provided, in order to show the influence of the boundary data and of the target distribution on the optimal exponent.

The final section then brings an existence result whose proof employs a compactness result (established in Lemma 6), which is based on Tikhonov's compactness theorem.

2. Functional analytic setting and results for the solution operator. The mathematical setting in which we work is the following: we consider an open and bounded domain $\Omega \subset \mathbb{R}^n$ and a differential operator $\mathcal{L}$ acting on functions mapping Ω into $\mathbb{R}$, together with appropriate boundary conditions. We generally assume that there exists a complete orthonormal system (i. e., an orthonormal basis) $\{e_j\}_{j\in\mathbb{N}}$ of $L^2(\Omega)$ having the property that each e_j lies in a suitable subspace $\mathcal{D}$ of $L^2(\Omega)$, and such that e_j is an eigenfunction of $\mathcal{L}$ with corresponding eigenvalue $\lambda_j \in \mathbb{R}$, for any $j \in \mathbb{N}$ (notice that in this way the boundary conditions of the differential operator $\mathcal{L}$ can be encoded in the functional space $\mathcal{D}$). In this setting, we may write, for any $j \in \mathbb{N}$,

$$\mathcal{L}e_j = \lambda_j e_j \text{ in } \Omega, \quad e_j \in \mathcal{D}.$$

We also generally assume that

$$\lambda_j \geqslant 0 \text{ for any } j \in \mathbb{N}.$$

The prototype of operator $\mathcal{L}$ that we have in mind is, of course, (minus) the Laplacian in a bounded and smooth domain Ω (possibly in the distributional sense), together with either Dirichlet or Neumann homogeneous boundary conditions (in these cases, for smooth domains, one can take, respectively, either $\mathcal{D} := H^2(\Omega) \cap H_0^1(\Omega)$ or $\mathcal{D} := H^2(\Omega)$).

For any v, $w \in L^2(\Omega)$, we consider the scalar product

$$\langle v, w \rangle := \int_\Omega v(x)\, w(x)\, dx.$$

In this way, we can write any function $v \in L^2(\Omega)$ in the form

$$v = \sum_{j \in \mathbb{N}} \langle v, e_j \rangle\, e_j,$$

where the equality is indented in the $L^2(\Omega)$-sense, and, if

$$v \in \mathcal{H}^1 := \Big\{ v \in L^2(\Omega) : \ \{\lambda_j \, \langle v, e_j \rangle\}_{j \in \mathbb{N}} \in \ell^2 \Big\}$$

then

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

For any $s > 0$, we define the s-power of the operator $\mathcal{L}$ in the following way. First, we consider the space

$$\mathcal{H}^s := \big\{ v \in L^2(\Omega) : \ \|v\|_{\mathcal{H}^s} < +\infty \big\}, \tag{2.1}$$

where we use the notation

$$\|v\|_{\mathcal{H}^s} := \Big(\sum_{j \in \mathbb{N}} \lambda_j^{2s} \, \big| \langle v, e_j \rangle \big|^2 \Big)^{1/2}. \tag{2.2}$$

Notice that the notation of the space $\mathcal{H}^s$ has been chosen in such a way that $\mathcal{H}^s$, for $s = 1$, reduces to the space $\mathcal{H}^1$ that was introduced above. This notation is reminiscent of, but different from, the notation for fractional Sobolev spaces (roughly speaking, $s = 1$ in our notation forces the Fourier coefficients to be in ℓ^2 weighted by one power of the eigenvalues; in the case of second order operators this would correspond to Sobolev spaces of order two, rather than one, and this difference in the notation is the main reason for which we chose to use calligraphic fonts for our functional spaces).

We then set, for any $v \in \mathcal{H}^s$,

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

We are ready now to define our notion of a solution to the state system: given $y_0 \in L^2(\Omega)$ and $f : \Omega \times [0, T] \to \mathbb{R}$ such that $f(\cdot, t) \in L^2(\Omega)$ for every $t \in [0, T]$, we say that $y : \Omega \times [0, T] \to \mathbb{R}$ is a solution to the state system (1.2), (1.3), if and only if the following conditions are satisfied:

$$y(\cdot, t) \in \mathcal{H}^s \ \text{ for any } \ t \in (0, T], \tag{2.4}$$

$$\lim_{t \searrow 0} \langle y(\cdot, t),\, e_j \rangle = \langle y_0,\, e_j \rangle \quad \text{for all } \ j \in \mathbb{N}, \tag{2.5}$$

for every $j \in \mathbb{N}$, the mapping $(0, T) \ni t \mapsto \langle y(\cdot, t),\, e_j \rangle$ is absolutely continuous, (2.6)

and it holds $\partial_t \langle y(\cdot, t),\, e_j \rangle + \lambda_j^s \langle y(\cdot, t),\, e_j \rangle = \langle f(\cdot, t),\, e_j \rangle$, (2.7)
for every $j \in \mathbb{N}$ and almost every $t \in (0, T)$.

We remark that conditions (2.4), (2.5), (2.6) and (2.7) are precisely the functional analytic translations of the functional identity in (1.2), (1.3).

We begin our analysis with a result that establishes existence, uniqueness and regularity of the solution to the state system (1.2), (1.3).

THEOREM 2.1. *Suppose that $f : \Omega \times [0,T] \to \mathbb{R}$ satisfies $f(\cdot,t) \in L^2(\Omega)$, for every $t \in [0,T]$, as well as*

$$\sum_{j\in\mathbb{N}} f_j^2 \,<\, +\infty, \quad \textit{where } f_j := \sup_{\theta\in(0,T)} \left|\langle f(\cdot,\theta), e_j\rangle\right| . \tag{2.8}$$

Then the following holds true:

(i) *If $y_0 \in L^2(\Omega)$, then there exists for every $s > 0$ a unique solution $y(s) := y$ to the state system* (1.2), (1.3) *that fulfills the conditions* (2.4)–(2.7) *and belongs to $L^2(Q)$. Moreover, with the control-to-state operator $\mathcal{S} : s \mapsto y(s)$, we have the explicit representation*

$$\mathcal{S}(s)(x,t) = y(s)(x,t) = \sum_{j\in\mathbb{N}} y_j(t,s)\, e_j(x) \quad \textit{a. e. in } Q, \tag{2.9}$$

where, for $j \in \mathbb{N}$ and $t \in [0,T]$, we have set

$$y_j(t,s) := \langle y_0,\, e_j\rangle\, e^{-\lambda_j^s t} + \int_0^t \langle f(\cdot,\tau),\, e_j\rangle\, e^{\lambda_j^s(\tau-t)}\, d\tau. \tag{2.10}$$

(ii) *If $y_0 \in \mathcal{H}^{s/2}$, then*

$$y(s) \in H^1(0,T;L^2(\Omega)) \cap L^\infty(0,T;\mathcal{H}^{s/2}) \cap L^2(0,T;\mathcal{H}^s)\,, \tag{2.11}$$

and

$$\partial_t y(s) \,=\, \sum_{j\in\mathbb{N}} \partial_t y_j(\cdot,s)\, e_j\,. \tag{2.12}$$

Moreover, we have the estimate

$$\begin{aligned}&\|\partial_t y(s)\|^2_{L^2(Q)} \,+\, \|y(s)\|^2_{L^\infty(0,T;\mathcal{H}^{s/2})} \,+\, \|y(s)\|^2_{L^2(0,T;\mathcal{H}^s)} \\ &\leqslant\, T \sum_{j\in\mathbb{N}} \sup_{\theta\in(0,T)} |\langle f(\cdot,\theta), e_j\rangle|^2 \,+\, \|y_0\|^2_{\mathcal{H}^{s/2}}.\end{aligned} \tag{2.13}$$

Remark: We point out that formula (2.10) is of classical flavor and related to Duhamel's Superposition Principle. In our setting, this kind of explicit representation is an auxiliary tool used to prove the regularity estimates with respect to the fractional parameter s that will be needed later in this paper.

PROOF OF THEOREM 2.1: **(i)**: We first prove that the series defined in (2.9) represents a function in $L^2(Q)$. To this end, we show that $\{\sum_{j=1}^n y_j(\cdot,s)\, e_j\}_{n\in\mathbb{N}}$

forms a Cauchy sequence in $L^2(Q)$. Indeed, we have, for every $n, p \in \mathbb{N}$, the identity

$$\begin{aligned}
&\Big\| \textstyle\sum_{j=1}^{n+p} y_j(\cdot,s)\, e_j \;-\; \sum_{j=1}^{n} y_j(\cdot,s)\, e_j \Big\|^2_{L^2(Q)} \qquad (2.14)\\
&= \int_0^T \Big\| \textstyle\sum_{j=n+1}^{n+p} y_j(t,s)\, e_j \Big\|^2_{L^2(\Omega)} dt \;=\; \int_0^T \sum_{j=n+1}^{n+p} |y_j(t,s)|^2 dt\,.
\end{aligned}$$

Now, for any $\tau \in (0,t)$, we have that $e^{\lambda_j^s(\tau-t)} \leqslant 1$, since $\lambda_j \geqslant 0$. Accordingly,

$$\begin{aligned}
\left| \int_0^t \langle f(\cdot,\tau),\, e_j\rangle\, e^{\lambda_j^s(\tau-t)}\, d\tau \right| &\leqslant \int_0^t \big|\langle f(\cdot,\tau),\, e_j\rangle\big|\, e^{\lambda_j^s(\tau-t)}\, d\tau\\
&\leqslant \int_0^t \big|\langle f(\cdot,\tau),\, e_j\rangle\big|\, d\tau \leqslant T \sup_{\theta\in(0,T)} |\langle f(\cdot,\theta), e_j\rangle|.
\end{aligned}$$

Thus, it follows from (2.10) that for every $j \in \mathbb{N}$ and $t \in [0,T]$ it holds

$$|y_j(t,s)| \;\leqslant\; |\langle y_0, e_j\rangle| \,+\, T \sup_{\theta\in(0,T)} |\langle f(\cdot,\theta), e_j\rangle|\,.$$

Since $y_0 \in L^2(\Omega)$, we have $\sum_{j\in\mathbb{N}} |\langle y_0, e_j\rangle|^2 = \|y_0\|^2_{L^2(\Omega)}$, and it readily follows from (2.8) that the sequence $\{\sum_{j=1}^n \int_0^T |y_j(t,s)|^2 dt\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, which proves the claim.

Next, we observe that

$$\begin{aligned}
\sup_{\theta\in(0,T)} \|f(\cdot,\theta)\|^2_{L^2(\Omega)} &= \sup_{\theta\in(0,T)} \sum_{j\in\mathbb{N}} \big|\langle f(\cdot,\theta), e_j\rangle\big|^2 \qquad (2.15)\\
&\leqslant \sum_{j\in\mathbb{N}} \sup_{\theta\in(0,T)} \big|\langle f(\cdot,\theta), e_j\rangle\big|^2,
\end{aligned}$$

which is finite, thanks to (2.8). Consequently,

$$\begin{aligned}
&\int_0^T \|f(\cdot,t)\|^2_{L^2(\Omega)}\, dt < +\infty\\
\text{and} \qquad &\int_0^T \|f(\cdot,t)\|_{L^2(\Omega)}\, dt < +\infty.
\end{aligned} \qquad (2.16)$$

Now, we prove the asserted existence result by showing that the function $y(s)$, which is explicitly defined by (2.9), (2.10) in the statement of the theorem, fulfills for every $s > 0$ all of the conditions (2.4)–(2.7). To this end, let $s > 0$ be fixed. We set, for $j \in \mathbb{N}$ and $t \in [0,T]$,

$$v_j(t,s) := \langle y_0,\, e_j\rangle\, e^{-\lambda_j^s t}, \quad w_j(t,s) := \int_0^t \langle f(\cdot,\tau),\, e_j\rangle\, e^{\lambda_j^s(\tau-t)}\, d\tau. \qquad (2.17)$$

Since $y(s) \in L^2(Q)$, we conclude from (2.9) and (2.10) that for every $j \in \mathbb{N}$ and $t \in [0,T]$ it holds that

$$\begin{aligned}
\langle y(s)(\cdot,t), e_j\rangle \;&=\; \lim_{n\to\infty} \sum_{k=1}^{n} \langle y_k(t,s)\, e_k, e_j\rangle \qquad (2.18)\\
&=\; y_j(t,s) \;=\; v_j(t,s) + w_j(t,s).
\end{aligned}$$

Moreover, for any $t \in (0,T]$, we set

$$\kappa(t) := \sup_{r \geqslant 0} \left(r e^{-rt} \right).$$

Notice that $\kappa(t) < +\infty$ for any $t \in (0,T]$, and

$$\lambda_j^s \, |v_j(t,s)| \leqslant \lambda_j^s \big| \langle y_0, \, e_j \rangle \big| \, e^{-\lambda_j^s t} \; \leqslant \; \kappa(t) \, \big| \langle y_0, \, e_j \rangle \big| \, .$$

Since $y_0 \in L^2(\Omega)$, we therefore have

$$\{\lambda_j^s \, v_j(t,s)\}_{j \in \mathbb{N}} \in \ell^2, \text{ for any } t \in (0,T]. \tag{2.19}$$

In addition, it holds that

$$\begin{aligned} \lambda_j^s \, |w_j(t,s)| &\leqslant \int_0^t \big| \langle f(\cdot, \tau), \, e_j \rangle \big| \, \lambda_j^s \, e^{\lambda_j^s (\tau - t)} \, d\tau \\ &\leqslant \sup_{\theta \in (0,T)} \big| \langle f(\cdot, \theta), \, e_j \rangle \big| \int_0^t \lambda_j^s \, e^{\lambda_j^s (\tau - t)} \, d\tau \\ &= \sup_{\theta \in (0,T)} \big| \langle f(\cdot, \theta), \, e_j \rangle \big| \, (1 - e^{-\lambda_j^s t}) \\ &\leqslant \sup_{\theta \in (0,T)} \big| \langle f(\cdot, \theta), \, e_j \rangle \big|, \end{aligned}$$

and we infer from (2.8) that also $\{\lambda_j^s \, w_j(t,s)\}_{j \in \mathbb{N}} \in \ell^2$, for any $t \in (0,T]$. Combining this with (2.18) and (2.19), we see that also the sequence $\{\lambda_j^s \, \langle y(s)(\cdot, t), e_j \rangle\}_{j \in \mathbb{N}}$ belongs to ℓ^2, for any $t \in (0,T]$. Thus, by (2.1) and (2.2), we conclude that $y(s)(\cdot, t) \in \mathcal{H}^s$ for any $t \in (0,T]$, and this proves (2.4).

Next, we point out that (2.5) follows directly from (2.10), and thus we focus on the proof of (2.6) and (2.7). To this end, fix $t \in (0,T)$. If $|h| > 0$ is so small that $t + h \in (0,T)$, then we observe that

$$\begin{aligned} &w_j(t+h,s) - w_j(t,s) \\ &= e^{-\lambda_j^s (t+h)} \int_t^{t+h} \langle f(\cdot, \tau), \, e_j \rangle \, e^{\lambda_j^s \tau} \, d\tau \\ &\quad + \left(e^{-\lambda_j^s h} - 1 \right) \int_0^t \langle f(\cdot, \tau), \, e_j \rangle \, e^{\lambda_j^s (\tau - t)} \, d\tau. \end{aligned} \tag{2.20}$$

On the other hand, if we set

$$g_j(t,s) := \langle f(\cdot, t), \, e_j \rangle \, e^{\lambda_j^s t},$$

then we have that

$$\begin{aligned} \|g_j(\cdot, s)\|_{L^1(0,T)} &\leqslant e^{\lambda_j^s T} \int_0^T \big| \langle f(\cdot, t), \, e_j \rangle \big| \, dt \\ &\leqslant e^{\lambda_j^s T} \int_0^T \|f(\cdot, t)\|_{L^2(\Omega)} \, dt, \end{aligned}$$

which is finite, thanks to (2.16). Hence,

$$g_j(\cdot, s) \in L^1(0,T),$$

and so $w_j(\cdot,s)$ is absolutely continuous, and, by the Lebesgue Differentiation Theorem (see e.g. [13] and the references therein),

$$\lim_{h\to 0}\frac{1}{h}\int_t^{t+h}\langle f(\cdot,\tau),\, e_j\rangle\, e^{\lambda_j^s\tau}\,d\tau = \lim_{h\to 0}\frac{1}{h}\int_t^{t+h} g_j(\tau,s)\,d\tau$$
$$= g_j(t,s) = \langle f(\cdot,t),\, e_j\rangle\, e^{\lambda_j^s t},$$

for almost every $t\in(0,T)$. From this and (2.20), we infer that

$$\lim_{h\to 0}\frac{w_j(t+h,s)-w_j(t,s)}{h} = \langle f(\cdot,t),\, e_j\rangle - \lambda_j^s\int_0^t\langle f(\cdot,\tau),\, e_j\rangle\, e^{\lambda_j^s(\tau-t)}\,d\tau,$$

for almost every $t\in(0,T)$. Since also $v_j(\cdot,s)$ is obviously absolutely continuous, we thus obtain that $y_j(\cdot,s)$ is absolutely continuous and thus differentiable almost everywhere in $(0,T)$, and we have the identity

$$\begin{aligned}
&\partial_t\langle y(s)(\cdot,t),e_j\rangle \;=\; \partial_t y_j(t,s)\\
&=\; -\lambda_j^s\,\langle y_0,\, e_j\rangle\, e^{-\lambda_j^s t} \;+\; \langle f(\cdot,t),\, e_j\rangle \;-\; \lambda_j^s\int_0^t\langle f(\cdot,\tau),\, e_j\rangle\, e^{\lambda_j^s(\tau-t)}\,d\tau\\
&=\; -\lambda_j^s\, y_j(t,s) + \langle f(\cdot,t),\, e_j\rangle\\
&=\; -\lambda_j^s\,\langle y(s)(\cdot,t),e_j\rangle \;+\; \langle f(\cdot,t),\, e_j\rangle, \quad \text{for almost every } \ t\in(0,T)\,.
\end{aligned}$$

This proves (2.6) and (2.7).

As for the uniqueness result, we again fix $s>0$ and assume that there are two solutions $y(s),\tilde{y}(s)\in L^2(Q)$. We put $y^*(s):=y(s)-\tilde{y}(s)$, and, adapting the notation of (2.10), $y_j^*(t,s):=\langle y^*(s)(\cdot,t),e_j\rangle$, for $j\in\mathbb{N}$. Then, using (2.5), (2.6), and (2.7), we infer that for every $j\in\mathbb{N}$ the mapping $t\mapsto y_j^*(t,s)$ is absolutely continuous in $(0,T)$, and it satisfies

$$\partial_t y_j^*(t,s) + \lambda_j^s\, y_j^*(t,s) = 0 \quad \text{for almost every } \ t\in(0,T), \tag{2.21}$$

as well as

$$\lim_{t\searrow 0} y_j^*(t,s) = 0. \tag{2.22}$$

Owing to the absolute continuity of $y_j^*(\cdot,s)$, we obtain (see, e.g., Remark 8 on page 206 of [5]) that $y_j^*(\cdot,s)\in W^{1,1}(0,T)$, so that we can use the chain rule (see, e.g., Corollary 8.11 in [5]). Thus, if we define $\zeta_j:=\ln\big(1+(y_j^*(\cdot,s))^2\big)$ and make use of (2.21), we have that

$$\partial_t\zeta_j = \frac{2y_j^*(\cdot,s)\,\partial_t y_j^*(\cdot,s)}{1+(y_j^*(\cdot,s))^2} = \frac{-2\lambda_j^s\,(y_j^*(\cdot,s))^2}{1+(y_j^*(\cdot,s))^2}\leqslant 0 \quad \text{a.e. in } (0,T).$$

Integrating this relation (see, e.g., Lemma 8.2 in [5]), we find that, for any $t_1<t_2\in(0,T)$,

$$\zeta_j(t_2)\leqslant\zeta_j(t_1).$$

Thus, from (2.22),

$$\zeta_j(t_2)\leqslant\lim_{t_1\searrow 0}\zeta_j(t_1) = \lim_{t_1\searrow 0}\ln\big(1+(y_j^*(t_1,s))^2\big) = \ln(1) = 0,$$

for any $t_2 \in (0,T)$. Since also $\zeta_j \geqslant 0$, we infer that ζ_j vanishes identically, and thus also $y_j^*(\cdot,s)$. This proves the uniqueness claim.

It remains to show the validity of the claim **(ii)**. To this end, let again $s>0$ be fixed and assume that $y_0 \in \mathcal{H}^{s/2}$, which means that $y_0 \in L^2(\Omega)$ and

$$\sum_{j\in\mathbb{N}} \lambda_j^s \, |\langle y_0, e_j\rangle|^2 < +\infty.$$

Now recall that $\partial_t y_j(t,s) + \lambda_j^s \, y_j(t,s) = \langle f(\cdot,t), e_j\rangle$, for every $j\in\mathbb{N}$ and almost every $t\in(0,T)$. Squaring this equality, we find that

$$|\partial_t y_j(t,s)|^2 \,+\, \lambda_j^s \frac{d}{dt}\, |y_j(t,s)|^2 \,+\, \lambda_j^{2s}\, |y_j(t,s)|^2 \;=\; |\langle f(\cdot,t), e_j\rangle|^2\,, \tag{2.23}$$

and integration over $[0,\tau]$, where $\tau \in [0,T]$, yields that for every $j\in\mathbb{N}$ we have the identity

$$\begin{aligned}&\int_0^\tau |\partial_t y_j(t,s)|^2 dt \,+\, \lambda_j^s\, |y_j(\tau,s)|^2 \,+\, \int_0^\tau \lambda_j^{2s}\, |y_j(t,s)|^2 dt \\ &= \lambda_j^s\, |\langle y_0, e_j\rangle|^2 \,+\, \int_0^\tau |\langle f(\cdot,t), e_j\rangle|^2 dt\,,\end{aligned} \tag{2.24}$$

whence, for every $n\in\mathbb{N}\cup\{0\}$, $p\in\mathbb{N}$, and $\tau\in[0,T]$,

$$\begin{aligned}&\int_0^\tau \sum_{j=n+1}^{n+p} |\partial_t y_j(t,s)|^2 dt \,+\, \sum_{j=n+1}^{n+p} \lambda_j^s\, |y_j(\tau,s)|^2 \,+\, \int_0^\tau \sum_{j=n+1}^{n+p} \lambda_j^{2s}\, |y_j(t,s)|^2 dt \\ &\leqslant \sum_{j=n+1}^{n+p} \lambda_j^s\, |\langle y_0, e_j\rangle|^2 \,+\, T \sum_{j=n+1}^{n+p} \sup_{\theta\in(0,T)} |\langle f(\cdot,\theta), e_j\rangle|^2\,.\end{aligned} \tag{2.25}$$

We remark that we exchanged here summations and integrals: since, up to now, we are only dealing with a finite summation, this exchange is valid due to the finite additivity of the integrals (in particular, we do not need here any fine result of measure theory).

Now, using the same Cauchy criterion argument as in the beginning of the proof of **(i)**, we can therefore infer that the series

$$\sum_{j\in\mathbb{N}} \partial_t y_j(\cdot,s)\, e_j, \quad \sum_{j\in\mathbb{N}} \lambda_j^{s/2}\, y_j(\cdot,s)\, e_j, \quad \text{and} \; \sum_{j\in\mathbb{N}} \lambda_j^s\, y_j(\cdot,s)\, e_j,$$

are strongly convergent in the spaces $L^2(Q)$, $L^\infty(0,T;L^2(\Omega))$, and $L^2(Q)$, in this order. Consequently, we have $y(s)\in L^\infty(0,T;\mathcal{H}^{s/2})\cap L^2(0,T;\mathcal{H}^s)$, and also $\partial_t y(s) \in L^2(0,T;L^2(\Omega))$.

We now show that (2.12) holds true, where we denote the limit of series on the right-hand side by z. From the above considerations, we know that, as $n\to\infty$,

$$\sum_{j=1}^n y_j(\cdot,s)\, e_j \;\to\; y(s), \quad \sum_{j=1}^n \partial_t y_j(\cdot,s)\, e_j \;\to\; z, \quad \text{strongly in } \; L^2(Q)\,.$$

Hence, there is a subsequence $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that, for every test function $\phi \in C_0^\infty(Q)$,

$$\phi \sum_{j=1}^{n_k} y_j(\cdot,s)\, e_j \;\to\; \phi\, y(s), \quad \phi \sum_{j=1}^{n_k} \partial_t y_j(\cdot,s)\, e_j \;\to\; \phi\, z, \quad \text{as } \; k\to\infty,$$

pointwise almost everywhere in Q. Using Lebesgue's Dominated Convergence Theorem and Fubini's Theorem twice, we therefore have the chain of equalities

$$
\begin{aligned}
\int_0^T\int_\Omega \phi(x,t)\, z(x,t)\, dx\, dt \;&=\; \lim_{k\to\infty}\int_\Omega \sum_{j=1}^{n_k} e_j(x)\int_0^T \phi(x,t)\, \partial_t y_j(t,s)\, dt\, dx \\
&= -\lim_{k\to\infty}\int_\Omega \sum_{j=1}^{n_k} e_j(x)\int_0^T \partial_t\phi(x,t)\, y_j(t,s)\, dt\, dx \\
&= -\int_0^T\int_\Omega \partial_t\phi(x,t)\, y(s)(x,t)\, dx\, dt\,,
\end{aligned}
$$

for every $\phi\in C_0^\infty(Q)$, that is, we have $z=\partial_t y(s)$ in the sense of distributions. Since $z\in L^2(Q)$, it therefore holds $y(s)\in H^1(0,T;L^2(\Omega))$ with $\partial_t y(s)=z$, as claimed.

Finally, we obtain the estimate (2.13) from choosing $n=0$ and letting $p\to\infty$ in (2.25), which concludes the proof of the assertion. $\square$

Next, we prove an auxiliary result on the derivatives of a function of exponential type that will play an important role in the subsequent analysis. To this end, we define, for fixed $\lambda>0$ and $t>0$, the real-valued function

$$E_{\lambda,t}(s) := e^{-\lambda^s t} \quad \text{for } \; s>0, \tag{2.26}$$

and denote its first, second, and third derivatives with respect to s by $E'_{\lambda,t}(s)$, $E''_{\lambda,t}(s)$, and $E'''_{\lambda,t}(s)$, respectively. We have the following result.

Lemma 2.2. *There exist constants $\widehat{C}_i>0$, $0\leqslant i\leqslant 3$, such that, for all $\lambda>0$, $t\in(0,T]$, and $s>0$,*

$$
\begin{aligned}
&\left|E_{\lambda,t}(s)\right| \leqslant \widehat{C}_0, \quad \left|E'_{\lambda,t}(s)\right| \leqslant s^{-1}\,\widehat{C}_1\left(1+|\ln(t)|\right),\\
&\left|E''_{\lambda,t}(s)\right| \leqslant s^{-2}\,\widehat{C}_2\left(1+|\ln(t)|^2\right), \quad \left|E'''_{\lambda,t}(s)\right| \leqslant s^{-3}\,\widehat{C}_3\left(1+|\ln(t)|^3\right).
\end{aligned}
$$

Proof: Obviously, we may choose $\widehat{C}_0=1$, and a simple differentiation exercise shows that the first three derivatives of $E_{\lambda,t}$ are given by

$$
\begin{aligned}
&E'_{\lambda,t}(s) = -\lambda^s\, t\, e^{-\lambda^s t}\, \ln(\lambda), \quad E''_{\lambda,t}(s) = \lambda^s\, t\, e^{-\lambda^s t}\,(\lambda^s t-1)\,(\ln(\lambda))^2,\\
&E'''_{\lambda,t}(s) = \lambda^s\, t\, e^{-\lambda^s t}\left(3\lambda^s t-1-(\lambda^s t)^2\right)(\ln(\lambda))^3.
\end{aligned}
$$

Now, observe that

$$\frac{\ln(\lambda^s t)-\ln(t)}{s} = \frac{\ln(\lambda^s)+\ln(t)-\ln(t)}{s} = \ln(\lambda).$$

Accordingly, we may substitute for $\ln(\lambda)$ in the above identities to obtain that

$$
\begin{aligned}
E'_{\lambda,t}(s) \;&= -s^{-1}\lambda^s t\, e^{-\lambda^s t}\left(\ln(\lambda^s t)-\ln(t)\right), \\
E''_{\lambda,t}(s) \;&= s^{-2}\lambda^s t\, e^{-\lambda^s t}\,(\lambda^s t-1)\left(\ln(\lambda^s t)-\ln(t)\right)^2,\\
E'''_{\lambda,t}(s) \;&= s^{-3}\lambda^s t\, e^{-\lambda^s t}\left(3\lambda^s t-1-(\lambda^s t)^2\right)\left(\ln(\lambda^s t)-\ln(t)\right)^3.
\end{aligned}
\tag{2.27}
$$

Thus, we may consider $r := \lambda^s t$ as a "free variable" in (2.27). Using the fact that

$$|\ln(r) - \ln(t)|^k \leqslant 2^k \left(|\ln(r)|^k + |\ln(t)|^k \right) \quad \text{for } 1 \leqslant k \leqslant 3,$$

and introducing the finite quantities

$$\begin{aligned} M_1 &:= \sup_{r>0} \left(r\, e^{-r}\, |\ln(r)| \right), \\ M_2 &:= \sup_{r>0} \left(r\, e^{-r} \right), \\ M_3 &:= \sup_{r>0} \left(r\, e^{-r}\, |r-1|\, 4\, |\ln(r)|^2 \right), \end{aligned}$$

$$\begin{aligned} M_4 &:= \sup_{r>0} \left(r\, e^{-r}\, 4\, |r-1| \right), \\ M_5 &:= \sup_{r>0} \left(r\, e^{-r} \left| 3r - 1 - r^2 \right| 8\, |\ln(r)|^3 \right), \\ M_6 &:= \sup_{r>0} \left(r\, e^{-r}\, 8 \left| 3r - 1 - r^2 \right| \right), \end{aligned}$$

we deduce from (2.27) the estimates

$$\begin{aligned} \left| E'_{\lambda,t}(s) \right| &\leqslant s^{-1} \big(M_1 + M_2\, |\ln(t)| \big), \\ \left| E''_{\lambda,t}(s) \right| &\leqslant s^{-2} \big(M_3 + M_4\, |\ln(t)|^2 \big), \\ \left| E'''_{\lambda,t}(s) \right| &\leqslant s^{-3} \big(M_5 + M_6\, |\ln(t)|^3 \big), \end{aligned}$$

whence the assertion follows. □

We are now in the position to derive differentiability properties for the control-to-state mapping $\mathcal{S}$. As a matter of fact, we will focus on the first and second derivatives, but derivatives of higher order may be taken into account with similar methods. In detail, we have the following result:

Theorem 2.3. *Suppose that that $f : \Omega \times [0,T] \to \mathbb{R}$ satisfies $f(\cdot,t) \in L^2(\Omega)$, for every $t \in [0,T]$, as well as the condition (2.8). Moreover, let $y_0 \in L^2(\Omega)$. Then the control-to-state mapping $\mathcal{S}$ is twice Fréchet differentiable on $(0,+\infty)$ when viewed as a mapping from $(0,+\infty)$ into $L^2(Q)$, and for every $\overline{s} \in (0,+\infty)$ the first and second Fréchet derivatives $D_s\mathcal{S}(\overline{s}) \in \mathcal{L}(\mathbb{R}, L^2(Q))$ and $D^2_{ss}\mathcal{S}(\overline{s}) \in \mathcal{L}(\mathbb{R}, \mathcal{L}(\mathbb{R}, L^2(Q)))$ can be identified with the $L^2(Q)$–functions*

$$\partial_s y(\overline{s}) := \sum_{j\in\mathbb{N}} \partial_s y_j(\cdot,\overline{s})\, e_j, \quad \partial^2_{ss} y(\overline{s}) := \sum_{j\in\mathbb{N}} \partial^2_{ss} y_j(\cdot,\overline{s})\, e_j\,, \tag{2.28}$$

respectively. More precisely, we have, for all $h,k \in \mathbb{R}$,

$$D_s\mathcal{S}(\overline{s})(h) = h\, \partial_s y(\overline{s}) \quad \textit{and} \quad D^2_{ss}\mathcal{S}(\overline{s})(h)(k) = h\, k\, \partial^2_{ss} y(\overline{s})\,. \tag{2.29}$$

Moreover, there is a constant $\widehat{C}_4 > 0$ such that for all $\overline{s} \in (0,+\infty)$ it holds that

$$\|D_s\mathcal{S}(\overline{s})\|_{\mathcal{L}(\mathbb{R},L^2(Q))} = \|\partial_s y(\overline{s})\|_{L^2(Q)} \leqslant \frac{\widehat{C}_4}{\overline{s}}\,, \tag{2.30}$$

$$\left\|D^2_{ss}\mathcal{S}(\overline{s})\right\|_{\mathcal{L}(\mathbb{R},\mathcal{L}(\mathbb{R},L^2(Q)))} = \|\partial^2_{ss} y(\overline{s})\|_{L^2(Q)} \leqslant \frac{\widehat{C}_4}{\overline{s}^2}\,. \tag{2.31}$$

PROOF: Let $\overline{s} \in (0, +\infty)$ be fixed. We first show that the functions defined in (2.28) do in fact belong to $L^2(Q)$. To this end, we first note that

$$e^{\lambda_j^s(\tau - t)} \leqslant e^{\lambda_j^s(t-\tau)} \quad \text{for } 0 \leqslant \tau < t,$$

and that for $1 \leqslant k \leqslant 3$ the functions

$$\phi_k(t) := 1 + |\ln(t)|^k, \quad \psi_k(t) := \int_0^t \big(1 + |\ln(t-\tau)|^k\big)\, d\tau, \quad t \in (0, T],$$

belong to $L^2(0,T)$.

To check this fact, we use the substitution $\theta = -\ln(t)$ and we observe that

$$\begin{aligned}\int_0^T |\ln(t)|^k\, dt &\leqslant \int_0^{T+1} |\ln(t)|^k\, dt \\ &\leqslant \int_0^1 |\ln(t)|^k\, dt + \int_1^{T+1} |\ln(T+1)|^k\, dt \\ &= \int_1^{+\infty} \theta^k\, e^{-\theta}\, d\theta + T\,|\ln(T+1)|^k \leqslant C(k,T),\end{aligned}$$

for some $C(k,T) \in (0, +\infty)$. Accordingly,

$$\int_0^T |\phi_k(t)|^2\, dt \leqslant 4 \int_0^T \big(1 + |\ln(t)|^{2k}\big)\, dt \leqslant 4T + 4C(2k, T),$$

hence $\phi_k \in L^2(0,T)$.

Similarly, for any $t \in (0, T]$,

$$\int_0^t |\ln(t-\tau)|^k\, d\tau = \int_0^t |\ln(\vartheta)|^k\, d\vartheta \leqslant \int_0^T |\ln(\vartheta)|^k\, d\vartheta \leqslant C(k,T),$$

and therefore

$$|\psi_k(t)| \leqslant T + C(k,T),$$

which gives that $\psi_k \in L^\infty(0,T) \subset L^2(0,T)$, as desired.

Next, we infer from (2.17) and Lemma 2.2 that, for every $t \in (0,T]$, $j \in \mathbb{N}$, and $1 \leqslant k \leqslant 3$, the estimates

$$\begin{aligned}\left|\frac{\partial^k}{\partial s^k}\, v_j(t, \overline{s})\right| &\leqslant |\langle y_0, e_j\rangle| \left|\frac{d^k}{ds^k}\, E_{\lambda_j, t}(\overline{s})\right| \leqslant \frac{\widehat{C}_k}{\overline{s}^{\,k}}\, \phi_k(t)\, |\langle y_0, e_j\rangle|\,, \\ \left|\frac{\partial^k}{\partial s^k}\, w_j(t, \overline{s})\right| &\leqslant \int_0^t |\langle f(\cdot, \tau), e_j\rangle| \left|\frac{d^k}{ds^k}\, E_{\lambda_j, \tau - t}(\overline{s})\right| d\tau \\ &\qquad \leqslant \widehat{C}_k\, \overline{s}^{\,-k}\, \psi_k(t) \sup_{\theta \in (0,T)} |\langle f(\cdot, \theta), e_j\rangle|\,.\end{aligned}$$

Therefore, recalling (2.18), we find that, for every $p \in \mathbb{N}$, $n \in \mathbb{N} \cup \{0\}$, and $1 \leqslant k \leqslant 2$,

$$
\begin{aligned}
&\left\| \sum_{j=n+1}^{n+p} \frac{\partial^k}{\partial s^k}\, y_j(t,\overline{s})\, e_j \right\|^2_{L^2(Q)} \;\leqslant\; \sum_{j=n+1}^{n+p} \int_0^T \left| \frac{\partial^k}{\partial s^k}\, y_j(t,\overline{s}) \right|^2 dt \\
&\leqslant\; 2 \sum_{j=n+1}^{n+p} \int_0^T \left| \frac{\partial^k}{\partial s^k}\, v_j(t,\overline{s}) \right|^2 dt \,+\, 2 \sum_{j=n+1}^{n+p} \int_0^T \left| \frac{\partial^k}{\partial s^k}\, w_j(t,\overline{s}) \right|^2 dt \\
&\leqslant\; 2\, \widehat{C}_k^2\, \overline{s}^{\,-2k} \Biggl(\int_0^T \phi_k^2(t)\, dt \sum_{j=n+1}^{n+p} |\langle y_0, e_j \rangle|^2 \\
&\qquad\qquad + \int_0^T \psi_k^2(t)\, dt \sum_{j=n+1}^{n+p} \sup_{\theta \in (0,T)} |\langle f(\cdot,\theta), e_j \rangle|^2 \Biggr) \;\longrightarrow 0\,,
\end{aligned}
$$

as $n \to \infty$. The Cauchy criterion for series then shows the validity of our claim. Moreover, taking $n = 0$ and letting $p \to \infty$ in the above estimate, we find that (2.30) and (2.31) are valid provided that (2.29) holds true.

It remains to show the differentiability results. To this end, let $0 < |h| < \overline{s}/2$. Then $\frac{1}{\overline{s}-|h|} < \frac{2}{\overline{s}}$, and, invoking Lemma 2 and Taylor's Theorem, we obtain for all $j \in \mathbb{N}$ and $t \in (0,T]$ the estimates

$$
\begin{aligned}
&\left| E_{\lambda_j,t}(\overline{s}+h) - E_{\lambda_j,t}(\overline{s}) - h\, E'_{\lambda_j,t}(\overline{s}) \right| \;=\; \frac{1}{2}\, h^2 \left| E''_{\lambda_j,t}(\xi_h) \right| \\
&\leqslant\; \frac{1}{2}\, \widehat{C}_2\, \xi_h^{-2}\, \phi_2(t)\, h^2 \;\leqslant\; 2\, \widehat{C}_2\, \overline{s}^{\,-2}\, \phi_2(t)\, h^2, \\
&\left| E'_{\lambda_j,t}(\overline{s}+h) - E'_{\lambda_j,t}(\overline{s}) - h\, E''_{\lambda_j,t}(\overline{s}) \right| \;=\; \frac{1}{2}\, h^2 \left| E'''_{\lambda_j,t}(\eta_h) \right| \\
&\leqslant\; 4\, \widehat{C}_3\, \overline{s}^{\,-3}\, \phi_3(t)\, h^2,
\end{aligned}
$$

with suitable points $\xi_h, \eta_h \in (\overline{s}-|h|, \overline{s}+|h|)$. By the same token,

$$
\begin{aligned}
&\int_0^t \left| E_{\lambda_j,\tau-t}(\overline{s}+h) - E_{\lambda_j,\tau-t}(\overline{s}) - h\, E'_{\lambda_j,\tau-t}(\overline{s}) \right| d\tau \\
&\leqslant\; 2\, \widehat{C}_2\, \overline{s}^{\,-2} \int_0^t \phi_2(t-\tau)\, d\tau\, h^2, \\
&\int_0^t \left| E'_{\lambda_j,\tau-t}(\overline{s}+h) - E'_{\lambda_j,\tau-t}(\overline{s}) - h\, E''_{\lambda_j,\tau-t}(\overline{s}) \right| d\tau \\
&\leqslant\; 4\, \widehat{C}_3\, \overline{s}^{\,-3} \int_0^t \phi_3(t-\tau)\, d\tau\, h^2\,.
\end{aligned}
$$

From this, we conclude that with suitable constants $K_i > 0$, $1 \leqslant i \leqslant 4$, which depend on $\overline{s}$ but not on $0<|h|<\overline{s}/2$, $j\in\mathbb{N}$, and $t\in(0,T]$, we have the estimates

$$|v_j(t,\overline{s}+h)-v_j(t,\overline{s})-h\,\partial_s v_j(t,\overline{s})|^2 \;\leqslant\; K_1\,\phi_2^2(t)\,|\langle y_0,e_j\rangle|^2\,h^4, \tag{2.32}$$

$$\begin{aligned}&\big|\partial_s v_j(t,\overline{s}+h)-\partial_s v_j(t,\overline{s})-h\,\partial_{ss}^2 v_j(t,\overline{s})\big|^2\\ &\leqslant\; K_2\,\phi_3^2(t)\,|\langle y_0,e_j\rangle|^2\,h^4,\end{aligned} \tag{2.33}$$

$$\begin{aligned}&|w_j(t,\overline{s}+h)-w_j(t,\overline{s})-h\,\partial_s w_j(t,\overline{s})|^2\\ &\leqslant\; K_3\int_0^T\phi_2^2(t)dt\,\sup_{\theta\in(0,T)}|\langle f(\cdot,\theta),e_j\rangle|^2\,h^4,\end{aligned} \tag{2.34}$$

$$\begin{aligned}&\big|\partial_s w_j(t,\overline{s}+h)-\partial_s w_j(t,\overline{s})-h\,\partial_{ss}^2 w_j(t,\overline{s})\big|^2\\ &\leqslant\; K_4\int_0^T\phi_3^2(t)dt\,\sup_{\theta\in(0,T)}|\langle f(\cdot,\theta),e_j\rangle|^2\,h^4\,.\end{aligned} \tag{2.35}$$

From (2.32) and (2.34), we infer that there is a constant $K_5>0$, which is independent of $0<|h|<\overline{s}/2$, such that

$$\begin{aligned}&\Big\|y(\overline{s}+h)-y(\overline{s})-h\sum_{j\in\mathbb{N}}\partial_s y_j(\cdot,\overline{s})\,e_j\Big\|^2_{L^2(Q)}\\ &\leqslant\;\lim_{n\to\infty}\sum_{j=1}^n\int_0^T|y_j(t,\overline{s}+h)-y_j(t,\overline{s})-h\,\partial_s y_j(t,\overline{s})|^2\,dt\\ &\leqslant\;K_5\Big(\sum_{j\in\mathbb{N}}|\langle y_0,e_j\rangle|^2\;+\;\sum_{j\in\mathbb{N}}f_j^2\Big)\,h^4\,.\end{aligned}$$

Hence, $\mathcal{S}$ is Fréchet differentiable at $\overline{s}$ as a mapping from $(0,+\infty)$ into $L^2(Q)$, and the Fréchet derivative is given by the linear mapping

$$h\mapsto D_s\mathcal{S}(\overline{s})(h)\;=\;h\sum_{j\in\mathbb{N}}\partial_s y_j(\cdot,\overline{s})\,e_j,$$

as claimed. The corresponding result for the second Fréchet derivative follows similarly employing the estimates (2.33) and (2.35). This concludes the proof of the assertion. □

3. Optimality conditions. In this section, we establish first-order necessary and second-order sufficient optimality conditions for the control problem **(IP)**. We do not address the question of existence of optimal controls, here; this will be the subject of the forthcoming section. We have the following result.

THEOREM 3.1. *Suppose that that $f:\Omega\times[0,T]\to\mathbb{R}$ satisfies $f(\cdot,t)\in L^2(\Omega)$, for every $t\in[0,T]$, as well as condition* (2.8). *Moreover, let $y_0\in L^2(\Omega)$ be given. Then the following holds true:*

(i) *If $\overline{s}\in(0,L)$ is an optimal parameter for* **(IP)** *and $y(\overline{s})$ is the associated (unique) solution to the state system* (1.2)–(1.3) *according to Theorem 1, then*

$$\int_0^T\!\!\int_\Omega(y(\overline{s})-y_Q)\,\partial_s y(\overline{s})\,dx\,dt\;+\;\varphi'(\overline{s})\;=\;0, \tag{3.1}$$

where $\partial_s y(\overline{s})$ *is given by* (2.28).

(ii) *If* $\overline{s} \in (0, L)$ *satisfies condition* (3.1) *and, in addition,*

$$\int_0^T \int_\Omega \left[(\partial_s y(\overline{s}))^2 \,+\, (y(\overline{s}) - y_Q)\, \partial^2_{ss} y(\overline{s})\right] dx\, dt \,+\, \varphi''(\overline{s}) \,>\, 0, \tag{3.2}$$

where $\partial^2_{ss} y(\overline{s})$ *is defined in* (2.28), *then* $\overline{s}$ *is optimal for* **(IP)**.

PROOF: By Theorem 2.3, the “reduced” cost functional $s \mapsto \mathcal{J}(s) := J(y(s), s)$ is twice differentiable on $(0, L)$, and it follows directly from the chain rule that

$$\begin{aligned}\mathcal{J}'(\overline{s}) \,&=\, \frac{d}{ds} J(y(\overline{s}), \overline{s}) \,=\, \partial_y J(y(\overline{s}), \overline{s}) \circ D_s \mathcal{S}(\overline{s}) + \partial_s J(y(\overline{s}), \overline{s}) \\ &=\, \int_0^T \int_\Omega (y(\overline{s}) - y_Q)\, \partial_s y(\overline{s})\, dx\, dt \,+\, \varphi'(\overline{s})\,.\end{aligned}$$

Moreover,

$$\mathcal{J}''(\overline{s}) \,=\, \int_0^T \int_\Omega \left[(\partial_s y(\overline{s}))^2 \,+\, (y(\overline{s}) - y_Q)\, \partial^2_{ss} y(\overline{s})\right]\, dx\, dt \,+\, \varphi''(\overline{s})\,.$$

The assertions **(i)** and **(ii)** then immediately follow. □

Remark: In our framework, optimizers $\overline{s}$ can be found by minimizing methods (see Theorem 4.1): in this setting, the conditions in (1.4) assure that the optimal parameter $\overline{s}$ lies in the open interval $(0, L)$. Also, if $\varphi'(s)$ blows up near 0 faster than $1/s$ (as it happens in the examples given in (1.5)), solutions of (3.1) do not accumulate near 0, since, by (2.13) and (2.30),

$$\left| \int_0^T \int_\Omega (y(s) - y_Q)\, \partial_s y(s)\, dx\, dt \right| \leqslant \|y(s) - y_Q\|_{L^2(Q)}\, \|\partial_s y(s)\|_{L^2(Q)} \leqslant \frac{C}{s},$$

for some $C > 0$.

Remark: It is customary in optimal control theory to formulate the first-order necessary optimality conditions in terms of a variational inequality (which encodes possible control constraints) and an adjoint state equation, while second-order sufficient condition also involve the so-called “τ–critical cone” (see, e. g., the textbook [17]). In our situation, we can avoid these abstract concepts, since we have explicit formulas for the relevant quantities at our disposal. Indeed, in order to evaluate $y(s)$, $\partial_s y(s)$, $\partial_{ss} y(s)$, we can use the series representations given in (2.9) and (2.28). In practice, this amounts to determining the eigenvalues λ_j and the associated eigenfunctions e_j up to a sufficiently large index j, and then to making use of the differentation formulas for the functions (2.26) for $\lambda = \lambda_j$ that are provided at the beginning of the proof of Lemma 2.2. Using a standard technique (say, Newton’s method), we then can easily find an approximate minimizer of the cost functional. Also in the case that control constraints $-\infty < a \leqslant s \leqslant b < +\infty$ are to be respected, this strategy would still work to find interior minimizers $\overline{s} \in (a, b)$, while the value of the cost at a and b can also be calculated.

Remark: We recall that in infinite dimensional setting conditions like (3.2) are not necessarily sufficient conditions, see Example 3.3 in [7]. On the other hand, this is the case in finite dimensions.

To clarify Theorem 3.1, we now present two simple explicit examples that outline the behavior of the optimal exponent $\overline{s}$ (recall (3.1) and (3.2)). To make the arguments as simple as possible, we assume that φ is strictly convex and that the forcing term f is identically zero (as a matter of fact, the functions φ presented in (1.5) as examples fulfill also this convexity assumtpion). Notice that under these assumptions on φ the function φ has a unique critical point $s_0 \in (0, +\infty)$, which is a minimum (see Figure 3.1). The examples are related to the fractional Laplacian in one variable,

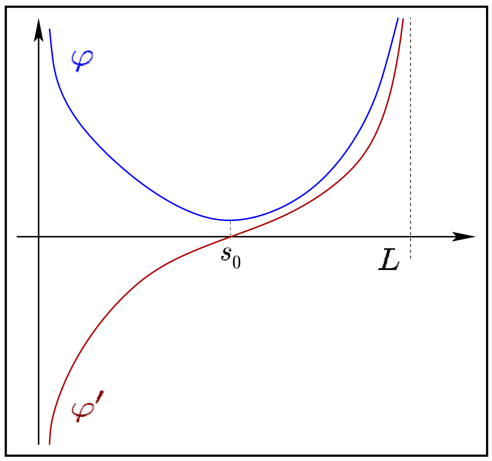

FIGURE 3.1. *The natural cost function φ and its derivative.*

namely, the case of homogeneous Neumann data and the case of homogeneous Dirichlet data on an interval. We will see that, in general, the optimal exponent $\overline{s}$ differs from the minimum s_0 of φ (and, in general, it can be both larger or smaller). In a sense, this shows that different boundary data and different target distributions y_Q influence the optimal exponent $\overline{s}$ and its relation with the minimum s_0 for φ.

Example 1. Consider as operator $\mathcal{L}$ the classical $-\Delta$ on the interval $(0, \pi)$ with homogeneous Neumann data. In this case, we can take as eigenfunctions $e_j(x) := c_j \cos(j\,x)$, where $c_j \in \mathbb{R} \setminus \{0\}$ is a normalizing constant, and $j = 0, 1, 2, 3, \ldots$. The eigenvalue corresponding to e_j is $\lambda_j = j^2$.

Now let, with a fixed $j_0 \in \mathbb{N}$, where $j_0 > 1$, and $\epsilon \in \mathbb{R}$,

$$y_0(x) := 1 + \epsilon\, e_{j_0}(x) \quad \forall\, x \in [0, \pi].$$

Then it is easily verified that for every $s > 0$ the unique solution to (1.2), (1.3) is given by

$$y(s)(x,t) = 1 + \epsilon\, e_{j_0}(x)\, e^{-j_0^{2s}\, t} \quad \forall\, (x,t) \in \overline{Q}.$$

We now make the special choice $y_Q(x,t) := 1$ for the target function. We then observe that

$$\partial_s y(s)(x,t) = -2\epsilon\, j_0^{2s}\, \ln(j_0)\, t\, e_{j_0}(x)\, e^{-j_0^{2s}\, t},$$

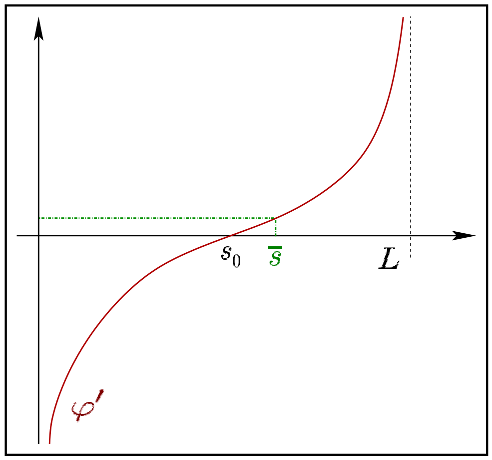

FIGURE 3.2. *The optimal exponent $\overline{s}$ in Example 1.*

and therefore, using the substitution $\vartheta := j_0^{2s}\,t$,

$$\begin{aligned}
&\int_0^T\int_\Omega (y(s)-y_Q)\,\partial_s y(s)\,dx\,dt\\
&= -2\epsilon^2\, j_0^{2s}\,\ln(j_0)\int_0^T\int_\Omega t\,e_{j_0}^2(x)\,e^{-2j_0^{2s}\,t}\,dx\,dt\\
&= -2\epsilon^2\, j_0^{2s}\,\ln(j_0)\int_0^T t\,e^{-j_0^{2s}\,t}\,dt\\
&= -2\epsilon^2\, j_0^{-2s}\,\ln(j_0)\int_0^{j_0^{2s}T}\vartheta\,e^{-2\vartheta}\,dt.
\end{aligned}$$

As a consequence, condition (3.1) becomes, in this case,

$$\varphi'(\overline{s}) = 2\epsilon^2\, j_0^{-2\overline{s}}\,\ln(j_0)\int_0^{j_0^{2\overline{s}}T}\vartheta\,e^{-2\vartheta}\,dt. \tag{3.3}$$

If $\epsilon = 0$ (and when $j_0 \to +\infty$), then the identity in (3.3) reduces to $\varphi'(\overline{s}) = 0$; that is, in this case the "natural" optimal exponent s_0 coincides with the optimal exponent $\overline{s}$ given by the full cost functional (that is, in this case the external conditions given by the exterior forcing term and the resources do not alter the natural diffusive inclination of the population).

But, in general, for fixed $\epsilon \neq 0$ and $j_0 > 1$, the identity in (3.3) gives that $\varphi'(\overline{s}) > 0$. This, given the convexity of φ, implies that $\overline{s} > s_0$, i.e., the optimal exponent given by the cost functional is larger than the natural one (see Figure 3.2).

Example 2. Now we consider as operator $\mathcal{L}$ the classical $-\Delta$ on the interval $(0,\pi)$ with homogeneous Dirichlet data. In this case, we can take as eigenfunctions $e_j(x) := c_j\,\sin(j\,x)$, where $c_j \in \mathbb{R}\setminus\{0\}$ is a normalizing constant, and $j = 1,2,3,\ldots$. The eigenvalue corresponding to e_j is $\lambda_j = j^2$.
For fixed $j_0 \in \mathbb{N}$ with $j_0 \geqslant 1$, and $\epsilon \in \mathbb{R}$, we set

$$y_0(x) := \epsilon\, e_{j_0}(x) \quad \forall\, x \in [0,\pi]\,.$$

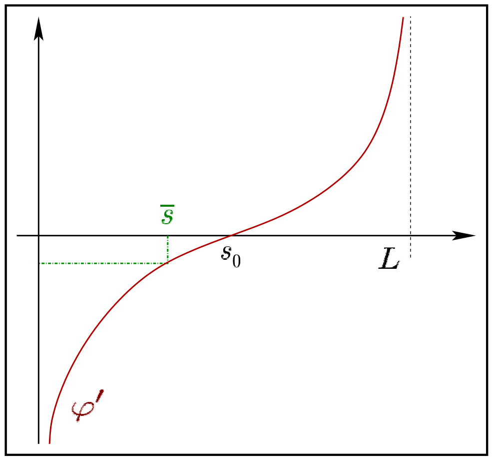

FIGURE 3.3. *The optimal exponent $\overline{s}$ in Example 2.*

Then, for every $s > 0$, the corresponding solution is given by

$$y(s)(x,t) = \epsilon\, e_{j_0}(x)\, e^{-j_0^{2s}\, t} \quad \forall\, (x,t) \in \overline{Q}\,.$$

Now, let $y_Q(x,t) := \epsilon\, e_{j_0}(x)$ for $(x,t) \in Q$. We have

$$\partial_s y(s)(x,t) = -2\epsilon\, j_0^{2s}\, \ln(j_0)\, t\, e_{j_0}(x)\, e^{-j_0^{2s}\, t},$$

and therefore, using the substitution $\vartheta := j_0^{2s}\, t$,

$$\begin{aligned}
&\int_0^T \int_\Omega (y(s) - y_Q)\, \partial_s y(s)\, dx\, dt \\
&= -2\epsilon^2\, j_0^{2s}\, \ln(j_0) \int_0^T \int_\Omega t\, e_{j_0}^2(x) \left(e^{-j_0^{2s}\, t} - 1\right) e^{-j_0^{2s}\, t}\, dx\, dt \\
&= -2\epsilon^2\, j_0^{2s}\, \ln(j_0) \int_0^T t \left(e^{-j_0^{2s}\, t} - 1\right) e^{-j_0^{2s}\, t}\, dt \\
&= -2\epsilon^2\, j_0^{-2s}\, \ln(j_0) \int_0^{j_0^{2s} T} \vartheta \left(e^{-\vartheta} - 1\right) e^{-\vartheta}\, d\vartheta.
\end{aligned}$$

So, in this case, condition (3.1) becomes

$$\varphi'(\overline{s}) = 2\epsilon^2\, j_0^{-2\overline{s}}\, \ln(j_0) \int_0^{j_0^{2\overline{s}} T} \vartheta \left(e^{-\vartheta} - 1\right) e^{-\vartheta}\, d\vartheta. \tag{3.4}$$

If $\epsilon = 0$ (and when $j_0 \to +\infty$), then the identity in (3.4) reduces to $\varphi'(\overline{s}) = 0$, which boils down to $\overline{s} = s_0$. But if $\epsilon \neq 0$ and $j_0 \geqslant 1$, then the identity in (3.4) gives that $\varphi'(\overline{s}) < 0$. By the convexity of φ, this implies that $\overline{s} < s_0$, i. e., the optimal exponent given by the full cost functional is in this case smaller than the natural one (see Figure 3.3).

We observe that, in the framework of Examples 1 and 2, the effect of a larger s is to "cancel faster" the higher order harmonics in the solution y; since these harmonics are related to "wilder oscillations", one may think that the higher s becomes, the bigger the smoothing effect is. In this regard, roughly speaking, a larger s "matches better" with a constant target function y_Q and a smaller s with an oscillating one (compare again Figures 3.2 and 3.3).

We also remark that when $j_0 \geqslant 2$ in Example 2 (or if ϵ is large in Example 1), the solution y is not positive. On the one hand, this seems to reduce the problem, in this case, to a purely mathematical question, since if y represents the density of a biological population, the assumption $y \geqslant 0$ seems to be a natural one. On the other hand, there are other models in applied mathematics in which the condition $y \geqslant 0$ is not assumed: for instance, if y represents the availability of specialized workforce in a given field, the fact that y becomes negative (in some regions of space, at some time) translates into the fact that there is a lack of this specialized workforce (and, for example, non-specialized workers have to be used to compensate this lack).

The use of mathematical models to deal with problems in the job market is indeed an important topic of contemporary research, see e.g. [16] and the references therein.

The models arising in the (short time) job market also provide natural examples in which the birth/death effects in the diffusion equations are negligible.

4. Existence and a compactness lemma. In this section, we establish an existence result for the identification problem **(IP)**. We make the following general assumption for the initial datum y_0:

$$\sup_{s\in(0,L)} \|y_0\|_{\mathcal{H}^s} < +\infty. \tag{4.1}$$

Remark: We remark that the condition (4.1) can be very restrictive if L is large. Indeed, we obviously have $\lambda_j^{2s} \leqslant 1$ for $\lambda_j \leqslant 1$, and for $\lambda_j > 1$ the function $s \mapsto \lambda_j^{2s}$ is strictly increasing. From this it follows that (4.1) is certainly fulfilled for a finite L if only $\|y_0\|_{\mathcal{H}^L} < +\infty$, that is, if $y_0 \in \mathcal{H}^L$.

For an example, consider the prototypical case when $\mathcal{L} = -\Delta$ with zero Dirichlet boundary condition. Then the choice $L = \frac{1}{2}$ leads to the requirement $y_0 \in H_0^1(\Omega)$, while for the choice $L = 1$ we must have $y_0 \in H^2(\Omega) \cap H_0^1(\Omega)$: indeed, if $\{\lambda_j\}_{j\in\mathbb{N}}$ are the corresponding eigenvalues with associated orthogonal eigenfunctions $\{e_j\}_{j\in\mathbb{N}}$, normalized by $\|e_j\|_{L^2(\Omega)} = 1$ for all $j \in \mathbb{N}$, then it is readily verified that the set $\{\lambda_j^{-1/2} e_j\}_{j\in\mathbb{N}}$ forms an orthonormal basis in the Hilbert space $\big(H_0^1(\Omega), \langle\cdot,\cdot\rangle_1\big)$ with respect to the inner product $\langle u, v\rangle_1 := \int_\Omega \nabla u \cdot \nabla v \, dx$. Therefore, if $y_0 \in H_0^1(\Omega)$, it follows from Parseval's identity and integration by parts that

$$\begin{aligned} +\infty \; > \; \|y_0\|^2_{H_0^1(\Omega)} \; &= \; \sum_{j\in\mathbb{N}} \Big| \langle y_0, \lambda_j^{-1/2}\, e_j\rangle_1 \Big|^2 \\ &= \sum_{j\in\mathbb{N}} \frac{1}{\lambda_j} \Big| \int_\Omega \nabla y_0 \cdot \nabla e_j \, dx \Big|^2 \; = \; \sum_{j\in\mathbb{N}} \frac{1}{\lambda_j} \Big| - \int_\Omega y_0 \, \Delta e_j \, dx \Big|^2 \\ &= \sum_{j\in\mathbb{N}} \lambda_j \, |\langle y_0, e_j\rangle|^2 \; = \; \|y_0\|^2_{\mathcal{H}^{1/2}} \, . \end{aligned}$$

The case $L = 1$ is handled similarly. It ought to be clear that with increasing L the condition (4.1) imposes ever higher regularity postulates on y_0. On the other hand, (4.1) is obviously satisfied for every finite $L > 0$ if y_0 belongs to the set of finite linear combinations of the eigenfunctions $\{e_j\}_{j\in\mathbb{N}}$, that is, on a dense subset of $L^2(\Omega)$.

We now give sufficient conditions that guarantee the existence of a solution to the optimal control problem **(IP)**.

Theorem 4.1. *Suppose that $f : \Omega \times [0,T] \to \mathbb{R}$ satisfies $f(\cdot,t) \in L^2(\Omega)$, for every $t \in [0,T]$, as well as condition (2.8). Moreover, let $y_0 \in L^2(\Omega)$ satisfy the condition (4.1). If $\lambda_j \nearrow +\infty$ as $j \to +\infty$, then the control problem* **(IP)** *has a solution, that is, $\mathcal{J}$ attains a minimum in $(0,+\infty)$.*

Before proving the existence result, we establish an auxiliary compactness lemma, which is of some interest in itself, since it acts between spaces with different fractional coefficients s.

Lemma 4.2. *Assume that the sequence $\{\lambda_k\}_{k\in\mathbb{N}}$ of eigenvalues of $\mathcal{L}$ satisfies $\lambda_k \nearrow +\infty$ as $k \to \infty$, and assume that the sequence $\{s_k\}_{k\in\mathbb{N}} \subset (0,+\infty)$ satisfies $s_k \to \overline{s}$ as $k \to \infty$, for some $\overline{s} \in (0,+\infty)\cup\{+\infty\}$. Moreover, let a sequence $\{y_k\}_{k\in\mathbb{N}}$ be given such that $y_k \in L^2(0,T;\mathcal{H}^{s_k})$ and $\partial_t y_k \in L^2(Q)$, for all $k \in \mathbb{N}$, as well as*

$$\sup_{k\in\mathbb{N}} \Big(\|y_k\|_{L^2(Q)} + \|y_k\|_{L^2(0,T;\mathcal{H}^{s_k})} \Big) < +\infty, \quad \textit{and} \tag{4.2}$$

$$\sup_{k\in\mathbb{N}} \|\partial_t y_k\|_{L^2(Q)} < +\infty.$$

Then $\{y_k\}_{k\in\mathbb{N}}$ contains a subsequence that converges strongly in $L^2(Q)$.

Proof: For fixed $j \in \mathbb{N}$, we define

$$y_{k,j}(t) := \int_\Omega y_k(x,t)\, e_j(x)\, dx.$$

Notice that

$$\begin{aligned}\int_0^T |\partial_t y_{k,j}(t)|^2\, dt &\leqslant \int_0^T \left(\int_\Omega |\partial_t y_k(x,t)|\, |e_j(x)|\, dx \right)^2 dt \\ &\leqslant \int_0^T \left(\int_\Omega |\partial_t y_k(x,t)|^2\, dx \right) dt = \|\partial_t y_k\|^2_{L^2(Q)},\end{aligned}$$

which is bounded uniformly in k, thanks to (4.2).

Hence, we obtain a bound in $H^1(0,T)$ for $y_{k,j}$, which is uniform with respect to $k \in \mathbb{N}$, for every $j \in \mathbb{N}$. Owing to the compactness of the embedding $H^1(0,T) \subset C^{1/4}([0,T])$, the sequence $\{y_{k,j}\}_{k\in\mathbb{N}}$ thus forms for every $j \in \mathbb{N}$ a compact subset C_j of $C^{1/4}([0,T])$.

Therefore, the infinite string $\big(\{y_{k,1}\}_{k\in\mathbb{N}}, \{y_{k,2}\}_{k\in\mathbb{N}}, \dots\big)$ lies in $C_1 \times C_2 \times \dots$, which, by virtue of Tikhonov's Theorem, is compact in the product space

$$C^{1/4}([0,T]) \times C^{1/4}([0,T]) \times \dots .$$

Hence, there is a subsequence (denoted by the index k_m), which converges in this product space to an infinite string of the form $\big(y_1^*, y_2^*, \dots\big)$. More explicitly, we have that $y_j^* \in C^{1/4}([0,T])$, for any $j \in \mathbb{N}$, and

$$\lim_{m\to\infty} \|y_{k_m,j} - y_j^*\|_{C^{1/4}([0,T])} = 0 \quad \text{for every } j \in \mathbb{N}. \tag{4.3}$$

We then define

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

and claim that

$$y_{k_m} \to y^* \quad \text{strongly in}\, L^2(Q). \tag{4.4}$$

To prove this claim, we fix $\epsilon \in (0,1)$ and choose $j_* \in \mathbb{N}$ so large that

$$\lambda_j \geqslant \epsilon^{-1} \text{ for any } j \geqslant j_*. \tag{4.5}$$

Then, by (4.3), we may also fix $m_* \in \mathbb{N}$ large enough, so that for any $m \geqslant m_*$ it holds that

$$s_{k_m} \geqslant \min\left\{1, \frac{\overline{s}}{2}\right\} =: \sigma,$$

as well as

$$\|y_{k_m,j} - y_j^*\|_{C^{1/4}([0,T])} \leqslant \frac{\epsilon}{j_*+1} \qquad \text{for every } j < j_*.$$

Now, let $t \in (0,T)$ be fixed. Then, for any $m \geqslant m_*$,

$$\begin{aligned} &\|y^*(\cdot,t) - y_{k_m}(\cdot,t)\|^2_{L^2(\Omega)} = \sum_{j\in\mathbb{N}} |y_j^*(t) - y_{k_m,j}(t)|^2 \\ &\leqslant \sum_{\substack{j\in\mathbb{N}\\ j<j_*}} |y_j^*(t) - y_{k_m,j}(t)|^2 + 4\sum_{\substack{j\in\mathbb{N}\\ j\geqslant j_*}} \big(|y_j^*(t)|^2 + |y_{k_m,j}(t)|^2\big) \\ &\leqslant \epsilon + 4\sum_{\substack{j\in\mathbb{N}\\ j\geqslant j_*}} \big(|y_j^*(t)|^2 + |y_{k_m,j}(t)|^2\big). \end{aligned} \tag{4.6}$$

Moreover, by (4.5), for any $\ell \in \mathbb{N}$,

$$\begin{aligned} &\sum_{\substack{j\in\mathbb{N}\\ j_*\leqslant j\leqslant j_*+\ell}} |y_{k_m,j}(t)|^2 \leqslant \sum_{\substack{j\in\mathbb{N}\\ j_*\leqslant j\leqslant j_*+\ell}} \epsilon^{2s_{k_m}} \lambda_j^{2s_{k_m}} |y_{k_m,j}(t)|^2 \\ &\leqslant \epsilon^{2\sigma} \|y_{k_m}\|^2_{\mathcal{H}^{s_{k_m}}} \leqslant \epsilon^{2\sigma} M, \end{aligned} \tag{4.7}$$

for some $M > 0$, where the last inequality follows from (4.2). Hence, by virtue of (4.3), taking limit as $m \to \infty$, we obtain that

$$\sum_{\substack{j\in\mathbb{N}\\ j_*\leqslant j\leqslant j_*+\ell}} |y_j^*(t)|^2 \leqslant \epsilon^{2\sigma} M. \tag{4.8}$$

Therefore, letting $\ell \to \infty$ in (4.7) and (4.8), we find that

$$\sum_{\substack{j\in\mathbb{N}\\ j\geqslant j_*}} |y_{k_m,j}(t)|^2 \leqslant \epsilon^{2\sigma} M \quad \text{and} \quad \sum_{\substack{j\in\mathbb{N}\\ j\geqslant j_*}} |y_j^*(t)|^2 \leqslant \epsilon^{2\sigma} M.$$

Insertion of these bounds in (4.6) then yields that

$$\|y^*(\cdot,t) - y_{k_m}(\cdot,t)\|^2_{L^2(\Omega)} \leqslant \epsilon + 8\epsilon^{2\sigma} M,$$

as long as $m \geqslant m_*$. By taking ϵ arbitrarily small, we conclude the validity of (4.4) and thus of the assertion of the lemma. $\square$

Proof of Theorem 4.1: The proof is a combination of the Direct Method with the regularity results proved in Theorem 1 and the compactness argument stated in Lemma 6. First of all, we observe that $\mathcal{J}(\frac{L}{2}) < +\infty$ if $0 < L < +\infty$, while $\mathcal{J}(\frac{1}{2}) < +\infty$ if $L = +\infty$. Hence, owing to (1.4), we have

$$0 < \inf_{0<s<L} \mathcal{J}(s) < +\infty.$$

Now, we pick a minimizing sequence $\{s_k\}_{k\in\mathbb{N}} \subset (0,L)$ and consider, for every $k \in \mathbb{N}$, the (unique) solution $y_k := \mathcal{S}(s_k) = y(s_k)$ to the state system (1.2), (1.3) associated with $s = s_k$. We may without loss of generality assume that

$$\mathcal{J}(s_k) \leqslant 1 + \mathcal{J}(s^*) \quad \forall\, k \in \mathbb{N},$$

where $s^* := \frac{L}{2}$ if $L < +\infty$ and $s^* := \frac{1}{2}$ otherwise. We then infer that

$$\|y_k\|_{L^2(Q)} + \varphi(s_k) \leqslant C_1 \quad \forall\, k \in \mathbb{N}, \tag{4.9}$$

where, here and in the following, we denote by C_i, $i \in \mathbb{N}$, constants that may depend on the data of the problem but not on k. In particular, by (1.4), the sequence $\{s_k\}_{k\in\mathbb{N}}$ is bounded, and we may without loss of generality assume that $s_k \to \overline{s}$ for some $\overline{s} \in (0,L)$.
Also, by virtue of (2.13) and (4.1), we obtain that

$$\|\partial_t y_k\|_{L^2(Q)} + \|y_k\|_{L^2(0,T;\mathcal{H}^s)} \leqslant C_2\,, \tag{4.10}$$

whence, in particular,

$$\sum_{j\in\mathbb{N}} \int_0^T |\langle \partial_t y_k(\cdot,t), e_j\rangle|^2\, dt \leqslant C_3 \quad \forall\, k \in \mathbb{N}. \tag{4.11}$$

Thus, using the compactness result of Lemma 6, we can select a subsequence, which is again indexed by k, such that there is some $\overline{y} \in H^1(0,T;L^2(\Omega))$ satisfying

$$\begin{aligned} &y_k \to \overline{y} \quad \text{strongly in } L^2(Q) \text{ and pointwise a.e. in } Q, \\ &y_k \to \overline{y} \quad \text{weakly in } H^1(0,T;L^2(\Omega)). \end{aligned} \tag{4.12}$$

Therefore, we can infer from (4.11) that

$$\sum_{j\in\mathbb{N}} \int_0^T |\langle \partial_t \overline{y}(\cdot,t), e_j\rangle|^2\, dt \leqslant C_3. \tag{4.13}$$

We now claim that $\overline{y} = y(\overline{s})$, that is, that $\overline{y}$ is the (unique) solution to the state system associated with $s = \overline{s}$. To this end, it suffices to show that $\overline{y}$ satisfies the conditions (2.5)–(2.7), since then the claim follows exactly in the same way as uniqueness was established in the proof of Theorem 1; in this connection, observe that for this argument the validity of (2.4) was not needed.

To begin with, we fix $j \in \mathbb{N}$. We conclude from (4.11) that it holds that

$$\int_0^T |\partial_t \langle y_k(\cdot,t), e_j\rangle|^2 \, dt \leqslant C_4 \quad \forall\, k \in \mathbb{N}\,.$$

Hence, the sequence formed by the mappings $t \mapsto \langle y_k(\cdot,t), e_j\rangle$ is a bounded subset of $H^1(0,T)$. Hence, its weak limit, which is given by the mapping $t \mapsto \langle \overline{y}(\cdot,t), e_j\rangle$, belongs to $H^1(0,T)$ and is thus absolutely continuous, which implies that (2.6) holds true for $\overline{y}$.

Moreover, by virtue of the continuity of the embedding $H^1(0,T) \subset C^{1/2}([0,T])$, we can infer from the Arzelà–Ascoli Theorem that the convergence of the sequence $\{\langle y_k(\cdot,t), e_j\rangle\}_{k\in\mathbb{N}}$ is uniform on $[0,T]$. Therefore, to any fixed $\epsilon > 0$ there exists some $k_\epsilon \in \mathbb{N}$ such that, for $k \geqslant k_\epsilon$,

$$\begin{aligned}
&\big|\langle \overline{y}(\cdot,t), e_j\rangle - \langle y_0, e_j\rangle\big| \\
&\leqslant \big|\langle \overline{y}(\cdot,t), e_j\rangle - \langle y_k(\cdot,t), e_j\rangle\big| + \big|\langle y_k(\cdot,t), e_j\rangle - \langle y_0, e_j\rangle\big| \\
&\leqslant \big|\langle y_k(\cdot,t), e_j\rangle - \langle y_0, e_j\rangle\big| + \epsilon.
\end{aligned}$$

Hence, taking the limit in t, and then letting $\epsilon \searrow 0$, we obtain that $\overline{y}$ fulfills (2.5).

Now we use the fact that the mapping $t \mapsto \langle y_k(\cdot,t), e_j\rangle$ belongs to $H^1(0,T)$ to write (2.7) in the weak sense. We have, for any test function $\Psi \in C_0^\infty(0,T)$,

$$\begin{aligned}
&-\int_0^T \langle y_k(\cdot,t),\, e_j\rangle\, \partial_t \Psi(t)\, dt + \lambda_j^{s_k} \int_0^T \langle y_k(\cdot,t),\, e_j\rangle\, \Psi(t)\, dt \\
&\quad = \int_0^T \langle f(\cdot,t),\, e_j\rangle\, \Psi(t)\, dt.
\end{aligned}$$

Passage to the limit as $k \to \infty$ then yields the identity

$$\begin{aligned}
&-\int_0^T \langle \overline{y}(\cdot,t),\, e_j\rangle\, \partial_t \Psi(t)\, dt + \lambda_j^{\overline{s}} \int_0^T \langle \overline{y}(\cdot,t),\, e_j\rangle\, \Psi(t)\, dt \\
&\quad = \int_0^T \langle f(\cdot,t),\, e_j\rangle\, \Psi(t)\, dt.
\end{aligned}$$

This, and the fact that the mapping $t \mapsto \langle \overline{y}(\cdot,t), e_j\rangle$ belongs to the space $H^1(0,T)$, give (2.7) (recall, for instance, Theorem 6.5 in [10]).

In conclusion, it holds $\overline{y} = y(\overline{s})$, and thus the pair $(\overline{s}, \overline{y})$ is admissible for the problem **(IP)**. By the weak sequential semicontinuity of the cost functional, $\overline{s}$ is a minimizer of $\mathcal{J}$. This concludes the proof of the assertion. □

Acknowledgments. EV was supported by ERC grant 277749 "EPSILON Elliptic Pde's and Symmetry of Interfaces and Layers for Odd Nonlinearities" and PRIN grant 201274FYK7 "Critical Point Theory and Perturbative Methods for Nonlinear Differential Equations".

REFERENCES

[1] H. Antil, E. Otárola, *A FEM for an optimal control problem of fractional powers of elliptic operators.* Preprint arXiv:1406.7460v3 [math.OC] 18 April 2015.

[2] H. Antil, E. Otárola, A. J. Salgado, *A fractional space-time optimal control problem: analysis and discretization*. Preprint arXiv:1504.00063v1 [math.OC] 31 March 2015.

[3] H. Antil, E. Otárola, A. J. Salgado, *Some applications of weighted norm inequalities to the analysis of optimal control problems*. Preprint arXiv:1505.03919v1 [math.OC] 14 May 2015.

[4] D. Bors, *Optimal control of nonlinear systems governed by Dirichlet fractional Laplacian in the minimax framework*. Preprint arXiv:1509.01283v1 [math.AP] 3 Sep 2015.

[5] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York (2011).

[6] R. S. Cantrell, C. Cosner, Y. Lou, *Advection-mediated coexistence of competing species*. Proc. Roy. Soc. Edinburgh Sect. A 137, no. 3, 497-518 (2007).

[7] E. Casas, F. Tröltzsch, *Second order optimality conditions and their role in PDE control*. Jahresber. Dtsch. Math.-Ver. 117, no. 1, 3-44 (2015).

[8] A. Friedman, *PDE problems arising in mathematical biology*. Netw. Heterog. Media 7, no. 4, 691-703 (2012).

[9] N. E. Humphries, N. Queiroz, J. R. M. Dyer, N. G. Pade, M. K. Musyl, K. M. Schaefer, D. W. Fuller, J. M. Brunnschweiler, T. K. Doyle, J. D. R. Houghton, G. C. Hays, C. S. Jones, L. R. Noble, V. J. Wearmouth, E. J. Southall, D. W. Sims, *Environmental context explains Lévy and Brownian movement patterns of marine predators*. Nature 465, 1066-1069 (2010).

[10] E. H. Lieb, M. Loss, *Analysis*. Second edition. Graduate Studies in Mathematics. American Mathematical Society, Providence (2001).

[11] A. Massaccesi, E. Valdinoci, *Is a nonlocal diffusion strategy convenient for biological populations in competition?* `http://arxiv.org/pdf/1503.01629.pdf`

[12] E. Montefusco, B. Pellacci, G. Verzini, *Fractional diffusion with Neumann boundary conditions: the logistic equation*. Discrete Contin. Dyn. Syst. Ser. B 18, no. 8, 2175-2202 (2013).

[13] J. J. Koliha, *A fundamental theorem of calculus for Lebesgue integration*. Am. Math. Mon. 113, no. 6, 551-555 (2006).

[14] R. Musina, A. I. Nazarov, *On fractional Laplacians - 3*. To appear on ESAIM Control Optim. Calc. Var.

[15] R. Servadei, E. Valdinoci, *On the spectrum of two different fractional operators*. Proc. Roy. Soc. Edinburgh Sect. A 144, no. 4, 831-855 (2014).

[16] B. D. Stewart, D. B. Webster, S. Ahmad, J. O. Matson, *Mathematical models for developing a flexible workforce*, Intern. Journ. of Production Economics 36, no. 3, 243-254 (1994).

[17] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*. Graduate Studies in Mathematics Vol. 112. American Mathematical Society, Providence, Rhode Island (2010).

[18] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, H. E. Stanley, *Lévy flight search patterns of wandering albatrosses*. Nature 381, 413-415 (1996).