

AN OPTIMIZATION-BASED APPROACH TO PARAMETER LEARNING FOR FRACTIONAL TYPE NONLOCAL MODELS

OLENA BURKOVSKA¹, CHRISTIAN GLUSA², AND MARTA D'ELIA³

ABSTRACT. Nonlocal operators of fractional type are a popular modeling choice for applications that do not adhere to classical diffusive behavior; however, one major challenge in nonlocal simulations is the selection of model parameters. In this work we propose an optimization-based approach to parameter identification for fractional models with an optional truncation radius. We formulate the inference problem as an optimal control problem where the objective is to minimize the discrepancy between observed data and an approximate solution of the model, and the control variables are the fractional order and the truncation length. For the numerical solution of the minimization problem we propose a gradient-based approach, where we enhance the numerical performance by an approximation of the bilinear form of the state equation and its derivative with respect to the fractional order. Several numerical tests in one and two dimensions illustrate the theoretical results and show the robustness and applicability of our method.

1. INTRODUCTION

Nonlocal models provide an improved predictive capability for several scientific and engineering applications thanks to their ability to describe effects that classical partial differential equations (PDEs) fail to capture. These effects include multiscale and anomalous behavior such as super- and sub-diffusion in, e.g., sub-surface modeling [7, 35, 43, 44] or sharp interface dynamics in phase field models [10, 26, 25]. Furthermore, applications in material science, corrosion, turbulent

¹ COMPUTER SCIENCE AND MATHEMATICS DIVISION, OAK RIDGE NATIONAL LABORATORY, USA

² CENTER FOR COMPUTING RESEARCH, SANDIA NATIONAL LABORATORIES, NM, USA

³ COMPUTATIONAL SCIENCE AND ANALYSIS, SANDIA NATIONAL LABORATORIES, CA, USA

E-mail address: burkovskao@ornl.gov, caglusa@sandia.gov, mdelia@sandia.gov.

Key words and phrases. Identification problem, optimal control, nonlocal diffusion, fractional Laplacian, affine approximation.

This manuscript has been authored in part by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE) and supported by the U.S. Department of Energy, Office of Advanced Scientific Computing Research, Applied Mathematics Program under the award numbers ERKJ345 and ERKJE45. The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

This work was partially supported by the Sandia National Laboratories (SNL) Laboratory-directed Research and Development program and by the U.S. Department of Energy, Office of Advanced Scientific Computing Research under the Collaboratory on Mathematics and Physics-Informed Learning Machines for Multiscale and Multiphysics Problems (PhILMs) project. SNL is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in this paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

flow, and geoscience exhibit hierarchical features that do not adhere to classical diffusive behavior. Fractional order models are one of the most common instances of the nonlocal models. Through fractional exponent derivatives, as opposed to integer ones in the case of PDEs, fractional order operators can accurately represent multiscale complex phenomena by incorporating long range interactions in the model itself in the form of an integral operator acting over the whole space.

Given a scalar quantity $u : \mathbb{R}^n \rightarrow \mathbb{R}$, the simplest form of a fractional operator is the fractional Laplacian; in its integral form, its action on u is defined as [31]

$$(-\Delta)^s u(x) = C_{n,s} \int_{\mathbb{R}^n} \frac{u(x) - u(y)}{|x - y|^{n+2s}} dy, \quad (1.1)$$

where $s \in (0, 1)$ is the *fractional order* and $C_{n,s}$ is a constant that depends on the dimension n and s :

$$C_{n,s} = \frac{2^{2s} s \Gamma(s + n/2)}{\pi^{n/2} \Gamma(1 - s)}. \quad (1.2)$$

Here, and in what follows, the integral in (1.1) should be understood in the Cauchy principal value sense for $s \in [1/2, 1)$. It follows from (1.1) that in fractional modeling the state of a system at a point depends on the value of the state at any other point in the space; in other words, fractional models are nonlocal. Specifically, fractional operators are special instances of more general nonlocal operators [13, 17, 18, 24, 39] of the following form:

$$-\mathcal{L}u(x) = 2 \int_{B_\delta(x)} (u(x) - u(y)) \gamma(x, y) dy. \quad (1.3)$$

Here, interactions are limited to a Euclidean ball $B_\delta(x)$ of radius δ , often referred to as horizon or interaction radius. The symmetric nonnegative kernel¹ $\gamma(x, y)$ is a modeling choice and determines regularity properties of the solution. First, for $\delta = \infty$ and for the fractional type kernel $\gamma(x, y) = (C_{n,s}/2)|x - y|^{-n-2s}$ the nonlocal operator in (1.3) is equivalent to the fractional Laplacian in (1.1); second, it has been shown [18] that for that choice of δ and γ solutions corresponding to the nonlocal operator (1.3) converge to the ones corresponding to the fractional operator (1.1) as $\delta \rightarrow \infty$. In this paper we consider nonlocal operators with finite-range interactions and characterized by fractional type kernels and we refer to them as *truncated fractional operators* (see [17] for a detailed classification of these operators and relationships between them).

The highly descriptive power of these operators comes at the price of several modeling and computational challenges, including the unresolved treatment of nonlocal interfaces [3, 11]; the prescription of nonlocal analogues of boundary conditions [12, 23, 34]; the increasing computational cost as the extent of the nonlocal interactions increases [1, 2, 14, 21, 42, 45, 49], i.e. as $\delta \rightarrow \infty$; and the uncertainty and sparsity of model parameters and data [5, 16, 19, 20, 22, 30, 33, 40, 41, 50]. In this work we focus on the latter challenge and design an optimization-based data-driven parameter identification strategy.

Parameter identification problems for nonlocal and fractional operators have been addressed in the literature both with an optimal control approach and a machine learning approach. The literature on uncertainty quantification in the context of parameter estimation is still limited, [41] being perhaps the only relevant paper. In several works the identification problem is formulated as an optimal control problem where the objective functional is the mismatch between the solution and a given target and the constraints are the nonlocal equations. As an example, in

¹For a discussion regarding nonpositive kernels we refer the reader to [36] and for non-symmetric kernels to [15, 28].

[20] the authors learn a variable diffusion coefficient characterizing the kernel γ for a class of nonlocal operators that includes truncated fractional operators. An approach similar to the one presented in this work for the estimation of the fractional order of the *spectral* fractional Laplacian [37] has been presented in [46] and further analyzed in [4]. Based on these works, applications to image processing have also been considered [6]. On the other hand, several other works tackled the identification problem with a machine learning approach; as an example, in [39, 40] the authors employ physics-informed neural networks to describe the nonlocal solution and learn model parameters such as s and δ through minimization of a loss function given by the solution mismatch and the residual of the state equation. A similar approach, referred to as *operator regression* and also based on minimization of the equation residual, is utilized in [50] where the authors learn nonlocal kernels, including fractional type kernels, in a least-squares regression setting. Due to the complexity of the functionals involved in their formulations, the machine learning approaches mentioned above do not provide any theoretical analysis, which is, instead, a central aspect of our work.

Our paper is aligned with the former set of approaches and formulates the learning strategy as an optimal control problem where the controls are the model parameters s and δ , the state equation is a Poisson problem for the truncated fractional Laplacian and the cost functional is a matching functional, i.e. a measure of the difference between the nonlocal solution and a given target. To the best of our knowledge, this is the first work that analyzes the optimal control problem with respect to the fractional order and the truncation length of the *integral* truncated fractional Laplacian. We also point out that our analysis holds for the case of infinite truncation length, for which the truncated nonlocal operator coincides with the fractional Laplacian itself; this was also missing in the literature. Additionally, we provide and analyze efficient approximations of the bilinear form induced by the variational formulation of the problem and of its derivatives, with the goal of improving the computational performance of the optimization algorithm. These approximations not only are beneficial in our optimal control setting, but have the potential to improve the performance of any algorithm that involves the evaluation of bilinear forms and corresponding derivatives when parameters, such as s and δ , vary throughout the algorithm. More specifically, our main contributions are:

- 1) the rigorous analysis of regularity properties of the bilinear form of the state equations and of the control-to-state map;
- 2) the formulation of a fractional-equation-constrained optimization problem and the derivation of optimality conditions;
- 3) the rigorous analysis, including error estimates, of efficient approximations of the bilinear form and its derivatives with respect to the fractional order (allowing for affine decompositions) and the truncation length;
- 4) the design of a gradient-based algorithm for the solution of the optimization problem, including the construction of an approximate gradient (based on 3) and its error analysis, that allows for fast computations and a complexity study for the resulting discretized formulation.

1.1. Outline of the paper. The paper is organized as follows. In Section 2 we introduce rigorous notation and recall relevant theoretical results. In Section 3 we analyze regularity properties of the bilinear form of the state equations with respect to the model parameters δ and s and of the control-to-state map. In Section 4 we introduce the optimal control problem, analyze the existence of solutions and propose an adjoint-based approach for its solution. In Section 5 we introduce approximations of the bilinear form and its derivatives and analyze their accuracy. In Section 6 we discuss the discretization of state and adjoint problems. We analyze

the convergence subject to discretization and interpolation error of state, adjoint and gradient of the cost functional. In Section 7 we illustrate our theoretical results with several one- and two-dimensional computational tests. Finally, in Section 8 we report concluding remarks and provide insights on future research directions.

2. PRELIMINARIES

Let $\Omega \subset \mathbb{R}^n$, $n \geq 1$ be a bounded domain with C^∞ -boundary $\partial\Omega$. We introduce a parameter $q = (s, \delta) \in Q$, where

$$Q := \{q \in \mathbb{R}^2 : 0 < s < 1, \ 0 < \delta < \infty\},$$

and we consider the parametrized nonlocal operator

$$-\mathcal{L}(q)u(x) := 2 \int_{\mathbb{R}^n} (u(x) - u(y))\gamma(x, y; q) \, dy, \quad x \in \mathbb{R}^n, \quad (2.1)$$

with a fractional type kernel function defined as

$$\gamma(x, y; q) := \frac{\mathcal{X}_{B_\delta(x)}(y)}{|x - y|^{n+2s}}. \quad (2.2)$$

Here, $\mathcal{X}_{B_\delta(x)}$ denotes the characteristic function over the Euclidean ball $B_\delta(x) := \{y \in \mathbb{R}^n : |x - y| \leq \delta\}$, the parameter δ defines the extent of the nonlocal interactions, and s represents the fractional order of the truncated fractional operator. We also introduce a vector $h := [h_s, h_\delta]^T \in Q$, and by a slight abuse of notation we denote by $|h|^2 := \|h\|^2 = |h_s|^2 + |h_\delta|^2$ the usual Euclidean norm.

Remark 2.1. *For convenience of notation we omit the scaling constant $C_{n,s}$ in (2.2), and note that the approaches developed in this work can be directly extended to the case of $C_{n,s}$ defined in (1.2).*

To capture nonlocal interactions outside of Ω , we introduce the so-called interaction domain $\Omega_\delta := \{y \in \mathbb{R}^n : \exists x \in \Omega : |x - y| < \delta\}$, and in terms of this notation the operator in (2.1) can also be expressed as

$$-\mathcal{L}(q)u(x) = 2 \int_{\Omega \cup \Omega_\delta} (u(x) - u(y))\gamma(x, y; q) \, dy = 2 \int_{(\Omega \cup \Omega_\delta) \cap B_\delta(x)} \frac{u(x) - u(y)}{|x - y|^{n+2s}} \, dy.$$

We consider the following nonlocal problem with homogeneous Dirichlet-type volume constraints:

$$\begin{aligned} -\mathcal{L}(q)u(x) &= f(x), & x \in \Omega, \\ u(x) &= 0, & x \in \Omega_\delta. \end{aligned} \quad (2.3)$$

For convenience of notation we suppose that u is extended by zero outside of $\Omega \cup \Omega_\delta$, and, by a slight abuse of notation, we still denoted it by u . When necessary, we chose the whole \mathbb{R}^n as a common spatial domain. In what follows we often use the relation $a \lesssim b$, implying that $a \leq Cb$, where $C > 0$ is some constant that may be different in every case.

2.1. Functional spaces. We denote by $L^2(\tilde{\Omega})$ the Lebesgue space of square-integrable functions and by $H^1(\tilde{\Omega})$ and $H_0^1(\tilde{\Omega})$ the usual Sobolev spaces, where $\tilde{\Omega}$ is given either by Ω , $\Omega \cup \Omega_\delta$ or \mathbb{R}^n . For $u, v \in L^2(\mathbb{R}^n)$ we define the bilinear form

$$a(u, v) := \int_{S_\delta} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{n+2s}} \, dy \, dx, \quad (2.4)$$

where S_δ is the strip

$$S_\delta = \{(x, y) \in \mathbb{R}^{2n} : |x - y| \leq \delta\}, \quad (2.5)$$

and for $s \in (0, 1)$, $0 < \delta \leq \infty$ we introduce the following nonlocal function spaces

$$V_\delta = \{v \in L^2(\Omega \cup \Omega_\delta) : a(v, v) < \infty, \ v = 0 \text{ a.e. on } \Omega_\delta\},$$

equipped with the inner product $(u, v)_{V_\delta} := a(u, v)$ and norm $\|u\|_{V_\delta}^2 = (u, u)_{V_\delta}$. Next, we define the fractional Sobolev spaces of order $s \in (0, 1)$

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

where the Gagliardo seminorm is defined as

$$|v|_{H^s(\Omega)}^2 := \int_{\Omega} \int_{\Omega} \frac{|v(x) - v(y)|^2}{|x - y|^{n+2s}} dy dx.$$

For $s > 1$ not an integer, we define $H^s(\Omega)$, $s = m + \sigma$ with $m \in \mathbb{N}$ and $\sigma \in (0, 1)$, as

$$H^s(\Omega) := \{v \in H^m(\Omega) : D^\alpha v \in H^\sigma(\Omega) \text{ for } |\alpha| = m\},$$

together with the semi-norm $|v|_{H^s(\Omega)}^2 = |v|_{H^m(\Omega)}^2 + \sum_{|\alpha|=m} |D^\alpha v|_{H^\sigma(\Omega)}^2$. The space $H^s(\Omega)$ is a Hilbert space, endowed with the norm $\|v\|_{H^s(\Omega)}^2 = \|v\|_{L^2(\Omega)}^2 + |v|_{H^s(\Omega)}^2$. Additionally, we define the space incorporating the volume constraints given by

$$H_\Omega^s(\mathbb{R}^n) := \{v \in H^s(\mathbb{R}^n) : v = 0 \text{ on } \mathbb{R}^n \setminus \Omega\},$$

that is endowed with the semi-norm of $H^s(\mathbb{R}^n)$, and by $(\cdot, \cdot)_{H_\Omega^s(\mathbb{R}^n)}$ we denote the corresponding inner product. For negative exponents, we define the corresponding spaces by duality $H^{-s}(\Omega) = (H_\Omega^s(\mathbb{R}^n))'$.

We point out that for $0 < \delta \leq \infty$ the nonlocal spaces V_δ are all equivalent, see [9, Proposition 2.2]. Furthermore, for V_δ is equivalent to $H_\Omega^s(\mathbb{R}^n)$, which implies that we can equivalently work either with V_δ or $H_\Omega^s(\mathbb{R}^n)$.

We also recall that for all $v \in H_\Omega^s(\mathbb{R}^n)$ and $s \in (0, 1)$ the following nonlocal Poincaré inequality holds

$$\|v\|_{L^2(\Omega)} \leq C_P \|v\|_{H_\Omega^s(\mathbb{R}^n)}, \quad C_P > 0. \quad (2.6)$$

2.2. Variational formulation. We consider the following weak formulation of problem (2.3). For $q \in Q$, $f \in H^r(\Omega)$, $r \geq -s$, $\varepsilon > 0$, find $u(q) \in H_\Omega^s(\mathbb{R}^n)$, such that

$$a(u, v; q) = (f, v)_{L^2(\Omega)}, \quad \forall v \in H_\Omega^s(\mathbb{R}^n). \quad (2.7)$$

The existence and uniqueness of the solution of (2.7) follows by the Lax-Milgram lemma. Moreover, the following improved regularity result holds true [8, 32].

Theorem 2.1. *Let Ω be a domain with C^∞ boundary $\partial\Omega$, and let $f \in H^r(\Omega)$, $r \geq -s$, and let $u \in H_\Omega^s(\mathbb{R}^n)$ be the solution of (2.7) with the kernel (2.2) for $\delta > 0$. Then, for any $\varepsilon > 0$ the following holds*

$$\|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \lesssim \|f\|_{H^r(\Omega)}, \quad (2.8)$$

where $\alpha = \min\{s + r, 1/2 - \varepsilon\}$.

Assumption 2.2. *We let the forcing term f be such that $f \in H^{1/2-\varepsilon}(\Omega)$.*

Note that Assumption 2.2 implies that $u \in H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)$, $\varepsilon > 0$; this is the highest regularity for the solution u .

To conclude this section, we introduce a solution operator $S : Q \rightarrow H_\Omega^s(\mathbb{R}^n) \subset L^2(\Omega)$, such that $u(q) = S(q)$ is a unique solution of (2.7), i.e., $S(q)$ is such that

$$a(S(q), v; q) = (f, v)_{L^2(\Omega)}, \quad \forall v \in H_\Omega^s(\mathbb{R}^n). \quad (2.9)$$

3. REGULARITY OF THE BILINEAR FORM AND OF THE CONTROL-TO-STATE MAP

In this section we recall from [9] the parametric regularity results with respect to s and δ for the bilinear form associated with the state equation and then analyze properties of the control-to-state map.

Proposition 3.1 (δ -regularity). *For $\delta \in (0, \infty)$, $s \in (0, 1)$, and γ defined in (2.2), the following hold true for all $u, v \in H_\Omega^s(\mathbb{R}^n)$:*

(i) *the bilinear form $a(\cdot, \cdot; \delta)$ is differentiable with respect to δ with*

$$a'_\delta(u, v; \delta) = \int_{\mathbb{R}^n} \int_{\partial B_\delta(x)} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{n+2s}} dy dx, \quad (3.1)$$

where $\partial B_\delta(x)$ is the surface of the ball of radius δ centered at x , i.e. $\partial B_\delta(x) = \{y \in \mathbb{R}^n : |x - y| = \delta\}$, and the inner integral is understood as a surface integral;

(ii) *the derivative $a'_\delta(\cdot, \cdot; \delta)$ is bounded*

$$|a'_\delta(u, v; \delta)| \lesssim \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}; \quad (3.2)$$

(iii) *the derivative $a'_\delta(\cdot, \cdot; \delta)$ is Hölder continuous, i.e., for $\theta \in [0, 1]$ and $\delta_1, \delta_2 \in (0, \infty)$ it holds*

$$|a'_\delta(u, v; \delta_1) - a'_\delta(u, v; \delta_2)| \lesssim |\delta_1 - \delta_2|^\theta \|u\|_{H_\Omega^\theta(\mathbb{R}^n)} \|v\|_{L^2(\Omega)}, \quad (3.3)$$

where $H_\Omega^0(\mathbb{R}^n) \equiv L^2(\Omega)$ and $H_\Omega^1(\mathbb{R}^n) \equiv H_0^1(\Omega)$.

Proof. We only need to prove the last bound, since the first two statements follow directly from [9, Theorem 3.3]. For $\theta = 1$ and $\theta = 0$, assertion (iii) has been proven in [9, Theorem 3.3]. By applying the interpolation theorem we obtain the desired estimate for $\theta \in (0, 1)$. \square

Proposition 3.2 (s -regularity). *Let $\delta \in (0, \infty)$ and $s \in (0, 1)$. Then, for $u \in H_\Omega^{s+\alpha}(\mathbb{R}^n)$, $0 < \alpha < 1 - s$, $v \in H_\Omega^s(\mathbb{R}^n)$, and γ defined in (2.2), the following hold true:*

(i) *the bilinear form $a(u, v; s)$ is infinitely many times differentiable, i.e., for $k = 1, 2, \dots$, it holds*

$$a_s^{(k)}(u, v; s) := (-2)^k \int_{S_\delta} \frac{(u(x) - u(y))(v(x) - v(y)) \log^k(|x - y|)}{|x - y|^{n+2s}} dy dx; \quad (3.4)$$

(ii) *the derivative $a_s^{(k)}(u, v; s)$ is bounded, i.e., for all $s_1, s_2 \in (0, 1)$, $s \in (0, (s_1 + s_2)/2) \subset (0, 1)$ it holds*

$$|a_s^{(k)}(u, v; s)| \leq C(k, \hat{\varepsilon}) \|u\|_{H_\Omega^{s_2}(\mathbb{R}^n)} \|v\|_{H_\Omega^{s_1}(\mathbb{R}^n)}, \quad (3.5)$$

where $C(k, \hat{\varepsilon}) = 2^k ((k/(e\hat{\varepsilon}))^k + \delta^{\hat{\varepsilon}} (\log(\delta))^k_+)$, $\hat{\varepsilon} = s_1 + s_2 - 2s > 0$;

(iii) *the derivative $a_s^{(k)}(u, v; s)$ is locally Lipschitz continuous:*

$$|a_s^{(k)}(u, v; s) - a_s^{(k)}(u, v; s + \eta)| \lesssim \eta \|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)}, \quad (3.6)$$

where $0 < \alpha < 1 - s$, $0 < \eta < \alpha/2$ and the hidden constant depends on Ω , and continuously on s and α .

Proof. Since, the first two statements follow directly from [9, Lemma 5.1], it remains to prove the validity of (3.6). Using the results of [9, Lemma 5.1] with $s_2 = s + \alpha$ and $s_1 = s$, we obtain

$$\begin{aligned} |a_s^{(k)}(u, v; s) - a_s^{(k)}(u, v; s + \eta)| &\leq \eta \max_{\xi \in [s, s+\eta]} |a_s^{(k+1)}(u, v; \xi)| \\ &\lesssim \eta \|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)}, \end{aligned}$$

which concludes the proof. \square

Combining the results of Proposition 3.1 and Proposition 3.2 we derive the following estimates for the mixed derivatives.

Proposition 3.3 (Mixed derivatives). *For $\delta \in (0, \infty)$, $s \in (0, 1)$, and γ defined in (2.2), the following hold true for all $u \in H_\Omega^{s+\alpha}(\mathbb{R}^n)$, $0 < \alpha < 1 - s$, $v \in H_\Omega^s(\mathbb{R}^n)$:*

(i) *there exists a second mixed derivative of $a(\cdot, \cdot; \delta)$ with respect to δ and s :*

$$\begin{aligned} a''_{s,\delta}(u, v; s, \delta) &:= \frac{\partial^2}{\partial s \partial \delta} (a(u, v; s, \delta)) \\ &= -2 \int_{\mathbb{R}^n} \int_{\partial B_\delta(x)} \frac{(u(x) - u(y))(v(x) - v(y)) \log(|x - y|)}{|x - y|^{n+2s}} dy dx, \end{aligned} \quad (3.7)$$

where $\partial B_\delta(x)$ is defined as in Proposition 3.1.

(ii) *the mixed derivative $a''_{s,\delta}(\cdot, \cdot; s, \delta)$ is bounded*

$$|a''_{s,\delta}(u, v; s, \delta)| \lesssim \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}, \quad (3.8)$$

with the hidden constant depending continuously on s and δ .

Next, we analyze the properties of the map $S : Q \rightarrow L^2(\Omega)$. The following proposition shows that S is Lipschitz continuous.

Proposition 3.4 (Lipschitz continuity of $S(q)$). *Let $\delta \in (0, \infty)$ and $s \in (0, 1)$. Then, for any $q \in Q$, $S(q) \in H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)$. Moreover, for $0 < \alpha < \min\{(1-s)/2 - \varepsilon, s\}$, $|h| < \alpha/2$, and γ defined in (2.2), $S(q)$ is locally Lipschitz continuous:*

$$\|S(q) - S(q+h)\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \lesssim |h| \|S(q+h)\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)}. \quad (3.9)$$

Proof. We note that the first statement $S(q) \in H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)$ follows directly from Theorem 2.1. Then, for $q, h \in Q$ using (2.9) we can express

$$\begin{aligned} a(S(q) - S(q+h), v; q) &= a(S(q), v; q) - a(S(q+h), v; q) \pm a(S(q+h)v; q+h) \\ &= a(S(q+h), v; q+h) - a(S(q+h), v; q). \end{aligned}$$

Let $\tilde{q}(\tau) = q + \tau h$, $\tau \in (0, 1)$, and $g(\tau) := a(S(\tilde{q}(1)), v; \tilde{q}(\tau))$, then we obtain that $g'(\tau) = a'_q(S(q+h), v; q + \tau h)(h)$ and the following holds

$$\begin{aligned} a(S(q+h), v; q+h) - a(S(q+h), v; q) &= g(1) - g(0) = \int_0^1 g'(\xi) d\xi \\ &= \int_0^1 a'_q(S(q+h), v; q + \xi h)(h) d\xi. \end{aligned}$$

Using (3.6) with $2\alpha - 1/2 + \varepsilon \leq h_s < \alpha/2$, where $0 < \alpha < \min\{(1-s)/2 - \varepsilon, s\}$, $\varepsilon > 0$, and (3.2) with (2.6) we obtain

$$\begin{aligned} &|a'_q(S(q+h), v; q + \xi h)(h)| \\ &\leq |h_s| |a'_s(S(q+h), v; q + \xi h)| + |h_\delta| |a'_\delta(S(q+h), v; q + \xi h)| \\ &\lesssim |h_s| \|S(q+h)\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^{s-\alpha}(\mathbb{R}^n)} + |h_\delta| \|S(q+h)\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\lesssim |h| \|S(q+h)\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^{s-\alpha}(\mathbb{R}^n)}. \end{aligned}$$

We note that under the above conditions on h_s and α , we have that $S(q+h) \in H_\Omega^{s+h_s+1/2-\varepsilon}(\mathbb{R}^n) \subset H_\Omega^{s+2\alpha}(\mathbb{R}^n)$, and hence $\|u\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)}$ is well-defined. Furthermore, if we introduce $F(v) := \int_0^1 a'_q(S(q+h), v; q + \xi h)(h) d\xi$, then the above also

implies that $F \in (H_\Omega^{s-\alpha}(\mathbb{R}^n))^* = H^{-s+\alpha}(\Omega)$. Now, applying Theorem 2.1 to the problem

$$a(S(q) - S(q+h), v; q) = F(v), \quad \forall v \in H_\Omega^s(\mathbb{R}^n), \quad (3.10)$$

we conclude that $S(q) - S(q+h) \in H_\Omega^{s+\beta}(\mathbb{R}^n)$, where $\beta = \min\{\alpha, 1/2 - \varepsilon\}$. Since, $\alpha < (1-s)/2 - \varepsilon$, it follows that $\beta = \alpha$ and from (2.8) we obtain

$$\|S(q) - S(q+h)\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \lesssim \|F\|_{H^{-s+\alpha}(\Omega)} \lesssim |h| \|S(q+h)\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)}$$

that concludes the proof. \square

We now investigate regularity properties. By differentiating the state equation (2.7) with respect to $q \in Q$ in the direction $h \in Q$, and taking into account that $a'_u(u, v; q)(hu) = a(hu, v; q)$ for any direction $hu \in H_\Omega^s(\mathbb{R}^n)$, we obtain the following *sensitivity equation*:

$$a(hu, v; q) = -a'_q(u, v; q)(h) \quad \forall v \in H_\Omega^s(\mathbb{R}^n). \quad (3.11)$$

Next, we establish the differentiability of the solution map $S : Q \rightarrow L^2(\Omega)$.

Theorem 3.5. *The map $S : Q \rightarrow L^2(\Omega)$ is Fréchet differentiable, i.e., for $q \in Q$ and $h \in Q$ we have*

$$\lim_{|h| \rightarrow 0} \frac{\|S(q+h) - S(q) - S'(q)(h)\|_{H_\Omega^s(\mathbb{R}^n)}}{|h|} = 0.$$

Proof. Let $z(q) := S(q+h) - S(q) - S'(q)(h)$, then using the similar arguments as in the proof of Proposition 3.4, and the sensitivity equation (3.11) with $hu := S'(q)(h)$, Lipschitz continuity and boundedness of $a'_q(\cdot, \cdot; q)$, we obtain:

$$\begin{aligned} a(z(q), v; q) &= a(S(q+h) - S(q) - S'(q)(h), v; q) \pm a(S(q+h), v; q+h) \\ &= a(S(q+h), v; q) - a(S(q+h), v; q+h) - a(S'(q)(h), v; q) \\ &\quad - \int_0^1 a'_q(S(q+h), v; q + \xi h)(h) d\xi + a'_q(S(q), v; q)(h) \\ &= \int_0^1 [a'_q(S(q+h), v; q) - a'_q(S(q+h), v; q + \xi h)](h) d\xi \\ &\quad + a'_q(S(q) - S(q+h), v; q)(h). \end{aligned} \quad (3.12)$$

Let $u := S(q+h)$, then using (3.6) with $\alpha - 1/2 + \varepsilon \leq h_s < \alpha/2$, where $0 < \alpha < \min\{1-s, 1/2 - \varepsilon\}$, $\varepsilon > 0$, and taking into account that $|h| \geq |h_s|$ we obtain

$$\begin{aligned} &|(a'_s(u, v; q) - a'_s(u, v; q + \xi h))(h_s)| \\ &\leq |a'_s(u, v; s, \delta) - a'_s(u, v; s + \xi h_s, \delta + \xi h_\delta) \pm a'_s(u, v; s + \xi h_s, \delta)| |h_s| \\ &\leq |a'_s(u, v; s, \delta) - a'_s(u, v; s + \xi h_s, \delta)| |h_s| \\ &\quad + |a'_s(u, v; s + \xi h_s, \delta) - a'_s(u, v; s + \xi h_s, \delta + \xi h_\delta)| |h_s| \\ &\lesssim \xi |h_s|^2 \|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)} + |h_s| \left| \int_{\delta+\xi h_\delta}^\delta a''_{s,\delta}(u, v; s + \xi h_s, \eta) d\eta \right| \\ &\lesssim \xi |h_s|^2 \|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)} + \xi |h_s| |h_\delta| \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\lesssim |h|^2 \|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)}, \end{aligned}$$

where in the last inequality, under the restrictions on h_s and α , stated above, we ensure that $u \in H_\Omega^{s+h_s+1/2-\varepsilon}(\mathbb{R}^n) \subset H_\Omega^{s+\alpha}(\mathbb{R}^n)$, and hence $\|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} < \infty$. In a

similar way as above using (3.3) with $\theta = s$, we obtain the estimate with respect to δ :

$$\begin{aligned}
& |(a'_\delta(u, v; q) - a'_\delta(u, v; q + \xi h))(h_\delta)| \\
& \leq |a'_\delta(u, v; s, \delta) - a'_\delta(u, v; s, \delta + \xi h_\delta)| |h_\delta| \\
& + |a'_\delta(u, v; s, \delta + \xi h_\delta) - a'_\delta(u, v; s + \xi h_s, \delta + \xi h_\delta)| |h_\delta| \\
& \lesssim \xi^s |h_\delta|^{s+1} \|u\|_{H_\Omega^s(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)} + |h_\delta| \left| \int_{s+\xi h_s}^s a''_{s,\delta}(u, v; \eta, \delta + \xi h_\delta) d\eta \right| \\
& \lesssim |h|^{1+s} \|u\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)}.
\end{aligned}$$

Using the boundedness of $a'_q(\cdot, \cdot; q)$, (3.2), (3.5), and the Lipschitz continuity of the map S (3.9), with further imposing additional conditions on h_s and α such as in Proposition 3.4, we obtain the following:

$$\begin{aligned}
& |a'_q(S(q) - S(q+h), v; q)(h)| \\
& \leq |a'_s(S(q) - S(q+h), v; q)| |h_s| + |a'_\delta(S(q) - S(q+h), v; q)| |h_\delta| \\
& \lesssim |h| \|S(q) - S(q+h)\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim |h|^2 \|S(q+h)\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)} \|v\|_{H_\Omega^s(\mathbb{R}^n)}.
\end{aligned}$$

Now, taking $v = S(q+h) - S(q) - S'(q)(h)$, and using the coercivity of $a(\cdot, \cdot; q)$, and combining the above estimates we obtain

$$\begin{aligned}
& \|S(q+h) - S(q) - S'(q)(h)\|_{H_\Omega^s(\mathbb{R}^n)} \\
& \lesssim |h| \left(\|h\| \|S(q+h)\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} + |h|^s \|S(q+h)\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} + |h| \|S(q+h)\|_{H_\Omega^{s+2\alpha}(\mathbb{R}^n)} \right).
\end{aligned}$$

Dividing the left-hand side by $|h|$, and taking the limit $|h| \rightarrow 0$ leads to the desired result. \square

4. OPTIMAL CONTROL PROBLEM

In this section we introduce the optimal control problem and analyze the existence of solutions. We augment the matching functional mentioned in the introduction with a regularization term that ensures that the parameters do not attain the values outside of the admissible set. Then, we propose a gradient-based approach that relies on the solution of an adjoint problem and that sets the groundwork for the solution of higher-dimensional inverse problems.

4.1. The minimization problem. We define the cost functional as follows:

$$J(u, q) := \frac{1}{2} \|u - u_d\|_{L^2(\Omega)}^2 + \mathcal{R}(q), \quad (4.1)$$

where $u_d \in H^{1/2-\varepsilon}(\Omega)$ is a given function and where $\mathcal{R} : Q \rightarrow \mathbb{R}_+$ is a regularization term. In particular, $\mathcal{R}(q) \in C^2(Q)$ is such that $\mathcal{R}(q) := \phi_1(s) + \phi_2(\delta)$ with ϕ_i , $i = 1, 2$, being given nonnegative convex functions that satisfy

$$\lim_{s \rightarrow 0} \phi_1(s) = +\infty = \lim_{s \rightarrow 1} \phi_1(s), \quad \lim_{\delta \rightarrow 0} \phi_2(\delta) = +\infty = \lim_{\delta \rightarrow \infty} \phi_2(\delta). \quad (4.2)$$

As such, \mathcal{R} prevents the fractional order s to be either 0 or 1. Note that such choice of regularization has been also employed in [4, 46]. We formulate the optimal control problem as

$$\min_{q \in Q} J(u, q) \quad \text{subject to (2.7)}. \quad (4.3)$$

By using the control-to-state map we define a reduced cost functional $j(q) = J(S(q), q)$ and restate (4.3) as an unconstrained minimization problem:

$$\min_{q \in Q} j(q) := \frac{1}{2} \|S(q) - u_d\|_{L^2(\Omega)}^2 + \mathcal{R}(q). \quad (4.4)$$

Proposition 3.4 and properties of the cost functional imply the existence of solutions to the optimization problem, as shown in the following proposition.

Proposition 4.1 (Existence). *There exists a solution to the minimization problem (4.4).*

Proof. Since, $j(q) \geq 0$, it follows that there exists a minimizing sequence $\{q_n\}_{n \in \mathbb{N}} \subset Q$, such that $j(q_n) \rightarrow \inf_{q \in Q} j(q)$, $n \rightarrow \infty$. Using the definition of \mathcal{R} and j , we obtain

$$\mathcal{R}(q_n) \leq \frac{1}{2} \|S(q_n) - u_d\|_{L^2(\Omega)}^2 + \mathcal{R}(q_n) \leq C, \quad C > 0,$$

and hence $\{q_n\}$ is bounded and we can extract a weakly-convergent sub-sequence $\{q_{n_k}\}_{n_k \in \mathbb{N}}$, such that $q_{n_k} \rightharpoonup \hat{q} \in \overline{Q}$, $k \rightarrow \infty$. Next, we argue that $\hat{q} \in Q$. Indeed, by continuity we have that $\mathcal{R}(\hat{q}) \leq C$, then if $\hat{q} \in \overline{Q} \setminus Q$, it follows, by the definition of \mathcal{R} , that $\mathcal{R}(q) = +\infty$, that contradicts the previous statement, and, hence $\hat{q} \in Q$. By the continuity of \mathcal{R} and S , and Proposition 3.4, it follows that $S(q_{n_k}) \rightarrow S(\hat{q})$ in $L^2(\Omega)$ as $k \rightarrow \infty$, and we obtain that $j(q_{n_k}) \rightarrow j(\hat{q})$, $k \rightarrow \infty$, and

$$j(\hat{q}) = \lim_{k \rightarrow \infty} j(q_{n_k}) = \inf_{q \in Q} j(q),$$

thus, q is the minimizer of (4.4). \square

4.2. Optimality conditions. Having established the existence of $S'(q)(h)$ for all $h \in Q$ in Theorem 3.5, by the chain rule we immediately deduce that $j(q)$ is Gâteaux differentiable, and the directional derivative $j'(q)(h)$ can be expressed as follows:

$$j'(q)(h) = (S(q) - u_d, hu)_{L^2(\Omega)} + \mathcal{R}'(q), \quad (4.5)$$

with $hu := S'(q)(h)$.

Proposition 4.2 (First order necessary optimality condition). *If $q \in Q$ is a local optimal solution of the optimization problem (4.4), then there holds the first order necessary optimality condition:*

$$j'(q)(h) = 0, \quad \forall h \in Q. \quad (4.6)$$

4.3. Adjoint approach. In this section we provide a computable representation of the derivative of $j(q)$, which is necessary for evaluating the optimality conditions stated in the previous section, and for employing a gradient-based optimization algorithm.

For a small dimensional control space, the most obvious choice to evaluate the derivative $j'(q)$ is the *sensitivity approach*, where the sensitivity needed to evaluate $j'(q)$ in (4.6), is obtained as a solution of the sensitivity equation (3.11). While in the present case, having only a two-dimensional control, would lead us to use the aforementioned procedure, to keep our approach general and set the groundwork for higher-dimensional problems, we resort to the so-called *adjoint method*.

We define the Lagrangian $\mathcal{L} : Q \times H_\Omega^s(\mathbb{R}^n) \times H_\Omega^s(\mathbb{R}^n) \rightarrow \mathbb{R}$ of the optimization problem (4.4):

$$\mathcal{L}(q, u, z) := J(u, q) + (f, v)_{L^2(\Omega)} - a(u, z; q). \quad (4.7)$$

Then, the following identity holds true for $u = S(q)$ and all $\phi \in H_\Omega^s(\mathbb{R}^n)$:

$$j(q) = J(u, q) = \mathcal{L}(q, u, \phi),$$

and for all $h \in Q$ we obtain

$$j'(q)(h) = \mathcal{L}'_q(q, u, z)(h) + \mathcal{L}'_u(q, u, z)(hu), \quad (4.8)$$

with $hu := S'(q)(h)$. Then the adjoint state $z \in H_\Omega^s(\mathbb{R}^n)$ is defined as a solution of the adjoint equation

$$\mathcal{L}'_u(q, u, z)(\phi) = 0, \quad \forall \phi \in H_\Omega^s(\mathbb{R}^n). \quad (4.9)$$

Taking into account that $\mathcal{L}'_u(q, u, z)(hu) = J'_u(u, q)(hu) - a(hu, z; q)$, the adjoint equation (4.9) can be written in the following explicit form:

$$a(\phi, z; q) = (u - u_d, \phi)_{L^2(\Omega)}, \quad \forall \phi \in H^s_\Omega(\mathbb{R}^n). \quad (4.10)$$

Combining (4.10) and (4.8), we obtain the following expression for $j'(q)$:

$$j'(q) = \mathcal{L}'_q(q, u, z) = \mathcal{R}'(q) - a'_q(u, z; q). \quad (4.11)$$

Invoking Theorem 2.1 for the state solution u , we obtain that $u - u_d \in H^{1/2-\varepsilon}_\Omega(\mathbb{R}^n)$. Then, following a boot-strapping procedure, similar to [8, Theorem 3.4], it can be shown the solution of the adjoint equation (4.10) admits an improved regularity $z \in H^{s+1/2-\varepsilon}_\Omega(\mathbb{R}^n)$, and the following holds

$$\|z\|_{H^{s+1/2-\varepsilon}_\Omega(\mathbb{R}^n)} \lesssim \|u - u_d\|_{H^{1/2-\varepsilon}(\Omega)} \lesssim \|f\|_{H^{1/2-\varepsilon}(\Omega)} + \|u_d\|_{H^{1/2-\varepsilon}(\Omega)}, \quad \varepsilon > 0. \quad (4.12)$$

5. APPROXIMATION OF THE BILINEAR FORM AND ITS DERIVATIVES

Based on the work introduced in [9], in this section we consider an approximation for the bilinear form based on Chebyshev interpolation. This allows for an affine decomposition of the bilinear form and its derivative with respect to the fractional order, which allows us to reduce the computational time needed for the assembly of these forms. At the end of the section we also introduce an approximation of the derivative with respect to the truncation length based on operator splitting.

5.1. Interpolation of the bilinear form. We let $\mathcal{S} = [s_{\min}, s_{\max}]$ and consider a subdivision of the interval \mathcal{S} into sub-intervals $\mathcal{S}_k := [s_{\min, k}, s_{\max, k}]$ with $1 \leq k \leq K$, $K \in \mathbb{N}$, and

$$\begin{aligned} s_{\min, 1} &= s_{\min}, & s_{\max, K} &= s_{\max}, \\ s_{\min, k} &= s_{\max, k-1}, & 2 \leq k &\leq K. \end{aligned} \quad (5.1)$$

Then, we approximate the parametrized bilinear form $a(\cdot, \cdot; \cdot, \cdot)$ with respect to s as follows:

$$a(u, v; s, \delta) \approx \tilde{a}(u, v; s, \delta) := \sum_{m=0}^{M_k} \Theta_{\mathcal{S}_k, m}(s) a(u, v; s_m, \delta),$$

where $\Theta_{\mathcal{S}_k, m}(s)$ are Lagrange polynomials with respect to the Chebyshev nodes on the interval \mathcal{S}_k .

Under a certain condition on the size of the interval, which is induced by the limited spatial regularity of the solution, it has been shown in [9] that the following error estimate for the interpolation of the bilinear form holds true.

Lemma 5.1 ([9]). *Let $u \in H^{s_{2,k}}_\Omega(\mathbb{R}^n)$, $v \in H^{s_{1,k}}_\Omega(\mathbb{R}^n)$ and $s \in \mathcal{S}_k = [s_{\min, k}, s_{\max, k}] \subset (0, 1)$, such that $(s_{1,k} + s_{2,k})/2 - 1/2 < s_{\max, k} < (s_{1,k} + s_{2,k})/2$, $s_{1,k}, s_{2,k} \in (0, 1)$. Then, for $\delta \in (0, \infty)$, we obtain*

$$|a(u, v; s, \delta) - \tilde{a}(u, v; s, \delta)| \leq \sigma_k^{M_k+1} C_k(\delta) \|u\|_{H^{s_{2,k}}_\Omega(\mathbb{R}^n)} \|v\|_{H^{s_{1,k}}_\Omega(\mathbb{R}^n)},$$

where $\sigma_k = (s_{\max, k} - s_{\min, k})/(2\hat{\varepsilon}(s_{\max, k}))$, $C_k(\delta) = 4(e^{-1} + \delta^{\hat{\varepsilon}(s_{\min, k})+1})$ if $\delta > 1$ and $C_k(\delta) = 4e^{-1}$ if $\delta \leq 1$, where $\hat{\varepsilon}(s) = s_{1,k} + s_{2,k} - 2s$.

Next, we investigate how to choose the intervals \mathcal{S}_k and the interpolation orders M_k in order to satisfy a tolerance $\eta > 0$.

Lemma 5.2. *Let $s \in [s_{\min}, s_{\max}] \subset (0, 1)$, $\delta \in (0, \infty)$, and let $\eta > 0$. Assume that $u(s) \in H_{\Omega}^{s+1/2-\varepsilon}(\mathbb{R}^n)$, $\varepsilon > 0$. Then, we define*

$$s_{\min,1} = s_{\min}, \quad s_{\max,K} = s_{\max}$$

$$s_{\max,k} = s_{\min,k} + (1/2 - \xi) \min\{1 - s_{\min,k}, \frac{1}{2} - \varepsilon\}, \quad 1 \leq k \leq K-1,$$

for some $\xi \in (1/10, 1/2)$. Furthermore, for $s \in \mathcal{S}_k$, we set

$$M_k = \left\lceil \frac{\log(\eta/C_k(\delta))}{\log \frac{1}{8}(\xi^{-2} - 2)} \right\rceil - 1.$$

Then, for a given $\xi \in (1/10, 1/2)$, $\bar{s}_2(s) := \min\{1, s_{\min,k} + \frac{1}{2}\} - \varepsilon$, $s \in \mathcal{S}_k$ and $v \in H_{\Omega}^s(\mathbb{R}^n)$, we obtain

$$|a(u(s), v; s, \delta) - \tilde{a}(u(s), v; s, \delta)| \leq \eta \|u(s)\|_{H_{\Omega}^{\bar{s}_2(s)}(\mathbb{R}^n)} \|v\|_{H_{\Omega}^s(\mathbb{R}^n)},$$

where the total number of interpolation nodes satisfies

$$\sum_{k=1}^K (M_k + 1) \leq C |\log \eta|,$$

and the constant C depends on ξ , δ and s_{\max} .

Proof. We will choose $s_{1,k}$, $s_{2,k}$, $s_{\min,k}$ and $s_{\max,k}$ so that we can apply Lemma 5.1 on each interval \mathcal{S}_k . Given $s_{\min,k}$ for some k , we set

$$s_{1,k} = s_{\min,k}, \quad s_{2,k} = \min\{1, s_{\min,k} + \frac{1}{2}\} - \varepsilon = s_{\min,k} + g_k,$$

with $g_k := \min\{1 - s_{\min,k}, \frac{1}{2}\} - \varepsilon$. We then choose

$$s_{\max,k} := \frac{s_{1,k} + s_{2,k}}{2} - \xi g_k = s_{\min,k} + \left(\frac{1}{2} - \xi\right) g_k.$$

In order to satisfy $s_{\min,k} < s_{\max,k}$, we need that $\xi < 1/2$. The contraction factor σ_k is then given by $\sigma_k = \frac{s_{\max,k} - s_{\min,k}}{2\varepsilon(s_{\max,k})} = (\xi^{-1} - 2)/8$. The condition $\sigma_k < 1$ implies that $\xi > 1/10$. Since $g_k \geq \min\{1 - s_{\max,k}, \frac{1}{2}\} - \varepsilon > 0$, the total number of intervals K is finite. Moreover, $g_k \leq 1$, so $\xi \in (1/10, 1/2)$ implies that $(s_{1,k} + s_{2,k})/2 - 1/2 < s_{\max,k} < (s_{1,k} + s_{2,k})/2$ is satisfied.

Now,

$$\eta \leq C_k(\delta) \sigma_k^{M_k+1} \quad \forall s \in \mathcal{S}_k$$

is satisfied by choosing the interpolation order

$$M_k = \left\lceil \frac{\log(\eta/C_k(\delta))}{\log \frac{1}{8}(\xi^{-2} - 2)} \right\rceil - 1.$$

The approximation results then follows from Lemma 5.1 and the continuous embeddings of $H_{\Omega}^s(\mathbb{R}^n) \hookrightarrow H_{\Omega}^{s_{1,k}}(\mathbb{R}^n)$ and $H_{\Omega}^{s+1/2-\varepsilon}(\mathbb{R}^n) \hookrightarrow H_{\Omega}^{s_{2,k}}(\mathbb{R}^n)$ for $s \in \mathcal{S}_k$.

We have that $\varepsilon(s_{\min,k}) = g_k \leq 1$, and hence $C_k(\delta)$ can be bounded independently of k . Hence, we can bound

$$M_k \lesssim \frac{|\log \eta|}{\left|\log \frac{1}{8}(\xi^{-2} - 2)\right|}$$

independently of k , where the constant depends on s_{\max} and δ . Since the choice of the intervals is independent of η and the number of intervals K is finite, the total number of interpolation nodes scales like $|\log \eta|$. \square

Remark 5.1. While Lemma 5.1 holds for $\delta \in (0, \infty)$, the result can be extended to the case of $\delta = \infty$. We can split

$$a(u, v; s, \infty) = a_\Omega(u, v; s) + a_{\mathbb{R}^n \setminus \Omega}(u, v; s),$$

with

$$\begin{aligned} a_\Omega(u, v; s) &= \frac{1}{2} \iint_{\Omega \times \Omega} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{n+2s}} dy dx, \\ a_{\mathbb{R}^n \setminus \Omega}(u, v; s) &= \int_\Omega u(x)v(x) \int_{\mathbb{R}^n \setminus \Omega} \frac{1}{|x - y|^{n+2s}} dy dx \\ &= \int_\Omega u(x)v(x) \int_{\partial\Omega} \frac{1}{2s} \frac{n \cdot (x - y)}{|x - y|^{n+2s}} dy dx. \end{aligned}$$

We can then apply the approximation to both a_Ω and $a_{\mathbb{R}^n \setminus \Omega}$ using $\delta = \text{diam } \Omega$. Hence, Lemma 5.1 and subsequent interpolation results also hold for $\delta = \infty$.

5.2. Interpolation of the s -derivative. We discuss the affine approximation of the derivative of the bilinear form with respect to s . That is, we consider for a given $u, v \in H_\Omega^s(\mathbb{R}^n)$ and $\delta \in (0, \infty)$

$$a'_s(u, v; s, \delta) \approx \tilde{a}'_s(u, v; s, \delta) = \sum_{m=0}^M \Theta'_{S_k, m}(s) a(u, v; s_m, \delta), \quad s \in S_k.$$

First, we provide an auxiliary result that establishes the error for the Chebyshev interpolation of the derivative of a function.

Lemma 5.3. Let $g \in C^\infty([-1, 1], \mathbb{R})$ and consider an expansion of g with respect to the Chebyshev polynomials T_m as

$$g(z) = \sum_{m=0}^{\infty} \gamma'_m T_m(z),$$

where the prime means that the first term is multiplied by $1/2$, and

$$\gamma_m = \frac{2}{\pi} \int_{-1}^1 \frac{g(z) T_m(z)}{\sqrt{1 - z^2}} dz.$$

Let

$$(I_M g)(z) := \sum_{m=0}^M \beta''_m T_m(z),$$

be its Chebyshev interpolant of rank M , where the double prime denotes a sum whose first and last terms are multiplied by $1/2$, and

$$\beta_m = \gamma_m + \sum_{j=1}^{\infty} (\gamma_{m+2jM} + \gamma_{-m+2jM}), \quad 0 \leq m \leq M.$$

Then, for $M > 2$ the following error estimate holds true

$$|g'(z) - (I_M g)'(z)| \leq \max_{\xi \in [-1, 1]} |g^{(M+1)}(\xi)| \frac{4}{(M-2)(M-2)!}.$$

Proof. We note that the following relation holds true, see, e.g., [48],

$$g(z) - (I_M g)(z) = \sum_{m=M+1}^{\infty} \gamma_m (T_m(z) - T_q(z)),$$

where $q = q(m, M) = |(m + M - 1) \pmod{2M} - (M - 1)|$, and $0 \leq q \leq M$. Then, using the above formula and the fact that $|T'_m(z)| \leq m^2$, we obtain that

$$\begin{aligned} |g'(z) - (I_M g)'(z)| &= \left| \sum_{m=M+1}^{\infty} \gamma_m (T'_m(z) - T'_q(z)) \right| \\ &\leq \sum_{m=M+1}^{\infty} |\gamma_m| (|T'_m(z)| + |T'_q(z)|) \\ &\leq 2 \sum_{m=M+1}^{\infty} |\gamma_m| m^2. \end{aligned}$$

We recall that from the aliasing formula the following bound on the Chebyshev coefficient holds true, see, e.g., [47, Theorem 4.2],

$$|\gamma_m| \leq \frac{2V_k}{\pi m(m-1) \dots (m-k)}, \quad m \geq k+1, \quad V_k := \|g^{(k)}\|_T = \int_{-1}^1 \frac{g^{(k+1)}(z)}{\sqrt{1-z^2}} dz,$$

where $k \geq 1$ denotes the k -th derivative of g . Using the above expressions and the fact that $m/(m-1) < (M+1)/M < 2$ for $M < m$ and

$$\sum_{j=M+1}^{\infty} \frac{1}{j(j-1) \dots (j-k)} = \frac{1}{kM(M-1) \dots (M-k+1)} = \frac{(M-k)!}{M!k},$$

we can estimate

$$\begin{aligned} |g'(z) - (I_M g)'(z)| &\leq 2 \sum_{m=M+1}^{\infty} |\gamma_m| m^2 \\ &\leq \frac{2V_k}{\pi} \sum_{m=M+1}^{\infty} \frac{m^2}{m(m-1) \dots (m-k)} \\ &\leq \frac{4V_k}{\pi} \sum_{m=M+1}^{\infty} \frac{1}{(m-2)(m-3) \dots (m-k)} \\ &\leq \frac{4V_k}{\pi} \sum_{m=M-1}^{\infty} \frac{1}{m(m-1) \dots (m-k+2)} \\ &= \frac{4V_k(M-k)!}{\pi(k-2)(M-2)!} \\ &\leq 4 \max_{\xi \in [-1,1]} |g^{(M+1)}(\xi)| \frac{1}{(M-2)(M-2)!}, \end{aligned}$$

where in the last inequality we set $k = M$ and used the fact that

$$|V_k| \leq \int_{-1}^1 \frac{|g^{(k+1)}(t)|}{\sqrt{1-t^2}} dt \leq \pi \max_{\xi \in [-1,1]} |g^{(k+1)}(\xi)|,$$

which concludes the proof. \square

Using the previous result we derive the error for the interpolation of the derivative of the bilinear form.

Lemma 5.4. *Let $u \in H_{\Omega}^{s_2, k}(\mathbb{R}^n)$, $v \in H_{\Omega}^{s_1, k}(\mathbb{R}^n)$, and $s \in \mathcal{S}_k$, where $s_{1,k}, s_{2,k}$ are such that the conditions of Lemma 5.1 hold true. Then, for $\delta \in (0, \infty)$, $M > 2$, we obtain*

$$|a'_s(u, v; s, \delta) - \tilde{a}'_s(u, v; s, \delta)| \lesssim \tilde{\eta}_k \|u\|_{H_{\Omega}^{s_2, k}(\mathbb{R}^n)} \|v\|_{H_{\Omega}^{s_1, k}(\mathbb{R}^n)}, \quad (5.2)$$

where $\tilde{\eta}_k = \tilde{\sigma}_k^M M^2$ with $\tilde{\sigma}_k = (s_{\max, k} - s_{\min, k})/\hat{\varepsilon}(s_{\max, k})$, and $\hat{\varepsilon}(s)$ is defined as in Lemma 5.1.

Proof. We present a proof for $\delta > \text{diam}|\Omega|$, the case $\delta \leq \text{diam}|\Omega|$ follows from the same arguments. Using the transformation $s = (s_{\min,k} + s_{\max,k})/2 + z(s_{\max,k} - s_{\min,k})/2$, $z \in [-1, 1]$, and applying Lemma 5.3, we obtain

$$|a'_s(u, v; s, \delta) - \tilde{a}'_s(u, v; s, \delta)| \leq \frac{4(s_{\max,k} - s_{\min,k})^M}{2^M(M-2)(M-2)!} \max_{\xi \in [s_{\min,k}, s_{\max,k}]} |a_s^{(M+1)}(u, v; \xi, \delta)|.$$

Using the same steps as in the proof of [9, Lemma 6.2], we obtain that

$$\max_{\xi \in [s_{\min,k}, s_{\max,k}]} |a_s^{(M+1)}(u, v; \xi, \delta)| \lesssim \frac{2^{M+1}(M+1)!}{\hat{\varepsilon}(s_{\max,k})^{M+1}} \|u\|_{H_\Omega^{s_{2,k}}(\mathbb{R}^n)} \|v\|_{H_\Omega^{s_{1,k}}(\mathbb{R}^n)}.$$

Then, combining the previous two estimates we obtain

$$\begin{aligned} |a'_s(u, v; s, \delta) - \tilde{a}'_s(u, v; s, \delta)| &\lesssim \left(\frac{s_{\max,k} - s_{\min,k}}{\hat{\varepsilon}(s_{\max,k})} \right)^M \cdot \frac{(M+1)!}{(M-2)(M-2)!\hat{\varepsilon}(s_{\max,k})} \|u\|_{H_\Omega^{s_{2,k}}(\mathbb{R}^n)} \|v\|_{H_\Omega^{s_{1,k}}(\mathbb{R}^n)} \\ &\lesssim \tilde{\sigma}_k^M M^2 \|u\|_{H_\Omega^{s_{2,k}}(\mathbb{R}^n)} \|v\|_{H_\Omega^{s_{1,k}}(\mathbb{R}^n)}, \end{aligned}$$

which concludes the proof. \square

5.3. Operator splitting for $\delta < \infty$ and δ -derivative. While the bilinear form $a(\cdot, \cdot; s, \delta)$ exhibits a smooth behavior with respect to s , as shown in Proposition 3.2, it is less regular with respect to δ , inherited by the limited spatial regularity of the solution; in fact, only a single derivative with respect to δ is available (Proposition 3.1). Therefore, δ -interpolation of the operator is bound to lead to unsatisfactory results. Instead, we split the operator into

$$a(u, v; s, \delta) = a(u, v; s, \infty) + c(u, v; s, \delta),$$

with correction term

$$\begin{aligned} c(u, v; s, \delta) &= a(u, v; s, \delta) - a(u, v; s, \infty) \\ &= - \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{n+2s}} \mathcal{X}_{|x-y|>\delta} dy dx \\ &= -2(u, v)_{L^2(\Omega)} \int_{\mathbb{R}^n \setminus B_\delta(0)} \frac{1}{|y|^{n+2s}} dy + 2 \int_{\Omega} \int_{\Omega} \frac{u(x)v(y)}{|x - y|^{n+2s}} \mathcal{X}_{|x-y|>\delta} dy dx \\ &= -\frac{2\pi^{n/2}}{\Gamma(n/2)} \frac{\delta^{-2s}}{s} (u, v)_{L^2(\Omega)} + 2 \int_{\Omega} \int_{\Omega} \frac{u(x)v(y)}{|x - y|^{n+2s}} \mathcal{X}_{|x-y|>\delta} dy dx. \end{aligned} \tag{5.3}$$

Here, we have implicitly extended u and v by zero to all of \mathbb{R}^n . For $\delta \geq \text{diam} \Omega$ the last term of (5.3) is zero.

As before, we can interpolate in s

$$\begin{aligned} a(u, v; s, \delta) &\approx \tilde{a}(u, v; s, \infty) + c(u, v; s, \delta) \\ &= \sum_{m=0}^{M_k} \Theta_{\mathcal{S}_k, m}(s) a(u, v; s_m, \infty) + c(u, v; s, \delta), \quad s \in \mathcal{S}_k. \end{aligned}$$

Changing the value of δ only requires reassembly of the correction term $c(\cdot, \cdot; s, \delta)$, which is significantly less expensive than assembly of $a(\cdot, \cdot; s_m, \delta)$, since integration over the singularity $x = y$ is avoided.

Moreover, we find that the δ -derivative is given by

$$\begin{aligned} a'_\delta(u, v; s, \delta) &= c'_\delta(u, v; s, \delta) \\ &= \frac{4\pi^{n/2}}{\Gamma(n/2)} \delta^{-1-2s} (u, v)_{L^2(\Omega)} - 2\delta^{-n-2s} \int_{\Omega} u(x) \int_{\partial B_\delta(x)} v(y) \, dy \, dx \\ &= \frac{4\pi^{n/2}}{\Gamma(n/2)} \delta^{-1-2s} (u, v - \bar{v})_{L^2(\Omega)}, \end{aligned}$$

where

$$\bar{v}(x) = \frac{1}{|\partial B_\delta(x)|} \int_{\partial B_\delta(x)} v(y) \, dy.$$

Note that, as described in the following section, we rely on finite element discretizations and that for finite element functions \bar{v} can be evaluated exactly. We also observe that $\bar{v} = 0$ for $\delta \geq \text{diam } \Omega$.

6. DISCRETIZATION AND ERROR ESTIMATES

In this section we first recall results of finite element approximations for the state and adjoint variables and then provide discretization error estimates for the state and adjoint solutions obtained via interpolated forms. We also provide, for the interpolated problem, error estimates for the gradient of the cost functional.

In what follows, we let $\Omega \subset \mathbb{R}^n$, $n = 1, 2$ to be convex, and let \mathcal{T}_h be a family of shape-regular and locally quasi-uniform triangulations of Ω [27], and let \mathcal{N}_h be the set of vertices of \mathcal{T}_h , h_K be the diameter of the element $K \in \mathcal{T}_h$. Let ϕ_i be the usual piecewise linear Lagrange basis function associated with a node $\tilde{z}_i \in \mathcal{N}_h$, satisfying $\phi_i(\tilde{z}_j) = \delta_{ij}$ for $\tilde{z}_j \in \mathcal{N}_h$, and let $X_h := \text{span} \{\phi_i \mid \tilde{z}_i \in \mathcal{N}_h\}$. The finite element subspace $V_h \subset H_\Omega^s(\mathbb{R}^n)$ is given by

$$V_h = \{v_h \in X_h \mid v_h = 0 \text{ on } \partial\Omega\} = \text{span} \{\phi_i \mid \tilde{z}_i \notin \partial\Omega\}.$$

6.1. Error estimates for the state and adjoint equations. It is well known that the following inverse inequality holds for $0 \leq \beta \leq \alpha$, see, e.g., [27, Corollary 1.141],

$$\|v_h\|_{H_\Omega^\alpha(\mathbb{R}^n)} \leq C_I h^{\beta-\alpha} \|v_h\|_{H_\Omega^\beta(\mathbb{R}^n)} \quad \forall v_h \in V_h. \quad (6.1)$$

Let I_h be the Scott-Zhang interpolation operator [27, Section 1.6.2] satisfying for $\beta \in [0, 2-s]$ the approximation property

$$\|v - I_h v\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim h^\beta \|v\|_{H_\Omega^{s+\beta}(\mathbb{R}^n)} \quad \forall v \in H_\Omega^{s+\beta}(\mathbb{R}^n) \quad (6.2)$$

and for $\beta \in [0, 1]$ the stability property

$$\|I_h v\|_{H_\Omega^\beta(\mathbb{R}^n)} \lesssim \|v\|_{H_\Omega^\beta(\mathbb{R}^n)} \quad \forall v \in H_\Omega^\beta(\mathbb{R}^n). \quad (6.3)$$

Then, for $f \in H^{1/2-\varepsilon}(\Omega)$, the discrete version of (2.7) reads as follows: find $u_h \in V_h$ such that

$$a(u_h, v_h; q) = (f, v_h)_{L^2(\Omega)}, \quad \forall v_h \in V_h. \quad (6.4)$$

From [8, Proposition 3.5], we obtain the following a-priori error estimate

$$\|u - u_h\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim h^{1/2-\varepsilon} \|f\|_{H_\Omega^{1/2-\varepsilon}(\mathbb{R}^n)}, \quad (6.5)$$

$$\|u - u_h\|_{L^2(\Omega)} \lesssim h^{1/2-\varepsilon+\beta} \|f\|_{H_\Omega^{1/2-\varepsilon}(\mathbb{R}^n)}, \quad (6.6)$$

where $\beta = \min\{s, 1/2 - \varepsilon\}$ and the constant C is independent of h . Similarly, we consider a discrete version of the adjoint equation (4.10): find $z_h \in V_h$ such that

$$a(\phi_h, z_h; q) = (u_h - u_d, \phi_h)_{L^2(\Omega)}, \quad \forall \phi_h \in V_h. \quad (6.7)$$

The following proposition establishes the a priori error estimates for the adjoint equation.

Proposition 6.1. *For $q \in Q$, let $z(q) \in H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)$ and $z_h(q) \in V_h$ be the solutions of (4.10) and (6.7), respectively. Then, the following holds*

$$\|z - z_h\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim h^{1/2-\varepsilon} \left(\|f\|_{H^{1/2-\varepsilon}(\Omega)} + \|u_d\|_{H^{1/2-\varepsilon}(\Omega)} \right). \quad (6.8)$$

Proof. Let $u_d \in H_\Omega^{1/2-\varepsilon}(\mathbb{R}^n)$ and let $u \in H_\Omega^s(\mathbb{R}^n)$ be the solution of (2.7); we denote by $\hat{z}_h \in H_\Omega^s(\mathbb{R}^n)$ the solution of

$$a(\phi_h, \hat{z}_h; q) = (u - u_d, \phi_h)_{L^2(\Omega)}, \quad \forall \phi_h \in V_h.$$

By the standard Lax-Milgram argument the above problem is well-posed. Furthermore, since $u - u_d \in H_\Omega^{1/2-\varepsilon}(\mathbb{R}^n)$, by invoking the regularity result (4.12), it follows from [8, Proposition 3.5] that

$$\|z - \hat{z}_h\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim h^{1/2-\varepsilon} \|u - u_d\|_{H^{1/2-\varepsilon}(\Omega)}.$$

By applying the above estimate and (6.6) we obtain

$$\begin{aligned} \|z - z_h\|_{H_\Omega^s(\mathbb{R}^n)} &\leq \|z - \hat{z}_h\|_{H_\Omega^s(\mathbb{R}^n)} + \|\hat{z}_h - z_h\|_{H_\Omega^s(\mathbb{R}^n)} \\ &\lesssim h^{1/2-\varepsilon} \|u - u_d\|_{H^{1/2-\varepsilon}(\Omega)} + \|u - u_d\|_{H^{-s}(\Omega)} \\ &\lesssim h^{1/2-\varepsilon} \|u - u_d\|_{H^{1/2-\varepsilon}(\Omega)} + h^{1/2-\varepsilon+\beta} \|f\|_{H^{1/2-\varepsilon}(\Omega)} \\ &\lesssim h^{1/2-\varepsilon} \left(\|f\|_{H^{1/2-\varepsilon}(\Omega)} + \|u_d\|_{H^{1/2-\varepsilon}(\Omega)} \right). \end{aligned}$$

This concludes the proof. \square

6.2. Error estimates for the interpolated state and adjoint equations.

Note that, in practice, instead of solving (6.4) and (6.7), which involve the exact bilinear form, we consider the approximated problem where the bilinear form is replaced by its affine interpolant. That is, we solve the discretized problems of finding, respectively, $\tilde{u}_h(q) \in V_h$, such that

$$\tilde{a}(\tilde{u}_h(q), v_h; q) = (f, v_h)_{L^2(\Omega)}, \quad \forall v_h \in V_h, \quad (6.9)$$

and \tilde{z}_h such that

$$\tilde{a}(\phi_h, \tilde{z}_h; q) = (\tilde{u}_h - u_d, \phi_h)_{L^2(\Omega)}, \quad \forall \phi_h \in V_h. \quad (6.10)$$

The next theorem provides the error estimate for the state solution of the interpolated problem.

Theorem 6.2. *Let $q = (s, \delta) \in [s_{\min}, s_{\max}] \times (0, \infty)$, and let $u(q)$ be the solution of (2.7). We consider an operator interpolation as given by Lemma 5.2 for the tolerance*

$$\eta \leq \frac{\min_{s \in [s_{\min}, s_{\max}]} \alpha(s, \delta)}{2C_I} h^{1/2-\varepsilon}, \quad (6.11)$$

where $\alpha(s, \delta)$ is the coercivity constant of $a(\cdot, \cdot; q)$. Let also \tilde{u}_h be the solution of the interpolated problem (6.9), then

$$\|u - \tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim \left(h^{1/2-\varepsilon} + \eta \right) \|f\|_{H^{1/2-\varepsilon}(\Omega)} \lesssim h^{1/2-\varepsilon} \|f\|_{H^{1/2-\varepsilon}(\Omega)}. \quad (6.12)$$

Proof. For $\tilde{u}_h \in V_h$

$$\begin{aligned}
\tilde{a}(\tilde{u}_h, \tilde{u}_h; q) &\geq a(\tilde{u}_h, \tilde{u}_h; q) - |a(\tilde{u}_h, \tilde{u}_h; q) - \tilde{a}(\tilde{u}_h, \tilde{u}_h; q)| \\
&\geq \alpha(s, \delta) \|\tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)}^2 - \eta \|\tilde{u}_h\|_{H_\Omega^{\bar{s}_2(s)}(\mathbb{R}^n)} \|\tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} \\
&\geq \left(\alpha(s, \delta) - \frac{C_I \eta}{h^{\bar{s}_2(s)-s}} \right) \|\tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)}^2 \\
&= \left(\alpha(s, \delta) - \frac{\min \alpha(s, \delta)}{2} h^{s+1/2-\varepsilon-\bar{s}_2(s)} \right) \|\tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)}^2 \\
&\geq \alpha(s, \delta)/2 \|\tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)}^2,
\end{aligned}$$

where we have used the coercivity of $a(\cdot, \cdot; q)$, Lemma 5.2, the inverse inequality (6.1), and $s+1/2-\varepsilon-\bar{s}_2(s) \geq 0$, where $\bar{s}_2(s)$ is defined as in Lemma 5.2. This shows that $\tilde{a}(\cdot, \cdot; q)$ is coercive on the discrete space V_h .

By Strang's first lemma [27], we have that

$$\begin{aligned}
\|u - \tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} &\lesssim \inf_{v_h} \left\{ \|u - v_h\|_{H_\Omega^s(\mathbb{R}^n)} + \sup_{w_h} \frac{|a(v_h, w_h; q) - \tilde{a}_\xi(v_h, w_h; q)|}{\|w_h\|_{H_\Omega^s(\mathbb{R}^n)}} \right\} \\
&\lesssim \inf_{v_h} \left\{ \|u - v_h\|_{H_\Omega^s(\mathbb{R}^n)} + \eta \|v_h\|_{H_\Omega^{\bar{s}_2(s)}(\mathbb{R}^n)} \right\} \\
&\lesssim \left\{ \|u - I_h u\|_{H_\Omega^s(\mathbb{R}^n)} + \eta \|I_h u\|_{H_\Omega^{\bar{s}_2(s)}(\mathbb{R}^n)} \right\}.
\end{aligned}$$

Using (6.2) and (6.3) we obtain

$$\begin{aligned}
\|u - \tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} &\lesssim h^{1/2-\varepsilon} \|u\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} + \eta \|u\|_{H_\Omega^{\bar{s}_2}(\mathbb{R}^n)} \\
&\lesssim (h^{1/2-\varepsilon} + \eta) \|u\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)}.
\end{aligned}$$

Finally, using (6.11), Assumption 2.2 and Theorem 2.1 we obtain the desired result. \square

Next, we derive the error estimates for the interpolated adjoint problem.

Theorem 6.3. For $q \in Q$, let $z(q) \in H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)$ and $z_h(q) \in V_h$ be the solutions of (4.10) and (6.7), respectively, and the tolerance η for the operator interpolation is chosen as in Theorem 6.2. Then, the following holds

$$\|z - \tilde{z}_h\|_{H_\Omega^s} \lesssim (h^{1/2-\varepsilon} + \eta) \left(\|f\|_{H^{1/2-\varepsilon}(\Omega)} + \|u_d\|_{H^{1/2-\varepsilon}(\Omega)} \right). \quad (6.13)$$

Proof. The proof follows similar arguments as in Theorem 6.2 by using Strang's first lemma and invoking (6.1), (6.2), and (6.12). \square

6.3. Error estimates for the gradient of the cost functional. First, we prove auxilliary convergence results, which will be needed later to derive the error estimate for the derivative of the reduced cost functional.

Lemma 6.4. Let u, \tilde{u}_h and z, \tilde{z}_h be the continuous and discrete solutions of (2.7), (6.9) and (4.10), (6.10), respectively, and let the conditions of Theorem 6.2 and Theorem 6.3 hold true. Then, for $\mu \in [0, s+1/2)$, $\varepsilon > 0$, and $\alpha := \min\{s-\mu, 1/2-\varepsilon\}$, $s \in (0, 1)$ the following holds true

$$\|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} \lesssim \left(h^{\min\{s+1/2-\mu, 1-\varepsilon\}-\varepsilon} + \eta \right) \|f\|_{H^{1/2-\varepsilon}(\Omega)} \quad (6.14)$$

and

$$\|z - \tilde{z}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} \lesssim \left(h^{\min\{s+1/2-\mu, 1-\varepsilon\}-\varepsilon} + \eta \right) C_d \quad (6.15)$$

where $C_d := \|f\|_{H^{1/2-\varepsilon}(\Omega)} + \|u_d\|_{H^{1/2-\varepsilon}(\Omega)}$, and $H_\Omega^0(\mathbb{R}^n) \equiv L^2(\Omega)$. Furthermore, we have the following norm bound

$$\|\tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} \lesssim \|f\|_{H^{1/2-\varepsilon}(\Omega)}, \quad (6.16)$$

for $\mu \in [0, s + 1/2)$, with the hidden constant independent of h .

Proof. We present a proof for the state solutions only, since for the adjoint solutions it follows similar arguments and we omit it for compactness purposes. First, we consider the case $\mu \in [s, s + 1/2)$. Using (6.1), (6.2), (2.8), (6.12) we obtain

$$\begin{aligned} \|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} &\leq \|u - I_h u\|_{H_\Omega^\mu(\mathbb{R}^n)} + \|I_h u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} \\ &\lesssim h^{s+1/2-\varepsilon-\mu} \|u\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} + h^{s-\mu} \|I_h u - \tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} \\ &\lesssim h^{s+1/2-\varepsilon-\mu} \|f\|_{H^{1/2-\varepsilon}(\Omega)} + h^{s-\mu} \left(\|I_h u - u\|_{H_\Omega^s(\mathbb{R}^n)} + \|u - \tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} \right) \\ &\lesssim h^{s+1/2-\varepsilon-\mu} \|f\|_{H^{1/2-\varepsilon}(\Omega)}. \end{aligned}$$

Then, it immediately follows that

$$\|\tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} \leq \|\tilde{u}_h - u\|_{H_\Omega^\mu(\mathbb{R}^n)} + \|u\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim (h^{s+1/2-\varepsilon-\mu} + 1) \|f\|_{H^{1/2-\varepsilon}(\Omega)},$$

and using $h \lesssim 1$ we conclude (6.16). Next, using a duality argument we prove the case $\mu \in [0, s)$. For $\chi \in H_\Omega^s(\mathbb{R}^n)$ we consider the following problem

$$a(v, \chi; q) = l(v) \quad \forall v \in H_\Omega^s(\mathbb{R}^n),$$

where $l(v) := (u - \tilde{u}_h, v)_{H_\Omega^\mu(\mathbb{R}^n)}$. It is clear that $l \in H^{-\mu}(\Omega)$, and applying Theorem 2.1 we obtain that $\chi \in H_\Omega^{s+\alpha}(\mathbb{R}^n)$ with $\alpha = \min\{s - \mu, 1/2 - \varepsilon\}$, and

$$\|\chi\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \lesssim \|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)}. \quad (6.17)$$

Now choosing $v = u - \tilde{u}_h \in H_\Omega^s(\mathbb{R}^n)$, and using Lemma 5.2, Theorem 6.2, (6.2), (6.3), and taking into account that $a(u, I_h \chi; q) - \tilde{a}(\tilde{u}_h, I_h \chi) = 0$, we obtain

$$\begin{aligned} \|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)}^2 &= a(u - \tilde{u}_h, \chi; q) - a(u, I_h \chi; q) + \tilde{a}(\tilde{u}_h, I_h \chi; q) \\ &\leq |a(u - \tilde{u}_h, \chi - I_h \chi; q)| + |a(\tilde{u}_h, I_h \chi; q) - \tilde{a}(\tilde{u}_h, I_h \chi; q)| \\ &\leq \|u - \tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} \|\chi - I_h \chi\|_{H_\Omega^s(\mathbb{R}^n)} + \eta \|\tilde{u}_h\|_{H_\Omega^{\bar{s}_2(s)}(\mathbb{R}^n)} \|I_h \chi\|_{H_\Omega^s(\mathbb{R}^n)} \\ &\lesssim h^{\alpha+1/2-\varepsilon} \|\chi\|_{H_\Omega^{s+\alpha}(\mathbb{R}^n)} \|f\|_{H^{1/2-\varepsilon}(\Omega)} + \eta \|\tilde{u}_h\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} \|\chi\|_{H_\Omega^s(\mathbb{R}^n)} \\ &\lesssim \left(h^{\alpha+1/2-\varepsilon} + \eta \right) \|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)} \|f\|_{H^{1/2-\varepsilon}(\Omega)}, \end{aligned}$$

where the last two inequalities have been obtained by using that $\tilde{u}_h \in H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n) \hookrightarrow H_\Omega^{\bar{s}_2(s)}(\mathbb{R}^n)$, where $\bar{s}_2(s)$ such as in Lemma 5.2, and $\|\tilde{u}_h\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} \lesssim \|f\|_{H^{1/2-\varepsilon}}$ by (6.16), and invoking the stability estimate

$$\|\chi\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim \|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)}$$

together with (6.17). Then, dividing by $\|u - \tilde{u}_h\|_{H_\Omega^\mu(\mathbb{R}^n)}$ we obtain the desired result (6.14) for $\mu \in [0, s)$. The remaining norm bound (6.16) is obtained by the similar arguments as before. This concludes the proof. \square

For $\tilde{u}_h, \tilde{z}_h \in V_h$ the solutions of (6.9) and (6.10), respectively, we define

$$\tilde{j}_h'(q) := \mathcal{R}'(q) - \tilde{a}'_q(\tilde{u}_h, \tilde{z}_h; q).$$

Then, we obtain the following error estimate for the derivative of the reduced cost functional.

Theorem 6.5. For $q \in Q$, $s \in (0, 1)$, $\varepsilon > 0$ the following estimate holds

$$|j'(q) - \tilde{j}'_h(q)| \lesssim \begin{cases} |\log h| (h^{s+1/2-\varepsilon} + \eta) + \tilde{\eta} & s < 1/2, \\ h^{3/2-s-3\varepsilon} + \eta + \tilde{\eta}, & s \geq 1/2, \end{cases} \quad (6.18)$$

where η is defined as in Theorem 6.2, $\tilde{\eta} = \max_k \tilde{\eta}_k$ and $\tilde{\eta}_k$ is defined as in (5.2).

Proof. We consider

$$\begin{aligned} |j'(q) - \tilde{j}'_h(q)| &\leq |a'_q(u, z; q) - \tilde{a}'_q(\tilde{u}_h, \tilde{z}_h; q)| \\ &\leq |a'_s(u - \tilde{u}_h, z; q)| + |a'_s(\tilde{u}_h, z - \tilde{z}_h; q)| + |a'_s(\tilde{u}_h, \tilde{z}_h; q) - \tilde{a}'_s(\tilde{u}_h, \tilde{z}_h; q)| \\ &\quad + |a'_\delta(u - \tilde{u}_h, z; q)| + |a'_\delta(\tilde{u}_h, z - \tilde{z}_h; q)|. \end{aligned} \quad (6.19)$$

Next, we estimate each term from the above. For the first two terms we distinguish between two cases $s < 1/2$ and $s \geq 1/2$. In particular, for $s \in (0, 1/2)$, using (3.5) with $s_1 = 2s$ and $s_2 = \xi$, $0 < \xi < 2(1-s)$, (4.12), (6.1), and Lemma 6.4, we can estimate

$$\begin{aligned} |a'_s(u - \tilde{u}_h, z; q)| &\leq C(1, \xi) \|u - \tilde{u}_h\|_{H_\Omega^\xi(\mathbb{R}^n)} \|z\|_{H_\Omega^{2s}(\mathbb{R}^n)} \\ &\leq C(1, \xi) \left(\|u - I_h u\|_{H_\Omega^\xi(\mathbb{R}^n)} + \|I_h u - \tilde{u}_h\|_{H_\Omega^\xi(\mathbb{R}^n)} \right) \|z\|_{H_\Omega^{2s}(\mathbb{R}^n)} \\ &\lesssim C(1, \xi) \left(h^{s+1/2-\varepsilon-\xi} + h^{-\xi} \|I_h u - \tilde{u}_h\|_{L^2(\Omega)} \right) \|z\|_{H_\Omega^{2s}(\mathbb{R}^n)} \\ &\lesssim C(1, \xi) h^{-\xi} \left(h^{s+1/2-\varepsilon} + \|I_h u - u\|_{L^2(\Omega)} + \|u - \tilde{u}_h\|_{L^2(\Omega)} \right) \|z\|_{H_\Omega^{2s}(\mathbb{R}^n)} \\ &\lesssim |\log h| \left(h^{s+1/2-\varepsilon} + \eta \right), \end{aligned}$$

where the last inequality has been obtained by taking $\xi \sim -1/\log h$, and noting that $C(1, \beta)$ in (3.5) scales like $C(1, \beta) \sim \xi^{-1}$, and, hence, $C(1, \xi) h^{-\xi} \sim |\log h|$. Similarly, we estimate the second term in (6.19) for $s \in (0, 1/2)$ and obtain

$$\begin{aligned} |a'_s(\tilde{u}_h, z - \tilde{z}_h; q)| &\leq C(1, \xi) \|z - \tilde{z}_h\|_{H_\Omega^\xi(\mathbb{R}^n)} \|\tilde{u}_h\|_{H_\Omega^{2s}(\mathbb{R}^n)} \\ &\lesssim C(1, \xi) h^{-\xi} \left(h^{s+1/2-\varepsilon} \|z\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} + \|I_h z - \tilde{z}_h\|_{L^2(\Omega)} \right) \|\tilde{u}_h\|_{H_\Omega^{2s}(\mathbb{R}^n)} \\ &\lesssim |\log h| \left(h^{s+1/2-\varepsilon} + \eta \right), \end{aligned}$$

where we have invoked Lemma 6.4 to estimate $\|\tilde{u}_h\|_{H_\Omega^{2s}(\mathbb{R}^n)} \lesssim \|f\|_{H^{1/2-\varepsilon}(\Omega)}$.

Next, we consider the case $s \in [1/2, 1)$. As previously, using (3.5) with $s_1 = 1 - \varepsilon$ and $s_2 = 2s - 1 + 2\varepsilon$, (4.12), (6.1), and Lemma 6.4 for $\mu = 2s - 1 + 2\varepsilon$ we obtain

$$|a'_s(u - \tilde{u}_h, z; q)| \lesssim \|u - \tilde{u}_h\|_{H_\Omega^{2s-1+2\varepsilon}(\mathbb{R}^n)} \|z\|_{H_\Omega^{1-\varepsilon}(\mathbb{R}^n)} \lesssim h^{3/2-s-3\varepsilon} + \eta.$$

Similarly, it follows for the second term in (6.19)

$$|a'_s(\tilde{u}_h, z - \tilde{z}_h; q)| \lesssim \|z - \tilde{z}_h\|_{H_\Omega^{2s-1+2\varepsilon}(\mathbb{R}^n)} \|\tilde{u}_h\|_{H_\Omega^{1-\varepsilon}(\mathbb{R}^n)} \lesssim h^{3/2-s-3\varepsilon} + \eta.$$

Next, using (3.2), (2.6) and Lemma 6.4 we can estimate the last two terms in (6.19)

$$\begin{aligned} |a'_\delta(u - \tilde{u}_h, z; q)| + |a'_\delta(\tilde{u}_h, z - \tilde{z}_h; q)| &\lesssim \|u - \tilde{u}_h\|_{L^2(\Omega)} \|z\|_{H_\Omega^s(\mathbb{R}^n)} + \|z - \tilde{z}_h\|_{L^2(\Omega)} \|\tilde{u}_h\|_{H_\Omega^s(\mathbb{R}^n)} \\ &\lesssim \left(h^{\alpha+1/2-\varepsilon} + \eta \right) (\|f\|_{H^{1/2-\varepsilon}} + \|u_d\|_{H^{1/2-\varepsilon}}), \end{aligned}$$

where $\alpha = \min\{s, 1/2 - \varepsilon\}$. Eventually, we estimate the error due to the affine interpolation of the bilinear form. In particular, invoking Lemma 5.4 with $s_{2,k} =$

$\bar{s}_2(s)$ and $s_{1,k} = s_{\min,k}$, where we recall $\bar{s}_2(s) = \min\{1, s_{\min,k} + 1/2\} - \varepsilon$ for $s \in \mathcal{S}_k$, and using the continuous embedding $H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n) \hookrightarrow H_\Omega^{\bar{s}_2(s)}(\mathbb{R}^n)$, we deduce

$$|a'_s(\tilde{u}_h, \tilde{z}_h; q) - \tilde{a}'_s(\tilde{u}_h, \tilde{z}_h; q)| \lesssim \tilde{\eta} \|\tilde{u}_h\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} \|\tilde{z}_h\|_{H_\Omega^s(\mathbb{R}^n)} \lesssim \tilde{\eta},$$

where, as before, we used the fact that $\|\tilde{u}_h\|_{H_\Omega^{s+1/2-\varepsilon}(\mathbb{R}^n)} \lesssim 1$, thanks to (6.16), and $\tilde{\eta} = \max_k \tilde{\eta}_k$. Then, combining all above estimates we obtain the desired result. \square

7. NUMERICAL RESULTS

In this section, we illustrate the theoretical results via numerical experiments. We first report on the accuracy of the interpolant of the bilinear form $a(\cdot, \cdot)$ and its s -derivative, and then solve identification problems either for s only or s and δ jointly.

In all cases, a panel clustering approach is used to avoid prohibitively expensive dense matrices [2]. In a nutshell, assembly is split into near field and far field interactions. Far field interactions are then approximated using Chebyshev interpolation, where order and interpolation domain depend on the distance of the interaction. Specially designed quadrature rules account for the singular behavior of the kernel, leading to assembly cost of $\mathcal{O}(N \log^{2d} N)$ for fixed δ and s , where $N = \dim V_h$ [1]. The error due to quadrature is always dominated by the discretization error, which is why we do not explicitly keep track of it and why we still denote the bilinear form involving quadrature as $a(\cdot, \cdot; \cdot)$. The equations for state and adjoint are solved using the conjugate gradient method, preconditioned by geometric multigrid, resulting in quasi-optimal complexity for both problem setup and solution. The discretized control problems are solved using the BFGS algorithm [38].

We will consider the following test problems:

I The fractional Laplacian of order s on the n -dimensional ball.

Let $\Omega = B_1(0) \subset \mathbb{R}^n$, $\delta = \infty$ and $(-\Delta)^s := C_{n,s} \mathcal{L}$, where \mathcal{L} is given by (2.1) and $C_{n,s}$ is given by (1.2). The analytic solution of

$$\begin{cases} (-\Delta)^s u &= 1 & \text{in } \Omega, \\ u &= 0 & \text{in } \mathbb{R}^n \setminus \Omega. \end{cases}$$

is given by [29]

$$u_{ex}(x; s) = c_{n,s} (1 - |x|^2)_+^s \in H_\Omega^{s+1/2-\varepsilon}, \quad \forall \varepsilon > 0,$$

with $c_{n,s} = \frac{\Gamma(n/2)}{2^{2s} \Gamma(\frac{n+2s}{2}) \Gamma(1+s)}$.

II The nonlocal operator \mathcal{L} of order s and truncation length δ on the n -dimensional ball.

We solve (2.3) with right-hand side $f = 1$, for which no analytic solution is known.

7.1. Convergence of the interpolation. In this section we illustrate the convergence of the interpolation for both the operator and the gradient of the functional.

7.1.1. Convergence of the operator interpolation. In order to illustrate the results of Theorem 6.2, we solve problem **I** for $s = \frac{1}{10}, \frac{2}{10}, \dots, \frac{9}{10}$ and dimension $n = 1$ for mesh sizes $h = 2^{-j}$, $j = 4, \dots, 10$, using the operator interpolation $\tilde{a}(\cdot, \cdot; s, \infty)$ and the exact (up to quadrature error) operator $a(\cdot, \cdot; s, \infty)$. The free parameter ξ is chosen via brute-force minimization with respect to the total number of interpolation nodes. We note that this procedure is inexpensive, since only intervals \mathcal{S}_k and interpolation orders M_k are manipulated.

In Figure 7.1 we plot the errors $\|\tilde{u}_h(s) - u(s)\|_{H_\Omega^s}$ and $\|u_h(s) - u(s)\|_{H_\Omega^s}$, where $\tilde{u}_h(s)$ and $u_h(s)$ are the solutions obtained with and without interpolation. It

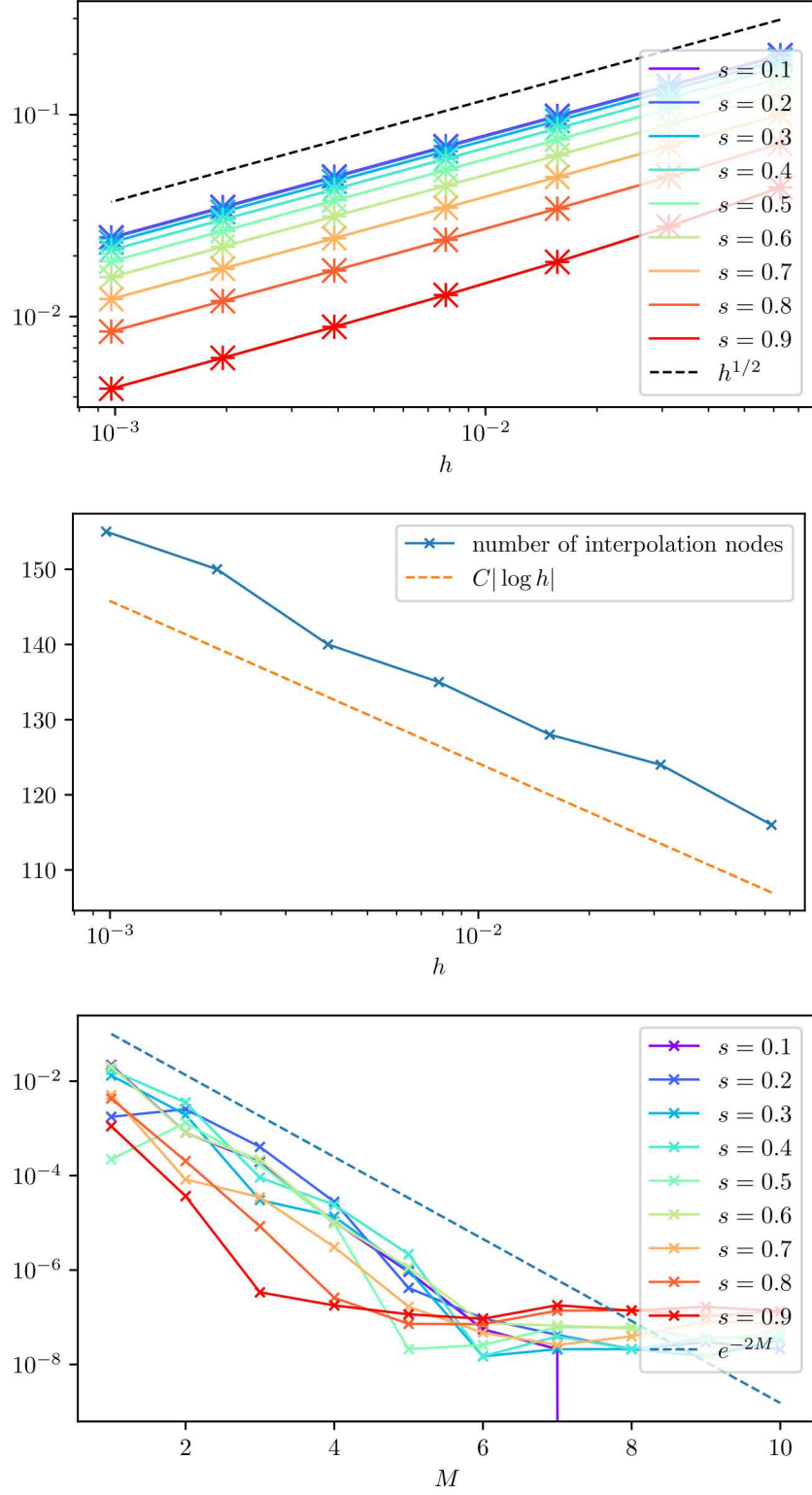


FIGURE 7.1. Solution errors with interpolation, $\|\tilde{u}_h(s) - u(s)\|_{H_\Omega^s}$ (\times), and without interpolation, $\|u_h(s) - u(s)\|_{H_\Omega^s}$ (+), for $h = 2^{-j}$, $j = 4, \dots, 10$ (top). Total number of interpolation nodes (middle). Exponential convergence of the interpolation error $\|\tilde{u}_{h,M}(s) - u_h(s)\|_{H_\Omega^s}$ with respect to M (bottom).

can be observed that both methods lead to virtually identical error. We also plot the total number of interpolation nodes and observe that it indeed is proportional to $|\log h|$. To further illustrate the exponential convergence with respect to the interpolation order, we also plot $\|\tilde{u}_{h,M}(s) - u_h(s)\|_{H^s_\Omega}$ for $h = 2^{-10}$, a fixed value of ξ and $1 \leq M \leq 12$ interpolation nodes on each interval \mathcal{S}_k . The observed plateau is due to quadrature error that impacts both $u_{h,M}$ and u_h . However, we note that the magnitude of this error is negligible compared to the discretization error, as can be seen from Figure 7.1.

7.1.2. Convergence of the gradient approximation. We evaluate the convergence of the approximation of the s -derivative. We set $u_d = 1 - x^2$ and consider the optimal control problem with respect to s for problem **I**. For $s = \frac{1}{10}, \frac{2}{10}, \dots, \frac{9}{10}$, we compute

$$\left| \tilde{j}'_h(s, \infty) - \tilde{j}'_{\underline{h}}(s, \infty) \right|,$$

where \tilde{j}_h is the reduced cost functional, evaluated on a mesh of size h (Figure 7.2) and $\underline{h} \ll h$. We notice that the observed convergence rates are of order $\mathcal{O}(h)$, which indicate that the derived error estimates (6.18) are nearly optimal for $s = 1/2$. We believe that an improved order of convergence could be obtained theoretically also for $s > 1/2$, however this would require to re-derive the estimates (3.5) for $s_2, s_1 \geq 1$, which is out of the scope of the current work. We also report, for fixed mesh size $h = 2^{-10}$,

$$\left| \tilde{j}'_M(s, \infty) - \tilde{j}'_{\overline{M}}(s, \infty) \right|,$$

where \tilde{j}_M is the reduced cost functional using M interpolation nodes on each interval \mathcal{S}_k and $M \ll \overline{M}$. Again, we observe exponential convergence until the quadrature error is reached, which is in agreement with the theoretical results.

7.2. Complexity and memory requirements. For convenience, we summarize the complexity and memory requirements of the operator approximation for $a(\cdot, \cdot; s, \delta)$. In a preprocessing step, $a(\cdot, \cdot; s_m, \infty)$ is assembled for all Chebyshev nodes $s_m \in \mathbb{S}_k$, $k = 1, \dots, K$. By combining Lemma 5.2 with Theorem 6.2, the number of required Chebyshev nodes scales like $|\log h| \sim \log N$, where $N = \dim V_h$. By using the panel clustering approach of [2], each operator evaluation costs $\mathcal{O}(N \log^{2n} N)$ operations and memory. Hence, the overall cost of the preprocessing step scales as $\mathcal{O}(N \log^{2n+1} N)$ in complexity and memory.

The assembly of the correction terms $c(\cdot, \cdot; s_m, \delta)$ for $\delta < \infty$ costs $\mathcal{O}(N \log^{2n} N)$ when using panel clustering. Since matrix-vector products involving a and c scale in the same fashion as the assembly, solving linear systems involving $a(\cdot, \cdot; s, \delta)$ for state and adjoint solution using multigrid is achieved in $\mathcal{O}(N \log^{2n+1} N)$ operations. Overall, each BFGS iteration has $\mathcal{O}(N \log^{2n+1} N)$ complexity. This shows that the presented approach is quasi-optimal in the number of unknowns.

We illustrate the efficiency of the approach by constructing all required operators of problem **II** for $s \in \{0.25, 0.75\}$ and $\delta \in \{0.5, 1.5, 2.5\}$ and different mesh sizes h . All computations were performed on a single core of a Intel Xeon E5-2650 processor. The timings for assembly of $\tilde{a}(\cdot, \cdot; s, \infty)$ and $c(\cdot, \cdot; s, \delta)$ given in Figure 7.3 show that we indeed recover quasi-optimal complexity. We observe that assembly of the corrections for $\delta \geq \text{diam } \Omega$ is significantly cheaper, since the last term of (5.3) drops out, and assembly of the purely local mass terms only has $\mathcal{O}(N)$ complexity.

7.3. Consistency of the identification strategy. For one- and two-dimensional problems we test the robustness of our optimization approach by using the manufactured solutions described in **I** and **II**.

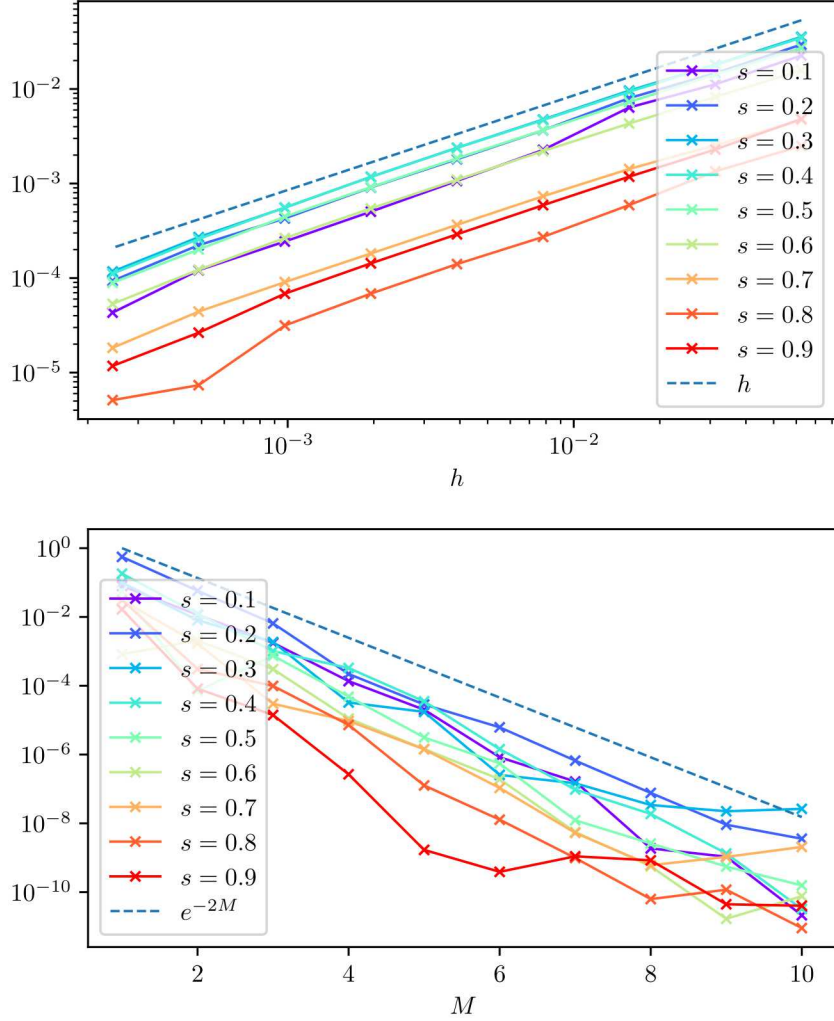


FIGURE 7.2. Convergence of the derivative of the reduced cost functional (*top*). Exponential convergence of the derivative of the reduced cost functional with respect to M (*bottom*).

7.3.1. The fractional Laplacian case. We consider test problem **I**, and set $\mathcal{R}(s) = \frac{\alpha}{s(1-s)}$, $\alpha = 5 \times 10^{-7}$ in order to satisfy conditions (4.2). Set $s^* = 0.5$ and $u_d = u_{ex}(s^*)$. The exact solution of the minimization problem (4.3) is given by $(s^*, u_{ex}(s^*))$, since s^* minimizes the regularization. In $n = 1$ dimensions, the mesh size is given by $h = 2.4 \times 10^{-4}$ and in $n = 2$ dimensions by $h = 0.04$, and the number of unknowns is 8191 and 3969 respectively. The initial guess is $s_0 = 0.1$. The linear solver tolerance is 10^{-10} , and the BFGS iteration is terminated if the norm of the gradient drops below 10^{-8} . In Figure 7.4 we display the values \tilde{j}_h of the cost functional for all functional evaluations (\bullet). In both cases, the exact value of s is recovered.

In a second example we take $s^* = 0.75$ and set $u_d = u_{ex}(s^*)$. Since s^* no longer coincides with the minimum of the regularization, no exact solution of the minimization problem (4.3) is known. The initial guess is again chosen to be $s_0 = 0.1$. The values of the BFGS iterates and the regularization \mathcal{R} are shown in Figure 7.5.

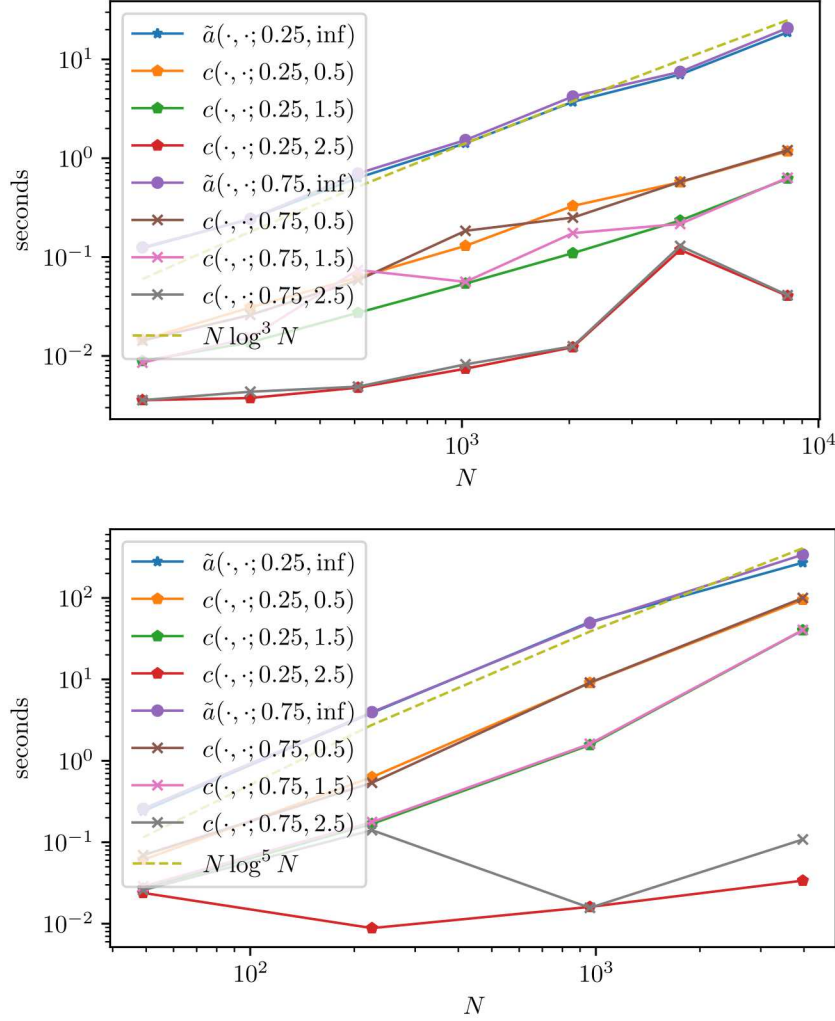


FIGURE 7.3. Assembly timings for $\tilde{a}(\cdot, \cdot; s, \infty)$ and correction terms $c(\cdot, \cdot; s, \delta)$ in 1D (*top*) and 2D (*bottom*).

While the exact solution of the minimization problem (4.3) is unknown, it can be seen that in both cases s_{final} , the final value of s , is close to s^* and the final value of the cost functional is quite close to $J(u_d, s^*) = \mathcal{R}(s^*)$.

7.3.2. The truncated fractional case. In a final set of examples, we solve the identification problem for $q = (s, \delta)$. We set $q^* = (s^*, \delta^*) = (0.75, 0.9)$ and solve the optimal control problem for \mathbf{II} in $n = 1$ and $n = 2$ dimensions. We set $u_d = u_h(q^*)$ and use the regularization $\mathcal{R}(s, \delta) = \frac{\alpha}{s(1-s)} + \beta \frac{e^\delta}{\delta}$ with $\alpha = 5 \times 10^{-7}$ and $\beta = 10^{-6}$. The initial guess is $q_0 = (0.1, 0.5)$. Figure 7.6 shows the (s, δ) -values of the BFGS iterates. We observe that the method indeed recovers q_{final} in proximity to q^* . In Figure 7.7, we display the corresponding values of the cost functional. We observe that $j(q_{\text{final}}) < j(q^*) = \mathcal{R}(q^*) \approx 5.4 \times 10^{-4}$.

7.4. Convergence with respect to the mesh size. In the same setting as in the previous section, we solve test problem \mathbf{II} in $n = 1$ dimensions for a sequence of meshes with varying mesh size h . The linear solver tolerance is 10^{-10} , and the

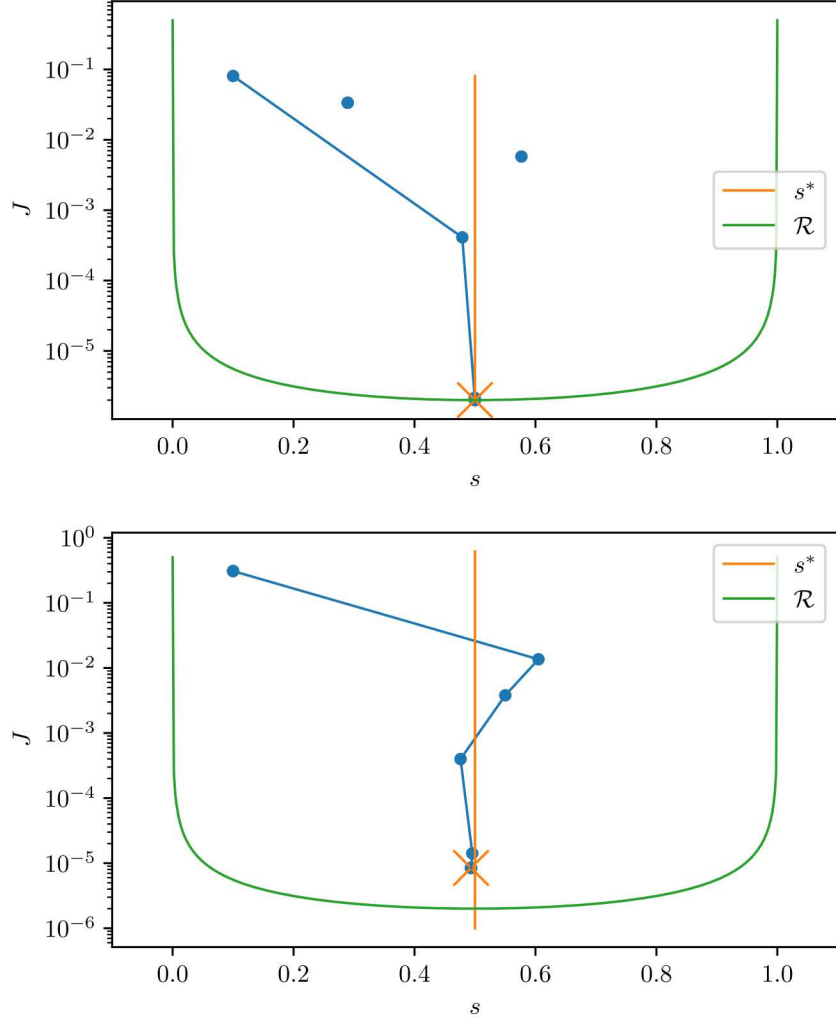


FIGURE 7.4. Value of the cost functional at all functional evaluations (\bullet) and for BFGS iterates ($-$) in 1D (*top*) and 2D (*bottom*) for test problem **I** and $u_d = u_{ex}(0.5)$. The final iterate is marked with \times .

BFGS iteration is terminated if the norm of the gradient drops below 10^{-8} . The obtained values for $q = (s, \delta)$, the number of BFGS iterations and the number of evaluations of the functional are shown in Table 1. We observe that the number of required iterations barely changes as the mesh is refined.

7.5. Convergence with respect to regularization. In the same setting as in Section 7.3.2, we solve test problem **II** in $n = 1$ dimensions for a sequence of regularization parameters α and β . The obtained values for $q = (s, \delta)$, the number of BFGS iterations and the number of evaluations of the functional are shown in Table 2. We observe that as $\alpha, \beta \rightarrow 0$, the recovered values for s and δ tend towards (s^*, δ^*) .

7.6. Noisy data. We again solve test problem **II** in 1D, but, as opposed to Section 7.3.2, augment the data u_d by a pointwise normal distributed noise: $u_d(x) =$

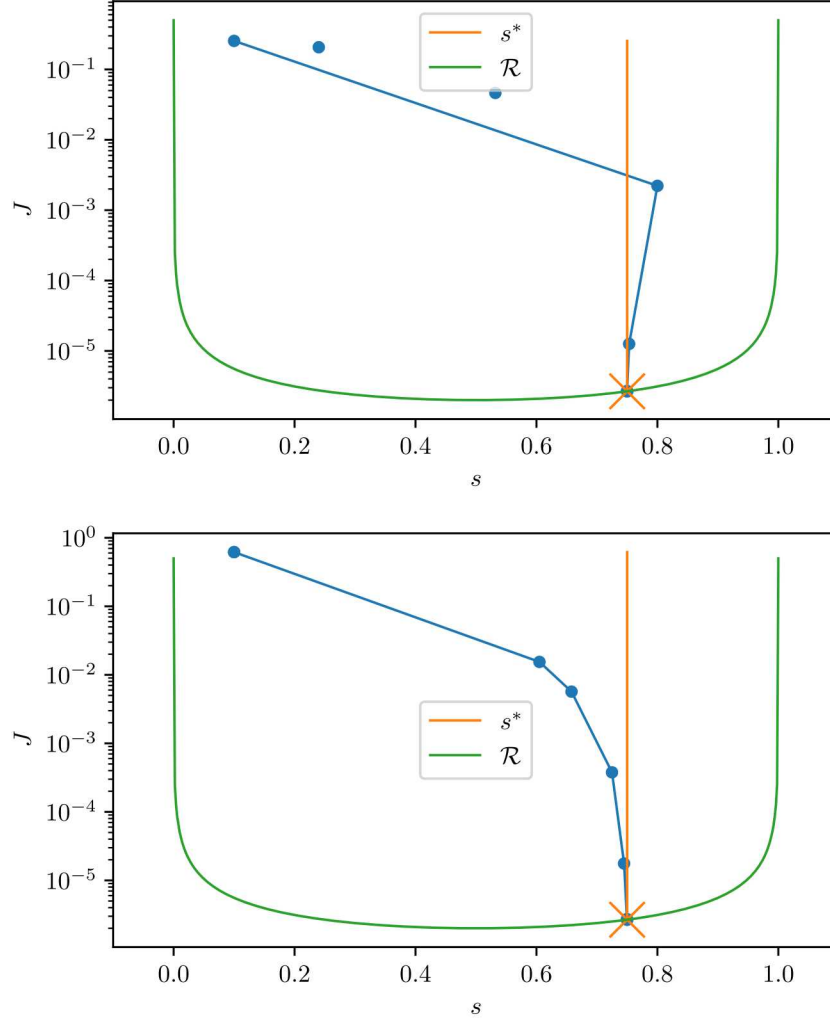


FIGURE 7.5. Value of the cost functional at all functional evaluations (\bullet) and for BFGS iterates ($-$) in 1D (*top*) and 2D (*bottom*) for test problem **I** and $u_d = u_{ex}(0.75)$. The final iterate is marked with \times .

h	N	s	δ	iterations	evaluations
2^{-10}	2047	0.74976	0.90306	24	77
2^{-11}	4095	0.74976	0.90332	28	87
2^{-12}	8191	0.74976	0.90329	23	79
2^{-13}	16383	0.74975	0.90331	25	63
2^{-14}	32767	0.74975	0.90331	23	73

TABLE 1. Convergence study for test problem **II** in 1D, $\mathcal{R}(s, \delta) = \frac{\alpha}{s(1-s)} + \beta \frac{e^\delta}{\delta}$ with $\alpha = 5 \times 10^{-7}$ and $\beta = 10^{-6}$. The optimal parameter values without regularization are $s^* = 0.75$ and $\delta^* = 0.9$.

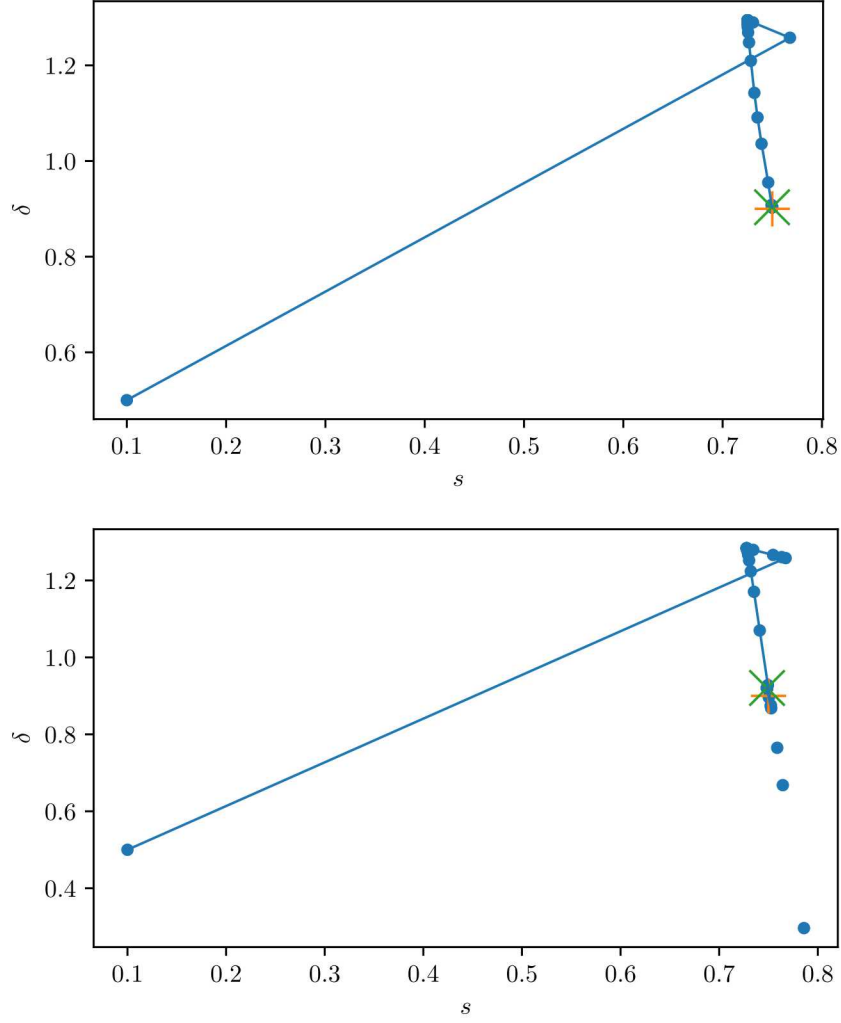


FIGURE 7.6. Cost functional evaluations (\bullet) and BFGS iterates ($-$) in 1D (*top*) and 2D (*bottom*) for test problem **II** and $\mathcal{R}(s, \delta) = \frac{\alpha}{s(1-s)} + \beta \frac{e^\delta}{\delta}$ with $\alpha = 5 \times 10^{-7}$ and $\beta = 10^{-6}$. The final iterate is marked with \times , the parameters q^* used to generate the data u_d is marked with $+$.

α	β	s	δ	iterations	evaluations
5e-04	1e-03	0.73096	1.1373	24	67
5e-05	1e-04	0.7394	1.0478	21	55
5e-06	1e-05	0.7478	0.92978	25	74
5e-07	1e-06	0.74976	0.90329	23	79
5e-08	1e-07	0.74998	0.90034	25	29
5e-09	1e-08	0.75	0.90004	25	87

TABLE 2. Convergence with respect to the regularization for test problem **II** in 1D, $\mathcal{R}(s, \delta) = \frac{\alpha}{s(1-s)} + \beta \frac{e^\delta}{\delta}$. The optimal parameter values without regularization are $s^* = 0.75$ and $\delta^* = 0.9$.

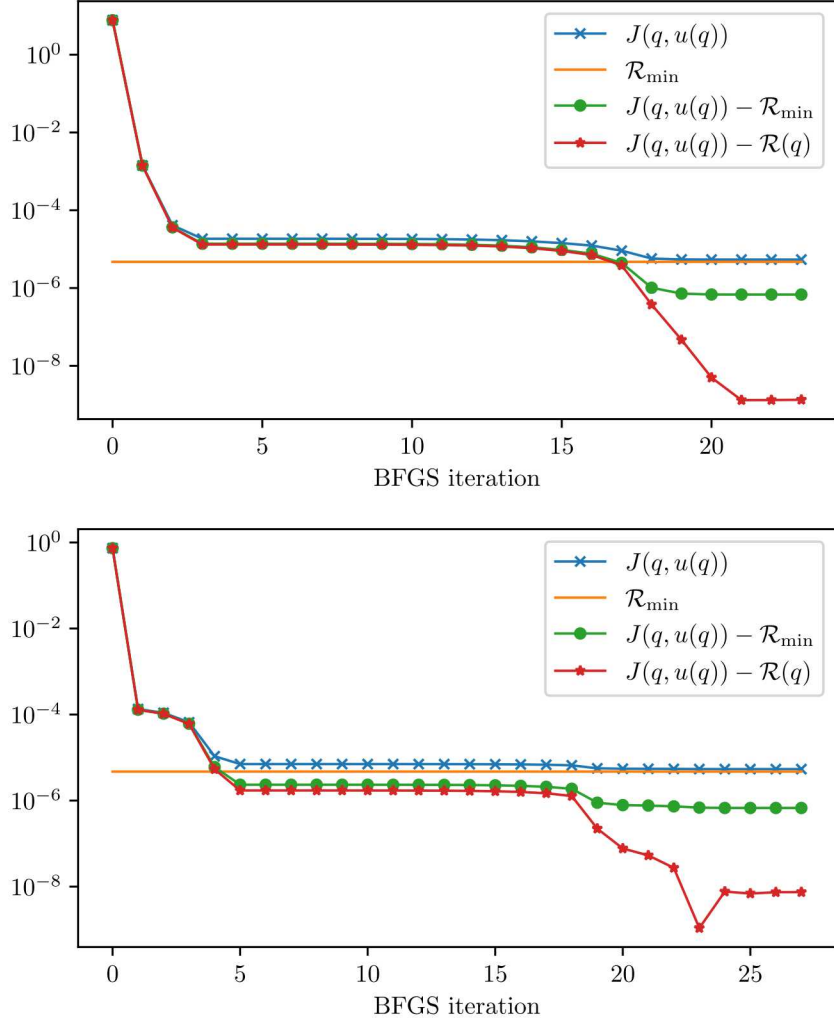


FIGURE 7.7. Values of the cost functional for the BFGS iterates in 1D (*top*) and 2D (*bottom*) for test problem **II** and $\mathcal{R}(s, \delta) = \frac{\alpha}{s(1-s)} + \beta \frac{e^\delta}{\delta}$ with $\alpha = 5 \times 10^{-7}$ and $\beta = 10^{-6}$. $\mathcal{R}_{\min} = \min_q \mathcal{R}(q)$.

$u_h(q^*)(x) + \mathcal{N}(0, \sigma^2)$. We observe that as $\sigma \rightarrow 0$, both s and δ converge towards their respective values obtained without noise.

8. CONCLUSION

We presented a mathematically rigorous approach to parameter identification for nonlocal models featuring kernels of fractional type. Our careful analysis of well-posedness and regularity properties of state and adjoint equations allows for a deeper understanding of the identification problem and for the design of suitable optimization techniques. More specifically, by introducing an approximation, via interpolation, of the bilinear form and its derivative, we are able to obtain highly accurate approximations of the gradients with a nearly optimal complexity. We stress that the impact of this approximation goes beyond the scope of this paper and can potentially impact any numerical framework that involves parametrized nonlocal

σ	s	δ	iterations	evaluations
1	0.8486	0.10847	35	105
2^{-1}	0.79461	0.45007	32	136
2^{-2}	0.76898	0.69291	28	67
2^{-3}	0.75936	0.79458	27	69
2^{-4}	0.75456	0.84792	23	71
2^{-5}	0.75217	0.87517	20	69
0	0.74976	0.90329	23	79

TABLE 3. Test problem **II** in 1D and $\mathcal{R}(s, \delta) = \frac{\alpha}{s(1-s)} + \beta \frac{e^\delta}{\delta}$ with $\alpha = 5 \times 10^{-7}$ and $\beta = 10^{-6}$ where the data u_d is pointwise disturbed by normally distributed noise with variance σ . The optimal parameter values without regularization are $s^* = 0.75$ and $\delta^* = 0.9$.

operators and their derivatives. As such, our approximation can be adapted to machine learning algorithms for the discovery of model parameters.

A natural follow up of this work is to consider a higher-dimensional parameter space and compare our approach with alternative identification techniques. Non-trivial extensions of this work include the generalization to the variable horizon case and the variable fractional order case. The latter problem is particularly challenging, as regularity properties of the associated state equation have not been fully analyzed.

REFERENCES

- [1] M. AINSWORTH AND C. GLUSA, *Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver*, Computer Methods in Applied Mechanics and Engineering, 327 (2017), pp. 4–35.
- [2] M. AINSWORTH AND C. GLUSA, *Towards an efficient finite element method for the integral fractional Laplacian on polygonal domains*, in Contemporary Computational Mathematics-A Celebration of the 80th Birthday of Ian Sloan, Springer, 2018, pp. 17–57.
- [3] B. ALALI AND M. GUNZBURGER, *Peridynamics and material interfaces*, Journal of Elasticity, 120 (2015), pp. 225–248.
- [4] H. ANTIL, E. OTAROLA, AND A. SALGADO, *Optimization with respect to order in a fractional diffusion model: Analysis, approximation and algorithmic aspects*, Journal of Scientific Computing, 77 (2018), pp. 204 – 224.
- [5] H. ANTIL AND M. WARMA, *Optimal control of fractional semilinear PDEs*, ESAIM Control Optimisation and Calculus of Variations, (2019). To appear.
- [6] S. BARTELS AND N. WEBER, *Parameter learning and fractional differential operators: application in image regularization and decomposition*, arXiv preprint arXiv:2001.03394, (2020).
- [7] D. A. BENSON, S. W. WHEATCRAFT, AND M. M. MEERSCHAERT, *Application of a fractional advection-dispersion equation*, Water Resources Research, 36 (2000), pp. 1403–1412.
- [8] O. BURKOVSKA AND M. GUNZBURGER, *Regularity analyses and approximation of nonlocal variational equality and inequality problems*, Journal of Mathematical Analysis and Applications, 478 (2019), pp. 1027 – 1048.
- [9] O. BURKOVSKA AND M. GUNZBURGER, *Affine Approximation of Parametrized Kernels and Model Order Reduction for Nonlocal and Fractional Laplace Models*, SIAM Journal on Numerical Analysis, 58 (2020), pp. 1469–1494.
- [10] O. BURKOVSKA AND M. GUNZBURGER, *On a nonlocal Cahn–Hilliard model permitting sharp interfaces*. arXiv:2004.14379, 2020.
- [11] G. CAPODAGLIO, M. D’ELIA, P. BOCHEV, AND M. GUNZBURGER, *An energy-based coupling approach to nonlocal interface problems*, Computers and Fluids, (2019). To appear.
- [12] C. CORTAZAR, M. ELGUETA, J. ROSSI, AND N. WOLANSKI, *How to approximate the heat equation with Neumann boundary conditions by nonlocal diffusion problems*, Archive for Rational Mechanics and Analysis, 187 (2008), pp. 137–156.

- [13] O. DEFETERLI, M. D'ELIA, Q. DU, M. GUNZBURGER, R. LEHOUCQ, AND M. M. MEERSCHAERT, *Fractional diffusion on bounded domains*, Fractional Calculus and Applied Analysis, 18 (2015), pp. 342–360.
- [14] M. D'ELIA, Q. DU, C. GLUSA, M. GUNZBURGER, X. TIAN, AND Z. ZHOU, *Numerical methods for nonlocal and fractional models*, Acta Numerica, (2020). To appear.
- [15] M. D'ELIA, Q. DU, M. GUNZBURGER, AND R. LEHOUCQ, *Nonlocal convection-diffusion problems on bounded domains and finite-range jump processes*, Computational Methods in Applied Mathematics, 29 (2017), pp. 71–103.
- [16] M. D'ELIA, C. GLUSA, AND E. OTÁROLA, *A priori error estimates for the optimal control of the integral fractional Laplacian*, SIAM Journal on Control and Optimization, 57 (2019), pp. 2775–2798.
- [17] M. D'ELIA, M. GULIAN, H. OLSON, AND G. E. KARNIADAKIS, *A unified theory of fractional, nonlocal, and weighted nonlocal vector calculus*. arXiv:2005.07686, 2020.
- [18] M. D'ELIA AND M. GUNZBURGER, *The fractional Laplacian operator on bounded domains as a special case of the nonlocal diffusion operator*, Computers and Mathematics with applications, 66 (2013), pp. 1245–1260.
- [19] ———, *Optimal distributed control of nonlocal steady diffusion problems*, SIAM Journal on Control and Optimization, 55 (2014), pp. 667–696.
- [20] ———, *Identification of the diffusion parameter in nonlocal steady diffusion problems*, Applied Mathematics and Optimization, 73 (2016), pp. 227–249.
- [21] M. D'ELIA, M. GUNZBURGER, AND C. VOLLMAN, *A cookbook for finite element methods for nonlocal problems, including quadrature rule choices and the use of approximate neighborhoods*. arXiv:2005.10775, 2020.
- [22] M. D'ELIA, J.-C. D. LOS REYES, AND A. M. TRUJILLO, *Bilevel parameter optimization for nonlocal image denoising models*. arXiv:1912.02347, 2019.
- [23] M. D'ELIA, X. TIAN, AND Y. YU, *A physically-consistent, flexible and efficient strategy to convert local boundary conditions into nonlocal volume constraints*, Accepted for publication in *SIAM Journal of Scientific Computing*, (2020).
- [24] Q. DU, M. GUNZBURGER, R. LEHOUCQ, AND K. ZHOU, *Analysis and approximation of nonlocal diffusion problems with volume constraints*, SIAM Review, 54 (2012), pp. 667–696.
- [25] Q. DU, L. JU, X. LI, AND Z. QIAO, *Stabilized linear semi-implicit schemes for the nonlocal Cahn–Hilliard equation*, Journal of Computational Physics, 363 (2018), pp. 39 – 54.
- [26] Q. DU AND J. YANG, *Asymptotically Compatible Fourier Spectral Approximations of Nonlocal Allen–Cahn Equations*, SIAM Journal on Numerical Analysis, 54 (2016), pp. 1899–1919.
- [27] A. ERN AND J.-L. GUERMOND, *Theory and Practice of Finite Elements.*, Applied Mathematical Sciences 159. New York, NY: Springer, 2004.
- [28] M. FELSINGER, M. KASSMANN, AND P. VOIGT, *The Dirichlet problem for nonlocal operators*, Mathematische Zeitschrift, 279 (2015), pp. 779–809.
- [29] R. K. GETTOOR, *First passage times for symmetric stable processes in space*, Transactions of the American Mathematical Society, 101 (1961), pp. 75–90.
- [30] C. GLUSA AND E. OTÁROLA, *Optimal control of a parabolic fractional pde: analysis and discretization*. arXiv:1905.10002, 2019.
- [31] R. GORENFLO AND F. MAINARDI, *Fractional calculus: integral and differential equations of fractional order*, Fractals and Fractional Calculus in Continuum Mechanics, (1997), pp. 223–276.
- [32] G. GRUBB, *Fractional Laplacians on domains, a development of Hörmander's theory of μ -transmission pseudodifferential operators*, Adv. Math., 268 (2015), pp. 478–528.
- [33] M. GULIAN, M. RAISSI, P. PERDIKARIS, AND G. E. KARNIADAKIS, *Machine learning of space-fractional differential equations*, SIAM Journal on Scientific Computing, 41 (2019), pp. A2485–A2509.
- [34] A. LISCHKE, G. PANG, M. GULIAN, F. SONG, C. GLUSA, X. ZHENG, Z. MAO, W. CAI, M. M. MEERSCHAERT, M. AINSWORTH, AND G. E. KARNIADAKIS, *What is the fractional Laplacian? A comparative review with new results*, Journal of Computational Physics, 404 (2020).
- [35] M. M. MEERSCHAERT, J. MORTENSEN, AND S. W. WHEATCRAFT, *Fractional vector calculus for fractional advection–dispersion*, Physica A: Statistical Mechanics and its Applications, 367 (2006), pp. 181–190.
- [36] T. MENGESHA AND Q. DU, *Analysis of a scalar nonlocal peridynamic model with a sign changing kernel*, Discrete & Continuous Dynamical Systems-B, 18 (2013), pp. 1415–1437.
- [37] R. MUSINA AND A. I. NAZAROV, *On fractional Laplacians*, Communications in Partial Differential Equations, 39 (2014), pp. 1780–1790.
- [38] J. NOCEDAL AND S. J. WRIGHT, *Numerical optimization*, Springer Series in Operations Research and Financial Engineering, Springer, New York, second ed., 2006.

- [39] G. PANG, M. D'ELIA, M. PARKS, AND G. E. KARNIADAKIS, *nPINNs: nonlocal Physics-Informed Neural Networks for a parametrized nonlocal universal Laplacian operator. Algorithms and Applications*, Journal on Computational Physics, (2020). To appear.
- [40] G. PANG, L. LU, AND G. E. KARNIADAKIS, *fPINNs: Fractional physics-informed neural networks*, SIAM Journal on Scientific Computing, 41 (2019), pp. A2603–A2626.
- [41] G. PANG, P. PERDIKARIS, W. CAI, AND G. E. KARNIADAKIS, *Discovering variable fractional orders of advection–dispersion equations from field data using multi-fidelity Bayesian optimization*, Journal of Computational Physics, 348 (2017), pp. 694 – 714.
- [42] M. PASETTO, *Enhanced Meshfree Methods for Numerical Solution of Local and Nonlocal Theories of Solid Mechanics*, PhD thesis, UC San Diego, 2019.
- [43] R. SCHUMER, D. BENSON, M. MEERSCHAERT, AND S. WHEATCRAFT, *Eulerian derivation of the fractional advection-dispersion equation*, Journal of Contaminant Hydrology, 48 (2001), pp. 69–88.
- [44] R. SCHUMER, D. A. BENSON, M. M. MEERSCHAERT, AND B. BAEUMER, *Multiscaling fractional advection-dispersion equations and their solutions*, Water Resources Research, 39 (2003), pp. 1022–1032.
- [45] S. A. SILLING AND E. ASKARI, *A meshfree method based on the peridynamic model of solid mechanics*, Computers & structures, 83 (2005), pp. 1526–1535.
- [46] 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 (2017), pp. 70–93.
- [47] L. N. TREFETHEN, *Is Gauss Quadrature Better than Clenshaw–Curtis?*, SIAM Review, 50 (2008), pp. 67–87.
- [48] L. N. TREFETHEN, *Approximation theory and approximation practice*, vol. 164, Siam, 2019.
- [49] H. WANG, K. WANG, AND T. SIRCAR, *A direct $\mathcal{O}(N \log^2 N)$ finite difference method for fractional diffusion equations*, Journal of Computational Physics, 229 (2010), pp. 8095–8104.
- [50] H. YOU, Y. YU, N. TRASK, M. GULIAN, AND M. D'ELIA, *Data-driven learning of robust nonlocal physics from high-fidelity synthetic data*. arXiv:2005.10076, 2020.