OPTIMIZATION WITH RESPECT TO ORDER IN A FRACTIONAL DIFFUSION MODEL: ANALYSIS, APPROXIMATION AND ALGORITHMIC ASPECTS*

HARBIR ANTIL†, ENRIQUE OTÁROLA‡, AND ABNER J. SALGADO§

Abstract. We consider an identification problem, where the state u is governed by a fractional elliptic equation and the unknown variable corresponds to the order $s \in (0,1)$ of the underlying operator. We study the existence of an optimal pair $(\bar{s}, \bar{\mathsf{u}})$ and provide sufficient conditions for its local uniqueness. We develop semi-discrete and fully discrete algorithms to approximate the solutions to our identification problem and provide a convergence analysis. We present numerical illustrations that confirm and extend our theory.

Key words. optimal control problems, identification problems, fractional diffusion, bisection algorithm, finite elements, stability, fully–discrete methods, convergence.

AMS subject classifications. 26A33, 35J70, 49J20, 49K21, 49M25, 65M12, 65M15, 65M60.

1. Introduction.

Supported by the claim that they seem to better describe many processes; nonlocal models have recently become of great interest in the applied sciences and engineering. This is specially the case when long range (i.e., nonlocal) interactions are to be taken into consideration; we refer the reader to [2] for a far from exhaustive list of examples where such phenomena take place. However, the actual range and scaling laws of these interactions — which determines the order of the model— cannot always be directly determined from physical considerations. This is in stark contrast with models governed by partial differential equations (PDEs), which usually arise from a conservation law. This justifies the need to, on the basis of physical observations, identify the order of a fractional model.

In [12], for the first time, this problem was addressed. The authors studied the optimization with respect to the order of the spatial operator in a nonlocal evolution equation; existence of solutions as well as first and second order optimality conditions were addressed. The present work is a natural extension of these results under the stationary regime: we address the local uniqueness of minimizers and propose a numerical algorithm to approximate them. In addition, we study the convergence rates of our method.

To make matters precise, let Ω be an open and bounded domain in $\mathbb{R}^n$ $(n \geq 1)$ with Lipschitz boundary $\partial\Omega$. Given a desired state $\mathsf{u}_d : \Omega \to \mathbb{R}$ (the observations), we define the cost functional

$$J(s, \mathsf{u}) = \frac{1}{2}\|\mathsf{u} - \mathsf{u}_d\|^2_{L^2(\Omega)} + \varphi(s), \tag{1}$$

where, for some a and b satisfying that $0 \leq a < b \leq 1$, $s \in (a,b)$ and, $\varphi \in C^2(a,b)$

*HA has been supported in part by NSF grant DMS-1521590. EO has been supported in part by CONICYT through FONDECYT project 3160201. AJS has been supported in part by NSF grant DMS-1418784.

†Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, USA. (hantil@gmu.edu, http://math.gmu.edu/~hantil/).

‡Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. (enrique.otarola@usm.cl, http://eotarola.mat.utfsm.cl/).

§Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA. (asalgad1@utk.edu, http://www.math.utk.edu/~abnersg)

denotes a nonnegative convex function that satisfies

$$\lim_{s \downarrow a} \varphi(s) = +\infty = \lim_{s \uparrow b} \varphi(s). \tag{2}$$

Examples of functions with these properties are

$$\varphi(s) = \frac{1}{(s-a)(b-s)}, \qquad \varphi(s) = \frac{e^{\frac{1}{(b-s)}}}{s-a}.$$

We shall thus be interested in the following identification problem: Find $(\bar{s}, \bar{\mathsf{u}})$ such that

$$J(\bar{s}, \bar{\mathsf{u}}) = \min J(s, \mathsf{u}) \tag{3}$$

subject to the fractional state equation

$$(-\Delta)^s \mathsf{u} = \mathsf{f} \text{ in } \Omega, \tag{4}$$

where $(-\Delta)^s$ denotes a fractional power of the Dirichlet Laplace operator $-\Delta$. We immediately remark that, with no modification, our approach can be extended to problems where the state equation is $L^s \mathsf{u} = \mathsf{f}$, where $L\mathsf{w} = -\text{div}(A \nabla \mathsf{w})$, supplemented with homogeneous Dirichlet boundary conditions, as long as the diffusion coefficient A is fixed, bounded and symmetric. In principle, one could also consider optimization with respect to order s and the diffusion A, as this could accommodate for anisotropies in the diffusion process. We refer the reader to [8], and the references therein, for the case when $s = 1$ is fixed and the optimization is carried out with respect to A.

We now comment on the choice of a and b. The practical situation can be envisioned as the following: from measurements or physical considerations we have an expected range for the order of the operator, and we want to optimize within that range to best fit the observations. From the existence and optimality conditions point of view, there is no limitation on their values, as long as $0 \leq a < b \leq 1$. However, when we discuss the convergence of numerical algorithms, many of the estimates and arguments that we shall make blow up as $s \downarrow 0$ or $s \uparrow 1$ so we shall assume that $a > 0$ and $b < 1$. How to treat numerically the full range of s is currently under investigation.

Our presentation is organized as follows. The notation and functional setting is introduced in section 2, where we also briefly describe, in section 2.1, the definition of the fractional Laplacian. In section 3, we study the fractional identification problem (3)–(4). We analyze the differentiability properties of the associated control to state map (section 3.1) and derive existence results as well as first and second order optimality conditions and a local uniqueness result (section 3.2). Section 4 is dedicated to the design and analysis of a numerical algorithm to approximate the solution to (3)–(4). Finally, in section 5 we illustrate the performance of our algorithm on several examples.

2. Notation and preliminaries. Throughout this work Ω is an open, bounded and convex polytopal subset of $\mathbb{R}^n$ $(n \geq 1)$ with boundary $\partial\Omega$. The relation $X \lesssim Y$ indicates that $X \leq CY$, with a nonessential constant C that might change at each occurrence.

2.1. The fractional Laplacian. Spectral theory for the operator $-\Delta$ yields the existence of a countable collection of eigenpairs $\{\lambda_k, \varphi_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^+ \times H_0^1(\Omega)$ such that $\{\varphi_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of $L^2(\Omega)$ and an orthogonal basis of $H_0^1(\Omega)$ and

$$-\Delta\varphi_k = \lambda_k\varphi_k \text{ in } \Omega, \qquad \varphi_k = 0 \text{ on } \partial\Omega, \qquad k \in \mathbb{N}. \tag{5}$$

With this spectral decomposition at hand, we define the fractional powers of the Dirichlet Laplace operator, which for convenience we simply call the fractional Laplacian, as follows: For any $s \in (0,1)$ and $w \in C_0^\infty(\Omega)$,

$$(-\Delta)^s w := \sum_{k\in\mathbb{N}} \lambda_k^s w_k \varphi_k, \quad w_k = (w, \varphi_k)_{L^2(\Omega)} := \int_\Omega w\varphi_k \,\mathrm{d}x. \tag{6}$$

By density, this definition can be extended to the space

$$\mathbb{H}^s(\Omega) = \left\{ w = \sum_{k\in\mathbb{N}} w_k\varphi_k \in L^2(\Omega) : \sum_{k\in\mathbb{N}} \lambda_k^s w_k^2 < \infty \right\}, \tag{7}$$

which we endow with the norm

$$\|w\|_{\mathbb{H}^s(\Omega)} = \left(\sum_{k\in\mathbb{N}} \lambda_k^s w_k^2 \right)^{\frac{1}{2}}; \tag{8}$$

see [5, 6, 9] for details. The space $\mathbb{H}^s(\Omega)$ coincides with $[L^2(\Omega), H_0^1(\Omega)]_s$, i.e., the interpolation space between $L^2(\Omega)$ and $H_0^1(\Omega)$; see [1, Chapter 7]. For $s \in (0,1)$, we denote by $\mathbb{H}^{-s}(\Omega)$ the dual space to $\mathbb{H}^s(\Omega)$ and remark that it admits the following characterization:

$$\mathbb{H}^{-s}(\Omega) = \left\{ w = \sum_{k\in\mathbb{N}} w_k\varphi_k \in \mathcal{D}'(\Omega) : \sum_{k\in\mathbb{N}} \lambda_k^{-s} w_k^2 < \infty \right\}, \tag{9}$$

where $\mathcal{D}'(\Omega)$ denotes the space of distributions on Ω. Finally, we denote by $\langle\cdot,\cdot\rangle$ the duality pairing between $\mathbb{H}^s(\Omega)$ and $\mathbb{H}^{-s}(\Omega)$.

3. The fractional identification problem.

3. The fractional identification problem. In this section we study the existence of minimizers for the fractional identification problem (3)–(4), as well as optimality conditions. We begin by introducing the so-called control to state map associated with problem (3)–(4) and studying its differentiability properties. This will allow us to derive first order necessary and second order sufficient optimality conditions for our identification problem, as well as existence results.

3.1. The control to state map. In this subsection we study the differentiability properties of the control to state map $\mathcal{S}$ associated with (3)–(4), which we define as follows: Given a control $s \in (0,1)$, the map $\mathcal{S}$ associates to it the state $\mathsf{u} = \mathsf{u}(s)$ that solves problem (4) with the forcing term $\mathsf{f} \in \mathbb{H}^{-s}(\Omega)$. In other words,

$$\mathcal{S} : (0,1) \to \mathbb{H}^s(\Omega), \qquad s \mapsto \mathcal{S}(s) = \sum_{k\in\mathbb{N}} \lambda_k^{-s} \mathsf{f}_k \varphi_k, \tag{10}$$

where $\mathsf{f}_k = \langle \mathsf{f}, \varphi_k \rangle$ and $\{\lambda_k, \varphi_k\}_{k\in\mathbb{N}}$ are defined by (5). Since $\mathsf{f} \in \mathbb{H}^{-s}(\Omega)$, the characterization of the space $\mathbb{H}^{-s}(\Omega)$, given in (9), allows us to immediately conclude that the map $\mathcal{S}$ is well–defined; see also [6, Lemma 2.2].

Before embarking on the study of the smoothness properties of the map $\mathcal{S}$ we define, for $\lambda > 0$, the function $E_\lambda : (0,1) \to \mathbb{R}^+$ by

$$E_\lambda(s) = \lambda^{-s}, \qquad s \in (0,1). \tag{11}$$

A trivial computation reveals that

$$D_s^m E_\lambda(s) = (-1)^m \ln^m(\lambda) E_\lambda(s), \quad m \in \mathbb{N}, \tag{12}$$

from which immediately follows that, for $m \in \mathbb{N}$, we have the estimate

$$|D_s^m E_\lambda(s)| \lesssim s^{-m}, \tag{13}$$

where the hidden constant is independent of s, it remains bounded as $\lambda \uparrow \infty$, but blows up as $\lambda \downarrow 0$; compare with [12, eq. (2.27)].

With this auxiliary function at hand we proceed, following [12], to study the differentiability properties of the map $\mathcal{S}$. To begin we notice the inclusion $\mathcal{S}((0,1)) \subset L^2(\Omega)$ so we consider $\mathcal{S}$ as a map with range in $L^2(\Omega)$ and we will denote by $|||\cdot|||$ the norm of $\mathcal{L}(\mathbb{R}, L^2(\Omega))$.

THEOREM 1 (properties of $\mathcal{S}$). *Let $\mathcal{S} : (0,1) \to L^2(\Omega)$ be the control to state map, defined in (10), and assume that $\mathsf{f} \in L^2(\Omega)$. For every $s \in (0,1)$ we have that*

$$\|\mathcal{S}(s)\|_{L^2(\Omega)} \lesssim 1, \tag{14}$$

where the hidden constant depends on Ω and $\|\mathsf{f}\|_{L^2(\Omega)}$, but not on s. In addition, $\mathcal{S}$ is three times Fréchet differentiable; the first and second derivatives of $\mathcal{S}$ are characterized as follows: for $h_1, h_2 \in \mathbb{R}$, we have that

$$D_s\mathcal{S}(s)[h_1] = h_1 D_s \mathsf{u}(s), \qquad D_s^2\mathcal{S}(s)[h_1,h_2] = h_1 h_2 D_s^2 \mathsf{u}(s), \tag{15}$$

where

$$D_s \mathsf{u}(s) = -\sum_{k\in\mathbb{N}} \lambda_k^{-s} \ln(\lambda_k) \mathsf{f}_k \varphi_k, \qquad D_s^2 \mathsf{u}(s) = \sum_{k\in\mathbb{N}} \lambda_k^{-s} \ln^2(\lambda_k) \mathsf{f}_k \varphi_k.$$

Finally, for $m = 1,2,3$, we have

$$|||\, D_s^m \mathcal{S}(s)||| \lesssim s^{-m}, \tag{16}$$

where the hidden constants are independent of s.

Proof. Let $s \in (0,1)$. To shorten notation we set $\mathsf{u} = \mathcal{S}(s)$. Using (10) we have that

$$\|\mathsf{u}\|_{L^2(\Omega)}^2 = \sum_{k\in\mathbb{N}} \lambda_k^{-2s} \mathsf{f}_k^2 \le \lambda_1^{-2s} \|\mathsf{f}\|_{L^2(\Omega)}^2, \tag{17}$$

where we used that, for all $k \in \mathbb{N}$, $0 < \lambda_1 \le \lambda_k$. Since $\sup_{s\in[0,1]} \lambda_1^{-2s}$ is bounded, we obtain (14).

We now define, for $N \in \mathbb{N}$, the partial sum $w_N = \sum_{k=1}^N \lambda_k^{-s} \mathsf{f}_k \varphi_k$. Evidently, as $N \to \infty$, we have that $w_N \to \mathsf{u}$ in $L^2(\Omega)$. Moreover, differentiating with respect to s we immediately obtain, in light of (12), the expression

$$D_s w_N = -\sum_{k\le N} \lambda_k^{-s} \ln(\lambda_k) \mathsf{f}_k \varphi_k,$$

and, using (12) and (13), that

$$\|D_s w_N\|^2_{L^2(\Omega)} = \sum_{k\le N} |D_s E_{\lambda_k}(s)|^2 \mathsf{f}_k^2 \lesssim \frac{1}{s^2}\|\mathsf{f}\|^2_{L^2(\Omega)},$$

where we used, again, that the eigenvalues are strictly away from zero. This estimate allows us to conclude that, as $N \to \infty$, we have $D_s w_N \to D_s \mathsf{u}$ in $L^2(\Omega)$ and the bound

$$\|D_s \mathsf{u}(s)\|_{L^2(\Omega)} \lesssim s^{-1}\|\mathsf{f}\|_{L^2(\Omega)}. \tag{18}$$

Let us now prove that $\mathcal{S}$ is Fréchet differentiable and that (15) holds. Taylor's theorem, in conjunction with (12), yields that, for every $l \in \mathbb{N}$ and $h_1 \in \mathbb{R}$, we have

$$e_{l,s} := |E_{\lambda_l}(s+h_1) - E_{\lambda_l}(s) - D_s E_{\lambda_l}(s) h_1| = \frac{1}{2} h_1^2 |D_s^2 E_{\lambda_l}(\theta)|,$$

for some $\theta \in (s-|h_1|, s+|h_1|)$. Now, if $|h_1| < s/2$, we have that $\theta^{-2} < 4s^{-2}$, and thus, in view of estimate (13), that

$$e_{l,s} = \frac{1}{2} h_1^2 |D_s^2 E_{\lambda_l}(\theta)| \lesssim h_1^2 s^{-2}.$$

This last estimate allows us to write

$$\|\mathcal{S}(s+h_1) - \mathcal{S}(s) - D_s \mathsf{u}(s) h_1\|^2_{L^2(\Omega)} = \sum_{k\in\mathbb{N}} e_{k,s}^2 \mathsf{f}_k^2 \lesssim h_1^4 s^{-4} \|\mathsf{f}\|^2_{L^2(\Omega)},$$

where the hidden constant is independent of h_1 and s. The previous estimate shows that $\mathcal{S} : (0,1) \to L^2(\Omega)$ is Fréchet differentiable and that $D_s\mathcal{S}(s)[h_1] = h_1 D_s \mathsf{u}(s)$. Finally, using (18), we conclude, estimate (16) for $m = 1$.

Similar arguments can be applied to show the higher order Fréchet differentiability of $\mathcal{S}$ and to derive estimate (16) for $m = 2, 3$. For brevity, we skip the details. □

3.2. Existence and optimality conditions. We now proceed to study the existence of a solution to problem (3)–(4) as well as to characterize it via first and second order optimality conditions. We begin by defining the reduced cost functional

$$f(s) = J(s, \mathcal{S}(s)), \tag{19}$$

where $\mathcal{S}$ denotes the control to state map defined in (10) and J is defined as in (1); we recall that $\varphi \in C^2(a,b)$. Notice that, owing to Theorem 1, $\mathcal{S}$ is three times Fréchet differentiable. Consequently, $f \in C^2(a,b)$ and, moreover, it verifies conditions similar to (2). These properties will allow us to show existence of an optimal control. We begin with a definition.

DEFINITION 2 (optimal pair). *The pair $(\bar{s}, \bar{\mathsf{u}}(\bar{s})) \in (a,b) \times \mathbb{H}^{\bar{s}}(\Omega)$ is called optimal for problem (3)–(4) if $\bar{\mathsf{u}}(\bar{s}) = \mathcal{S}(\bar{s})$ and*

$$f(\bar{s}) \le f(s),$$

for all $(s, \mathsf{u}(s)) \in (a,b) \times \mathbb{H}^s(\Omega)$ such that $\mathsf{u}(s) = \mathcal{S}(s)$.

THEOREM 3 (existence). *There is an optimal pair $(\bar{s}, \bar{\mathsf{u}}(\bar{s})) \in (a,b) \times \mathbb{H}^{\bar{s}}(\Omega)$ for problem (3)–(4).*

Proof. Let $\{a_l\}_{l\in\mathbb{N}}, \{b_l\}_{l\in\mathbb{N}} \subset (a,b)$ be such that, for every $l \in \mathbb{N}$, $a < a_{l+1} < a_l < b_l < b_{l+1} < b$ and $a_l \to a$, $b_l \to b$ as $l \to \infty$. Denote $I_l = [a_l, b_l]$ and consider the problem of finding

$$s_l = \operatorname*{argmin}_{s\in I_l} f(s).$$

The properties of f guarantee its existence. Notice that, since the intervals I_l are nested, we have

$$f(s_m) \leq f(s_l), \qquad m \geq l.$$

We have thus constructed a sequence $\{s_l\}_{l\in\mathbb{N}} \subset (a,b)$ from which we can extract a convergent subsequence, which we still denote by the same $\{s_l\}_{l\in\mathbb{N}}$, such that $s_l \to \bar{s} \in [a,b]$. We claim that f attains its infimum, over (a,b), at the point $\bar{s}$.

Let us begin by showing that, in fact, $\bar{s} \in (a,b)$. The decreasing property of $\{f(s_l)\}_{l\in\mathbb{N}}$ shows that

$$f(\bar{s}) \leq f(s_l), \quad \forall l \in \mathbb{N},$$

which, if $\bar{s} = a$ or $\bar{s} = b$, would lead to a contradiction with the fact that $f(s) \geq \varphi(s)$ and (2).

Let $s_\star$ be any point of (a,b). The construction of the intervals I_l guarantee that there is $L \in \mathbb{N}$ for which $s_\star \in I_l$ whenever $l > L$. Therefore, we have

$$f(\bar{s}) \leq f(s_l) = \min_{s\in I_l} f(s) \leq f(s_\star).$$

Which shows that $\bar{s}$ is a minimizer.

Since $\mathcal{S}$, as a map from (a,b) to $L^2(\Omega)$, is continuous — even differentiable — we see that there is $\bar{\mathsf{u}} \in L^2(\Omega)$, for which $\mathcal{S}(s_l) \to \bar{\mathsf{u}}$ in $L^2(\Omega)$ as $l \to \infty$. Let us now show that, indeed, $\bar{\mathsf{u}} \in \mathbb{H}^{\bar{s}}(\Omega)$ and that it satisfies the state equation.

Set $\bar{\mathsf{u}} = \sum_{k\in\mathbb{N}} \bar{\mathsf{u}}_k \varphi_k$ and notice that, as $l \to \infty$,

$$(\mathcal{S}(s_l) - \bar{\mathsf{u}}, \varphi_m)_{L^2(\Omega)} = \lambda_m^{-s_l}\mathsf{f}_m - \bar{\mathsf{u}}_m \to \lambda_m^{-\bar{s}}\mathsf{f}_m - \bar{\mathsf{u}}_m.$$

Therefore $\bar{\mathsf{u}}_m = \lambda_m^{-\bar{s}}\mathsf{f}_m$. This shows that $\bar{\mathsf{u}} \in \mathbb{H}^{\bar{s}}(\Omega)$ and that $\bar{\mathsf{u}}$ solves (4).

The result is thus proved. ∎

We now provide first order necessary and second order sufficient optimality conditions for the identification problem (3)–(4).

THEOREM 4 (optimality conditions). *Let $(\bar{s}, \bar{\mathsf{u}})$ be an optimal pair for problem (3)–(4). Then it satisfies the following first order necessary optimality condition*

$$(\bar{\mathsf{u}} - \mathsf{u}_d, D_s\bar{\mathsf{u}})_{L^2(\Omega)} + \varphi'(\bar{s}) = 0. \tag{20}$$

On the other hand, if $(\bar{s}, \bar{\mathsf{u}})$, with $\bar{\mathsf{u}} = \mathcal{S}(\bar{s})$, satisfies (20) and, in addition, the second order optimality condition

$$(D_s\bar{\mathsf{u}}, D_s\bar{\mathsf{u}})_{L^2(\Omega)} + (\bar{\mathsf{u}} - \mathsf{u}_d, D_s^2\bar{\mathsf{u}})_{L^2(\Omega)} + \varphi''(\bar{s}) > 0 \tag{21}$$

holds, then $(\bar{s}, \bar{\mathsf{u}})$ is an optimal pair.

Proof. Since, as shown in Theorem 3, $\bar{s} \in (a,b)$, the first order optimality condition reads:

$$f'(\bar{s}) = (\mathcal{S}(\bar{s}) - \mathsf{u}_d, D_s\mathcal{S}(\bar{s}))_{L^2(\Omega)} + \varphi'(\bar{s}) = 0. \tag{22}$$

The characterization of the first order derivative of $\mathcal{S}$, given in Theorem 1, allows us to conclude (20). A similar computation reveals that

$$f''(\bar{s}) = (D_s\mathcal{S}(\bar{s}), D_s\mathcal{S}(\bar{s}))_{L^2(\Omega)} + (\mathcal{S}(\bar{s}) - \mathsf{u}_d, D_s^2\mathcal{S}(\bar{s}))_{L^2(\Omega)} + \varphi''(\bar{s}). \tag{23}$$

Using, again, the characterization for the first and second order derivatives of $\mathcal{S}$ given in Theorem 1 we obtain (21). This concludes the proof. ∎

Let us now provide a sufficient condition for local uniqueness of the optimal identification parameter $\bar{s}$. To accomplish this task we assume that the function φ, that defines the functional J in (1), is strongly convex with parameter ξ, i.e., for all points s_1, s_2 in (a,b), we have that

$$(\varphi'(s_1) - \varphi'(s_2)) \cdot (s_1 - s_2) \geq \xi|s_1 - s_2|^2. \tag{24}$$

We thus present the following result.

Lemma 5 (second–order sufficient conditions). *Let $\bar{s}$ be optimal for problem (3)–(4) and f be defined as in (19). If $\|\mathsf{f}\|_{L^2(\Omega)}$ and $\|\mathsf{u}_d\|_{L^2(\Omega)}$ are sufficiently small, then there exist a constant $\vartheta > 0$ such that*

$$f''(\bar{s}) \geq \vartheta. \tag{25}$$

Proof. On the basis of (23), we invoke the strong convexity of φ to conclude that

$$f''(\bar{s}) \geq \|D_s\mathcal{S}(\bar{s})\|_{L^2(\Omega)}^2 + (\mathcal{S}(\bar{s}) - \mathsf{u}_d, D_s^2\mathcal{S}(\bar{s}))_{L^2(\Omega)} + \xi.$$

It thus suffices to control the term $(\mathcal{S}(\bar{s}) - \mathsf{u}_d, D_s^2\mathcal{S}(\bar{s}))_{L^2(\Omega)}$; and to do so we use the estimates of Theorem 1. In fact, we have that

$$|(\mathcal{S}(\bar{s}) - \mathsf{u}_d, D_s^2\mathcal{S}(\bar{s}))_{L^2(\Omega)}| \leq C_1 \left(C_2\|\mathsf{f}\|_{L^2(\Omega)} + \|\mathsf{u}_d\|_{L^2(\Omega)}\right) \bar{s}^{-2}\|\mathsf{f}\|_{L^2(\Omega)},$$

where C_1 and C_2 depend on Ω and the operator $-\Delta$ but are independent of $\bar{s}$, f and u_d. Since Theorem 3 guarantees that $\bar{s} \in (a,b)$, we conclude that the right hand side of the previous expression is bounded. This, in view of the fact that $\|\mathsf{f}\|_{L^2(\Omega)}$ and $\|\mathsf{u}_d\|_{L^2(\Omega)}$ are sufficiently small, concludes the proof. ∎

As a consequence of the previous Lemma we derive, for the reduced cost functional f, a convexity property that will be important to analyze the fully discrete scheme of section 4, and a quadratic growth condition that implies the local uniqueness of $\bar{s}$.

Corollary 6 (convexity and quadratic growth). *Let $\bar{s}$ be optimal for problem (3)–(4) and f be defined as in (19). If $\|\mathsf{f}\|_{L^2(\Omega)}$ and $\|\mathsf{u}_d\|_{L^2(\Omega)}$ are sufficiently small, then there exists $\delta > 0$ such that*

$$(f'(s) - f'(\bar{s})) \cdot (s - \bar{s}) \geq \frac{\vartheta}{2}|s - \bar{s}|^2 \qquad \forall s \in (a,b) \cap (\bar{s} - \delta, \bar{s} + \delta), \tag{26}$$

where ϑ is the constant that appears in (25). In addition, we have the quadratic growth condition

$$f(s) \geq f(\bar{s}) + \frac{\vartheta}{4}|s - \bar{s}|^2 \qquad \forall s \in (a,b) \cap (\bar{s} - \delta, \bar{s} + \delta). \tag{27}$$

In particular, f has a local minimum at $\bar{s}$. Moreover, this minimum is unique in $(\bar{s} - \delta, \bar{s} + \delta) \cap (a,b)$.

Proof. Estimates (26) and (27) follow immediately from an application of Taylor's theorem and estimate (25); see [14, Theorem 4.23] for details. The local uniqueness follows immediately from (27). ∎

4. A numerical scheme for the fractional identification problem. In this section we propose a numerical method that approximates the solution to the fractional identification problem (3)–(4). To be able to provide a convergence analysis of the proposed method we make the following assumption.

Assumption 7 (compact subinterval). The optimization bounds a and b satisfy

$$0 < a < b < 1.$$

The scheme that we propose below is based on the discretization of the first order optimality condition (20): we discretize the first derivative $D_s \mathsf{u}(s)$ in (20) using a centered difference and then we approximate the solution to the state equation (4) with the finite element techniques introduced in [9].

4.1. Discretization in s. To set the ideas, we first propose a scheme that only discretizes the variable s and analyze its convergence properties. We begin by introducing some terminology. Let $\sigma > 0$ and $s \in (a,b)$ such that $s \pm \sigma \in (a,b)$. We thus define, for $\psi : (a,b) \to \mathbb{R}$, the centered difference approximation of $D_s\psi$ at s by

$$d_\sigma \psi(s) := \frac{\psi(s+\sigma) - \psi(s-\sigma)}{2\sigma}. \tag{28}$$

If $\psi \in C^3(a,b)$, a basic application of Taylor's theorem immediately yields the estimate

$$|D_s\psi(s) - d_\sigma\psi(s)| \leq \frac{\sigma^2}{3}\|D_s^3\psi\|_{L^\infty(s-\sigma,s+\sigma)}. \tag{29}$$

We also define the function $j_\sigma : (a,b) \to \mathbb{R}$ by

$$j_\sigma(s) = (\mathsf{u}(s) - \mathsf{u}_d, d_\sigma \mathsf{u}(s))_{L^2(\Omega)} + \varphi'(s), \tag{30}$$

where $\mathsf{u}(s)$ denotes the solution to (4). Finally, a point $s_\sigma \in (a,b)$ for which

$$j_\sigma(s_\sigma) = 0, \tag{31}$$

will serve as an approximation of the optimal parameter $\bar{s}$.

Notice that, in (30), the definition of j_σ coincides with the first order optimality condition (20), when we replace the derivative of the state, i.e., $D_s\mathsf{u}$, by its centered difference approximation, as defined in (28). The existence of s_σ will be shown by proving convergence of Algorithm 1 which, essentially, is a bisection algorithm. In addition, if the algorithm reaches line 14, since $j_\sigma \in C([s_l, s_r])$ and it takes values of different signs at the endpoints, the intermediate value theorem guarantees that the bisection step will produce a sequence of values that we use to approximate the root of j_σ. It remains then to show that we can eventually find the requisite interval $[s_l, s_r] \subset (a,b)$. This is the content of the following result.

LEMMA 8 (root isolation). *If σ is sufficiently small, there exist s_l and s_r in (a,b) such that $j_\sigma(s_l) < 0$ and $j_\sigma(s_r) > 0$, i.e., the root isolation step in Algorithm 1 terminates.*

Proof. We begin the proof by noticing that, for $s \in (\sigma, 1-\sigma) \subset (a,b)$, the estimates of Theorem 1 immediately yield the existence of a constant $C > 0$ such that

$$\left|(\mathsf{u}(s) - \mathsf{u}_d, d_\sigma\mathsf{u}(s))_{L^2(\Omega)}\right| \leq \frac{C}{\sigma}, \tag{32}$$

Algorithm 1 Bisection algorithm.

$0 < \sigma \ll 1$ and set $s_l, s_r \in (a, b)$, with $s_l < s_r$.; ▷ Initialization
▷ We take care of possible degenerate cases
if $j_\sigma(s_l) = 0$ **then**
  $s_\sigma = s_l$;
end if
if $j_\sigma(s_r) = 0$ **then**
  $s_\sigma = s_r$;
end if
▷ Root isolation
while $j_\sigma(s_r) < 0$ **do**
  $s_r := s_r + \sigma$;
end while
while $j_\sigma(s_l) > 0$ **do**
  $s_l := s_l - \sigma$;
end while
▷ Bisection
$k = 1$;
repeat
  $s_k = \frac{1}{2}(s_l + s_r)$;
  if $j_\sigma(s_k) = 0$ **then**
    $s_\sigma = s_k$;
    break; ▷ The solution has been found
  end if
  if $j_\sigma(s_l) j_\sigma(s_k) > 0$ **then** ▷ Sign check
    $s_l = s_k$;
  else
    $s_r = s_k$;
  end if
  $k = k + 1$;
until forever

where C depends on Ω, u_d and f but not on s or σ.

On the other hand, since property (2) implies that $\varphi'(s) \to -\infty$ as $s \downarrow a$, we deduce the existence of $\epsilon_l > 0$ such that, if $s \in (a, a+\epsilon_l)$ then $\varphi'(s) < -C/\sigma$. Assume that $\sigma < \epsilon_l$. Consequently, in view of the bound (32), definition (30) immediately implies that, for every $s \in (a+\sigma, a+\epsilon_l)$, we have the estimate

$$j_\sigma(s) \leq \frac{C}{\sigma} + \varphi'(s) < 0.$$

Similar arguments allow us to conclude the existence of $\epsilon_r > 0$ such that, if $s \in (b - \epsilon_r, b)$ then $\varphi'(s) > C/\sigma$. Assume that $\sigma < \epsilon_r$. We thus conclude that, for every $s \in (b - \epsilon_r, b - \sigma)$, we have the bound

$$j_\sigma(s) \geq -\frac{C}{\sigma} + \varphi'(s) > 0.$$

In light of the previous estimates we thus conclude that, for $\sigma < \min\{\epsilon_l, \epsilon_r\}$, we can find s_l and s_r in (a, b) such that $j_\sigma(s_l) < 0$ and $j_\sigma(s_r) > 0$. This concludes the proof. □

From Lemma 8 we immediately conclude that the bisection algorithm can be performed and exhibits the following convergence property.

LEMMA 9 (convergence rate: bisection method). *The sequence* $\{s_k\}_{k\geq 1}$ *generated by the bisection algorithm satisfies*

$$|s_\sigma - s_k| \lesssim 2^{-k}. \tag{33}$$

In addition, there exists s_l *and* s_r *such that* $a < s_l < s_r < b$ *and* $s_\sigma \in (s_l, s_r)$.

The results of Lemmas 8 and 9 guarantee that, for a fixed σ, the bisection algorithm can be performed and exhibits a convergence rate dictated by (33). Let us now discuss the convergence properties, as $\sigma \to 0$, of this semi-discrete method. We begin with two technical lemmas.

LEMMA 10 (convergence of j_σ). *Let* $j_\sigma : (a,b) \to \mathbb{R}$ *be defined as in* (30)*, then,* $j_\sigma \rightrightarrows f'$ *on* (a,b) *as* $\sigma \to 0$.

Proof. From the definitions we obtain that, whenever $s \in (a,b)$

$$\begin{aligned} |f'(s) - j_\sigma(s)| &= \left|(\mathsf{u}(s) - \mathsf{u}_d, D_s\mathsf{u}(s) - d_\sigma \mathsf{u}(s))_{L^2(\Omega)}\right| \\ &\lesssim \sup_{s\in[a,b]} \|D_s\mathsf{u}(s) - d_\sigma\mathsf{u}(s)\|_{L^2(\Omega)}, \end{aligned}$$

where the hidden constant depends on u_d and estimate (14). Since, from Theorem 1 we know that the control to state map is three times differentiable, we can conclude that

$$\|D_s\mathsf{u}(s) - d_\sigma\mathsf{u}(s)\|_{L^2(\Omega)} \lesssim \frac{\sigma^2}{a^3},$$

where we used a formula analogous to (29) and estimate (16). The fact that $a > 0$ (Assumption 7) allows us to conclude. ◻

With the uniform convergence of j_σ at hand, we can obtain the convergence of its roots to parameters that are optimal.

LEMMA 11 (convergence of s_σ). *The family* $\{s_\sigma\}_{\sigma>0}$ *contains a convergent subsequence. Moreover, the limit of any convergent subsequence satisfies* (20).

Proof. The existence of a convergent subsequence follows from the fact that $\{s_\sigma\}_{\sigma>0} \subset [a,b]$. Moreover, as in Theorem 3, we conclude that the limit is in (a,b). Let us now show that any limit satisfies (20).

By Lemma 10, for any $\varepsilon > 0$, if σ is sufficiently small, we have that

$$|f'(s_\sigma)| = |f'(s_\sigma) - j_\sigma(s_\sigma)| < \varepsilon$$

which implies that $f'(s_\sigma) \to 0$ as $\sigma \to 0$. Let now $\{s_{\sigma_k}\}_{k\in\mathbb{N}} \subset \{s_\sigma\}$ be a convergent subsequence. Denote the limit point by $\underline{s} \in (a,b)$. By continuity of f' we have $f'(s_{\sigma_k}) \to f'(\underline{s})$ which implies that

$$f'(\underline{s}) = 0. \qquad ◻$$

Remark 12 (stronger convergence). It is expected that we cannot prove more than convergence up to subsequences, since there might be more than one s that satisfies (20). If there is a unique optimal s, then the previous result implies that the family $\{s_\sigma\}_{\sigma>0}$ converges to it.

In what follows, to simplify notation, we denote by $\{s_\sigma\}_{\sigma>0}$ any convergent subfamily. The next result then provides a rate of convergence.

THEOREM 13 (convergence rate in σ). *Let $\bar{s}$ denote a solution to the identification problem (3)–(4) and let s_σ be its approximation defined as the solution to equation (31). If σ is sufficiently small then we have*

$$|\bar{s} - s_\sigma| \lesssim \frac{\sigma^2}{a^3} \left(\|\mathsf{f}\|_{L^2(\Omega)} + \|\mathsf{u}_d\|_{L^2(\Omega)} \right),$$

where the hidden constant is independent of $\bar{s}$, s_σ, σ, f and u_d.

Proof. We begin by considering the parameter σ sufficiently small such that $s_\sigma \in (\bar{s} - \delta, \bar{s} + \delta)$, where $\delta > 0$ is as in the statement of Corollary 6. Thus, an application of the estimate (26) in conjunction with the fact that $j_\sigma(s_\sigma) = 0$ allow us to conclude that

$$\begin{aligned}\tfrac{\vartheta}{2}|\bar{s} - s_\sigma|^2 &\le (f'(\bar{s}) - f'(s_\sigma)) \cdot (\bar{s} - s_\sigma) = f'(s_\sigma)(s_\sigma - \bar{s}) \\ &= (f'(s_\sigma) - j_\sigma(s_\sigma)) \cdot (s_\sigma - \bar{s}).\end{aligned}$$

Consequently, following Lemma 10 we obtain that

$$\begin{aligned}\frac{\vartheta}{2}|\bar{s} - s_\sigma| &\le \left| (\mathsf{u}(s_\sigma) - \mathsf{u}_d, D_s\mathsf{u}(s_\sigma) - d_\sigma \mathsf{u}(s_\sigma))_{L^2(\Omega)} \right| \\ &\lesssim \frac{\sigma^2}{a^3} \left(\|\mathsf{f}\|_{L^2(\Omega)} + \|\mathsf{u}_d\|_{L^2(\Omega)} \right).\end{aligned} \tag{34}$$

The theorem is thus proved. ∎

4.2. Space discretization. The goal of this subsection is to propose, on the basis of the bisection algorithm of section 4.1, a fully discrete scheme that approximates the solution to problem (3)–(4). To accomplish this task we will utilize the discretization techniques introduced in [9] that provides an approximation to the solution to the fractional diffusion problem (4). In order to make the exposition as clear as possible, we briefly review these aforementioned techniques below.

4.2.1. A discretization technique for fractional diffusion. Exploiting the cylindrical extension proposed and investigated in [3, 6, 13], that is in turn inspired in the breakthrough by L. Caffarelli and L. Silvestre analyzed in [4], the authors of [9] have proposed a numerical technique to approximate the solution to problem (4) that is based on an anisotropic finite element discretization of the following local and nonuniformly elliptic PDE:

$$\operatorname{div}(y^\alpha \nabla \mathscr{U}) = 0 \text{ in } \mathcal{C}, \qquad \mathscr{U} = 0 \text{ on } \partial_L \mathcal{C}, \qquad \partial_{\nu^\alpha} \mathscr{U} = d_s \mathsf{f} \text{ in } \Omega. \tag{35}$$

Here, $\mathcal{C}$ denotes the semi–infinite cylinder with base Ω defined by

$$\mathcal{C} = \Omega \times (0, \infty) \subset \mathbb{R}^{n+1}_+ = \{(x', y) : x' \in \mathbb{R}^n, y > 0\},$$

and $\partial_L \mathcal{C} = \partial\Omega \times [0, \infty)$ its lateral boundary. In addition, $d_s = 2^\alpha \Gamma(1-s)/\Gamma(s)$ and

$$\partial_{\nu^\alpha} \mathscr{U} = -\lim_{y \to 0+} y^\alpha \mathscr{U}_y.$$

Finally, $\alpha = 1 - 2s \in (-1, 1)$. Although degenerate or singular, the variable coefficient y^α satisfies a key property. Namely, it belongs to the Muckenhoupt class $A_2(\mathbb{R}^{n+1})$. This allows for an optimal piecewise polynomial interpolation theory [9].

To state the results of [3, 4, 6, 13], we define the weighted Sobolev space

$$\mathring{H}^1_L(y^\alpha,\mathcal{C}) = \left\{ w \in H^1(y^\alpha,\mathcal{C}) : w = 0 \text{ on } \partial_L \mathcal{C} \right\},$$

and the trace operator

$$\operatorname{tr}_\Omega : \mathring{H}^1_L(y^\alpha,\mathcal{C}) \to \mathbb{H}^s(\Omega), \qquad w \mapsto \operatorname{tr}_\Omega w, \tag{36}$$

where $\operatorname{tr}_\Omega w$ denotes the trace of w onto $\Omega \times \{0\}$.

The results of [3, 4, 6, 13] thus read as follows: Let $\mathscr{U} \in \mathring{H}^1_L(y^\alpha,\mathcal{C})$ and $\mathsf{u} \in \mathbb{H}^s(\Omega)$ be the solutions to (35) and (4), respectively, then

$$\mathsf{u} = \operatorname{tr}_\Omega \mathscr{U}. \tag{37}$$

A first step toward a discretization scheme is to truncate, for a given truncation parameter $\mathcal{Y} > 0$, the semi–infinite cylinder $\mathcal{C}$ to $\mathcal{C}_\mathcal{Y} := \Omega \times (0,\mathcal{Y})$ and seek solutions in this bounded domain. In fact, let $v \in \mathring{H}^1_L(y^\alpha,\mathcal{C}_\mathcal{Y})$ be the solution to

$$\int_{\mathcal{C}_\mathcal{Y}} y^\alpha \nabla v \cdot \nabla \phi = d_s \langle \mathsf{f}, \operatorname{tr}_\Omega \phi \rangle \qquad \forall \phi \in \mathring{H}^1_L(y^\alpha,\mathcal{C}_\mathcal{Y}), \tag{38}$$

where $\mathring{H}^1_L(y^\alpha,\mathcal{C}_\mathcal{Y}) = \left\{ w \in H^1(y^\alpha,\mathcal{C}_\mathcal{Y}) : w = 0 \text{ on } \partial_L\mathcal{C}_\mathcal{Y} \cup \Omega \times \{\mathcal{Y}\} \right\}$. Then the exponential decay of $\mathscr{U}$ in the extended variable y implies the following error estimate

$$\|\nabla(\mathscr{U} - v)\|_{L^2(y^\alpha,\mathcal{C})} \lesssim e^{-\sqrt{\lambda_1}\mathcal{Y}/4} \|\mathsf{f}\|_{\mathbb{H}^{-s}(\Omega)},$$

provided $\mathcal{Y} \geq 1$, and the hidden constant depends on s, but is bounded on compact subsets of $(0,1)$. We refer the reader to [9, Section 3] for details. With this truncation at hand, we thus recall the finite element discretization techniques of [9, Section 4].

To avoid technical difficulties, we assume that Ω is a convex polytopal subset of $\mathbb{R}^n$ and refer the reader to [11] for results involving curved domains. Let $\mathscr{T}_\Omega = \{K\}$ be a conforming and shape regular triangulation of Ω into cells K that are isoparametrically equivalent to either a simplex or a cube. Let $\mathcal{I}_\mathcal{Y} = \{I\}$ be a partition of the interval $[0,\mathcal{Y}]$ with mesh points

$$y_j = \left(\frac{j}{M}\right)^\gamma \mathcal{Y}, \quad j = 0,\dots,M, \quad \gamma > \frac{3}{1-\alpha} = \frac{3}{2s} > 1. \tag{39}$$

We then construct a mesh of the cylinder $\mathcal{C}_\mathcal{Y}$ by $\mathscr{T}_\mathcal{Y} = \mathscr{T}_\Omega \otimes \mathcal{I}_\mathcal{Y}$, i.e., each cell $T \in \mathscr{T}_\mathcal{Y}$ is of the form $T = K \times I$ where $K \in \mathscr{T}_\Omega$ and $I \in \mathcal{I}_\mathcal{Y}$. We note that, by construction, $\#\mathscr{T}_\mathcal{Y} = M \#\mathscr{T}_\Omega$. When $\mathscr{T}_\Omega$ is quasiuniform with $\#\mathscr{T}_\Omega \approx M^n$ we have $\#\mathscr{T}_\mathcal{Y} \approx M^{n+1}$ and, if $h_{\mathscr{T}_\Omega} = \max\{\operatorname{diam}(K) : K \in \mathscr{T}_\Omega\}$, then $M \approx h_{\mathscr{T}_\Omega}^{-1}$. Having constructed the mesh $\mathscr{T}_\mathcal{Y}$ we define the finite element space

$$\mathbb{V}(\mathscr{T}_\mathcal{Y}) := \left\{ W \in C^0(\bar{\mathcal{C}}_\mathcal{Y}) : W_{|T} \in \mathcal{P}(K) \otimes \mathbb{P}_1(I) \ \forall T \in \mathscr{T}_\mathcal{Y},\ W_{|\Gamma_D} = 0 \right\},$$

where, $\Gamma_D = \partial\Omega \times [0,\mathcal{Y}) \cup \Omega \times \{\mathcal{Y}\}$, and if K is isoparametrically equivalent to a simplex, $\mathcal{P}(K) = \mathbb{P}_1(K)$ i.e., the set of polynomials of degree at most one. If K is a cube $\mathcal{P}(K) = \mathbb{Q}_1(K)$, that is, the set of polynomials of degree at most one in each variable. We must immediately comment that, owing to (39), the meshes $\mathscr{T}_\mathcal{Y}$ are not shape regular but satisfy: if $T_1 = K_1 \times I_1$ and $T_2 = K_2 \times I_2$ are neighbors, then there is $\kappa > 0$ such that

$$h_{I_1} \leq \kappa h_{I_2}, \qquad h_I = |I|.$$

The use of anisotropic meshes in the extended direction y is imperative if one wishes to obtain a quasi-optimal approximation error since $\mathscr{U}$, the solution to (35), possesses a singularity as $y \downarrow 0$; see [9, Theorem 2.7].

We thus define a finite element approximation of the solution to the truncated problem (38): Find $V_{\mathscr{T}_\mathcal{Y}} \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$ such that

$$(40) \qquad \int_{\mathcal{C}_\mathcal{Y}} y^\alpha \nabla V_{\mathscr{T}_\mathcal{Y}} \cdot \nabla W = d_s \langle \mathsf{f}, \operatorname{tr}_\Omega W \rangle \quad \forall W \in \mathbb{V}(\mathscr{T}_\mathcal{Y}).$$

With this discrete function at hand, and on the basis of the localization results of Caffarelli and Silvestre, we define an approximation $U_{\mathscr{T}_\Omega} \in \mathbb{U}(\mathscr{T}_\Omega) = \operatorname{tr}_\Omega \mathbb{V}(\mathscr{T}_\mathcal{Y})$ of the solution u to problem (4) as follows:

$$(41) \qquad U_{\mathscr{T}_\Omega} := \operatorname{tr}_\Omega V_{\mathscr{T}_\mathcal{Y}}.$$

4.2.2. A fully discrete scheme for the fractional identification problem. Following the discussion in [9] one observes that many of the stability and error estimates in this work contain constants that depend on s. While these remain bounded in compact subsets of $(0,1)$ many of these degenerate or blow up as $s \downarrow 0$ or $s \uparrow 1$. In fact, it is not clear if the PDE (35) is well under the passage of these limits. Even if this problem made sense, the Caffarelli-Silvestre extension property (37) does not hold as we take the limits mentioned above. For this reason, we continue to work under Assumption 7. We begin by defining the discrete control to state map $S_\mathscr{T}$ as follows:

$$S_\mathscr{T} : (a,b) \to \mathbb{U}(\mathscr{T}_\Omega), \quad s \mapsto U_{\mathscr{T}_\Omega}(s),$$

where $U_{\mathscr{T}_\Omega}(s)$ is defined as in (41). We also define the function $j_{\sigma,\mathscr{T}} : (a,b) \to \mathbb{R}$ as

$$(42) \qquad j_{\sigma,\mathscr{T}}(s) = (U_{\mathscr{T}_\Omega}(s) - \mathsf{u}_d, d_\sigma U_{\mathscr{T}_\Omega}(s))_{L^2(\Omega)} + \varphi'(s),$$

where the centered difference d_σ is defined as in (28). With these elements at hand, we thus define a fully discrete approximation of the optimal identification parameter $\bar{s}$ as the solution to the following problem: Find $s_{\sigma,\mathscr{T}} \in (a,b)$ such that

$$(43) \qquad j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}}) = 0.$$

We notice that, under the assumption that the map $S_\mathscr{T}$ is continuous in (a,b), the same arguments developed in the proof of Lemma 8 yield the existence of $s_{r,\mathscr{T}}$ and $s_{l,\mathscr{T}}$ in (a,b) such that $j_{\sigma,\mathscr{T}}(s_{r,\mathscr{T}}) < 0$ and $j_{\sigma,\mathscr{T}}(s_{l,\mathscr{T}}) > 0$. This implies that, if in the bisection algorithm of section 4.1 we replace j_σ by $j_{\sigma,\mathscr{T}}$, the step **Root isolation** can be performed. Consequently, we deduce the convergence of the bisection algorithm and thus the existence of a solution $s_{\sigma,\mathscr{T}} \in (a,b)$ to problem (43).

It is then necessary to study the continuity of $S_\mathscr{T}$, but this can be easily achieved because we are in finite dimensions and the problem is linear.

PROPOSITION 14 (continuity of $S_\mathscr{T}$). *For every mesh $\mathscr{T}_\mathcal{Y}$, defined as in Section 4.2.1, the map $S_\mathscr{T}$ is continuous on (a,b).*

Proof. Let $\{s_k\}_{k\in\mathbb{N}} \subset (a,b)$ be such that $s_k \to s \in (a,b)$. Since the operator tr_Ω, defined as in (36), is continuous [9, Proposition 2.5], it suffices to show that the application $s \mapsto V_{\mathscr{T}_\mathcal{Y}}(s)$ is continuous. Consider

$$V_{\mathscr{T}_\mathcal{Y}}(s) \in \mathbb{V}(\mathscr{T}_\mathcal{Y}) : \quad \int_{\mathcal{C}_\mathcal{Y}} y^{1-2s} \nabla V_{\mathscr{T}_\mathcal{Y}}(s) \cdot \nabla W_s = d_s \langle \mathsf{f}, \operatorname{tr}_\Omega W_s \rangle \quad \forall W_s \in \mathbb{V}(\mathscr{T}_\mathcal{Y}),$$

and

$$V_{\mathscr{T}_{\mathscr{Y}}}(s_k) \in \mathbb{V}(\mathscr{T}_{\mathscr{Y}}): \quad \int_{\mathcal{C}_{\mathscr{Y}}} y^{1-2s_k} \nabla V_{\mathscr{T}_{\mathscr{Y}}}(s_k) \cdot \nabla W_k = d_{s_k} \langle \mathsf{f}, \operatorname{tr}_\Omega W_k \rangle \quad \forall W_k \in \mathbb{V}(\mathscr{T}_{\mathscr{Y}}).$$

Set $W_s = V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k)$ and $W_k = V_{\mathscr{T}_{\mathscr{Y}}}(s_k) - V_{\mathscr{T}_{\mathscr{Y}}}(s)$ and add these two identities to obtain

$$\begin{aligned} \|\nabla(V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k))\|^2_{L^2(y^{1-2s},\mathcal{C}_{\mathscr{Y}})} &= (d_s - d_{s_k})\langle \mathsf{f}, \operatorname{tr}_\Omega(V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k))\rangle \\ &+ \int_{\mathcal{C}_{\mathscr{Y}}} (y^{1-2s_k} - y^{1-2s}) \nabla V_{\mathscr{T}_{\mathscr{Y}}}(s_k) \cdot \nabla(V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k)) = \mathrm{I} + \mathrm{II}. \end{aligned}$$

We now proceed to estimate each one of these terms.

For the first term we have

$$|\mathrm{I}| \le |d_s - d_{s_k}| \|\mathsf{f}\|_{L^2(\Omega)} \| \operatorname{tr}_\Omega(V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k)\|_{L^2(\Omega)} \to 0$$

as $k \to \infty$. This is the case because $\| \operatorname{tr}_\Omega(V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k)\|_{L^2(\Omega)}$ is uniformly bounded [9, Proposition 2.5] and, by Assumption 7, we have that $d_{s_k} \to d_s$.

We estimate the second term as follows

$$|\mathrm{II}| \le |\Omega| \|\nabla V_{\mathscr{T}_{\mathscr{Y}}}(s_k)\|_{L^\infty(\mathcal{C}_{\mathscr{Y}})} \|\nabla(V_{\mathscr{T}_{\mathscr{Y}}}(s) - V_{\mathscr{T}_{\mathscr{Y}}}(s_k))\|_{L^\infty(\mathcal{C}_{\mathscr{Y}})} \int_0^{\mathscr{Y}} |y^{1-2s} - y^{1-2s_k}|.$$

Using that we are in finite dimensions, the question reduces to the convergence

$$\int_0^{\mathscr{Y}} |y^{1-2s} - y^{1-2s_k}| \to 0,$$

which follows from the a.e. convergence of y^{1-2s_k} to y^{1-2s}, the fact that, for $0 < y < 1$, we have $0 < y^{1-2s_k} \le y^{1-2a} \in L^1(0,1)$ and an application of the dominated convergence theorem.

This concludes the proof. ◻

We now proceed to derive an a priori error bound for the error between the exact identification parameter $\bar{s}$ and its approximation $s_{\sigma,\mathscr{T}}$ given as the solution (43). We begin by noticing that, following the proof of Lemma 10, using [10, Proposition 28] and Assumption 7 we have

$$(44) \qquad |j_\sigma(s) - j_{\sigma,\mathscr{T}}(s)| \lesssim \frac{1}{\sigma} |\log(\#\mathscr{T}_{\mathscr{Y}})|^{2b} (\#\mathscr{T}_{\mathscr{Y}})^{-(1+a)/(n+1)},$$

where the hidden constant depends on a and b but is uniform in (a, b). Clearly, for fixed σ, this implies the uniform convergence of $j_{\sigma,\mathscr{T}}$ to j_σ as we refine the mesh. By repeating the arguments of Lemma 11 we conclude the convergence, up to subsequences, of $\{s_{\sigma,\mathscr{T}}\}_{\mathscr{T}}$ to s_σ, a root of j_σ. Arguing as in Remark 12, we see that we cannot expect convergence of the entire family.

Finally, we denote one of these convergent subsequences by $\{s_{\sigma,\mathscr{T}}\}_{\mathscr{T}}$ and provide an error estimate.

Theorem 15 (Error estimate: discretization in s and space). *Let $\bar{s}$ be optimal for the identification problem (3)–(4) and $s_{\sigma,\mathscr{T}}$ its approximation defined as the solution to (43). If σ is sufficiently small, $\#\mathscr{T}_{\mathscr{Y}}$ is sufficiently large and, $\mathsf{f} \in \mathbb{H}^{1-a}(\Omega)$, then*

$$(45) \qquad |\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim \sigma^{-1} |\log(\#\mathscr{T}_{\mathscr{Y}})|^{2b} (\#\mathscr{T}_{\mathscr{Y}})^{-(1+a)/(n+1)} \|\mathsf{f}\|_{\mathbb{H}^{1-a}(\Omega)} + \sigma^2,$$

where the hidden constant is independent of $\bar{s}$, $s_{\sigma,\mathscr{T}}$, f and the mesh $\mathscr{T}_{\mathscr{Y}}$.

Proof. We begin by remarking that, by setting σ sufficiently small and $\#\mathscr{T}_\mathcal{Y}$ sufficiently large, respectively, we can assert that $s_{\sigma,\mathscr{T}} \in (\bar{s}-\delta, \bar{s}+\delta)$ with δ being the parameter of Corollary 6. By invoking the estimate (26) and in view of the fact that $f'(\bar{s}) = 0 = j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$, we deduce the following estimate:

$$\frac{\vartheta}{2}|\bar{s}-s_{\sigma,\mathscr{T}}|^2 \leq (f'(\bar{s}) - f'(s_{\sigma,\mathscr{T}}))\cdot(\bar{s}-s_{\sigma,\mathscr{T}}) = (j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}}) - f'(s_{\sigma,\mathscr{T}}))\cdot(\bar{s}-s_{\sigma,\mathscr{T}}).$$

We proceed to bound the right hand side of the previous expression. To accomplish this task, we invoke the definition (42) of $j_{\sigma,\mathscr{T}}$ and repeating the arguments of Lemma 10 we obtain that

$$(46)\qquad \begin{aligned} |j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}}) - f'(s_{\sigma,\mathscr{T}})| \leq & \left|\left(U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - \mathsf{u}_d, d_\sigma U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - D_s\mathsf{u}(s_{\sigma,\mathscr{T}})\right)_{L^2(\Omega)}\right| \\ & + \left|\left(U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - \mathsf{u}(s_{\sigma,\mathscr{T}}), D_s\mathsf{u}(s_{\sigma,\mathscr{T}})\right)_{L^2(\Omega)}\right| = \mathrm{I} + \mathrm{II}. \end{aligned}$$

We thus examine each term separately. We start with II: its control relies on the a priori error estimates of [9, 10]. In fact, combining the results of [10, Proposition 28] with the estimate (16) for $m = 1$, we arrive at

$$\begin{aligned} |\mathrm{II}| &\leq \|D_s\mathsf{u}(s_{\sigma,\mathscr{T}})\|_{L^2(\Omega)}\|U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - \mathsf{u}(s_{\sigma,\mathscr{T}})\|_{L^2(\Omega)} \\ &\lesssim s_{\sigma,\mathscr{T}}^{-1}|\log(\#\mathscr{T}_\mathcal{Y})|^{2s_{\sigma,\mathscr{T}}}(\#\mathscr{T}_\mathcal{Y})^{-(1+s_{\sigma,\mathscr{T}})/(n+1)}\|\mathsf{f}\|_{\mathbb{H}^{1-s_{\sigma,\mathscr{T}}}(\Omega)} \\ &\lesssim |\log(\#\mathscr{T}_\mathcal{Y})|^{2b}(\#\mathscr{T}_\mathcal{Y})^{-(1+a)/(n+1)}\|\mathsf{f}\|_{\mathbb{H}^{1-a}(\Omega)} \end{aligned}$$

where the hidden constant depends on a and b but is independent of $\bar{s}$, $s_{\sigma,\mathscr{T}}$, f and $\mathscr{T}_\mathcal{Y}$. Notice that here we used Assumption 7 to, for instance, control the term $s_{\sigma,\mathscr{T}}^{-1}$.

We now proceed to control the term I in (46). A basic application of the Cauchy–Schwarz inequality yields

$$|\mathrm{I}| \leq \|U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - \mathsf{u}_d\|_{L^2(\Omega)}\|d_\sigma U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - D_s\mathsf{u}(s_{\sigma,\mathscr{T}})\|_{L^2(\Omega)}.$$

We thus apply the estimate (14) and the triangle inequality to obtain that

$$|\mathrm{I}| \lesssim \|d_\sigma\left(U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - \mathsf{u}(s_{\sigma,\mathscr{T}})\right)\|_{L^2(\Omega)} + \|d_\sigma\mathsf{u}(s_{\sigma,\mathscr{T}}) - D_s\mathsf{u}(s_{\sigma,\mathscr{T}})\|_{L^2(\Omega)}.$$

We estimate the first term on the right hand side of the previous expression: the definition (28) of d_σ and [10, Proposition 28] imply that

$$\begin{aligned} \|d_\sigma\left(U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}) - \mathsf{u}(s_{\sigma,\mathscr{T}})\right)\|_{L^2(\Omega)} \leq \frac{1}{2\sigma}\Big(\|U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}+\sigma) - \mathsf{u}(s_{\sigma,\mathscr{T}}+\sigma)\|_{L^2(\Omega)} \\ + \|U_{\mathscr{T}_\Omega}(s_{\sigma,\mathscr{T}}-\sigma) - \mathsf{u}(s_{\sigma,\mathscr{T}}-\sigma)\|_{L^2(\Omega)}\Big) \lesssim \frac{1}{\sigma}|\log(\#\mathscr{T}_\mathcal{Y})|^{2b}(\#\mathscr{T}_\mathcal{Y})^{-\frac{1+a}{n+1}}\|\mathsf{f}\|_{\mathbb{H}^{1-a}(\Omega)}; \end{aligned}$$

we notice that σ is small enough such that $s_{\sigma,\mathscr{T}} \pm \sigma \in (a,b)$. On the other hand, an estimate similar to (29) yields that

$$\|D_s\mathsf{u}(s_{\sigma,\mathscr{T}}) - d_\sigma\mathsf{u}(s_{\sigma,\mathscr{T}})\|_{L^2(\Omega)} \lesssim \sigma^2 a^{-3}.$$

Collecting the previous estimates we arrive at the following bound for the term I:

$$(47)\qquad |\mathrm{I}| \lesssim \sigma^{-1}|\log(\#\mathscr{T}_\mathcal{Y})|^{2b}(\#\mathscr{T}_\mathcal{Y})^{-(1+a)/(n+1)}\|\mathsf{f}\|_{\mathbb{H}^{1-a}(\Omega)} + \sigma^2 a^{-3}.$$

On the basis of (46), this bound, and the estimate for the term II yield

$$|\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim \sigma^{-1} |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2b} (\#\mathscr{T}_{\mathcal{Y}})^{-(1+a)/(n+1)} \|\mathsf{f}\|_{\mathbb{H}^{1-a}(\Omega)} + \sigma^2,$$

where the hidden constant depends on a and b, but is independent of σ and $\#\mathscr{T}_{\mathcal{Y}}$. This concludes the proof. ∎

A natural choice of σ comes from equilibrating the terms on the right–hand side of (45): $\sigma \approx |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2b/3} (\#\mathscr{T}_{\mathcal{Y}})^{-(1+a)/3(n+1)}$. This implies the following error estimate.

COROLLARY 16 (error estimate: discretization in s and space). *Let $\bar{s}$ be optimal for the identification problem (3)–(4) and $s_{\sigma,\mathscr{T}}$ be its approximation defined as the solution to (43). If $\#\mathscr{T}_{\mathcal{Y}}$ is sufficiently large, the parameter σ is chosen as*

$$\sigma \approx |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2b/3} (\#\mathscr{T}_{\mathcal{Y}})^{-(1+a)/3(n+1)},$$

and $\mathsf{f} \in \mathbb{H}^{1-a}(\Omega)$ then

$$|\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim |\log(\#\mathscr{T}_{\mathcal{Y}})|^{4b/3} (\#\mathscr{T}_{\mathcal{Y}})^{-\frac{2(1+a)}{3(n+1)}}, \tag{48}$$

where the hidden constant depends on a and b but is independent of $\bar{s}$, $s_{\sigma,\mathscr{T}}$, and the mesh $\mathscr{T}_{\mathcal{Y}}$.

5. Numerical examples. In this section, we study the performance of the proposed bisection algorithm of section 4 when applied to the fully discrete parameter identification problem of section 4.2.2 with the help of four numerical examples.

The implementation has been carried out within the MATLAB software library *i*FEM [7]. The stiffness matrices of the discrete system (40) are assembled exactly and the forcing terms are computed by a quadrature rule which is exact for polynomials up to degree 4. Additionally, the first term in (42) is computed by a quadrature formula which is exact for polynomials of degree 7. All the linear systems are solved exactly using MATLAB's built-in direct solver.

In all examples, $n = 2$, $\Omega = (0,1)^2$, $\mathsf{TOL} = 2.2204e\text{-}16$, and the initial value of s_l, s_r is 0.3, and 0.9, respectively. The truncation parameter for the cylinder $\mathcal{C}_{\mathcal{Y}}$ is $\mathcal{Y} = 1 + \frac{1}{3}(\#\mathscr{T}_{\Omega})$ which allows balancing the approximation and truncation errors for our state equation, see [9, Remark 5.5]. Moreover,

$$\sigma = \frac{1}{2.5} (\#\mathscr{T}_{\mathcal{Y}})^{-\frac{(1+\epsilon)}{9}},$$

with $\epsilon = 10^{-10}$.

Under the above setting, the eigenvalues and eigenvectors of $-\Delta$ are:

$$\lambda_{k,l} = \pi^2(k^2 + l^2), \quad \varphi_{k,l}(x_1, x_2) = \sin(k\pi x_1)\sin(l\pi x_2), \quad k, l \in \mathbb{N}.$$

Consequently, by letting $\mathsf{f} = \lambda_{2,2}^s \varphi_{2,2}$ for any $s \in (0,1)$ we obtain $\bar{\mathsf{u}} = \varphi_{2,2}$.

In what follows we will consider four examples. In the first one we set $\bar{s} = 1/2$, f and $\bar{\mathsf{u}}$ as above and we set $\mathsf{u}_d = \bar{\mathsf{u}}$. The second one differs from the first one in that we set $\bar{s} = (3 - \sqrt{5})/2$. In our third example, the exact solution is not known. Finally, in our last example we explore the robustness of our algorithm with respect to perturbations in the data. We accomplish this by considering the same setting as in the first example but we add a random perturbation $r \in (-e, e)$ to the right hand side f. We then explore the behavior of the optimal parameter $\bar{s}$ as the size of the perturbation e varies.

5.1. Example 1. We recall the definition of the cost function $J(\mathsf{u}, s)$ from (1) and set $\varphi(s) = \frac{1}{s(1-s)}$. The latter is strictly convex over the interval $(0, 1)$ and fulfills the conditions in (2). The optimal solution $\bar{s}$ to (3)–(4) is given by $\bar{s} = 1/2$.

Table 1 illustrates the performance of our optimization solver. The first column indicates the degrees of freedom $\#\mathscr{T}_{\mathcal{Y}}$, the second column shows the value of $s_{\sigma,\mathscr{T}}$ obtained by solving (43), and the third column shows the corresponding value $j_{\sigma,\mathscr{T}}$ at $s_{\sigma,\mathscr{T}}$. The final column shows the total number of optimization iterations N taken, for the bisection algorithm to converge. We notice that the observed values of $s_{\sigma,\mathscr{T}}$ matches almost perfectly with $\bar{s}$. In addition, the pattern in N, as we refine the mesh, indicates a mesh-independent behavior.

$\#\mathscr{T}_{\mathcal{Y}}$	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
3146	4.96572e-01	-8.89011e-14	53
10496	4.98371e-01	-8.38218e-14	53
25137	4.99069e-01	3.49235e-14	53
49348	4.99402e-01	1.52327e-12	53
85529	4.99585e-01	6.28221e-12	53

TABLE 1

The first column indicates the degrees of freedom, the second one corresponds to the solution $s_{\sigma,\mathscr{T}}$ of our discrete optimality system (43) and the third column illustrates the corresponding value of $j_{\sigma,\mathscr{T}}$ at $s_{\sigma,\mathscr{T}}$. The final column shows, N, the number of iterations taken by the bisection algorithm to converge. The values of N are moderate. Additionally, we observe that $s_{\sigma,\mathscr{T}}$ matches with the exact solution $\bar{s} = 1/2$ and the pattern in N shows a mesh independent behavior upon mesh refinement.

Figure 1 (left panel) shows the computational rate of convergence. We observe that

$$|\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim (\#\mathscr{T}_{\mathcal{Y}})^{-0.6}$$

which is significantly better than the predicated rate of $(\#\mathscr{T}_{\mathcal{Y}})^{-0.22}$ by the Corollary 16. Indeed this suggests that our theoretical rates are pessimistic and in practice, our algorithm works much better.

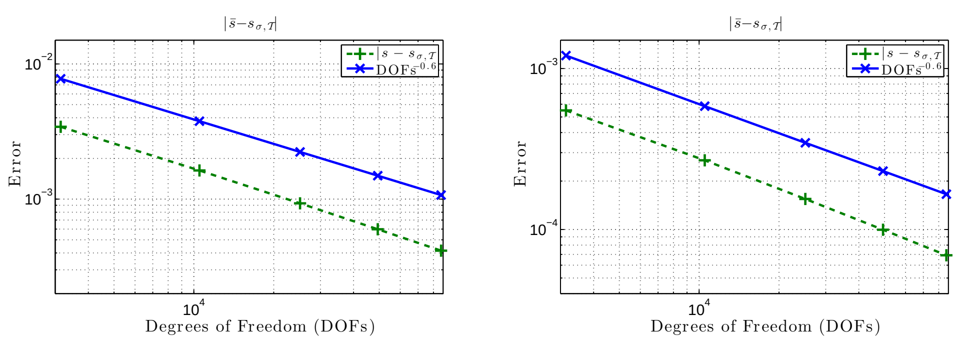

FIG. 1. *The left panel (dotted curve) shows the convergence rate for Example 1 and the right one for Example 2. The solid line is the reference line. We notice that the computational rates of convergence, in both examples, are much higher than the theoretically predicted rates in Corollary 16.*

5.2. Example 2. We set $\varphi(s) = s^{-1}e^{\frac{1}{(1-s)}}$ which is again strictly convex over the interval $(0, 1)$ and fulfills the conditions in (2). The optimal solution $\bar{s}$ to (3)–(4) is given by $\bar{s} = (3 - \sqrt{5})/2$.

Table 2 illustrates the performance of our optimization solver. As we noted in section 5.1, the numerically computed solution $s_{\sigma,\mathscr{T}}$ matches almost perfectly with $\bar{s}$ and the pattern of N, with mesh refinement, again indicates a mesh independent behavior.

$\#\mathscr{T}_{\mathcal{Y}}$	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
3146	3.81417e-01	9.99201e-16	46
10496	3.81697e-01	-2.52812e-13	53
25137	3.81811e-01	1.36418e-12	53
49348	3.81866e-01	2.66251e-12	53
85529	3.81897e-01	3.53083e-12	53

TABLE 2

The first column indicates the degrees of freedom, the second one corresponds to the solution $s_{\sigma,\mathscr{T}}$ of our discrete optimality system (43) and the third column illustrates the corresponding value of $j_{\sigma,\mathscr{T}}$ at $s_{\sigma,\mathscr{T}}$. The final column shows, N, the number of iterations taken by the bisection algorithm to converge. The values of N are moderate. Additionally, we observe that $s_{\sigma,\mathscr{T}}$ matches with the exact solution $\bar{s} = (3-\sqrt{5})/2$ and the pattern in N shows a mesh independent behavior upon mesh refinement.

Figure 1 (right panel) shows the computational rate of convergence. We again see that

$$|\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim (\#\mathscr{T}_{\mathcal{Y}})^{-0.6}$$

Thus the observed rate is far superior than the theoretically predicted rate in Corollary 16.

5.3. Example 3. In our third example, we take $\varphi(s) = s^{-1}e^{\frac{1}{(1-s)}}$, $\mathsf{f} = 10$, and $\mathsf{u}_d = \max\left\{0.5 - \sqrt{|x_1 - 0.5|^2 + |x_2 - 0.5|^2}, 0\right\}$. We notice that f is large, thus the requirements of Theorem 13 are not necessarily fulfilled. In addition, for $\mu \leq 1/2$, $\mathsf{f} \notin \mathbb{H}^{1-\mu}(\Omega)$ thus the requirements of Corollary 16 are not fulfilled. Nevertheless, as we illustrate in Table 3, we can still solve the problem. We again notice a mesh independent behavior in the number of iterations (N) taken by the bisection algorithm to converge.

$\#\mathscr{T}_{\mathcal{Y}}$	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
3146	4.44005e-01	4.22951e-12	53
10496	4.47239e-01	2.97451e-11	53
25137	4.48182e-01	-3.20792e-11	53
49348	4.48544e-01	4.83542e-11	53
85529	4.48690e-01	2.68390e-10	53

TABLE 3

The first column indicates the degrees of freedom, the second one corresponds to the solution $s_{\sigma,\mathscr{T}}$ of our discrete optimality system (43) and the third column illustrates the corresponding value $j_{\sigma,\mathscr{T}}$ at $s_{\sigma,\mathscr{T}}$. The final column shows, N, the number of iterations taken by the bisection algorithm to converge. The values of N are moderate and show a mesh independent character.

5.4. Example 4. In our final example we consider a similar setup to subsection 5.1. We modify the right hand side $\mathsf{f} = \lambda_{2,2}^{\bar{s}} \sin(2\pi x_1)\sin(2\pi x_2)$, with $\bar{s} = 1/2$, by adding a uniformly distributed random parameter $r \in (-e, e)$. We fix the spatial mesh to $\#\mathscr{T}_{\mathcal{Y}} = 85,529$.

At first we set $e = 200$, as a result r is more than 200 times the actual signal f, see the first row on Table 4. Despite such a large noise, the recovery of $\bar{s}$ is reasonable. Letting $e \downarrow 0$, we can recover $\bar{s}$ almost perfectly.

e	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
200	6.33937e-01	7.28484e-12	53
20	5.06469e-01	-5.17408e-12	53
2	4.99341e-01	-7.37949e-12	53
0.5	4.99581e-01	-5.68941e-12	53
0.25	4.99586e-01	3.64379e-12	53
0.125	4.99584e-01	3.33318e-13	53

TABLE 4

Robustness of our algorithm with respect to noisy data. The number of spatial degrees of freedom is fixed to $\#\mathscr{T}_{\mathcal{Y}} = 85,529$. The first column indicates the range of the uniformly distributed parameter r which is added to the right hand side f, the second one corresponds to the solution $s_{\sigma,\mathscr{T}}$ of our discrete optimality system (43) and the third column illustrates the corresponding value $j_{\sigma,\mathscr{T}}$ at $s_{\sigma,\mathscr{T}}$. The final column shows N, the number of iterations taken by the bisection algorithm to converge. Notice that even with a noise which is 200 times more than the actual signal f the recovery of $\bar{s}$ is reasonable (first row). If the noise is of the same order as f we can recover $\bar{s}$ perfectly. The values of N are moderate and show a mesh independent character.

REFERENCES

[1] R. Adams, *Sobolev spaces*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[2] H. Antil and E. Otárola, *A FEM for an optimal control problem of fractional powers of elliptic operators*, SIAM J. Control Optim., 53 (2015), pp. 3432–3456, http://dx.doi.org/10.1137/140975061.

[3] X. Cabré and J. Tan, *Positive solutions of nonlinear problems involving the square root of the Laplacian*, Adv. Math., 224 (2010), pp. 2052–2093, http://dx.doi.org/10.1016/j.aim.2010.01.025.

[4] L. Caffarelli and L. Silvestre, *An extension problem related to the fractional Laplacian*, Comm. Part. Diff. Eqs., 32 (2007), pp. 1245–1260, http://dx.doi.org/10.1080/03605300600987306.

[5] L. Caffarelli and P. Stinga, *Fractional elliptic equations, Caccioppoli estimates and regularity*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 33 (2016), pp. 767–807, http://dx.doi.org/10.1016/j.anihpc.2015.01.004.

[6] A. Capella, J. Dávila, L. Dupaigne, and Y. Sire, *Regularity of radial extremal solutions for some non-local semilinear equations*, Comm. Partial Differential Equations, 36 (2011), pp. 1353–1384, http://dx.doi.org/10.1080/03605302.2011.562954.

[7] L. Chen, *iFEM: an integrated finite element methods package in matlab*, tech. report, Technical Report, University of California at Irvine, 2009.

[8] K. Deckelnick and M. Hinze, *Convergence and error analysis of a numerical method for the identification of matrix parameters in elliptic PDEs*, Inverse Problems, 28 (2012), pp. 115015, 15, http://dx.doi.org/10.1088/0266-5611/28/11/115015.

[9] R. Nochetto, E. Otárola, and A. Salgado, *A PDE approach to fractional diffusion in general domains: A priori error analysis*, Found. Comput. Math., 15 (2015), pp. 733–791, http://dx.doi.org/10.1007/s10208-014-9208-x.

[10] R. Nochetto, E. Otárola, and A. Salgado, *A PDE approach to space-time fractional parabolic problems*, SIAM J. Numer. Anal., 54 (2016), pp. 848–873, http://dx.doi.org/10.1137/14096308X.

[11] E. Otárola, *A piecewise linear FEM for an optimal control problem of fractional operators: error analysis on curved domains*, ESAIM Math. Model. Numer. Anal., (2016), http://dx.doi.org/10.1051/m2an/2016065. (to appear).

[12] J. Sprekels and E. Valdinoci, *A new type of identification problems: optimizing the fractional order in a nonlocal evolution equation.* arXiv:1601.00568, 2016.

[13] P. Stinga and J. Torrea, *Extension problem and Harnack's inequality for some fractional operators*, Comm. Part. Diff. Eqs., 35 (2010), pp. 2092–2122, http://dx.doi.org/10.1080/03605301003735680.

[14] F. Tröltzsch, *Optimal control of partial differential equations*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010, http://dx.doi.org/10.1090/gsm/112. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.