Results in Applied Mathematics 9 (2021) 100133

Contents lists available at ScienceDirect

Results in Applied Mathematics

journal homepage: www.elsevier.com/locate/results-in-applied-mathematics

A note on multigrid preconditioning for fractional PDE-constrained optimization problems

Harbir Antil [a,*], Andrei Drăgănescu [b], Kiefer Green [a]

[a] Department of Mathematical Sciences and the Center for Mathematics and Artificial Intelligence (CMAI), George Mason University, Fairfax, VA 22030, USA
[b] Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA

ARTICLE INFO

Article history:
Received 9 June 2020
Received in revised form 17 November 2020
Accepted 17 November 2020
Available online 4 December 2020

Keywords:
Optimal control
Fractional diffusion
Multigrid
Preconditioner

ABSTRACT

In this note we present a multigrid preconditioning method for solving quadratic optimization problems constrained by a fractional diffusion equation. Multigrid methods within the all-at-once approach to solve the first order optimality Karush–Kuhn–Tucker (KKT) systems are widely popular, but their development have relied on the underlying systems being sparse. On the other hand, for most discretizations, the matrix representation of fractional operators is expected to be dense. We develop a preconditioning strategy for our problem based on a reduced approach, namely we eliminate the state constraint using the control-to-state map. Our multigrid preconditioning approach shows a dramatic reduction in the number of CG iterations. We assess the quality of preconditioner in terms of the spectral distance. Finally, we provide a partial theoretical analysis for this preconditioner, and we formulate a conjecture which is clearly supported by our numerical experiments.

1. Introduction

Let $\Omega \subset \mathbb{R}^N$ be an open bounded Lipschitz polygonal domain with boundary $\partial\Omega$. The goal of this paper is to develop an efficient multigrid based solver for the following optimal control problem: Given datum $u_d \in L^2(\Omega)$ and a regularization parameter $\beta > 0$, solve

$$\min_{z\in L^2(\Omega)} \frac{1}{2}\|u - u_d\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|z\|^2_{L^2(\Omega)}, \tag{1a}$$

subject to the constraints posed by the fractional partial differential equation (PDE)

$$\begin{cases} (-\Delta)^s u &= z \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{1b}$$

Here, u and z denote the state and control variables, respectively. Moreover, $(-\Delta)^s$, with $0 < s < 1$, denotes the s powers of the $L^2(\Omega)$ realization of the Laplace's operator $-\Delta$, with the Dirichlet boundary condition $u = 0$ on $\partial\Omega$. This is the so-called spectral fractional Laplacian. We refer to [1] for the case of non-zero boundary conditions.

* Corresponding author.
E-mail address: hantil@gmu.edu (H. Antil).

https://doi.org/10.1016/j.rinam.2020.100133

The rising interest of the community in fractional operators has been motivated by their ever-growing applicability. In [2] (see also [3] for an efficient solver), a fractional Helmholtz equation is derived using first principle arguments in-conjunction with a constitutive relationship. It also shows a direct qualitative match between numerical simulations and experimental data. In the classical setting, it is well-known that constrained optimization problems with the Helmholtz equation as constraint arise naturally in various applications. Examples include direct-field acoustic testing [4] and remote sensing applications such as source inversion in seismology [5]. A natural first step to create efficient solvers for these optimization problems is to begin with optimization problems constrained by Poisson type equations. Following this line of argument, we are hereby creating an efficient solver for (1). Fractional operators have also received a significant attention due to their applicability in imaging science [6,7].

Problem (1) was introduced in [8], and has attracted significant attention ever since. While it is a natural extension of the standard elliptic control problem corresponding to the case $s = 1$, it leads to a number of challenging questions, beginning with the definition and the numerical representation of the fractional operator. In [8], problem (1) was formulated and analyzed using the extension approach [9,10]. An alternative numerical analysis for (1) was provided in [11]. The latter used a numerical scheme to approximate (3), based on Kato's formula [12], originally introduced in [13]. See also [14] for a tensor based method to solve (1). For completeness, we also refer to related optimal control problems corresponding to integral fractional Laplacian where the control is distributed [15,16], or it is in the coefficient [17,18], or it is in the exterior [19,20]. We also refer to [21] for an efficient multigrid solver for fractional PDEs with integral fractional Laplacian.

The majority of efficient solution methods for solving PDE-constrained optimization problems focus on the first order optimality conditions, namely the Karush–Kuhn–Tucker (KKT) system [22]. The KKT system couples the PDE (1b) (the state equation) and the adjoint equation, the latter being a linear PDE with a similar character to the state equation. Hence, for the case of classical PDE constraints with finite element discretizations, the KKT system – albeit indefinite – will have a sparse structure, and solvers and preconditioners used for the state equation can play an important role for the KKT system as well. However, for most discretizations the matrix representation of discrete fractional operators is expected to be dense, therefore the all-at-once approach of solving the KKT system loses its main attractiveness, namely sparsity.

In this work we use a reduced approach, namely we eliminate the state constraint from the optimization problem (1) using the control-to-state map. Using the discretization from [11], we introduce a multigrid based preconditioner to solve (1). Multigrid methods, traditionally known as some of the most efficient solvers of discretizations of PDEs, has been employed in recent times with great success in PDE-constrained optimization [22] as well. Our approach is motivated by [23], and we develop a multigrid preconditioner for the reduced system of (1).

2. The fractional operator and the optimality conditions

2.1. Continuous optimality conditions

For $s \geq 0$, we define the fractional order Sobolev space

$$\mathbb{H}^s(\Omega) := \left\{ u = \sum_{k=1}^{\infty} u_k \varphi_k \in L^2(\Omega) \,:\, \|u\|^2_{\mathbb{H}^s(\Omega)} := \sum_{k=1}^{\infty} \lambda_k^s u_k^2 \ < \infty \right\}, \tag{2}$$

where λ_k are the eigenvalues of $-\Delta$ and φ_k the corresponding eigenfunctions with zero Dirichlet boundary conditions and $\|\varphi_k\|_{L^2(\Omega)} = 1$, and

$$u_k = (u, \varphi_k)_{L^2(\Omega)} = \int_\Omega u \varphi_k \, .$$

By now, it is well-known that the definition of $\mathbb{H}^s(\Omega)$ in (2) is equivalent to $H_0^s(\Omega)$ for $s > 1/2$, and $\mathbb{H}^s(\Omega) = H^s(\Omega) = H_0^s(\Omega)$ when $s < 1/2$, while $\mathbb{H}^{\frac{1}{2}}(\Omega) = H_{00}^{\frac{1}{2}}(\Omega)$, i.e., the Lions–Magenes space [24]. Recall that, for $0 < s < 1$, $\mathbb{H}^s(\Omega)$ is the interpolation space between $L^2(\Omega)$ and $H_0^1(\Omega)$ [24], a fact that is relevant for the analysis below. Let $\mathbb{H}^{-s}(\Omega)$ be the dual space of $\mathbb{H}^s(\Omega)$.

For $s \geq 0$, the spectral fractional Laplacian is defined on the space $C_0^\infty(\Omega)$ by

$$(-\Delta)^s u := \sum_{k=1}^{\infty} \lambda_k^s u_k \varphi_k \quad \text{with} \quad u_k = \int_\Omega u \varphi_k \, .$$

Notice that, for any $w = \sum_{k=1}^{\infty} w_k \varphi_k \in \mathbb{H}^s(\Omega)$, we have that

$$\left| \int_\Omega (-\Delta)^s u w \right| = \left| \sum_{k=1}^{\infty} \lambda_k^s u_k w_k \right| = \left| \sum_{k=1}^{\infty} \lambda_k^{\frac{s}{2}} u_k \lambda_k^{\frac{s}{2}} w_k \right| \leq \|u\|_{\mathbb{H}^s(\Omega)} \|w\|_{\mathbb{H}^s(\Omega)},$$

and thus $(-\Delta)^s$ extends as an operator mapping from $\mathbb{H}^s(\Omega)$ to $\mathbb{H}^{-s}(\Omega)$ due to density. In addition, we have that

$$\|u\|_{\mathbb{H}^s(\Omega)} = \|(-\Delta)^{\frac{s}{2}} u\|_{L^2(\Omega)}.$$

Cf. [9], for every $z \in \mathbb{H}^{-s}(\Omega)$ there exists a unique $u \in \mathbb{H}^s(\Omega)$ that solves (1b). Using Kato's formula (see [13,25] for a derivation), the solution u can be explicitly written as

$$u = (-\Delta)^{-s} z = \frac{\sin s\pi}{\pi} \int_{-\infty}^{\infty} e^{(1-s)y} (e^y - \Delta)^{-1} z \, dy. \tag{3}$$

Notice that $(-\Delta)^{-s} : \mathbb{H}^{-s}(\Omega) \to \mathbb{H}^s(\Omega)$ is bounded and linear. By restricting $(-\Delta)^{-s}$ to $L^2(\Omega)$, and using the compact embedding $\mathbb{H}^s(\Omega) \hookrightarrow L^2(\Omega)$, we can treat the solution map $\mathcal{K}^s := (-\Delta)^{-s}$ as a bounded linear operator on $L^2(\Omega)$. Hence, the adjoint operator $(\mathcal{K}^s)^* : L^2(\Omega) \to L^2(\Omega)$ is well-defined, and is equal to $\mathcal{K}^s$. Using $\mathcal{K}^s$, the reduced form of problem (1) is given by

$$\min_{z \in L^2(\Omega)} \frac{1}{2} \|\mathcal{K}^s z - u_d\|^2_{L^2(\Omega)} + \frac{\beta}{2} \|z\|^2_{L^2(\Omega)}. \tag{4}$$

Problem (4) has a unique solution that satisfies the following first-order necessary and sufficient optimality conditions

$$\mathcal{H}^s z \overset{\text{def}}{=} \left((\mathcal{K}^s)^* \mathcal{K}^s + \beta I\right) z = (\mathcal{K}^s)^* u_d. \tag{5}$$

Notice that (5) follows immediately after differentiating twice the functional in (4). The operator $\mathcal{H}^s$ in (5) is the continuous reduced Hessian operator. Next we shall discretize (5).

2.2. Discrete optimality conditions

We consider a quasi-uniform discretization $\mathcal{T}_h$ of Ω and the spaces of continuous piecewise linear functions $\mathcal{V}_h$ and $\mathcal{V}_h^0 = \mathcal{V}_h \cap H_0^1(\Omega)$. The control z is discretized using $\mathcal{V}_h$, while the state u is discretized using $\mathcal{V}_h^0$. According to [13], the discrete solution operator $\mathcal{K}_h^s : \mathcal{V}_h \to \mathcal{V}_h^0$ is defined as

$$\mathcal{K}_h^s := \frac{\sin s\pi}{\pi} m \sum_{\ell=-N^-}^{N^+} e^{(1-s)y_\ell} (e^{y_\ell} - \Delta_h)^{-1},$$

where the quadrature nodes are uniformly distributed as $y_\ell = m\ell$. This quadrature rule has been shown to be exponentially convergent (see [13]) to the continuous integral in (3). The underlying constants N^- and N^+ are chosen to balance the quadrature error and spatial discretization error. In our case they are: $m \sim (\ln \frac{1}{h})^{-1}$, $N^+ = \lceil \frac{\pi^2}{4sm^2} \rceil$, and $N^- = \lceil \frac{\pi^2}{4(1-s)m^2} \rceil$. Finally, we shall denote by $\pi_h : L^2(\Omega) \to \mathcal{V}_h$, the L^2-orthogonal projection.

Using the above discretization, the discrete form of (5) is given by

$$\mathcal{H}_h^s z_h \overset{\text{def}}{=} ((\mathcal{K}_h^s)^* \mathcal{K}_h^s + \beta I) z_h = (\mathcal{K}_h^s)^* u_{d,h}, \tag{6}$$

where $u_{d,h} = \pi_h u_d$. This work is concerned with a multigrid preconditioning approach to efficiently solve (6).

3. Two-grid and multigrid preconditioner

3.1. Preconditioner description

Following [23], assuming $\mathcal{T}_h$ is a refinement of $\mathcal{T}_{2h}$, we define the two-grid preconditioner:

$$\mathcal{G}_h^s = \beta(I - \mathcal{E}^h \pi_{2h}) + \mathcal{E}^h \mathcal{H}_{2h}^s \pi_{2h}, \tag{7}$$

where $\mathcal{E}^h : \mathcal{V}_{2h} \to \mathcal{V}_h$ is the natural embedding operator.

The extension of the preconditioners from two-grid to multigrid is a streamlined process that is presented in full detail in [23,26]. It is sufficient to say that the multigrid version has a W-cycle structure, and that the coarsest grid has to be sufficiently fine. Hence, it may be that the coarsest level used in the multigrid version of $\mathcal{G}_h^s$ is not the coarsest that is in principle available by the existing geometric framework. The number of levels that can be used is problem dependent, and depends also of the quality of the two-grid preconditioner, as described below.

3.2. Analysis and conjecture

We assess the quality of the preconditioner $\mathcal{G}_h^s$ by estimating the spectral distance (see [23]) $d(\mathcal{H}_h^s, \mathcal{G}_h^s)$, where for two symmetric positive definite operators $A, B \in \mathfrak{L}(L^2(\mathcal{V}_h))$

$$d(A, B) = \max_{u \in \mathcal{V}_h} |\ln(Au, u) - \ln(Bu, u)| = \max\{|\ln \lambda| \ : \ \lambda \in \sigma(A, B)\}. \tag{8}$$

For the optimal control of elliptic PDEs (the case $s = 1$), and under maximum regularity assumptions, it is known that

$$d(\mathcal{H}_h^1, \mathcal{G}_h^1) \leq C\frac{h^2}{\beta}. \tag{9}$$

Consequently, when solving (6) using multigrid preconditioned conjugate gradient (CG), the number of iterations will decrease with increasing resolution at the optimal rate. This is significant, since at higher resolutions the most expensive operation is precisely the Hessian-vector multiplication. A decrease in the power of h in (9), which can occur in a number of instances (boundary control, loss of elliptic regularity), results in (7) becoming a less efficient preconditioner.

We conduct our analysis using Lemma 1 in [26], which requires estimating the operator $L^2(\Omega)$-norm

$$\|\mathcal{K}_h^s - \mathcal{E}^h\mathcal{K}_{2h}^s\pi_{2h}\| = \sup_{z\in\mathcal{V}_h} \frac{\|(\mathcal{K}_h^s - \mathcal{E}^h\mathcal{K}_{2h}^s\pi_{2h})z\|}{\|z\|}, \tag{10}$$

where $\|\cdot\|$ on the right-hand-side denotes the norm in $L^2(\Omega)$, and $\mathcal{E}^h$ also denotes the restriction of $\mathcal{E}^h$ to $\mathcal{V}_h^0$. From here on, $\|\cdot\|$ without subscripts represents either the vector or the operator L^2-norm, depending on the context. Notice that only the control-to-state solution operators play a role in (10). The estimation process is based on the following apriori estimate in Corollary 2 from [11], which assumes Ω to be convex in $\mathbb{R}^2$ or $\mathbb{R}^3$: for any $s \in (0, 1)$ and $\varepsilon' > 0$, there exists $C = C(\varepsilon', s)$ so that

$$\|\mathcal{K}_h^s z - \mathcal{K}^s z\| \leq Ch^{2s-\varepsilon'}\|z\|. \tag{11}$$

We also recall the following regularity estimate: for $s \in (0, 1)$ there exists C (uniformly bounded in s) so that:

$$\|\mathcal{K}^s z\|_{\mathbb{H}^{2s}(\Omega)} \leq C\|z\|. \tag{12}$$

To show (12), let $z = \sum_{k=1}^{\infty} z_k\varphi_k$ in $L^2(\Omega)$ and $u = \sum_{k=1}^{\infty} u_k\varphi_k = \mathcal{K}^s z$ be the solution of the state equation (1b). Then

$$u_k = \lambda_k^{-s} z_k.$$

From the definition of $\mathbb{H}^s$-norm in (2), we have that

$$\|u\|^2_{\mathbb{H}^{2s}(\Omega)} = \sum_{k=1}^{\infty} \lambda_k^{2s} u_k^2\varphi_k = \sum_{k=1}^{\infty} \lambda_k^{2s}\lambda_k^{-2s} z_k^2\varphi_k = \|z\|^2.$$

As a consequence of convergence (11) and regularity (12) we obtain the following uniform bound (with respect to h) of the operator norm of $\mathcal{K}_h^s$: there exists L_s independent of h so that

$$\|\mathcal{K}_h^s z\| \leq L_s\|z\|, \quad \forall z \in \mathcal{V}_h. \tag{13}$$

Lemma 3.1. *Assume $\Omega \subset \mathbb{R}^N$ with $N = 2, 3$ be a convex polygonal bounded domain. Then for any $\varepsilon' > 0$ and $s \in (0, 1)$ there is a constant $C_s > 0$ so that*

$$\|\mathcal{K}_h^s - \mathcal{E}^h\mathcal{K}_{2h}^s\pi_{2h}\| \leq C_s h^{2s-\varepsilon'}. \tag{14}$$

Proof. For $z \in \mathcal{V}_h$ we have

$$\|(\mathcal{K}_h^s - \mathcal{E}^h\mathcal{K}_{2h}^s\pi_{2h})z\| \leq \|(\mathcal{K}_h^s - \mathcal{K}^s)z\| + \|\mathcal{K}^s(I - \pi_{2h})z\| + \|(\mathcal{K}^s - \mathcal{K}_{2h}^s\pi_{2h})z\|, \tag{15}$$

where we omitted the embedding operators. Using (11) we can bound the first and third terms on the right-hand side of (15) by

$$\|(\mathcal{K}_h^s - \mathcal{K}^s)z\| \leq Ch^{2s-\varepsilon'}\|z\|, \quad \|(\mathcal{K}^s - \mathcal{K}_{2h}^s)\pi_{2h}z\| \leq C(2h)^{2s-\varepsilon'}\|z\|, \tag{16}$$

where we have also used $\|\pi_{2h}z\| \leq \|z\|$. For the middle term in (15) we interpolate between the inequalities (see [27])

$$\|u - \pi_h u\| \leq Ch\|u\|_{H_0^1(\Omega)}, \quad \|u - \pi_h u\| \leq \|u\|_{L^2(\Omega)}, \tag{17}$$

that hold for all $u \in H_0^1(\Omega)$, respectively $u \in L^2(\Omega)$. It follows that

$$\|u - \pi_h u\| \leq Ch^s\|u\|_{\mathbb{H}^s(\Omega)}, \quad \forall u \in \mathbb{H}^s(\Omega). \tag{18}$$

Hence,

$$\begin{aligned}\|\mathcal{K}^s(I - \pi_{2h})z\| &= \sup_{v\in L^2(\Omega)} \frac{|(\mathcal{K}^s(I - \pi_{2h})z, v)|}{\|v\|} = \sup_{v\in L^2(\Omega)} \frac{|((I - \pi_{2h})z, \mathcal{K}^s v)|}{\|v\|} \\ &= \sup_{v\in L^2(\Omega)} \frac{|((I - \pi_{2h})z, \mathcal{K}^s v - \pi_{2h}\mathcal{K}^s v)|}{\|v\|}\end{aligned}$$

Table 1
Direct measurements of spectral distances in one spatial dimension for two different regularization parameters β and different mesh sizes. The dependence on β is according to the predicted theory in Theorem 3.1 and the dependence on mesh size h is according to the Conjecture 3.1.

$\beta = 1$							
N	16	32	64	128	256	512	1024
$s = 0.25$	3.51e−2	1.78e−2	8.97e−03	4.50e−3	2.25e−3	1.13e−3	5.64e−4
$\log_2(d_i/d_{i+1})$	0.9771	0.9910	0.9961	0.9982	0.9991	0.9996	
$s = 0.3$	1.82e−2	8.02e−3	3.51e−3	1.53e−3	6.66e−4	2.90e−4	1.26e−4
$\log_2(d_i/d_{i+1})$	1.1807	1.1931	1.1976	1.1991	1.1997	1.1999	
$s = 0.4$	4.81e−3	1.61e−3	5.34e−4	1.76e−4	5.82e−5	1.92e−5	6.34e−6
$\log_2(d_i/d_{i+1})$	1.5780	1.5934	1.5976	1.5993	1.5998	1.5999	
N	64	128	256	512	1024	2048	4096
$s = 0.5$	1.20e−4	2.71e−5	6.16e−6	1.46e−6	3.43e−7	8.24e−8	2.03e−08
$\log_2(d_i/d_{i+1})$	2.1432	2.1386	2.0742	2.0949	2.0566	2.0194	
$s = 0.6$	8.53e−5	1.89e−5	4.45e−6	1.02e−6	2.40e−7	5.81e−8	1.40e−8
$\log_2(d_i/d_{i+1})$	2.1730	2.0865	2.1167	2.0946	2.0486	2.0536	
$s = 0.7$	5.83e−5	1.37e−5	3.15e−6	7.25e−7	1.71e−7	4.13e−8	1.01e−8
$\log_2(d_i/d_{i+1})$	2.0930	2.1184	2.1195	2.0805	2.0524	2.0350	
$\beta = 0.1$							
N	16	32	64	128	256	512	1024
$s = 0.25$	3.06e−1	1.66e−1	8.63e−2	4.41e−2	2.23e−2	1.12e−2	5.62e−3
$\log_2(d_i/d_{i+1})$	0.8846	0.9391	0.9685	0.9840	0.9919	0.9959	
$s = 0.3$	1.68e−1	7.75–2	3.45e−2	1.52e−2	6.64e−3	2.90e−3	1.26.e-3
$\log_2(d_i/d_{i+1})$	1.1210	1.1651	1.1850	1.1935	1.1972	1.1988	
$s = 0.4$	4.71e−2	1.60e−2	5.33e−3	1.76e−3	5.82–4	1.92e−4	6.34e−5
$\log_2(d_i/d_{i+1})$	1.5578	1.5864	1.5953	1.5985	1.5995	1.5998	
N	64	128	256	512	1024	2048	4096
$s = 0.5$	8.74e−4	2.12e−4	5.20e−5	1.29e−5	3.20e−6	7.98e−7	1.99e−8
$\log_2(d_i/d_{i+1})$	2.0416	2.0299	2.0118	2.0111	2.0045	2.0019	
$s = 0.6$	5.72e−4	1.28e−4	3.02e−5	7.01e−6	1.65e−6	4.00e−7	9.73e−8
$\log_2(d_i/d_{i+1})$	2.1625	2.0807	2.1077	2.0866	2.0445	2.0400	
$s = 0.7$	4.40e−4	1.03e−4	2.39e−5	5.51e−6	1.31e−6	3.15e−7	7.72e−8
$\log_2(d_i/d_{i+1})$	2.0900	2.1138	2.1141	2.0764	2.0496	2.0330	

$$\overset{(18)}{\leq} (2h)^{2s} \sup_{v\in L^2(\Omega)} \frac{\|(I-\pi_{2h})z\|\,\|\mathcal{K}^s v\|_{\mathbb{H}^{2s}(\Omega)}}{\|v\|}$$
$$\overset{(12)}{\leq} C(2h)^{2s}\|(I-\pi_{2h})z\| \leq C' h^{2s}\|z\|. \tag{19}$$

The conclusion follows from (16) and (19). ■

The next theorem follows from Lemma 3.1 and Lemma 1 in [26].

Theorem 3.1. *If $C_s h^{2s-\varepsilon'} \leq \beta/(4L_s)$, then*

$$d(\mathcal{H}_h^s, \mathcal{G}_h^s) \leq 4L_s\beta^{-1}h^{2s-\varepsilon'}. \tag{20}$$

This result certifies that the quality of the two-grid (and hence multigrid) preconditioner is improving with increasing resolution, as in the elliptic case, but at a rate that is degrading as s decreases to 0. Consequently, the preconditioner is expected to be less efficient as s decreases. At the same time, the coarsest mesh that can be used may also need to be finer and finer as s decreases due to the hypothesis in Theorem 3.1; hence, the number of levels that can be used at some point will necessarily be smaller. Remarkably, the numerical results in Section 4 show an improved picture: they suggest that in fact a significantly stronger estimate holds. Hence, we formulate the following conjecture.

Conjecture 3.1. *Assuming the domain is convex, there is a constant $\tilde{C}_s$ independent of h, so that, if h is sufficiently small,*

$$d(\mathcal{H}_h^s, \mathcal{G}_h^s) \leq \begin{cases} \tilde{C}_s\beta^{-1}h^{4s}, & \text{if } 0 < s < 1/2 \\ \tilde{C}_s\beta^{-1}h^{2}, & \text{if } 1/2 \leq s < 1. \end{cases} \tag{21}$$

It is notable that (21) is consistent with the classical result (9) for $s = 1$, and also with Theorem 3.1 as $s \approx 1$. However, it shows that the preconditioner is uniformly very good when $1/2 \leq s < 1$ and even with the classical case $s = 1$, and is twice as efficient compared to what the analysis predicts for $0 < s < 1/2$. Proving Conjecture 3.1 requires a different approach from proving Theorem 3.1, since we do not expect any superconvergence to hold in (14). Instead, we expect the proof of the conjecture to involve higher order estimates in weaker norms for the control-to-state map, in addition to more refined regularity results.

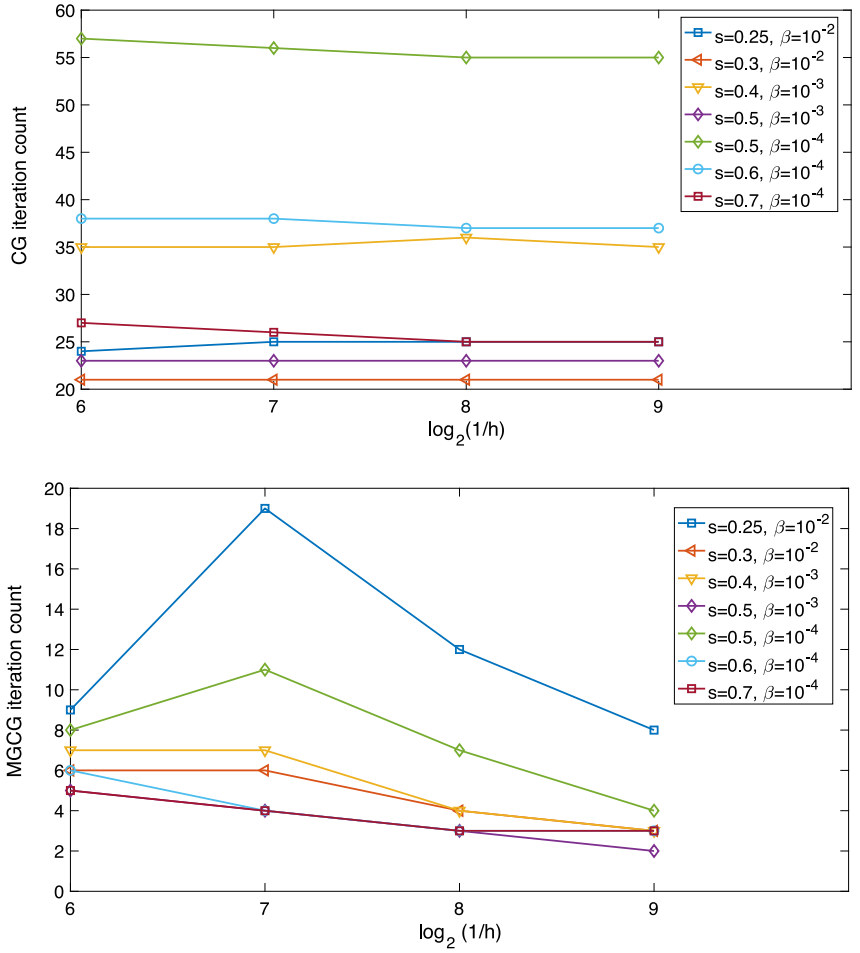

Fig. 1. The top and bottom panels respectively show the iteration counts for CG and MGCG with respect to the mesh-size. As expected, we observe the iteration count to be roughly independent of the mesh-size for CG; for MGCG it is expected to decrease as the mesh-size decreases. An exception may be the transition from two to three grids, where the number of iterations may increase. With the base case being $h = 2^{-5}$, we see in two of the cases an increase in the MGCG iteration count when refining from $h = 2^{-6}$ (two grids) to $h = 2^{-7}$ (three grids).

4. Numerical experiments

We have performed two kinds of numerical experiments. First we aim to verify (21) directly by building matrices corresponding to the Hessian and the two-grid preconditioner for a set of grids with $h_j = 2^{-j}, j = j_{\min}, \ldots, j_{\max}$, followed by a direct computation of $d^s_{h_j} := d(\mathcal{H}^s_{h_j}, \mathcal{G}^s_{h_j})$ using generalized eigenvalues ($\mathcal{H}^s_{h_j} u = \lambda \mathcal{G}^s_{h_j} u$). Then we form the ratios $d^s_{h_{j-1}}/d^s_{h_j}$ to confirm the formula (21). We show results for $\beta = 1$ and $\beta = 0.1$ for this purpose. However, due to the sizes of the matrices involved, these computations are limited to the one-dimensional case $\Omega = (0, 1)$. The results for $s = 0.25, 0.3, 0.4, 0.5, 0.6, 0.7$ are shown in Table 1, and they strongly support Conjecture 3.1. The precise values of $j_{\min}, j_{\max}$ vary with s, due primarily to memory limitation (smaller s requires more memory). It is notable that the spectral distances in the lower part of the table are approximately ten times larger than their counterparts in the upper half (for a value of β that is ten times smaller), thus also supporting the dependence on β in (21).

The second kind of numerical results are actual two-dimensional solves in $\Omega = (0, 1)^2$ of (6), i.e., our optimal control problem with a multigrid version of the preconditioner. The data is $u_d(x, y) = \sin(4\pi x)\sin(3\pi y)$. For each case considered, we compare the number of unpreconditioned CG iterations to the number of multigrid preconditioned CG (MGCG) iterations, and we report the wall-clock times. The results are reported in Table 2 and also in Figs. 1 and 2. The solvers are all matrix-based, in the sense that the sparse matrices implementing the operators $\mathcal{K}^s_h$ are formed in block-diagonal form and prefactored. Only the coarsest Hessian is formed at resolution 32×32, which is used as the base case for all cases considered. The effect of decreasing the value of the regularizer β and/or that of the parameter s is an increase in the number of CG iterations. In order to maintain the number of unpreconditioned CG iterations between 20 and 50 (for illustration purposes) we have chosen slightly larger values of β as we decreased s in the experiments

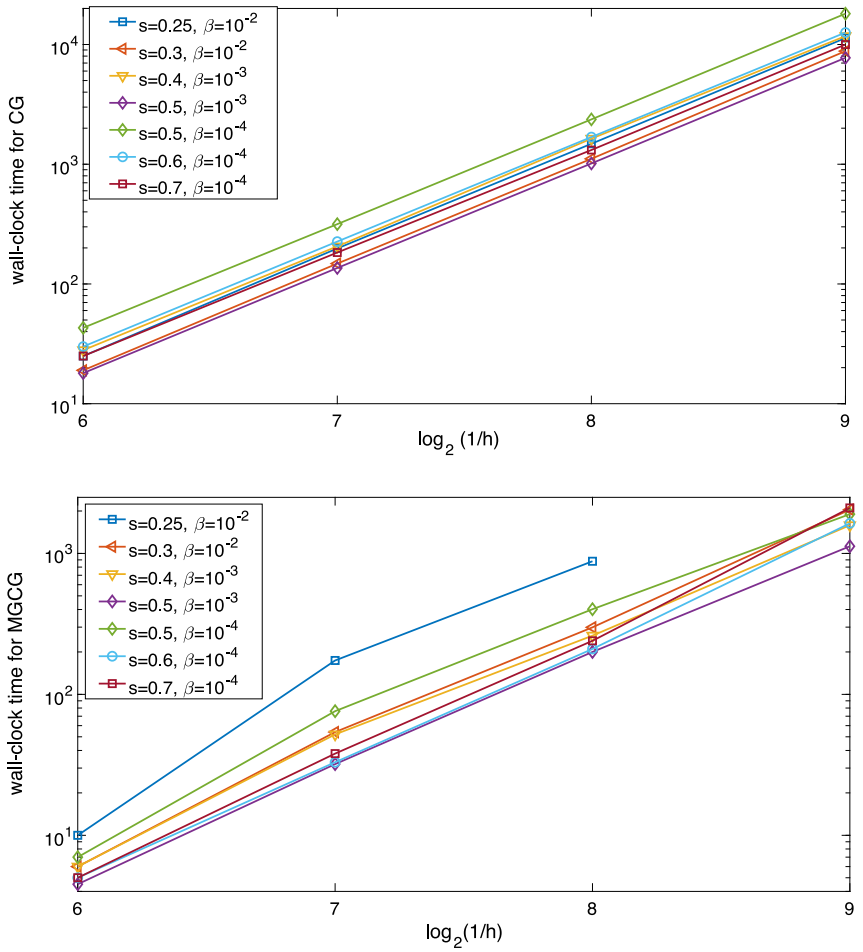

Fig. 2. The top and bottom panels respectively show the wall clock time for CG and MGCG with respect to the mesh-size. As expected, on the log–log plot we observe a perfectly linear behavior for CG. For MGCG, the behavior is sublinear for all the cases for which transitioning to a finer grid results in a drop of the number of iterations. We show two cases for which there is no drop: for $s = 0.6, 0.7$, when refining the grid from $h = 2^{-8}$ to $h = 2^{-9}$, the number of iterations remains 3, which is very low. The outlier value marked with superscript "a" in Table 2 was excluded ($s = 0.25, h = 2^{-9}$), since it only reflects that the machine was forced into much slower swap space.

described below. The number of CG iterations also indicates the difficulty of the problem at hand, as it corresponds to the number of relevant eigenmodes that can be recovered for the control for a given problem setting. All computations were performed using MATLAB on a system with two eight-core 2.9 GHz Intel Xeon E5-2690 CPUs and 256 GB memory.

The cases include $s = 0.25, 0.3, 0.4, 0.5, 0.6, 0.7$. The results show a dramatic reduction in the number of MGCG iterations compared to unpreconditioned CG, as well as a reduction in computing time. It is notable that for each case, the number of MGCG iterations is ultimately decreasing with increasing resolution. E.g., for $s = 0.4$ the number of multigrid CG iterations, decreases from 7 on a 64×64 grid to 3 on a 512×512 grid, while the number of unpreconditioned CG iterations is virtually constant. However, for $s = 0.25$ the decrease is less dramatic. It is expected that for a regular, iterative or parallel implementation of the matrix–vector product of the Hessian, the dramatic decrease in number of iterations will be reflected in the decrease of computing time, since the most expensive iteration remains at the finest-level fractional Poisson solve.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

H. Antil and K. Green are partially supported by NSF, USA DMS-1913004 and DMS-1818772. A. Drăgănescu is partially supported by NSF, USA DMS-1913201 and the U.S. Department of Energy Office of Science, Office of Advanced Scientific

Table 2
Iteration counts for unpreconditioned CG vs. MGCG with base case 32 × 32; wall-clock times are shown in seconds in parenthesis.

64 × 64		128 × 128		256 × 256		512 × 512	
CG	MGCG	CG	MGCG	CG	MGCG	CG	MGCG
$s = 0.25, \quad \beta = 10^{-2}$							
24 (25)	9 (10)	25 (198)	19 (174)	25 (1482)	12 (880)	25 (11326)	8 (23876[a])
$s = 0.3, \quad \beta = 10^{-2}$							
21 (19)	6 (6)	21 (148)	6 (54)	21 (1110)	4 (299)	21 (8805)	3 (2045)
$s = 0.4, \quad \beta = 10^{-3}$							
35 (28)	7 (6)	35 (206)	7 (52)	36 (1635)	4 (261)	35 (11753)	3 (1585)
$s = 0.5, \quad \beta = 10^{-3}$							
23 (18)	5 (4.5)	23 (136)	4 (32)	23 (1015)	3 (200)	23 (7743)	2 (1124)
$s = 0.5, \quad \beta = 10^{-4}$							
57 (43)	8 (7)	56 (316)	11 (76)	55 (2369)	7 (402)	55 (18090)	4 (1900)
$s = 0.6, \quad \beta = 10^{-4}$							
38 (30)	6 (5)	38 (226)	4 (33)	37 (1687)	3 (210)	37 (12591)	3 (1635)
$s = 0.7, \quad \beta = 10^{-4}$							
27 (25)	5 (5)	26 (183)	4 (38)	25 (1315)	3 (240)	25 (9991)	3 (2103)

[a]The time marked is not relevant, because the computation was forced into much slower swap space due to memory limitations.

Computing Research under Award Number DE-SC0005455. The authors also thank the anonymous reviewers for the careful reading and useful suggestions.

References

[1] Antil H, Pfefferer J, Rogovs S. Fractional operators with inhomogeneous boundary conditions: analysis, control, and discretization. Commun Math Sci 2018;16(5):1395–426. http://dx.doi.org/10.4310/CMS.2018.v16.n5.a11.
[2] Weiss CJ, van Bloemen Waanders BG, Antil H. Fractional operators applied to geophysical electromagnetics. Geophys J Int 2020;220(2):1242–59.
[3] Glusa CA, Antil H, D'Elia M, Weiss CJ, van Bloemen Waanders BG. A fast solver for the fractional Helmholtz equation. Tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States); 2019.
[4] Larkin P. Developments in direct-field acoustic testing. Sound Vib Mag 2014.
[5] Auger E, D'Auria L, Martini M, Chouet B, Dawson P. Real-time monitoring and massive inversion of source parameters of very long period seismic signals: An application to Stromboli Volcano, Italy. Geophys Res Lett 2006;33(4).
[6] Antil H, Bartels S. Spectral approximation of fractional PDEs in image processing and phase field modeling. Comput Methods Appl Math 2017;17(4):661–78.
[7] Antil H, Di ZW, Khatri R. Bilevel optimization, deep learning and fractional Laplacian regularizatin with applications in tomography. Inverse Problems 2020;36(6):064001. http://dx.doi.org/10.1088/1361-6420/ab80d7.
[8] Antil H, Otárola E. A FEM for an optimal control problem of fractional powers of elliptic operators. SIAM J Control Optim 2015;53(6):3432–56. http://dx.doi.org/10.1137/140975061.
[9] Stinga P, Torrea J. Extension problem and Harnack's inequality for some fractional operators. Comm Partial Differential Equations 2010;35(11):2092–122. http://dx.doi.org/10.1080/03605301003735680.
[10] Caffarelli L, Silvestre L. An extension problem related to the fractional Laplacian. Comm Partial Differential Equations 2007;32(7–9):1245–60. http://dx.doi.org/10.1080/03605300600987306.
[11] Dohr S, Kahle C, Rogovs S, Swierczynski P. A FEM for an optimal control problem subject to the fractional Laplace equation. Calcolo 2019;56(4):37.
[12] Kato T. Note on fractional powers of linear operators. Proc Japan Acad 1960;36(3):94–6.
[13] Bonito A, Pasciak J. Numerical approximation of fractional powers of elliptic operators. Math Comp 2015;84(295):2083–110. http://dx.doi.org/10.1090/S0025-5718-2015-02937-8.
[14] Heidel G, Khoromskaia V, Khoromskij BN, Schulz V. Tensor approach to optimal control problems with fractional d-dimensional elliptic operator in constraints. 2018, arXiv preprint arXiv:1809.01971.
[15] Antil H, Warma M. Optimal control of fractional semilinear PDEs. ESAIM Control Optim Calc Var 2020;26. http://dx.doi.org/10.1051/cocv/2019003, Paper No. 5, 30.
[16] D'Elia M, Glusa C, Otárola E. A priori error estimates for the optimal control of the integral fractional Laplacian. SIAM J Control Optim 2019;57(4):2775–98. http://dx.doi.org/10.1137/18M1219989.
[17] Antil H, Warma M. Optimal control of the coefficient for the regional fractional p-Laplace equation: Approximation and convergence. Math Control Relat Fields 2019;9(1):1–38. http://dx.doi.org/10.3934/mcrf.2019001.
[18] Antil H, Warma M. Optimal control of the coefficient for fractional $\{p\}$-Laplace equation: Approximation and convergence. RIMS Kôkyûroku Bessatsu 2018;2090:102–16.
[19] Antil H, Khatri R, Warma M. External optimal control of nonlocal PDEs. Inverse Problems 2019;35(8):084003, 35. http://dx.doi.org/10.1088/1361-6420/ab1299.
[20] Antil H, Verma D, Warma M. External optimal control of fractional parabolic PDEs. ESAIM Control Optim Calc Var 2020;26. http://dx.doi.org/10.1051/cocv/2020005.
[21] Ainsworth M, Glusa C. Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver. Comput Methods Appl Mech Engrg 2017;327:4–35. http://dx.doi.org/10.1016/j.cma.2017.08.019.

[22] Borzì A, Schulz V. Computational optimization of systems governed by partial differential equations. Computational science & engineering, vol. 8, Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM); 2012.
[23] Drăgănescu A, Dupont T. Optimal order multilevel preconditioners for regularized ill-posed problems. Math Comput 2008;77(264):2001–38.
[24] Tartar L. An introduction to Sobolev spaces and interpolation spaces. Lecture notes of the unione matematica italiana, vol. 3, Berlin: Springer; 2007.
[25] Antil H, Pfefferer J. A short matlab implementation of fractional Poisson equation with nonzero boundary conditions. Tech. rep., 2017.
[26] Barker AT, Drăgănescu A. Algebraic multigrid preconditioning of the Hessian in optimization constrained by a partial differential equation. Numer Linear Algebra Appl 2020. http://dx.doi.org/10.1002/nla.2333.
[27] Brenner SC, Scott L. The mathematical theory of finite element methods, In: Texts in applied mathematics, 3rd ed.. vol. 15, New York: Springer; 2008, p. xviii+397. http://dx.doi.org/10.1007/978-0-387-75934-0.