A FEM for an optimal control problem subject to the fractional Laplace equation

Stefan Dohr · Christian Kahle · Sergejs Rogovs · Piotr Swierczynski

Received: 5 November 2018 / Revised: 10 September 2019 / Accepted: 13 September 2019 / Published online: 19 September 2019
© Istituto di Informatica e Telematica (IIT) 2019

Abstract
We study the numerical approximation of linear–quadratic optimal control problems subject to the fractional Laplace equation with its spectral definition. We compute an approximation of the state equation using a discretization of the Balakrishnan formula that is based on a finite element discretization in space and a sinc quadrature approximation of the additionally involved integral. A tailored approach for the numerical solution of the resulting linear systems is proposed. Concerning the discretization of the optimal control problem we consider two schemes. The first one is the variational approach, where the control set is not discretized, and the second one is the fully discrete scheme where the control is discretized by piecewise constant functions. We derive finite element error estimates for both methods and illustrate our results by numerical experiments.

Keywords Fractional Laplacian · Linear–quadratic optimal control problem · Finite element method · A priori error estimates · Dunford–Taylor integral

This work was supported by the International Research Training Group 1754, funded by the German Research Foundation (DFG), and the Austrian Science Fund (FWF).

Christian Kahle
christian.kahle@ma.tum.de
Stefan Dohr
stefan.dohr@tugraz.at
Sergejs Rogovs
sergejs.rogovs@unibw.de
Piotr Swierczynski
piotr.swierczynski@ma.tum.de

1 Institut für Angewandte Mathematik, Technische Universität Graz, 8010 Graz, Austria
2 Zentrum Mathematik, Technische Universität München, 85748 Garching bei München, Germany
3 Institut für Mathematik und Computergestützte Simulation, Universität der Bundeswehr München, 85577 Neubiberg, Germany
Mathematics Subject Classification 65N30 · 35J15 · 49K20

1 Introduction

Let $\Omega \subset \mathbb{R}^n$ ($n \in \{2, 3\}$) be a polygonally ($n = 2$)/polyhedrally ($n = 3$) bounded Lipschitz domain with boundary $\Gamma := \partial \Omega$ and $s \in (0, 1)$. For $u_d : \Omega \to \mathbb{R}$ we define the objective functional

$$J(u, z) := \frac{1}{2} \| u - u_d \|^2_{L^2(\Omega)} + \frac{\mu}{2} \| z \|^2_{L^2(\Omega)},$$

where $\mu > 0$ denotes a regularization parameter. In this work we consider the optimal control problem of finding

$$\arg \min_z J(u, z),$$

subject to the fractional state equation

$$(-\Delta)^s u = z\text{ in } \Omega, \quad u = 0\text{ on } \Gamma,$$

and the control constraints

$$a \leq z(x) \leq b \quad \text{a.e. in } \Omega,$$

with constants $a, b \in \mathbb{R}$ satisfying $a \leq 0 \leq b$. Here, we understand the operator $(-\Delta)^s$ in the sense of its spectral definition, compare e.g. [4,15,17,33]. The main difficulty in studying this problem is the nonlocality of the fractional Laplace operator [15].

One way to overcome this issue is based on the Cafarelli–Silvestre extension [16] on unbounded domains and its extension to bounded domains [15,17,41]. Let us summarize the results on the numerical approximation of (3) with this approach. Here, an auxiliary problem in an extended domain $\mathcal{C} := \Omega \times (0, \infty)$ is introduced and the solution of the state equation (3) is then given as the Dirichlet trace on $\Omega \times \{0\}$ of the solution to the extended problem. Exponential decay of the solution in the artificial dimension allows construction of different numerical methods, see e.g. [10,29,33]. In these publications, the problem is discretized by introducing a tensor product mesh of the domain $\mathcal{C}_Y = \Omega \times (0, Y)$, which is constructed by a conformal triangulation of $\Omega$ and a graded mesh in the artificial direction, see e.g. [33, Section 5.1]. A convergence rate of $h^{1+s}$ (up to some logarithmic term) in the $L^2(\Omega)$-norm can be obtained [3,34], provided that $z \in H^{1-s}(\Omega)$, where $h$ denotes the global mesh parameter. However, numerical experiments show that this convergence rate is not optimal in a specific range of fractional powers $s$. The cost of solving the problem is related to the number of elements in $\mathcal{C}_Y$, and not only to the number of elements in $\Omega$, resulting in an increased computational complexity. This issue is first overcome in [29] by exploiting $p$-finite elements in the extended direction. The same approach is also proposed by [8]. In the latter work, additionally a sparse tensor product FEM based on multilevel FEM in $\Omega$ and on geometric meshes in the extended variable is proposed, that leads to the
same asymptotic complexity. In addition, in [8] the authors present exponential rates of convergence of a $hp$-FEM for spectral, fractional diffusion operators in energy norm under sufficient regularity assumptions. This is achieved by a combined tensor product $hp$-FEM for the Caffarelli–Silvestre extension in the truncated cylinder $\Omega \times (0, \mathcal{Y})$ with anisotropic geometric meshes that are refined towards $\partial \Omega$.

An alternative approach for solving (3) uses the Balakrishnan representation formula [43, IX. 11.], namely for $s \in (0, 1)$ and $z \in \mathbb{H}^{-s}(\Omega)$

$$(-\Delta)^{-s}z = \frac{\sin (s\pi)}{\pi} \int_0^\infty v^{-s}(vI - \Delta)^{-1}zd\nu. \quad (5)$$

Numerical approximation of (5) is then based on a suitable quadrature formula for (5) with respect to $\nu$ and a discretization of the operator $vI - \Delta$ using the finite element method, see [10].

Concerning literature on optimal control of the fractional Laplace equation we also refer to [2–5,21,25,36].

While the numerical analysis of the optimal control problem (2)–(4) using an equivalent formulation with the Caffarelli–Silvestre extension is well established, the numerical analysis using the Balakrishnan formula is still open.

In this article we propose and analyze two discrete schemes for the approximation of the solution to the optimal control problem (2)–(4) using the Balakrishnan representation of the solution $u$ of the state equation (3). Both schemes rely on a finite element discretization of the operator $vI - \Delta$ in (5) and a sinc quadrature approximation [11] of the integral in (5). The first method is the variational discretization approach [27], where the set of controls is not discretized a priori. However, it inherits its approximation properties from the approximation of the adjoint state. The second one uses a fully discrete setting, where the set of controls is discretized by piecewise constant functions [6,18,37]. We derive $L^2(\Omega)$-error estimates for the state and control for both types of the FE discretization of the optimal control problem.

Regarding the variational approach for the discretization of the optimal control problem (2)–(4) we show an optimal convergence rate of $h^{\min(2,3/2+2s)}$ (up to some logarithmic term) for the control and the state in the $L^2(\Omega)$-norm, whereas using the extension approach [3] yields a convergence rate of $h^{1+s}$ (up to some logarithmic term).

In the case of the fully discrete scheme we show the expected linear convergence for the control in the $L^2(\Omega)$-norm and for the state in the $\mathbb{H}^s(\Omega)$-norm. Numerically we also consider the post-processing approach [30] for the optimal control and measure again the same rate of $h^{\min(2,3/2+2s)}$ for the post-processed optimal control. Similar results are shown for the extension approach in [3]. While the convergence rate for the optimal control is optimal, as confirmed by numerical experiments, there is still a gap between the theoretical and practical rates for the state, which will be addressed in future work.

The outline of this paper is as follows. In Sect. 2 we review existence and uniqueness results for the fractional optimal control problem based on [3] as well as regularity properties of the optimal control problem. The numerical analysis of the mentioned discretization methods is conducted in Sect. 3, starting with the derivation of the
error estimates for the discretization of the state equation (3) using the Balakrishnan formula. In Sect. 3.2 we study the convergence properties of the optimal control and state using the semidiscrete approach, while Sect. 3.3 is devoted to the numerical analysis of the fully discrete scheme. In Sect. 4 we introduce a solver for the finite element approximation of the problem. Numerical results validating the theoretical convergence results for the proposed discretization techniques are presented in Sect. 5.

2 Existence and regularity of optimal controls

In this section we review existence and uniqueness as well as the regularity results for the optimal control problem (2)–(4) based on [3, Sec. 3]. We start this section with a brief introduction of the spectral definition of the fractional operator $(-\Delta)^s$ following [15,17].

The eigenfunctions $\{\varphi_k\}_{k \in \mathbb{N}}$ with eigenvalues $\{\lambda_k\}_{k \in \mathbb{N}}$ of the Laplace operator, i.e.,

$$-\Delta \varphi_k = \lambda_k \varphi_k \text{ in } \Omega, \quad \varphi_k = 0 \text{ on } \Gamma, \quad k \in \mathbb{N}$$

form an orthonormal basis of $L^2(\Omega)$. The spectral fractional Laplace operator (also called Navier Laplacian) for $w \in C_0^\infty(\Omega)$ is then defined as

$$(-\Delta)^s w := \sum_{k=1}^\infty \lambda_k^s w_k \varphi_k, \quad w_k := \int_\Omega w \varphi_k dx, \quad k \in \mathbb{N}. $$

This definition can be extended by density to the space $H^s(\Omega)$ [3,33] defined as

$$H^s(\Omega) := \left\{ w = \sum_{k=1}^\infty w_k \varphi_k : \sum_{k=1}^\infty \lambda_k^2 w_k^2 < \infty \right\} = \begin{cases} H^s(\Omega) \equiv H_0^s(\Omega) & \text{if } s \in (0, \frac{1}{2}], \\ H^{1/2}_0(\Omega) & \text{if } s = \frac{1}{2}, \\ H_0(\Omega) & \text{if } s \in (\frac{1}{2}, 1]. \\ \end{cases} \quad (6)$$

The characterization of the fractional Sobolev–Slobodeckij spaces on the right hand side in (6) can be found, e.g., in [3]. It further holds $H^s(\Omega) = H^s(\Omega) \cap H_0^1(\Omega)$ for $s \in (1, \frac{3}{2})$ as stated in [10,32]. If additionally $\Omega$ is convex, then additionally $H^s(\Omega) = H^s(\Omega) \cap H_0^1(\Omega)$ for $s \in (\frac{3}{2}, 2)$, see [36, Section 2.2].

The dual space of $H^s(\Omega)$ we denote by $H^{-s}(\Omega)$. We stress that this definition of $(-\Delta)^s$ inherently assumes homogeneous Dirichlet boundary data (in a suitable sense). For a generalization to inhomogeneous Dirichlet boundary data in combination with the spectral definition of the fractional Laplacian we refer to [1,20] and especially [4].

Let $u_d \in L^2(\Omega)$ and $a, b \in \mathbb{R}$ with $a \leq 0 \leq b$ be given. We define the set of admissible controls $Z_{ad}$ by

$$Z_{ad} := \left\{ w \in L^2(\Omega) : a \leq w(x) \leq b \text{ a.e. in } \Omega \right\}. $$
Let $S : H^{-s}(Ω) \rightarrow H^s(Ω)$ denote the control to state operator defined as $Sz := u$, where $u \in H^s(Ω)$ is the unique solution of the state equation (3). It holds, that for any $z \in H^{-s}(Ω)$, the boundary value problem (3) has a unique solution $u \in H^s(Ω)$, see e.g. [29]. We may also consider the operator $S$ acting on $L^2(Ω)$ with range in $L^2(Ω)$. Note also that $S$ is self-adjoint, since the operator $(-\Delta)^s$ is self-adjoint. The adjoint state $p \in H^s(Ω)$ for $z \in H^{-s}(Ω)$ is then given by $p = S(Sz - u_d)$. In [3, Section 3.1] the existence and uniqueness of a solution to the optimal control problem (2)–(4) is shown. Let us recall the main result from that reference.

**Theorem 1** (Existence, uniqueness, optimality conditions [3, Section 3.1]) The optimal control problem (2)–(4) has a unique optimal solution $(\bar{u}, \bar{z}) \in H^s(Ω) \times Z_{ad}$. These fulfill the necessary and sufficient optimality conditions

\begin{align}
\bar{u} = S\bar{z} \in H^s(Ω), \quad & (7) \\
\bar{p} = S(\bar{u} - u_d) \in H^s(Ω), \quad & (8) \\
\bar{z} \in Z_{ad}, \quad (\mu\bar{z} + \bar{p}, z - \bar{z})_{L^2(Ω)} \geq 0 \quad \text{for all } z \in Z_{ad}. \quad & (9)
\end{align}

From [3,36] we additionally obtain the following regularity results.

**Lemma 1** ($H^1$-regularity of the optimal control, [3, Lemma 3.5])

Let $Ω$ be polygonally/polyhedrally bounded, $u_d \in H^{1-s}(Ω)$, and $a, b \in H^1(Ω)$. Let $\bar{z} \in Z_{ad}$ be the optimal control. If $s \in [\frac{1}{2}, 1)$ then $\bar{z} \in H^1(Ω)$. If $s \in (0, \frac{1}{4})$ and $a \leq 0 \leq b$ hold, then $\bar{z} \in H^1_0(Ω)$

**Lemma 2** (Higher regularity of state and adjoint on convex domains, [36, Lemma 3.4]) Let $Ω$ be convex and $u_d \in H^{1-s}(Ω)$, $a < 0 < b$. Then $\bar{u} \in H^{\min(1+2s,2)}(Ω)$. If additionally $u_d \in H^1_0(Ω)$, then $\bar{p} \in H^{\min(1+2s,2)}(Ω)$.

For $\mu > 0$ and $\bar{p} = S(\bar{u} - u_d)$ the variational inequality (9) is equivalent to the projection formula [42]

$$\bar{z}(x) = \text{proj}_{[a,b]} \left( -\frac{1}{\mu} \bar{p}(x) \right) \quad (10)$$

where $\text{proj}_{[a,b]}(v) := \min \{b, \max \{a, v\}\}$.

Using the projection operator, we can obtain even more regularity for the optimal control $\bar{z}$ and start with a useful lemma on the continuity of the $\text{proj}_{[a,b]}$ operator.

**Lemma 3** Let $v \in H^t(Ω)$ then $\text{proj}_{[a,b]}(v) \in H^t(Ω)$ for $0 < t < r < \frac{3}{2}$ and there exists $C > 0$ such that $\| \text{proj}_{[a,b]}(v) \|_{H^t(Ω)} \leq C \| v \|_{H^r(Ω)}$.

**Proof** Before we show the desired result, we introduce the space $H^t_{\bar{p}}(\mathbb{R}^n)$ as the Sobolev–Slobodeckij space, where the norm is defined based on the Fourier transformation, see e.g. [31] and note, that $H^1(\mathbb{R}^n) = H^1_{\bar{p}}(\mathbb{R}^n)$, meaning that the corresponding norms are equivalent. Further it is sufficient to investigate the term $\max(0, p)$ instead of $\text{proj}_{[a,b]}(p)$.

From [26, Theorem 1.4.3.1] we obtain, that there exists an extension $\tilde{p} \in H^t(\mathbb{R}^n)$ such that $\tilde{p}|_{Ω} \equiv p$ and $\| \tilde{p} \|_{H^t(\mathbb{R}^n)} \leq C \| p \|_{H^t(Ω)}$ with a constant $C > 0$. This yields
with another generic constant $C > 0$

$$\| \max(0, p) \|_{H^s(\Omega)} = \| \max(0, \tilde{p}) \|_{H^s(\Omega)} \leq \| \max(0, \tilde{p}) \|_{H^s(\mathbb{R}^n)} \leq C \| \max(0, \tilde{p}) \|_{H^s_{\mathcal{P}}(\mathbb{R}^n)}.$$  

Next we note, that $\max(0, x) = \frac{1}{2}(x + |x|)$ and obtain from [31, Proposition 1]

$$\| \max(0, \tilde{p}) \|_{H^s_{\mathcal{P}}(\mathbb{R}^n)} \leq \| \tilde{p} \|_{H^s_{\mathcal{P}}(\mathbb{R}^n)},$$

where [31, Proposition 1] shows this result for the absolute value instead of the max operator. We finish with

$$\| \tilde{p} \|_{H^s_{\mathcal{P}}(\mathbb{R}^n)} \leq C \| \tilde{p} \|_{H^s(\mathbb{R}^n)} = C \| p \|_{H^s(\mathbb{R}^n)}. \tag*{\Box}$$

Remark 1 For related results we also refer to [35, Theorem 1], [38, Theorem 5.4.1], [13, Theorem 2], and [14, Theorem 21].

Lemma 4 Let $\bar{z} \in Z_{ad}$ be the solution of the optimal control problem (2)–(4) with $u_d \in H^{\max(0,3/2-2s)}(\Omega)$. Then $\bar{z} \in H^{3/2-\varepsilon}(\Omega)$, where $\varepsilon$ is a positive, arbitrary small number.

Proof By Lemma 1 we already know $\bar{z} \in H^1_0(\Omega)$, thus $z \in H^1(\Omega)$ as $a \leq 0 \leq b$ is satisfied. Further, by elementary calculation, for $t \geq 0$ and $f \in H^t(\Omega)$ it holds $(-\Delta)^{\frac{1}{2}} f \in H^{t+2s}(\Omega)$. Note, that there is no upper limit for $t$.

Thus, starting with $\bar{z} \in H^1(\Omega)$ we obtain that $\bar{u} - u_d \in H^{1+2s}(\Omega)$ and thus $p \in H^{1+4s}(\Omega)$ is satisfied. As in [3, Lemma 3.5] we need to consider two cases.

Case I: $1 + 4s \geq 3/2$ We restrict the considered regularity to $p \in H^{3/2-\varepsilon}(\Omega) = H^{3/2-\varepsilon}(\Omega)$ for some $\varepsilon > 0$. Lemma 3 yields the result.

Case II: $1 + 4s < 3/2$: In this case we observe $\bar{z} \in H^{1+4s-\varepsilon}(\Omega)$ by Lemma 3 for some $\varepsilon > 0$ and thus $\tilde{p} \in H^{1+8s-\varepsilon}(\Omega)$. If $1 + 8s - \varepsilon \geq 3/2$ we proceed as in case I, else we repeat the argumentation to obtain $p \in H^{1+12s-2\varepsilon}(\Omega)$. By a finite number of bootstrapping arguments, we arrive at $p \in H^{3/2-\varepsilon}(\Omega)$ and finish as in case I.  \( \Box \)

3 A priori error estimates

In this section, we analyse two finite element approximations of the fractional optimal control problem (2)–(4). First, we investigate the variational approach [27], where the control set is not discretized, and then move to a fully discrete scheme. Both techniques are based on a finite element discretization of the state equation (3) using the Balakrishnan formula (5). In the following subsection, we review the resulting FE error estimates, based on [10, Sec. 4].

Assumption 1 In the following we assume, that $\Omega$ is convex. Further we assume that the regularity result from Lemma 4 holds and, for notational convenience, we assume that $u_d \in H^{3/2}(\Omega)$. 

Springer
3.1 A finite element method for the state equation

Let \( \mathcal{U}(\mathcal{T}_h) \) be the space of piecewise linear and globally continuous functions vanishing on the boundary \( \partial \Omega \), defined with respect to a conforming quasi-uniform triangulation \( \mathcal{T}_h \) of the domain \( \Omega \). A FE approximation of problem (3) for \( z \in L^2(\Omega) \) is given by

\[
uh = \sin \left( \frac{s \pi}{\pi} \right) \left( I - \Delta_h \right)^{-1} z = \int_0^\infty \nu^{-s} (\nu I - \Delta_h)^{-1} z \, d\nu,
\]

where \( \Delta_h \) denotes the discrete Laplace operator.

For \( k > 0 \) we define the numbers \( N_+ := \left\lceil \frac{\pi^2}{4sk^2} \right\rceil \) and \( N_- := \left\lceil \frac{\pi^2}{4(1-s)k^2} \right\rceil \). The sinc quadrature approximation of \( uh \) is then given by

\[
uk_h := \sin \left( \frac{s \pi}{\pi} \right) \frac{1}{k} \sum_{l=-N_-}^{N_+} e^{(1-s)kl} (e^{kl} I - \Delta_h)^{-1} z.
\]

In the following, we denote by \( S_h \) the discrete, self-adjoint solution operator defined by (12). Practical aspects of the numerical implementation of this method are discussed in Sect. 4.

For this numerical approach discretization error estimates are provided in [12]. We follow the presentation from [10, Section 4.1]. In Theorem 2 the parameter \( \alpha \in (0, 1) \) denotes the so-called pick-up regularity, see [10, Definition 4.1]. If \( \Omega \) is a Lipschitz domain \( \alpha = \frac{1}{2} + \varepsilon \) is satisfied, if \( \Omega \) is additionally convex we have \( \alpha = 1 \).

**Theorem 2** (Finite element approximation, [10, Theorem 4.2]) Given \( r \in [0, 1] \) with \( r \leq 2s \), set \( \gamma := \max (r + 2\alpha_\ast - 2s, 0) \) and \( \alpha_\ast := \frac{1}{2}(\alpha + \min (1-r, \alpha)) \) with \( \alpha \in (0, 1] \). If \( z \in H^{\gamma}(\Omega) \) for \( \delta \geq \gamma \), then

\[
\| u - u_h \|_{H^r(\Omega)} \leq C_h h^{2\alpha_\ast} \| z \|_{H^{\gamma}(\Omega)},
\]

where \( C_h \leq c \log (2/h) \) if \( \delta = \gamma \) and \( r + 2\alpha_\ast \geq 2s \), and \( C_h \leq c \) otherwise.

Assumption 1 on the quality of \( \Omega \) leads to the following estimates in Theorem 2.

**Corollary 1** For \( z \in H^{\delta+\varepsilon'}(\Omega) \) with \( \varepsilon' > 0 \) arbitrary small and \( \delta \geq -\varepsilon' \), there holds

\[
\| u - u_h \|_{L^2(\Omega)} \leq c h^\min(2, \delta + 2s) \| z \|_{H^{\delta+\varepsilon'}(\Omega)},
\]

\[
\| u - u_h \|_{H^r(\Omega)} \leq c h^\min(2-r, \delta + s) \| z \|_{H^{\delta+\varepsilon'}(\Omega)}.
\]

The quadrature formula (12) possesses the following approximation property.

**Theorem 3** (Sinc quadrature approximation, [11, Theorem 4.3]) For \( r \in [0, 1] \) and \( z \in H^r(\Omega) \) there holds

\[
\| u_h - u_h^k \|_{H^r(\Omega)} \leq c e^{-\pi^2/(2k)} \| z \|_{H^\max(0, r-2s+\varepsilon)(\Omega)} \leq c e^{-\pi^2/(2k)} \| z \|_{H^r(\Omega)}.
\]
Now, if we choose $k$ appropriately, we can balance the sinc quadrature and the finite element errors.

**Corollary 2** Assume that the number of integration points in Theorem 3 is balanced with the FE errors in Theorem 2, i.e., $k \in \mathcal{O}(\ln h^{-1})$. For $z \in H^{d+\varepsilon'}(\Omega)$ with $\varepsilon' > 0$ and $\delta \geq -\varepsilon'$ we obtain

$$
\begin{align*}
\| u - u_h^k \|_{L^2(\Omega)} &\leq c h^{\min(2, d+2s)} \| z \|_{H^{d+\varepsilon'}(\Omega)}, \\
\| u - u_h^k \|_{H^s(\Omega)} &\leq c h^{\min(2-d, d+s)} \| z \|_{H^{d+\varepsilon'}(\Omega)}.
\end{align*}
$$

(14)

Given the regularity results of Lemma 4 together with Assumption 1 we conclude the following error estimates.

**Corollary 3** For $z \in H^{3/2-\varepsilon}(\Omega)$ there holds

$$
\begin{align*}
\| u - u_h^k \|_{L^2(\Omega)} &\leq c h^{\min(2, 3/2+2s-\varepsilon')} \| z \|_{H^{3/2-\varepsilon}(\Omega)} \\
\| u - u_h^k \|_{H^s(\Omega)} &\leq c h^{\min(2-s, 3/2+s-\varepsilon')} \| z \|_{H^{3/2-\varepsilon}(\Omega)}
\end{align*}
$$

(14)

with $\varepsilon' > 0$ and $\varepsilon > 0$ arbitrary small and $\varepsilon < \varepsilon'$. Note, that the approximation $u_h^k$ converges quadratically in the $L^2(\Omega)$-norm, provided that $s > 1/4$.

In the following, we drop the superscript $k$ and write $z_h$, $u_h$, and $p_h$ for the discrete approximations of $z$, $u$, and $p$.

Before we proceed with deriving error estimates for the discrete approximation of the optimal control problem in Sects. 3.2 and 3.3, we state some general stability results for the solution operator $S$ and the discrete solution operator $S_h$.

**Lemma 5** Let $v \in L^2(\Omega)$ be given. Then the following stability estimates hold

$$
\begin{align*}
\| Sv \|_{H^s(\Omega)} &\leq c \| v \|_{L^2(\Omega)}, & \| Sv \|_{L^2(\Omega)} &\leq c \| v \|_{L^2(\Omega)}, \\
\| S_h v \|_{H^s(\Omega)} &\leq c \| v \|_{L^2(\Omega)}, & \| S_h v \|_{L^2(\Omega)} &\leq c \| v \|_{L^2(\Omega)}.
\end{align*}
$$

(15) (16)

**Proof** The first estimate of (15) directly follows from Lax–Milgram’s theorem, while the second is a direct consequence thereof

$$
\| Sv \|_{L^2(\Omega)} \leq c \| Sv \|_{H^s(\Omega)} \leq c \| v \|_{L^2(\Omega)}.
$$

To prove the first estimate of (16) we introduce the intermediate function $Sv$ to obtain

$$
\| S_h v \|_{H^s(\Omega)} \leq \| (S_h - S) v \|_{H^s(\Omega)} + \| Sv \|_{H^s(\Omega)},
$$

and apply the a priori error estimate (14) with $\delta = -\varepsilon'$ as well as the first estimate in (15). The proof of the second estimate of (16) follows the same path.  

\[ \square \]
3.2 Variational discretization

We define the variational discretization of the optimal control problem (2)–(4) as finding \( \bar{z}_h \in Z_{ad} \) such that

\[
\bar{z}_h := \arg \min_{z \in Z_{ad}} J_h(z) = \arg \min_{z \in Z_{ad}} \frac{1}{2} \| S_h z - u_d \|_{L^2(\Omega)}^2 + \frac{\mu}{2} \| z \|_{L^2(\Omega)}^2. \tag{17}
\]

Similarly, as in the continuous setting, the discrete optimal control problem (17) has a unique solution \( \bar{z}_h \in Z_{ad} \). We denote by \( \bar{u}_h := S_h \bar{z}_h \) the optimal discrete state and by \( \bar{p}_h := S_h(S_h \bar{z}_h - u_d) \) the optimal discrete adjoint state. Here \( S_h \) denotes the discrete solution operator defined in Sect. 3.1. In this case the variational inequality reads as

\[
(\bar{p}_h + \mu \bar{z}_h, z - \bar{z}_h)_{L^2(\Omega)} \geq 0 \quad \forall z \in Z_{ad}, \tag{18}
\]

which implies

\[
\bar{z}_h = \text{proj}_{[a,b]} \left( -\frac{1}{\mu} \bar{p}_h \right). \tag{19}
\]

**Theorem 4** Let the pairs \((\bar{u}(\bar{z}), \bar{z})\) and \((\bar{u}_h(\bar{z}_h), \bar{z}_h)\) be the solutions to problems (2) and (17), respectively. Then the estimates

\[
\| \bar{z} - \bar{z}_h \|_{L^2(\Omega)} \leq c h^{\min(2,3/2+2s-\varepsilon')} \left( \| \bar{z} \|_{H^{3/2-\varepsilon}(\Omega)} + \| u_d \|_{H^{3/2}(\Omega)} \right), \tag{20}
\]

\[
\| \bar{u} - \bar{u}_h \|_{L^2(\Omega)} \leq c h^{\min(2,3/2+2s-\varepsilon')} \left( \| \bar{z} \|_{H^{3/2-\varepsilon}(\Omega)} + \| u_d \|_{H^{3/2}(\Omega)} \right), \tag{21}
\]

\[
\| \bar{u} - \bar{u}_h \|_{H^s(\Omega)} \leq c h^{\min(2-s,3/2+s-\varepsilon')} \left( \| \bar{z} \|_{H^{3/2-\varepsilon}(\Omega)} + \| u_d \|_{H^{3/2}(\Omega)} \right) \tag{22}
\]

hold, provided \( \varepsilon < \varepsilon' \).

**Proof** We begin by showing the first estimate. The proof is similar to the proof of [3, Theorem 5.10] based on ideas introduced in [27]. Testing variational inequalities (9) and (18) with \( \bar{z}_h \in Z_{ad} \) and \( \bar{z} \in Z_{ad} \), respectively, and adding both expressions, we arrive at

\[
\mu \| \bar{z} - \bar{z}_h \|_{L^2(\Omega)}^2 \leq (\bar{p} - \bar{p}_h, \bar{z}_h - \bar{z}) \leq ((S - S_h)S\bar{z}, \bar{z}_h - \bar{z}) + (S_h(S - S_h)\bar{z}, \bar{z}_h - \bar{z}) + ((S_h - S)u_d, \bar{z}_h - \bar{z}) + (S_h^2(\bar{z} - \bar{z}_h), \bar{z}_h - \bar{z}). \tag{23}
\]

The first two terms can be estimated using the Cauchy–Schwarz inequality, Lemma 5 and the a priori estimate (14)

\[
((S - S_h)S\bar{z}, \bar{z}_h - \bar{z}) \leq c h^{\min(2,3/2+2s-\varepsilon')} \| \bar{z} \|_{H^{3/2-\varepsilon}(\Omega)} \| \bar{z} - \bar{z}_h \|_{L^2(\Omega)}, \tag{24}
\]

\[
(S_h(S - S_h)\bar{z}, \bar{z}_h - \bar{z}) \leq c h^{\min(2,3/2+2s-\varepsilon')} \| \bar{z} \|_{H^{3/2-\varepsilon}(\Omega)} \| \bar{z} - \bar{z}_h \|_{L^2(\Omega)}. \tag{25}
\]
The estimate of the third term follows from the Cauchy–Schwarz inequality and the estimate (14)

\[(S_h - S) u_d, \tilde{z}_h - \bar{z}\] \[\leq c h^{\min(2, 3/2 + 2s - \epsilon')} \| u_d \|_{H^{3/2}(\Omega)} \| \tilde{z} - \bar{z}_h \|_{L^2(\Omega)},\] (26)
and the last term is non-positive, since \(S_h\) is self-adjoint and therefore

\[(S_h^2 (\bar{z} - \bar{z}_h), \bar{z}_h - \bar{z}) \leq -\| S_h(\bar{z} - \bar{z}_h) \|^2_{L^2(\Omega)} \leq 0.\] (27)

The desired estimate follows from estimates (24)–(27).

Application of Lemma 2, Lemma 5 and (20) leads to

\[\| \bar{u} - \bar{u}_h \|_{L^2(\Omega)} \leq \| S\bar{z} - S_h \bar{z} \|_{L^2(\Omega)} + \| S_h \bar{z} - S_h \bar{z}_h \|_{L^2(\Omega)} \leq c h^{\min(2, 3/2 + 2s - \epsilon')} \left( \| \bar{z} \|_{H^{3/2-\epsilon}(\Omega)} + \| u_d \|_{H^{3/2}(\Omega)} \right),\]

and this proves (21). The proof of (22) follows the same path. \(\square\)

### 3.3 A fully discrete scheme

In this section we consider a fully discrete scheme for the optimal control problem (2)–(4). We discretize the set of admissible controls with piecewise constant functions

\[Z_h := \{ z_h \in L^\infty(\Omega) : z_h|_T \in \mathcal{P}_0 \text{ for all } T \in \mathcal{T}_h \}, \quad \text{and } Z_{ad}^h := Z_h \cap Z_{ad}.\]

The discretized optimal control problem reads as: find \(\bar{z}_h \in Z_{ad}^h\) such that

\[\bar{z}_h = \arg \min_{z_h \in Z_{ad}^h} J_h(z_h) = \arg \min_{z_h \in Z_{ad}^h} \frac{1}{2} \| S_h z_h - u_d \|_{L^2(\Omega)}^2 + \frac{\mu}{2} \| z_h \|_{L^2(\Omega)}^2.\] (28)

Using the same argumentation as in the continuous case, it can be shown that the optimal control problem (28) has a unique solution \(\bar{z}_h \in Z_{ad}^h\). Let \(\bar{u}_h = S_h \bar{z}_h\) and \(\bar{p}_h = S_h(\bar{z}_h - u_d)\) be the optimal discrete state and optimal discrete adjoint state, respectively, associated with \(\bar{z}_h\). Then the discrete optimality condition reads as

\[(\bar{p}_h + \mu \bar{z}_h, z_h - \bar{z}_h)_{L^2(\Omega)} \geq 0 \quad \forall z_h \in Z_{ad}^h.\] (29)

Before we state the main result of this section, we define the \(L^2(\Omega)\)-projection operator \(Q_h : L^2(\Omega) \rightarrow Z_h\) by

\[\int_{\Omega} (z - Q_h z) v_h = 0 \quad \forall v_h \in Z_h,\]

which has the following properties

(L1) \(\| Q_h v \|_{L^2(\Omega)} \leq c \| v \|_{L^2(\Omega)} \quad \forall v \in L^2(\Omega),\)
(L2) \( \|v - Q_h v\|_{L^2(\Omega)} \leq c \ h \|v\|_{H^1(\Omega)} \forall v \in H^1(\Omega) \).

**Theorem 5** Let the pairs \( (\bar{u}(\bar{z}), \bar{z}) \) and \( (\bar{u}_h(\bar{z}_h), \bar{z}_h) \) be the solutions to problems (2) and (28), respectively. Then the estimates

\[
\|\bar{z} - \bar{z}_h\|_{L^2(\Omega)} \leq c \ h \left( \|\bar{z}\|_{H^1(\Omega)} + \|u_d\|_{H^2(\Omega)} \right), \tag{30}
\]

\[
\|\bar{u} - \bar{u}_h\|_{H^1(\Omega)} \leq c \ h \left( \|\bar{z}\|_{H^1(\Omega)} + \|u_d\|_{H^2(\Omega)} \right), \tag{31}
\]

\[
\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} \leq c \ h \left( \|\bar{z}\|_{H^1(\Omega)} + \|u_d\|_{H^2(\Omega)} \right). \tag{32}
\]

hold.

**Proof** The proof is similar to the proof of [3, Theorem 5.16]. First, we use \( z = \bar{z}_h \in Z_{ad} \) in the continuous optimality condition (9) to get

\[
(\bar{p} + \mu \bar{z}, \bar{z}_h - \bar{z}) \geq 0. 
\]

Second, using \( z_h = Q_h \bar{z} \in Z_{ad} \) in the discrete optimality condition (29) and introducing \( \bar{z} \), we arrive at

\[
(\bar{p}_h + \mu \bar{z}_h, Q_h \bar{z} - \bar{z}) + (\bar{p}_h + \mu \bar{z}_h, \bar{z} - \bar{z}_h) \geq 0.
\]

Consequently, adding the previous two inequalities together we get

\[
(\bar{p} - \bar{p}_h + \mu (\bar{z} - \bar{z}_h), \bar{z}_h - \bar{z}) + (\bar{p}_h + \mu \bar{z}_h, Q_h \bar{z} - \bar{z}) \geq 0.
\]

Hence, we can conclude

\[
\mu \|\bar{z} - \bar{z}_h\|^2_{L^2(\Omega)} \leq (\bar{p} - \bar{p}_h, \bar{z}_h - \bar{z}) + (\bar{p}_h + \mu \bar{z}_h, Q_h \bar{z} - \bar{z}). \tag{33}
\]

The estimate for the first term on the right hand side of (33) is obtained in the same way as the estimate for (23). However, here we use reduced regularity for \( \bar{z} \) and \( u_d \) in the application of (14) because the final rate is dictated by the second term in (33) that only leads to a reduced rate. In the application of (14) we use \( \delta = \max(-\varepsilon, 1 - 2s) \). Following (23) and the subsequent estimates, we state the estimates

\[
(S - S_h)\bar{z}, \bar{z}_h - \bar{z}) \leq ch \|\bar{z}\|_{H^2(\Omega)} \|\bar{z}_h - \bar{z}\|_{L^2(\Omega)},
\]

\[
(S_h - S)\bar{z}, \bar{z}_h - \bar{z}) \leq ch \|\bar{z}\|_{H^2(\Omega)} \|\bar{z}_h - \bar{z}\|_{L^2(\Omega)},
\]

\[
((S - S_h)u_d, \bar{z}_h - \bar{z}) \leq ch \|u_d\|_{H^2(\Omega)} \|\bar{z}_h - \bar{z}\|_{L^2(\Omega)},
\]

which leads to the first estimate

\[
(\bar{p} - \bar{p}_h, \bar{z}_h - \bar{z}) \leq ch \left( \|\bar{z}\|_{H^2(\Omega)} + \|u_d\|_{H^2(\Omega)} \right) \|\bar{z}_h - \bar{z}\|_{L^2(\Omega)}. \tag{34}
\]
To estimate the second term in (33) we add and subtract \( \bar{p} \) and \( \mu \bar{z} \) and get
\[
(\bar{p}_h + \mu \bar{z}_h, Q_h \bar{z} - \bar{z}) = (\bar{p} + \mu \bar{z}, Q_h \bar{z} - \bar{z}) + \mu (\bar{z}_h - \bar{z}, Q_h \bar{z} - \bar{z}) + (\bar{p}_h - \bar{p}, Q_h \bar{z} - \bar{z}).
\]
(35)

To estimate the first term on the right hand side of (35) we use the definition of the operator \( Q_h \) and obtain
\[
(\bar{p} + \mu \bar{z}, Q_h \bar{z} - \bar{z}) = (\bar{p} + \mu \bar{z} - Q_h (\bar{p} + \mu \bar{z}), Q_h \bar{z} - \bar{z}) \leq c h^2 \| \bar{p} + \mu \bar{z} \|_{H^1(\Omega)} \| \bar{z} \|_{H^1(\Omega)}.
\]
where the last inequality follows from property (L2) of the \( L^2 \)-projection. The application of the Cauchy–Schwarz inequality yields the desired estimate of the second term
\[
\mu (\bar{z}_h - \bar{z}, Q_h \bar{z} - \bar{z}) \leq c h \| \bar{z} - \bar{z}_h \|_{L^2(\Omega)} \| \bar{z} \|_{H^1(\Omega)}.
\]
(36)
The estimate of the third term can be shown analogous to (34) with an application of (L2) and yields
\[
(\bar{p}_h - \bar{p}, Q_h \bar{z} - \bar{z}) \leq c h \left( \| \bar{z} \|_{H^{\max}(0,1-2s)}(\Omega) + \| u_d \|_{H^{\max}(0,1-2s)}(\Omega) \right) \| Q_h \bar{z} - \bar{z} \|_{L^2(\Omega)}
\]
\[
\leq c h^2 \left( \| \bar{z} \|_{H^{\max}(0,1-2s)}(\Omega) + \| u_d \|_{H^{\max}(0,1-2s)}(\Omega) \right) \| \bar{z} \|_{H^1(\Omega)}
\]
(37)

Estimates (34)–(37) together with an appropriate application of Hölder’s inequality and Young’s inequality yield the desired estimate (30).

In order to prove estimate (31) we proceed as follows. Introducing intermediate functions, applying the triangle inequality and using the stability results from Lemma 5 yields
\[
\| \bar{u} - \bar{u}_h \|_{H^s(\Omega)} \leq \| (S - S_h) \bar{z} \|_{H^s(\Omega)} + \| S_h (\bar{z} - \bar{z}_h) \|_{H^s(\Omega)}
\]
\[
\leq \| (S - S_h) \bar{z} \|_{H^s(\Omega)} + c \| \bar{z} - \bar{z}_h \|_{L^2(\Omega)}.
\]
(38)

Hence an application of the a priori error estimate (14) with \( \delta = 1 - \epsilon' \) and estimate (30) proves (31). The third estimate is obtained in the same way. \( \square \)

**Remark 2** We see, that the rates, that are obtained by Theorem 5 are not optimal with respect to the state and that they are dictated by the optimal linear rate, that we obtain for the control. In Sect. 5 we numerically measure higher rates for the optimal state, namely the same as for the variational discretization. The proof of higher rates will be addressed in future work and is mainly based on supercloseness results [30] for the control. These convergence rates carry over to the rates for the discrete control computed by the so-called post-processing step, i.e. using the projection formula (19) to obtain a new, piecewise linear approximation of the control. Numerical experiments for this post-processing approach are also contained in Sect. 5. The theoretical analysis is left for future work.
4 Implementation

In this section, we introduce a solver for the finite element approximation (11). The use of the Balakrishnan formula for inverting the fractional operator leads to the necessity of solving a large number of independent linear systems of equations to obtain an accurate solution. However, these systems carry a lot of structure that can be used to design efficient iterative schemes based on tailored Krylov subspace methods.

Following [10], application of the sinc quadrature to the Balakrishnan representation (5) gives rise to the discretization of the state equation (3). For convenience, we repeat the resulting approximation here.

\[ u^k_h = \frac{\sin (s\pi)}{\pi} k \sum_{l=-N_-}^{N_+} e^{(1-s)kl} v^l_h \]  

(39)

where \( v^l_h \in \mathbb{U}(\mathcal{T}_h) \) is the unique solution of the Galerkin variational problem

\[ \int_{\Omega} \nabla v^l_h \cdot \nabla w^l_h \, dx + e^{kl} \int_{\Omega} v^l_h w^l_h \, dx = \int_{\Omega} z w^l_h \, dx \quad \forall w^l_h \in \mathbb{U}(\mathcal{T}_h). \]  

(40)

The evaluation of (39) requires the solution of \( N_+ + N_- + 1 \) linear systems of the form

\[ (A + \alpha^l M) v^l = Z, \quad -N_- \leq l \leq N_. \]  

(41)

Here, \( \alpha^l = e^{kl} \) and \( A, M \) denote respectively the corresponding stiffness and mass matrices of the system, \( Z \) denotes the load vector, while \( v^l \) denotes the node vector for \( v^l_h \). Notice that the \( N_- + N_+ + 1 \) linear systems in (41) are independent for different values of \( l \). Hence, a first approach for solving systems (41) might be the use of massive parallelization. However, we shall follow a more efficient approach that exploits the structure of the linear systems and uses tailored conjugated gradients solvers.

We start by normalizing the systems. Application of a standard mass-lumping strategy results in a diagonal mass matrix \( M_h \). Then we define \( \rho := \|M_h^{-1/2} AM_h^{-1/2}\|_\infty \), \( \tilde{A} = \frac{1}{\rho} M_h^{-1/2} AM_h^{-1/2}, \tilde{\alpha}^l = \frac{1}{\rho} \alpha^l \), \( \tilde{V}^l = M_h^{1/2} V^l \) and \( \tilde{Z} = \frac{1}{\rho} M_h^{-1/2} Z \). Hence the linear systems (41) can be reformulated as

\[ (\tilde{A} + \tilde{\alpha}^l I) \tilde{V}^l = \tilde{Z}, \quad -N_- \leq l \leq N_. \]  

(42)

We can estimate the 2-condition number of system \( l \) in (42) by

\[ \kappa \left( \tilde{A} + \tilde{\alpha}^l I \right) = \frac{\lambda_{\max}(\tilde{A} + \tilde{\alpha}^l I)}{\lambda_{\min}(\tilde{A} + \tilde{\alpha}^l I)} = \frac{\lambda_{\max}(\tilde{A}) + \tilde{\alpha}^l}{\lambda_{\min}(\tilde{A}) + \tilde{\alpha}^l} \leq 1 + \min \left( \frac{\lambda_{\max}(\tilde{A})}{\tilde{\alpha}^l}, \kappa(\tilde{A}) \right) \]  

(43)

where \( \lambda_{\max}(\tilde{A}) \) and \( \lambda_{\min}(\tilde{A}) \) denote the largest and smallest eigenvalue of the symmetric positive definite matrix \( \tilde{A} \), respectively. From (43) we observe, that for small
the condition number of $\tilde{A} + \tilde{\alpha} I$ is close to the condition number of $\tilde{A}$, which is a scaled stiffness matrix, while for large $\tilde{\alpha}$ the condition number converges to 1. By introducing the scaling with $\rho$, we fix $\lambda_{\text{max}}(\tilde{A}) \leq 1$.

Thus for $l$ decreasing from $N_+$ to $N_-$ the condition number of the linear system $\tilde{A} + \tilde{\alpha} I$ is increasing. While for $l \equiv N_+$ conjugated gradients without preconditioning is a well suited solver, for $l \equiv N_-$ preconditioning in general is required. Due to this observation, we consider two adapted linear solvers.

- Linear problems, for which $l$ is sufficiently large, are considered to be well-conditioned, and no further preconditioning is needed to obtain fast convergence of the conjugate gradient solver. Thanks to the shift-invariance property of Krylov subspace methods the Krylov spaces that are generated during the conjugate gradients method are independent of $l$. As the build-up of the Krylov space contains the only matrix-vector multiplication in the conjugated gradients method, the dimension of the space is equal to the number of matrix-vector multiplications. We fix a number $N_{\text{max}}$ of multiplications and proceed as follows. Starting with $l \equiv N_+$ we solve linear systems for decreasing $l$, where we reuse the Krylov spaces from previous solutions. We stop at $l \equiv N_0$ as soon as the required Krylov space has reached the dimension of $N_{\text{max}}$.

For the implementation, we use a variant of the conjugate gradient method proposed in [24]. In Algorithm 1 we summarize the pseudo code.

- For the resulting systems $N_-, \ldots, N_0$ preconditioning is necessary which conflicts with the shift invariant property of Krylov methods. Note that subsequent systems are still similar, and thus, that a preconditioner for system $l$ is also a (worse, but not necessarily bad) preconditioner for system $l + 1$. Therefore, we use the standard approach for solving system (42) sequentially and recalculating a new preconditioner whenever the old one is no longer good enough, i.e. as soon as a given maximum number of iterations is exceeded in a conjugated gradients method, see Algorithm 2.

In Sect. 5.1 we report on the behaviour of the proposed solver.

\begin{algorithm}
\textbf{Input:} $\tilde{A}, \tilde{\alpha}, \tilde{Z}, N_{\text{max}}$
\textbf{Data:} Set: $\mathcal{K} = \emptyset, l = N_+$
\textbf{Output:} $N_0$
\begin{algorithmic}
\State \textbf{while} $\dim(\mathcal{K}) < N_{\text{max}}$ \textbf{do}
\State \hspace{1em} \textbf{for} $k = 1, \ldots$ \textbf{do} // cg-iteration for system $l$
\State \hspace{2em} \textbf{if} $k > \dim(\mathcal{K})$ \textbf{then}
\State \hspace{3em} Calculate basis vector for $k$-th Krylov space and store into $\mathcal{K}$;
\State \hspace{2em} \textbf{end}
\State \hspace{1em} Solve linear system $l$ in space $\mathcal{K}_k$ using [24, Alg. 4];
\State \hspace{1em} $l := l - 1$;
\State \textbf{end}
\State $N_0 := l$\end{algorithmic}
\textbf{Algorithm 1:} Pseudo code for solving the well-conditioned systems. The only matrix-vector multiplication appears in line 4. Here the space $\mathcal{K}_k$ in line 6 is the span over the first $k$ basis elements of $\mathcal{K}$.
Input: $\tilde{A}, \tilde{\alpha}, \tilde{Z}, N_0$

Data: Set: $N_{\text{max}} > 0$, $N_{\text{iter}} = N_{\text{max}} + 1$,

1. for $l = -N_{\text{max}} \ldots N_0$ do
2.   if $N_{\text{iter}} > N_{\text{max}}$ then
3.     Build up amg preconditioner $P$ for system $l$;
4.   end
5. Solve systems $l$ with preconditioned conjugate gradients method using preconditioner $P$ with $N_{\text{iter}}$ iterations;
6. end

Algorithm 2: Pseudo code for solving the not well-conditioned systems

Remark 3 (An alternative solver) In [19] a conjugate gradients method is proposed that uses Krylov spaces generated for one of the linear systems (42), called seed system, to generate good initial values, or even solutions, for the other systems. Thanks to the particular structure of the systems (42) this can be done without additional matrix-vector multiplications, see [19, Sec. 3.1]. Upon convergence, another system is chosen as a seed system, for which the Krylov spaces are generated. In our implementation, we combine this approach with algebraic multigrid (amg) preconditioning, and choose that system as the next seed system, that currently has the largest residuum.

This approach requires storing the solution to all systems in memory to apply the Krylov spaces and to find the next seed system. Unfortunately, this turned out to be not feasible for fine meshes in 3D, but we obtained very fast convergence of the method, when applicable.

A combination of the proposed sequential solver as in Algorithm 2 and the solver proposed in [19] seems possible (at least with some restrictions) and will be subject to future work.

Finally we note that a large number of tailored Krylov methods is proposed to deal with shifted systems that require preconditioning and we only refer to [9,23,39,40,44].

5 Numerical results

In this section, we validate the theoretical rates of convergence derived in Sect. 3. In Sect. 5.1, we investigate the solver for the fractional Laplace proposed in Sect. 4. In Sect. 5.2, we present the convergence rates of the fully discrete finite element scheme for the approximation of the optimal control problem. All numerical experiments are conducted for a range of values of the fractional exponent $s$.

We implement the solver proposed in Sect. 4 in C++ using the PETSc linear algebra package [7] and solve the optimal control problem using the TAO package of PETSc using the bound-constrained limited-memory variable-metric method (tao_blmvm), which is a limited memory BFGS method. We generate meshes, finite element functions and assemble matrices using FEniCS [28] through the C++ interface.

For the solver of the fractional operator proposed in Sect. 4 we fix $N_{\text{max}} = 500$ for 2D simulations and $N_{\text{max}} = 250$ for 3D simulations. The individual linear systems are solved up to a relative accuracy of $10^{-8}$. For Algorithm 2, we calculate a new preconditioner as soon as more than $N_{\text{iter}} = 20$ iterations are taken in the preconditioned
Table 1  The behavior of the proposed solver for the fractional operator for 2D simulation

| $N_\Omega$ | $N_\alpha(s = 0.05)$ | $N_\alpha(s = 0.5)$ | $N_\alpha(s = 0.95)$ | #amg setup |
|------------|---------------------|---------------------|---------------------|------------|
| 25         | 58 (58/0)           | 13 (13/0)           | 58 (58/0)           | 0          |
| 81         | 158 (158/0)         | 31 (31/0)           | 158 (158/0)         | 0          |
| 289        | 308 (308/0)         | 61 (61/0)           | 308 (308/0)         | 0          |
| 1089       | 508 (508/0)         | 99 (99/0)           | 508 (508/0)         | 0          |
| 4225       | 757 (757/0)         | 145 (145/0)         | 757 (757/0)         | 0          |
| 16,641     | 1056 (1056/0)       | 203 (203/0)         | 1056 (1056/0)       | 0          |
| 66,049     | 1406 (1406/0)       | 269 (269/0)         | 1406 (1406/0)       | 0          |
| 263,169    | 1806 (1677/129)     | 345 (135/210)       | 1806 (54/1752)      | 1          |
| 1,050,625  | 2254 (2089/165)     | 429 (163/266)       | 2254 (62/2192)      | 2          |
| 4,198,401  | 2753 (2548/205)     | 525 (196/329)       | 2753 (72/2681)      | 2          |

$N_\Omega$ denotes the number of degrees of freedom in $\Omega$, $N_\alpha = N_- + N_+ + 1$ denotes the number of linear systems to solve. In brackets, we show how many systems are solved by Algorithms 1 and 2 respectively. We give results for $s = 0.05$, $s = 0.5$, and $s = 0.95$. The number of amg setups in Algorithm 2 are equal in all cases.

conjugate gradients method. As amg preconditioner we use 2 V-cycles of Hypre [22] that is accessed through the PETSc interface. We stop the optimization as soon as the $l^2$-norm of the projected gradient is smaller or equal to $10^{-5} \sqrt{h_n}$, where $h$ is the length of the longest edge in the finite element mesh. Here the scaling with $h$ mimics the different scaling of the $l^2$-norm and the $L^2(\Omega)$-norm.

5.1 The solver for the fractional operator

Let us first report on the performance of the proposed solver for the systems (42). As a test example we use $\Omega = (0, 1)^n, n \in \{2, 3\}$ and set $f = \min(0.25, f_0)$, where $f_0(m) = 0.5$, with $m$ the center of $\Omega$, $f_0|_{\partial\Omega} = 0$ and $f_0$ is linearly interpolated between these values. Note that no analytical solution is known for this right-hand side $f$, and that $f$ enjoys $H^{3/2 - \epsilon}(\Omega)$ regularity, which is the maximal regularity of the optimal control $z$. We solve the equation $(-\Delta)^s u_h^k = f$ on a sequence of homogeneously refined meshes and use the solution on the finest mesh (with $N_\Omega = 4,198,401$ nodes for $n = 2$) as the reference solution. These meshes are chosen, such that all kinks in $f$ are resolved, and the integration of $f$ is done with no numerical error.

In Tables 1 and 2 we report on the solver for the cases $n = 2$ and $n = 3$, respectively, and for $s = 0.05$ and $s = 0.5$.

We observe that in fact, the number of amg setups is very small or a set up is not even necessary, which indicates, how closely related the systems are.

Finally, for small $s$ the operator is closer to identity, and thus, more systems are well-conditioned, which can be seen by the number of systems that are solved by Algorithm 1 in comparison to the number of systems solved by Algorithm 2.

Let us briefly comment on the convergence rate from Corollary 3 for $u_h^k$. As the above defined right hand side enjoys $f \in H^{3/2 - \epsilon}(\Omega)$, we expect a an order of
Table 2  Behavior of the proposed solver for the fractional operator for 3D simulation

| $N_{\Omega}$ | $N_\alpha(s = 0.05)$ | $N_\alpha(s = 0.5)$ | $N_\alpha(s = 0.95)$ | #amg setup |
|--------------|----------------------|---------------------|----------------------|-------------|
| 125          | 38 (38/0)            | 9 (9/0)             | 38 (38/0)            | 0           |
| 729          | 124 (124/0)          | 25 (25/0)           | 124 (124/0)          | 0           |
| 4913         | 258 (258/0)          | 51 (51/0)           | 258 (258/0)          | 0           |
| 35,937       | 444 (444/0)          | 85 (85/0)           | 444 (444/0)          | 0           |
| 274,625      | 678 (678/0)          | 131 (131/0)         | 678 (678/0)          | 0           |
| 2,146,689    | 964 (895/69)         | 185 (73/112)        | 964 (30/934)         | 1           |

For an explanation of the abbreviations, see Table 1

Table 3  Experimental convergence rates for the numerical solution of the fractional Laplace equation with $f \in H^{3/2-\epsilon}(\Omega)$

| $N_{\Omega}$ | $h$ | $\eta_{L^2(\Omega)}$ | $\eta_{L^2(\Omega)}^{0.05}$ | $\eta_{L^2(\Omega)}^{0.10}$ | $\eta_{L^2(\Omega)}^{0.25}$ |
|--------------|-----|----------------------|-----------------------------|-----------------------------|-----------------------------|
| 25           | 0.3536 | 0.038866 | 0.00 | 0.033601 | 0.00 | 0.009917 | 0.00 |
| 81           | 0.1768 | 0.012021 | 1.69 | 0.009930 | 1.76 | 0.002924 | 1.76 |
| 289          | 0.0884 | 0.003728 | 1.69 | 0.002913 | 1.77 | 0.000662 | 2.14 |
| 1089         | 0.0442 | 0.001190 | 1.65 | 0.000871 | 1.74 | 0.000158 | 2.07 |
| 4225         | 0.0221 | 0.000387 | 1.62 | 0.000265 | 1.72 | 0.000038 | 2.05 |
| 16,641       | 0.0110 | 0.000127 | 1.61 | 0.000081 | 1.71 | 0.000009 | 2.09 |
| 66,049       | 0.0055 | 0.000041 | 1.62 | 0.000025 | 1.72 | 0.000002 | 2.06 |
| 263,169      | 0.0028 | 0.000013 | 1.68 | 0.000007 | 1.77 | 0.000001 | 2.10 |

We observe, that the expected rate $h^{\min(2,3/2+2s-\epsilon')}$ is fulfilled for all examples

$h^{\min(2,3/2+2s-\epsilon')}$. In Table 3 we show the observed convergence rates for $n = 2$ and $s \in \{0.05, 0.1, 0.25\}$, which indeed confirm the theoretical predictions. For $n = 3$ memory consumption restricts the quality of the reference solution, such that we do not measure a rate for $n = 3$.

5.2 Optimal control problem

To verify the theoretical convergence rates of the finite element discretization of the optimal control problem, we perform numerical experiments without a known optimal solution. We use the domain $\Omega = (0, 1)^2$ and set the desired state to be equal to an eigenfunction of the Laplacian on the square, namely $u_d = \sin(2\pi x) \sin(2\pi y)$ and $f \equiv 0$. We consider three different values of the fractional parameter, namely $s \in \{0.05, 0.25, 0.5\}$ and choose $a = -0.8, b = 0.8$, such that the box-constraints are attained in some subdomain of $\Omega$. The optimal solution for $h = 0.0014$ is considered as reference solution.

Results of the numerical tests are summarized in Figs. 1 and 2. First order convergence of the approximation of the control is obtained, which is in line with (30).
Fig. 1 Convergence rates for the discretization of the control $z$ in 2D. First order of convergence of the piecewise constant approximation can be observed. Application of the additional post-processing significantly improves the convergence properties, and we observe $h^{\min(2,3/2+2s)}$ convergence.

Fig. 2 Convergence rates for the discretization of the state $u$ in 2D. Convergence order of the piecewise linear finite element method measured in the $L^2(\Omega)$-norm depends on the choice of $s$ and varies between $3/2$ and 2, which can be attained for sufficiently large $s$. Convergence order in $H^s(\Omega)$-norm is included for completeness.
We also report on results using the post-processing approach [30]. Here the projection formula (19) is used to obtain a higher order approximation for the optimal control. Higher order means, that instead of an approximation with piecewise constant functions, a piecewise linear approximation is obtained, thus we define \( \bar{z}_{h}^{PP} := \text{proj}_{[a,b]}(-\frac{1}{h} p_{h}) \). We expect thus the same optimal rate of convergence for this post-processed optimal control \( \bar{z}_{h}^{PP} \) as for the variational discretization approach, namely \( \| z - \bar{z}_{h}^{PP} \|_{L^{2}(\Omega)} \leq c h^{\min(2,3/2+2s-\epsilon')} \). In Fig. 1 we observe the expected higher rates for the optimal controls with post-processing approach for \( n = 2 \).

We also investigate the finite element approximation of the state in the \( L^{2}(\Omega) \)- and \( H^{s}(\Omega) \)-norms, the latter being estimated using the Gagliardo–Nirenberg interpolation inequality

\[
\| \bar{u} - \bar{u}_{h} \|_{H^{s}(\Omega)} \leq c \| \bar{u} - \bar{u}_{h} \|_{L^{2}(\Omega)}^{1-s} \| \bar{u} - \bar{u}_{h} \|_{H^{1}(\Omega)}^{s}.
\]

We observe \( h^{\min(2,3/2+2s)} \) order of convergence in the \( L^{2}(\Omega) \)-norm and \( h^{\min(2-s,3/2+s)} \) in the \( H^{s}(\Omega) \)-norm. The theoretical justification as well as the analysis of the post-processing approach is left for future work.

Acknowledgements We would like to thank Harbir Antil for giving a compact course on fractional PDEs as a preparation for this work. Furthermore, we would like to thank Johannes Pfefferer for many fruitful discussions during the preparation of this work. Further we thank Johannes Pfefferer and Constantin Christof for many helpful discussions during the revision of this work. Finally, we would like to thank the anonymous reviewers for their valuable suggestions to improve this work.

References

1. Abatangelo, N., Dupaigne, L.: Nonhomogeneous boundary conditions for the spectral fractional Laplacian. Ann. l’Inst. Henri Poincare Non Linear Anal. 34(2), 439–467 (2017). https://doi.org/10.1016/j.anihpc.2016.02.001
2. Antil, H., Khatri, E.R., Warma, M.: External optimal control of nonlocal PDEs. Inverse Problems (2019, to appear). arXiv:1811.04515
3. Antil, H., Otárola, E.: A FEM for an optimal control problem of fractional powers of elliptic operators. SIAM J. Control Optim. 53(6), 3432–3456 (2015)
4. Antil, H., Pfefferer, J., Rogovs, S.: Fractional operators with inhomogeneous boundary conditions: analysis, control, and discretization. Commun. Math. Sci. 16(5), 1395–1426 (2018)
5. Antil, H., Warma, M.: Optimal control of fractional semilinear PDEs. ESAIM Control Optim. Calc. Var. (2019). https://doi.org/10.1051/cocv/2019003
6. Arada, N., Casas, E., Tröltzsch, F.: Error estimates for the numerical approximation of a semilinear elliptic control problem. Comput. Optim. Appl. 23(2), 201–229 (2002). https://doi.org/10.1023/A:1020576801966
7. Balay, S., Abhyankar, S., Adams, M.F., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Eijkhout, V., Gropp, W.D., Kaushik, D., Knepley, M.G., McInnes, L.C., Rupp, K., Smith, B.F., Zhang, H., Zhang, H.: PETSc Web page (2017). http://www.mcs.anl.gov/petsc. Accessed 10 Sept 2019
8. Banjai, L., Melenk, J.M., Nochetto, R.H., Otárola, E., Salgado, A.J., Schwab, C.: Tensor FEM for spectral fractional diffusion. Found. Comput. Math. (2018). https://doi.org/10.1007/s10208-018-9402-3
9. Benzi, M., Bertaccini, D.: Approximate inverse preconditioning for shifted linear systems. BIT Numer. Math. 43, 231–244 (2003)
10. Bonito, A., Borghagaray, J.P., Nochetto, R.H., Otárola, E., Salgado, A.J.: Numerical methods for fractional diffusion. Comput. Vis. Sci. 19(5), 19–46 (2018). https://doi.org/10.1007/s00791-018-0289-y
11. Bonito, A., Lei, W., Pasciak, J.: On Sinc quadrature approximations of fractional powers of regularly accretive operators. J. Numer. Math. (2018, to appear). arXiv:1709.06619
12. Bonito, A., Pasciak, J.: Numerical approximation of fractional powers of regularly accretive operators.IMA J. Numer. Anal. 37(3), 1245–1273 (2016)
13. Bourdaud, G., Meyer, Y.: Fonctions qui opèrent sur les espaces de Sobolev. J. Funct. Anal. 97(2), 351–360 (1991). https://doi.org/10.1016/0022-1236(91)90006-Q
14. Bourdaud, G., Sickel, W.: Composition operators on function spaces with fractional order of smoothness (harmonic analysis and nonlinear partial differential equations) (2014). https://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/187881. Accessed 10 Sept 2019
15. Cabré, X., Tan, J.: Positive solutions of nonlinear problems involving the square root of the Laplacian. Adv. Math. 224(5), 2052–2093 (2010)
16. Caffarelli, L., Silvestre, L.: An extension problem related to the fractional Laplacian. Commun. Part. Diff. Equ. 32(7–9), 1245–1260 (2007)
17. Capella, A., Dávila, J., Dupaigne, L., Sire, Y.: Regularity of radial extremal solutions for some non-local semilinear equations. Commun. Part. Differ. Equ. 36(8), 1353–1384 (2011)
18. Casas, E., Mateos, M., Tröltzsch, F.: Error estimates for the numerical approximation of boundary semilinear elliptic control problems. Comput. Optim. Appl. 31(2), 193–219 (2005). https://doi.org/10.1007/s10589-005-2180-2
19. Chan, T., Ng, M.: Galerkin projection methods for solving multiple linear systems. SIAM J. Sci. Comput. 21(3), 836–850 (1999)
20. Cusimano, N., del Teso, F., Gerardo-Giorda, L., Pagnini, G.: Discretizations of the spectral fractional laplacian on general domains with Dirichlet, Neumann, and Robin boundary conditions. SIAM J. Numer. Anal. 56(3), 1243–1272 (2018). https://doi.org/10.1137/17M1128010
21. D’Elia, M., Glusa, C., Otárola, E.: A priori error estimates for the optimal control of the integral fractional Laplacian. SIAM J. Control Optim. 57(4), 2775–2798 (2019). https://doi.org/10.1137/18M1219989
22. Falgout, R., Barker, A., Gahvari, H., Kolev, T., Li, R., Osei-Kuffuor, D., Schroder, J., Vassilevski, P., Wang, L., Yang, U.M.: Hypre webpage (2017). https://computation.llnl.gov/projects/hypre-scalable-linear-solvers-multigrid-methods. Accessed 10 Sept 2019
23. Frommer, A.: BiCGStab(l) for families of shifted linear systems. Computing 70, 87–109 (2003)
24. Frommer, A., Maass, P.: Fast CG-based methods for Tikhonov–Phillips regularization. SIAM J. Sci. Comput. 20(5), 1831–1850 (1999)
25. Glusa, C., Otárola, E.: Optimal control of a parabolic fractional PDE: analysis and discretization. (2019) arXiv e-prints arXiv:1905.10002
26. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Pitman, Boston (1985)
27. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. Comput. Optim. Appl. 30, 45–61 (2005)
28. Logg, A., Mardal, K.A., Wells, G.N., et al.: Automated Solution of Differential Equations by the Finite Element Method. Springer, Berlin (2012). https://doi.org/10.1007/978-3-642-23099-8
29. Meyn, J., Pfefferer, J., Schürholz, K., Vexler, B.: hp-finite elements for fractional diffusion. SIAM J. Numer. Anal. 56(4), 2345–2374 (2018). https://doi.org/10.1137/17M1135517
30. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. SIAM J. Control Optim. 43(3), 970–985 (2004)
31. Musina, R., Nazarov, A.I.: A note on truncations in fractional Sobolev spaces. Bull. Math. Sci. (2017). https://doi.org/10.1007/s13373-017-0107-8
32. Nazarov, A.I.: Remark on fractional Laplacians. Preprints of the St. Petersburg Mathematical Society 2016-01 (2016). http://www.mathsoc.spb.ru/preprint/2016/index.html. Accessed 10 Sept 2019
33. Nochetto, R., Otárola, E., Salgado, A.: A PDE approach to fractional diffusion in general domains: a priori error analysis. Found. Comput. Math. 15(3), 848–873 (2015)
34. Nochetto, R., Otárola, E., Salgado, A.: A PDE approach to space–time fractional parabolic problems. SIAM J. Numer. Anal. 54(2), 848–873 (2016)
35. Oswald, P.: On the boundedness of the mapping $f \rightarrow |f|$ in Besov spaces. Comment. Math. Univ. Carolin. 33(1), 57–66 (1992)
36. Otárola, Enrique: A piecewise linear FEM for an optimal control problem of fractional operators: error analysis on curved domains. ESAIM: M2AN 51(4), 1473–1500 (2017). https://doi.org/10.1051/m2an/2016065
37. Rösch, A.: Error estimates for linear–quadratic control problems with control constraints. Optim. Methods Softw. 21(1), 121–134 (2006). https://doi.org/10.1080/10556780500094945
38. Runst, T., Sickel, W.: Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations, De Gruyter Series in Nonlinear Analysis and Applications, vol. 3. de Gruyter, Berlin (2011)
39. Soodhalter, K.: Two recursive GMRES-type methods for shifted linear systems with general preconditioning. Electron. Trans. Numer. Anal. 45, 499–523 (2016)
40. Soodhalter, K., Szyld, D., Xue, F.: Krylov subspace recycling for sequences of shifted linear systems. Appl. Numer. Math. 81, 105–118 (2014)
41. Stinga, P., Torrea, J.: Extension problem and Harnack’s inequality for some fractional operators. Commun. Part. Differ. Equ. 35(11), 2092–2122 (2010)
42. Tröltzsch, F.: Optimale Steuerung partieller Differentialgleichungen: Theorie, Verfahren und Anwendungen. Vieweg, Wiesbaden (2005)
43. Yosida, K.: Functional Analysis. Springer, Berlin (1980)
44. Zhong, H., Gu, X.: A flexible and adaptive Simpler GMRES with deflated restarting for shifted linear systems. Comput. Math. Appl. 78(3), 997–1007 (2019). https://doi.org/10.1016/j.camwa.2019.03.017

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.