SIAM J. CONTROL OPTIM.
Vol. 0, No. 0, pp. 000–000

A SPACE-TIME FRACTIONAL OPTIMAL CONTROL PROBLEM: ANALYSIS AND DISCRETIZATION*

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

Abstract. We study a linear-quadratic optimal control problem involving a parabolic equation with fractional diffusion and Caputo fractional time derivative of orders $s \in (0,1)$ and $\gamma \in (0,1]$, respectively. The spatial fractional diffusion is realized as the Dirichlet-to-Neumann map for a nonuniformly elliptic operator. Thus, we consider an equivalent formulation with a quasi-stationary elliptic problem with a dynamic boundary condition as state equation. The rapid decay of the solution to this problem suggests a truncation that is suitable for numerical approximation. We consider a fully discrete scheme: piecewise constant functions for the control and, for the state, first-degree tensor product finite elements in space and a finite difference discretization in time. We show convergence of this scheme and, under additional data regularity, derive a priori error estimates for the case $s \in (0,1)$ and $\gamma = 1$.

Key words. linear-quadratic optimal control problem, fractional derivatives and integrals, fractional diffusion, weighted Sobolev spaces, finite elements, stability, fully discrete methods

AMS subject classifications. 26A33, 35J70, 49J20, 49M25, 65M12, 65M15, 65M60, 65R10

DOI. 10.1137/15M1014991

1. Introduction.

We are interested in the design and analysis of efficient solution techniques for a linear-quadratic optimal control problem involving an initial boundary value problem for a space-time fractional parabolic equation. Let Ω be an open and bounded domain in $\mathbb{R}^n$ $(n \geq 1)$, with boundary $\partial\Omega$. Given $s \in (0,1)$, $\gamma \in (0,1]$, and a desired state $\mathsf{u}_d : \Omega \times (0,T) \to \mathbb{R}$, we define

$$J(\mathsf{u},\mathsf{z}) = \frac{1}{2}\int_0^T \left(\|\mathsf{u} - \mathsf{u}_d\|^2_{L^2(\Omega)} + \mu\|\mathsf{z}\|^2_{L^2(\Omega)} \right) \mathrm{d}t, \tag{1.1}$$

where $\mu > 0$ is the so-called regularization parameter. Let $\mathsf{f} : \Omega \times (0,T) \to \mathbb{R}$ and $\mathsf{u}_0 : \Omega \to \mathbb{R}$ be fixed functions. We will call them the right-hand side and initial datum, respectively. We shall be concerned with the following optimal control problem: Find

$$\min J(\mathsf{u},\mathsf{z}) \tag{1.2}$$

subject to the *space-time fractional state equation*

$$\partial_t^\gamma \mathsf{u} + \mathcal{L}^s \mathsf{u} = \mathsf{f} + \mathsf{z} \ \text{ in } \Omega \times (0,T), \qquad \mathsf{u}(0) = \mathsf{u}_0 \ \text{ in } \Omega, \tag{1.3}$$

and the *control constraints*

$$\mathsf{a}(x',t) \leq \mathsf{z}(x',t) \leq \mathsf{b}(x',t) \quad \text{a.e. } (x',t) \in Q := \Omega \times (0,T). \tag{1.4}$$

*Received by the editors April 1, 2015; accepted for publication (in revised form) February 2, 2016; published electronically DATE.
http://www.siam.org/journals/sicon/x-x/M101499.html

†Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030 (hantil@gmu.edu). This author's research was supported in part by NSF grant DMS-1521590.

‡Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile (enrique.otarola@usm.cl). This author's research was supported in part by NSF grant DMS-1411808 and by CONICYT through FONDECYT project 3160201 and Anillo ACT1106.

§Department of Mathematics, University of Tennessee, Knoxville, TN 37996 (asalgad1@utk.edu). This author's research was supported in part by NSF grant DMS-1418784.

The functions a and b both belong to $L^2(Q)$ and satisfy the property $\mathsf{a}(x',t) \leq \mathsf{b}(x',t)$ for almost every $(x',t) \in Q$. The operator $\mathcal{L}^s$, with $s \in (0,1)$, is the fractional power of the second order elliptic operator

$$\mathcal{L}w = -\mathrm{div}_{x'}(A\nabla_{x'}w) + cw \ \text{ in } \Omega, \qquad w = 0 \ \text{ on } \partial\Omega,$$

where $0 \leq c \in L^\infty(\Omega)$ and $A \in C^{0,1}(\Omega, \mathsf{GL}(n,\mathbb{R}))$ is symmetric and positive definite.

The fractional derivative in time ∂_t^γ for $\gamma \in (0,1)$ is understood as *the left-sided Caputo fractional derivative of order* γ with respect to t, which is formally defined by

$$\partial_t^\gamma \mathsf{u}(x',t) = \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{1}{(t-r)^\gamma} \frac{\partial \mathsf{u}(x',r)}{\partial r} \, \mathrm{d}r, \tag{1.5}$$

where Γ is the Gamma function. For $\gamma = 1$, we consider the usual derivative ∂_t.

For convenience, we will refer to the optimal control problem (1.2)–(1.4) as the *space-time fractional optimal control problem*; see section 3 for its precise description and analysis. One of the main difficulties in the study of the state equation (1.3) is the nonlocality of the fractional time derivative and the fractional space operator (see [14, 15, 34, 54, 55]). A possible approach to overcome the nonlocality in space is given by the result of Caffarelli and Silvestre in $\mathbb{R}^n$ [14] and its extensions to bounded domains [15, 55]: Fractional powers of the spatial operator $\mathcal{L}$ can be realized as an operator that maps a Dirichlet boundary condition to a Neumann condition via an extension problem on the semi-infinite cylinder $\mathcal{C} = \Omega \times (0,\infty)$. Therefore, we shall use the Caffarelli–Silvestre extension to rewrite the space-time fractional state equation (1.3) as a quasi-stationary elliptic problem with a dynamic boundary condition:

$$\begin{aligned} -\,\mathrm{div}\,(y^\alpha \mathbf{A} \nabla \mathscr{U}) + y^\alpha c \mathscr{U} = 0 \ &\text{ in } \mathcal{C} \times (0,T), \qquad \mathscr{U} = 0 \ \text{ on } \partial_L \mathcal{C} \times (0,T), \\ \partial_t^\gamma \mathscr{U} + \tfrac{1}{d_s} \partial_\nu^\alpha \mathscr{U} = \mathsf{f} + \mathsf{z} \ &\text{ on } (\Omega \times \{0\}) \times (0,T), \quad \mathscr{U} = \mathsf{u}_0 \ \text{ on } \Omega \times \{0\}, \ t = 0, \end{aligned} \tag{1.6}$$

where $\partial_L \mathcal{C} = \partial\Omega \times [0,\infty)$ is the lateral boundary of $\mathcal{C}$, $\alpha = 1 - 2s \in (-1,1)$, $d_s = 2^\alpha \Gamma(1-s)/\Gamma(s)$, and the conormal exterior derivative of $\mathscr{U}$ at $\Omega \times \{0\}$ is

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

where the limit must be understood in the distributional sense [14, 15, 55]. Finally, $\mathbf{A}(x',y) = \mathrm{diag}\{A(x'),1\} \in C^{0,1}(\mathcal{C}, \mathsf{GL}(n{+}1,\mathbb{R}))$. We will call y the *extended variable*, and the dimension $n+1$ in $\mathbb{R}^{n+1}_+$ the *extended dimension* of problem (1.6). As noted in [14, 15, 55], $\mathcal{L}^s$ and the Dirichlet-to-Neumann operator of (1.6) are related by

$$d_s \mathcal{L}^s \mathsf{u} = \partial_\nu^\alpha \mathscr{U} \quad \text{in } (\Omega \times \{0\}) \times (0,T).$$

We briefly elaborate on these ideas in section 2.5. A rigorous analysis is provided in [48, 49].

Motivated by applications, the study of elliptic and parabolic problems involving fractional diffusion and fractional derivatives has received considerable attention in recent years. A rather incomplete list of problems where fractional diffusion and fractional derivatives appear includes mechanics [6], where they are used to model viscoelastic behavior [19], turbulence [16, 22], and the hereditary properties of materials [27]; diffusion processes [1, 46], in particular processes in disordered media, where the disorder may change the laws of Brownian motion and thus lead to anomalous diffusion [8, 11]; biophysics [13]; image processing [25]; nonlocal electrostatics [31];

finance [37]; chaotic dynamical systems [52]; and many others [12, 20]. It is then only natural that interest in efficient approximation schemes for these problems arises and that one might be interested in their control.

The study of solution techniques for problems involving fractional derivatives is a relatively new but rapidly growing area of research, and thus it is impossible to provide a complete overview of the available results and limitations. We restrict ourselves to referring the interested reader to [2, 10, 45, 48, 49, 51] for a survey. Concerning the analysis and discretization of optimal control problems with PDE constraints, we suggest [29, 30, 44, 59]. Mainly, these references are concerned with control problems governed by elliptic and parabolic PDEs, both linear and semilinear. The common feature here is that, in contrast to (1.2)–(1.4), the state equation is local.

The numerical analysis of optimal control problems involving evolution equations with fractional diffusion and fractional time derivative is still in its infancy. To the best of our knowledge, the first work that provides a comprehensive treatment of an optimal control problem involving fractional elliptic operators in space is [5]. Concerning fractional derivatives in time, the first work that attempts to study an optimization problem constrained by a fractional order ODE is [3] where, through completely formal calculations, the author derives optimality conditions and a numerical scheme. However, no justification is provided for either the optimality conditions or the numerical scheme. Later, similar optimization problems have been discretized via a finite element method [4], a modified Grünwald–Letnikov approach [7, 21], and a rational approximation approach [58]. However, fundamental mathematical results such as stability and convergence of the proposed numerical schemes are missing in these works. Recently, convergence of spectral based techniques has been explored in [41, 42] for an optimization problem restricted to fractional order ODEs. Optimal control problems for one dimensional evolution equations with only fractional time derivatives have been recently studied in [61, 62]. In these references, the authors rigorously derive first order necessary optimality conditions, propose numerical schemes based on spectral methods, and obtain a priori error estimates. These error estimates, however, require a rather strong regularity assumption in time, namely $\mathsf{u}_{tt} \in L^\infty(0,T)$, which is problematic [32, 38, 39]. Since $0 < \gamma < 1$, derivatives of the solution u of (1.3) with respect to t are unbounded as $t \downarrow 0$ [43, 49].

We provide a comprehensive treatment of a linear-quadratic optimal control problem involving evolution equations with fractional diffusion and fractional time derivative: $s \in (0,1)$ and $\gamma \in (0,1]$. To the best of our knowledge this is the first work addressing such a problem from a mathematical point of view. We rigorously derive optimality conditions, present a numerical scheme, and prove that the solution to the discrete problem converges to the exact solution for all values of $s \in (0,1)$ and $\gamma \in (0,1]$. For $\gamma = 1$ and under additional smoothness, we derive a priori error estimates. We overcome the nonlocality of $\mathcal{L}^s$ by using the results of Caffarelli and Silvestre [14]. We realize the state equation (1.3) by (1.6), so that our problem can be equivalently written as: Minimize J subject to the *extended state equation* (1.6) and the control constraints (1.4).

Inspired by [5, 49, 48], we propose a simple strategy for finding the solution to the space-time fractional optimal control problem (1.2)–(1.4): Given f and u_d, we realize (1.3) by (1.6) and apply standard techniques to solve this problem. We thus obtain an optimal control $\bar{\mathsf{z}} : \Omega \times (0,T) \to \mathbb{R}$ and an optimal state $\bar{\mathscr{U}} : \mathcal{C} \times (0,T) \to \mathbb{R}$. Setting $\bar{\mathsf{u}}(x',t) = \bar{\mathscr{U}}(x',0,t)$ for all $(x',t) \in \Omega \times (0,T)$, we obtain $(\bar{\mathsf{u}}, \bar{\mathsf{z}})$ that solves (1.2)–(1.4).

The outline of this paper is as follows. In section 2 we introduce notation, recall elements from fractional calculus, define fractional powers of elliptic operators via spectral theory, and show that this definition is equivalent with the Caffarelli–Silvestre extension. This allows us to study (1.6) and provide some energy estimates. On the basis of this, in section 3, we study the *space-time fractional optimal control problem*. We derive existence and uniqueness results together with first order sufficient and necessary optimality conditions. In section 4, we begin the numerical analysis of our problem. We introduce a truncation of the state equation and derive approximation properties of its solution. In section 5, we recall the fully discrete scheme of [49] that approximates the solution to the state equation (1.3). For $s \in (0,1)$ and $\gamma = 1$, we derive a novel $L^2(Q)$-error estimate in section 5.4. Section 6.1 is devoted to the design of a numerical scheme to approximate the control problem (1.2)–(1.4), and in section 6.2 we derive a priori error estimates for $s \in (0,1)$ and $\gamma = 1$. The convergence of the scheme is analyzed in section 6.3 for $s \in (0,1)$ and $\gamma \in (0,1]$. Finally, section 7 presents numerical experiments that illustrate the theory developed in section 6.2.

2. Notation and preliminaries. Let us set notation and recall some facts that will be useful later.

2.1. Notation. Throughout this work Ω is a convex polytopal subset of $\mathbb{R}^n$, $n \geq 1$, with boundary $\partial\Omega$. If $T > 0$ is a fixed time, we set $Q = \Omega \times (0,T)$. We will follow the notation of [48, 49] and define the semi-infinite cylinder with base Ω and its lateral boundary, respectively, by $\mathcal{C} = \Omega \times (0,\infty)$ and $\partial_L \mathcal{C} = \partial\Omega \times [0,\infty)$. For $\mathscr{Y} > 0$, we define the truncated cylinder $\mathcal{C}_\mathscr{Y} = \Omega \times (0,\mathscr{Y})$, and $\partial_L \mathcal{C}_\mathscr{Y}$ accordingly. Since we will be dealing with objects defined on $\mathbb{R}^n$ and $\mathbb{R}^{n+1}$, it will be convenient to distinguish the extended $(n+1)$ dimension. If $x \in \mathbb{R}^{n+1}$, we write $x = (x', y)$, with $x' \in \mathbb{R}^n$ and $y \in \mathbb{R}$.

If $\mathcal{X}$ and $\mathcal{Y}$ are normed spaces, $\mathcal{X} \hookrightarrow \mathcal{Y}$ means that $\mathcal{X}$ is continuously embedded in $\mathcal{Y}$. We denote by $\mathcal{X}'$ and $\|\cdot\|_\mathcal{X}$ the dual and norm, respectively, of $\mathcal{X}$. The relation $a \lesssim b$ indicates that $a \leq Cb$, with a nonessential constant C that might change at each occurrence.

If $D \subset \mathbb{R}^n$ is open, $n \geq 1$, and $\phi : D \times (0,T) \to \mathbb{R}$, we will regard ϕ as a function of t with values in a Banach space $\mathcal{X}$; i.e., $\phi : (0,T) \ni t \mapsto \phi(t) \equiv \phi(\cdot,t) \in \mathcal{X}$. For $1 \leq p \leq \infty$, $L^p(0,T;\mathcal{X})$ is the space of $\mathcal{X}$-valued functions whose $\mathcal{X}$-norm is in $L^p(0,T)$. This is a Banach space for the norm

$$\|\phi\|_{L^p(0,T;\mathcal{X})} = \left(\int_0^T \|\phi(t)\|_\mathcal{X}^p \, \mathrm{d}t \right)^{\frac{1}{p}}, \; 1 \leq p < \infty, \qquad \|\phi\|_{L^\infty(0,T;\mathcal{X})} = \operatorname*{esssup}_{t \in (0,T)} \|\phi(t)\|_\mathcal{X}.$$

We denote by $\mathcal{L}^s$, $s \in (0,1)$, a fractional power of the second order, symmetric, and uniformly elliptic operator $\mathcal{L}$. The parameter α belongs to $(-1,1)$ and is related to the power s of the fractional operator $\mathcal{L}^s$ by the formula $\alpha = 1 - 2s$.

In what follows we will, without explicit mention, make use of the following regularity result [28]:

$$\|w\|_{H^2(\Omega)} \lesssim \|\mathcal{L}w\|_{L^2(\Omega)} \quad \forall w \in H^2(\Omega) \cap H_0^1(\Omega). \tag{2.1}$$

2.2. Fractional derivatives and integrals. The left Caputo fractional derivative is defined in (1.5). The *right-sided Caputo fractional derivative* is [34, 54]

$$\partial_{T-t}^\gamma g(t) = -\frac{1}{\Gamma(1-\gamma)} \int_t^T \frac{g'(\xi)}{(\xi - t)^\gamma} \, \mathrm{d}\xi, \qquad \gamma \in (0,1). \tag{2.2}$$

For $g \in L^1(0,T)$ and $\sigma > 0$, the *left* and *right Riemann–Liouville fractional integrals* of order σ are, respectively, (see [54, section 2, Definition 2.1])

$$(2.3)\quad (I_t^\sigma g)(t) = \frac{1}{\Gamma(\sigma)} \int_0^t \frac{g(\xi)}{(t-\xi)^{1-\sigma}}\,\mathrm{d}\xi, \qquad (I_{T-t}^\sigma g)(t) = \frac{1}{\Gamma(\sigma)} \int_t^T \frac{g(\xi)}{(\xi-t)^{1-\sigma}}\,\mathrm{d}\xi.$$

Further, [54, sections 2.2–2.3] provides a motivation for these definitions inspired by the Abel equation.

PROPOSITION 1 (continuity of fractional integrals). *If $\sigma > 0$ and $1 \le p \le \infty$, then I_t^σ and I_{T-t}^σ are continuous from $L^p(0,T)$ into itself and*

$$\|I_t^\sigma g\|_{L^p(0,T)} \le \frac{T^\sigma}{\Gamma(\sigma+1)}\|g\|_{L^p(0,T)}, \qquad \|I_{T-t}^\sigma g\|_{L^p(0,T)} \le \frac{T^\sigma}{\Gamma(\sigma+1)}\|g\|_{L^p(0,T)}$$

for all $g \in L^p(0,T)$. These maps also are continuous from $C([0,T])$ into itself.

Proof. For the proof of continuity in $L^p(0,T)$ see [54, section 2, Theorem 2.6]. To obtain the continuity in $C([0,T])$ we use the continuity in $L^\infty(0,T)$, together with the fact that if $g \in C([0,T])$, then its fractional integrals are continuous as well. This can be easily shown by recalling that g is also uniformly continuous. ∎

We also define the *left* and *right Riemann–Liouville fractional derivatives* of order $\gamma \in (0,1)$, respectively, by (see [54, section 2.3, Definition 2.2])

$$D_t^\gamma g(t) = \frac{1}{\Gamma(1-\gamma)} \frac{\mathrm{d}}{\mathrm{d}t} \int_0^t \frac{g(\xi)}{(t-\xi)^\gamma}\,\mathrm{d}\xi, \qquad D_{T-t}^\gamma g(t) = \frac{-1}{\Gamma(1-\gamma)} \frac{\mathrm{d}}{\mathrm{d}t} \int_t^T \frac{g(\xi)}{(\xi-t)^\gamma}\,\mathrm{d}\xi.$$

A relation between the Caputo and Riemann–Liouville derivatives is given below.

LEMMA 2 (relation between fractional derivatives). *Let $\gamma \in (0,1)$ and $g \in W_1^1(0,T)$; then $D_t^\gamma g$ and $D_{T-t}^\gamma g$ exist almost everywhere on $[0,T]$. In addition, $D_t^\gamma g, D_{T-t}^\gamma g \in L^r(0,T)$ for $1 \le r < \frac{1}{\gamma}$, and*

$$(2.4)\quad D_t^\gamma g(t) = \partial_t^\gamma g(t) + \frac{1}{\Gamma(1-\gamma)} \frac{g(0)}{t^\gamma}, \quad D_{T-t}^\gamma g(t) = \partial_{T-t}^\gamma g(t) + \frac{1}{\Gamma(1-\gamma)} \frac{g(T)}{(T-t)^\gamma}.$$

Proof. See [54, section 2.3, Lemma 2.2]. ∎

We now derive an integration-by-parts formula for Caputo derivatives that will be fundamental in our analysis. For $\gamma \in (0,1)$ we define

$$\mathbb{L}_\gamma = \{f \in C([0,T]) : \partial_t^\gamma f \in L^2(0,T)\}, \qquad \mathbb{R}_\gamma = \{g \in C([0,T]) : \partial_{T-t}^\gamma g \in L^2(0,T)\}.$$

LEMMA 3 (fractional integration-by-parts formula). *If $f \in \mathbb{L}_\gamma$ and $g \in \mathbb{R}_\gamma$, then the following fractional integration by parts holds:*

$$(2.5)\quad \int_0^T \partial_t^\gamma f(t) g(t)\,\mathrm{d}t + f(0)(I_{T-t}^{1-\gamma} g)(0) = \int_0^T f(t)\partial_{T-t}^\gamma g(t)\,\mathrm{d}t + g(T)(I_t^{1-\gamma} f)(T).$$

Proof. If f and g are smooth, recall that [54, section 2.6, Corollary 2]

$$\int_0^T D_t^\gamma f(t) g(t)\,\mathrm{d}t = \int_0^T f(t) D_{T-t}^\gamma g(t)\,\mathrm{d}t;$$

(2.5) now follows from (2.4) and (2.3), and the point values are well-defined since $f \in C([0,T])$ implies $I_t^\gamma f, I_{T-t}^\gamma f \in C([0,T])$; see Proposition 1. Conclude by density. ∎

It is important to note that there is another definition, not completely equivalent, of fractional derivatives: the so-called Grünwald–Letnikov derivative [34]. Among all possible definitions of fractional derivatives, we adopt the left-sided Caputo fractional derivative as ∂_t^γ in problem (1.3): The Caputo approach leads to an initial condition of the form $\mathsf{u} = \mathsf{u}_0$, which is *physically meaningful.* The Riemann–Liouville approach leads to initial conditions containing the limit values of the Riemann–Liouville fractional integrals at $t = 0$, something that does not have a clear physical meaning.

2.3. Fractional powers of second order elliptic operators. Spectral theory for the operator $\mathcal{L}$ yields the existence of $\{(\lambda_k, \varphi_k)\}_{k\in\mathbb{N}} \subset \mathbb{R}_+ \times H_0^1(\Omega)$ such that

$$\mathcal{L}\varphi_k = \lambda_k \varphi_k \;\text{ in } \Omega, \qquad \varphi_k = 0 \;\text{ on } \partial\Omega, \qquad k \in \mathbb{N}.$$

$\{\varphi_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of $L^2(\Omega)$. Fractional powers of $\mathcal{L}$ are defined by

$$\mathcal{L}^s w := \sum_{k=1}^{\infty} \lambda_k^s w_k \varphi_k \quad \forall w \in C_0^\infty(\Omega),\; s \in (0,1), \qquad w_k = \int_\Omega w \varphi_k \,\mathrm{d}x'. \tag{2.6}$$

By density, (2.6) can be extended to

$$\mathbb{H}^s(\Omega) = \left\{ w = \sum_{k=1}^{\infty} w_k \varphi_k \in L^2(\Omega) : \|w\|_{\mathbb{H}^s(\Omega)}^2 := \sum_{k=1}^{\infty} \lambda_k^s |w_k|^2 < \infty \right\}, \tag{2.7}$$

which corresponds to $[L^2(\Omega), H_0^1(\Omega)]_s$ (see [40, Chapter 1]). In addition, if $s \in (\frac{1}{2}, 1)$, this space can be characterized by

$$\mathbb{H}^s(\Omega) = \left\{ w \in H^s(\Omega) : w = 0 \text{ on } \partial\Omega \right\},$$

and if $s \in (0, \frac{1}{2})$, then $\mathbb{H}^s(\Omega) = H^s(\Omega) = H_0^s(\Omega)$. The space $\mathbb{H}^{1/2}(\Omega)$ corresponds to the so-called *Lions–Magenes* space and is characterized as (see [40, Theorem 11.7] and [56, Chapter 33])

$$\mathbb{H}^{1/2}(\Omega) = \left\{ w \in H^{1/2}(\Omega) : \int_\Omega \frac{w^2(x')}{\operatorname{dist}(x', \partial\Omega)} \,\mathrm{d}x' < \infty \right\}.$$

If $s \in (1,2]$, owing to (2.1), we have that $\mathbb{H}^s(\Omega) = H^s(\Omega) \cap H_0^1(\Omega)$; see [24].

For $s \in (0,1)$ we denote by $\mathbb{H}^{-s}(\Omega)$ the dual of $\mathbb{H}^s(\Omega)$. With this notation, $\mathcal{L}^s : \mathbb{H}^s(\Omega) \to \mathbb{H}^{-s}(\Omega)$ is an isomorphism.

2.4. Weighted Sobolev spaces. To study (1.6) we consider Lebesgue and Sobolev spaces with the weight $|y|^\alpha$, $\alpha \in (-1,1)$. For $D \subset \mathbb{R}^{n+1}$ we define

$$L^2(|y|^\alpha, D) = \left\{ w \in L^1_{\mathrm{loc}}(D) : \|w\|_{L^2(|y|^\alpha, D)}^2 = \int_D |y|^\alpha |w|^2 \,\mathrm{d}x < \infty \right\}$$

and

$$H^1(|y|^\alpha, D) = \left\{ w \in L^2(|y|^\alpha, D) : |\nabla w| \in L^2(|y|^\alpha, D) \right\},$$

with norm

$$\|w\|_{H^1(|y|^\alpha, D)} = \left(\|w\|_{L^2(|y|^\alpha, D)}^2 + \|\nabla w\|_{L^2(|y|^\alpha, D)}^2 \right)^{1/2}. \tag{2.8}$$

Since $\alpha \in (-1,1)$, $|y|^\alpha$ belongs to the Muckenhoupt class $A_2(\mathbb{R}^{n+1})$; see [26, 60]. Then, $H^1(|y|^\alpha, D)$ is Hilbert, and $C^\infty(D) \cap H^1(|y|^\alpha, D)$ is dense in $H^1(|y|^\alpha, D)$ (cf. [60, Proposition 2.1.2, Corollary 2.1.6], [36], and [26, Theorem 1]).

We also define the weighted Sobolev space

$$\mathring{H}^1_L(y^\alpha, \mathcal{C}) = \left\{ w \in H^1(y^\alpha, \mathcal{C}) : w = 0 \text{ on } \partial_L \mathcal{C} \right\}. \tag{2.9}$$

As [48, inequality (2.21)] shows, the following *weighted Poincaré inequality* holds,

$$\|w\|_{L^2(y^\alpha,\mathcal{C})} \lesssim \|\nabla w\|_{L^2(y^\alpha,\mathcal{C})} \quad \forall w \in \mathring{H}^1_L(y^\alpha, \mathcal{C}), \tag{2.10}$$

and thus $\|\nabla w\|_{L^2(y^\alpha,\mathcal{C})}$ is equivalent to (2.8) in $\mathring{H}^1_L(y^\alpha, \mathcal{C})$. For $w \in H^1(y^\alpha, \mathcal{C})$, $\operatorname{tr}_\Omega w$ denotes its trace onto $\Omega \times \{0\}$. We recall that, for $\alpha = 1-2s$, [48, Proposition 2.5] yields

$$\operatorname{tr}_\Omega \mathring{H}^1_L(y^\alpha, \mathcal{C}) = \mathbb{H}^s(\Omega), \qquad \|\operatorname{tr}_\Omega w\|_{\mathbb{H}^s(\Omega)} \lesssim \|w\|_{\mathring{H}^1_L(y^\alpha,\mathcal{C})}. \tag{2.11}$$

2.5. The state equation. We follow [49] and define

$$\mathbb{W} := \{ w \in L^\infty(0,T; L^2(\Omega)) \cap L^2(0,T; \mathbb{H}^s(\Omega)) : \partial_t^\gamma w \in L^2(0,T;\mathbb{H}^{-s}(\Omega)) \},$$
$$\mathbb{V} := \{ w \in L^2(0,T; \mathring{H}^1_L(y^\alpha, \mathcal{C})) : \partial_t^\gamma \operatorname{tr}_\Omega w \in L^2(0,T;\mathbb{H}^{-s}(\Omega)) \}.$$

The Caffarelli–Silvestre extension result for problem (1.3) reads [14, 15, 49, 55]: Given $\mathsf{f}, \mathsf{z} \in L^2(0,T;\mathbb{H}^{-s}(\Omega))$, the function $\mathsf{u} \in \mathbb{W}$ solves (1.3) if and only if its harmonic extension $\mathscr{U} \in \mathbb{V}$ solves the following weak version of (1.6): Find $\mathscr{U} \in \mathbb{V}$ such that $\operatorname{tr}_\Omega \mathscr{U}(0) = \mathsf{u}_0$ and for a.e. $t \in (0,T)$

$$\langle \operatorname{tr}_\Omega \partial_t^\gamma \mathscr{U}, \operatorname{tr}_\Omega \phi \rangle + a(\mathscr{U}, \phi) = \langle \mathsf{f} + \mathsf{z}, \operatorname{tr}_\Omega \phi \rangle \quad \forall \phi \in \mathring{H}^1_L(y^\alpha, \mathcal{C}), \tag{2.12}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between $\mathbb{H}^s(\Omega)$ and $\mathbb{H}^{-s}(\Omega)$ and

$$a(w,\phi) = \frac{1}{d_s} \int_{\mathcal{C}} y^\alpha \left(\mathbf{A}(x') \nabla w \cdot \nabla \phi + c(x') w \phi \right) \mathrm{d}x' \, \mathrm{d}y. \tag{2.13}$$

The aforementioned equivalence between (1.3) and (2.12) results in $\mathsf{u} = \operatorname{tr}_\Omega \mathscr{U}$.

The regularity of $\mathbf{A}$ and c implies that a is bounded and coercive in $\mathring{H}^1_L(y^\alpha, \mathcal{C})$. In what follows, we shall use that $a(w,w)^{1/2}$ is a norm equivalent to $|\cdot|_{\mathring{H}^1_L(y^\alpha,\mathcal{C})}$.

Define

$$\Lambda^2_\gamma(\mathsf{v}, \mathsf{g}) := I_t^{1-\gamma} \|\mathsf{v}\|^2_{L^2(\Omega)}(T) + \|\mathsf{g}\|^2_{L^2(0,T;\mathbb{H}^{-s}(\Omega))}, \tag{2.14}$$

$$\Sigma^2(\mathsf{v}, \mathsf{g}) := \|\mathsf{v}\|^2_{L^2(\Omega)} + \|\mathsf{g}\|^2_{L^2(0,T;\mathbb{H}^{-s}(\Omega))}, \tag{2.15}$$

where $I_t^{1-\gamma}$ is the left fractional integral of order $1-\gamma$ defined in (2.3).

THEOREM 4 (existence and uniqueness of u and $\mathscr{U}$). *Given $s \in (0,1)$, $\gamma \in (0,1]$, $\mathsf{f}, \mathsf{z} \in L^2(0,T;\mathbb{H}^{-s}(\Omega))$, and $\mathsf{u}_0 \in L^2(\Omega)$, problems* (1.3) *and* (2.12) *have a unique solution. In addition, we have the following energy estimates for u, solution to* (1.3):

$$I_t^{1-\gamma} \|\mathsf{u}\|^2_{L^2(\Omega)}(T) + \|\mathsf{u}\|^2_{L^2(0,T;\mathbb{H}^s(\Omega))} \lesssim \Lambda^2_\gamma(\mathsf{u}_0, \mathsf{f}+\mathsf{z}), \qquad \gamma < 1, \tag{2.16}$$

$$\|\mathsf{u}\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\mathsf{u}\|^2_{L^2(0,T;\mathbb{H}^s(\Omega))} \lesssim \Sigma^2(\mathsf{u}_0, \mathsf{f}+\mathsf{z}), \qquad \gamma = 1. \tag{2.17}$$

In addition, we have following energy estimates for $\mathscr{U}$*, solution to* (2.12):

$$I_t^{1-\gamma}\|\operatorname{tr}_\Omega \mathscr{U}\|^2_{L^2(\Omega)}(T) + \|\mathscr{U}\|^2_{L^2(0,T;\mathring{H}^1_L(y^\alpha,\mathcal{C}))} \lesssim \Lambda^2_\gamma(\mathsf{u}_0, \mathsf{f}+\mathsf{z}), \qquad \gamma < 1, \tag{2.18}$$

$$\|\operatorname{tr}_\Omega \mathscr{U}\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\mathscr{U}\|^2_{L^2(0,T;\mathring{H}^1_L(y^\alpha,\mathcal{C}))} \lesssim \Sigma^2(\mathsf{u}_0, \mathsf{f}+\mathsf{z}), \qquad \gamma = 1, \tag{2.19}$$

where the hidden constants do not depend on u*,* $\mathscr{U}$*, or the problem data.*

Proof. The well-posedness of (1.3) and (1.6), together with (2.16) and (2.18), is presented in [49, Theorem 6 and Corollary 17]. The estimates (2.17) and (2.19) follow from the arguments developed in [49, 53]. ∎

Remark 5 ($\gamma = 1$). Given $g \in L^p(0,T)$, we have $I_t^\sigma g \to g$ in $L^p(0,T)$ as $\sigma \downarrow 0$ [54, Theorem 2.6]. Take the limit as $\gamma \uparrow 1$ in (2.16) and (2.18) to recover the well-known energy estimates, i.e., (2.17) and (2.19), for parabolic equations with first order derivative in time.

Remark 6 (continuity in time). An adaption of [53, Theorems 2.1–2.2] shows that, for every $\gamma \in (0,1]$ and $s \in (0,1)$, the solution $\operatorname{tr}_\Omega \mathscr{U} = \mathsf{u} \in C([0,T];L^2(\Omega))$. This is not only necessary to make sense of the initial condition, but also to derive optimality conditions, as we will see in section 3.

We conclude with an elementary extension of Lemma 3.

LEMMA 7 (fractional integration by parts). *Let* $\gamma \in (0,1]$*. If* $\mathsf{v}, \mathsf{w} \in \mathbb{W} \cap C([0,T];L^2(\Omega))$*, then we have the following integration-by-parts formula:*

$$\int_0^T \langle \partial_t^\gamma \mathsf{v}(t), \mathsf{w}(t)\rangle - \langle \partial_{T-t}^\gamma \mathsf{w}(t), \mathsf{v}(t)\rangle \,\mathrm{d}t = (\mathsf{w}(T), (I_t^{1-\gamma}\mathsf{v})(T)) - (\mathsf{v}(0), (I_{T-t}^{1-\gamma}\mathsf{w})(0)),$$

where $(\cdot,\cdot) = (\cdot,\cdot)_{L^2(\Omega)}$ *and* $\langle\cdot,\cdot\rangle$ *is the duality pairing between* $\mathbb{H}^s(\Omega)$ *and* $\mathbb{H}^{-s}(\Omega)$*.*

Proof. When v and w are smooth, we integrate (2.5). Conclude by density. ∎

3. The fractional control problem. In this section, we analyze the *space-time fractional optimal control problem*. We derive existence and uniqueness results together with first order necessary and sufficient optimality conditions.

For J defined in (1.1), the fractional control problem reads: Find min $J(\mathsf{u},\mathsf{z})$ subject to the state equation (1.3) and the control constraints (1.4). The set of *admissible controls* is defined by

$$\mathsf{Z}_{\mathrm{ad}} := \left\{\mathsf{w} \in L^2(Q) : \ \mathsf{a}(x',t) \leq \mathsf{w}(x',t) \leq \mathsf{b}(x',t) \ \text{ a.e. } (x',t) \in Q\right\}, \tag{3.1}$$

which is a nonempty, bounded, closed, and convex subset of $L^2(Q)$. To study this problem, following [59, section 3], we introduce the control to state operator.

DEFINITION 8 (control to state operator). *The map* $\mathbf{S} : L^2(0,T;\mathbb{H}^{-s}(\Omega)) \ni \mathsf{z} \mapsto \mathsf{u}(\mathsf{z}) \in \mathbb{W}$*, where* $\mathsf{u}(\mathsf{z})$ *solves* (1.3)*, is called the fractional control to state operator.*

$\mathbf{S}$ is an affine and, by the estimates of Theorem 4, continuous operator. Moreover, since $\mathbb{W} \hookrightarrow L^2(Q) \hookrightarrow L^2(0,T;\mathbb{H}^{-s}(\Omega))$, we may also consider the operator $\mathbf{S}$ as acting from $L^2(Q)$ into itself. For simplicity, we keep the notation $\mathbf{S}$. We now define the optimal fractional state-control pair.

DEFINITION 9 (optimal fractional state-control pair). *A state-control pair* $(\bar{\mathsf{u}}(\bar{\mathsf{z}}), \bar{\mathsf{z}}) \in \mathbb{W} \times \mathsf{Z}_{ad}$ *is called optimal for the problem* (1.2)–(1.4) *if* $\bar{\mathsf{u}}(\bar{\mathsf{z}}) = \mathbf{S}\bar{\mathsf{z}}$ *and*

$$J(\bar{\mathsf{u}}(\bar{\mathsf{z}}), \bar{\mathsf{z}}) \leq J(\mathsf{u}(\mathsf{z}), \mathsf{z})$$

for all $(\mathsf{u}(\mathsf{z}), \mathsf{z}) \in \mathbb{W} \times \mathsf{Z}_{ad}$ *such that* $\mathsf{u}(\mathsf{z}) = \mathbf{S}\mathsf{z}$*.*

The existence and uniqueness of an optimal state-control pair is as follows.

THEOREM 10 (existence and uniqueness). *The optimal control problem* (1.2)–(1.4) *has a unique solution* $(\bar{\mathsf{u}}(\bar{\mathsf{z}}), \bar{\mathsf{z}}) \in \mathbb{W} \times \mathsf{Z}_{ad}$.

Proof. Using the operator $\mathbf{S}$, problem (1.2)–(1.4) reduces to: Minimize

$$f(\mathsf{z}) := \frac{1}{2}\|\mathbf{S}\mathsf{z} - \mathsf{u}_d\|^2_{L^2(Q)} + \frac{\mu}{2}\|\mathsf{z}\|^2_{L^2(Q)} \tag{3.2}$$

over Z_{ad}. Since $\mu > 0$, the strict convexity of f is immediate. $\mathbf{S}$ is continuous, so f is weakly lower semicontinuous. Z_{ad} is weakly sequentially compact. The direct method of the calculus of variations [18, Theorem 1.15] allows us to conclude. ∎

3.1. Formal Lagrangian formulation. We now formally derive first order necessary and sufficient optimality conditions for the optimal control problem (1.2)–(1.4). We proceed via the *Lagrangian approach* described in [59, section 3.1]. We must emphasize that, although these computations are merely formal, they are quite insightful, as they allow us to determine what the correct form of the optimality conditions is with a simple and intuitive procedure.

Letting p denote the adjoint variable, the Lagrangian $\mathbb{L} : \mathbb{W} \times \mathsf{Z}_{\mathrm{ad}} \times \mathbb{W} \to \mathbb{R}$ is

$$\mathbb{L}(\mathsf{u}, \mathsf{z}, \mathsf{p}) = J(\mathsf{u}, \mathsf{z}) - \int_Q (\partial_t^\gamma \mathsf{u} + \mathcal{L}^s \mathsf{u} - \mathsf{f} - \mathsf{z})\, \mathsf{p} \, \mathrm{d}x' \, \mathrm{d}t.$$

We expect the following necessary and sufficient optimality conditions [59, section 3.1]:

$$D_{\mathsf{p}}\mathbb{L}(\bar{\mathsf{u}}, \bar{\mathsf{z}}, \bar{\mathsf{p}})h = 0 \quad \forall h \in \mathbb{W}, \tag{3.3}$$

$$D_{\mathsf{u}}\mathbb{L}(\bar{\mathsf{u}}, \bar{\mathsf{z}}, \bar{\mathsf{p}})h = 0 \quad \forall h \in \mathbb{W}, \text{ with } h(0) = 0, \tag{3.4}$$

$$D_{\mathsf{z}}\mathbb{L}(\bar{\mathsf{u}}, \bar{\mathsf{z}}, \mathsf{p})(\mathsf{z} - \bar{\mathsf{z}}) \geq 0 \quad \forall \mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}. \tag{3.5}$$

We start with a formal computation which uses the integration-by-parts formula (2.5):

$$\begin{aligned}\int_Q (\partial_t^\gamma \bar{\mathsf{u}} + \mathcal{L}^s \bar{\mathsf{u}} - \mathsf{f} - \bar{\mathsf{z}})\, \bar{\mathsf{p}} \, \mathrm{d}x' \, \mathrm{d}t = \int_Q \left(\partial_{T-t}^\gamma \bar{\mathsf{p}} + \mathcal{L}^s \bar{\mathsf{p}}\right) \bar{\mathsf{u}} \, \mathrm{d}x' \, \mathrm{d}t - \int_Q (\mathsf{f} + \mathsf{z})\, \bar{\mathsf{p}} \, \mathrm{d}x' \, \mathrm{d}t \\ - \int_\Omega \bar{\mathsf{u}}(0)(I_{T-t}^{1-\gamma} \bar{\mathsf{p}})(0) \, \mathrm{d}x' + \int_\Omega \bar{\mathsf{p}}(T)(I_t^{1-\gamma} \bar{\mathsf{u}})(T) \, \mathrm{d}x'.\end{aligned}$$

Based on the previous computation, we rewrite (3.4) as follows:

$$\begin{aligned}D_{\mathsf{u}}\mathbb{L}(\bar{\mathsf{u}}, \bar{\mathsf{z}}, \bar{\mathsf{p}})h &= (\bar{\mathsf{u}} - \mathsf{u}_d, h)_{L^2(Q)} - (\partial_{T-t}^\gamma \bar{\mathsf{p}} + \mathcal{L}^s \bar{\mathsf{p}}, h)_{L^2(Q)} \\ &\quad - (\bar{\mathsf{p}}(T), (I_t^{1-\gamma} h)(T))_{L^2(\Omega)} = 0\end{aligned} \tag{3.6}$$

for all $h \in \mathbb{W}$ such that $h(0) = 0$.

Let $\phi \in C_0^\infty(0, T)$, $\psi \in C_0^\infty(\Omega)$ be arbitrary, and define $h = \psi\varphi$, where φ solves the Abel equation $(I_t^{1-\gamma}\varphi)(t) = \phi(t)$. This is possible because of the unique solvability of the Abel equation given in [54, section 2.2, Theorem 2.1]. Notice that with this definition, $(I_t^{1-\gamma} h)(T) = 0$. Using this particular choice of h in (3.6) yields

$$\int_0^T \varphi(t) \left(\partial_{T-t}^\gamma \bar{\mathsf{p}} + \mathcal{L}^s \bar{\mathsf{p}} - (\bar{\mathsf{u}} - \mathsf{u}_d), \psi\right)_{L^2(\Omega)} \mathrm{d}t = 0.$$

Owing to the results of [54, Theorems 13.2 and 13.5], the range of the fractional integral $I_t^{1-\gamma}$ contains all smooth functions. In other words the relation above must hold for all smooth and compactly supported φ, which implies

$$\partial_{T-t}^{\gamma}\bar{\mathsf{p}} + \mathcal{L}^s\bar{\mathsf{p}} = \bar{\mathsf{u}} - \mathsf{u}_d. \tag{3.7}$$

It remains to obtain a terminal condition for $\bar{\mathsf{p}}$. To do so, we notice that we have

$$\big(\bar{\mathsf{p}}(T), (I_t^{1-\gamma}h)(T)\big)_{L^2(\Omega)} = 0.$$

If we were allowed to set h constant in time, this would yield $\bar{\mathsf{p}}(T) = 0$. However, since $h(0) = 0$, the only admissible and constant-in-time function is $h \equiv 0$. To circumvent this we set $h = \ell_\epsilon(t)\chi$ with $\chi \in C_0^\infty(\Omega)$ arbitrary, and $\ell_\epsilon(t) \in C^{0,1}([0,T])$ given by

$$\ell_\epsilon(t) = \epsilon^{-\gamma}T^{-\gamma}t^\gamma, \quad 0 < t \le \epsilon T, \qquad \ell_\epsilon(t) = 1, \quad \epsilon T < t \le T.$$

This particular choice of h yields

$$(\bar{\mathsf{p}}(T), \chi)_{L^2(\Omega)}\,(I_t^{1-\gamma}\ell_\epsilon)(T) = 0.$$

To conclude, it remains to notice that $\lim_{\epsilon\to 0}(I_t^{1-\gamma}\ell_\epsilon)(T) = (I_t^{1-\gamma}1)(T) > 0$. Collecting the derived equations, our formal argument yields the following strong system for the adjoint variable p.

Definition 11 (fractional adjoint state). *The solution* $\mathsf{p} = \mathsf{p}(\mathsf{z}) \in \mathbb{W}$ *of*

$$\partial_{T-t}^{\gamma}\mathsf{p} + \mathcal{L}^s\mathsf{p} = \mathsf{u} - \mathsf{u}_d \ \text{ in } Q, \qquad \mathsf{p}(T) = 0 \ \text{ in } \Omega, \tag{3.8}$$

for $\mathsf{z} \in L^2(0,T;\mathbb{H}^{-s}(\Omega))$, *is called the fractional adjoint state associated to* $\mathsf{u} = \mathsf{u}(\mathsf{z})$.

Remark 12 ($\gamma \to 1$). For $g \in W_1^1(0,T)$, $\partial_{T-t}^{\gamma}g \to -\partial_t g$ as $\gamma \uparrow 1$, and we recover

$$-\partial_t\mathsf{p} + \mathcal{L}^s\mathsf{p} = \mathsf{u} - \mathsf{u}_d \ \text{ in } Q, \qquad \mathsf{p}(T) = 0 \ \text{ in } \Omega,$$

a standard backward parabolic problem with terminal condition.

Well-posedness of (3.8) follows from a change of variables. If $\psi : [0,T] \to \mathbb{R}$, define $\tilde{\psi}(t) := \psi(T-t)$ and notice that $\tilde{\psi}'(t) = -\psi'(T-t)$. Therefore, with $c_\gamma = \Gamma(1-\gamma)$,

$$c_\gamma\partial_t^\gamma\tilde{\psi}(T-t) = \int_0^{T-t}\frac{-\psi'(T-\xi)}{((T-t)-\xi)^\gamma}\,\mathrm{d}\xi = -\int_t^T\frac{\psi'(\mu)}{(\mu-t)^\gamma}\,\mathrm{d}\mu = c_\gamma\partial_{T-t}^{\gamma}\psi(t).$$

As a consequence, the backward in time problem (3.8) with a right Caputo fractional derivative can be equivalently written as a forward in time problem with a left Caputo fractional derivative as (1.3). The well-posedness of (3.8) then follows from the results of section 2.5.

We conclude this formal analysis with the following variational inequality:

$$(\mu\bar{\mathsf{z}} + \bar{\mathsf{p}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} \ge 0 \qquad \forall \mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}, \tag{3.9}$$

which follows from (3.5).

Remark 13 (Lagrangian approach). Although formal, this approach is systematic and useful for deriving optimality conditions of a control problem, especially in our case, where the state equation (1.3) involves fractional derivatives in time and space.

3.2. Optimality conditions. We begin with a classical result.

LEMMA 14 (variational inequality). *Let f be defined by* (3.2). *The function $\bar{\mathsf{z}} \in \mathsf{Z}_{ad}$ minimizes the functional f if and only if*

$$(f'(\bar{\mathsf{z}}), \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} \geq 0 \tag{3.10}$$

for every $\mathsf{z} \in \mathsf{Z}_{ad}$.

Proof. See [59, Lemma 2.21]. ∎

To derive first order optimality conditions, we need the following result.

LEMMA 15 (auxiliary result I). *Let $\bar{\mathsf{z}}$ denote the optimal control given by Theorem* 10 *and $\bar{\mathsf{u}} = \mathbf{S}\bar{\mathsf{z}}$. Then, for every $\mathsf{z} \in \mathsf{Z}_{ad}$, we have*

$$(\bar{\mathsf{u}} - \mathsf{u}_d, \mathsf{u} - \bar{\mathsf{u}})_{L^2(Q)} = (\bar{\mathsf{p}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)}, \tag{3.11}$$

where $\mathsf{u} = \mathbf{S}\mathsf{z} \in \mathbb{W}$ and $\mathsf{p} = \mathsf{p}(\mathsf{z}) \in \mathbb{W}$ solve problems (1.3) *and* (3.8), *respectively.*

Proof. Define $\phi := \mathsf{u} - \bar{\mathsf{u}} \in \mathbb{W}$, and notice that $\phi(0) = 0$ in Ω. Moreover,

$$\langle \partial_t^\gamma \phi + \mathcal{L}^s \phi, w \rangle = (\mathsf{z} - \bar{\mathsf{z}}, w)_{L^2(\Omega)} \quad \forall w \in \mathbb{H}^s(\Omega), \text{ a.e. } (0,T). \tag{3.12}$$

Since $\bar{\mathsf{p}} \in \mathbb{W}$, setting $w = \bar{\mathsf{p}}(t)$ in (3.12) and integrating over time yields

$$(\bar{\mathsf{p}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} = \int_0^T \langle \partial_t^\gamma \phi + \mathcal{L}^s \phi, \bar{\mathsf{p}} \rangle \, \mathrm{d}t.$$

Lemma 7 and the fact that the operator $\mathcal{L}^s$ is self-adjoint allow us to write

$$(\bar{\mathsf{p}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} = \int_0^T \langle \partial_{T-t}^\gamma \bar{\mathsf{p}} + \mathcal{L}^s \bar{\mathsf{p}}, \phi \rangle,$$

where we used the terminal and initial conditions $\bar{\mathsf{p}}(T) = 0$ and $\phi(0) = 0$, respectively, which are well defined in view of Remark 6. On the other hand, setting ϕ as a test function in the weak version of (3.8) and integrating in time yields

$$\int_0^T \langle \partial_{T-t}^\gamma \bar{\mathsf{p}} + \mathcal{L}^s \bar{\mathsf{p}}, \phi \rangle = (\bar{\mathsf{u}} - \mathsf{u}_d, \phi)_{L^2(Q)}.$$

The desired identity (3.11) follows easily from the derived expressions. ∎

We now prove necessary and sufficient optimality conditions for (1.2)–(1.4).

THEOREM 16 (first order optimality conditions). *$\bar{\mathsf{z}} \in \mathsf{Z}_{ad}$ is the optimal control of problem* (1.2)–(1.4) *if and only if it solves* (3.9), *where $\bar{\mathsf{p}} = \bar{\mathsf{p}}(\bar{\mathsf{z}})$ solves* (3.8).

Proof. We recall the control to state operator $\mathbf{S} : L^2(0,T,\mathbb{H}^{-s}(\Omega)) \to \mathbb{W}$ defined by $\mathbf{S}(\mathsf{z}) = \mathsf{u}(\mathsf{z})$, where $\mathsf{u}(\mathsf{z}) \in \mathbb{W}$ solves problem (1.3). Next, we write $\mathbf{S}(\mathsf{z}) = \mathbf{S}_0(\mathsf{z}) + \psi_0$, where $\mathbf{S}_0(\mathsf{z})$ denotes the solution to (1.3) with $\mathsf{f} = 0$ and $\mathsf{u}_0 = 0$, while ψ_0 solves (1.3) with $\mathsf{z} = 0$. Since $\mathbf{S}_0$ is linear, in our setting the variational inequality (3.10) reads

$$(\bar{\mathsf{u}} - \mathsf{u}_d, \mathbf{S}_0(\mathsf{z} - \bar{\mathsf{z}}))_{L^2(Q)} + \mu(\bar{\mathsf{z}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} \geq 0.$$

Since $\mathbf{S}_0(\mathsf{z} - \bar{\mathsf{z}}) = \mathbf{S}_0\mathsf{z} + \psi_0 - (\psi_0 + \mathbf{S}_0\bar{\mathsf{z}}) = \mathsf{u}(\mathsf{z}) - \bar{\mathsf{u}}$, the previous expression becomes

$$(\bar{\mathsf{u}} - \mathsf{u}_d, \mathsf{u}(\mathsf{z}) - \bar{\mathsf{u}})_{L^2(Q)} + \mu(\bar{\mathsf{z}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} \geq 0.$$

Using identity (3.11) of Lemma 15, we arrive at

$$(\bar{\mathsf{p}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} + \mu(\bar{\mathsf{z}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} \geq 0,$$

which is (3.9) and concludes the proof. ∎

3.3. Regularity of the optimal control. Since we shall be concerned with approximating the solution to the control problem (1.2)–(1.4), it is essential to study its regularity. Here, on the basis of a bootstrap argument, we obtain such results. We will need the following assumption on a and b defining the set $\mathsf{Z}_{\rm ad}$:

$$\mathsf{a} \leq 0 \leq \mathsf{b} \quad \text{on } \partial\Omega \times (0,T). \tag{3.13}$$

In what follows, to shorten the exposition, we define

$$\mathfrak{C} := \|\mathsf{u}_0\|_{L^2(\Omega)} + \|\mathsf{f}\|_{L^2(Q)} + \|\mathsf{u}_d\|_{L^2(Q)} + \|\mathsf{a}\|_{H^1(0,T;L^2(\Omega))} + \|\mathsf{b}\|_{H^1(0,T;L^2(\Omega))} \tag{3.14}$$

and, for $\epsilon > 0$,

$$\begin{aligned}\mathfrak{D} := \|\mathsf{f}\|_{L^2(0,T;\mathbb{H}^{1-\epsilon}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{1-\epsilon}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{1-\epsilon}(\Omega)} \\ + \|\mathsf{a}\|_{L^2(0,T;H^1(\Omega))} + \|\mathsf{b}\|_{L^2(0,T;H^1(\Omega))}.\end{aligned} \tag{3.15}$$

THEOREM 17 (regularity of $\bar{\mathsf{z}}$). *Let $\gamma = 1$. Assume that, for every $\epsilon > 0$, we have $\mathsf{f}, \mathsf{u}_d \in L^2(0,T;\mathbb{H}^{1-\epsilon}(\Omega))$ and that $\mathsf{u}_0 \in \mathbb{H}^{1-\epsilon}(\Omega)$. If $\mathsf{a}, \mathsf{b} \in H^1(Q)$ and* (3.13) *holds, then $\bar{\mathsf{z}}$, the solution to the optimal control problem* (1.2)–(1.4)*, satisfies*

$$\|\bar{\mathsf{z}}\|_{H^1(0,T;L^2(\Omega))} \lesssim \mathfrak{C} \tag{3.16}$$

and

$$\|\bar{\mathsf{z}}\|_{L^2(0,T;H^1(\Omega))} \lesssim \mathfrak{D}, \tag{3.17}$$

where in (3.16) *the constant is independent of s, while in* (3.17) *it is inversely proportional to s. In both cases the constants are independent of f, u_d, u_0, a, and b. Moreover, $\bar{\mathsf{z}} \in L^2(0,T;H^1_0(\Omega))$.*

Proof. Let us first prove the regularity in time, (3.16). By assumption, the right-hand side of the state equation (1.3) satisfies $\mathsf{f} + \bar{\mathsf{z}} \in L^2(Q)$, while the initial condition satisfies $\mathsf{u}_0 \in \mathbb{H}^s(\Omega)$. Standard arguments, which heuristically entail multiplying (1.3) by the derivative of the solution, yield the energy estimate

$$\|\bar{\mathsf{u}}\|_{H^1(0,T;L^2(\Omega))} + \|\bar{\mathsf{u}}\|_{L^\infty(0,T;\mathbb{H}^s(\Omega))} \lesssim \|\mathsf{f} + \bar{\mathsf{z}}\|_{L^2(Q)} + \|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)}, \tag{3.18}$$

where we have used the definition of the $\mathbb{H}^s(\Omega)$-norm given in (2.7). We now write (1.3) as $\mathcal{L}^s\bar{\mathsf{u}} = \mathsf{f} + \bar{\mathsf{z}} - \partial_t\bar{\mathsf{u}} \in L^2(Q)$ to conclude that $\bar{\mathsf{u}} \in L^2(0,T;\mathbb{H}^{2s}(\Omega))$ with

$$\|\bar{\mathsf{u}}\|_{L^2(0,T;\mathbb{H}^{2s}(\Omega))} \lesssim \|\mathsf{f} + \bar{\mathsf{z}}\|_{L^2(Q)} + \|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)}. \tag{3.19}$$

We recall that, for $s \in (0,1)$, $\mathbb{H}^{2s}(\Omega)$ is defined in section 2.3. The right-hand side of (3.8) verifies $\bar{\mathsf{u}} - \mathsf{u}_d \in L^2(Q)$. Repeating the arguments that led to (3.18) gives $\bar{\mathsf{p}} \in H^1(0,T;L^2(\Omega)) \cap L^\infty(0,T;\mathbb{H}^s(\Omega))$ with

$$\|\bar{\mathsf{p}}\|_{H^1(0,T;L^2(\Omega))} + \|\bar{\mathsf{p}}\|_{L^\infty(0,T;\mathbb{H}^s(\Omega))} \lesssim \|\bar{\mathsf{u}}\|_{L^2(Q)} + \|\mathsf{u}_d\|_{L^2(Q)}.$$

A basic energy estimate for $\bar{\mathsf{u}}$ then yields

$$\|\bar{\mathsf{p}}\|_{H^1(0,T;L^2(\Omega))} + \|\bar{\mathsf{p}}\|_{L^\infty(0,T;\mathbb{H}^s(\Omega))} \lesssim \|\mathsf{f} + \bar{\mathsf{z}}\|_{L^2(Q)} + \|\mathsf{u}_0\|_{L^2(\Omega)} + \|\mathsf{u}_d\|_{L^2(Q)}. \tag{3.20}$$

From [59, section 3.6.3] we have

$$\bar{\mathsf{z}} = \max\left\{\mathsf{a}, \min\left\{\mathsf{b}, -\frac{1}{\mu}\bar{\mathsf{p}}\right\}\right\}, \tag{3.21}$$

which, together with the fact that $H^1(0,T;L^2(\Omega))$ is a lattice (see [35, Theorem A.1]) and (3.20), allows us to conclude that $\bar{\mathsf{z}} \in H^1(0,T;L^2(\Omega))$ and

$$\|\bar{\mathsf{z}}\|_{H^1(0,T;L^2(\Omega))} \lesssim \|\bar{\mathsf{p}}\|_{H^1(0,T;L^2(\Omega))} + \|\mathsf{a}\|_{H^1(0,T;L^2(\Omega))} + \|\mathsf{b}\|_{H^1(0,T;L^2(\Omega))} \lesssim \mathfrak{C},$$

where $\mathfrak{C}$ is defined in (3.14).

On the basis of a bootstrap argument like that in [5, Lemma 3.5], we now proceed to prove the regularity in space, (3.17). In light of (3.19) and the fact that $\mathsf{u}_d \in L^2(0,T;\mathbb{H}^{1-\epsilon}(\Omega))$, we have $\bar{\mathsf{u}} - \mathsf{u}_d \in L^2(0,T;\mathbb{H}^s(\Omega))$. We define $q_3 = \mathcal{L}^{s/2}\bar{\mathsf{p}}$ and notice that q_3 solves $-\partial_t q_3 + \mathcal{L}^s q_3 = \mathcal{L}^{s/2}(\bar{\mathsf{u}} - \mathsf{u}_d) \in L^2(Q)$ and $q_3(T) = 0$. Thus, in complete analogy to (3.19), we have $q_3 \in L^2(0,T;\mathbb{H}^{2s}(\Omega))$ with

$$\|q_3\|_{L^2(0,T;\mathbb{H}^{2s}(\Omega))} \lesssim \|\mathcal{L}^{s/2}(\bar{\mathsf{u}} - \mathsf{u}_d)\|_{L^2(Q)},$$

which in turn implies that $\bar{\mathsf{p}} = \mathcal{L}^{-s/2}q_3 \in L^2(0,T;\mathbb{H}^{3s}(\Omega))$ with the estimate

$$\begin{aligned} \|\bar{\mathsf{p}}\|_{L^2(0,T;\mathbb{H}^{3s}(\Omega))} &\lesssim \|\bar{\mathsf{u}} - \mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^s(\Omega))} \\ &\lesssim \|\mathsf{f} + \bar{\mathsf{z}}\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^s(\Omega))} + \|\mathsf{u}_0\|_{L^2(\Omega)}, \end{aligned} \tag{3.22}$$

where the last inequality follows from a basic energy estimate for $\bar{\mathsf{u}}$.

We now consider the following two cases:

[1] $s \in [\frac{1}{3}, 1)$: Since $\bar{\mathsf{p}} \in L^2(0,T;H_0^1(\Omega))$, formula (3.21) yields $\bar{\mathsf{z}} \in L^2(0,T;H_0^1(\Omega))$. Notice that assumption (3.13) is needed here to preserve the boundary values. In addition, the H^1-continuity of the projection (3.21) and (3.22) imply

$$\|\bar{\mathsf{z}}\|_{L^2(0,T;H^1(\Omega))} \lesssim \mathfrak{D},$$

where $\mathfrak{D}$ is defined in (3.15).

[2] $s \in (0, \frac{1}{3})$: We now begin the bootstrapping argument. A nonlinear operator interpolation result as in [5, Lemma 3.5] yields that $\bar{\mathsf{z}} \in L^2(0,T;\mathbb{H}^{3s}(\Omega))$. We note that (3.13) is needed to preserve the boundary values. Moreover, in light of (3.22), we have

$$\|\bar{\mathsf{z}}\|_{L^2(0,T;\mathbb{H}^{3s}(\Omega))} \lesssim \|\bar{\mathsf{p}}\|_{L^2(0,T;\mathbb{H}^{3s}(\Omega))} + \|\mathsf{a}\|_{L^2(0,T;H^1(\Omega))} + \|\mathsf{b}\|_{L^2(0,T;H^1(\Omega))} \lesssim \mathfrak{D}.$$

Next, we define $w_4 = \mathcal{L}^{3s/2}\bar{\mathsf{u}}$ and notice that, since $\mathsf{f} + \bar{\mathsf{z}} \in L^2(0,T;\mathbb{H}^{3s}(\Omega))$, w_4 solves $\partial_t w_4 + \mathcal{L}^s w_4 = \mathcal{L}^{3s/2}(\mathsf{f} + \bar{\mathsf{z}}) \in L^2(Q)$ with initial condition $w_4(0) = \mathcal{L}^{3s/2}\mathsf{u}_0 \in L^2(\Omega)$. Therefore $w_4 \in L^2(0,T;\mathbb{H}^s(\Omega))$ with

$$\|w_4\|_{L^2(0,T;\mathbb{H}^s(\Omega))} \lesssim \|\mathcal{L}^{3s/2}(\mathsf{f} + \bar{\mathsf{z}})\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))} + \|\mathcal{L}^{3s/2}\mathsf{u}_0\|_{L^2(\Omega)},$$

which in turn implies that $\bar{\mathsf{u}} = \mathcal{L}^{-3s/2}w_4 \in L^2(0,T;\mathbb{H}^{4s}(\Omega))$ with the estimate

$$\|\bar{\mathsf{u}}\|_{L^2(0,T;\mathbb{H}^{4s}(\Omega))} \lesssim \|\mathsf{f} + \bar{\mathsf{z}}\|_{L^2(0,T;\mathbb{H}^{2s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{3s}(\Omega)}. \tag{3.23}$$

Now define $q_5 = \mathcal{L}^{3s/2}\bar{\mathsf{p}}$. Notice that the same arguments that provide (3.19) yield $q_5 \in L^2(0,T;\mathbb{H}^{2s}(\Omega))$ and, therefore, $\bar{\mathsf{p}} \in L^2(0,T;\mathbb{H}^{5s}(\Omega))$. In addition,

$$\begin{aligned} \|\bar{\mathsf{p}}\|_{L^2(0,T;\mathbb{H}^{5s}(\Omega))} &\lesssim \|\bar{\mathsf{u}}\|_{L^2(0,T;\mathbb{H}^{3s}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{3s}(\Omega))} \\ &\lesssim \|\mathsf{f}\|_{L^2(0,T;\mathbb{H}^{2s}(\Omega))} + \|\bar{\mathsf{z}}\|_{L^2(0,T;\mathbb{H}^{2s}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{3s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{3s}(\Omega)}, \end{aligned} \tag{3.24}$$

where we have used (3.23).

We now consider, again, two cases:

$\boxed{2.1}$ $s \in [\frac{1}{5}, \frac{1}{3})$: As in case $\boxed{1}$, we have that $\bar{\mathsf{z}} \in L^2(0,T;H_0^1(\Omega))$ with

$$\|\bar{\mathsf{z}}\|_{L^2(0,T;H_0^1(\Omega))} \lesssim \mathfrak{D}.$$

$\boxed{2.2}$ $s \in (0, \frac{1}{5})$: Again nonlinear operator interpolation, (3.13), and (3.24) give us that $\bar{\mathsf{z}} \in L^2(0,T;\mathbb{H}^{5s}(\Omega))$. Next define $w_6 = \mathcal{L}^{5s/2}\bar{\mathsf{u}}$, which solves $\partial_t w_6 + \mathcal{L}^s w_6 = \mathcal{L}^{5s/2}(\mathsf{f} + \bar{\mathsf{z}}) \in L^2(Q)$ with $w_6(0) = \mathcal{L}^{5s/2}\mathsf{u}_0 \in L^2(\Omega)$. This yields that $\bar{\mathsf{u}} \in L^2(0,T;\mathbb{H}^{6s}(\Omega))$ and $\bar{\mathsf{p}} \in L^2(0,T;\mathbb{H}^{7s}(\Omega))$. In addition,

$$\begin{aligned}\|\bar{\mathsf{p}}\|_{L^2(0,T;\mathbb{H}^{7s}(\Omega))} &\lesssim \|\bar{\mathsf{u}}\|_{L^2(0,T;\mathbb{H}^{5s}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{5s}(\Omega))} \\ &\lesssim \|\mathsf{f}\|_{L^2(0,T;\mathbb{H}^{4s}(\Omega))} + \|\bar{\mathsf{z}}\|_{L^2(0,T;\mathbb{H}^{4s}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{5s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{5s}(\Omega)}.\end{aligned}$$

We consider, one more time, two cases:

$\boxed{2.2.1}$ $s \in [\frac{1}{7}, \frac{1}{5})$: In this case $\bar{\mathsf{z}} \in L^2(0,T;H_0^1(\Omega))$ with

$$\|\bar{\mathsf{z}}\|_{L^2(0,T;H_0^1(\Omega))} \lesssim \mathfrak{D}.$$

$\boxed{2.2.2}$ $s \in (0, \frac{1}{7})$: Define $w_8 = \mathcal{L}^{7s/2}\bar{\mathsf{u}}$ and argue as before to obtain that, for $s \in [\frac{1}{9}, \frac{1}{7})$,

$$\|\bar{\mathsf{z}}\|_{L^2(0,T;H_0^1(\Omega))} \lesssim \mathfrak{D}.$$

From this procedure we note that, at every step, there is a regularity gain of $2s$. Consequently, after a finite number of steps (which is proportional to s^{-1}) we can conclude that $\bar{\mathsf{z}} \in L^2(0,T;H_0^1(\Omega))$ and that (3.17) holds. This concludes the proof. ∎

Remark 18 (regularity of $\bar{\mathsf{u}}$ and $\bar{\mathsf{p}}$). Notice that, while proving Theorem 17, we have also shown that $\bar{\mathsf{u}}, \bar{\mathsf{p}} \in H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$.

Remark 19 (regularity as $s \downarrow 0$). The constant in the estimate (3.17) blows up as $s \downarrow 0$. This is expected since, at least heuristically, $\mathcal{L}^s w \to w$ as $s \downarrow 0$. Thus, in the limit, there is no spatial regularity gain for $\bar{\mathsf{u}}$ or $\bar{\mathsf{p}}$.

3.4. The extended control problem. To circumvent the nonlocality of the operator $\mathcal{L}^s$ in problem (1.2)–(1.4) we realize it using the Caffarelli–Silvestre extension. In what follows we consider the *equivalent* problem: Find $\min\{J(\operatorname{tr}_\Omega \mathscr{U}, \mathsf{z}) : \mathscr{U} \in \mathbb{V}, \mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}\}$ subject to the *extended state equation*: Find $\mathscr{U} \in \mathbb{V}$ such that $\operatorname{tr}_\Omega \mathscr{U}(0) = \mathsf{u}_0$ in Ω and, for a.e. $t \in (0,T)$,

$$\langle \operatorname{tr}_\Omega \partial_t^\gamma \mathscr{U}, \operatorname{tr}_\Omega \phi \rangle + a(\mathscr{U}, \phi) = \langle \mathsf{f} + \mathsf{z}, \operatorname{tr}_\Omega \phi \rangle \qquad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}), \tag{3.25}$$

where a and $\mathring{H}_L^1(y^\alpha, \mathcal{C})$ are defined in (2.13) and (2.9), respectively.

To describe the optimality conditions we introduce the *extended adjoint problem*: Find $\mathscr{P} \in \mathbb{V}$ such that $\operatorname{tr}_\Omega \mathscr{P}(T) = 0$ in Ω and, for a.e. $t \in (0,T)$,

$$\langle \operatorname{tr}_\Omega \partial_{T-t}^\gamma \mathscr{P}, \operatorname{tr}_\Omega \phi \rangle + a(\mathscr{P}, \phi) = (\operatorname{tr}_\Omega \mathscr{U} - \mathsf{u}_d, \operatorname{tr}_\Omega \phi)_{L^2(\Omega)} \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}). \tag{3.26}$$

The optimality conditions in this setting now read as follows: The pair $(\bar{\mathscr{U}}(\bar{\mathsf{z}}), \bar{\mathsf{z}}) \in \mathbb{V} \times \mathsf{Z}_{\mathrm{ad}}$ is optimal if and only if $\bar{\mathscr{U}}(\bar{\mathsf{z}})$ solves (3.25) and

$$(\operatorname{tr}_\Omega \bar{\mathscr{P}} + \mu \bar{\mathsf{z}}, \mathsf{z} - \bar{\mathsf{z}})_{L^2(Q)} \geq 0 \quad \forall \mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}, \tag{3.27}$$

where $\bar{\mathscr{P}} = \bar{\mathscr{P}}(\bar{\mathsf{z}}) \in \mathbb{V}$ solves (3.26).

We conclude by noticing that the results of Caffarelli and Silvestre [14] (see also [55]) yield that $\operatorname{tr}_\Omega \mathscr{U} = \mathsf{u}$ and $\operatorname{tr}_\Omega \mathscr{P} = \mathsf{p}$, where u and p solve (1.3) and (3.8), respectively.

4. A truncated optimal control problem. The state equation (3.25) is posed on the infinite cylinder $\mathcal{C} = \Omega \times (0, \infty)$; therefore it cannot be directly approximated with finite element–like techniques. The first step towards discretization is to truncate $\mathcal{C}$ to a bounded cylinder $\mathcal{C}_{\mathscr{Y}} = \Omega \times (0, \mathscr{Y})$, which is possible because a suitable norm of $\mathscr{U}$ decreases exponentially in $\mathscr{Y}$; see [49, Proposition 22] for details.

PROPOSITION 20 (exponential decay). *If, for a given* $\gamma \in (0, 1]$ *and* $s \in (0, 1)$, $\mathscr{U} = \mathscr{U}(\mathsf{z}) \in \mathbb{V}$ *solves* (3.25)*, then for every* $\mathscr{Y} > 1$ *we have*

$$\|\nabla \mathscr{U}\|_{L^2(0,T;L^2(y^\alpha,\Omega\times(\mathscr{Y},\infty)))} \lesssim e^{-\sqrt{\lambda_1}\mathscr{Y}/2}\Lambda_\gamma(\mathsf{u}_0, \mathsf{f} + \mathsf{z}), \tag{4.1}$$

where Λ_γ *is defined in* (2.14).

Proposition 20 motivates a *truncated control problem* as follows. We first define

$$\mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathscr{Y}}) = \left\{ w \in H^1(y^\alpha, \mathcal{C}_{\mathscr{Y}}) : w = 0 \text{ on } \partial_L \mathcal{C}_{\mathscr{Y}} \cup \Omega \times \{\mathscr{Y}\} \right\},$$
$$\mathbb{V}_{\mathscr{Y}} = \{ w \in L^2(0, T; \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathscr{Y}})) : \partial_t^\gamma \operatorname{tr}_\Omega w \in L^2(0, T; \mathbb{H}^{-s}(\Omega)) \},$$

and, for $w, \phi \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathscr{Y}})$, the bilinear form

$$a_{\mathscr{Y}}(w, \phi) = \frac{1}{d_s} \int_{\mathcal{C}_{\mathscr{Y}}} y^\alpha \left(\mathbf{A}(x') \nabla w \cdot \nabla \phi + c(x') w \phi \right) \mathrm{d}x' \, \mathrm{d}y. \tag{4.2}$$

We define the *truncated control problem* as: Find $\min\{J(\operatorname{tr}_\Omega v, \mathsf{r}) : v \in \mathbb{V}_{\mathscr{Y}}, \mathsf{r} \in \mathsf{Z}_{\mathrm{ad}}\}$, subject to the *truncated state equation*: Find $v \in \mathbb{V}_{\mathscr{Y}}$ with $\operatorname{tr}_\Omega v(0) = \mathsf{u}_0$ in Ω and

$$\langle \operatorname{tr}_\Omega \partial_t^\gamma v, \operatorname{tr}_\Omega \phi \rangle + a_{\mathscr{Y}}(v, \phi) = \langle \mathsf{f} + \mathsf{r}, \operatorname{tr}_\Omega \phi \rangle \quad \forall \phi \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathscr{Y}}), \text{ a.e. } t \in (0, T). \tag{4.3}$$

Before analyzing the *truncated control problem*, we present the following result.

LEMMA 21 (exponential convergence). *Let* $\mathscr{U}(\mathsf{r})$ *be the solution to* (3.25) *with* z *replaced by* r *and* v *the solution to* (4.3)*. Then, for* $\gamma \in (0, 1]$ *and* $\mathscr{Y} \geq 1$*, we have*

$$\begin{aligned} I^{1-\gamma} \| \operatorname{tr}_\Omega (\mathscr{U}(\mathsf{r}) - v) \|^2_{L^2(\Omega)}(T) + \| \nabla (\mathscr{U}(\mathsf{r}) - v) \|^2_{L^2(0,T;L^2(y^\alpha,\mathcal{C}_{\mathscr{Y}}))} \\ \lesssim e^{-\sqrt{\lambda_1}\mathscr{Y}} \Lambda^2_\gamma(\mathsf{u}_0, \mathsf{r} + \mathsf{f}), \end{aligned} \tag{4.4}$$

where Λ_γ *is defined in* (2.14).

As an instrument we define $\mathcal{H}_\alpha : \mathbb{H}^s(\Omega) \to \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathscr{Y}})$, the α-harmonic extension to $\mathcal{C}_{\mathscr{Y}}$; i.e., if $\mathsf{w} \in \mathbb{H}^s(\Omega)$, then $w = \mathcal{H}_\alpha \mathsf{w}$ solves

$$\operatorname{div}(y^\alpha \nabla w) = 0 \text{ in } \mathcal{C}_{\mathscr{Y}}, \quad w = 0 \text{ on } \partial_L \mathcal{C}_{\mathscr{Y}} \cup \Omega \times \{\mathscr{Y}\}, \quad w = \mathsf{w} \text{ on } \Omega \times \{0\}. \tag{4.5}$$

Remark 22 (initial datum). The initial datum u_0 of (1.3) determines $v(0)$ only on $\Omega \times \{0\}$ in a trace sense. We thus define $v(0) = \mathcal{H}_\alpha \mathsf{u}_0$. Remark 3.4 in [48] provides the estimate $\|\nabla v(0)\|_{L^2(y^\alpha,\mathcal{C}_{\mathscr{Y}})} \lesssim \|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)}$.

Let us now provide, for $\gamma = 1$, an energy estimate that will be useful in deriving an $L^2(Q)$ error estimate for the fully discrete scheme of subsection 5.3.

THEOREM 23 (energy estimate: $\gamma = 1$). *Let* $s \in (0, 1)$ *and* $\gamma = 1$*, and denote by* $v \in \mathbb{V}_{\mathscr{Y}}$ *the solution to* (4.3)*. If* f *and* r *belong to* $L^2(Q)$ *and* $\mathsf{u}_0 \in \mathbb{H}^s(\Omega)$*, then*

$$\| \operatorname{tr}_\Omega \partial_t v \|_{L^2(Q)} + \| \nabla v \|_{L^\infty(0,T;L^2(y^\alpha,\mathcal{C}_{\mathscr{Y}}))} \lesssim \| \mathsf{f} + \mathsf{r} \|_{L^2(Q)} + \| \mathsf{u}_0 \|_{\mathbb{H}^s(\Omega)}, \tag{4.6}$$

where the hidden constant is independent of the data.

Proof. Set $\phi = \partial_t v$ in (4.3), integrate over time, and use the estimate of Remark 22: $\|\nabla v(0)\|_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)}$. ∎

We introduce the *truncated adjoint problem*: Find $p \in \mathbb{V}_\mathscr{Y}$ such that $\operatorname{tr}_\Omega p(T) = 0$ and, for a.e. $t \in (0,T)$,

$$\langle \operatorname{tr}_\Omega \partial^\gamma_{T-t} p, \operatorname{tr}_\Omega \phi\rangle + a_\mathscr{Y}(p,\phi) = \langle \operatorname{tr}_\Omega v - \mathsf{u}_d, \operatorname{tr}_\Omega \phi\rangle \quad \forall \phi \in \mathring{H}^1_L(y^\alpha,\mathcal{C}_\mathscr{Y}). \tag{4.7}$$

The same arguments provided in Theorem 16 allow us to conclude that the pair $(\bar{v}(\bar{\mathsf{r}}),\bar{\mathsf{r}}) \in \mathbb{V}_\mathscr{Y} \times \mathsf{Z}_{\mathrm{ad}}$ is optimal if and only if $\bar{v} = \bar{v}(\bar{\mathsf{r}})$ solves (4.3) and $\bar{\mathsf{r}}$ satisfies

$$(\operatorname{tr}_\Omega \bar{p} + \mu\bar{\mathsf{r}}, \mathsf{r} - \bar{\mathsf{r}})_{L^2(Q)} \geq 0 \quad \forall \mathsf{r} \in \mathsf{Z}_{\mathrm{ad}}, \tag{4.8}$$

where $\bar{p} = \bar{p}(\bar{\mathsf{r}}) \in \mathbb{V}_\mathscr{Y}$ solves (4.7).

The next result shows how $(\bar{v}(\bar{\mathsf{r}}),\bar{\mathsf{r}})$ approximates $(\bar{\mathscr{U}}(\bar{\mathsf{z}}),\bar{\mathsf{z}})$.

Lemma 24 (exponential convergence). *For every $\mathscr{Y} \geq 1$ we have*

$$\|\bar{\mathsf{r}} - \bar{\mathsf{z}}\|_{L^2(Q)} \lesssim e^{-\frac{\sqrt{\lambda_1}}{2}\mathscr{Y}}\big(\Lambda_\gamma(\mathsf{u}_0,\mathsf{f}) + \|\bar{\mathsf{r}}\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))} + \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))}\big) \tag{4.9}$$

and

$$\begin{aligned}\|\bar{\mathscr{U}} - \bar{v}\|_{L^2(0,T;\mathring{H}^1_L(y^\alpha,\mathcal{C}))} \lesssim e^{-\frac{\sqrt{\lambda_1}}{2}\mathscr{Y}}\big(&\Lambda_\gamma(\mathsf{u}_0,\mathsf{f}) + \|\bar{\mathsf{r}}\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))} \\ &+ \|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))}\big),\end{aligned} \tag{4.10}$$

where Λ_γ is defined in (2.14).

Proof. We proceed in four steps:

[1] Set $\mathsf{z} = \bar{\mathsf{r}} \in \mathsf{Z}_{\mathrm{ad}}$ and $\mathsf{r} = \bar{\mathsf{z}} \in \mathsf{Z}_{\mathrm{ad}}$ in the variational inequalities (3.27) and (4.8), respectively, and add the obtained inequalities to arrive at

$$\begin{aligned}\mu\|\bar{\mathsf{r}} - \bar{\mathsf{z}}\|^2_{L^2(Q)} &\leq (\operatorname{tr}_\Omega(\bar{\mathscr{P}} - \bar{p}), \bar{\mathsf{r}} - \bar{\mathsf{z}})_{L^2(Q)} \\ &= (\operatorname{tr}_\Omega(\bar{\mathscr{P}} - \mathscr{P}(\bar{\mathsf{r}})), \bar{\mathsf{r}} - \bar{\mathsf{z}})_{L^2(Q)} + (\operatorname{tr}_\Omega(\mathscr{P}(\bar{\mathsf{r}}) - \bar{p}), \bar{\mathsf{r}} - \bar{\mathsf{z}})_{L^2(Q)}.\end{aligned}$$

[2] Consider $(\operatorname{tr}_\Omega(\bar{\mathscr{P}} - \mathscr{P}(\bar{\mathsf{r}})), \bar{\mathsf{r}} - \bar{\mathsf{z}})_{L^2(Q)}$. Define $\psi := \bar{\mathscr{P}} - \mathscr{P}(\bar{\mathsf{r}}) \in \mathbb{V}$ and observe that $\operatorname{tr}_\Omega \psi(T) = 0$ and, for all $\phi_p \in \mathring{H}^1_L(y^\alpha,\mathcal{C})$, we have

$$\int_0^T \big(\langle \operatorname{tr}_\Omega \partial^\gamma_{T-t}\psi, \operatorname{tr}_\Omega \phi_p\rangle + a(\psi,\phi_p)\big)\,\mathrm{d}t = \int_0^T \big(\operatorname{tr}_\Omega(\bar{\mathscr{U}} - \mathscr{U}(\bar{\mathsf{r}})), \operatorname{tr}_\Omega \phi_p\big)_{L^2(\Omega)}\,\mathrm{d}t.$$

Analogously, define $\varphi := \bar{\mathscr{U}} - \mathscr{U}(\bar{\mathsf{r}}) \in \mathbb{V}$, which satisfies $\operatorname{tr}_\Omega \varphi(0) = 0$ and

$$\int_0^T (\langle \operatorname{tr}_\Omega \partial^\gamma_t \varphi, \operatorname{tr}_\Omega \phi_u\rangle + a(\varphi,\phi_u))\,\mathrm{d}t = \int_0^T (\bar{\mathsf{z}} - \bar{\mathsf{r}}, \operatorname{tr}_\Omega \phi_u)_{L^2(\Omega)}\,\mathrm{d}t \quad \forall \phi_u \in \mathring{H}^1_L(y^\alpha,\mathcal{C}).$$

Set $\phi_p = \varphi$, $\phi_u = \psi$ and apply Lemma 7. Since $\operatorname{tr}_\Omega \psi(T) = \operatorname{tr}_\Omega \varphi(0) = 0$ we get

$$(\operatorname{tr}_\Omega(\bar{\mathscr{P}} - \mathscr{P}(\bar{\mathsf{r}})), \bar{\mathsf{r}} - \bar{\mathsf{z}})_{L^2(Q)} = -\|\operatorname{tr}_\Omega(\bar{\mathscr{U}} - \mathscr{U}(\bar{\mathsf{r}}))\|^2_{L^2(Q)} \leq 0.$$

[3] The previous step shows that proving (4.9) reduces to obtaining a bound for $\|\operatorname{tr}_\Omega(\mathscr{P}(\bar{\mathsf{r}}) - \bar{p})\|_{L^2(Q)}$, and Lemma 21 yields such a bound; see [5, Lemma 4.6] for details.

[4] In order to prove (4.10) we write

$$\bar{\mathscr{U}}(\bar{\mathsf{z}}) - \bar{v}(\bar{\mathsf{r}}) = \big(\bar{\mathscr{U}}(\bar{\mathsf{z}}) - \mathscr{U}(\bar{\mathsf{r}})\big) + (\mathscr{U}(\bar{\mathsf{r}}) - \bar{v}(\bar{\mathsf{r}}))\,.$$

The first term satisfies (2.12) with right-hand side $\bar{\mathsf{z}} - \bar{\mathsf{r}}$, so that by (4.9) this term is bounded. For the second term we again apply Lemma 21. ∎

Remark 25 (regularity of $\bar{\mathsf{r}}$ versus $\bar{\mathsf{z}}$). In Theorem 17 we studied the regularity of $\bar{\mathsf{z}}$. The techniques of [49, Remark 25] allow us to transfer these results to $\bar{\mathsf{r}}$, the solution of the truncated optimal control problem. In a similar fashion, we can establish the regularity results of Remark 18 for $\operatorname{tr}_\Omega \bar{v}$ and $\operatorname{tr}_\Omega \bar{p}$. For brevity we skip the details.

5. Approximation of the state equation. We recall the numerical approximation of the state equation (2.12) developed in [49]. The scheme employs first-degree tensor product finite elements in space and finite differences in time. The latter is the backward Euler scheme for $\gamma = 1$, whereas for $\gamma \in (0,1)$ it is the scheme of [38, 39], which was studied under appropriate time-regularity conditions on the solution $\mathscr{U}$ in [49]. We also derive a novel $L^2(Q)$ a priori error estimate for the fully discrete approximation of the state equation (2.12) with $\gamma = 1$ and $s \in (0,1)$.

5.1. Time discretization. Let $\mathcal{K} \in \mathbb{N}$ denote the number of time steps. Define the uniform time step $\tau = T/\mathcal{K} > 0$, and set $t_k = k\tau$ for $0 \leq k \leq \mathcal{K}$. We denote the time partition by $\mathcal{T} := \{t_k\}_{k=0}^{\mathcal{K}}$. If $\phi \in C([0,T],\mathcal{X})$, we denote $\phi^k = \phi(t_k)$ and $\phi^\tau = \{\phi^k\}_{k=0}^{\mathcal{K}}$. Over sequences $\phi^\tau \subset \mathcal{X}$, we define the norms

$$\|\phi^\tau\|_{\ell^\infty(\mathcal{X})} = \max_{0 \leq k \leq \mathcal{K}} \|\phi^k\|_{\mathcal{X}}, \qquad \|\phi^\tau\|^2_{\ell^2(\mathcal{X})} = \sum_{k=1}^{\mathcal{K}} \tau \|\phi^k\|^2_{\mathcal{X}},$$

and the discrete time derivative δ^1 by

$$\delta^1 \phi^{k+1} = \tau^{-1}(\phi^{k+1} - \phi^k), \quad k = 0, \dots, \mathcal{K}-1. \tag{5.1}$$

As in [49, section 3.2], we also define, for $\gamma \in (0,1)$, the discrete fractional derivative δ^γ as

$$\Gamma(2-\gamma)\delta^\gamma \phi^{k+1} := \sum_{j=0}^{k} \frac{a_j}{\tau^{\gamma-1}} \delta^1 \phi^{k+1-j} = \frac{\phi^{k+1}}{\tau^\gamma} - \sum_{j=0}^{k-1} \frac{a_j - a_{j+1}}{\tau^\gamma} \phi^{k-j} - \frac{a_k}{\tau^\gamma}\phi^0, \tag{5.2}$$

where $a_j = (j+1)^{1-\gamma} - j^{1-\gamma}$ and provided the sum for $k=0$ is defined to be zero.

We note that any sequence $\phi^\tau \subset \mathcal{X}$ can be equivalently understood as a piecewise constant, in time, function $\phi \in L^\infty(0,T;\mathcal{X})$:

$$\phi(t) = \phi^k \qquad \forall t \in (t_{k-1}, t_k], \ k = 1, \dots, \mathcal{K}. \tag{5.3}$$

This identification will be very useful, and in what follows, we will use it repeatedly and without explicit mention.

5.2. Space discretization. The space discretization is based on truncation and the finite element method. The truncation is as in [49, Lemma 24], which shows that truncating $\mathcal{C}$ to $\mathcal{C}_\mathcal{Y}$ induces an exponentially small error. Since we are now dealing with the bounded domain $\mathcal{C}_\mathcal{Y}$, we can discretize using finite elements.

The finite element discretization follows [48, section 4]. Let $\mathscr{T}_\Omega = \{K\}$ be a conforming triangulation of Ω into cells K (simplices or n-rectangles). We denote by $\mathbb{T}_\Omega$ the collection of all conforming refinements of an original mesh $\mathscr{T}_\Omega^0$, and assume $\mathbb{T}_\Omega$ is shape regular [17]. If $\mathscr{T}_\Omega \in \mathbb{T}_\Omega$, we define $h_{\mathscr{T}_\Omega} = \max_{K \in \mathscr{T}_\Omega} h_K$. We define $\mathscr{T}_\mathcal{Y}$ to be a partition of $\mathcal{C}_\mathcal{Y}$ into cells of the form $T = K \times I$, where $K \in \mathscr{T}_\Omega$, and I is an interval that comes from the partition $\{y_m\}_{m=0}^M$ of $[0,\mathcal{Y}]$ defined by

$$y_m = \left(\frac{m}{M}\right)^\zeta \mathcal{Y}, \quad m = 0, \dots, M, \tag{5.4}$$

where $\zeta = \zeta(\alpha) > 3/(1-\alpha) > 1$. The set of all such triangulations is denoted by $\mathbb{T}$. Note that the following weak regularity condition is valid: There is a constant σ such that, for all $\mathscr{T}_\mathcal{Y} \in \mathbb{T}$, if $T_1 = K_1 \times I_1$, $T_2 = K_2 \times I_2 \in \mathscr{T}_\mathcal{Y}$ have nonempty intersection, then $h_{I_1}/h_{I_2} \le \sigma$, where $h_I = |I|$; see [48, 50]. For $\mathscr{T}_\mathcal{Y} \in \mathbb{T}$, we denote by $\mathcal{N}(\mathscr{T}_\mathcal{Y})$ the set of its nodes and by $\mathring{\mathcal{N}}(\mathscr{T}_\mathcal{Y})$ the set of its interior and Neumann nodes. We also denote by $N = \# \mathring{\mathcal{N}}(\mathscr{T}_\mathcal{Y})$ the number of degrees of freedom of $\mathscr{T}_\mathcal{Y}$. We assume that $\#\mathscr{T}_\Omega \approx M^n$ so that $N \approx M^{n+1}$.

The main motivation to consider elements as in (5.4) is to compensate for the rather singular behavior of $\mathscr{U}$, solution to problem (2.12), as $y \approx 0^+$; see [49] for details.

For $\mathscr{T}_\mathcal{Y} \in \mathbb{T}$ and $\Gamma_D = \partial_L \mathcal{C}_\mathcal{Y} \cup \Omega \times \{\mathcal{Y}\}$ we define the finite element space

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

If K is a simplex, then $\mathcal{P}_1(K) = \mathbb{P}_1(K)$, whereas if K is a cube, then $\mathcal{P}_1(K) = \mathbb{Q}_1(K)$. We also define $\mathbb{U}(\mathscr{T}_\Omega) = \operatorname{tr}_\Omega \mathbb{V}(\mathscr{T}_\mathcal{Y})$, i.e., a $\mathcal{P}_1$ finite element space over the mesh $\mathscr{T}_\Omega$.

5.3. A fully discrete scheme. The fully discrete scheme to solve (1.6) combines the space discretization of section 5.2 with the time discretization of section 5.1. To define it, we first consider the weighted elliptic projector $G_{\mathscr{T}_\mathcal{Y}}$ studied in [49, section 4.3]:

$$w \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathcal{Y}) : \qquad a_\mathcal{Y}\big(G_{\mathscr{T}_\mathcal{Y}} w, W\big) = a_\mathcal{Y}(w, W) \quad \forall W \in \mathbb{V}(\mathscr{T}_\mathcal{Y}). \tag{5.5}$$

The fully discrete scheme computes $V^\tau_{\mathscr{T}_\mathcal{Y}} \subset \mathbb{V}(\mathscr{T}_\mathcal{Y})$, an approximation of the solution to (4.3), with $\mathsf{r} = 0$, at each time step. We initialize the scheme by setting

$$V^0_{\mathscr{T}_\mathcal{Y}} = \mathcal{I}_{\mathscr{T}_\Omega} \mathsf{u}_0 = G_{\mathscr{T}_\mathcal{Y}} \circ \mathcal{H}_\alpha \mathsf{u}_0, \tag{5.6}$$

where $\mathcal{I}_{\mathscr{T}_\Omega} = G_{\mathscr{T}_\mathcal{Y}} \circ \mathcal{H}_\alpha$ and $\mathcal{H}_\alpha$ is the α-harmonic extension operator defined in section 4; notice that $\operatorname{tr}_\Omega V^0_{\mathscr{T}_\mathcal{Y}} = \operatorname{tr}_\Omega G_{\mathscr{T}_\mathcal{Y}} v(0)$, where $v(0)$ solves (4.5) with $\mathsf{w} = \mathsf{u}_0$.

For $k = 0, \dots, \mathcal{K}-1$, $V^{k+1}_{\mathscr{T}_\mathcal{Y}} \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$ solves

$$(\delta^\gamma \operatorname{tr}_\Omega V^{k+1}_{\mathscr{T}_\mathcal{Y}}, \operatorname{tr}_\Omega W)_{L^2(\Omega)} + a_\mathcal{Y}(V^{k+1}_{\mathscr{T}_\mathcal{Y}}, W) = \langle \mathsf{f}^{k+1}, \operatorname{tr}_\Omega W \rangle \quad \forall W \in \mathbb{V}(\mathscr{T}_\mathcal{Y}), \tag{5.7}$$

where δ^γ is defined by (5.2) for $\gamma \in (0,1)$ and by (5.1) for $\gamma = 1$ and $\mathsf{f}^{k+1} = \tau^{-1} \int_{t_k}^{t_{k+1}} \mathsf{f} \, \mathrm{d}t$. An approximate solution of problem (1.3) is $U^\tau_{\mathscr{T}_\Omega} \subset \mathbb{U}(\mathscr{T}_\Omega)$ with

$$U^\tau_{\mathscr{T}_\Omega} = \operatorname{tr}_\Omega V^\tau_{\mathscr{T}_\mathcal{Y}}. \tag{5.8}$$

The stability of this scheme is as follows (see [49, Lemma 29]).

THEOREM 26 (unconditional stability: $s \in (0,1)$, $\gamma \in (0,1]$). *Let $V^\tau_{\mathscr{T}_\mathcal{Y}} \subset \mathbb{V}(\mathscr{T}_\mathcal{Y})$ solve* (5.6)–(5.7)*; then we have*

$$I_t^{1-\gamma} \| \operatorname{tr}_\Omega V^\tau_{\mathscr{T}_\mathcal{Y}} \|^2_{L^2(\Omega)}(T) + \| V^\tau_{\mathscr{T}_\mathcal{Y}} \|^2_{\ell^2(\mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathcal{Y}))} \lesssim \Lambda_\gamma(V^0_{\mathscr{T}_\mathcal{Y}}, f^\tau)^2,$$

where I_t^0 is the identity.

To present error estimates for scheme (5.6)–(5.7) for $\gamma \in (0,1)$, we introduce

$$\mathcal{A} = \mathcal{A}(\mathsf{v}, \mathsf{g}) = \|\mathsf{v}\|_{\mathbb{H}^s(\Omega)} + \|\mathsf{g}\|_{H^2(0,T;\mathbb{H}^{-s}(\Omega))},$$
$$\mathcal{B} = \mathcal{B}(\mathsf{v}, \mathsf{g}) = \|\mathsf{v}\|_{\mathbb{H}^{1+3s}(\Omega)} + \|\mathsf{g}|_{t=0}\|_{\mathbb{H}^{1+s}(\Omega)} + \|\mathsf{g}\|_{W^1_\infty(0,T;\mathbb{H}^{1-(1-2\nu)s}(\Omega))},$$

where $\nu > 0$ is arbitrary. Theorem 30 in [49] provides the following error estimates for the scheme (5.6)–(5.7) with $\gamma \in (0,1)$ and $s \in (0,1)$.

THEOREM 27 (error estimates: $s, \gamma \in (0,1)$). *Let* $\mathsf{r} = 0$ *and* $s, \gamma \in (0,1)$. *If* v *solves* (4.3), $V_{\mathscr{T}_{\mathcal{Y}}}^{\tau}$ *solves* (5.6)–(5.7), $\mathcal{A}(\mathsf{u}_0, \mathsf{f}), \mathcal{B}(\mathsf{u}_0, \mathsf{f}) < \infty$, *and* $\mathscr{T}_{\mathcal{Y}}$ *verifies* (5.4), *then*

$$\left[I_t^{1-\gamma} \| \operatorname{tr}_\Omega (v^\tau - V_{\mathscr{T}_{\mathcal{Y}}}^\tau) \|_{L^2(\Omega)}^2 (T) \right]^{\frac{1}{2}} \lesssim \tau^\theta \mathcal{A}(\mathsf{u}_0, f) + |\log N|^{2s} N^{\frac{-(1+s)}{n+1}} \mathcal{B}(\mathsf{u}_0, f) \tag{5.9}$$

and

$$\| v^\tau - V_{\mathscr{T}_{\mathcal{Y}}}^\tau \|_{\ell^2(\mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}))} \lesssim \tau^\theta \mathcal{A}(\mathsf{u}_0, f) + |\log N|^{s} N^{\frac{-1}{n+1}} \mathcal{B}(\mathsf{u}_0, f), \tag{5.10}$$

where $\theta \in (0, \frac{1}{2})$, *and the hidden constant does not depend on* v, $V_{\mathscr{T}_{\mathcal{Y}}}^\tau$, *or the problem data but blows up as* $\theta \uparrow \frac{1}{2}$.

5.4. $L^2(Q)$-error estimate: $s \in (0,1)$ and $\gamma = 1$. We now derive a novel $L^2(Q)$-error estimate for (5.6)–(5.7) with $\gamma = 1$, which is inspired by classical techniques developed, for instance, in [9, 47]. To obtain it, we set $\mathsf{r} = 0$ and consider, as a technical instrument, a semidiscrete approximation to (4.3): Set $V^0 = \mathcal{H}_\alpha \mathsf{u}_0$ and, for $k = 0, \dots, \mathcal{K} - 1$, compute $V^{k+1} \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}})$ that solves

$$(\delta^1 \operatorname{tr}_\Omega V^{k+1}, \operatorname{tr}_\Omega \phi)_{L^2(\Omega)} + a_{\mathcal{Y}}(V^{k+1}, \phi) = \langle \mathsf{f}^{k+1}, \operatorname{tr}_\Omega \phi \rangle \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}), \tag{5.11}$$

where δ^1 and $a_{\mathcal{Y}}$ are defined by (5.1) and (4.2), respectively. We present the following stability result.

LEMMA 28 (improved stability). *Let* $V^\tau \subset \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}})$ *solve* (5.11). *If* $\mathsf{f} \in L^2(Q)$ *and* $\mathsf{u}_0 \in \mathbb{H}^s(\Omega)$, *then we have*

$$\| \operatorname{tr}_\Omega \delta^1 V^\tau \|_{\ell^2(L^2(\Omega))}^2 + \| \nabla V^\tau \|_{\ell^\infty(L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}}))}^2 \lesssim \| \mathsf{f}^\tau \|_{\ell^2(L^2(\Omega))}^2 + \| \mathsf{u}_0 \|_{\mathbb{H}^s(\Omega)}^2, \tag{5.12}$$

where the hidden constant is independent of the data, N, τ, *and* V^τ.

Proof. Set $\phi = V^{k+1} - V^k$ and use the estimate of Remark 22. ∎

Define the piecewise linear function $\hat{V} \in C^{0,1}([0,T]; \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}))$ by $\hat{V}(0) = V^0$ and

$$\hat{V}(t) = V^k + (t - t_k) \delta^1 V^{k+1}, \quad t \in (t_k, t_{k+1}], \tag{5.13}$$

for $k = 0, \dots, \mathcal{K} - 1$. Using this notation, we rewrite (5.11) as

$$(\operatorname{tr}_\Omega \partial_t \hat{V}(t), \operatorname{tr}_\Omega \phi)_{L^2(\Omega)} + a_{\mathcal{Y}}(V^\tau(t), \phi) = \langle \mathsf{f}^\tau(t), \operatorname{tr}_\Omega \phi \rangle \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \tag{5.14}$$

for a.e. $t \in (0,T)$. We are now in position to derive an error estimate for (5.11).

THEOREM 29 (semidiscrete error estimate: $\gamma = 1$). *Let* v *and* V^τ *solve* (4.3) *and* (5.11), *respectively. If* $\mathsf{f} \in L^\infty(0,T; L^2(\Omega))$ *and* $\mathsf{u}_0 \in \mathbb{H}^s(\Omega)$, *then*

$$\| \operatorname{tr}_\Omega (v - V^\tau) \|_{L^2(0,T;L^2(\Omega))} \lesssim \tau \left(\| \mathsf{f} \|_{L^\infty(0,T;L^2(\Omega))} + \| \mathsf{u}_0 \|_{\mathbb{H}^s(\Omega)} \right), \tag{5.15}$$

where the hidden constant is independent of the data, N, τ, v, *and* V^τ.

Proof. Define $\hat{e} = v - \hat{V}$ and $\bar{e} = v - V^\tau$. Set $\gamma = 1$ and $\mathsf{r} = 0$ in (4.3) and then subtract from it (5.14). Integrating the result with respect to time, we obtain

$$\begin{aligned}(\operatorname{tr}_\Omega \bar{e}(t), \operatorname{tr}_\Omega \phi)_{L^2(\Omega)} + a_{\mathcal{Y}} \left(\int_0^t \bar{e}(\xi) \, \mathrm{d}\xi, \phi \right) &= \left\langle \int_0^t (\mathsf{f}(\xi) - \mathsf{f}^\tau(\xi)) \, \mathrm{d}\xi, \operatorname{tr}_\Omega \phi \right\rangle \\ &+ (\operatorname{tr}_\Omega (\bar{e}(t) - \hat{e}(t)), \operatorname{tr}_\Omega \phi)_{L^2(\Omega)} \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathcal{Y}}), \text{ a.e. } t \in (0,T).\end{aligned}$$

Set, for a.e. $t \in (0,T)$, $\phi = \bar{e}(t) \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathscr{Y})$. Integrating over time once more yields

$$\int_0^T \| \operatorname{tr}_\Omega \bar{e}(t) \|^2_{L^2(\Omega)} \,\mathrm{d}t \le \left| \int_0^T \left\langle \int_0^t (\mathsf{f}(\xi) - \mathsf{f}^\tau(\xi)) \,\mathrm{d}\xi, \operatorname{tr}_\Omega \bar{e}(t) \right\rangle \mathrm{d}t \right| + \left| \int_0^T (\operatorname{tr}_\Omega(\bar{e}(t) - \hat{e}(t)), \operatorname{tr}_\Omega \bar{e}(t))_{L^2(\Omega)} \,\mathrm{d}t \right|, \tag{5.16}$$

where we have used that

$$\int_0^T a_\mathscr{Y}\left(\int_0^t \bar{e}(\xi)\,\mathrm{d}\xi, \bar{e}(t) \right) \mathrm{d}t = \frac{1}{2} a_\mathscr{Y}\left(\int_0^T \bar{e}(t)\,\mathrm{d}t, \int_0^T \bar{e}(t)\,\mathrm{d}t \right) \ge 0.$$

Notice that, since $\mathsf{f}^{k+1} = \tau^{-1} \int_{t_k}^{t_{k+1}} \mathsf{f}(t)\,\mathrm{d}t$,

$$\int_0^{t_l} (\mathsf{f}(\xi) - \mathsf{f}^\tau(\xi))\,\mathrm{d}\xi = \sum_{k=1}^{l} \int_{t_{k-1}}^{t_k} \left(\mathsf{f}(\xi) - \mathsf{f}^k \right) \mathrm{d}\xi = 0.$$

Consequently, if $t_l \le t < t_{l+1}$, we have

$$\int_0^t (\mathsf{f}(\xi) - \mathsf{f}^\tau(\xi))\,\mathrm{d}\xi = \int_{t_l}^t (\mathsf{f}(\xi) - \mathsf{f}^\tau(\xi))\,\mathrm{d}\xi \lesssim \tau \|\mathsf{f}\|_{L^\infty(0,t)}.$$

In conclusion, the first term on the right-hand side of (5.16) can be bounded by

$$\left| \int_0^T \left\langle \int_0^t (\mathsf{f}(\xi) - \mathsf{f}^\tau(\xi))\,\mathrm{d}s, \operatorname{tr}_\Omega \bar{e}(t) \right\rangle \mathrm{d}t \right| \lesssim \tau^2 \|\mathsf{f}\|^2_{L^\infty(0,T,L^2(\Omega))} + \frac{1}{4} \| \operatorname{tr}_\Omega \bar{e} \|^2_{L^2(Q)}.$$

Since, on $(t_k, t_{k+1}]$, we have that $|\bar{e}(t) - \hat{e}(t)| \le \tau |\delta^1 V^{k+1}|$, estimate (5.12) yields

$$\int_0^T \| \operatorname{tr}_\Omega(\hat{e}(t) - \bar{e}(t)) \|^2_{L^2(\Omega)} \,\mathrm{d}t \le \tau^2 \left\| \delta^1 V^\tau \right\|^2_{\ell^2(L^2(\Omega))} \lesssim \tau^2 \left(\|\mathsf{f}^\tau\|^2_{\ell^2(L^2(\Omega))} + \|\mathsf{u}_0\|^2_{\mathbb{H}^s(\Omega)} \right),$$

and therefore the second term on the right-hand side of (5.16) can be bounded by

$$\left| \int_0^T (\operatorname{tr}_\Omega(\bar{e}(t) - \hat{e}(t)), \operatorname{tr}_\Omega \bar{e}(t))_{L^2(\Omega)} \,\mathrm{d}t \right| \le \frac{1}{4} \| \operatorname{tr}_\Omega \bar{e} \|^2_{L^2(0,T;L^2(\Omega))} + C\tau^2 \left(\|\mathsf{f}^\tau\|^2_{\ell^2(L^2(\Omega))} + \|\mathsf{u}_0\|^2_{\mathbb{H}^s(\Omega)} \right).$$

Collecting all the derived bounds, we arrive at the desired error estimate (5.15). ◻

With this estimate in hand we can control the difference between the fully and the semidiscrete problems.

THEOREM 30 (auxiliary error estimate: $\gamma = 1$). *Let $\gamma = 1$, and assume that $\mathsf{u}_0 \in \mathbb{H}^{1+s}(\Omega)$ and $\mathsf{f} \in L^2(0,T;\mathbb{H}^{1-s}(\Omega))$. If V^τ and $V^\tau_{\mathscr{T}_\mathscr{Y}}$ solve (5.11) and (5.7), respectively, then*

$$\| \operatorname{tr}_\Omega (V^\tau - V^\tau_{\mathscr{T}_\mathscr{Y}}) \|_{\ell^2(L^2(\Omega))} \lesssim |\log N|^{2s} N^{-\frac{1+s}{n+1}} \left(\|\mathsf{f}^\tau\|_{\ell^2(0,T;\mathbb{H}^{1-s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{1+s}(\Omega)} \right),$$

where the hidden constant does not depend on v, V^τ, or the problem data.

Proof. We start by defining the error

$$E^\tau = (V^\tau - G_{\mathscr{T}_\mathscr{Y}} V^\tau) + (G_{\mathscr{T}_\mathscr{Y}} V^\tau - V^\tau_{\mathscr{T}_\mathscr{Y}}) = \theta^\tau + \rho^\tau_{\mathscr{T}_\mathscr{Y}}, \tag{5.17}$$

where $G_{\mathscr{T}_\mathscr{Y}}$ is defined in (5.5). We estimate θ^τ by invoking the approximation properties [49, Proposition 28] of $G_{\mathscr{T}_\mathscr{Y}}$ and the regularity results of [49, Theorem 7]:

$$\| \operatorname{tr}_\Omega \theta^\tau \|_{\ell^2(L^2(\Omega))} \lesssim |\log N|^{2s} N^{-\frac{1+s}{n+1}} \left(\|\mathsf{f}^\tau\|_{\ell^2(\mathbb{H}^{1-s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{1+s}(\Omega)} \right).$$

The estimate of $\rho^\tau_{\mathscr{T}_\mathscr{Y}}$ follows along the lines of [9, Lemma 5.6]. For brevity, we skip the details. ∎

We collect the estimates of Theorems 29 and 30 to derive a $L^2(Q)$-error estimate.

THEOREM 31 (error estimate for v: $\gamma = 1$). *Assume that $\gamma = 1$, and let v and $V^\tau_{\mathscr{T}_\mathscr{Y}}$ solve* (4.3) *and* (5.7)*, respectively. If $\mathsf{f} \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;\mathbb{H}^{1-s}(\Omega))$ and $\mathsf{u}_0 \in \mathbb{H}^{1+s}(\Omega)$, then*

$$\begin{aligned}\| \operatorname{tr}_\Omega (v - V^\tau_{\mathscr{T}_\mathscr{Y}})\|_{L^2(Q)} \lesssim\ & \tau \left(\|\mathsf{f}\|_{L^\infty(0,T;L^2(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)} \right) \\ & + |\log N|^{2s} N^{-\frac{1+s}{n+1}} \left(\|\mathsf{f}\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{1+s}(\Omega)} \right),\end{aligned}$$

where the hidden constant does not depend on N, τ, v, V^τ, or the problem data.

COROLLARY 32 (error estimate for u: $\gamma = 1$). *Assume that $\gamma = 1$, and let u solve* (1.3) *with $\mathsf{z} = 0$, and $U^\tau_{\mathscr{T}_\Omega}$ be defined by* (5.8)*. If $\mathsf{f} \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;\mathbb{H}^{1-s}(\Omega))$ and $\mathsf{u}_0 \in \mathbb{H}^{1+s}(\Omega)$, then*

$$\begin{aligned}\|\mathsf{u} - U^\tau_{\mathscr{T}_\Omega}\|_{L^2(0,T;L^2(\Omega))} \lesssim\ & \tau \left(\|\mathsf{f}\|_{L^\infty(0,T;L^2(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)} \right) \\ & + |\log N|^{2s} N^{-\frac{1+s}{n+1}} \left(\|\mathsf{f}\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))} + \|\mathsf{u}_0\|_{\mathbb{H}^{1+s}(\Omega)} \right),\end{aligned}$$

where the hidden constant does not depend on N, τ, v, V^τ, or the problem data.

Proof. Recalling that v solves (4.3) with $\mathsf{r} = 0$, a simple application of the triangle inequality allows us to write

$$\|\mathsf{u} - U^\tau_{\mathscr{T}_\Omega}\|_{L^2(0,T;L^2(\Omega))} \le \| \operatorname{tr}_\Omega (\mathscr{U} - v)\|_{L^2(0,T;L^2(\Omega))} + \| \operatorname{tr}_\Omega (v - V^\tau_{\mathscr{T}_\mathscr{Y}})\|_{L^2(0,T;L^2(\Omega))}.$$

Lemma 21 yields $\| \operatorname{tr}_\Omega (\mathscr{U} - v)\|_{L^2(0,T;L^2(\Omega))} \lesssim e^{-\sqrt{\lambda_1}\mathscr{Y}/2} \Lambda_\gamma(\mathsf{u}_0, \mathsf{f})$. The bound of the second term follows from Theorem 31. Finally, the desired estimate is a consequence of the natural choice of the truncation parameter, $\mathscr{Y}$, as $\mathscr{Y} \approx |\log N|$; see [48, Remark 5.5]. ∎

6. Approximation of the fractional control problem. We propose an implicit fully discrete scheme to approximate the solution of the fractional control problem (1.2)–(1.4): piecewise constant functions for the control and, for the state, first-degree tensor product finite elements in space, as described in section 5.2, and the finite difference discretization in time detailed in section 5.1.

As stated in Theorem 27, in order to have the error estimates (5.9) and (5.10) for the approximation of the state equation (1.3), we require that $\mathcal{A}(\mathsf{u}_0, \mathsf{f} + \bar{\mathsf{r}}) < \infty$. This strong H^2 in time regularity assumption is not satisfied by the optimal control $\bar{\mathsf{r}}$, meaning that we are not able to apply the results of Theorem 27. This is in sharp contrast with the case $\gamma = 1$ which, according to Theorem 31, requires only $\mathsf{f}, \bar{\mathsf{r}} \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;\mathbb{H}^{1-s}(\Omega))$, which, by imposing (3.13) and invoking

Remark 25 and Theorem 17, is satisfied by the optimal control $\bar{\mathsf{r}}$. Due to this regularity restriction, we can obtain an error analysis for $\gamma = 1$ only. We remark that $L^2(Q)$-error estimates for $s, \gamma \in (0,1)$ are not available in the literature, especially under the correct regularity assumptions. In section 6.2 we will present error estimates for $\gamma = 1$ and $s \in (0,1)$, and in section 6.3 we will show the convergence, without rates, for the remaining range of parameters.

Finally, to simplify the exposition, in what follows we assume that a and b are constants that satisfy (3.13).

6.1. An implicit fully discrete scheme. To discretize the control we introduce the finite element space of piecewise constant functions over $\mathscr{T}_\Omega$,

$$\mathbb{Z}(\mathscr{T}_\Omega) = \left\{ Z \in L^\infty(\Omega) : Z|_K \in \mathbb{P}_0(K) \quad \forall K \in \mathscr{T}_\Omega \right\},$$

and the space of piecewise constant functions in time and space,

$$\mathbb{Z}(\mathcal{T}, \mathscr{T}_\Omega) = \left\{ Z^\tau \subset L^\infty(Q) : Z^k \in \mathbb{Z}(\mathscr{T}_\Omega) \right\}. \tag{6.1}$$

We define the space of discrete admissible controls as follows:

$$\mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega) = \mathsf{Z}_{\mathrm{ad}} \cap \mathbb{Z}(\mathcal{T}, \mathscr{T}_\Omega), \tag{6.2}$$

where Z_{ad} is defined in (3.1). It will be useful to introduce the $L^2(Q)$-orthogonal projection onto $\mathbb{Z}(\mathcal{T}, \mathscr{T}_\Omega)$. The operator $\Pi_{\mathscr{T}_\Omega}^{\mathcal{T}} : L^2(Q) \to \mathbb{Z}(\mathcal{T}, \mathscr{T}_\Omega)$ is defined by

$$r \in L^2(Q) : \quad (r - \Pi_{\mathscr{T}_\Omega}^{\mathcal{T}} r, Z)_{L^2(Q)} = 0 \quad \forall Z \in \mathbb{Z}(\mathcal{T}, \mathscr{T}_\Omega) \tag{6.3}$$

and, for all $r \in H^1(0,T; L^2(\Omega)) \cap L^2(0,T; H^1(\Omega))$, satisfies

$$\| r - \Pi_{\mathscr{T}_\Omega}^{\mathcal{T}} r \|_{L^2(Q)} \lesssim h_{\mathscr{T}_\Omega} \| \nabla_{x'} r \|_{L^2(Q)} + \tau \| \partial_t r \|_{L^2(Q)}. \tag{6.4}$$

Notice also that, since a and b are constant, $\Pi_{\mathscr{T}_\Omega}^{\mathcal{T}} \mathsf{Z}_{\mathrm{ad}} \subset \mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega)$.

We define the discrete functional $J_{\mathscr{T}_\Omega}^{\mathcal{T}} : \mathbb{U}(\mathscr{T}_\Omega)^{\mathcal{K}} \times \mathbb{Z}(\mathcal{T}, \mathscr{T}_\Omega) \to \mathbb{R}$ by

$$J_{\mathscr{T}_\Omega}^{\mathcal{T}}(U_{\mathscr{T}_\Omega}^\tau, Z_{\mathscr{T}_\Omega}^\tau) = \frac{1}{2} \| U_{\mathscr{T}_\Omega}^\tau - \mathsf{u}_d^\tau \|_{\ell^2(L^2(\Omega))}^2 + \frac{\mu}{2} \| Z_{\mathscr{T}_\Omega}^\tau \|_{\ell^2(L^2(\Omega))}^2,$$

where the ℓ^2-norm is defined in section 5.1. The identification between a sequence ϕ^τ and the piecewise constant function (5.3) will be used repeatedly below. For instance, if $\mathsf{u}_d^\tau = \mathsf{u}_d$, we would have that $J_{\mathscr{T}_\Omega}^{\mathcal{T}}(w, r) = J(w, r)$ whenever $w^\tau = w$ and $r^\tau = r$; that is, the arguments are piecewise constant over $\mathcal{T}$. This was already implicitly used in (6.2), when we defined $\mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega)$.

The numerical scheme reads: Find $\min J_{\mathscr{T}_\Omega}^{\mathcal{T}}(\operatorname{tr}_\Omega V_{\mathscr{T}_\mathcal{Y}}^\tau, Z_{\mathscr{T}_\Omega}^\tau)$ subject to the discrete state equation: Initialize as in (5.6), and for $k = 0, \dots, \mathcal{K} - 1$ let $V_{\mathscr{T}_\mathcal{Y}}^{k+1} \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$ solve

$$(\delta^\gamma \operatorname{tr}_\Omega V_{\mathscr{T}_\mathcal{Y}}^{k+1}, \operatorname{tr}_\Omega W)_{L^2(\Omega)} + a_\mathcal{Y}(V_{\mathscr{T}_\mathcal{Y}}^{k+1}, W) = \left\langle \mathsf{f}^{k+1} + Z_{\mathscr{T}_\Omega}^{k+1}, \operatorname{tr}_\Omega W \right\rangle \tag{6.5}$$

for all $W \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$ and the control constraints $Z_{\mathscr{T}_\Omega}^\tau \in \mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega)$. If $(\bar{V}_{\mathscr{T}_\mathcal{Y}}^\tau, \bar{Z}_{\mathscr{T}_\Omega}^\tau)$ denotes the solution to this problem, setting

$$\bar{U}_{\mathscr{T}_\Omega}^\tau = \operatorname{tr}_\Omega \bar{V}_{\mathscr{T}_\mathcal{Y}}^\tau, \tag{6.6}$$

we obtain a fully discrete approximation $(\bar{U}_{\mathscr{T}_\Omega}^\tau, \bar{Z}_{\mathscr{T}_\Omega}^\tau) \in \mathbb{U}(\mathscr{T}_\Omega)^{\mathcal{K}} \times \mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega)$ to the fractional control problem (1.2)–(1.4).

Remark 33 (locality). The main advantage of the scheme (6.5) approximating the fractional control problem (1.2)–(1.4) via (6.6) is its *local nature*, meaning that the solution is found by solving a local PDE.

6.2. A priori error analysis: $\gamma = 1$ and $s \in (0,1)$. Let us consider $s \in (0,1)$ and $\gamma = 1$ in (3.25)–(3.26) and provide an a priori error analysis for the fully discrete scheme proposed in section 6.1. To do so, we provide first order necessary and sufficient optimality conditions for the fully discrete problem. We define the discrete adjoint problem: Find $P^{\tau}_{\mathscr{T}_{\mathcal{Y}}} \subset \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ such that $\operatorname{tr}_{\Omega} P^{\mathcal{K}}_{\mathscr{T}_{\mathcal{Y}}} = 0$, and for $k = \mathcal{K}-1, \dots, 0$, $P^{k}_{\mathscr{T}_{\mathcal{Y}}} \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ solves

$$(6.7) \qquad (\bar{\delta}^1 \operatorname{tr}_{\Omega} P^{k}_{\mathscr{T}_{\mathcal{Y}}}, \operatorname{tr}_{\Omega} W)_{L^2(\Omega)} + a_{\mathcal{Y}}(P^{k}_{\mathscr{T}_{\mathcal{Y}}}, W) = \left\langle \operatorname{tr}_{\Omega} V^{k+1}_{\mathscr{T}_{\mathcal{Y}}} - \mathsf{u}_d^{k+1}, \operatorname{tr}_{\Omega} W \right\rangle$$

for all $W \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$. Here $\bar{\delta}^1$ denotes $\bar{\delta}^1 \phi^k = -\tau^{-1}(\phi^{k+1} - \phi^k)$. The optimality condition reads: $(\bar{V}^{\tau}_{\mathscr{T}_{\mathcal{Y}}}, \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}})$ is optimal if and only if $\bar{V}^{\tau}_{\mathscr{T}_{\mathcal{Y}}}$ solves (5.6) and (6.5) and

$$(6.8) \qquad (\operatorname{tr}_{\Omega} \bar{P}^{\tau}_{\mathscr{T}_{\mathcal{Y}}} + \mu \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}}, Z - \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}})_{L^2(Q)} \geq 0 \qquad \forall Z \in \mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_{\Omega}),$$

where $\bar{P}^{\tau}_{\mathscr{T}_{\mathcal{Y}}}$ solves (6.7). Notice that (6.8) can be equivalently written as

$$(\operatorname{tr}_{\Omega} \bar{P}^{k}_{\mathscr{T}_{\mathcal{Y}}} + \mu \bar{Z}^{k}_{\mathscr{T}_{\Omega}}, Z - \bar{Z}^{k}_{\mathscr{T}_{\Omega}})_{L^2(\Omega)} \geq 0 \quad \forall Z \in \mathbb{Z}(\mathscr{T}_{\Omega}), \quad \mathsf{a} \leq Z \leq \mathsf{b}, \quad \forall k = 1, \dots, \mathcal{K}.$$

To see this, it suffices to set $Z^{\tau} = Z\chi_{(t_{k-1}, t_k]}$, with $Z \in \mathbb{Z}(\mathscr{T}_{\Omega})$ and $\mathsf{a} \leq Z \leq \mathsf{b}$. This greatly simplifies the implementation.

Let us now introduce two auxiliary problems. The first one reads: Find $Q^{\tau}_{\mathscr{T}_{\mathcal{Y}}} \subset \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ such that $\operatorname{tr}_{\Omega} Q^{\mathcal{K}}_{\mathscr{T}_{\mathcal{Y}}} = 0$ and, for $k = \mathcal{K}-1, \dots, 0$, $Q^{k}_{\mathscr{T}_{\mathcal{Y}}} \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ solves

$$(6.9) \qquad (\bar{\delta}^1 \operatorname{tr}_{\Omega} Q^{k}_{\mathscr{T}_{\mathcal{Y}}}, \operatorname{tr}_{\Omega} W)_{L^2(\Omega)} + a_{\mathcal{Y}}(Q^{k}_{\mathscr{T}_{\mathcal{Y}}}, W) = \left\langle \operatorname{tr}_{\Omega} \bar{v}^{k+1} - \mathsf{u}_d^{k+1}, \operatorname{tr}_{\Omega} W \right\rangle$$

for all $W \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$, and where $\bar{v} = \bar{v}(\bar{\mathsf{r}})$ solves (4.3). The second one is: Find $R^{\tau}_{\mathscr{T}_{\mathcal{Y}}} \subset \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ such that $\operatorname{tr}_{\Omega} R^{\mathcal{K}}_{\mathscr{T}_{\mathcal{Y}}} = 0$, and for $k = \mathcal{K}-1, \dots, 0$, $R^{k}_{\mathscr{T}_{\mathcal{Y}}} \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ solves

$$(6.10) \qquad (\bar{\delta}^1 \operatorname{tr}_{\Omega} R^{k}_{\mathscr{T}_{\mathcal{Y}}}, \operatorname{tr}_{\Omega} W)_{L^2(\Omega)} + a_{\mathcal{Y}}(R^{k}_{\mathscr{T}_{\mathcal{Y}}}, W) = \left\langle \operatorname{tr}_{\Omega} V^{k+1}_{\mathscr{T}_{\mathcal{Y}}}(\bar{\mathsf{r}}) - \mathsf{u}_d^{k+1}, \operatorname{tr}_{\Omega} W \right\rangle$$

for all $W \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$. These auxiliary problems will allow us to derive error estimates for the fully discrete scheme proposed in section 6.1.

LEMMA 34 (error estimate for the control: $\gamma = 1$ and $s \in (0,1)$). *Let $\bar{\mathsf{r}}$ be the solution to the truncated optimal control problem of section* 4*, and let $\bar{Z}^{\tau}_{\mathscr{T}_{\Omega}}$ be the solution to the fully discrete optimal control problem of section* 6.1*. Assume that $\mathsf{u}_0 \in \mathbb{H}^{1+s}(\Omega)$ and, for every $\epsilon > 0$, u_d and f belong to $L^2(0,T;\mathbb{H}^{1-\epsilon}(\Omega)) \cap L^{\infty}(0,T;L^2(\Omega))$. Then*

$$\|\bar{\mathsf{r}} - \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}}\|_{L^2(Q)} \lesssim \tau + |\log N|^{2s} N^{-\frac{1}{n+1}},$$

where the hidden constant is independent of the discretization parameters but depends on the problem data.

Proof. We proceed in several steps:

[1] Setting $\mathsf{r} = \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}}$ and $Z = \Pi^{\mathcal{T}}_{\mathscr{T}_{\Omega}} \bar{\mathsf{r}}$ in (4.8) and (6.8), respectively, and adding the derived inequalities, we arrive at

$$\mu \|\bar{\mathsf{r}} - \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}}\|^2_{L^2(Q)} \leq (\operatorname{tr}_{\Omega}(\bar{p} - \bar{P}^{\tau}_{\mathscr{T}_{\mathcal{Y}}}), \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}} - \bar{\mathsf{r}})_{L^2(Q)} + (\operatorname{tr}_{\Omega} \bar{P}^{\tau}_{\mathscr{T}_{\mathcal{Y}}} + \mu \bar{Z}^{\tau}_{\mathscr{T}_{\Omega}}, \Pi^{\mathcal{T}}_{\mathscr{T}_{\Omega}} \bar{\mathsf{r}} - \bar{\mathsf{r}})_{L^2(Q)},$$

where $\Pi^{\mathcal{T}}_{\mathscr{T}_{\Omega}}$ is defined in (6.3).

$\boxed{2}$ Using the solution $Q^\tau_{\mathscr{T}_\mathcal{Y}}$ to (6.9), we write $\bar{p}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}} = (\bar{p}-Q^\tau_{\mathscr{T}_\mathcal{Y}})+(Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}})$. The first term is estimated by using the result of Theorem 31 as follows:

$$\|\operatorname{tr}_\Omega(\bar{p}-Q^\tau_{\mathscr{T}_\mathcal{Y}})\|_{L^2(Q)} \lesssim \tau\left(\|\operatorname{tr}_\Omega \bar{v}\|_{L^\infty(0,T;L^2(\Omega))}+\|\mathsf{u}_d\|_{L^\infty(0,T;L^2(\Omega))}\right) + |\log N|^{2s}N^{-\frac{1+s}{n+1}}\left(\|\operatorname{tr}_\Omega \bar{v}\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))}+\|\mathsf{u}_d\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))}\right),$$

where we used that $\operatorname{tr}_\Omega \bar{v}\in L^\infty(0,T;L^2(\Omega))\cap L^2(0,T;\mathbb{H}^{1-s}(\Omega))$, which follows from Remarks 18 and 25.

$\boxed{3}$ To estimate the difference $\operatorname{tr}_\Omega(Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}})$, we write $Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}} = (Q^\tau_{\mathscr{T}_\mathcal{Y}}-R^\tau_{\mathscr{T}_\mathcal{Y}})+(R^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}})$, where $R^\tau_{\mathscr{T}_\mathcal{Y}}$ solves (6.10). Employing the stability estimate established in Theorem 26 (see also [49, Lemma 29]) and Theorem 31, we arrive at

$$\begin{aligned}\|\operatorname{tr}_\Omega(Q^\tau_{\mathscr{T}_\mathcal{Y}}-R^\tau_{\mathscr{T}_\mathcal{Y}})\|_{L^2(Q)} &\lesssim \|\operatorname{tr}_\Omega(\bar{v}-V^\tau_{\mathscr{T}_\mathcal{Y}}(\bar{\mathsf{r}}))\|_{L^2(Q)}\\ &\lesssim \tau\left(\|\bar{\mathsf{r}}\|_{L^\infty(0,T;L^2(\Omega))}+\|\mathsf{f}\|_{L^\infty(0,T;L^2(\Omega))}+\|\mathsf{u}_0\|_{\mathbb{H}^s(\Omega)}\right)\\ &\quad + |\log N|^{2s}N^{-\frac{1+s}{n+1}}\left(\|\bar{\mathsf{r}}\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))}+\|\mathsf{f}\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))}\right.\\ &\qquad\left.+\|\mathsf{u}_0\|_{\mathbb{H}^{1+s}(\Omega)}\right).\end{aligned}$$

The results of Theorem 17 and Remark 25 imply that the norms $\|\bar{\mathsf{r}}\|_{L^\infty(0,T;L^2(\Omega))}$ and $\|\bar{\mathsf{r}}\|_{L^2(0,T;\mathbb{H}^{1-s}(\Omega))}$ are uniformly controlled by the problem data.

To handle the term $R^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}}$ we invoke the discrete counterpart of step $\boxed{2}$ in Lemma 24, that is, an argument based on summation by parts, to arrive at

$$(\operatorname{tr}_\Omega(R^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}}),\bar{Z}^\tau_{\mathscr{T}_\Omega}-\bar{\mathsf{r}}^\tau)_{L^2(Q)}\leq 0.$$

$\boxed{4}$ Using the solutions to problems (6.9) and (6.10), we write

$$\begin{aligned}&(\operatorname{tr}_\Omega \bar{P}^\tau_{\mathscr{T}_\mathcal{Y}}+\mu\bar{Z}^\tau_{\mathscr{T}_\Omega},\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}\bar{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(Q)} = (\operatorname{tr}_\Omega \bar{p}+\mu\bar{\mathsf{r}},\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}\bar{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(Q)}\\ &+(\operatorname{tr}_\Omega(\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}}\pm Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{p}),\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}\bar{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(Q)}+\mu(\bar{Z}^\tau_{\mathscr{T}_\Omega}-\bar{\mathsf{r}},\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}\bar{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(Q)} = \mathrm{I}+\mathrm{II}+\mathrm{III}.\end{aligned}$$

Using the properties of the projector $\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}$ and the smoothness properties of $\bar{p}$ and $\bar{\mathsf{r}}$, derived in Theorem 17 and Remark 25, we have

$$\begin{aligned}\mathrm{I} = (\operatorname{tr}_\Omega \bar{p}+\mu\bar{\mathsf{r}}-\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}(\operatorname{tr}_\Omega \bar{p}+\mu\bar{\mathsf{r}}),\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}\bar{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(Q)} &\lesssim \big(\tau\|\operatorname{tr}_\Omega \bar{p}+\mu\bar{\mathsf{r}}\|_{H^1(0,T;L^2(\Omega))}\\ + h_{\mathscr{T}_\Omega}\|\operatorname{tr}_\Omega \bar{p}+\mu\bar{\mathsf{r}}\|_{L^2(0,T;H^1(\Omega))}\big)&(\tau\|\bar{\mathsf{r}}\|_{H^1(0,T;L^2(\Omega))}+h_{\mathscr{T}_\Omega}\|\bar{\mathsf{r}}\|_{L^2(0,T;H^1(\Omega))}).\end{aligned}$$

The term II can be handled by repeating the arguments of steps $\boxed{2}$ and $\boxed{3}$. We write $\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}}\pm Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{p} = (\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}}-R^\tau_{\mathscr{T}_\mathcal{Y}})+(R^\tau_{\mathscr{T}_\mathcal{Y}}-Q^\tau_{\mathscr{T}_\mathcal{Y}})+(Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{p})$. The stability result of Theorem 26 yields the following bound for the first term:

$$\left\|\operatorname{tr}_\Omega(R^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{P}^\tau_{\mathscr{T}_\mathcal{Y}})\right\|_{L^2(Q)} \lesssim \left\|\operatorname{tr}_\Omega(\bar{V}^\tau_{\mathscr{T}_\mathcal{Y}}-V^\tau_{\mathscr{T}_\mathcal{Y}}(\bar{\mathsf{r}}))\right\|_{L^2(Q)} \lesssim \|\bar{Z}^\tau_{\mathscr{T}_\Omega}-\bar{\mathsf{r}}\|_{L^2(Q)}.$$

The second term, $R^\tau_{\mathscr{T}_\mathcal{Y}}-Q^\tau_{\mathscr{T}_\mathcal{Y}}$, is estimated in step $\boxed{3}$, while the third one, i.e., $Q^\tau_{\mathscr{T}_\mathcal{Y}}-\bar{p}$, is estimated in step $\boxed{2}$. An application of the Cauchy–Schwarz and Young's inequalities yields the desired bound. The term III is controlled by a trivial application of the Cauchy–Schwarz inequality.

$\boxed{5}$ The assertion follows from collecting all the estimates we obtained in previous steps and recalling that $h_{\mathscr{T}_\Omega}\approx N^{-1/(n+1)}$. ∎

On the basis of of Lemma 34 we derive the following important result.

THEOREM 35 (control error estimates: $s \in (0,1)$ and $\gamma = 1$). *Let* $\bar{\mathsf{z}}$ *be the solution to the space-time fractional optimal control problem* (1.2)–(1.4)*, and let* $\bar{Z}^{\tau}_{\mathscr{T}_\Omega}$ *be the solution to the fully discrete optimal control problem of section* 6.1*. In the framework of Lemma* 34*, we have the following error estimate:*

$$\|\bar{\mathsf{z}} - \bar{Z}^{\tau}_{\mathscr{T}_\Omega}\|_{L^2(Q)} \lesssim \tau + |\log N|^{2s} N^{-\frac{1}{n+1}},$$

where the hidden constant is independent of the discretization parameters but depends on the problem data.

Proof. The result follows from Lemmas 24 and 34, in conjunction with an appropriate selection of the truncation parameter, $\mathcal{Y}$, as $\mathcal{Y} \approx |\log N|$ (see [48, Remark 5.5] for details). ∎

We conclude with an error estimate for the state in the $L^2(0,T;\mathbb{H}^s(\Omega))$-norm.

THEOREM 36 (state error estimates: $s \in (0,1)$ and $\gamma = 1$). *Let* $\bar{\mathsf{u}}$ *be the optimal state of the space-time fractional optimal control problem* (1.2)–(1.4)*, and let* $\bar{U}^{\tau}_{\mathscr{T}_\Omega}$ *be defined as in* (6.6)*. If, in addition to the conditions of Lemma* 34*, we assume that* $\mathsf{f} \in BV(0,T;L^2(\Omega))$*, then we have the following error estimate:*

$$\|\bar{\mathsf{u}} - \bar{U}^{\tau}_{\mathscr{T}_\Omega}\|_{L^2(0,T;\mathbb{H}^s(\Omega))} \lesssim \tau + |\log N|^{2s} N^{\frac{-1}{n+1}},$$

where the hidden constant is independent of the discretization parameters but depends on the problem data.

Proof. We first write

$$\|\bar{\mathsf{u}} - \bar{U}^{\tau}_{\mathscr{T}_\Omega}\|_{L^2(0,T;\mathbb{H}^s(\Omega))} \leq \|\operatorname{tr}_\Omega(\bar{\mathscr{U}} - \bar{v})\|_{L^2(0,T;\mathbb{H}^s(\Omega))} + \|\operatorname{tr}_\Omega \bar{v} - \bar{U}^{\tau}_{\mathscr{T}_\Omega}\|_{L^2(0,T;\mathbb{H}^s(\Omega))},$$

and note that, by using the trace inequality (2.11), the first term is controlled in (4.10). The second term is handled by noticing that

$$\begin{aligned}\|\operatorname{tr}_\Omega \bar{v} - \bar{U}^{\tau}_{\mathscr{T}_\Omega}\|_{L^2(0,T;\mathbb{H}^s(\Omega))} &= \|\operatorname{tr}_\Omega(\bar{v} - \bar{V}^{\tau}_{\mathscr{T}_\mathcal{Y}})\|_{L^2(0,T;\mathbb{H}^s(\Omega))} \\ &\leq \|\operatorname{tr}_\Omega(\bar{v} - V^{\tau}_{\mathscr{T}_\mathcal{Y}}(\bar{\mathsf{r}}))\|_{L^2(0,T;\mathbb{H}^s(\Omega))} + \|\operatorname{tr}_\Omega(V^{\tau}_{\mathscr{T}_\mathcal{Y}}(\bar{\mathsf{r}}) - \bar{V}^{\tau}_{\mathscr{T}_\mathcal{Y}})\|_{L^2(0,T;\mathbb{H}^s(\Omega))}.\end{aligned}$$

Since $\mathsf{f} + \bar{\mathsf{r}} \in BV(0,T;L^2(\Omega))$, the first term on the right-hand side of this inequality is estimated using the error estimates for the discrete scheme presented in [49, Theorem 33]. The second can be handled by invoking the stability of the discrete scheme and the error estimates of Theorem 35. Collecting these bounds, we obtain the result. ∎

6.3. Convergence. Let us now consider the case when either $\gamma, s \in (0,1)$ or the problem data is not smooth enough to yield the error estimates of section 6.2 and elucidate the general convergence properties of the fully discrete scheme. Notice that we not only are approximating the state equation via discretization, but also are approximating the cost, so convergence of discrete optimal controls to the continuous control is not immediate. To avoid irrelevant technicalities let us here assume that $\mathsf{f} \equiv 0$ and $\mathsf{u}_0 \equiv 0$. Having them nonzero would imply changing the function u_d in the discussion that follows. In this setting $\mathbf{S}$ is linear. To begin, as in Definition 8, we introduce the discrete control to state operator

$$\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}} : L^2(Q) \ni z \mapsto \operatorname{tr}_\Omega V^{\tau}_{\mathscr{T}_\mathcal{Y}} \subset \mathbb{U}(\mathscr{T}_\Omega), \tag{6.11}$$

where $V^{\tau}_{\mathscr{T}_{\mathcal{Y}}}$ solves (5.7) with f replaced by $\Pi^{\mathcal{T}} z$ and where $\Pi^{\mathcal{T}}$ is the $L^2(0,T)$-orthogonal projection onto piecewise constants. Notice that the stability estimates of Theorem 26 combined with the uniform boundedness principle yield that for all $\gamma, s \in (0,1)$ the family $\{\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_{\mathcal{Y}}}\}$ is uniformly bounded in $\mathcal{B}(L^2(Q))$. Moreover, since the error estimates are valid for right-hand sides that range over a dense subset of $L^2(Q)$, we obtain the pointwise convergence of these operators to $\mathbf{S}$; see [23, Proposition 5.17].

We also introduce $\mathbf{S}^{\mathcal{T},\star}_{\mathscr{T}_{\mathcal{Y}}}$, the adjoint of $\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_{\mathcal{Y}}}$. A simple but tedious calculation, shown in Appendix A, reveals that for $\zeta \in L^2(Q)$ we have that

$$\mathbf{S}^{\mathcal{T},\star}_{\mathscr{T}_{\mathcal{Y}}}\zeta = \operatorname{tr}_\Omega P^{\tau}_{\mathscr{T}_{\mathcal{Y}}} \subset \mathbb{U}(\mathscr{T}_\Omega),$$

where $P^{\tau}_{\mathscr{T}_{\mathcal{Y}}}$ solves

$$(\operatorname{tr}_\Omega \bar{\delta}^\gamma P^k_{\mathscr{T}_{\mathcal{Y}}}, \operatorname{tr}_\Omega W)_{L^2(\Omega)} + a_{\mathcal{Y}}(P^k_{\mathscr{T}_{\mathcal{Y}}}, W) = \left(\Pi^{\mathcal{T}}\zeta^k, \operatorname{tr}_\Omega W\right)_{L^2(\Omega)} \quad \forall W \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}}), \tag{6.12}$$

with $P^{\mathcal{K}}_{\mathscr{T}_{\mathcal{Y}}} = 0$. Here $\bar{\delta}^\gamma P^k_{\mathscr{T}_{\mathcal{Y}}}$ is defined, for $\gamma < 1$, by

$$\bar{\delta}^\gamma P^k_{\mathscr{T}_{\mathcal{Y}}} = \frac{1}{\Gamma(2-\gamma)} \sum_{i=k}^{\mathcal{K}-1} \frac{a_{i-k}}{\tau^{\gamma-1}} \bar{\delta}^1 P^i_{\mathscr{T}_{\mathcal{Y}}},$$

where the coefficients a_j are defined in section 5.1.

Problem (6.12) is a stable and consistent approximation of the backward fractional parabolic equation: Find $p \in \mathbb{V}$ such that $\operatorname{tr}_\Omega p(T) = 0$ and for a.e. $t \in (0,T)$

$$\langle \operatorname{tr}_\Omega \partial^\gamma_{T-t} p, \operatorname{tr}_\Omega \phi\rangle + a_{\mathcal{Y}}(\phi, p) = \langle \zeta, \operatorname{tr}_\Omega \phi\rangle \quad \forall \phi \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathcal{Y}}).$$

Consequently, we also have that the family $\{\mathbf{S}^{\mathcal{T},\star}_{\mathscr{T}_{\mathcal{Y}}}\}$ is uniformly bounded in $\mathcal{B}(L^2(Q))$ and pointwise convergent. These observations will be crucial in showing convergence.

With the discrete control to state operator at hand, as in (3.2), we define the reduced cost functional by

$$\begin{aligned} F^{\mathcal{T}}_{\mathscr{T}_\Omega}(Z^\tau_{\mathscr{T}_\Omega}) &= J^{\mathcal{T}}_{\mathscr{T}_\Omega}(\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_{\mathcal{Y}}} Z^\tau_{\mathscr{T}_\Omega}, Z^\tau_{\mathscr{T}_\Omega}) \\ &= \frac{1}{2}\|\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_{\mathcal{Y}}} Z^\tau_{\mathscr{T}_\Omega} - \mathsf{u}^\tau_d\|^2_{\ell^2(L^2(\Omega))} + \frac{\mu}{2}\|Z^\tau_{\mathscr{T}_\Omega}\|^2_{\ell^2(L^2(\Omega))}. \end{aligned} \tag{6.13}$$

The convergence of the fully discrete scheme under minimal regularity assumptions is the content of the next result.

THEOREM 37 (convergence). *The family $\{\bar{Z}^\tau_{\mathscr{T}_\Omega}\}_{\mathscr{T}_\Omega \in \mathbb{T}_\Omega, \tau > 0}$ is uniformly bounded and contains a subsequence that converges in $L^2(Q)$ to $\bar{\mathsf{r}}$, the solution to the truncated optimal control problem.*

Proof. Boundedness follows immediately from the fact that $\bar{Z}^\tau_{\mathscr{T}_\Omega}$ minimizes $F^{\mathcal{T}}_{\mathscr{T}_{\mathcal{Y}}}$. If $z_0 \in \mathsf{Z}_{\mathrm{ad}}$, then

$$F^{\mathcal{T}}_{\mathscr{T}_\Omega}(\bar{Z}^\tau_{\mathscr{T}_\Omega}) \leq F^{\mathcal{T}}_{\mathscr{T}_\Omega}(\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega} z_0) \lesssim \|z_0\|^2_{L^2(Q)} + \|\mathsf{u}_d\|^2_{L^2(Q)},$$

where we used the uniform boundedness of $\Pi^{\mathcal{T}}$, $\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega}$, and $\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_{\mathcal{Y}}}$. This implies the existence of a (not relabeled) weakly convergent subsequence.

To show convergence of this subsequence to $\bar{\mathsf{r}}$ we appeal to the theory of Γ-convergence, for which we need to verify several assumptions:

$\boxed{1}$ *Lower bound inequality.* Assume that $Z^{\tau}_{\mathscr{T}_\Omega} \rightharpoonup z$ in $L^2(Q)$. For $w \in L^2(Q)$, we have

$$(\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}} Z^{\tau}_{\mathscr{T}_\Omega} - \mathbf{S}z, w)_{L^2(Q)} = (\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}} z - \mathbf{S}z, w)_{L^2(Q)} + (\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}} (Z^{\tau}_{\mathscr{T}_\Omega} - z), w)_{L^2(Q)} = \mathrm{I} + \mathrm{II}.$$

The pointwise convergence of $\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}}$ to $\mathbf{S}$ shows that $\mathrm{I} \to 0$, while the the pointwise convergence of $\mathbf{S}^{\mathcal{T},\star}_{\mathscr{T}_\mathcal{Y}}$ shows that $\mathrm{II} = (Z^{\tau}_{\mathscr{T}_\Omega} - z, \mathbf{S}^{\mathcal{T},\star}_{\mathscr{T}_\mathcal{Y}} w)_{L^2(Q)} \to 0$. In conclusion, $\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}} Z^{\tau}_{\mathscr{T}_\Omega} \rightharpoonup \mathbf{S}z$. Lower semicontinuity of the norms and $\mathsf{u}^{\tau}_d \to \mathsf{u}_d$ in $L^2(Q)$ imply

$$f(z) \le \liminf F^{\mathcal{T}}_{\mathscr{T}_\Omega}(Z^{\tau}_{\mathscr{T}_\Omega}),$$

which is what we needed to show.

$\boxed{2}$ *Existence of a recovery sequence.* Let $z \in \mathsf{Z}_{\mathrm{ad}}$; then $\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega} z \in \mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega)$ converges strongly to z in $L^2(Q)$. Consequently, $\mathbf{S}^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}} \Pi^{\mathcal{T}}_{\mathscr{T}_\Omega} z \to \mathbf{S}z$ in $L^2(Q)$ as well. The continuity of $F^{\mathcal{T}}_{\mathscr{T}_\Omega}$ then implies

$$f(z) \ge \limsup F^{\mathcal{T}}_{\mathscr{T}_\Omega}(\Pi^{\mathcal{T}}_{\mathscr{T}_\Omega} z).$$

$\boxed{3}$ *Equicoerciveness.* Since

$$F^{\mathcal{T}}_{\mathscr{T}_\Omega}(Z) \ge \frac{\mu}{2}\|Z\|^2_{L^2(Q)} \quad \forall Z \in \mathsf{Z}_{\mathrm{ad}}(\mathcal{T}, \mathscr{T}_\Omega),$$

we have, by [18, Proposition 7.7], that the family $\{F^{\mathcal{T}}_{\mathscr{T}_\Omega}\}$ is equicoercive.

Assumptions $\boxed{1}$ and $\boxed{2}$ show the Γ-convergence of the discrete reduced costs $F^{\mathcal{T}}_{\mathscr{T}_\Omega}$ to the reduced cost f. This implies, using [18, Corollary 7.20], that minimizers of $F^{\mathcal{T}}_{\mathscr{T}_\mathcal{Y}}$, if they converge, must do so to a minimizer of f. Assumption $\boxed{3}$ and the uniqueness of the minimizer of the reduced cost f are the conditions for the fundamental lemma of Γ-convergence [18, Corollary 7.24]. In conclusion, $\{\bar{Z}^{\tau}_{\mathscr{T}_\Omega}\}$ converges weakly to $\bar{\mathsf{r}}$, the minimum of the truncated cost functional.

We conclude with the strong convergence. The weak convergence of $\bar{Z}^{\tau}_{\mathscr{T}_\Omega}$ to $\bar{\mathsf{r}}$ implies, using [18, equation (7.32)], that $F^{\mathcal{T}}_{\mathscr{T}_\Omega}(\bar{Z}^{\tau}_{\mathscr{T}_\Omega}) \to f(\bar{\mathsf{r}})$. Consequently, we have

$$\begin{aligned} &\frac{1}{2}\left\|\mathbf{S}^{\tau}_{\mathscr{T}_\mathcal{Y}} \bar{Z}^{\tau}_{\mathscr{T}_\Omega} - \mathbf{S}\bar{\mathsf{r}}\right\|^2_{L^2(Q)} + \frac{\mu}{2}\|\bar{Z}^{\tau}_{\mathscr{T}_\Omega} - \bar{\mathsf{r}}\|^2_{L^2(Q)} \\ &= F^{\mathcal{T}}_{\mathscr{T}_\Omega}(\bar{Z}^{\tau}_{\mathscr{T}_\Omega}) + f(\bar{\mathsf{r}}) - \left(\mathbf{S}^{\tau}_{\mathscr{T}_\mathcal{Y}} \bar{Z}^{\tau}_{\mathscr{T}_\Omega}, \mathbf{S}\bar{\mathsf{r}} - \mathsf{u}^{\tau}_d\right)_{L^2(Q)} + (\mathsf{u}_d, \mathbf{S}\bar{\mathsf{r}} - \mathsf{u}^{\tau}_d)_{L^2(Q)} - \mu(\bar{Z}^{\tau}_{\mathscr{T}_\Omega}, \bar{\mathsf{r}})_{L^2(Q)} \\ &\qquad \to f(\bar{\mathsf{r}}) + f(\bar{\mathsf{r}}) - 2f(\bar{\mathsf{r}}) = 0, \end{aligned}$$

where the passage to the limit can be obtained by observing that, since $\mathbf{S}^{\tau,\star}_{\mathscr{T}_\mathcal{Y}}$ converges pointwise,

$$\left(\mathbf{S}^{\tau}_{\mathscr{T}_\mathcal{Y}} \bar{Z}^{\tau}_{\mathscr{T}_\Omega}, \mathbf{S}\bar{\mathsf{r}} - \mathsf{u}^{\tau}_d\right)_{L^2(Q)} = \left(\bar{Z}^{\tau}_{\mathscr{T}_\Omega}, \mathbf{S}^{\tau,\star}_{\mathscr{T}_\mathcal{Y}}(\mathbf{S}\bar{\mathsf{r}} - \mathsf{u}^{\tau}_d)\right)_{L^2(Q)} \to \left(\bar{\mathsf{r}}, \mathbf{S}^{\star}(\mathbf{S}\bar{\mathsf{r}} - \mathsf{u}_d)\right)_{L^2(Q)}.$$

This concludes the proof. ∎

7. Numerical experiments. Let us illustrate the performance of the fully discrete scheme proposed in section 6.1 for $\gamma = 1$ and the error estimates derived in section 6.2.

7.1. Implementation. The implementation has been developed in MATLAB. The stiffness and mass matrices of the discrete system (6.5) are assembled exactly, and the respective forcing boundary terms are computed by a quadrature formula which is exact for polynomials of degree 4. The resulting linear system is solved by using the built-in *direct solver* of MATLAB. To solve the minimization problem, we use the projected Broyden–Fletcher–Goldfrab–Shanno (BFGS) method with Armijo line search; see [33]. The optimization algorithm is terminated when the ℓ^2-norm of the projected gradient is less than or equal to 10^{-9}. Convergence rates in space (cf. section 7.2) are obtained by using a desktop computer with a 2.5 GHz Intel Core i5 processor and 20 GB of memory. Convergence rates in time (cf. section 7.3) are obtained using ARGO, which is a research computing cluster provided by the Office of Research Computing at George Mason University (see http://orc.gmu.edu) and XSEDE [57].

To illustrate the error estimates of section 6.2 we need an exact solution to the fractional control problem (1.2)–(1.4). Let $n = 2$, $\mu = 1$, $\Omega = (0,1)^2$, and $\mathcal{L} = -\Delta$. In this setting, the eigenpairs of $\mathcal{L}$ are

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

Set $\bar{\mathsf{u}} = e^t \sin(2\pi x_1')\sin(2\pi x_2')$, which yields $\mathsf{f} = (1 + \lambda_{2,2}^s)e^t \sin(2\pi x_1')\sin(2\pi x_2') - \bar{\mathsf{z}}$. Set also $\bar{\mathsf{p}} = -\mu(T-t)e^t \sin(2\pi x_1')\sin(2\pi x_2')$. Definition 11 then yields $\mathsf{u}_d = \big[1 - \mu\{-1 + (1 - \lambda_{2,2}^s)(T-t)\}\big]e^t \sin(2\pi x_1')\sin(2\pi x_2')$. Finally, we set $\mathsf{a} = 0$ and $\mathsf{b} = 0.5$. The projection formula (3.21) gives the value of $\bar{\mathsf{z}}$. This defines, for any $s \in (0,1)$, the data and solution to the optimal control problem (1.2)–(1.4).

7.2. Convergence rates in space. Let $\mathcal{K} = 1400$. The experimental rates of convergence for the optimal control and state are shown in Figure 1. The top left panel illustrates the rate of convergence for the optimal control in the $L^2(0,T;L^2(\Omega))$-norm, and the top right panel that for the optimal state in the $L^2(0,T;\mathbb{H}^s(\Omega))$-norm with respect to the number of degrees of freedom N for all choices of the parameter s considered. We see that the rate for the optimal control is optimal and that for the state agrees with our theory. As noted in [5, 49], in order to recover optimality for the control, the state and adjoint equations must be discretized with the anisotropic refinement, in the extended dimension, dictated by (5.4).

The bottom panel of Figure 1 illustrates that the approximate optimal state in the $L^2(0,T;L^2(\Omega))$-norm converges with a rate $N^{-\frac{2}{3}}$. This rate is not discussed in this paper and will be part of a future work. We remark that since the theoretical rate of convergence for the approximation of the state without optimization is, according to Corollary 32, $N^{-\frac{1+s}{3}}$, we expect the same to hold in this case.

7.3. Convergence rates in time. Let $N = 927828$. The asymptotic relation

$$\|\bar{\mathsf{z}} - \bar{Z}_{\mathscr{T}_\Omega}^{\tau}\|_{L^2(0,T;L^2(\Omega))} \approx \mathcal{K}^{-1}$$

is shown in the left panel of Figure 2 and illustrates the optimal decay rate in the control with respect to $\mathcal{K}$ for all choices of the parameter s considered. We observe that the convergence rate deteriorates as τ becomes smaller. This is due to the fact that the space discretization error saturates the total error. To explore this saturation effect computationally, in Figure 2(right) we have also displayed the error decay when $N = 1764909$. This shows a clear improvement over the case $N = 927828$.

Appendix A. The adjoint of the operator $\mathbf{S}_{\mathscr{T}_\mathcal{Y}}^{\mathcal{T}}$. Here we compute $\mathbf{S}_{\mathscr{T}_\mathcal{Y}}^{\mathcal{T},\star}$, the $L^2(Q)$-adjoint of the discrete control to state operator. Recall that $\mathbf{S}_{\mathscr{T}_\mathcal{Y}}^{\mathcal{T}} z = \operatorname{tr}_\Omega V_{\mathscr{T}_\mathcal{Y}}^{\tau}$,

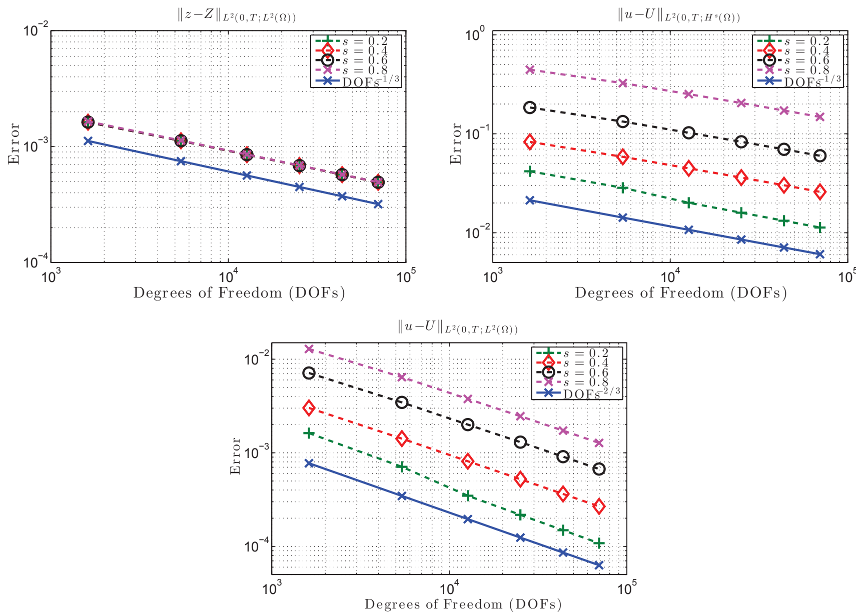

FIG. 1. *Computational rates of convergence for the control and state on anisotropic meshes for $n = 2$ and $s = 0.2$, 0.4, 0.6, and 0.8. For a fixed number of time steps, $\mathcal{K} = 1400$, the top left panel shows the decrease of the $L^2(0,T;L^2(\Omega))$-control error with respect to N, and the top right one that of the $L^2(0,T;\mathbb{H}^s(\Omega))$-state error. In both cases we recover a rate of $N^{-1/3}$. The bottom panel illustrates the decrease of the $L^2(0,T;L^2(\Omega))$-state error, where we observe a rate of $N^{-2/3}$. This last, while heuristically expected (due to Corollary 32), is beyond the scope of our theory.*

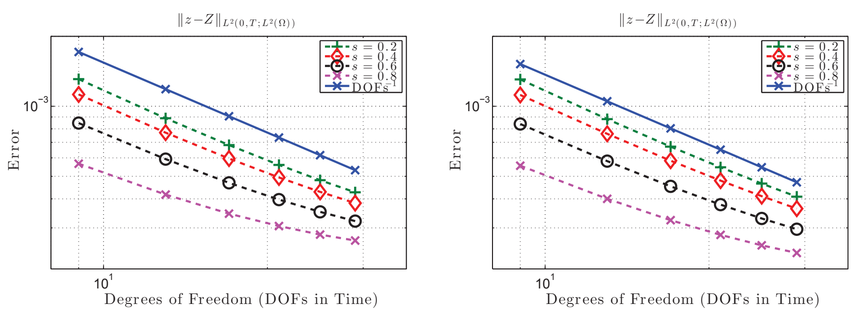

FIG. 2. *Computational rates of convergence for the control on anisotropic meshes for $n = 2$ and $s = 0.2$, 0.4, 0.6, and 0.8. For a fixed number of degrees of freedom in space, $N = 927828$, the left panel shows the decrease of the $L^2(0,T;L^2(\Omega))$-control error with respect to $\mathcal{K}$. The right panel shows that when the number of degrees of freedom is increased to $N = 1764909$. In both cases we recover the rate $\mathcal{K}^{-1}$. We notice that for $N = 927828$ and small time steps the displayed error deteriorates. This is attributed to the saturation effect caused by space discretization error. An improvement is shown in the right panel when $N = 1764909$.*

where $V_{\mathscr{T}_\mathcal{Y}}^\tau$ is defined as the solution to

$$\text{(A.1)} \qquad \left(\operatorname{tr}_\Omega \delta^\gamma V_{\mathscr{T}_\mathcal{Y}}^{k+1}, \operatorname{tr}_\Omega W\right)_{L^2(\Omega)} + a_\mathcal{Y}(V_{\mathscr{T}_\mathcal{Y}}^{k+1}, W) = \left(\Pi^\mathcal{T} z^{k+1}, \operatorname{tr}_\Omega W\right)_{L^2(\Omega)},$$

with $V^0_{\mathscr{T}_\mathscr{Y}} = 0$. Multiply (A.1) by τ and add over k. Using the identification of a sequence ϕ^τ with a piecewise constant (in time) function, we obtain

$$\text{(A.2)} \quad \int_0^T \left(\operatorname{tr}_\Omega \delta^\gamma V^\tau_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega W^\tau_V\right)_{L^2(\Omega)} \,\mathrm{d}t + \int_0^T a_\mathscr{Y}(V^\tau_{\mathscr{T}_\mathscr{Y}}, W^\tau_V)\,\mathrm{d}t = \int_0^T (z, \operatorname{tr}_\Omega W^\tau_V)_{L^2(\Omega)} \,\mathrm{d}t$$

for every W^τ_V, which takes values on $\mathbb{V}(\mathscr{T}_\mathscr{Y})$ and, on each time interval $(t_k, t_{k+1}]$, is constant.

Define, for $\zeta \in L^2(Q)$, the sequence $P^\tau_{\mathscr{T}_\mathscr{Y}} \subset \mathbb{V}(\mathscr{T}_\mathscr{Y})$ as $P^\mathcal{K}_{\mathscr{T}_\mathscr{Y}} = 0$ and, for $k = \mathcal{K}-1, \dots, 0$, set $P^k_{\mathscr{T}_\mathscr{Y}}$ as the solution of

$$\left(\operatorname{tr}_\Omega \bar{\delta}^\gamma P^k_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega W\right)_{L^2(\Omega)} + a_\mathscr{Y}(P^k_{\mathscr{T}_\mathscr{Y}}, W) = \left(\Pi^\mathcal{T} \zeta^k, \operatorname{tr}_\Omega W\right)_{L^2(\Omega)}.$$

Multiply this identity by τ and add over k. Using, again, the identification between sequences and piecewise constant (in time) functions, we obtain

$$\text{(A.3)} \quad \int_0^T \left(\operatorname{tr}_\Omega \bar{\delta}^\gamma P^\tau_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega W^\tau_P\right)_{L^2(\Omega)} \,\mathrm{d}t + \int_0^T a_\mathscr{Y}(P^\tau_{\mathscr{T}_\mathscr{Y}}, W^\tau_P)\,\mathrm{d}t = \int_0^T (\zeta, \operatorname{tr}_\Omega W^\tau_P)_{L^2(\Omega)} \,\mathrm{d}t.$$

Set $W^\tau_V = P^\tau_{\mathscr{T}_\mathscr{Y}}$ in (A.2) and $W^\tau_P = V^\tau_{\mathscr{T}_\mathscr{Y}}$ in (A.3), respectively, and subtract the equations. This yields

$$\text{(A.4)} \quad \left(\zeta, \operatorname{tr}_\Omega V^\tau_{\mathscr{T}_\mathscr{Y}}\right)_{L^2(Q)} - \left(z, \operatorname{tr}_\Omega P^\tau_{\mathscr{T}_\mathscr{Y}}\right)_{L^2(Q)} = \int_0^T \left(\operatorname{tr}_\Omega \delta^\gamma V^\tau_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega P^\tau_{\mathscr{T}_\mathscr{Y}}\right)_{L^2(\Omega)} \,\mathrm{d}t - \int_0^T \left(\operatorname{tr}_\Omega \bar{\delta}^\gamma P^\tau_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega V^\tau_{\mathscr{T}_\mathscr{Y}}\right)_{L^2(\Omega)} \,\mathrm{d}t.$$

Recalling that $\operatorname{tr}_\Omega V^\tau_{\mathscr{T}_\mathscr{Y}} = \mathbf{S}^\mathcal{T}_{\mathscr{T}_\mathscr{Y}} z$, we see that if the right-hand side of the identity above were zero, we would obtain that $\mathbf{S}^{\mathcal{T},\star}_{\mathscr{T}_\mathscr{Y}} \zeta = \operatorname{tr}_\Omega P^\tau_{\mathscr{T}_\mathscr{Y}}$. We accomplish this by proving a sort of fractional summation-by-parts formula.

Let us now compute the right-hand side of (A.4) explicitly, using the fact $V^0_{\mathscr{T}_\mathscr{Y}} = P^\mathcal{K}_{\mathscr{T}_\mathscr{Y}} = 0$. We have

$$\mathcal{D} := \int_0^T \left(\operatorname{tr}_\Omega \delta^\gamma V^\tau_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega P^\tau_{\mathscr{T}_\mathscr{Y}}\right)_{L^2(\Omega)} \,\mathrm{d}t = \tau \sum_{k=0}^{\mathcal{K}-1} (\operatorname{tr}_\Omega \delta^\gamma V^{k+1}_{\mathscr{T}_\mathscr{Y}}, \operatorname{tr}_\Omega P^k_{\mathscr{T}_\mathscr{Y}})_{L^2(\Omega)} = \frac{\tau}{\Gamma(2-\gamma)} \sum_{k=0}^{\mathcal{K}-1} \sum_{i=0}^{k} a_i \left(\frac{V^{k+1-i}_{\mathscr{T}_\mathscr{Y}} - V^{k-i}_{\mathscr{T}_\mathscr{Y}}}{\tau^\gamma}, \operatorname{tr}_\Omega P^k_{\mathscr{T}_\mathscr{Y}} \right)_{L^2(\Omega)}.$$

Expanding the sums yields that $\tilde{\mathcal{D}} = \frac{\Gamma(2-\gamma)}{\tau^{1-\gamma}}\mathcal{D}$ verifies

$$\begin{aligned}\tilde{\mathcal{D}} := {} & a_0(V^1_{\mathscr{T}_\gamma} - V^0_{\mathscr{T}_\gamma}, P^0_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & + a_0(V^2_{\mathscr{T}_\gamma} - V^1_{\mathscr{T}_\gamma}, P^1_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_1(V^1_{\mathscr{T}_\gamma} - V^0_{\mathscr{T}_\gamma}, P^1_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & + a_0(V^3_{\mathscr{T}_\gamma} - V^2_{\mathscr{T}_\gamma}, P^2_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_1(V^2_{\mathscr{T}_\gamma} - V^1_{\mathscr{T}_\gamma}, P^2_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_2(V^1_{\mathscr{T}_\gamma} - V^0_{\mathscr{T}_\gamma}, P^2_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & \vdots \\ & + a_0(V^{\mathcal{K}}_{\mathscr{T}_\gamma} - V^{\mathcal{K}-1}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_1(V^{\mathcal{K}-1}_{\mathscr{T}_\gamma} - V^{\mathcal{K}-2}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma})_{L^2(\Omega)} + \cdots \\ & \quad + a_{\mathcal{K}-1}(V^1_{\mathscr{T}_\gamma} - V^0_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma})_{L^2(\Omega)}.\end{aligned}$$

Rearranging terms and using that $V^0_{\mathscr{T}_\gamma} = P^{\mathcal{K}}_{\mathscr{T}_\gamma} = 0$ yields

$$\begin{aligned}\tilde{\mathcal{D}} = {} & a_0(V^1_{\mathscr{T}_\gamma}, P^0_{\mathscr{T}_\gamma} - P^1_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_1(V^1_{\mathscr{T}_\gamma}, P^1_{\mathscr{T}_\gamma} - P^2_{\mathscr{T}_\gamma})_{L^2(\Omega)} + \cdots \\ & \quad + a_{\mathcal{K}-1}(V^1_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma} - P^{\mathcal{K}}_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & \vdots \\ & + a_0(V^{\mathcal{K}-2}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-3}_{\mathscr{T}_\gamma} - P^{\mathcal{K}-2}_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_1(V^{\mathcal{K}-2}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-2}_{\mathscr{T}_\gamma} - P^{\mathcal{K}-1}_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & \quad + a_2(V^{\mathcal{K}-2}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma} - P^{\mathcal{K}}_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & + a_0(V^{\mathcal{K}-1}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-2}_{\mathscr{T}_\gamma} - P^{\mathcal{K}-1}_{\mathscr{T}_\gamma})_{L^2(\Omega)} + a_1(V^{\mathcal{K}-1}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma} - P^{\mathcal{K}}_{\mathscr{T}_\gamma})_{L^2(\Omega)} \\ & + a_0(V^{\mathcal{K}}_{\mathscr{T}_\gamma}, P^{\mathcal{K}-1}_{\mathscr{T}_\gamma} - P^{\mathcal{K}}_{\mathscr{T}_\gamma})_{L^2(\Omega)}.\end{aligned}$$

This shows that

$$\tilde{\mathcal{D}} = \sum_{k=0}^{\mathcal{K}-1}\sum_{i=k}^{\mathcal{K}-1} a_{i-k}(V^{k+1}_{\mathscr{T}_\gamma}, P^i_{\mathscr{T}_\gamma} - P^{i+1}_{\mathscr{T}_\gamma})_{L^2(\Omega)}.$$

Consequently,

$$\int_0^T \left(\operatorname{tr}_\Omega \delta^\gamma V^\tau_{\mathscr{T}_\gamma}, \operatorname{tr}_\Omega P^\tau_{\mathscr{T}_\gamma}\right)_{L^2(\Omega)} \mathrm{d}t - \int_0^T \left(\operatorname{tr}_\Omega \bar{\delta}^\gamma P^\tau_{\mathscr{T}_\gamma}, \operatorname{tr}_\Omega V^\tau_{\mathscr{T}_\gamma}\right)_{L^2(\Omega)} \mathrm{d}t = 0,$$

which is what we needed to show.

Let us conclude by observing that $\bar{\delta}^\gamma$ is a stable and consistent approximation of the right Caputo derivative ∂^γ_{T-t}, defined in (2.2). This can be obtained by following *mutatis mutandis* the arguments of [49, sections 3.2–3.4].

Acknowledgments. We would like to thank the referees for their insightful comments and suggestions. We also wish to thank T. Stephens for his technical support.

REFERENCES

[1] S. Abe and S. Thurner, *Anomalous diffusion in view of Einstein's 1905 theory of Brownian motion*, Phys. A, 356 (2005), pp. 403–407, doi:10.1016/ j.physa.2005.03.035; available online at http://www.sciencedirect.com/science/article/pii/S0378437105003432.

[2] G. Acosta and J. Borthagaray, *A Fractional Laplace Equation: Regularity of Solutions and Finite Element Approximations*, preprint, arXiv:1507.08970, 2015.

[3] O. Agrawal, *A general formulation and solution scheme for fractional optimal control problems*, Nonlinear Dynam., 38 (2004), pp. 323–337, doi:10.1023/B:NODY. 0000045544.96418.bf.

[4] O. Agrawal, *A general finite element formulation for fractional variational problems*, J. Math. Anal. Appl., 337 (2008), pp. 1–12, doi:10.1016/j.jmaa.2007.03.105.

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

[6] T. M. Atanacković, S. Pilipović, B. Stanković, and D. Zorica, *Fractional Calculus with Applications in Mechanics*, Mech. Eng. Solid Mech. Ser., John Wiley & Sons, New York, 2014.

[7] D. Baleanu, O. Defterli, and O. Agrawal, *A central difference numerical scheme for fractional optimal control problems*, J. Vib. Control, 15 (2009), pp. 583–597, doi:10.1177/1077546308088565.

[8] E. Barkai, R. Metzler, and J. Klafter, *From continuous time random walks to the fractional Fokker-Planck equation*, Phys. Rev. E (3), 61 (2000), pp. 132–138, doi:10.1103/PhysRevE.61.132.

[9] A. Bermúdez, D. Gómez, R. Rodríguez, and P. Venegas, *Numerical analysis of a transient non-linear axisymmetric eddy current model*, Comput. Math. Appl., 70 (2015), pp. 1984–2005, doi:10.1016/j.camwa.2015.08.017.

[10] A. Bonito and J. Pasciak, *Numerical approximation of fractional powers of elliptic operators*, Math. Comp., 84 (2015), pp. 2083–2110, doi:10.1090/S0025-5718-2015-02937-8.

[11] J.-P. Bouchaud and A. Georges, *Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications*, Phys. Rep., 195 (1990), pp. 127–293, doi:10.1016/0370-1573(90)90099-N.

[12] C. Bucur and E. Valdinoci, *Nonlocal Diffusion and Applications*, Springer, New York, 2016.

[13] A. Bueno-Orovio, D. Kay, V. Grau, B. Rodriguez, and K. Burrage, *Fractional diffusion models of cardiac electrical propagation: Role of structural heterogeneity in dispersion of repolarization*, J. R. Soc. Interface, 11 (2014), doi:10.1098/rsif.2014.0352.

[14] L. Caffarelli and L. Silvestre, *An extension problem related to the fractional Laplacian*, Comm. Partial Differential Equations, 32 (2007), pp. 1245–1260, doi:10.1080/03605300600987306.

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

[16] W. Chen, *A speculative study of 2/3-order fractional Laplacian modeling of turbulence: Some thoughts and conjectures*, Chaos, 16 (2006), 023126, doi:10.1063/1.2208452.

[17] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Appl. Math. 40, SIAM, Philadelphia, 2002, doi:10.1137/1.9780898719208.

[18] G. Dal Maso, *An Introduction to Γ-Convergence*, Birkhäuser Boston, Cambridge, MA, 1993, doi:10.1007/978-1-4612-0327-8.

[19] L. Debnath, *Fractional integral and fractional differential equations in fluid mechanics*, Fract. Calc. Appl. Anal., 6 (2003), pp. 119–155.

[20] L. Debnath, *Recent applications of fractional calculus to science and engineering*, Int. J. Math. Math. Sci., 2003 (2003), pp. 3413–3442, doi:10.1155/S0161171203301486.

[21] O. Defterli, *A numerical scheme for two-dimensional optimal control problems with memory effect*, Comput. Math. Appl., 59 (2010), pp. 1630–1636, doi:10.1016/j.camwa.2009.08.005.

[22] D. del Castillo-Negrete, B. A. Carreras, and V. E. Lynch, *Fractional diffusion in plasma turbulence*, Phys. Plasmas, 11 (2004), pp. 3854–3864, doi:10.1063/1.1767097.

[23] G. Folland, *Real Analysis*, John Wiley & Sons, New York, 1984.

[24] D. Fujiwara, *Concrete characterization of the domains of fractional powers of some elliptic differential operators of the second order*, Proc. Japan Acad., 43 (1967), pp. 82–86.

[25] P. Gatto and J. Hesthaven, *Numerical approximation of the fractional Laplacian via hp-finite elements, with an application to image denoising*, J. Sci. Comput., 65 (2015), pp. 249–270, doi:10.1007/s10915-014-9959-1.

[26] V. Gol′dshtein and A. Ukhlov, *Weighted Sobolev spaces and embedding theorems*, Trans. Amer. Math. Soc., 361 (2009), pp. 3829–3850, doi:10.1090/S0002-9947-09-04615-7.

[27] R. Gorenflo, F. Mainardi, D. Moretti, and P. Paradisi, *Time fractional diffusion: A discrete random walk approach*, Nonlinear Dynam., 29 (2002), pp. 129–143, doi:10.1023/A:1016547232119.

[28] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monogr. Stud. Math. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[29] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Math. Model. Theory Appl. 23, Springer, New York, 2009.

[30] M. Hinze and F. Tröltzsch, *Discrete concepts versus error analysis in PDE-constrained optimization*, GAMM-Mitt., 33 (2010), pp. 148–162, doi:10.1002/gamm.201010012.

[31] R. Ishizuka, S.-H. Chong, and F. Hirata, *An integral equation theory for inhomogeneous molecular fluids: The reference interaction site model approach*, J. Chem. Phys, 128 (2008), 034504, doi:10.1063/1.2819487.
[32] B. Jin, R. Lazarov, Y. Liu, and Z. Zhou, *The Galerkin finite element method for a multi-term time-fractional diffusion equation*, J. Comput. Phys., 281 (2015), pp. 825–843, doi:10.1016/j.jcp.2014.10.051.
[33] C. T. Kelley, *Iterative Methods for Optimization*, Frontiers in Appl. Math. 18, SIAM, Philadelphia, 1999, doi:10.1137/1.9781611970920.
[34] A. Kilbas, H. Srivastava, and J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier Science B.V., Amsterdam, 2006.
[35] G. Kinderlehrer and D. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Pure Appl. Math. 88, Academic Press (Harcourt Brace Jovanovich), New York, London, 1980.
[36] A. Kufner and B. Opic, *How to define reasonably weighted Sobolev spaces*, Comment. Math. Univ. Carolin., 25 (1984), pp. 537–554.
[37] S. Levendorskiĭ, *Pricing of the American put under Lévy processes*, Int. J. Theor. Appl. Finance, 7 (2004), pp. 303–335, doi:10.1142/S0219024904002463.
[38] Y. Lin, X. Li, and C. Xu, *Finite difference/spectral approximations for the fractional cable equation*, Math. Comp., 80 (2011), pp. 1369–1396, doi:10.1090/S0025-5718-2010-02438-X.
[39] Y. Lin and C. Xu, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225 (2007), pp. 1533–1552, doi:10.1016/j.jcp.2007.02.001.
[40] J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications. Vol. I*, Springer-Verlag, New York, 1972.
[41] A. Lotfi, M. Dehghan, and S. Yousefi, *A numerical technique for solving fractional optimal control problems*, Comput. Math. Appl., 62 (2011), pp. 1055–1067, doi:10.1016/j.camwa.2011.03.044.
[42] A. Lotfi, S. Yousefi, and M. Dehghan, *Numerical solution of a class of fractional optimal control problems via the Legendre orthonormal basis combined with the operational matrix and the Gauss quadrature rule*, J. Comput. Appl. Math., 250 (2013), pp. 143–160, doi:10.1016/j.cam.2013.03.003.
[43] W. McLean, *Regularity of solutions to a time-fractional diffusion equation*, ANZIAM J., 52 (2010), pp. 123–138, doi:10.1017/S1446181111000617.
[44] D. Meidner and B. Vexler, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems part* ii: *Problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329, doi:10.1137/070694028.
[45] K. Mustapha and W. McLean, *Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations*, SIAM J. Numer. Anal., 51 (2013), pp. 491–515, doi:10.1137/120880719.
[46] R. Nigmatullin, *The realization of the generalized transfer equation in a medium with fractal geometry*, Phys. Status Solidi (B), 133 (1986), pp. 425–430, doi:10.1002/pssb.2221330150.
[47] R. Nochetto, *Error estimates for two-phase Stefan problems in several space variables,* I: *Linear boundary conditions*, Calcolo, 22 (1985), pp. 457–499, doi:10.1007/BF02575898.
[48] R. Nochetto, E. Otárola, and A. Salgado, *A PDE approach to fractional diffusion in general domains: A priori error analysis*, Found. Comput. Math., 15 (2015), pp. 733–791, doi:10.1007/s10208-014-9208-x.
[49] R. H. Nochetto, E. Otárola, and A. J. Salgado, *A PDE approach to space-time fractional parabolic problems*, SIAM J. Numer. Anal., 54 (2016), pp. 848–873, doi:10.1137/14096308x.
[50] R. Nochetto, E. Otárola, and A. Salgado, *Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications*, Numer. Math., 132 (2016), pp. 85–130, doi:10.1007/s00211-015-0709-6.
[51] E. Otárola, *A PDE Approach to Numerical Fractional Diffusion*, Ph.D. thesis, Department of Mathematics, University of Maryland, College Park, MD, 2014.
[52] A. Saichev and G. Zaslavsky, *Fractional kinetic equations: Solutions and applications*, Chaos, 7 (1997), pp. 753–764, doi:10.1063/1.166272.
[53] K. Sakamoto and M. Yamamoto, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, J. Math. Anal. Appl., 382 (2011), pp. 426–447, doi:10.1016/j.jmaa.2011.04.058.
[54] S. Samko, A. Kilbas, and O. Marichev, *Fractional Integrals and Derivatives*, Gordon and Breach Science Publishers, Yverdon, Switzerland, 1993.
[55] P. Stinga and J. Torrea, *Extension problem and Harnack's inequality for some fractional operators*, Comm. Partial Differential Equations, 35 (2010), pp. 2092–2122, doi:10.1080/03605301003735680.

[56] L. Tartar, *An Introduction to Sobolev Spaces and interpolation Spaces*, Lecture Notes Unione Mat. Ital. 3, Springer, Berlin, 2007.

[57] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. Peterson, R. Roskies, J. Scott, and N. Wilkins-Diehr, *XSEDE: Accelerating scientific discovery*, Comput. Sci. Engrg., 16 (2014), pp. 62–74.

[58] C. Tricaud and Y. Chen, *An approximate method for numerically solving fractional order optimal control problems of general form*, Comput. Math. Appl., 59 (2010), pp. 1644–1655, doi:10.1016/j.camwa.2009.08.006.

[59] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, AMS, Providence, RI, 2010.

[60] B. Turesson, *Nonlinear Potential Theory and Weighted Sobolev Spaces*, Springer, New York, 2000, doi:10.1007/BFb0103908.

[61] X. Ye and C. Xu, *Spectral optimization methods for the time fractional diffusion inverse problem*, Numer. Math. Theory Methods Appl., 6 (2013), pp. 499–516.

[62] X. Ye and C. Xu, *A spectral method for optimal control problems governed by the time fractional diffusion equation with control constraints*, in Spectral and High Order Methods for Partial Differential Equations—ICOSAHOM 2012, Springer, Berlin, 2014, pp. 403–414, doi:10.1007/978-3-319-01601-6_33.