AN ADAPTIVE FINITE ELEMENT METHOD FOR THE SPARSE OPTIMAL CONTROL OF FRACTIONAL DIFFUSION*

ENRIQUE OTÁROLA†

Abstract. We propose and analyze an a posteriori error estimator for a PDE–constrained optimization problem involving a nondifferentiable cost functional, the spectral fractional powers of the Dirichlet Laplace operator as state equation and control–constraints. We realize fractional diffusion as the Dirichlet-to-Neumann map for a nonuniformly elliptic equation and propose an equivalent problem with a local state equation. For such an equivalent problem, we design an a posteriori error estimator which can be defined as the sum of four contributions: two contributions related to the finite element approximation of the state and adjoint equations and two contributions that account for the discretization of the control variable and its associated subgradient. The contributions related to the approximation of the state and adjoint equations rely on anisotropic error estimators in Muckenhoupt Sobolev spaces. We prove that the proposed a posteriori error estimator is locally efficient and, under suitable assumptions, reliable. We design an adaptive scheme that yields, for the examples that we perform, optimal experimental rates of convergence.

Key words. PDE–constrained optimization, nonsmooth objectives, sparse controls, the spectral fractional Laplacian, nonlocal operators, a posteriori error analysis, anisotropic estimates, adaptive loop.

AMS subject classifications. 35R11, 35J70, 49J20, 49M25, 65N12, 65N30, 65N50.

1. Introduction. The main goal of this work is the design and study of a posteriori error estimates for an optimal control problem that entails the minimization of a nondifferentiable cost functional, a state equation that involves the spectral fractional Laplacian and constraints on the control variable. To be precise, let $\Omega \subset \mathbb{R}^n$ $(n \geq 1)$ be a bounded and open polytopal domain with Lipschitz boundary $\partial\Omega$, $s \in (0,1)$ and $\mathsf{u}_d : \Omega \to \mathbb{R}$ be a desired state. For parameters $\sigma > 0$ and $\nu > 0$, we define the nonsmooth cost functional

$$J(\mathsf{u},\mathsf{z}) := \frac{1}{2}\|\mathsf{u} - \mathsf{u}_d\|^2_{L^2(\Omega)} + \frac{\sigma}{2}\|\mathsf{z}\|^2_{L^2(\Omega)} + \nu\|\mathsf{z}\|_{L^1(\Omega)}. \tag{1.1}$$

We will be interested in the numerical approximation of the following nondifferentiable PDE–constrained optimization problem: Find

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

subject to the *nonlocal state equation*

$$(-\Delta)^s \mathsf{u} = \mathsf{z} \text{ in } \Omega, \qquad \mathsf{u} = 0 \text{ on } \partial\Omega, \tag{1.3}$$

and the *constraints*

$$\mathsf{a} \leq \mathsf{z}(x') \leq \mathsf{b} \quad \text{a.a. } x' \in \Omega. \tag{1.4}$$

For $s \in (0,1)$, the operator $(-\Delta)^s$ denotes the *spectral* fractional powers of the Dirichlet Laplace operator, the so–called spectral fractional Laplacian. We must immediately comment that this definition and the classical one that is based on a pointwise

*EO is partially supported by CONICYT through FONDECYT project 3160201.

†Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. enrique.otarola@usm.cl.

integral formula [36, 53, 55] do not coincide; their difference is positive and positivity preserving [42]. In addition, we also comment that, since we are interested in the nonsmooth scenario, we will assume, in (1.4), that the control bounds $\mathsf{a}, \mathsf{b} \in \mathbb{R}$ satisfy that $\mathsf{a} < 0 < \mathsf{b}$. We refer the reader to [16, Remark 2.1] for a discussion.

The efficient approximation of problems involving the spectral fractional Laplacian carries two essential difficulties. The first, and most important, is that $(-\Delta)^s$ is a nonlocal operator [10, 11, 14, 56]. The second feature is the lack of boundary regularity [13], which leads to reduced convergence rates [6, 43]. In fact, as [13, Theorem 1.3] shows, if $\partial\Omega$ is sufficiently smooth, then the solution u to problem (1.3) behaves like

$$\begin{aligned} \mathsf{u}(x') &\approx \operatorname{dist}(x', \partial\Omega)^{2s} + \mathsf{v}(x'), \qquad && s \in (0, \tfrac{1}{2}), \\ \mathsf{u}(x') &\approx \operatorname{dist}(x', \partial\Omega) + \mathsf{v}(x'), \qquad && s \in (\tfrac{1}{2}, 1), \end{aligned} \tag{1.5}$$

where $\operatorname{dist}(x', \partial\Omega)$ denotes the distance from x' to $\partial\Omega$ and v is a smooth function. The case $s = \frac{1}{2}$ is exceptional:

$$\mathsf{u}(x') \approx \operatorname{dist}(x', \partial\Omega) |\log(\operatorname{dist}(x', \partial\Omega))| + \mathsf{v}(x'), \quad s = \tfrac{1}{2}, \tag{1.6}$$

where v is, again, a smooth function; see [24] for $\Omega \subset \mathbb{R}^2$ and $\partial\Omega$ smooth.

The aforementioned nonlocality difficulty can be overcame with the localization results by Caffarelli and Silvestre [11]. When $\Omega = \mathbb{R}^n$, the authors of [11] proved that any power of the fractional Laplacian can be realized as the Dirichlet-to-Neumann map for an extension problem posed in the upper half–space $\mathbb{R}^{n+1}_+$. A similar extension property is valid for the spectral fractional Laplacian in a bounded domain Ω [10, 14, 56]. The latter extension involves a local but nonuniformly elliptic PDE formulated in the semi–infinite cylinder $\mathcal{C} = \Omega \times (0, \infty)$:

$$\operatorname{div}(y^\alpha \nabla \mathscr{U}) = 0 \text{ in } \mathcal{C}, \quad \mathscr{U} = 0 \text{ on } \partial_L \mathcal{C}, \quad \partial_{\nu^\alpha} \mathscr{U} = d_s \mathsf{z} \quad \text{on } \Omega \times \{0\}, \tag{1.7}$$

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

$$\partial_{\nu^\alpha} \mathscr{U} = - \lim_{y \to 0^+} y^\alpha \mathscr{U}_y; \tag{1.8}$$

the limit is understood in the distributional sense. We shall refer to y as the *extended variable* and to the dimension $n + 1$, in $\mathbb{R}^{n+1}_+$, the *extended dimension* of problem (1.7). With the extension $\mathscr{U}$ at hand, we thus introduce the fundamental result by Caffarelli and Silvestre [10, 11, 14, 56]: the Dirichlet-to-Neumann map of problem (1.7) and the spectral fractional Laplacian are related by $d_s(-\Delta)^s \mathsf{u} = \partial_{\nu^\alpha} \mathscr{U}$ in Ω.

The use of the extension problem (1.7) for the numerical approximation of the spectral fractional Laplacian was first used in [43]. In such a work the authors introduce the following solution scheme: Given a datum z, solve, on the basis of a finite element discretization, the extension problem (1.7) and obtain V, thus set $U = \operatorname{tr}_\Omega V$ and arrive at an approximation of the solution u of problem (1.3). The main advantage of the scheme proposed in [43] is that it involves the resolution of the local problem (1.7) and thus its implementation uses basic ingredients of finite element analysis; its analysis, however, involves asymptotic estimates of Bessel functions [1], to derive regularity estimates in weighted Sobolev spaces, elements of harmonic analysis [26, 41], and an anisotropic polynomial interpolation theory in weighted Sobolev spaces [27, 44]. Such an interpolation theory allows for tensor product elements that exhibit an anisotropic feature in the extended dimension, that is in turn needed to compensate the singular behavior of the solution $\mathscr{U}$ in the extended variable y [43, Theorem 2.7], [6, Theorem 4.7].

Recently, and on the basis of the aforementioned numerical scheme, the authors of [48] have provided an a priori error analysis for the optimal control problem (1.2)–(1.4). This was mainly motivated by the following considerations:

- The fractional Laplacian has recently become of great interest in the applied sciences and engineering: practitioners claim that it seems to better describe many processes. A rather incomplete list of problems where fractional diffusion appears includes finance [37, 49], turbulent flow [20], quasi–geostrophic flows models [12], models of anomalous thermoviscous behaviors [21], biophysics [9], nonlocal electrostatics [34], image processing [30], peridynamics [25, 52] and many others. 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 cost functional J involves an $L^1(\Omega)$–control cost term that leads to sparsely supported optimal controls [16, 54, 60]; a desirable feature, for instance, in the optimal placement of discrete actuators [54].

In [48], on the basis of the localization results of [10, 11, 14, 56], the authors first consider an equivalent optimal control problem that involves the local elliptic PDE (1.7) as state equation. Second, since (1.7) is posed on the semi–infinite cylinder $\mathcal{C} = \Omega \times (0, \infty)$, they propose a truncated optimal control problem on $\mathcal{C}_{\mathcal{Y}} = \Omega \times (0, \mathcal{Y})$ and derive an exponential error estimate with respect to the truncation parameter $\mathcal{Y}$. Then, they propose a scheme to approximate the truncated optimal control problem: the first–degree finite element approximation on anisotropic meshes of [43] for the state and adjoint equations and piecewise constant approximation for the optimal control variable. The derived a priori error estimate reads as follows: Given $\mathsf{u}_d \in \mathbb{H}^{1-s}(\Omega)$ and $\mathsf{a}, \mathsf{b} \in \mathbb{R}$ such that $\mathsf{a} < 0 < \mathsf{b}$, if Ω is convex, then

$$\|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)} \lesssim |\log N|^{2s} N^{-\frac{1}{n+1}}, \tag{1.9}$$

where $\bar{Z}$ corresponds to the optimal control variable of the fully discrete scheme of [48, Section 6] and N denotes the total number of degrees of freedom of the underlying mesh. The adaptive finite element method (AFEM) that we propose in our work is thus motivated, in addition to the search of a numerical scheme that efficiently solves (1.2)–(1.4) with relatively modest computational resources, by the following considerations:

- The a priori error estimate (1.9) requires the convexity of the domain Ω and compatibility conditions on the desired state u_d that are expressed as $\mathsf{u}_d \in \mathbb{H}^{1-s}(\Omega)$: *it is required that* u_d *vanishes on* $\partial\Omega$ *for* $s \in (0, \frac{1}{2}]$. If these conditions do not hold then, the a priori error estimate (1.9) is not longer valid. In particular, the violation of the latter condition implies that the adjoint state $\bar{\mathsf{p}}$ will behave as (1.5)–(1.6). On the other hand, if the first condition (the convexity of the domain) is violated, then both the state and adjoint state variables may exhibit singularities in the direction of the x' variables. An efficient technique to solve (1.2)–(1.4) must thus resolve both of the aforementioned approximation issues.
- The sparsity term $\|\mathsf{u}\|_{L^1(\Omega)}$, in the cost functional (1.2), yield an optimal control that is non–zero only in sets of small support in Ω [16, 54, 60]. It is then natural, to efficiently resolve such a behavior, the consideration of AFEMs.

We organize our exposition as follows: In Section 2 we recall the definition of the spectral fractional Laplacian, present the fundamental result by Caffarelli and Silvestre [11] and recall elements from convex analysis. In Section 3 we recall the numerical scheme proposed in [48] and review its a priori error analysis. Section 4, that is a highlight of our contribution, is dedicated to the design and analysis of an ideal a posteriori error estimator for our optimal control problem (1.2)–(1.4) that is equivalent to the error. Since, the aforementioned estimator is not computable, we propose, in Section 5 a computable one and show that is equivalent, under suitable assumptions, to the error up to oscillation terms. Finally, in Section 6 we design an AFEM, comment on some implementation details pertinent to the problem and present numerical experiments that yield optimal experimental rates of convergence.

2. Notation and preliminaries. We adopt the notation of [43, 47]. Besides the semi–infinite cylinder $\mathcal{C} = \Omega \times (0,\infty)$, we introduce, for $\mathcal{Y} > 0$, the truncated cylinder with base Ω and height $\mathcal{Y}$ and its lateral boundary

$$\mathcal{C}_{\mathcal{Y}} = \Omega \times (0,\mathcal{Y}), \qquad \partial_L \mathcal{C}_{\mathcal{Y}} = \partial\Omega \times (0,\mathcal{Y}),$$

respectively. Since we will be dealing with objects defined in $\mathbb{R}^n$ and $\mathbb{R}^{n+1}_+$, it will be convenient to distinguish the extended variable y. For $x \in \mathbb{R}^{n+1}_+$, we write

$$x = (x', y), \quad x' \in \mathbb{R}^n, \quad y \in (0,\infty). \tag{2.1}$$

The parameter $\alpha \in (-1,1)$ and the power s of the spectral fractional Laplacian $(-\Delta)^s$ are related by the formula $\alpha = 1 - 2s$.

The relation $a \lesssim b$ indicates that $a \le Cb$, with a constant C which is independent of a and b and the size of the elements in the mesh. The value of the constant C might change at each occurrence.

2.1. The fractional Laplace operator. To define $(-\Delta)^s$ we invoke spectral theory [7]. Since $-\Delta : \mathcal{D}(-\Delta) \subset L^2(\Omega) \to L^2(\Omega)$ is an unbounded, positive and closed operator with dense domain $\mathcal{D}(-\Delta) = H^2(\Omega) \cap H^1_0(\Omega)$ and its inverse is compact, the eigenvalue problem: Find $(\lambda, \varphi) \in \mathbb{R} \times H^1_0(\Omega) \setminus \{0\}$ such that

$$(\nabla \varphi, \nabla v)_{L^2(\Omega)} = \lambda (\varphi, v)_{L^2(\Omega)} \quad \forall v \in H^1_0(\Omega)$$

has a countable collection of eigenpairs $\{\lambda_k, \varphi_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^+ \times H^1_0(\Omega)$, with real eigenvalues enumerated in increasing order, counting multiplicities. In addition, $\{\varphi_k\}_{k\in\mathbb{N}}$ is an orthogonal basis of $H^1_0(\Omega)$ and an orthonormal basis of $L^2(\Omega)$. With these eigenpairs at hand, we define, for $s \ge 0$, the fractional Sobolev space

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

We denote by $\mathbb{H}^{-s}(\Omega)$ the dual space of $\mathbb{H}^s(\Omega)$. The duality pairing between $\mathbb{H}^s(\Omega)$ and $\mathbb{H}^{-s}(\Omega)$ will be denoted by $\langle \cdot, \cdot \rangle$. We thus define, with this functional space setting at hand, and for $s \in (0,1)$, the *spectral* fractional Laplacian $(-\Delta)^s$ as follows:

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

2.2. An extension property. The operator in (1.7) is in divergence form and thus amenable to variational techniques. However, since the weight y^α either blows up, for $-1 < \alpha < 0$, or degenerates, for $0 < \alpha < 1$, as $y \downarrow 0$, such a local operator is nonuniformly elliptic; the case $\alpha = 0$ is exceptional and corresponds to the regular harmonic extension [10]. This entails dealing with Lebesgue and Sobolev spaces with the weight y^α for $\alpha \in (-1,1)$ [10, 11, 14].

Let $D \subset \mathbb{R}^{n+1}$ be open, and define $L^2(|y|^\alpha, D)$ as the Lebesgue space for the measure $|y|^\alpha \, \mathrm{d}x$. We also define the weighted Sobolev space $H^1(|y|^\alpha, D) := \{w \in L^2(|y|^\alpha, D) : |\nabla w| \in L^2(|y|^\alpha, D)\}$, and the norm

$$\|w\|_{H^1(|y|^\alpha,D)} = \left(\|w\|^2_{L^2(|y|^\alpha,D)} + \|\nabla w\|^2_{L^2(|y|^\alpha,D)} \right)^{\frac{1}{2}}. \tag{2.2}$$

Since $\alpha = 1-2s \in (-1,1)$, $|y|^\alpha$ belongs to Muckenhoupt class $A_2(\mathbb{R}^{n+1})$ [26, 29, 31, 41, 58]. This, in particular, implies that $H^1(|y|^\alpha, D)$ is Hilbert and that $C^\infty(D) \cap H^1(|y|^\alpha, D)$ is dense in $H^1(|y|^\alpha, D)$ (cf. [58, Proposition 2.1.2, Corollary 2.1.6] and [31, Theorem 1]).

To seek for a weak solution to problem (1.7), we introduce the weighted space

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

We have the *weighted Poincaré inequality* [43, ineq. (2.21)],

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

This implies that the seminorm on $\mathring{H}^1_L(y^\alpha, \mathcal{C})$ is equivalent to (2.2). For $w \in H^1(y^\alpha, \mathcal{C})$, $\operatorname{tr}_\Omega w$ denotes its trace onto $\Omega \times \{0\}$. We recall ([43, Prop. 2.5] and [14, Prop. 2.1])

$$\operatorname{tr}_\Omega \mathring{H}^1_L(y^\alpha, \mathcal{C}) = \mathbb{H}^s(\Omega), \qquad \|\operatorname{tr}_\Omega w\|_{\mathbb{H}^s(\Omega)} \leq C_{\operatorname{tr}_\Omega} \|w\|_{\mathring{H}^1_L(y^\alpha,\mathcal{C})}, \quad C_{\operatorname{tr}_\Omega} > 0. \tag{2.3}$$

We mention that $C_{\operatorname{tr}_\Omega} \leq d_s^{-\frac{1}{2}}$ [19, Section 2.3] with $d_s = 2^\alpha \Gamma(1-s)/\Gamma(s)$. This property will be of importance in the a posteriori error analysis that we will perform.

Define the bilinear form

$$a : \mathring{H}^1_L(y^\alpha, \mathcal{C}) \times \mathring{H}^1_L(y^\alpha, \mathcal{C}) \to \mathbb{R}, \quad a(w, \phi) := \frac{1}{d_s} \int_{\mathcal{C}} y^\alpha \nabla w \cdot \nabla \phi \, \mathrm{d}x. \tag{2.4}$$

The weak formulation of problem (1.7) thus reads: Find $\mathscr{U} \in \mathring{H}^1_L(y^\alpha, \mathcal{C})$ such that

$$a(\mathscr{U}, w) = \langle \mathsf{z}, w \rangle \quad \forall w \in \mathring{H}^1_L(y^\alpha, \mathcal{C}). \tag{2.5}$$

We conclude this section with the fundamental result by Caffarelli and Silvestre [10, 11, 14, 56]: Let u solve (1.3). If $\mathscr{U} \in \mathring{H}^1_L(y^\alpha, \mathcal{C})$ solves (2.5), then $\mathsf{u} = \operatorname{tr}_\Omega \mathscr{U}$ and

$$d_s(-\Delta)^s \mathsf{u} = \partial_{\nu^\alpha} \mathscr{U} \text{ in } \Omega.$$

2.3. Convex functions and subdifferentials. Let S be a real and normed vector space. Let $\chi : S \to \mathbb{R} \cup \{\infty\}$ be a convex and proper function and let $v \in S$ be such that $\chi(v) < \infty$. A *subgradient* of χ at v is an element $v^\star \in S^*$ that satisfies

$$(v^\star, w - v)_{S^\star,S} \leq \chi(w) - \chi(v) \quad \forall w \in S, \tag{2.6}$$

where $(\cdot,\cdot)_{S^\star,S}$ denotes the duality pairing between S^* and S. We denote by $\partial\chi(v)$ the set of all subgradients of χ at v; the so–called *subdifferential* of χ at v. As a consequence of the convexity of χ, $\partial\chi(v) \neq \emptyset$ for all points v in the interior of the effective domain of χ.

We now present the following property which will be essential in our analysis: the subdifferential is monotone, i.e.,

$$\langle v^\star - w^\star, v - w\rangle_{S^\star,S} \geq 0 \quad \forall v^\star \in \partial\chi(v),\ \forall w^\star \in \partial\chi(w). \tag{2.7}$$

The reader is referred to [23, 50] for a detailed treatment on convex analysis.

3. A priori error estimates. In this section we briefly review the a priori error analysis for the fully discrete approximation of the optimal control problem (1.1)–(1.4) proposed and investigated in [48]. We also make clear the limitation of such a priori error analysis, thereby justifying the quest for an a posteriori error analysis.

3.1. The extended optimal control problem. In view of the localization results of [10, 11, 14, 56], we will consider a solution technique for (1.1)–(1.4) that relies on the discretization of an equivalent problem: the *extended optimal control problem*, which has as a main advantage its local nature. To present it, we first define the set of *admissible controls*:

$$\mathsf{Z}_{\mathrm{ad}} = \{\mathsf{w} \in L^2(\Omega) : \mathsf{a} \leq \mathsf{w}(x') \leq \mathsf{b} \ \text{ a.a } \ x' \in \Omega\}, \tag{3.1}$$

where a and b are real and satisfy the property $\mathsf{a} < 0 < \mathsf{b}$; see[16, Remark 2.1]. The extended optimal control problem reads as follows: Find

$$\min\{J(\mathrm{tr}_\Omega \mathscr{U}, \mathsf{z}) : \mathscr{U} \in \mathring{H}^1_L(y^\alpha, \mathcal{C}), \mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}\} \tag{3.2}$$

subject to the *linear* and nonuniformly elliptic state equation

$$a(\mathscr{U}, \phi) = \langle \mathsf{z}, \mathrm{tr}_\Omega \phi\rangle \quad \forall \phi \in \mathring{H}^1_L(y^\alpha, \mathcal{C}). \tag{3.3}$$

We recall that J is defined as in (1.2) with $\sigma, \nu > 0$, and $\mathsf{u}_d \in L^2(\Omega)$. This problem admits a unique optimal pair $(\bar{\mathscr{U}}, \bar{\mathsf{z}}) \in \mathring{H}^1_L(y^\alpha, \mathcal{C}) \times \mathsf{Z}_{\mathrm{ad}}$. More importantly, it is equivalent to the optimal control problem (1.2)–(1.4): $\mathrm{tr}_\Omega \bar{\mathscr{U}} = \bar{\mathsf{u}}$.

3.2. The truncated optimal control problem. The previously defined PDE–constrained optimization problem involves the state equation (3.3), which is posed on the semi–infinite cylinder $\mathcal{C} = \Omega \times (0, \infty)$. Consequently, it cannot be directly approximated with standard finite element techniques. However, in view of the exponential decayment of the optimal state in the extended variable [43, Proposition 3.1], it is suitable to propose the following *truncated optimal control problem.* Find

$$\min\{J(\mathrm{tr}_\Omega v, \mathsf{r}) : v \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathscr{Y}), \mathsf{r} \in \mathsf{Z}_{\mathrm{ad}}\}$$

subject to the *truncated* state equation

$$a_\mathscr{Y}(v, \phi) = \langle \mathsf{r}, \mathrm{tr}_\Omega \phi\rangle \quad \forall \phi \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathscr{Y}). \tag{3.4}$$

Here,

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

and the bilinear form $a_{\mathscr{Y}} : \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}}) \times \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}}) \to \mathbb{R}$ is defined by

$$a_{\mathscr{Y}}(w, \phi) = \frac{1}{d_s} \int_{\mathcal{C}_{\mathscr{Y}}} y^\alpha \nabla w \cdot \nabla \phi \, \mathrm{d}x. \tag{3.5}$$

This optimal control problem admits a unique optimal solution $(\bar{v}, \bar{\mathsf{r}}) \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}}) \times \mathsf{Z}_{\mathrm{ad}}$. In addition, such a pair is optimal if and only if $\bar{v}$ solves (3.4) and

$$(\mathrm{tr}_\Omega \, \bar{p} + \sigma \bar{\mathsf{r}} + \nu \bar{t}, \mathsf{r} - \bar{\mathsf{r}})_{L^2(\Omega)} \geq 0 \quad \forall \mathsf{r} \in \mathsf{Z}_{\mathrm{ad}}, \tag{3.6}$$

where $\bar{t} \in \partial\psi(\bar{\mathsf{r}})$ and $\bar{p} \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}})$ solves the truncated adjoint problem

$$a_{\mathscr{Y}}(\phi, \bar{p}) = (\mathrm{tr}_\Omega \, \bar{v} - \mathsf{u}_{\mathrm{d}}, \mathrm{tr}_\Omega \, \phi)_{L^2(\Omega)} \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}}). \tag{3.7}$$

The convex and Lipschitz function ψ is defined as follows:

$$\psi : L^1(\Omega) \to \mathbb{R}, \quad \psi(\mathsf{r}) := \int_\Omega |\mathsf{r}(x')| \, \mathrm{d}x'. \tag{3.8}$$

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

PROPOSITION 3.1 (exponential convergence). *Let $(\bar{\mathscr{U}}, \bar{\mathsf{z}}) \in \mathring{H}_L^1(y^\alpha, \mathcal{C}) \times \mathsf{Z}_{ad}$ and $(\bar{v}, \bar{\mathsf{r}}) \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}}) \times \mathsf{Z}_{ad}$ be the solutions to the extended and truncated optimal control problems, respectively. Then,*

$$\|\bar{\mathsf{z}} - \bar{\mathsf{r}}\|_{L^2(\Omega)} \lesssim e^{-\sqrt{\lambda_1}\mathscr{Y}/4} \left(\|\bar{\mathsf{r}}\|_{L^2(\Omega)} + \|\mathsf{u}_d\|_{L^2(\Omega)} \right),$$
$$\|\nabla \left(\bar{\mathscr{U}} - \bar{v} \right) \|_{L^2(y^\alpha, \mathcal{C})} \lesssim e^{-\sqrt{\lambda_1}\mathscr{Y}/4} \left(\|\bar{\mathsf{r}}\|_{L^2(\Omega)} + \|\mathsf{u}_d\|_{L^2(\Omega)} \right);$$

λ_1 corresponds to the first eigenvalue of $-\Delta$.

Proof. See [48, Theorem 5.2]. ◻

We conclude the section with the following projection formulas that follow from [48, Corollary 3.7]. To present them, we first introduce the following nonlinear projection operator [57, Section 2.8]

$$\Pi_{[\mathfrak{a},\mathfrak{b}]} : L^2(\Omega) \to \mathsf{Z}_{\mathrm{ad}}, \quad \Pi_{[\mathfrak{a},\mathfrak{b}]}(x') = \min\{\mathfrak{b}, \max\{\mathfrak{a}, x'\}\}, \tag{3.9}$$

where $\mathfrak{a}$ and $\mathfrak{b}$ denote real constants.

PROPOSITION 3.2 (projection formulas). *If $\bar{\mathsf{r}}$, $\bar{v}$, $\bar{p}$ and $\bar{t}$ denote the optimal variables associated to the truncated optimal control problem, then*

$$\bar{\mathsf{r}}(x') = \Pi_{[\mathsf{a},\mathsf{b}]} \left(-\frac{1}{\sigma} \left(\mathrm{tr}_\Omega \, \bar{p}(x') + \nu \bar{t}(x') \right) \right), \tag{3.10}$$

$$\bar{\mathsf{r}}(x') = 0 \quad \Leftrightarrow \quad |\, \mathrm{tr}_\Omega \, \bar{p}(x')| \leq \nu, \tag{3.11}$$

$$\bar{t}(x') = \Pi_{[-1,1]} \left(-\frac{1}{\nu} \, \mathrm{tr}_\Omega \, \bar{p}(x') \right). \tag{3.12}$$

3.3. A fully discrete scheme for the fractional optimal control problem. In what follows we briefly recall the fully discrete scheme proposed in [48] and review its a priori error analysis. To accomplish this task, we will assume in this section that

$$\|w\|_{H^2(\Omega)} \lesssim \|\Delta_{x'} w\|_{L^2(\Omega)} \quad \forall w \in H^2(\Omega) \cap H_0^1(\Omega). \tag{3.13}$$

This regularity assumption holds if, for instance, Ω is convex [32].

Before describing the aforementioned solution technique, we briefly recall the finite element approximation of [43] for the state equation (2.5).

Let $\mathscr{T}_\Omega = \{K\}$ be a conforming and shape regular mesh of Ω into cells K that are isoparametrically equivalent either to the unit cube $[0,1]^n$ or the unit simplex in $\mathbb{R}^n$ [8, 22, 28]. Let $\mathcal{I}_\mathcal{Y}$ be a partition of $[0,\mathcal{Y}]$ with mesh points

$$y_\ell = \left(\frac{\ell}{M}\right)^\gamma \mathcal{Y}, \quad \gamma > \frac{3}{(1-\alpha)} = \frac{3}{2s} > 1, \quad \ell = 0,\dots,M. \tag{3.14}$$

We construct a mesh $\mathscr{T}_\mathcal{Y}$ over the cylinder $\mathcal{C}_\mathcal{Y}$ as $\mathscr{T}_\mathcal{Y} = \mathscr{T}_\Omega \otimes I_\mathcal{Y}$, the tensor product triangulation of $\mathscr{T}_\Omega$ and $\mathcal{I}_\mathcal{Y}$. The set of all the obtained meshes is denoted by $\mathbb{T}$. Notice that, owing to (3.14), the meshes $\mathscr{T}_\mathcal{Y}$ are not shape regular but satisfy: if $T_1 = K_1 \times I_1$ and $T_2 = K_2 \times I_2$ are neighbors, then there is $\mu > 0$ such that

$$h_{I_1} h_{I_2}^{-1} \le \mu, \qquad h_I = |I|.$$

This condition allows for anisotropy in the extended variable y [27, 43, 44], which is needed to compensate the rather singular behavior of $\mathscr{U}$, solution to (3.3). We refer the reader to [43] for details.

With the mesh $\mathscr{T}_\mathcal{Y} \in \mathbb{T}$ at hand, we define the finite element space

$$\mathbb{V}(\mathscr{T}_\mathcal{Y}) = \left\{ W \in C^0(\overline{\mathcal{C}_\mathcal{Y}}) : W|_T \in \mathcal{P}_1(K) \otimes \mathbb{P}_1(I) \ \forall T \in \mathscr{T}_\mathcal{Y},\ W|_{\Gamma_D} = 0 \right\}, \tag{3.15}$$

where $\Gamma_D = \partial_L \mathcal{C}_\mathcal{Y} \cup \Omega \times \{\mathcal{Y}\}$ is the Dirichlet boundary. If the base K of the element $T = K \times I$ is a cube, $\mathcal{P}_1(K)$ stand for $\mathbb{Q}_1(K)$ – the space of polynomials of degree not larger that one in each variable. When K is a simplex, the space $\mathcal{P}_1(K)$ is $\mathbb{P}_1(K)$, i.e., the set of polynomials of degree at most one. We also define $\mathbb{U}(\mathscr{T}_\Omega) = \operatorname{tr}_\Omega \mathbb{V}(\mathscr{T}_\mathcal{Y})$. Notice that $\mathbb{U}(\mathscr{T}_\Omega)$ corresponds to a $\mathcal{P}_1$ finite element space over $\mathscr{T}_\Omega$.

We now describe the *fully discrete optimal control problem.* To accomplish this task, we first introduce the discrete sets

$$\mathbb{Z}_{ad}(\mathscr{T}_\Omega) = \mathsf{Z}_{\mathrm{ad}} \cap \mathbb{P}_0(\mathscr{T}_\Omega), \quad \mathbb{P}_0(\mathscr{T}_\Omega) = \{ Z \in L^\infty(\Omega) : Z|_K \in \mathbb{P}_0(K) \quad \forall K \in \mathscr{T}_\Omega \}.$$

The fully discrete optimal control problem thus reads as follows: Find

$$\min\{ J(\operatorname{tr}_\Omega V, Z) : V \in \mathbb{V}(\mathscr{T}_\mathcal{Y}), Z \in \mathbb{Z}_{ad}(\mathscr{T}_\Omega) \}$$

subject to

$$a_\mathcal{Y}(V,W) = (Z, \operatorname{tr}_\Omega W)_{L^2(\Omega)} \quad \forall W \in \mathbb{V}(\mathscr{T}_\mathcal{Y}). \tag{3.16}$$

We recall that J and $a_\mathcal{Y}$ are defined by (1.1) and (3.5), respectively. Standard arguments guarantee the existence of a unique optimal pair $(\bar{V}, \bar{Z}) \in \mathbb{V}(\mathscr{T}_\mathcal{Y}) \times \mathbb{Z}_{ad}(\mathscr{T}_\Omega)$. In view of the results of [43, 47], we invoke the discrete solution $\bar{V} \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$ and set

$$\bar{U} := \operatorname{tr}_\Omega \bar{V}. \tag{3.17}$$

We have thus obtained a fully discrete approximation $(\bar{U}, \bar{Z}) \in \mathbb{U}(\mathscr{T}_\Omega) \times \mathbb{Z}_{ad}(\mathscr{T}_\Omega)$ of $(\bar{\mathsf{u}}, \bar{\mathsf{z}}) \in \mathbb{H}^s(\Omega) \times \mathsf{Z}_{\mathrm{ad}}$, the solution to the fractional optimal control problem.

The optimality conditions for the fully discrete optimal control problem read as follows: the pair $(\bar{V}, \bar{Z}) \in \mathbb{V}(\mathscr{T}_\mathcal{Y}) \times \mathbb{Z}_{ad}(\mathscr{T}_\Omega)$ is optimal if and only if $\bar{V} \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$ solves (3.16) and

$$(\operatorname{tr}_\Omega \bar{P} + \sigma \bar{Z} + \nu \bar{\Lambda}, Z - \bar{Z})_{L^2(\Omega)} \ge 0 \quad \forall Z \in \mathbb{Z}_{ad}(\mathscr{T}_\Omega), \tag{3.18}$$

where $\bar{\Lambda} \in \partial\psi(\bar{Z})$ and the optimal discrete adjoint state $\bar{P} \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$ solves

$$a_{\mathcal{Y}}(W, \bar{P}) = (\mathrm{tr}_{\Omega} \bar{V} - \mathsf{u}_{\mathrm{d}}, \mathrm{tr}_{\Omega} W)_{L^2(\Omega)} \quad \forall W \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}}). \tag{3.19}$$

To write an priori error estimates for the aforementioned scheme, we first observe that $\#\mathscr{T}_{\mathcal{Y}} = M\,\#\mathscr{T}_{\Omega}$, and that $\#\mathscr{T}_{\Omega} \approx M^n$. Consequently, $\#\mathscr{T}_{\mathcal{Y}} \approx M^{n+1}$. Thus, if $\mathscr{T}_{\Omega}$ is quasi–uniform, we have that $h_{\mathscr{T}_{\Omega}} \approx (\#\mathscr{T}_{\Omega})^{-1/n}$.

THEOREM 3.3 (fractional control problem: error estimate). *Let $(\bar{V}, \bar{Z}) \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}}) \times \mathbb{Z}_{ad}(\mathscr{T}_{\Omega})$ be the optimal pair for the fully discrete optimal control problem. Let $\bar{U} \in \mathbb{U}(\mathscr{T}_{\Omega})$ be defined as in* (3.17)*. If Ω verifies* (3.13) *and $\mathsf{u}_d \in \mathbb{H}^{1-s}(\Omega)$, then*

$$\|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)} \lesssim |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2s}(\#\mathscr{T}_{\mathcal{Y}})^{\frac{-1}{n+1}} \left(\|\bar{\mathsf{r}}\|_{H^1(\Omega)} + \|\mathsf{u}_d\|_{\mathbb{H}^{1-s}(\Omega)}\right), \tag{3.20}$$

and

$$\|\bar{\mathsf{u}} - \bar{U}\|_{\mathbb{H}^s(\Omega)} \lesssim |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2s}(\#\mathscr{T}_{\mathcal{Y}})^{\frac{-1}{n+1}} \left(\|\bar{\mathsf{r}}\|_{H^1(\Omega)} + \|\mathsf{u}_d\|_{\mathbb{H}^{1-s}(\Omega)}\right), \tag{3.21}$$

where the truncation parameter $\mathcal{Y}$, in the truncated optimal control problem, is chosen such that $\mathcal{Y} \approx \log(\#\mathscr{T}_{\mathcal{Y}})$. The hidden constants in both inequalities are independent of the discretization parameters and the continuous and discrete optimal variables.

Proof. See [48, Theorem 6.4]. □

We stress that the results of Theorem 3.3 are valid if and only if Ω satisfies (3.13) and $\mathsf{u}_d \in \mathbb{H}^{1-s}(\Omega)$. These conditions guarantee that the optimal control variable $\bar{\mathsf{z}}$ and $\bar{\mathsf{r}}$ belong to $H_0^1(\Omega)$; see [48, Theorem 3.9].

We conclude this section by defining the following auxiliary variables that will be of importance in the a posteriori error analysis that we will perform:

$$v \in \mathring{H}_L^1(y^{\alpha}, \mathcal{C}_{\mathcal{Y}}) : \quad a_{\mathcal{Y}}(v, \phi) = (\bar{Z}, \mathrm{tr}_{\Omega}\,\phi)_{L^2(\Omega)} \quad \forall \phi \in \mathring{H}_L^1(y^{\alpha}, \mathcal{C}_{\mathcal{Y}}), \tag{3.22}$$

and

$$p \in \mathring{H}_L^1(y^{\alpha}, \mathcal{C}_{\mathcal{Y}}) : \quad a_{\mathcal{Y}}(\phi, p) = (\mathrm{tr}_{\Omega} \bar{V} - \mathsf{u}_{\mathrm{d}}, \mathrm{tr}_{\Omega}\,\phi)_{L^2(\Omega)} \quad \forall \phi \in \mathring{H}_L^1(y^{\alpha}, \mathcal{C}_{\mathcal{Y}}). \tag{3.23}$$

4. An *ideal* a posteriori error estimator.

The main goal of this work is the derivation and analysis of a computable a posteriori error estimator for problem (1.2)–(1.4). An a posteriori error estimator is a computable quantity that provides information about the local quality of the underlying approximated solution. It is an essential ingredient of AFEMs, which are iterative methods that improve the quality of the approximated solution and are based on loops of the form

$$\textsf{SOLVE} \rightarrow \textsf{ESTIMATE} \rightarrow \textsf{MARK} \rightarrow \textsf{REFINE}. \tag{4.1}$$

A posteriori error estimators are the heart of the step ESTIMATE. The theory for linear and second–order elliptic boundary value problems is well–established. We refer the reader to [39, 45, 46, 59] for an up to date discussion that also includes the design of AFEMs, convergence results, and optimal complexity.

The a posteriori error analysis for finite element approximations of constrained optimal control problems is currently under development; the main source of difficulty being its inherent nonlinear feature. Starting with the pioneering work [38], several authors have contributed to its advancement. For an up to date survey on a posteriori error analysis for optimal control problems we refer the reader to [3, 35, 51]. In contrast to these advances, the theory for optimal control problems involving a sparsity functional, as (1.2), is much less developed. To the best of our knowledge the only works that provides an advance concerning this matter are [60] and [2]. In [60], the authors propose a residual–type a posteriori error estimator and prove that it yields an upper bound for the approximation error of the state and control variables [60, Theorem 6.2]. These results have been recently complemented in [2], by deriving local efficiency estimates, and extended, by also deriving an a posteriori error analysis when the optimal control is discretized with piecewise linear functions and when the so–called variational discretization approach is considered.

In this work, we follow [33, 35, 38] and design an a posteriori error estimator for problem (1.2)–(1.4). To accomplish this task, it is essential to have at hand an error estimator for the truncated state equation (3.4). Such an equation involves a nonuniformly elliptic operator with a variable coefficient y^α that vanishes $(0 < \alpha < 1)$ or blows up $(-1 < \alpha < 0)$ as $y \downarrow 0$. Consequently, the design of error estimators for (3.4) is far from being trivial: In fact, a simple computation reveals that the usual residual estimator does not apply. In addition, numerical evidence shows that anisotropic refinement in the extended dimension is essential to observe experimental rates of convergence. In [19] an a posteriori error estimator based on the solution of weighted local problems on cylindrical stars is proposed for (3.4). It is inspired by the ideas that lead to the estimators introduced in [15, 40] that are in turn inspired by the work of Babuška and Miller [5].

As a first step, in the derivation of an error estimator for (1.2)–(1.4), we explore an *ideal* anisotropic estimator that can be decomposed as the sum of four contributions:

$$\mathscr{E}_{\mathrm{ocp}}^2 = \mathscr{E}_V^2 + \mathscr{E}_P^2 + \mathscr{E}_Z^2 + \mathscr{E}_\Lambda^2. \tag{4.2}$$

The error indicators $\mathscr{E}_V$ and $\mathscr{E}_P$, that are related to the finite element approximation of the state and adjoint equations, (3.4) and (3.7) respectively, follow from [19] and are constructed upon solving local problems on *cylindrical stars*. They allow for the nonuniform coefficient y^α, in the state and adjoint equations, and the anisotropic meshes in the family $\mathbb{T}$. The error indicators $\mathscr{E}_Z$ and $\mathscr{E}_\Lambda$ are related to the discretization of the control variable and the associated subgradient. We refer to this estimator as *ideal* since the computation of $\mathscr{E}_V$ and $\mathscr{E}_P$ involve the resolution of problems in infinite dimensional spaces. We derive, in Section 4, the equivalence between the ideal estimator (4.2) and the error without oscillation terms. Such an equivalence relies on a geometric condition imposed on the mesh that is independent of the exact optimal variables and is computationally implementable. This ideal estimator sets the basis to define, in Section 5, a computable error estimator, which is also decomposed as the sum of four contributions. This computable estimator is, under certain assumptions, equivalent, up to data oscillations terms, to the error.

Finally, we would like to mention that the aforementioned a posteriori analysis is of value not only for (1.2)–(1.4) but also for optimal control problems that requires the consideration of anisotropic meshes since rigorous anisotropic error estimates for such problems are scarce in the literature.

4.1. Preliminaries. We follow [19, Section 5.1] and introduce some notation and terminology. Given a node z on $\mathscr{T}_\mathcal{Y}$, in view of (2.1), we write $z = (z', z'')$ where z' and z'' correspond to nodes on $\mathscr{T}_\Omega$ and $\mathcal{I}_\mathcal{Y}$ respectively.

Given $K \in \mathscr{T}_\Omega$, we denote by $\mathcal{N}(K)$ and $\mathring{\mathcal{N}}(K)$ the set of nodes and interior nodes of K, respectively. We also define

$$\mathcal{N}(\mathscr{T}_\Omega) = \cup\{\mathcal{N}(K) : K \in \mathscr{T}_\Omega\}, \qquad \mathring{\mathcal{N}}(\mathscr{T}_\Omega) = \cup\{\mathring{\mathcal{N}}(K) : K \in \mathscr{T}_\Omega\}.$$

Given $T \in \mathscr{T}_{\mathcal{Y}}$, we define $\mathcal{N}(T)$, $\mathring{\mathcal{N}}(T)$, and then $\mathcal{N}(\mathscr{T}_{\mathcal{Y}})$ and $\mathring{\mathcal{N}}(\mathscr{T}_{\mathcal{Y}})$ accordingly.

Given $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, we define

$$S_{z'} := \bigcup \{K \in \mathscr{T}_\Omega : \ K \ni z'\} \subset \Omega$$

and the *cylindrical star* around z' as

$$\mathcal{C}_{z'} := \bigcup \{T \in \mathscr{T}_{\mathcal{Y}} : T = K \times I, \ K \ni z'\} = S_{z'} \times (0, \mathcal{Y}) \subset \mathcal{C}_{\mathcal{Y}}. \tag{4.3}$$

Given a cell $K \in \mathscr{T}_\Omega$ we define the *patch* S_K as $S_K := \bigcup_{z' \in K} S_{z'}$. For $T \in \mathscr{T}_{\mathcal{Y}}$, we define the patch S_T similarly. Given $z' \in \mathcal{N}(\mathscr{T}_\Omega)$ we define its *cylindrical patch* as

$$\mathcal{D}_{z'} := \bigcup \{\mathcal{C}_{w'} : w' \in S_{z'}\} \subset \mathcal{C}_{\mathcal{Y}}.$$

Finally, for each $z' \in \mathcal{N}(\mathscr{T}_\Omega)$ we set $h_{z'} := \min\{h_K : K \ni z'\}$.

4.2. Local weighted Sobolev spaces. In what follows we define the local weighted Sobolev spaces that will be useful for our analysis.

Definition 4.1 (local weighted Sobolev spaces). *Given $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, we define*

$$\mathbb{W}(\mathcal{C}_{z'}) = \left\{ w \in H^1(y^\alpha, \mathcal{C}_{z'}) : w = 0 \ \textit{on} \ \partial \mathcal{C}_{z'} \setminus \Omega \times \{0\} \right\}, \tag{4.4}$$

where $\mathcal{C}_{z'}$ denotes the cylindrical star around z' and is defined as in (4.3).

In view of the fact that y^α belongs to the class $A_2(\mathbb{R}^{n+1})$, the space $\mathbb{W}(\mathcal{C}_{z'})$ is Hilbert [26, 29, 31, 41, 58]. In addition, we have that [14, Proposition 2.1]

$$\| \operatorname{tr}_\Omega w \|_{L^2(S_{z'})} \leq C_{\operatorname{tr}_\Omega} \| \nabla w \|_{L^2(y^\alpha, \mathcal{C}_{z'})} \quad \forall w \in \mathbb{W}(\mathcal{C}_{z'}), \quad C_{\operatorname{tr}_\Omega} \leq d_s^{-\frac{1}{2}} \tag{4.5}$$

4.3. Auxiliary variables. In the next section we will propose an error estimator whose analysis relies on two auxiliary variables, that we will denote by $\tilde{\mathsf{r}}$ and $\tilde{\lambda}$ and define in what follows. With the operator $\Pi_{[\mathsf{a},\mathsf{b}]}$, defined as in (3.9), at hand, we define

$$\tilde{\lambda}(x') := \Pi_{[-1,1]} \left(-\frac{1}{\nu} \operatorname{tr}_\Omega \bar{P}(x') \right), \qquad \tilde{\mathsf{r}} := \Pi_{[\mathsf{a},\mathsf{b}]} \left(-\frac{1}{\sigma} \left(\operatorname{tr}_\Omega \bar{P}(x') + \nu \tilde{\lambda}(x') \right) \right). \tag{4.6}$$

In the next result we derive two important properties that will be essential in the a posteriori error analysis that we will perform.

Lemma 4.2. *Let $\tilde{\mathsf{r}} \in \mathsf{Z}_{ad}$ and $\tilde{\lambda} \in L^\infty(\Omega)$ be defined as in* (4.6). *Then $\tilde{\mathsf{r}}$ can be characterized by the variational inequality*

$$(\operatorname{tr}_\Omega \bar{P} + \sigma \tilde{\mathsf{r}} + \nu \tilde{\lambda}, \mathsf{r} - \tilde{\mathsf{r}})_{L^2(\Omega)} \geq 0 \quad \forall \mathsf{r} \in \mathsf{Z}_{ad}, \tag{4.7}$$

and $\tilde{\lambda} \in \partial \psi(\tilde{\mathsf{r}})$.

Proof. The variational inequality (4.7) follows immediately from the arguments elaborated in the proof of [57, Lemma 2.26]. It thus suffices to prove that $\tilde{\lambda} \in \partial\psi(\tilde{\mathsf{r}})$. To accomplish this task, we first assume that $\tilde{\mathsf{r}}(x') = 0$. In view of the definition of $\tilde{\lambda}$, we immediately conclude that $\tilde{\lambda}(x') \in [-1, 1]$. Let us now assume that $\tilde{\mathsf{r}}(x') > 0$. This, and (4.6), implies that

$$\operatorname{tr}_\Omega \bar{P}(x') + \nu \tilde{\lambda}(x') < 0 \iff \tilde{\lambda}(x') < -\frac{1}{\nu} \operatorname{tr}_\Omega \bar{P}(x') \implies \tilde{\lambda}(x') = 1.$$

Analogously, we obtain that, if $\tilde{\mathsf{r}}(x') < 0$, then $\tilde{\lambda}(x') = -1$. In conclusion, $\tilde{\lambda} \in L^\infty(\Omega)$ satisfies that

$$\tilde{\lambda}(x') = 1, \quad \tilde{\mathsf{r}}(x') > 0, \qquad \tilde{\lambda}(x') = -1, \quad \tilde{\mathsf{r}}(x') < 0, \quad \tilde{\lambda}(x') \in [-1,1], \quad \tilde{\mathsf{r}}(x') = 0.$$

This, that is equivalent to $\tilde{\lambda} \in \partial\psi(\tilde{\mathsf{r}})$, concludes the proof. ◻

4.4. A posteriori error analysis. In this section we design and study an *ideal* a posteriori error estimator for (1.2)–(1.4). We refer to such an estimator as ideal since it involves the resolution of local problems on the infinite dimensional spaces $\mathbb{W}(\mathcal{C}_{z'})$; the estimator is therefore not computable. However, it will provide the intuition to propose, in Section 5, a computable error estimator. We prove that it is equivalent to the error without oscillation terms; see Theorems 4.4 and 4.5 below.

We define the ideal a posteriori error estimator as the sum of four contributions:

$$\begin{aligned}\mathscr{E}^2_{\text{ocp}}(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathscr{T}_\mathscr{Y}) = \mathscr{E}^2_V(\bar{V},\bar{Z};\mathcal{N}(\mathscr{T}_\Omega)) + \mathscr{E}^2_P(\bar{P},\bar{V};\mathcal{N}(\mathscr{T}_\Omega)) \\ + \mathscr{E}^2_Z(\bar{Z},\bar{P};\mathscr{T}_\Omega) + \mathscr{E}^2_\Lambda(\bar{\Lambda},\bar{P};\mathscr{T}_\Omega), \quad (4.8)\end{aligned}$$

where $\bar{V}$, $\bar{P}$, $\bar{Z}$ and $\bar{\Lambda}$ denote the optimal variables associated to the fully discrete optimal control problem defined in Section 3.3. In what follows, we describe each contribution in (4.8) separately. To accomplish this task, we define the bilinear form

$$a_{z'} : \mathbb{W}(\mathcal{C}_{z'}) \times \mathbb{W}(\mathcal{C}_{z'}) \to \mathbb{R}, \quad a_{z'}(w,\phi) = \frac{1}{d_s}\int_{\mathcal{C}_{z'}} y^\alpha \nabla w \cdot \nabla \phi \,\mathrm{d}x. \qquad (4.9)$$

The first contribution in (4.8) corresponds to the a posteriori error estimator of [19, Section 5.3]. Let us define, for $z' \in \mathcal{N}(\mathscr{T}_\Omega)$,

$$\zeta_{z'} \in \mathbb{W}(\mathcal{C}_{z'}) : \quad a_{z'}(\zeta_{z'},\psi) = \langle \bar{Z}, \operatorname{tr}_\Omega \psi\rangle - a_{z'}(\bar{V},\psi) \quad \forall \psi \in \mathbb{W}(\mathcal{C}_{z'}). \qquad (4.10)$$

With this definition at hand, we define the posteriori error indicators and estimator

$$\mathscr{E}^2_V(\bar{V},\bar{Z};\mathcal{C}_{z'}) = \|\nabla \zeta_{z'}\|^2_{L^2(y^\alpha,\mathcal{C}_{z'})},\ \mathscr{E}^2_V(\bar{V},\bar{Z};\mathcal{N}(\mathscr{T}_\Omega)) = \sum_{z'\in\mathcal{N}(\mathscr{T}_\Omega)} \mathscr{E}^2_V(\bar{V},\bar{Z};\mathcal{C}_{z'}). \quad (4.11)$$

We now describe the second contribution in (4.8). Let us define, for $z' \in \mathcal{N}(\mathscr{T}_\Omega)$,

$$\chi_{z'} \in \mathbb{W}(\mathcal{C}_{z'}) : \quad a_{z'}(\chi_{z'},\psi) = \langle \operatorname{tr}_\Omega \bar{V} - \mathsf{u}_d, \operatorname{tr}_\Omega \psi\rangle - a_{z'}(\psi,\bar{P}) \quad \forall \psi \in \mathbb{W}(\mathcal{C}_{z'}). \quad (4.12)$$

We thus define the posteriori error indicators and estimator

$$\mathscr{E}^2_P(\bar{P},\bar{V};\mathcal{C}_{z'}) = \|\nabla \chi_{z'}\|^2_{L^2(y^\alpha,\mathcal{C}_{z'})}, \mathscr{E}^2_P(\bar{P},\bar{V};\mathcal{N}(\mathscr{T}_\Omega)) = \sum_{z'\in\mathcal{N}(\mathscr{T}_\Omega)} \mathscr{E}^2_P(\bar{P},\bar{V};\mathcal{C}_{z'}). \quad (4.13)$$

The third contribution in (4.8) is defined as follows:

$$\mathscr{E}_Z(\bar{Z},\bar{P};K) = \|\bar{Z} - \tilde{\mathsf{r}}\|_{L^2(K)}, \quad \mathscr{E}^2_Z(\bar{Z},\bar{P};\mathscr{T}_\Omega) = \sum_{K\in\mathscr{T}_\Omega} \mathscr{E}^2_Z(\bar{Z},\bar{P};K), \qquad (4.14)$$

where the auxiliary variable $\tilde{\mathsf{r}}$ is defined as in (4.6).

Finally, and having in mind the definition of the auxiliary variable $\tilde{\lambda}$, given as in (4.6), we define the fourth contribution in (4.8) as follows:

$$\mathscr{E}_\Lambda(\bar{\Lambda},\bar{P};K) = \|\bar{\Lambda} - \tilde{\lambda}\|_{L^2(K)}, \quad \mathscr{E}^2_\Lambda(\bar{\Lambda},\bar{P};\mathscr{T}_\Omega) = \sum_{K\in\mathscr{T}_\Omega} \mathscr{E}^2_Z(\bar{\Lambda},\bar{P};K). \qquad (4.15)$$

Since $\bar{V}$ can be seen as the finite element approximation of v, defined in (3.22), within the space $\mathbb{V}(\mathscr{T}_\mathscr{Y})$, the following result follows from [19, Proposition 5.14]:

$$\|\nabla(v - \bar{V})\|_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \mathscr{E}_V(\bar{V},\bar{Z};\mathcal{N}(\mathscr{T}_\Omega)).$$

Similarly,

$$\|\nabla(p-\bar{P})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim \mathscr{E}_P(\bar{P},\bar{V};\mathcal{N}(\mathscr{T}_\Omega)).$$

These estimates are essential to prove the estimate (4.22) below.

REMARK 4.3 (implementable geometric condition). The results of [19, Section 5.3] are valid under an implementable geometric condition that allows for the consideration of graded meshes in Ω and anisotropy in the extended variable y; the latter being necessary for optimal approximation. Having graded meshes in Ω is essential to compensate the singularities in the x'–variables that appear due to the violation of compatibility conditions on the forcing terms. In what follows we will assume the following *condition* over $\mathbb{T}$: there exists $C_\mathbb{T}>0$ such that, for $\mathscr{T}_\mathcal{Y}\in\mathbb{T}$, we have

$$h_\mathcal{Y} \leq C_\mathbb{T} h_{z'}, \tag{4.16}$$

for all interior nodes z' of $\mathscr{T}_\Omega$. The term $h_\mathcal{Y}$ denotes the largest size among all the elements in the mesh $I_\mathcal{Y}$. We finally mention that this condition is fully implementable.

To present the following result we define the errors $e_V := \bar{v}-\bar{V}$, $e_P := \bar{p}-\bar{P}$, $e_Z = \bar{\mathsf{r}}-\bar{Z}$, $e_\Lambda = \bar{t}-\bar{\Lambda}$, the vector $e=(e_V,e_P,e_Z,e_\Lambda)$, and the norm

$$|\!|\!| e |\!|\!|^2 := \|\nabla e_V\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + \|\nabla e_P\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + \|e_Z\|^2_{L^2(\Omega)} + \|e_\Lambda\|^2_{L^2(\Omega)}. \tag{4.17}$$

4.4.1. Reliability. We now prove a global reliability property for $\mathscr{E}_{\mathrm{ocp}}$, which is defined as in (4.8); the proof combines the arguments of [4] with elements of Sections 2.3 and 4.3.

THEOREM 4.4 (global upper bound). *Let* $(\bar{v},\bar{p},\bar{\mathsf{r}},\bar{t}) \in \mathring{H}^1_L(y^\alpha,\mathcal{C}_\mathcal{Y})\times\mathring{H}^1_L(y^\alpha,\mathcal{C}_\mathcal{Y})\times\mathsf{Z}_{ad}\times L^\infty(\Omega)$ *be the optimal variables associated to the truncated optimal control problem of Section 3.2 and* $(\bar{V},\bar{P},\bar{Z},\bar{\Lambda})\in\mathbb{V}(\mathscr{T}_\mathcal{Y})\times\mathbb{V}(\mathscr{T}_\mathcal{Y})\times\mathbb{Z}_{ad}(\mathscr{T}_\Omega)\times\mathbb{P}_0(\mathscr{T}_\Omega)$ *its numerical approximation as is described in Section 3.3. If* (4.16) *holds, then*

$$|\!|\!| e |\!|\!| \lesssim \mathscr{E}_{ocp}(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathscr{T}_\mathcal{Y}), \tag{4.18}$$

where the hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the meshes $\mathscr{T}_\Omega$ *and* $\mathscr{T}_\mathcal{Y}$, *and* $\#\mathscr{T}_\Omega$ *and* $\#\mathscr{T}_\mathcal{Y}$.

Proof. We proceed in five steps.

Step 1. Applying the triangle inequality we immediately observe that

$$\|\bar{\mathsf{r}}-\bar{Z}\|^2_{L^2(\Omega)} \leq 2\|\bar{\mathsf{r}}-\tilde{\mathsf{r}}\|^2_{L^2(\Omega)} + 2\|\tilde{\mathsf{r}}-\bar{Z}\|^2_{L^2(\Omega)} = 2\|\bar{\mathsf{r}}-\tilde{\mathsf{r}}\|^2_{L^2(\Omega)} + 2\mathscr{E}_Z^2, \tag{4.19}$$

where $\mathscr{E}_Z = \mathscr{E}_Z(\bar{Z},\bar{P};\mathscr{T}_\Omega)$ corresponds to the error estimator associated to the optimal control variable, defined in (4.14), and $\tilde{\mathsf{r}}$ denotes the auxiliary control variable defined in (4.6).

Step 2. The previous estimate reveals that it thus suffices to bound $\|\bar{\mathsf{r}}-\tilde{\mathsf{r}}\|_{L^2(\Omega)}$. To accomplish this task, we set $\mathsf{r}=\tilde{\mathsf{r}}$ in (3.6) and $\mathsf{r}=\bar{\mathsf{r}}$ in (4.7) and add the obtained inequalities to arrive at

$$\sigma\|\bar{\mathsf{r}}-\tilde{\mathsf{r}}\|^2_{L^2(\Omega)} \leq (\mathrm{tr}_\Omega(\bar{p}-\bar{P}),\tilde{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(\Omega)} + \nu(\bar{t}-\tilde{\lambda},\tilde{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(\Omega)}, \tag{4.20}$$

where $\tilde{\lambda}$ is defined as in (4.6), $\bar{t}\in\partial\psi(\bar{\mathsf{r}})$ and $\bar{p}$ and $\bar{P}$ solve (3.7) and (3.19), respectively. Invoking the results of Lemma 4.2 we obtain that $\tilde{\lambda}\in\partial\psi(\tilde{\mathsf{r}})$. This, in view of (2.7), implies that

$$(\bar{t}-\tilde{\lambda},\tilde{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(\Omega)} \leq 0,$$

and thus that

$$\sigma\|\bar{\mathsf{r}}-\tilde{\mathsf{r}}\|_{L^2(\Omega)}^2 \leq (\mathrm{tr}_\Omega(\bar{p}-\bar{P}),\tilde{\mathsf{r}}-\bar{\mathsf{r}})_{L^2(\Omega)}. \tag{4.21}$$

Step 3. The control of the right hand side of (4.21) follows from [4, Theorem 4.2]. In fact, we have that

$$\|\nabla e_V\|_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} + \|\nabla e_P\|_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} + \|e_Z\|_{L^2(\Omega)} \lesssim \mathscr{E}_{\mathrm{ocp}}(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathscr{T}_\mathscr{Y}). \tag{4.22}$$

Step 4. We now control the error in the approximation of the subgradient, i.e., $\|\bar{t}-\bar{\Lambda}\|_{L^2(\Omega)}$. To accomplish this task, we first apply the triangle inequality and write

$$\|\bar{t}-\bar{\Lambda}\|_{L^2(\Omega)} \leq \|\bar{t}-\tilde{\lambda}\|_{L^2(\Omega)} + \|\tilde{\lambda}-\bar{\Lambda}\|_{L^2(\Omega)}.$$

We now use the definition of $\tilde{\lambda}$, given by (4.6), and the the fact that $\Pi_{[-1,1]}$, defined as in (3.9), is a Lipschitz continuous map, to arrive at

$$\begin{aligned}\|\bar{t}-\bar{\Lambda}\|_{L^2(\Omega)}^2 &\lesssim \|\Pi_{[-1,1]}(-\tfrac{1}{\nu}\,\mathrm{tr}_\Omega\,\bar{p}) - \Pi_{[-1,1]}(-\tfrac{1}{\nu}\,\mathrm{tr}_\Omega\,\bar{P})\|_{L^2(\Omega)}^2 + \mathscr{E}_\Lambda^2(\bar{\Lambda},\bar{P};\mathscr{T}_\Omega)\\ &\lesssim \|\,\mathrm{tr}_\Omega(\bar{p}-\bar{P})\|_{L^2(\Omega)}^2 + \mathscr{E}_\Lambda^2(\bar{\Lambda},\bar{P};\mathscr{T}_\Omega) \lesssim \mathscr{E}_{\mathrm{ocp}}(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathscr{T}_\mathscr{Y}).\end{aligned} \tag{4.23}$$

Step 5. The desired estimate (4.18) follows from (4.22) and (4.23). This concludes the proof. ☐

4.4.2. Efficiency. We now continue with the study of the a *ideal* posteriori error estimator $\mathscr{E}_{\mathrm{ocp}}$ and derive its local efficiency. In order to present the next result we define the constant $\mathfrak{C}(d_s,\sigma,\nu)$, that depends only on d_s, the regularization parameter σ and the sparsity parameter ν, by

$$\mathfrak{C}(d_s,\sigma,\nu) = \max\{2d_s^{-1}, d_s^{-\frac{1}{2}}+1, d_s^{-\frac{1}{2}}(\nu^{-1}+2\sigma^{-1}+d_s^{-\frac{1}{2}}), 1\}. \tag{4.24}$$

THEOREM 4.5 (local lower bound). *Let* $(\bar{v},\bar{p},\bar{\mathsf{r}},\bar{t}) \in \mathring{H}_L^1(y^\alpha,\mathcal{C}_\mathscr{Y}) \times \mathring{H}_L^1(y^\alpha,\mathcal{C}_\mathscr{Y}) \times \mathsf{Z}_{ad}\times L^\infty(\Omega)$ *be the optimal variables associated to the truncated optimal control problem of Section 3.2 and* $(\bar{V},\bar{P},\bar{Z},\bar{\Lambda}) \in \mathbb{V}(\mathscr{T}_\mathscr{Y})\times\mathbb{V}(\mathscr{T}_\mathscr{Y})\times\mathbb{Z}_{ad}(\mathscr{T}_\Omega)\times\mathbb{P}_0(\mathscr{T}_\Omega)$ *its numerical approximation as is described in Section 3.3. If* $z' \in \mathcal{N}(\mathscr{T}_\Omega)$*, then*

$$\begin{aligned}\mathscr{E}_V(\bar{V},\bar{Z};\mathcal{C}_{z'}) + \mathscr{E}_P(\bar{P},\bar{V};\mathcal{C}_{z'}) + \mathscr{E}_Z(\bar{Z},\bar{P};S_{z'}) + \mathscr{E}_\Lambda(\bar{\Lambda},\bar{P};S_{z'}) \leq \mathfrak{C}(d_s,\sigma,\nu)\\ \cdot\left(\|\nabla e_V\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + \|\nabla e_P\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + \|e_Z\|_{L^2(S_{z'})} + \|e_\Lambda\|_{L^2(S_{z'})}\right),\end{aligned} \tag{4.25}$$

where $\mathfrak{C}(d_s,\sigma,\nu)$ *depends only on* d_s*,* σ*, and* ν *and is defined as in* (4.24).

Proof. We proceed in five steps.

Step 1. The objective of this step is to study the efficiency properties of the local indicator $\mathscr{E}_V$ defined in (4.11). Let $z' \in \mathcal{N}(\mathscr{T}_\Omega)$. Thus, on the basis of the fact $\zeta_{z'} \in \mathbb{W}(\mathcal{C}_{z'})$ solves (4.10) and that $\langle\bar{\mathsf{r}},\mathrm{tr}_\Omega\,\zeta_{z'}\rangle = a_\mathscr{Y}(\bar{v},\zeta_{z'}) = a_{z'}(\bar{v},\zeta_{z'})$, which follows from setting $\phi=\zeta_{z'}$ in (3.4), we obtain that

$$\begin{aligned}\mathscr{E}_V^2(\bar{V},\bar{Z};\mathcal{C}_{z'}) = a_{z'}(\zeta_{z'},\zeta_{z'}) &= \langle\bar{\mathsf{r}},\mathrm{tr}_\Omega\,\zeta_{z'}\rangle + \langle\bar{Z}-\bar{\mathsf{r}},\mathrm{tr}_\Omega\,\zeta_{z'}\rangle - a_{z'}(\bar{V},\zeta_{z'})\\ &= a_{z'}(\bar{v}-\bar{V},\zeta_{z'}) + \langle\bar{Z}-\bar{\mathsf{r}},\mathrm{tr}_\Omega\,\zeta_{z'}\rangle.\end{aligned} \tag{4.26}$$

Use now the local version of the trace estimate (4.5) with $C_{\mathrm{tr}_\Omega} \leq d_s^{-\frac{1}{2}}$ and the definition of the local bilinear form $a_{z'}$, given by (4.9), to conclude that

$$\mathscr{E}_V^2(\bar{V},\bar{Z};\mathcal{C}_{z'}) \leq \left(d_s^{-1}\|\nabla e_V\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + d_s^{-\frac{1}{2}}\|e_Z\|_{L^2(S_{z'})}\right)\|\nabla\zeta_{z'}\|_{L^2(y^\alpha,\mathcal{C}_{z'})},$$

which, on the basis of (4.11), immediately yields the local efficiency of $\mathscr{E}_V$:

$$\mathscr{E}_V(\bar{V},\bar{Z};\mathcal{C}_{z'}) \le d_s^{-1}\|\nabla e_V\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + d_s^{-\frac{1}{2}}\|e_Z\|_{L^2(S_{z'})}. \tag{4.27}$$

Step 2. Similar arguments to the ones that lead to (4.26) reveal that

$$\mathscr{E}_P^2(\bar{P},\bar{V};\mathcal{C}_{z'}) = a_{z'}(\chi_{z'},e_P) - \langle \operatorname{tr}_\Omega e_V, \operatorname{tr}_\Omega \chi_{z'}\rangle,$$

where $\chi_{z'} \in \mathbb{W}(\mathcal{C}_{z'})$ solves (4.12). Thus,

$$\mathscr{E}_P(\bar{P},\bar{V};\mathcal{C}_{z'}) \le d_s^{-1}\|\nabla e_P\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + d_s^{-1}\|\nabla e_V\|_{L^2(y^\alpha,\mathcal{C}_{z'})}. \tag{4.28}$$

Step 3. The goal of this step is to obtain local efficiency properties for the indicator $\mathscr{E}_\Lambda$, given by (4.15). To accomplish this task, we invoke the projection formula (3.12) for $\bar{t}$ and the definition of $\tilde{\lambda}$, given by (4.6), to obtain, for $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, that

$$\mathscr{E}_\Lambda(\bar{\Lambda},\bar{P};S_{z'}) \le \|\bar{\Lambda}-\bar{t}\|_{L^2(S_{z'})} + \|\Pi_{[-1,1]}(-\tfrac{1}{\nu}\operatorname{tr}_\Omega \bar{p}) - \Pi_{[-1,1]}(-\tfrac{1}{\nu}\operatorname{tr}_\Omega \bar{P})\|_{L^2(S_{z'})}.$$

This, in view of the Lipschitz continuity of the operator $\Pi_{[-1,1]}$ and the local trace estimate (4.5), implies that

$$\mathscr{E}_\Lambda(\bar{\Lambda},\bar{P};S_{z'}) \le \|e_\Lambda\|_{L^2(S_{z'})} + \nu^{-1}d_s^{-\frac{1}{2}}\|\nabla e_P\|_{L^2(y^\alpha,\mathcal{C}_{z'})}. \tag{4.29}$$

Step 4. In this step we study the efficiency properties of $\mathscr{E}_Z$, which is defined by (4.14). We proceed using similar arguments to the ones that lead to (4.29). In fact, let $z' \in \mathcal{N}(\mathscr{T}_\Omega)$ and invoke the projection formula (3.10), for $\bar{\mathsf{r}}$, and the definition of $\tilde{\mathsf{r}}$, given by (4.6). We thus obtain that

$$\begin{aligned}\mathscr{E}_Z(\bar{Z},\bar{P};S_{z'}) \le{}& \|\bar{Z}-\bar{\mathsf{r}}\|_{L^2(S_{z'})} \\ &+ \left\|\Pi_{[\mathsf{a},\mathsf{b}]}\left[-\tfrac{1}{\sigma}\left(\operatorname{tr}_\Omega \bar{p}+\nu\bar{t}\right)\right] - \Pi_{[\mathsf{a},\mathsf{b}]}\left[-\tfrac{1}{\sigma}\left(\operatorname{tr}_\Omega \bar{P}+\nu\tilde{\lambda}\right)\right]\right\|_{L^2(S_{z'})},\end{aligned}$$

which, in view of the Lipschitz continuity of $\Pi_{[\mathsf{a},\mathsf{b}]}$, implies that

$$\mathscr{E}_Z(\bar{Z},\bar{P};S_{z'}) \le \|\bar{Z}-\bar{\mathsf{r}}\|_{L^2(S_{z'})} + \frac{1}{\sigma}\|\operatorname{tr}_\Omega(\bar{p}-\bar{P})\|_{L^2(S_{z'})} + \frac{\nu}{\sigma}\|\bar{t}-\tilde{\lambda}\|_{L^2(S_{z'})}.$$

We now invoke the local trace estimate (4.5) and the fact that $\tilde{\lambda} = \Pi_{[-1,1]}(-\frac{1}{\nu}\operatorname{tr}_\Omega \bar{P})$ to arrive at

$$\mathscr{E}_Z(\bar{Z},\bar{P};S_{z'}) \le \|e_Z\|_{L^2(S_{z'})} + 2\sigma^{-1}d_s^{-\frac{1}{2}}\|\nabla e_P\|_{L^2(y^\alpha,\mathcal{C}_{z'})}. \tag{4.30}$$

Step 5. We thus conclude the proof of the local efficiency property (4.25) by collecting the estimates (4.27), (4.28), (4.29), and (4.30). □

Remark 4.6 (Local efficiency of each contribution). The arguments elaborated in the proof of Theorem 4.5 reveal that each contribution $\mathscr{E}_V$, $\mathscr{E}_P$, $\mathscr{E}_\Lambda$ and $\mathscr{E}_Z$ is locally efficient; see estimates (4.27), (4.28), (4.29), and (4.30). In addition, we mention that the constants involved in such efficiency results are fully computable and depend only on d_s, ν, and σ.

5. A computable a posteriori error estimator. The results of Theorems 4.4 and 4.5 show that the a posteriori error estimator $\mathscr{E}_{ocp}$, defined as in (4.8), is globally reliable and locally efficient. Notice that both properties do not involve any oscillation term. Consequently, the error estimator $\mathscr{E}_{ocp}$ is equivalent to the error $\|\!|e|\!\|$. However, it has an insurmountable drawback: it requires, for each node z', the computation of the contributions $\mathscr{E}_V$ and $\mathscr{E}_P$ that in turn requires the resolution of the local problems (4.10) and (4.12), respectively. Since the latter problems are posed on the infinite–dimensional space $\mathbb{W}(\mathcal{C}_{z'})$, the error estimator (4.8) is thus not computable. In spite of this fact, and as we will see in this section, it provides the intuition to define an anisotropic and computable posteriori error estimator that allows for the nonuniform coefficient y^α in both, the state and the adjoint equations.

Let us begin our discussion with the following definition.

DEFINITION 5.1 (discrete local spaces). *For $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, we define*

$$\begin{aligned}\mathcal{W}(\mathcal{C}_{z'}) = \big\{ &W \in C^0(\overline{\mathcal{C}_{z'}}) : W|_T \in \mathcal{P}_2(K) \otimes \mathbb{P}_2(I)\ \forall T = K \times I \in \mathcal{C}_{z'},\\ &W|_{\partial\mathcal{C}_{z'} \setminus \Omega \times \{0\}} = 0 \big\}.\end{aligned}$$

If K is a quadrilateral, $\mathcal{P}_2(K)$ stands for $\mathbb{Q}_2(K)$. When K is a simplex, $\mathcal{P}_2(K)$ corresponds to $\mathbb{P}_2(K) \oplus \mathbb{B}(K)$, where $\mathbb{B}(K)$ stands for the space spanned by a local cubic bubble function.

With the discrete space $\mathcal{W}(\mathcal{C}_{z'})$ at hand, we thus define the discrete functions $\eta_{z'}$ and $\theta_{z'}$ as follows:

$$\eta_{z'} \in \mathcal{W}(\mathcal{C}_{z'}): \quad a_{z'}(\eta_{z'}, W) = \langle \bar{Z}, \operatorname{tr}_\Omega W \rangle - a_{z'}(\bar{V}, W) \tag{5.1}$$

for all $W \in \mathcal{W}(\mathcal{C}_{z'})$, and

$$\theta_{z'} \in \mathcal{W}(\mathcal{C}_{z'}): \quad a_{z'}(W, \theta_{z'}) = \langle \operatorname{tr}_\Omega \bar{V} - \mathsf{u}_d, \operatorname{tr}_\Omega W \rangle - a_{z'}(W, \bar{P}) \tag{5.2}$$

for all $W \in \mathcal{W}(\mathcal{C}_{z'})$. Notice that problems (5.1) and (5.2) correspond to the Galerkin approximations of problems (4.10) and (4.12), respectively.

Inspired in the definition of the ideal estimator $\mathscr{E}_{\mathrm{ocp}}$, given by (4.8), we define its computable counterpart, which we denote by $\mathcal{E}_{\mathrm{ocp}}$, as follows:

$$\begin{aligned}\mathcal{E}^2_{\mathrm{ocp}}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; \mathscr{T}_{\mathcal{Y}}) = \mathcal{E}^2_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathscr{T}_\Omega)) + \mathcal{E}^2_P(\bar{P}, \bar{V}; \mathcal{N}(\mathscr{T}_\Omega))&\\ + \mathcal{E}^2_Z(\bar{Z}, \bar{P}; \mathscr{T}_\Omega) + \mathcal{E}^2_\Lambda(\bar{\Lambda}, \bar{P}; \mathscr{T}_\Omega).&\end{aligned} \tag{5.3}$$

We recall that $\bar{V}$, $\bar{P}$, $\bar{Z}$, and $\bar{\Lambda}$ denote the optimal variables associated to the fully discrete optimal control problem of Section 3.3.

In what follows we describe each contribution in (5.3). First, having $\eta_{z'} \in \mathcal{W}(\mathcal{C}_{z'})$ at hand, defined as in (5.1), we define the local error indicators and error estimator, associated to the state equation, as

$$\mathcal{E}^2_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) = \|\nabla \eta_{z'}\|^2_{L^2(y^\alpha, \mathcal{C}_{z'})} \tag{5.4}$$

and

$$\mathcal{E}^2_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathscr{T}_\Omega)) = \sum_{z' \in \mathcal{N}(\mathscr{T}_\Omega)} \mathcal{E}^2_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}), \tag{5.5}$$

respectively. The second contribution in (5.3), that is associated to the discretization of the adjoint equation, is defined similarly. We define the local error indicators and error estimator as

$$\mathcal{E}^2_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) = \|\nabla \theta_{z'}\|^2_{L^2(y^\alpha, \mathcal{C}_{z'})} \tag{5.6}$$

and

$$\mathcal{E}_P^2(\bar{P},\bar{V};\mathcal{N}(\mathscr{T}_\Omega)) = \sum_{z'\in\mathcal{N}(\mathscr{T}_\Omega)} \mathcal{E}_P^2(\bar{P},\bar{V};\mathcal{C}_{z'}), \tag{5.7}$$

respectively. The third and fourth contributions, $\mathcal{E}_Z$ and $\mathcal{E}_\Lambda$, associated to the discretization of the control and its subgradient, are defined exactly as in (4.14) and (4.15), respectively.

5.1. Efficiency. The next result shows the local efficiency of the computable a posteriori error estimator $\mathcal{E}_{\text{ocp}}$.

Theorem 5.2 (local lower bound). *Let $(\bar{v},\bar{p},\bar{r},\bar{t}) \in \mathring{H}_L^1(y^\alpha,\mathcal{C}_\mathscr{Y}) \times \mathring{H}_L^1(y^\alpha,\mathcal{C}_\mathscr{Y}) \times \mathsf{Z}_{ad} \times L^\infty(\Omega)$ be the optimal variables associated to the truncated optimal control problem of Section 3.2 and $(\bar{V},\bar{P},\bar{Z},\bar{\Lambda}) \in \mathbb{V}(\mathscr{T}_\mathscr{Y})\times\mathbb{V}(\mathscr{T}_\mathscr{Y})\times\mathbb{Z}_{ad}(\mathscr{T}_\Omega)\times\mathbb{P}_0(\mathscr{T}_\Omega)$ its numerical approximation as is described in Section 3.3. If $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, then*

$$\begin{aligned}\mathcal{E}_V(\bar{V},\bar{Z};\mathcal{C}_{z'}) + \mathcal{E}_P(\bar{P},\bar{V};\mathcal{C}_{z'}) + \mathcal{E}_Z(\bar{Z},\bar{P};S_{z'}) + \mathcal{E}_\Lambda(\bar{\Lambda},\bar{P};S_{z'}) \leq \mathfrak{C}(d_s,\sigma,\nu)\\ \cdot\left(\|\nabla e_V\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + \|\nabla e_P\|_{L^2(y^\alpha,\mathcal{C}_{z'})} + \|e_Z\|_{L^2(S_{z'})} + \|e_\Lambda\|_{L^2(S_{z'})}\right),\end{aligned} \tag{5.8}$$

where $\mathfrak{C}(d_s,\sigma,\nu)$ depends only on d_s, σ and ν and is defined as in (4.24).

Proof. The local efficiency properties of the estimators $\mathcal{E}_\Lambda$ and $\mathcal{E}_Z$ are derived in (4.29) and (4.30), respectively. The local efficiency of $\mathcal{E}_V$ and $\mathcal{E}_P$ follow similar arguments to the ones that lead to (4.27) and (4.28). For brevity, we skip details. ☐

Remark 5.3 (strong efficiency). The lower bound (5.8) does not involve any oscillation term and, in addition, it is fully computable in the sense that the involved constant, $\mathfrak{C}(d_s,\sigma,\nu)$ is known. These properties imply a strong concept of efficiency: the relative size of the local error indicator dictates mesh refinement, which is independent of fine structure of the data.

5.2. Reliability. We now analyze the reliability properties of the computable a posteriori error estimator $\mathcal{E}_{ocp}$. To accomplish this task, we introduce, for $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, the local *oscillation* of $f \in L^2(\Omega)$ as

$$\mathsf{osc}(f;S_{z'}) = h_{z'}^s\|f - f_{z'}\|_{L^2(S_{z'})}, \quad f_{z'}|_K = \fint_K f, \tag{5.9}$$

where $h_{z'} = \min\{h_K : K \ni z'\}$. We also define the global oscillation of f as

$$\mathsf{osc}^2(f;\mathscr{T}_\Omega) = \sum_{z'\in\mathcal{N}(\mathscr{T}_\Omega)} \mathsf{osc}^2(f;S_{z'}). \tag{5.10}$$

With these definition at hand, we define, for $z' \in \mathcal{N}(\mathscr{T}_\Omega)$, the *total error indicator*

$$\mathfrak{E}^2(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathcal{C}_{z'}) = \mathcal{E}_{\text{ocp}}^2(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathcal{C}_{z'}) + \mathsf{osc}^2(\text{tr}_\Omega\,\bar{V} - \mathsf{u}_d;S_{z'}), \tag{5.11}$$

which will be essential in the MARK module of the AFEM that we will propose in Section 6.

Let $\mathscr{K}_{\mathscr{T}_\Omega} = \{S_{z'} : z' \in \mathcal{N}(\mathscr{T}_\Omega)\}$. We set $\mathscr{K}_\mathscr{Y} = \mathscr{K}_{\mathscr{T}_\Omega} \times (0,\mathscr{Y})$ and, for any $\mathscr{M} \subset \mathscr{K}_{\mathscr{T}_\Omega}$, $\mathscr{M}_\mathscr{Y} = \mathscr{M} \times (0,\mathscr{Y})$ and

$$\mathfrak{E}^2(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathscr{M}_\mathscr{Y}) = \sum_{S_{z'}\in\mathscr{M}} \mathfrak{E}^2(\bar{V},\bar{P},\bar{Z},\bar{\Lambda};\mathcal{C}_{z'}), \tag{5.12}$$

where, we recall that $\mathcal{C}_{z'} = S_{z'} \times (0, \mathcal{Y})$.

REMARK 5.4 (Marking). We follow [17, Remark 4.4] and comment that, the AFEM of Section 6 will utilize the total error indicator (5.11), namely the sum of the computable error estimator and the oscillation term, for marking. If the estimate

$$\mathcal{E}_{\text{ocp}}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; \mathscr{K}_{\mathcal{Y}}) \geq C \mathsf{osc}(\operatorname{tr}_\Omega \bar{V} - \mathsf{u}_d; \mathscr{K}_{\mathscr{T}_\Omega})$$

holds for $C > 0$, marking with respect to the total error indicator could be avoided. This estimate holds for the residual estimator with $C = 1$. However, it does not hold for other error estimators as the one that we are analyzing in this work. The oscillation cannot be removed for marking without further assumptions. We refer the reader to [17] for a complete discussion on this matter.

Notice that, since $\bar{V}$ corresponds to the finite element approximation of v, defined in (3.22), within the space $\mathbb{V}(\mathscr{T}_{\mathcal{Y}})$, we have that [19, Theorem 5.37]:

$$\|\nabla(v - \bar{V})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} \lesssim \mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{N}(\mathscr{T}_\Omega)).$$

Notice that the oscillation of $\bar{Z}$ vanishes since it is a piecewise constant function on $\mathscr{T}_\Omega$. Similarly, we have that

$$\|\nabla(p - \bar{P})\|_{L^2(y^\alpha, \mathcal{C}_{\mathcal{Y}})} \lesssim \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{N}(\mathscr{T}_\Omega)) + \mathsf{osc}(\operatorname{tr}_\Omega \bar{V} - \mathsf{u}_d; \mathscr{T}_\Omega).$$

We now state the global reliability of the computable error estimator $\mathcal{E}_{ocp}$.

THEOREM 5.5 (global upper bound). *Let $(\bar{v}, \bar{p}, \bar{\mathsf{r}}, \bar{t}) \in \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \times \mathring{H}^1_L(y^\alpha, \mathcal{C}_{\mathcal{Y}}) \times \mathsf{Z}_{ad} \times L^\infty(\Omega)$ be the optimal variables associated to the truncated optimal control problem of Section 3.2 and $(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}) \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}}) \times \mathbb{V}(\mathscr{T}_{\mathcal{Y}}) \times \mathbb{Z}_{ad}(\mathscr{T}_\Omega) \times \mathbb{P}_0(\mathscr{T}_\Omega)$ its numerical approximation as described in Section 3.3. If* (4.16) *and [19, Conjecture 5.28] hold, then*

$$|||e||| \lesssim \mathfrak{E}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; \mathscr{K}_{\mathcal{Y}}), \tag{5.13}$$

where the hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the meshes $\mathscr{T}_\Omega$ and $\mathscr{T}_{\mathcal{Y}}$, and $\#\mathscr{T}_\Omega$ and $\#\mathscr{T}_{\mathcal{Y}}$.

Proof. The proof of the estimate (5.13) follows closely the arguments developed in the proof of Theorem 4.4; the difference being the use of the computable error indicator $\mathcal{E}_{\text{ocp}}$ instead of the ideal estimator $\mathscr{E}_{\text{ocp}}$. ☐

REMARK 5.6 (Conjecture). The proof of Theorem 5.5 is valid under the assumption that there exists an operator $\mathcal{M}_{z'}$ that verifies the conditions stipulated in [19, Conjecture 5.28]. The construction of $\mathcal{M}_{z'}$ is an open problem. The numerical experiments performed for the state equation and the ones that we perform in Section 6 provide computational evidence that the bound (5.13) holds with no oscillation terms and thus indirect evidence of the existence of the aforementioned operator $\mathcal{M}_{z'}$ with the requisite properties [19, (5.29)–(5.31)].

6. Numerical Experiments.

In this section we explore computationally the performance of the computable a posteriori error estimator $\mathcal{E}_{ocp}$, defined as in (5.3), with a series of numerical examples. To accomplish this task, we start, in the next section, with the design of an AFEM that is based on iterations of the loop (4.1).

6.1. Design of the AFEM.

We describe the four modules in (4.1):

- SOLVE: Given the anisotropic mesh $\mathscr{T}_{\mathcal{Y}} \in \mathbb{T}$, we compute the unique solution $(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}) \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}}) \times \mathbb{V}(\mathscr{T}_{\mathcal{Y}}) \times \mathbb{Z}_{ad}(\mathscr{T}_{\Omega}) \times \mathbb{P}_0(\mathscr{T}_{\Omega})$ to the fully discrete optimal control problem of Section 3.3:

$$(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}) = \mathsf{SOLVE}(\mathscr{T}_{\mathcal{Y}}).$$

  The aforementioned discrete problem is solved by using the active set strategy of [54, Algorithm 2].
- $\mathsf{ESTIMATE}$: With the discrete solution at hand, we compute, for each $z' \in \mathcal{N}(\mathscr{T}_{\Omega})$, the local error indicator $\mathcal{E}_{\mathrm{ocp}}$, which we define as follows:

$$\begin{aligned}\mathcal{E}_{\mathrm{ocp}}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; \mathcal{C}_{z'}) = \mathcal{E}_V(\bar{V}, \bar{Z}; \mathcal{C}_{z'}) &+ \mathcal{E}_P(\bar{P}, \bar{V}; \mathcal{C}_{z'}) \\ &+ \mathcal{E}_Z(\bar{Z}, \bar{P}; S_{z'}) + \mathcal{E}_\Lambda(\bar{\Lambda}, \bar{P}; S_{z'}).\end{aligned}$$

  The local indicators $\mathcal{E}_V$, $\mathcal{E}_P$, $\mathcal{E}_Z$, and $\mathcal{E}_\Lambda$ are defined as in (5.4), (5.6), (4.14), and (4.15), respectively. Once the local indicator $\mathcal{E}_{\mathrm{ocp}}$ is obtained, we compute the local oscillation term

$$\mathsf{osc}(\mathrm{tr}_\Omega \bar{V} - \mathsf{u}_d; S_{z'})$$

  and then construct the total error indicator $\mathfrak{E}$, which is defined as in (5.11):

$$\left\{\mathfrak{E}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; S_{z'})\right\}_{S_{z'} \in \mathscr{K}_{\mathscr{T}_\Omega}} = \mathsf{ESTIMATE}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; \mathscr{T}_{\mathcal{Y}}), \tag{6.1}$$

  where $\mathscr{K}_{\mathscr{T}_\Omega} = \{S_{z'} : z' \in \mathcal{N}(\mathscr{T}_\Omega)\}$. Notice that, for notational convenience, we replaced $\mathcal{C}_{z'}$ by $S_{z'}$ in (6.1).
- MARK: We follow the numerical evidence presented in [35, Section 6] and use the *maximum strategy* as a marking technique for our AFEM: Given $\theta \in [0, 1]$ and the set of indicators obtained in the previous step, we select the set

$$\mathscr{M} = \mathsf{MARK}\left(\left\{\mathfrak{E}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; S_{z'})\right\}_{S_{z'} \in \mathscr{K}_{\mathscr{T}_\Omega}}\right) \subset \mathscr{K}_{\mathscr{T}_\Omega},$$

  that contains the stars $S_{z'}$ in $\mathscr{K}_{\mathscr{T}_\Omega}$ that satisfies

$$\mathfrak{E}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; S_{z'}) \geq \theta \mathfrak{E}_{\max}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}),$$

  where $\mathfrak{E}_{\max}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}) := \max\{\mathfrak{E}(\bar{V}, \bar{P}, \bar{Z}, \bar{\Lambda}; S_{z'}) : S_{z'} \in \mathscr{K}_{\mathscr{T}_\Omega}\}$.
- $\mathsf{REFINEMENT}$: We bisect all the elements $K \in \mathscr{T}_\Omega$ that are contained in $\mathscr{M}$ with the newest–vertex bisection method [45, 46] and create a new mesh $\mathscr{T}_\Omega'$. In the truncated optimal control problem, we choose the truncation parameter $\mathcal{Y}$ as $\mathcal{Y} = 1 + (1/3)\log(\#\mathscr{T}_\Omega')$. As a consequence of this election, the approximation and truncation errors are balanced [43, Remark 5.5]. Next, we generate a new mesh in the extended dimension, which we denote by $\mathcal{I}_{\mathcal{Y}}'$. The latter is constructed on the basis of (3.14): the number of degrees of freedom M is chosen to be a sufficiently large number in order to guarantee condition (4.16). This is achieved by first designing a partition $\mathcal{I}_{\mathcal{Y}}'$ with $M \approx (\#\mathscr{T}_\Omega')^{1/n}$ and checking (4.16). If this condition is not satisfied, we increase the number of mesh points until we obtain the desired result. Consequently,

$$\mathscr{T}_{\mathcal{Y}}' = \mathsf{REFINE}(\mathscr{M}), \qquad \mathscr{T}_{\mathcal{Y}}' := \mathscr{T}_\Omega' \otimes \mathcal{I}_{\mathcal{Y}}'.$$

6.2. Implementation. The proposed AFEM is implemented within the MATLAB software library *i*FEM [18]. The stiffness matrices associated to the finite element approximation of the state and adjoint equations are assembled exactly, while the forcing terms, in both equations, are computed with the help of a quadrature formula that is exact for polynomials of degree 4. The optimality system associated to the fully discrete control problem of Section 3.3 is solved by using the active–set strategy of [54, Algorithm 2].

To compute $\eta_{z'}$, which is defined as in (5.1), we follow [19, Section 6]: we loop around each node $z' \in \mathcal{N}(\mathscr{T}_\Omega)$ and collect data about the cylindrical star $\mathcal{C}_{z'}$. We thus assemble the small linear system in (5.1) and solve it with the built-in *direct solver* of MATLAB. The integrals that involve the weight y^α and discrete functions are computed exactly, while the ones that also involve data functions are computed element–wise by a quadrature formula that is exact for polynomials of degree 7. The computation of $\theta_{z'}$, which is defined as in (5.2), is similar.

In the MARK step, we modify the estimator from star–wise to element–wise. To do this, we first scale the nodal–wise estimator as $\mathscr{E}_{z'}^2/(\#S_{z'})$ and then, compute

$$\mathscr{E}_K^2 := \sum_{z' \in K} \mathscr{E}_{z'}^2, \quad K \in \mathscr{T}_\Omega.$$

Consequently, we have that $\sum_{K \in \mathscr{T}_\Omega} \mathscr{E}_K^2 = \sum_{z' \in \mathcal{N}(\mathscr{T}_\Omega)} \mathscr{E}_{z'}^2$. The cell–wise data oscillation is thus defined as

$$\mathsf{osc}_K(f)^2 := h_K^{2s} \|f - \bar{f}_K\|_{L^2(K)}^2,$$

where $\bar{f}_K = |K|^{-1} \int_K f \, \mathrm{d}x'$. The data oscillation is computed using a quadrature formula that is exact for polynomials of degree 7.

We finally mention that all the computations that we present in the next section are done without explicitly enforcing the mesh restriction (4.16), which provide experimental evidence to support the fact that (4.16) is nothing but an artifact in our proofs.

6.3. L–shaped domain with incompatible data. The a priori theory of [48] relies on the assumptions that Ω is convex and u_d belongs to $\mathbb{H}^{1-s}(\Omega)$; the latter implies that u_d vanishes on $\partial\Omega$ for $s \in (0, \frac{1}{2}]$. Let us thus consider a case where these conditions are not met:

- The regularization and sparsity parameters are $\sigma = 0.1$ and $\nu = 0.5$ when $s \leq \frac{1}{2}$ and $\sigma = 0.05$ and $\nu = 0.1$ when $s > \frac{1}{2}$.
- The control bounds a and b are such that $\mathsf{a} = -0.3$ and $\mathsf{b} = 0.3$.
- The domain $\Omega = (-1,1)^2 \setminus (0,1) \times (-1,0)$ is non-convex.
- The desired state is $\mathsf{u}_d = 1$. Notice that u_d is not compatible, in the sense that $\mathsf{u}_d \notin \mathbb{H}^{1-s}(\Omega)$ for $s \in (0, \frac{1}{2}]$. This has as a consequence that the singular behavior (1.5)–(1.6) for the optimal adjoint variable $\bar{\mathsf{p}}$ is expected.
- The parameter θ that governs the module MARK is $\theta = 0.5$.

The results of Figure 6.1 show that the AFEM proposed in Section 6.1 delivers optimal experimental rates of convergence for the total error estimator $\mathfrak{E}$ and all choices of the parameter s considered.

In Figure 6.2 we present, for $s = 0.3$ (left) and $s = 0.9$ (right), the computational rates of convergence for each of the four contributions of the computable error estimator $\mathcal{E}_{ocp}$: $\mathcal{E}_V$, $\mathcal{E}_P$, $\mathcal{E}_Z$, and $\mathcal{E}_\Lambda$. It can be observed that, in both cases, each contribution decays with the optimal rate $\#(\mathscr{T}_\mathscr{Y})^{-1/3}$.

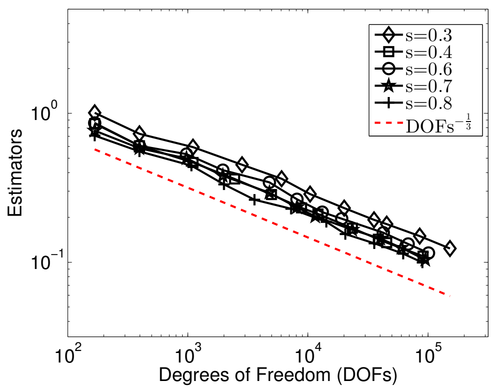


FIG. 6.1. *Computational rates of convergence for our anisotropic AFEM with incompatible desired data u_d and a non-convex domain $\Omega = (-1,1)^2 \setminus (0,1) \times (-1,0)$; $n = 2$ and $s = 0.3, 0.4, 0.6, 0.7,$ and $s = 0.8$. The panel shows the decrease of the total error indicator $\mathfrak{E}$ with respect to the number of degrees of freedom (DOFs). In all the presented cases the optimal rate $(\#\mathscr{T}_{\mathcal{Y}})^{-1/3}$ is achieved.*

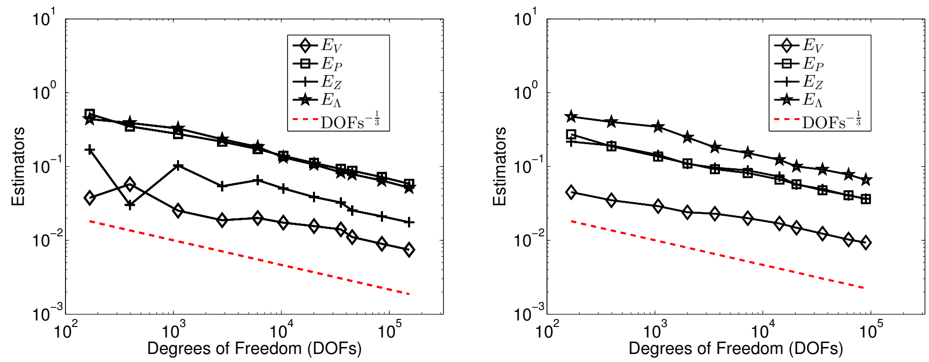


FIG. 6.2. *Computational rates of convergence for the contributions $\mathcal{E}_V$, $\mathcal{E}_P$, $\mathcal{E}_Z$, and $\mathcal{E}_\Lambda$ of the computable and anisotropic a posteriori error estimator $\mathcal{E}_{ocp}$. We have considered $n = 2$, an incompatible desired data u_d, and a non-convex domain $\Omega = (-1,1)^2 \setminus (0,1) \times (-1,0)$. The panels show the decrease of the contributions $\mathcal{E}_V$, $\mathcal{E}_P$, $\mathcal{E}_Z$, and $\mathcal{E}_\Lambda$ with respect to the number of degrees of freedom for $s = 0.3$ (left panel) and $s = 0.8$ (right panel). In both cases we recover the optimal rate $(\#\mathscr{T}_{\mathcal{Y}})^{-1/3}$ for each contribution.*

Finally, in Figure 6.3 we present the adaptive meshes obtained by our AFEM after 11 iterations and the corresponding finite element approximations of the optimal control variable $\bar{\mathsf{z}}$. Several conclusions can be drawn:

- *Geometric singularities and fractional diffusion behavior*: Since $\mathsf{u}_d = 1$, u_d is an incompatible forcing term for the adjoint equation $(-\Delta)^s \bar{\mathsf{p}} = \bar{\mathsf{u}} - \mathsf{u}_d$ when $s \in (0, \frac{1}{2}]$, As a consequence, it can be observed that, when $s = 0.2$ (Figure 6.3a) and $s = 0.3$ (Figure 6.3b), our AFEM localizes an important density of degrees of freedom near the boundary of the domain Ω. This is to compensate the singular behavior (1.5)–(1.6) that is inherent to fractional diffusion with incompatible data. When $s = 0.8$ (Figure 6.3c) and $s = 0.9$ (Figure 6.3d), the incompatibility of the desired data do not occur and then the optimal adjoint state do not exhibits boundary layers. Our AFEM is thus focused on resolving the reentrant corner.

- *Sparse optimal controls*: The fact that cost functional J involves the term $\|\mathsf{z}\|_{L^1(\Omega)}$ leads to sparsely supported optimal controls. Figures 6.3a–6.3d are particular instances of this feature. In addition, and in particular in Figure 6.3a, it can be observed that the error estimator $\mathcal{E}_{ocp}$ also focuses on refining the interface, i.e., the boundary of the active set.

REFERENCES

[1] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. 1964.

[2] A. Allendes, F. Fuica, and E. Otárola. Adaptive finite element methods for sparse PDE–constrained optimization. *arXiv:1712.00448*, 2017.

[3] A. Allendes, E. Otárola, R. Rankin, and A. J. Salgado. Adaptive finite element methods for an optimal control problem involving Dirac measures. *Numer. Math.*, 137(1):159–197, 2017.

[4] A. Antil and E. Otárola. An a *posteriori* error analysis for an optimal control problem involving the fractional laplacian. *IMA J. Numer. Anal.*, 38(1):198–226, 2018.

[5] I. Babuška and A. Miller. A feedback finite element method with a posteriori error estimation. I. The finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Engrg.*, 61(1):1–40, 1987.

[6] L. Banjai, J.M. Melenk, R.H. Nochetto, E. Otárola, A.J. Salgado, and Ch. Schwab. Tensor FEM for spectral fractional diffusion. *arXiv:1707.07367*, 2017.

[7] M. Š. Birman and M. Z. Solomjak. *Spektralnaya teoriya samosopryazhennykh operatorov v gilbertovom prostranstve*. Leningrad. Univ., Leningrad, 1980.

[8] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[9] 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(97), 2014.

[10] X. Cabré and J. Tan. Positive solutions of nonlinear problems involving the square root of the Laplacian. *Adv. Math.*, 224(5):2052–2093, 2010.

[11] L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Comm. Part. Diff. Eqs.*, 32(7-9):1245–1260, 2007.

[12] L. Caffarelli and A. Vasseur. Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math.*, 171:1903–1930, 2010.

[13] L. A. Caffarelli and P. R. Stinga. Fractional elliptic equations, Caccioppoli estimates and regularity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 33(3):767–807, 2016.

[14] A. Capella, J. Dávila, L. Dupaigne, and Y. Sire. Regularity of radial extremal solutions for some non-local semilinear equations. *Comm. Part. Diff. Eqs.*, 36(8):1353–1384, 2011.

[15] C. Carstensen and S. A. Funken. Fully reliable localized error control in the FEM. *SIAM J. Sci. Comput.*, 21(4):1465–1484 (electronic), 1999.

[16] E. Casas, R. Herzog, and G. Wachsmuth. Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. *SIAM J. Optim.*, 22(3):795–820, 2012.

[17] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[18] L. Chen. *i*FEM: An integrated finite element methods package in matlab. Technical report, University of California at Irvine, 2009.

[19] L. Chen, R. H. Nochetto, E. Otárola, and A. J. Salgado. A PDE approach to fractional diffusion: a posteriori error analysis. *J. Comput. Phys.*, 293:339–358, 2015.

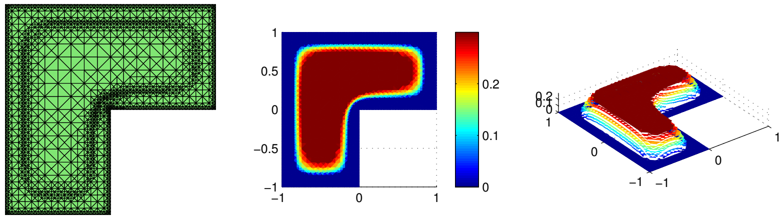
[20] W. Chen. A speculative study of 2/3-order fractional laplacian modeling of turbulence: Some thoughts and conjectures. *Chaos*, 16(2):1–11, 2006.

[21] W. Chen and S. Holm. Fractional laplacian time-space models for linear and nonlinear lossy media exhibiting arbitrary frequency power-law dependency. *The Journal of the Acoustical Society of America*, 115(4):1424–1430, 2004.

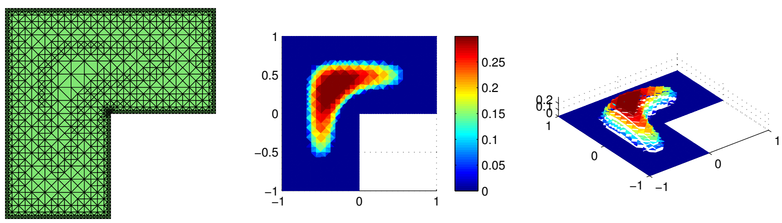
[22] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, 2002.

[23] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

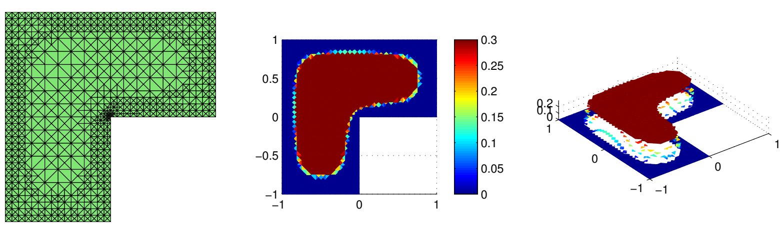
[24] M. Costabel and M. Dauge. General edge asymptotics of solutions of second-order elliptic



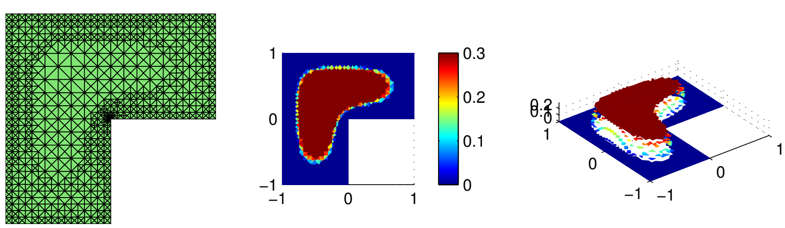
(a) $s = 0.2$ ($\sigma = 0.1$ and $\nu = 0.5$)



(b) $s = 0.3$ ($\sigma = 0.1$ and $\nu = 0.5$)



(c) $s = 0.8$ ($\sigma = 0.05$ and $\nu = 0.1$)



(d) $s = 0.9$ ($\sigma = 0.05$ and $\nu = 0.1$)

Fig. 6.3. *Adaptive meshes obtained by our anisotropic AFEM, after 11 iterations, with incompatible desired data* u_d *and a non-convex domain* $\Omega = (-1,1)^2 \setminus (0,1) \times (-1,0)$*, and the corresponding finite element approximations of the optimal control variable* $\bar{\mathsf{z}}$.

boundary value problems. I, II. *Proc. Roy. Soc. Edinburgh Sect. A*, 123(1):109–155, 157–184, 1993.
[25] Q. Du, M. Gunzburger, R.B. Lehoucq, and K. Zhou. Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.*, 54(4):667–696, 2012.

[26] J. Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[27] R. G. Durán and A. L. Lombardi. Error estimates on anisotropic Q_1 elements for functions in weighted Sobolev spaces. *Math. Comp.*, 74(252):1679–1706 (electronic), 2005.

[28] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.

[29] E. B. Fabes, C.E. Kenig, and R.P. Serapioni. The local regularity of solutions of degenerate elliptic equations. *Comm. Part. Diff. Eqs.*, 7(1):77–116, 1982.

[30] P. Gatto and J.S. Hesthaven. Numerical approximation of the fractional laplacian via hp-finite elements, with an application to image denoising. *J. Sci. Comp.*, 65(1):249–270, 2015.

[31] V. Gol′dshtein and A. Ukhlov. Weighted Sobolev spaces and embedding theorems. *Trans. Amer. Math. Soc.*, 361(7):3829–3850, 2009.

[32] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[33] M. Hintermüller, R. H. W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM Control Optim. Calc. Var.*, 14(3):540–560, 2008.

[34] 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(3), 2008.

[35] K. Kohls, A. Rösch, and K. G. Siebert. A posteriori error analysis of optimal control problems with control constraints. *SIAM J. Control Optim.*, 52(3):1832–1861, 2014.

[36] N. S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, New York, 1972. Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180.

[37] S. Z. Levendorskiĭ. Pricing of the American put under Lévy processes. *Int. J. Theor. Appl. Finance*, 7(3):303–335, 2004.

[38] W. Liu and N. Yan. A posteriori error estimates for distributed convex optimal control problems. *Adv. Comput. Math.*, 15(1-4):285–309 (2002), 2001. A posteriori error estimation and adaptive computational methods.

[39] P. Morin, R. H. Nochetto, and K. G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488 (electronic), 2000.

[40] P. Morin, R.H. Nochetto, and K.G. Siebert. Local problems on stars: a posteriori error estimators, convergence, and performance. *Math. Comp.*, 72(243):1067–1097 (electronic), 2003.

[41] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[42] R. Musina and A. I. Nazarov. On fractional Laplacians. *Comm. Partial Differential Equations*, 39(9):1780–1790, 2014.

[43] R. H. Nochetto, E. Otárola, and A. J. Salgado. A PDE approach to fractional diffusion in general domains: a priori error analysis. *Found. Comput. Math.*, 15(3):733–791, 2015.

[44] R. H. Nochetto, E. Otárola, and A. J. Salgado. Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications. *Numer. Math.*, 132(1):85–130, 2016.

[45] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, Berlin, 2009.

[46] R. H. Nochetto and A. Veeser. Primer of adaptive finite element methods. In *Multiscale and adaptivity: modeling, numerics and applications*, volume 2040 of *Lecture Notes in Math.*, pages 125–225. Springer, Heidelberg, 2012.

[47] E. Otárola. *A PDE approach to numerical fractional diffusion*. ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–University of Maryland, College Park.

[48] E. Otárola and Salgado A. J. Sparse optimal control for fractional diffusion. *Comput. Methods Appl. Math*, 2017. DOI: https://doi.org/10.1515/cmam-2017-0030.

[49] H. Pham. Optimal stopping, free boundary, and American option in a jump-diffusion model. *Appl. Math. Optim.*, 35(2):145–164, 1997.

[50] W. Schirotzek. *Nonsmooth analysis*. Universitext. Springer, Berlin, 2007.

[51] R. Schneider and G. Wachsmuth. A Posteriori Error Estimation for Control-Constrained, Linear-Quadratic Optimal Control Problems. *SIAM J. Numer. Anal.*, 54(2):1169–1192, 2016.

[52] S.A. Silling. Why peridynamics? In *Handbook of peridynamic modeling*, Adv. Appl. Math. CRC Press, 2017.

[53] L. Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator.

Comm. Pure Appl. Math., 60(1):67–112, 2007.
[54] G. Stadler. Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices. *Comput. Optim. Appl.*, 44(2):159–181, 2009.
[55] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
[56] P. R. Stinga and J. L. Torrea. Extension problem and Harnack's inequality for some fractional operators. *Comm. Part. Diff. Eqs.*, 35(11):2092–2122, 2010.
[57] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.
[58] B. O. Turesson. *Nonlinear potential theory and weighted Sobolev spaces*, volume 1736 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.
[59] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley, 1996.
[60] G. Wachsmuth and D. Wachsmuth. Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM Control Optim. Calc. Var.*, 17(3):858–886, 2011.