SIAM J. SCI. COMPUT.
Vol. 41, No. 5, pp. A2967–A2998

AN ADAPTIVE FEM FOR THE POINTWISE TRACKING OPTIMAL CONTROL PROBLEM OF THE STOKES EQUATIONS*

ALEJANDRO ALLENDES†, FRANCISCO FUICA†, ENRIQUE OTÁROLA†, AND DANIEL QUERO†

Abstract. We propose and analyze a reliable and efficient a posteriori error estimator for the pointwise tracking optimal control problem of the Stokes equations. This linear-quadratic optimal control problem entails the minimization of a cost functional that involves point evaluations of the velocity field that solves the state equations. This leads to an adjoint problem with a linear combination of Dirac measures as a forcing term and whose solution exhibits reduced regularity properties. We also consider constraints on the control variable. The proposed a posteriori error estimator can be decomposed as the sum of four contributions: three contributions related to the discretization of the state and adjoint equations and another contribution that accounts for the discretization of the control variable. On the basis of the devised a posteriori error estimator, we design a simple adaptive strategy that illustrates our theory and exhibits a competitive performance.

Key words. linear-quadratic optimal control problem, Stokes equations, a posteriori error estimates, Dirac measures, Muckenhoupt weights, weighted estimates, maximum-norm estimates

AMS subject classifications. 49K20, 49M25, 65K10, 65N15, 65N30, 65N50, 65Y20

DOI. 10.1137/18M1222363

1. Introduction.

In this work we shall be interested in the design and analysis of an a posteriori error estimator for the pointwise tracking optimal control problem of the Stokes equations; control constraints are also considered. To make matters precise, for $d \in \{2,3\}$, we let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytopal domain with Lipschitz boundary $\partial\Omega$ and let $\mathcal{D}$ be a finite ordered subset of Ω with cardinality $\#\mathcal{D} = m$. We will refer to $\mathcal{D}$ as the set of observable points. Given a set of desired states $\{\mathbf{y}_t\}_{t\in\mathcal{D}}$, a regularization parameter $\lambda > 0$, and the cost functional

$$J(\mathbf{y},\mathbf{u}) := \frac{1}{2}\sum_{t\in\mathcal{D}} |\mathbf{y}(t) - \mathbf{y}_t|^2 + \frac{\lambda}{2}\|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)}, \tag{1}$$

our problem reads as follows: Find $\min J(\mathbf{y},\mathbf{u})$ subject to the Stokes equations

$$-\Delta\mathbf{y} + \nabla p = \mathbf{u} \text{ in } \Omega, \quad \operatorname{div}\mathbf{y} = 0 \text{ in } \Omega, \quad \mathbf{y} = \mathbf{0} \text{ on } \partial\Omega \tag{2}$$

and the control constraints

$$\mathbf{u} \in \mathbb{U}_{ad}, \quad \mathbb{U}_{ad} := \{\mathbf{v} \in \mathbf{L}^2(\Omega) : \mathbf{a} \le \mathbf{v} \le \mathbf{b} \text{ a.e. in } \Omega\} \tag{3}$$

with $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$ satisfying $\mathbf{a} < \mathbf{b}$. We immediately comment that, throughout this work, vector inequalities must be understood componentwise. In (1), $|\cdot|$ denotes the euclidean norm.

*Submitted to the journal's Methods and Algorithms for Scientific Computing section October 23, 2018; accepted for publication (in revised form) June 25, 2019; published electronically October 1, 2019.

https://doi.org/10.1137/18M1222363

Funding: The work of the first author was supported by CONICYT through FONDECYT project 1170579. The work of the third author was partially supported by CONICYT through FONDECYT project 11180193. The work of the fourth author was supported by USM through Programa de Incentivos a la Iniciacion Cientifica (PIIC).

†Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile (alejandro.allendes@usm.cl, http://aallendes.mat.utfsm.cl/, francisco.fuica@sansano.usm.cl, enrique.otarola@usm.cl, http://eotarola.mat.utfsm.cl/, daniel.quero@alumnos.usm.cl).

The pointwise tracking optimal control problem for the Poisson equation has been considered in a number of works [8, 9, 11, 13]. In [8], the authors operate under the framework of Muckenhoupt weighted Sobolev spaces [26] and circumvent the difficulties associated with the underlying adjoint equation: a Poisson equation with a linear combination of Dirac deltas as a forcing term. Weighted Sobolev spaces allow for working under a Hilbert space-based framework in comparison to the non-Hilbertian setting of [9, 11, 13]. An a priori error analysis for a standard finite element approximation of the aforementioned problem can be found in [8, 11] while its a posteriori error analysis has been recently provided in [4, 13]. In contrast to these advances and to the best of our knowledge, the pointwise tracking optimal control problem for the Stokes equations has not been considered before. We immediately notice that, since the cost functional involves point evaluations of the velocity field that solves the state equations, the momentum equation of the adjoint equations reads

$$-\Delta \mathbf{z} - \nabla r = \sum_{t \in \mathcal{D}} (\mathbf{y} - \mathbf{y}_t)\delta_t. \tag{4}$$

Consequently, $\mathbf{z} \notin \mathbf{H}^1(\Omega)$ and $r \notin L^2(\Omega)/\mathbb{R}$ [18, section IV.2]. This complicates the analysis as well as the numerical approximation since, as observed in [5], standard a posteriori error estimation techniques [32] fail when solving (4).

Adaptive finite element methods (AFEMs) are iterative feedback procedures that improve the quality of the finite element approximation to a partial differential equation (PDE) while striving to keep an optimal distribution of computational resources measured in terms of degrees of freedom. An essential ingredient of AFEMs is an a posteriori error estimator which is of importance in computational practice because of its ability to provide computable information about errors and drive adaptive mesh refinement algorithms. The a posteriori error analysis for standard finite element approximations of linear second-order elliptic boundary value problems has a solid foundation [3, 27, 28, 32]. In contrast to this well-established theory, the a posteriori error analysis for finite element approximations of control constrained optimal control problems has not yet been fully understood. In view of their inherent nonlinear feature, which appears due to the control constraints, the analysis involves more arguments and technicalities [21, 25]. An attempt to present a unifying framework has been carried out recently in [22]. The authors derive an important error equivalence that simplifies the analysis to, simply put, provide estimators for the state and adjoint equations which satisfy a set of suitable assumptions [22, Theorem 3.2]. Unfortunately, this analysis relies fundamentally on a particular structure for the problem and the relations among the spaces for the state, adjoint state, and control. The analysis is based on the energy norm inherited by the state and adjoint equations. Many problems do not fit into this framework and thus one must either extend the theory or devise new estimators. The problem we consider in this work is an instance of this issue.

In this work we propose an a posteriori error estimator, for the pointwise tracking optimal control problem of the Stokes equations, that can be decomposed as the sum of four contributions: two related to the discretization of the state equations, one associated to the discretization of the adjoint equations, and another one that accounts for the discretization of the control variable. Since problem (4) involves point evaluations of the velocity field that solves the state equations, we consider, for such a variable, an a posteriori error estimator in maximum-norm [24] while a standard one is considered for the associated pressure [32]. For the adjoint variables we consider the a posteriori error estimator in Muckenhoupt weighted Sobolev spaces

of [5]. We obtain global reliability and local efficiency properties. On the basis of the devised a posteriori error estimator, we also design a simple adaptive strategy that exhibits optimal experimental rates of convergence for the state and adjoint variables. We remark that the analysis involves estimates in weighted Sobolev spaces and $\mathbf{L}^\infty$-norms, combined with having to deal with the variational inequality that characterizes the optimal control. This subtle intertwining of ideas is one of the highlights of this contribution.

The plan of the paper is as follows. In section 2, we introduce the notation and functional framework we shall work with. In section 3.1 we review basic results for the Stokes equations and provide a weighted integrability result on Lipschitz polytopes. In section 3.2, we recall one of the maximum-norm a posteriori error estimators for the Stokes equations developed in [24] and provide an efficiency analysis for it. Section 4 contains the description and analysis of the pointwise tracking optimal control problem for the Stokes equations. The core of our work is section 5, where we introduce a discrete scheme, devise an a posteriori error estimator, and show, in sections 5.3 and 5.4, its global reliability and local efficiency, respectively. We conclude, in section 6, with a series of numerical examples that illustrate and go beyond our theory.

2. Notation and preliminaries. Let us fix notation and the setting in which we will operate.

2.1. Notation. Throughout this work $d \in \{2,3\}$ and $\Omega \subset \mathbb{R}^d$ is an open and bounded polytopal domain with Lipschitz boundary $\partial\Omega$. If $\mathcal{X}$ and $\mathcal{Y}$ are normed vector spaces, we write $\mathcal{X} \hookrightarrow \mathcal{Y}$ to denote that $\mathcal{X}$ is continuously embedded in $\mathcal{Y}$. We denote by $\mathcal{X}'$ and $\|\cdot\|_{\mathcal{X}}$ the dual and the norm of $\mathcal{X}$, respectively.

We shall use lowercase bold letters to denote vector-valued functions whereas uppercase bold letters are used to denote function spaces. For a bounded domain $G \subset \mathbb{R}^d$, if $X(G)$ corresponds to a function space over G, we shall denote $\mathbf{X}(G) = [X(G)]^d$. In particular, we denote $\mathbf{L}^2(G) = [L^2(G)]^d$, which is equipped with the following inner product and norm:

$$(\mathbf{w},\mathbf{v})_{\mathbf{L}^2(G)} = \int_G \mathbf{w}\cdot\mathbf{v}, \qquad \|\mathbf{v}\|_{\mathbf{L}^2(G)} = (\mathbf{v},\mathbf{v})^{\frac{1}{2}}_{\mathbf{L}^2(G)} \qquad \forall\, \mathbf{w},\mathbf{v} \in \mathbf{L}^2(G).$$

The relation $a \lesssim b$ indicates that $a \le Cb$ with a positive constant that depends neither on a, b nor the discretization parameter. The value of C might change at each occurrence.

2.2. Weighted Sobolev spaces. A weight is an almost everywhere positive function $\omega \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. For a Borel set $G \subset \mathbb{R}^d$ and a weight ω, we define $\omega(G) = \int_G \omega$. A particular class of weights that will be of importance is the so-called Muckenhoupt class $A_2(\mathbb{R}^d)$ [15, 31]. The following example will be essential in our analysis: Let x_0 be an interior point of Ω and denote by $\mathrm{d}_{x_0}(x) = |x - x_0|$ the Euclidean distance to x_0; $\mathsf{d}^\alpha_{x_0}(x) = \mathsf{d}_{x_0}(x)^\alpha$ belongs to the Muckenhoupt class $A_2(\mathbb{R}^d)$ if and only if $\alpha \in (-d, d)$.

Let $\omega \in A_2(\mathbb{R}^d)$ and $G \subset \mathbb{R}^d$ be an open and bounded domain. We define $L^2(\omega, G)$ as the Lebesgue space of square-integrable functions with respect to the measure $\omega \mathrm{d}x$. We also define the weighted Sobolev space $H^1(\omega, G) := \{v \in L^2(\omega, G) : |\nabla v| \in L^2(\omega, G)\}$, which we equip with the norm

$$\|v\|_{H^1(\omega,G)} := \left(\|v\|^2_{L^2(\omega,G)} + \|\nabla v\|^2_{L^2(\omega,G)}\right)^{\frac{1}{2}}. \tag{5}$$

Since $\omega \in A_2(\mathbb{R}^d)$, the results of [31, Proposition 2.1.2 and Corollary 2.1.6] and [20, Theorem 1] guarantee that $L^2(\omega, \Omega)$ and $H^1(\omega, \Omega)$ are Hilbert spaces. Moreover, the

space $C^\infty(\Omega)$ is dense in $H^1(\omega,\Omega)$. We define $H^1_0(\omega,\Omega)$ as the closure of $C^\infty_0(\Omega)$ in $H^1(\omega,\Omega)$. Additionally, a Poincaré inequality holds for all $v \in H^1_0(\omega,G)$ [17, Theorem 1.3]. Consequently, in $H^1_0(\omega,G)$ the seminorm $\|\nabla v\|_{L^2(\omega,\Omega)}$ is equivalent to (5).

3. Pointwise a posteriori error estimation for the Stokes equations. In this section we review $\mathbf{L}^\infty$ a posteriori error estimates as well as standard results concerning regularity properties of the solution to the Stokes equations. In addition, we provide an efficiency analysis for one of the estimators proposed in [24].

3.1. The Stokes problem in Lipschitz polytopes. We begin by introducing the bilinear forms $a : \mathbf{H}^1_0(\Omega)\times\mathbf{H}^1_0(\Omega)\to\mathbb{R}$ and $b : \mathbf{H}^1_0(\Omega)\times L^2(\Omega)/\mathbb{R}\to\mathbb{R}$ by

$$a(\mathbf{w},\mathbf{v}) := \int_\Omega \nabla\mathbf{w} : \nabla\mathbf{v} = \sum_{i=1}^d \int_\Omega \nabla\mathbf{w}_i\cdot\nabla\mathbf{v}_i, \qquad b(\mathbf{v},q) := -\int_\Omega q\,\operatorname{div}\mathbf{v} \tag{6}$$

for all $\mathbf{v},\mathbf{w}\in\mathbf{H}^1_0(\Omega)$ and $q\in L^2(\Omega)/\mathbb{R}$, respectively.

Given $\mathbf{f}\in\mathbf{H}^{-1}(\Omega)$, we consider the following weak version of the Stokes equations: Find $(\mathbf{y},p)\in\mathbf{H}^1_0(\Omega)\times L^2(\Omega)/\mathbb{R}$ such that

$$a(\mathbf{y},\mathbf{v})+b(\mathbf{v},p) = \langle\mathbf{f},\mathbf{v}\rangle_{\mathbf{H}^{-1}(\Omega),\mathbf{H}^1_0(\Omega)} \quad \forall\mathbf{v}\in\mathbf{H}^1_0(\Omega), \quad b(\mathbf{y},q)=0 \quad \forall q\in L^2(\Omega)/\mathbb{R}. \tag{7}$$

We now provide a local and weighted integrability result for $\mathbf{y}$ and p.

PROPOSITION 1 (weighted integrability). *Let $(\mathbf{y},p)\in\mathbf{H}^1_0(\Omega)\times L^2(\Omega)/\mathbb{R}$ be the solution of (7) with $\mathbf{f}\in\mathbf{L}^l(\Omega)$ and $l>d$. Let $x\in\Omega$, $\zeta<\operatorname{dist}(x,\partial\Omega)$ and B denote the ball of radius ζ and center x. If $\omega\in A_2(\mathbb{R}^d)$, then we have that $(\mathbf{y},p)\in\mathbf{H}^1_0(\omega,B)\times L^2(\omega,B)/\mathbb{R}$. In addition, we have the estimate*

$$\|\nabla\mathbf{y}\|_{\mathbf{L}^2(\omega,B)} + \|p\|_{L^2(\omega,B)} \lesssim \|\mathbf{f}\|_{\mathbf{L}^l(\Omega)}, \tag{8}$$

where the hidden constant depends on $\omega(B)$, $\operatorname{dist}(B,\partial\Omega)$, ζ, and Ω, but is independent of $\mathbf{y}$, p, and $\mathbf{f}$.

Proof. Since $l>d$, the following embedding result holds: $\mathbf{W}^{1,l}(B)\hookrightarrow\mathbf{L}^\infty(B)$. This, combined with the fact that $\operatorname{dist}(B,\partial\Omega)>0$, and the estimate of [18, Theorem IV.4.1] reveal that $\|\nabla\mathbf{y}\|_{\mathbf{L}^\infty(B)}+\|p\|_{L^\infty(B)}\lesssim\|\mathbf{y}\|_{\mathbf{W}^{2,l}(B)}+\|p\|_{W^{1,l}(B)}\lesssim\|\mathbf{f}\|_{\mathbf{L}^l(\Omega)}$. This immediately implies the desired estimate. ∎

3.2. Pointwise a posteriori error estimates. In this section we briefly present one of the pointwise a posteriori error estimators introduced and analyzed in [24]. To accomplish this task, we assume that $\mathbf{f}\in\mathbf{L}^\infty(\Omega)$.

Let us start the discussion by introducing some standard finite element notation. Let $\mathscr{T}=\{T\}$ be a conforming partition of $\bar\Omega$ into simplices T with size $h_T:=\operatorname{diam}(T)$. We denote by $\mathbb{T}$ the collection of conforming and shape regular meshes that are refinements of an initial mesh $\mathscr{T}_0$. We define $\mathscr{S}$ as the set of internal $(d-1)$-dimensional interelement boundaries S of $\mathscr{T}$. For $T\in\mathscr{T}$, let $\mathscr{S}_T$ denote the subset of $\mathscr{S}$ that contains the sides in $\mathscr{S}$ which are sides of T. We also denote by $\mathcal{N}_S$ the subset of $\mathscr{T}$ that contains the two elements that have S as a side. In addition, we define the *stars* or *patches* associated with an element $T\in\mathscr{T}$ as

$$\mathcal{N}_T = \cup\{T'\in\mathscr{T} : \mathscr{S}_T\cap\mathscr{S}_{T'}\neq\emptyset\}, \quad \mathcal{N}^*_T = \cup\{T'\in\mathscr{T} : T\cap T'\neq\emptyset\}. \tag{9}$$

For a discrete tensor valued function $\mathbf{V}_{\mathscr{T}}$, we denote by $[\![\mathbf{V}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]$ the jump or interelement residual, which is defined, on the internal side $S \in \mathscr{S}$ shared by the distinct elements $T^+, T^- \in \mathcal{N}_S$, by

$$[\![\mathbf{V}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!] = \mathbf{V}_{\mathscr{T}}|_{T^+} \cdot \boldsymbol{\nu}^+ + \mathbf{V}_{\mathscr{T}}|_{T^-} \cdot \boldsymbol{\nu}^-. \tag{10}$$

Here, $\boldsymbol{\nu}^+, \boldsymbol{\nu}^-$ are unit normals on S pointing toward T^+, T^-, respectively.

Given a mesh $\mathscr{T} \in \mathbb{T}$, we denote by $\mathbf{V}(\mathscr{T})$ and $Q(\mathscr{T})$ the finite element spaces that approximate the velocity field and the pressure, respectively, based on the classical Taylor–Hood elements [16, section 4.2.5]:

$$\begin{aligned} Q(\mathscr{T}) &:= \left\{ q_{\mathscr{T}} \in C(\bar{\Omega}) \ : \ q_{\mathscr{T}}|_T \in \mathbb{P}_1(T) \, \forall \, T \in \mathscr{T} \right\} \cap L^2(\Omega)/\mathbb{R}, \\ \mathbf{V}(\mathscr{T}) &:= \left\{ \mathbf{v}_{\mathscr{T}} \in \mathbf{C}(\bar{\Omega}) \ : \ \mathbf{v}_{\mathscr{T}}|_T \in \mathbb{P}_2(T)^d \, \forall \, T \in \mathscr{T} \right\} \cap \mathbf{H}^1_0(\Omega). \end{aligned} \tag{11}$$

With these spaces at hand, we define the Galerkin approximation to (7) as the solution to the following problem: Find $(\mathbf{y}_{\mathscr{T}}, p_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T})$ that solves

$$a(\mathbf{y}_{\mathscr{T}}, \mathbf{v}_{\mathscr{T}}) + b(\mathbf{v}_{\mathscr{T}}, p_{\mathscr{T}}) = (\mathbf{f}, \mathbf{v}_{\mathscr{T}})_{\mathbf{L}^2(\Omega)}, \qquad b(\mathbf{y}_{\mathscr{T}}, q_{\mathscr{T}}) = 0 \tag{12}$$

for all $\mathbf{v}_{\mathscr{T}} \in \mathbf{V}(\mathscr{T})$ and $q_{\mathscr{T}} \in Q(\mathscr{T})$.

We introduce the pointwise a posteriori error estimator $\mathcal{E}_\infty$ as follows:

$$\mathcal{E}_\infty(\mathbf{y}_{\mathscr{T}}, p_{\mathscr{T}}, \mathbf{f}) := \max_{T \in \mathscr{T}} \mathcal{E}_{\infty,T}(\mathbf{y}_{\mathscr{T}}, p_{\mathscr{T}}, \mathbf{f}), \tag{13}$$

where, for every $T \in \mathscr{T}$, the local a posteriori error indicators $\mathcal{E}_{\infty,T}$ are given by

$$\begin{aligned} \mathcal{E}_{\infty,T}(\mathbf{y}_{\mathscr{T}}, p_{\mathscr{T}}, \mathbf{f}) := {} & h_T^2 \|\mathbf{f} + \Delta \mathbf{y}_{\mathscr{T}} - \nabla p_{\mathscr{T}}\|_{\mathbf{L}^\infty(T)} \\ & + \tfrac{h_T}{2} \|[\![\nabla \mathbf{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(\partial T \setminus \partial\Omega)} + h_T \|\operatorname{div} \mathbf{y}_{\mathscr{T}}\|_{L^\infty(T)}. \end{aligned} \tag{14}$$

3.2.1. Reliability. In order to present the global reliability of the a posteriori error estimator $\mathcal{E}_\infty$, and for future reference, we introduce

$$\ell_{\mathscr{T}} := \left| \log \left(\max_{T \in \mathscr{T}} h_T^{-1} \right) \right|. \tag{15}$$

We present the following result; see [24, Theorem 4.1] and [14, Lemma 3].

THEOREM 2 (global reliability of $\mathcal{E}_\infty$). *Let $(\mathbf{y}, p) \in \mathbf{H}^1_0(\Omega) \times L^2(\Omega)/\mathbb{R}$ be the solution to the Stokes equations* (7) *with $\mathbf{f} \in \mathbf{L}^\infty(\Omega)$ and $(\mathbf{y}_{\mathscr{T}}, p_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T})$ its numerical approximation obtained as the solution to* (12). *Then,*

$$\|\mathbf{y} - \mathbf{y}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim \ell_{\mathscr{T}}^{\beta_d} \mathcal{E}_\infty, \tag{16}$$

where $\beta_2 = 2$ and $\beta_3 = 4/3$.

3.2.2. Efficiency. We investigate the local efficiency properties of $\mathcal{E}_{\infty,T}$. To accomplish this task, we define, for $\mathscr{M} \subset \mathscr{T}$, $\mathfrak{m} \in \{2, \infty\}$, and $\mathbf{g} \in \mathbf{L}^{\mathfrak{m}}(\Omega)$,

$$\operatorname{osc}_2^2(\mathbf{g}; \mathscr{M}) := \sum_{T \in \mathscr{M}} h_T^2 \|\mathbf{g} - \Pi_{\mathscr{T}}(\mathbf{g})\|^2_{\mathbf{L}^2(T)}, \quad \operatorname{osc}_\infty(\mathbf{g}; \mathscr{M}) := \max_{T \in \mathscr{M}} h_T^2 \|\mathbf{g} - \Pi_{\mathscr{T}}(\mathbf{g})\|_{\mathbf{L}^\infty(T)},$$

where $\Pi_{\mathscr{T}}$ denotes the L^2-orthogonal projection operator onto piecewise linear functions over $\mathscr{T}$.

For an edge/face or triangle/tetrahedron G, we denote by $\mathscr{V}(G)$ the set of vertices of G. We introduce, for $T \in \mathscr{T}$, the standard bubble function [3, section 2.3.1] $\varphi_T = (d+1)^{d+1} \prod_{\mathrm{v} \in \mathscr{V}(T)} \phi_{\mathrm{v}}|_T$, where ϕ_{v} are the barycentric coordinates of T. The function φ_T satisfies the following properties:

$$|T| \lesssim \int_T \varphi_T, \quad \text{supp } \varphi_T = T, \quad \|\nabla^k \varphi_T\|_{L^2(T)} \lesssim h_T^{\frac{d}{2}-k}, \quad k = 0, 1, 2. \tag{17}$$

The local efficiency of the indicator (14) is as follows.

THEOREM 3 (local efficiency of $\mathcal{E}_{\infty,T}$). *Let $(\mathbf{y}, p) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R}$ be the solution to the Stokes equations (7) with $\mathbf{f} \in \mathbf{L}^\infty(\Omega)$ and $(\mathbf{y}_\mathscr{T}, p_\mathscr{T}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T})$ its numerical approximation obtained as the solution to (12). If Ω is convex, then, for $T \in \mathscr{T}$, the local error indicators $\mathcal{E}_{\infty,T}$ defined as in (14) satisfy that*

$$\mathcal{E}_{\infty,T}(\mathbf{y}_\mathscr{T}, p_\mathscr{T}) \lesssim \|\mathbf{y} - \mathbf{y}_\mathscr{T}\|_{\mathbf{L}^\infty(\mathcal{N}_T)} + h_T \|p - p_\mathscr{T}\|_{L^\infty(\mathcal{N}_T)} + \mathrm{osc}_\infty(\mathbf{f}; \mathcal{N}_T), \tag{18}$$

where $\mathcal{N}_T$ is defined as in (9). The hidden constant is independent of the continuous and discrete solutions, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. We proceed in four steps.

Step 1. To simplify the presentation of the material, we define $\mathbf{e}_\mathbf{y} = \mathbf{y} - \mathbf{y}_\mathscr{T}$ and $e_p = p - p_\mathscr{T}$. Let us consider $\mathbf{v} \in \mathbf{H}_0^1(\Omega)$ which is such that $\mathbf{v}|_T \in \mathbf{C}^2(T)$ for all $T \in \mathscr{T}$. We invoke the fact that $(\mathbf{y}, p)$ solves (7) to arrive at

$$a(\mathbf{e}_\mathbf{y}, \mathbf{v}) + b(\mathbf{v}, e_p) = (\mathbf{f}, \mathbf{v})_{\mathbf{L}^2(\Omega)} - (\nabla \mathbf{y}_\mathscr{T}, \nabla \mathbf{v})_{\mathbf{L}^2(\Omega)} + (p_\mathscr{T}, \operatorname{div} \mathbf{v})_{L^2(\Omega)}. \tag{19}$$

Consequently, (19) and an integration by parts formula allow us to conclude the following identity:

$$\begin{aligned} &- \sum_{T \in \mathscr{T}} \left((\Delta \mathbf{v}, \mathbf{e}_\mathbf{y})_{\mathbf{L}^2(T)} + (e_p, \operatorname{div} \mathbf{v})_{L^2(T)} \right) - \sum_{S \in \mathscr{S}} (\mathbf{e}_\mathbf{y}, [\![\nabla \mathbf{v} \cdot \boldsymbol{\nu}]\!])_{\mathbf{L}^2(S)} \\ &\quad = \sum_{T \in \mathscr{T}} \left((\Pi_\mathscr{T}(\mathbf{f}) + \Delta \mathbf{y}_\mathscr{T} - \nabla p_\mathscr{T}, \mathbf{v})_{\mathbf{L}^2(T)} + (\mathbf{f} - \Pi_\mathscr{T}(\mathbf{f}), \mathbf{v})_{\mathbf{L}^2(T)} \right) \\ &\qquad + \sum_{S \in \mathscr{S}} ([\![\nabla \mathbf{y}_\mathscr{T} \cdot \boldsymbol{\nu}]\!], \mathbf{v})_{\mathbf{L}^2(S)}. \end{aligned} \tag{20}$$

Notice that we have used that $p_\mathscr{T} \in Q(\mathscr{T})$, which implies that $[\![p_\mathscr{T}]\!] = 0$.

Step 2. Let $T \in \mathscr{T}$. We estimate the term $h_T^2 \|\mathbf{f} + \Delta \mathbf{y}_\mathscr{T} - \nabla p_\mathscr{T}\|_{\mathbf{L}^\infty(T)}$ in (14). To accomplish this task, we first invoke the triangle inequality and obtain that

$$h_T^2 \|\mathbf{f} + \Delta \mathbf{y}_\mathscr{T} - \nabla p_\mathscr{T}\|_{\mathbf{L}^\infty(T)} \leq h_T^2 \|\Pi_\mathscr{T}(\mathbf{f}) + \Delta \mathbf{y}_\mathscr{T} - \nabla p_\mathscr{T}\|_{\mathbf{L}^\infty(T)} + h_T^2 \|\mathbf{f} - \Pi_\mathscr{T}(\mathbf{f})\|_{\mathbf{L}^\infty(T)}.$$

To simplify the presentation of the material, we define $\mathbf{R}_T := (\Pi_\mathscr{T}(\mathbf{f}) + \Delta \mathbf{y}_\mathscr{T} - \nabla p_\mathscr{T})|_T$. It thus suffices to bound $h_T^2 \|\mathbf{R}_T\|_{\mathbf{L}^\infty(T)}$. To derive such a bound, we set $\mathbf{v} = \varphi_T^2 \mathbf{R}_T$ in (20) and invoke properties of the function φ_T. This yields

$$\begin{aligned} \|\mathbf{R}_T\|_{\mathbf{L}^2(T)}^2 &\lesssim \|\mathbf{f} - \Pi_\mathscr{T}(\mathbf{f})\|_{\mathbf{L}^\infty(T)} \|\varphi_T^2 \mathbf{R}_T\|_{\mathbf{L}^1(T)} \\ &\quad + \|\mathbf{e}_\mathbf{y}\|_{\mathbf{L}^\infty(T)} \|\Delta(\varphi_T^2 \mathbf{R}_T)\|_{\mathbf{L}^1(T)} + \|e_p\|_{L^\infty(T)} \|\operatorname{div}(\varphi_T^2 \mathbf{R}_T)\|_{L^1(T)}, \end{aligned} \tag{21}$$

where we have used that, for $S \in \mathscr{S}_T$, $\int_S \mathbf{e}_\mathbf{y} [\![\nabla(\varphi_T^2 \mathbf{R}_T) \cdot \boldsymbol{\nu}]\!] = 0$.

On the other hand, standard computations reveal that $\Delta(\varphi_T^2\mathbf{R}_T) = 2(\varphi_T\Delta\varphi_T + |\nabla\varphi_T|^2)\mathbf{R}_T + 4\varphi_T\nabla\mathbf{R}_T\nabla\varphi_T + \varphi_T^2\Delta\mathbf{R}_T$. This, in conjunction with the properties that φ_T satisfies, stated in (17), and the inverse estimates of [10, Lemma 4.5.3], imply that

$$\|\Delta(\varphi_T^2\mathbf{R}_T)\|_{\mathbf{L}^1(T)} \lesssim h_T^{\frac{d}{2}-2}\|\mathbf{R}_T\|_{\mathbf{L}^2(T)}. \tag{22}$$

Similar arguments to the ones that yield (22) allow us to derive

$$\|\mathrm{div}(\varphi_T^2\mathbf{R}_T)\|_{L^1(T)} \lesssim h_T^{\frac{d}{2}-1}\|\mathbf{R}_T\|_{\mathbf{L}^2(T)}, \quad \|\varphi_T^2\mathbf{R}_T\|_{\mathbf{L}^1(T)} \lesssim h_T^{\frac{d}{2}}\|\mathbf{R}_T\|_{\mathbf{L}^2(T)}. \tag{23}$$

We thus replace the estimates (22)–(23) into (21) to arrive at

$$h_T^2\|\mathbf{R}_T\|_{\mathbf{L}^2(T)} \lesssim h_T^{\frac{d}{2}+2}\|\mathbf{f} - \Pi_{\mathscr{T}}(\mathbf{f})\|_{\mathbf{L}^\infty(T)} + h_T^{\frac{d}{2}}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T)} + h_T^{\frac{d}{2}+1}\|e_p\|_{L^\infty(T)}.$$

The inverse estimate $\|\mathbf{R}_T\|_{\mathbf{L}^\infty(T)} \lesssim h_T^{-\frac{d}{2}}\|\mathbf{R}_T\|_{\mathbf{L}^2(T)}$ allows us to conclude.

Step 3. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We bound the term $\frac{h_T}{2}\|[\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(\partial T\setminus\partial\Omega)}$ in (14). We begin by invoking standard arguments to conclude the existence of an edge bubble function $\varphi_S \in \mathbb{P}_{(14d-19)}(\mathcal{N}_S)$ such that

$$\varphi_S = 0 \text{ on } \partial\mathcal{N}_S, \quad \nabla\varphi_S = \mathbf{0} \text{ on } \partial\mathcal{N}_S, \quad [\![\nabla\varphi_S\cdot\boldsymbol{\nu}]\!] = 0 \text{ on } S. \tag{24}$$

Define the vector-valued bubble function $\boldsymbol{\varphi}_S$ as $\boldsymbol{\varphi}_S := (2\varphi_S, \varphi_S)^T$, if $d = 2$, and $\boldsymbol{\varphi}_S := (9\varphi_S, \varphi_S, \varphi_S)^T$, if $d = 3$. Thus,

$$|S|\|[\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(S)} \lesssim \int_S [\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\boldsymbol{\varphi}_S. \tag{25}$$

We have assumed, without loss of generality, that $\|[\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(S)} = ([\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!])_1(\mathrm{v}) > 0$ with $\mathrm{v} \in \mathscr{V}(S)$.

Now, we set $\mathbf{v} = \boldsymbol{\varphi}_S$ in (20) and use (24) and the derived estimate for $\|\mathbf{R}_T\|_{\mathbf{L}^\infty(T)}$ to arrive at

$$\int_S [\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\boldsymbol{\varphi}_S \lesssim \sum_{T'\in\mathcal{N}_S}\left(h_T^{d-2}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T')} + h_T^{d-1}\|e_p\|_{L^\infty(T')} + h_T^d\|\mathbf{f}-\Pi_{\mathscr{T}}(\mathbf{f})\|_{\mathbf{L}^\infty(T')}\right).$$

We thus replace the previous estimate into (25) and use, in view of the mesh regularity assumptions, that $|T|/|S| \approx h_T$ to conclude that

$$h_T\|[\![\nabla\mathbf{y}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(S)} \lesssim \sum_{T'\in\mathcal{N}_S}\left(\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T')} + h_T\|e_p\|_{L^\infty(T')} + h_T^2\|\mathbf{f}-\Pi_{\mathscr{T}}(\mathbf{f})\|_{\mathbf{L}^\infty(T')}\right).$$

Step 4. Let $T \in \mathscr{T}$. We control the term $h_T\|\mathrm{div}\ \mathbf{y}_{\mathscr{T}}\|_{L^\infty(T)}$ in (14). To achieve this, we first use that $\mathrm{div}\ \mathbf{y} = 0$, and thus an integration by parts formula in conjunction with the properties (17) of φ_T to arrive at

$$\begin{aligned}\|\mathrm{div}\ \mathbf{y}_{\mathscr{T}}\|_{L^2(T)}^2 &\lesssim \int_T \mathrm{div}\ (\mathbf{y}_{\mathscr{T}} - \mathbf{y})\,(\varphi_T\ \mathrm{div}\ \mathbf{y}_{\mathscr{T}}) \lesssim \left|\int_T \mathbf{e_y}\cdot\nabla(\varphi_T\mathrm{div}\ \mathbf{y}_{\mathscr{T}})\right| \\ &\lesssim h_T^{\frac{d}{2}-1}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T)}\|\mathrm{div}\ \mathbf{y}_{\mathscr{T}}\|_{L^2(T)},\end{aligned} \tag{26}$$

where we also have used an inverse inequality. Consequently, using an inverse estimate, again, we conclude that $h_T\|\text{div } \mathbf{y}_{\mathscr{T}}\|_{L^\infty(T)} \lesssim h_T^{1-d/2}\|\text{div } \mathbf{y}_{\mathscr{T}}\|_{L^2(T)} \lesssim \|\mathbf{e}_{\mathbf{y}}\|_{\mathbf{L}^\infty(T)}$.

The collection of the estimates derived in Steps 2, 3, and 4 yields (18). This concludes the proof. ∎

Remark 4 (convexity of Ω). The convexity of the domain Ω is needed in Theorem 3 in order to guarantee that $p \in L^\infty(\Omega)$. This result follows from [14, Lemma 14]. In fact, it can be derived that the pressure $p \in C^\beta(\Omega)$ for some $\beta > 0$. The error $\|p - p_{\mathscr{T}}\|_{L^\infty(\mathcal{N}_T)}$ is thus well-defined. The convexity assumption on the domain, however, will not be necessary when deriving efficiency properties for the proposed a posteriori error estimator for the optimal control problem; see Theorem 14.

4. The pointwise tracking optimal control problem. In this section we analyze a weak version of the optimal control problem (1)–(3), which reads

$$\min\{J(\mathbf{y},\mathbf{u}) : (\mathbf{y},\mathbf{u}) \in \mathbf{H}_0^1(\Omega) \times \mathbb{U}_{ad}\} \tag{27}$$

subject to

$$a(\mathbf{y},\mathbf{v}) + b(\mathbf{v},p) = (\mathbf{u},\mathbf{v})_{\mathbf{L}^2(\Omega)} \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega), \quad b(\mathbf{y},q) = 0 \quad \forall q \in L^2(\Omega)/\mathbb{R}. \tag{28}$$

Since a is coercive on $\mathbf{H}_0^1(\Omega)$ and b satisfies an inf-sup condition, there is a unique solution $(\mathbf{y},p) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R}$ to problem (28) [16, Theorem 4.3]. In addition,

$$\|\nabla \mathbf{y}\|_{\mathbf{L}^2(\Omega)} + \|p\|_{L^2(\Omega)} \lesssim \|\mathbf{u}\|_{\mathbf{L}^2(\Omega)}. \tag{29}$$

Due to Rham's theorem [16, section 4.1.3] we can consider the following equivalent formulation of problem (28) [16, Proposition 4.6]: Find $\mathbf{y} \in \mathbf{X}$ such that

$$a(\mathbf{y},\mathbf{v}) = (\mathbf{u},\mathbf{v})_{\mathbf{L}^2(\Omega)} \quad \forall\, \mathbf{v} \in \mathbf{X}, \tag{30}$$

where $\mathbf{X} := \{\mathbf{v} \in \mathbf{H}_0^1(\Omega) :\ \text{div } \mathbf{v} = 0\}$.

To provide an analysis for (27)–(28), we introduce the control-to-state operator $\mathcal{S} : \mathbf{L}^2(\Omega) \to \mathbf{X}$ which, given a control $\mathbf{u}$, associates to it the unique state $\mathbf{y} \in \mathbf{X}$ that solves (30). With this operator at hand, we introduce the reduced cost functional

$$j(\mathbf{u}) := J(\mathcal{S}\mathbf{u},\mathbf{u}) = \frac{1}{2}\sum_{t\in\mathcal{D}} |\mathcal{S}\mathbf{u}(t) - \mathbf{y}_t|^2 + \frac{\lambda}{2}\|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)}. \tag{31}$$

We comment that, since the control variable $\mathbf{u} \in \mathbb{U}_{ad} \subset \mathbf{L}^\infty(\Omega)$ and $\partial\Omega$ is Lipschitz, the results of [12, Theorem 2.9] (see also [24, Theorem 1.1] and [14, Lemma 12]) guarantee the Hölder regularity of $\mathbf{y} = \mathcal{S}\mathbf{u}$; point evaluations of $\mathbf{y} = \mathcal{S}\mathbf{u}$ in (31) are thus well-defined.

We present the following existence and uniqueness result.

THEOREM 5 (existence and uniqueness). *The optimal control problem* (27)–(28) *admits a unique solution* $(\bar{\mathbf{y}},\bar{\mathbf{u}}) \in \mathbf{H}_0^1(\Omega) \times \mathbb{U}_{ad}$.

Proof. We begin by noticing that the reduced cost functional j is strictly convex and continuous. In addition, $\mathbb{U}_{ad}$ is a nonempty, bounded, convex, and closed subset of $\mathbf{L}^2(\Omega)$. We thus apply [30, Theorem 2.14] to conclude the desired result. ∎

The following result is standard [30, Lemma 2.21]: If $\bar{\mathbf{u}}$ denotes the optimal control of (27)–(28), then

$$j'(\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) \geq 0 \quad \forall\, \mathbf{u} \in \mathbb{U}_{ad}. \tag{32}$$

Here $j'(\bar{\mathbf{u}})$ denotes the Gateâux-derivative of the functional j in $\bar{\mathbf{u}}$. To explore this variational inequality and obtain optimality conditions we first shall state and derive some results on weighted Sobolev spaces.

Let us consider the ordered set of points $\mathcal{D} \subset \Omega$ with finite cardinality $m := \#\mathcal{D} < \infty$. We define

$$d_{\mathcal{D}} = \begin{cases} \operatorname{dist}(\mathcal{D}, \partial\Omega) & \text{if } m = 1, \\ \min\{\operatorname{dist}(\mathcal{D}, \partial\Omega), \min\{|t - t'| : t, t' \in \mathcal{D},\ t \neq t'\}\} & \text{otherwise.} \end{cases}$$

Since $\mathcal{D} \subset \Omega$ and $\mathcal{D}$ is finite, we immediately conclude that $d_{\mathcal{D}} > 0$. We now define the weight ρ that will be of importance for the analysis that we will perform:

$$\text{if } m = 1,\ \rho(x) = \mathsf{d}_t^{\alpha}(x),\ \text{otherwise, } \rho(x) = \begin{cases} \mathsf{d}_t^{\alpha}(x), & \exists t \in \mathcal{D} : \mathsf{d}_t(x) < \frac{d_{\mathcal{D}}}{2}, \\ 1, & \mathsf{d}_t(x) \geq \frac{d_{\mathcal{D}}}{2}\ \forall\, t \in \mathcal{D}, \end{cases} \tag{33}$$

where $\mathsf{d}_t(x) := |x - t|$ and $\alpha \in (d-2, d)$. The following comments provide a brief insight into the definition of ρ. When x is close to a point $t \in \mathcal{D}$, i.e., $\mathsf{d}_t(x) < d_{\mathcal{D}}/2$, the weight behaves as $\mathsf{d}_t^{\alpha}(x)$. We immediately notice that the weight ρ is well-defined in the following sense: if $\mathsf{d}_t(x) < d_{\mathcal{D}}/2$, then $\mathsf{d}_{t'}(x) > d_{\mathcal{D}}/2$ for $t' \in \mathcal{D} \setminus \{t\}$, the latter being a consequence of the definition of $d_{\mathcal{D}} > 0$. If x is far from all points $t \in \mathcal{D}$, then $\rho(x) = 1$. Finally, since $(d-2, d) \subset (-d, d)$, it can be proved that the weight ρ belongs to the Muckenhoupt class $A_2(\mathbb{R}^d)$ [2, Theorem 6].

We present the following embedding result.

THEOREM 6 ($\mathbf{H}^1_0(\rho, \Omega) \hookrightarrow \mathbf{L}^2(\Omega)$). *If $\alpha \in (d-2, 2)$, then $\mathbf{H}^1_0(\rho, \Omega) \hookrightarrow \mathbf{L}^2(\Omega)$. Moreover, the following weighted Poincaré inequality holds:*

$$\|\mathbf{v}\|_{\mathbf{L}^2(\Omega)} \lesssim \|\nabla \mathbf{v}\|_{\mathbf{L}^2(\rho,\Omega)} \quad \forall\, \mathbf{v} \in \mathbf{H}^1_0(\rho, \Omega), \tag{34}$$

where the hidden constant depends only on Ω and $d_{\mathcal{D}}$.

Proof. The proof follows from [4, Lemmas 1 and 2]. ∎

We now derive a regularity result in weighted Sobolev spaces.

LEMMA 7 (weighted regularity). *Let $(\mathbf{y}, p) \in \mathbf{H}^1_0(\Omega) \times L^2(\Omega)/\mathbb{R}$ be the solution to (28) with $\mathbf{u} \in \mathbb{U}_{ad}$. Thus, we have that $(\mathbf{y}, p) \in \mathbf{H}^1_0(\rho^{-1}, \Omega) \times L^2(\rho^{-1}, \Omega)/\mathbb{R}$.*

Proof. We prove that $\mathbf{y} \in \mathbf{H}^1_0(\rho^{-1}, \Omega)$; similar arguments can be used to obtain $p \in L^2(\rho^{-1}, \Omega)/\mathbb{R}$. For each $t \in \mathcal{D}$, we denote by $B(t)$ the ball of center t and radius $d_{\mathcal{D}}/2$ and set $A = \Omega \setminus \cup_{t \in \mathcal{D}} B(t)$. We thus have, for each $i \in \{1, \ldots, d\}$, that

$$\int_{\Omega} \rho^{-1} |\nabla \mathbf{y}_i|^2 = \int_A \rho^{-1} |\nabla \mathbf{y}_i|^2 + \sum_{t \in \mathcal{D}} \int_{B(t)} \rho^{-1} |\nabla \mathbf{y}_i|^2 = \mathrm{I}_i + \mathrm{II}_i.$$

We first estimate I_i. In view of definition (33), we conclude that there exists $a > 0$ such that $\rho(x) \geq a$ for every $x \in A$. Consequently, since $\mathbf{y} \in \mathbf{H}^1_0(\Omega)$, we conclude in view of (29) that

$$\mathrm{I}_i = \int_A \rho^{-1} |\nabla \mathbf{y}_i|^2 \lesssim \int_A |\nabla \mathbf{y}_i|^2 \leq \|\nabla \mathbf{y}\|^2_{\mathbf{L}^2(\Omega)} \lesssim \|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)}.$$

We now bound II_i. Since $B(t) \Subset \Omega$, $\rho \in A_2(\mathbb{R}^d)$, and $\mathbf{u} \in \mathbb{U}_{ad} \subset \mathbf{L}^{\infty}(\Omega)$, we can apply the results of Proposition 1 to arrive at the estimate

$$\int_{B(t)} \rho^{-1} |\nabla \mathbf{y}_i|^2 \lesssim \rho^{-1}(B(t)) \|\mathbf{u}\|^2_{\mathbf{L}^l(\Omega)}, \quad l > d, \quad i \in \{1, \ldots, d\},$$

which implies that $\mathrm{II}_i \lesssim \|\mathbf{u}\|^2_{\mathbf{L}^l(\Omega)}$ for $l > d$. This concludes the proof. ∎

To explore (32) we introduce the adjoint pair $(\mathbf{z}, r)$ as the unique solution to the problem: Find $(\mathbf{z}, r) \in \mathbf{H}^1_0(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R}$ such that

$$a(\mathbf{w}, \mathbf{z}) - b(\mathbf{w}, r) = \sum_{t \in \mathcal{D}} \langle (\mathbf{y} - \mathbf{y}_t)\delta_t, \mathbf{w} \rangle, \quad b(\mathbf{z}, s) = 0 \tag{35}$$

for all $(\mathbf{w}, s) \in \mathbf{H}^1_0(\rho^{-1}, \Omega) \times L^2(\rho^{-1}, \Omega)/\mathbb{R}$, where $\mathbf{y} = \mathcal{S}\mathbf{u}$ solves (28). The well-posedness of (35) follows from [29, section 4] combined with the fact that $\delta_t \in H^1_0(\rho^{-1}, \Omega)'$ [23, Lemma 7.1.3].

THEOREM 8 (optimality conditions). *Let $\alpha \in (d-2, d)$. The pair $(\bar{\mathbf{y}}, \bar{\mathbf{u}}) \in \mathbf{H}^1_0(\Omega) \times \mathbb{U}_{ad}$ is optimal for the pointwise tracking optimal control problem* (27)–(28) *if and only if $\bar{\mathbf{y}} = \mathcal{S}\bar{\mathbf{u}}$ and $\bar{\mathbf{u}} \in \mathbb{U}_{ad}$ satisfies the variational inequality*

$$(\bar{\mathbf{z}} + \lambda \bar{\mathbf{u}}, \mathbf{u} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)} \geq 0 \quad \forall\, \mathbf{u} \in \mathbb{U}_{ad}, \tag{36}$$

where $(\bar{\mathbf{z}}, \bar{r}) \in \mathbf{H}^1_0(\rho, \Omega) \times \mathbf{L}^2(\rho, \Omega)/\mathbb{R}$ corresponds to the optimal adjoint state, which solves (35) *with* $\mathbf{y}$ *replaced by $\bar{\mathbf{y}} = \mathcal{S}\bar{\mathbf{u}}$. We recall that $\lambda > 0$ denotes the regularization parameter involved in the definition* (1) *of the cost functional J.*

Proof. A simple computation shows that, for all $\mathbf{u} \in \mathbb{U}_{ad}$, (32) can be written as

$$\sum_{t \in \mathcal{D}} (\mathcal{S}\bar{\mathbf{u}}(t) - \mathbf{y}_t)\, (\mathcal{S}\mathbf{u} - \mathcal{S}\bar{\mathbf{u}})(t) + \lambda(\bar{\mathbf{u}}, \mathbf{u} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)} \geq 0. \tag{37}$$

In what follows, to simplify the presentation of the material, we set $\mathbf{y} = \mathcal{S}\mathbf{u}$. Let us focus on the first term on the left-hand side of the previous expression. To study such a term, we invoke the results of Lemma 7 to conclude that $\mathbf{y} - \bar{\mathbf{y}} \in \mathbf{H}^1_0(\Omega) \cap \mathbf{H}^1_0(\rho^{-1}, \Omega)$ and $p - \bar{p} \in L^2(\Omega)/\mathbb{R} \cap L^2(\rho^{-1}, \Omega)/\mathbb{R}$. We can thus consider $\mathbf{w} = \mathbf{y} - \bar{\mathbf{y}}$ and $s = p - \bar{p}$ as test functions in problem (35). This yields, on the basis of the fact that $\operatorname{div}(\mathbf{y} - \bar{\mathbf{y}}) = 0$ a.e. in Ω, that

$$a(\mathbf{y} - \bar{\mathbf{y}}, \bar{\mathbf{z}}) = \sum_{t \in \mathcal{D}} (\bar{\mathbf{y}}(t) - \mathbf{y}_t)(\mathbf{y} - \bar{\mathbf{y}})(t). \tag{38}$$

Now, notice that $(\mathbf{y} - \bar{\mathbf{y}}, p - \bar{p}) \in \mathbf{H}^1_0(\Omega) \times L^2(\Omega)/\mathbb{R}$ solves the problem

$$a(\mathbf{y} - \bar{\mathbf{y}}, \mathbf{v}) + b(\mathbf{v}, p - \bar{p}) = (\mathbf{u} - \bar{\mathbf{u}}, \mathbf{v})_{\mathbf{L}^2(\Omega)}, \quad b(\mathbf{y} - \bar{\mathbf{y}}, q) = 0 \tag{39}$$

for all $\mathbf{v} \in \mathbf{H}^1_0(\Omega)$ and $q \in L^2(\Omega)/\mathbb{R}$. A density argument thus yields

$$a(\mathbf{y} - \bar{\mathbf{y}}, \bar{\mathbf{z}}) = (\mathbf{u} - \bar{\mathbf{u}}, \bar{\mathbf{z}})_{\mathbf{L}^2(\Omega)}. \tag{40}$$

In fact, let $\{\mathbf{z}_n\}_{n \in \mathbb{N}} \subset \mathbf{C}^\infty_0(\Omega)$ be such that $\mathbf{z}_n \to \bar{\mathbf{z}}$ in $\mathbf{H}^1_0(\rho, \Omega)$. We can thus set, for $n \in \mathbb{N}$, $\mathbf{v} = \mathbf{z}_n$ and $q = 0$ in (39). This yields $a(\mathbf{y} - \bar{\mathbf{y}}, \mathbf{z}_n) + b(\mathbf{z}_n, p - \bar{p}) = (\mathbf{u} - \bar{\mathbf{u}}, \mathbf{z}_n)_{\mathbf{L}^2(\Omega)}$. We now observe that

$$\left| \left(\mathbf{u} - \bar{\mathbf{u}}, \bar{\mathbf{z}}\right)_{\mathbf{L}^2(\Omega)} - (\mathbf{u} - \bar{\mathbf{u}}, \mathbf{z}_n)_{\mathbf{L}^2(\Omega)} \right| \leq \rho^{-1}(\Omega)^{\frac{1}{2}} \|\mathbf{u} - \bar{\mathbf{u}}\|_{\mathbf{L}^\infty(\Omega)} \|\bar{\mathbf{z}} - \mathbf{z}_n\|_{\mathbf{L}^2(\rho,\Omega)} \to 0$$

as $n \to \infty$ upon using a Poincaré inequality. The continuity of the bilinear form b on $\mathbf{H}^1_0(\rho, \Omega) \times L^2(\rho^{-1}, \Omega)$ immediately implies that $b(\mathbf{z}_n, p - \bar{p})$ converges to 0. Finally, the continuity of the bilinear form a on $\mathbf{H}^1_0(\rho^{-1}, \Omega) \times \mathbf{H}^1_0(\rho, \Omega)$ and the fact that $\mathbf{y} - \bar{\mathbf{y}} \in \mathbf{H}^1_0(\rho^{-1}, \Omega)$ allow us to obtain the required expression (40). This, (37), and (38) allow us to conclude. ∎

In order to obtain an explicit characterization for the optimal control variable $\bar{\mathbf{u}}$, we introduce the projection operator $\Pi : \mathbf{L}^1(\Omega) \to \mathbb{U}_{ad}$ as $\Pi(\mathbf{v}) := \min\{\mathbf{b}, \max\{\mathbf{v}, \mathbf{a}\}\}$. With this projector at hand, we recall the so-called projection formula (see [30, Lemma 2.26]): The optimal control $\bar{\mathbf{u}}$ satisfies (36) if and only if

$$\bar{\mathbf{u}} = \Pi\left(-\lambda^{-1}\bar{\mathbf{z}}\right) \text{ a.e. in } \Omega. \tag{41}$$

To summarize, the pair $(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ is optimal for the pointwise tracking optimal control problem (27)–(28) if and only if $(\bar{\mathbf{y}}, \bar{p}, \bar{\mathbf{z}}, \bar{r}, \bar{\mathbf{u}}) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R} \times \mathbf{H}_0^1(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R} \times \mathbb{U}_{ad}$ solves (28), (35), and (36).

5. A posteriori error analysis for the optimal control problem. The optimal adjoint pair $(\bar{\mathbf{z}}, \bar{r})$, which solves (35), exhibits reduced regularity properties. In fact, near a point $t \in \mathcal{D}$, $|\nabla \bar{\mathbf{z}}(x)| \approx |x - t|^{1-d}$ and $|\bar{r}(x)| \approx |x - t|^{1-d}$ [18, section IV.2]. This implies that $(\bar{\mathbf{z}}, \bar{r}) \notin \mathbf{H}^2(\Omega) \times H^1(\Omega)$ and, as a consequence, optimal error estimates for a standard a priori error analysis of (27)–(28) cannot be expected. This motivates the development and analysis of AFEMs for problem (27)–(28). In addition, as is customary in a posteriori error analysis, the study of AFEMs is also motivated by restrictions on the domain Ω that are needed to perform an optimal a priori error analysis. In the following section we will propose and analyze a reliable and efficient a posteriori error estimator for the optimal control problem (27)–(28). To accomplish this task, we begin by introducing a discrete scheme.

5.1. Finite element discretization. In order to propose a solution technique for problem (27)–(28), we define

$$\mathbb{U}_{ad}(\mathscr{T}) := \mathbf{U}(\mathscr{T}) \cap \mathbb{U}_{ad}, \quad \mathbf{U}(\mathscr{T}) := \{\mathbf{u} \in \mathbf{C}(\bar{\Omega}) \ : \ \mathbf{u}|_T \in \mathbb{P}_2(T)^d \ \forall\, T \in \mathscr{T}\}.$$

The discrete counterpart of (27)–(28) thus reads as follows: Find min $J(\mathbf{y}_{\mathscr{T}}, \mathbf{u}_{\mathscr{T}})$ subject to the discrete state equations

$$a(\mathbf{y}_{\mathscr{T}}, \mathbf{v}_{\mathscr{T}}) + b(\mathbf{v}_{\mathscr{T}}, p_{\mathscr{T}}) = (\mathbf{u}_{\mathscr{T}}, \mathbf{v}_{\mathscr{T}})_{\mathbf{L}^2(\Omega)}, \qquad b(\mathbf{y}_{\mathscr{T}}, q_{\mathscr{T}}) = 0 \tag{42}$$

for all $\mathbf{v}_{\mathscr{T}} \in \mathbf{V}(\mathscr{T})$ and $q_{\mathscr{T}} \in Q(\mathscr{T})$, and the discrete control constraints $\mathbf{u}_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T})$. Standard arguments reveal the existence of a unique optimal pair $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}})$. In addition, the pair $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}})$ is optimal if and only if $\bar{\mathbf{y}}_{\mathscr{T}}$ solves (42) and $\bar{\mathbf{u}}_{\mathscr{T}}$ satisfies

$$(\bar{\mathbf{z}}_{\mathscr{T}} + \lambda \bar{\mathbf{u}}_{\mathscr{T}}, \mathbf{u}_{\mathscr{T}} - \bar{\mathbf{u}}_{\mathscr{T}})_{\mathbf{L}^2(\Omega)} \geq 0 \quad \forall\, \mathbf{u}_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T}), \tag{43}$$

where $(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}})$ solves

$$a(\mathbf{w}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}) - b(\mathbf{w}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}) = \sum_{t \in \mathcal{D}} \langle (\bar{\mathbf{y}}_{\mathscr{T}} - \mathbf{y}_t)\delta_t, \mathbf{w}_{\mathscr{T}} \rangle, \qquad b(\bar{\mathbf{z}}_{\mathscr{T}}, s_{\mathscr{T}}) = 0 \tag{44}$$

for all $\mathbf{w}_{\mathscr{T}} \in \mathbf{V}(\mathscr{T})$ and $s_{\mathscr{T}} \in Q(\mathscr{T})$.

5.2. A posteriori error estimates. We now construct the error estimators associated with the state and adjoint equations, (28) and (35), respectively. To accomplish this task, we introduce the following auxiliary variables: Let $(\hat{\mathbf{y}}, \hat{p}) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R}$ and $(\hat{\mathbf{z}}, \hat{r}) \in \mathbf{H}_0^1(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R}$ be the solutions to

$$a(\hat{\mathbf{y}}, \mathbf{v}) + b(\mathbf{v}, \hat{p}) = (\bar{\mathbf{u}}_{\mathscr{T}}, \mathbf{v})_{\mathbf{L}^2(\Omega)} \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega), \quad b(\hat{\mathbf{y}}, q) = 0 \quad \forall q \in L^2(\Omega)/\mathbb{R}, \tag{45}$$

and

$$a(\mathbf{w}, \hat{\mathbf{z}}) - b(\mathbf{w}, \hat{r}) = \sum_{t \in \mathcal{D}} \langle (\bar{\mathbf{y}}_{\mathscr{T}} - \mathbf{y}_t)\delta_t, \mathbf{w} \rangle, \qquad b(\hat{\mathbf{z}}, s) = 0 \tag{46}$$

for all $(\mathbf{w}, s) \in \mathbf{H}^1_0(\rho^{-1}, \Omega) \times L^2(\rho^{-1}, \Omega)/\mathbb{R}$, respectively. We immediately notice that $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}})$ and $(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}})$ can be seen as finite element approximations of $(\hat{\mathbf{y}}, \hat{p})$ and $(\hat{\mathbf{z}}, \hat{r})$, respectively. We thus introduce the following local error indicators:

$$
\begin{aligned}
E_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) := \Big(& h_T^2 \|\bar{\mathbf{u}}_{\mathscr{T}} + \Delta \bar{\mathbf{y}}_{\mathscr{T}} - \nabla \bar{p}_{\mathscr{T}}\|^2_{\mathbf{L}^2(T)} \\
& + \tfrac{h_T}{2} \|[\![\nabla \bar{\mathbf{y}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(\partial T \setminus \partial\Omega)} + \|\operatorname{div} \bar{\mathbf{y}}_{\mathscr{T}}\|^2_{L^2(T)} \Big)^{\frac{1}{2}},
\end{aligned} \tag{47}
$$

$$
\begin{aligned}
\mathcal{E}_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) := & h_T^2 \|\bar{\mathbf{u}}_{\mathscr{T}} + \Delta \bar{\mathbf{y}}_{\mathscr{T}} - \nabla \bar{p}_{\mathscr{T}}\|_{\mathbf{L}^\infty(T)} \\
& + \tfrac{h_T}{2} \|[\![\nabla \bar{\mathbf{y}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(\partial T \setminus \partial\Omega)} + h_T \|\operatorname{div} \bar{\mathbf{y}}_{\mathscr{T}}\|_{L^\infty(T)},
\end{aligned} \tag{48}
$$

$$
\begin{aligned}
\mathcal{E}_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}) := \Bigg(& h_T^2 D_T^\alpha \|\Delta \bar{\mathbf{z}}_{\mathscr{T}} + \nabla \bar{r}_{\mathscr{T}}\|^2_{\mathbf{L}^2(T)} + \|\operatorname{div} \bar{\mathbf{z}}_{\mathscr{T}}\|^2_{L^2(\rho,T)} \\
& + h_T D_T^\alpha \|[\![\nabla \bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(\partial T \setminus \partial\Omega)} \\
& + \sum_{t \in \mathcal{D}} h_T^{\alpha+2-d} |\bar{\mathbf{y}}_{\mathscr{T}}(t) - \mathbf{y}_t|^2 \chi(\{t \in T\}) \Bigg)^{\frac{1}{2}},
\end{aligned} \tag{49}
$$

where $D_T = \min_{t \in \mathcal{D}} \{\max_{x \in T} |x - t|\}$. With these local error indicators at hand, we introduce the following a posteriori error estimators:

$$
E_{st}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) := \left(\sum_{T \in \mathscr{T}} E^2_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \right)^{\frac{1}{2}}, \tag{50}
$$

$$
\mathcal{E}_{st}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) := \max_{T \in \mathscr{T}} \mathcal{E}_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}), \tag{51}
$$

$$
\mathcal{E}_{ad}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}) := \left(\sum_{T \in \mathscr{T}} \mathcal{E}^2_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}) \right)^{\frac{1}{2}}. \tag{52}
$$

We assume that

$$
\forall\, T \in \mathscr{T},\ \#(\mathcal{N}^*_T \cap \mathcal{D}) \leq 1, \tag{53}
$$

that is, for every element $T \in \mathscr{T}$ its patch $\mathcal{N}^*_T$ contains at most one observable point. This is not a restrictive assumption, as it can always be satisfied by starting with a suitably refined mesh $\mathscr{T}_0$.

In view of [32, Theorem 4.70], Theorem 2, and [5, Theorem 7], we can immediately conclude the following estimates:

$$
\|\hat{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim E_{st}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}), \tag{54}
$$

$$
\|\hat{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim \ell_{\mathscr{T}}^{\beta_d} \mathcal{E}_{st}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}), \tag{55}
$$

and

$$
\|\nabla(\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)} + \|\hat{r} - \bar{r}_{\mathscr{T}}\|_{L^2(\rho,\Omega)} \lesssim \mathcal{E}_{ad}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}), \tag{56}
$$

where $\ell_{\mathscr{T}}$ is defined in (15), and β_d is provided in the statement of Theorem 2.

We now define the a posteriori error estimator associated to the discretization of the optimal control variable

$$\mathcal{E}_{ct}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) = \left(\sum_{T \in \mathscr{T}} \mathcal{E}_{ct,T}^2(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \right)^{\frac{1}{2}}, \tag{57}$$

based on the local error indicators

$$\mathcal{E}_{ct,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) = \|\bar{\mathbf{u}}_{\mathscr{T}} - \Pi(-\lambda^{-1}\bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(T)}. \tag{58}$$

On the basis of the previously introduced estimators, we define the global a posteriori error estimator for the optimal control problem (27)–(28) as the sum of four contributions:

$$\mathcal{E}_{ocp}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) := \left(\mathcal{E}_{st}^2 + \mathcal{E}_{ad}^2 + \mathcal{E}_{ct}^2 + E_{st}^2\right)^{\frac{1}{2}}. \tag{59}$$

In order to prove a reliability result for the error estimator (59), we introduce the following auxiliary variables: Let $(\tilde{\mathbf{y}}, \tilde{p}) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R}$ and $(\tilde{\mathbf{z}}, \tilde{r}) \in \mathbf{H}_0^1(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R}$ be the solutions to

$$a(\tilde{\mathbf{y}}, \mathbf{v}) + b(\mathbf{v}, \tilde{p}) = (\tilde{\mathbf{u}}, \mathbf{v})_{\mathbf{L}^2(\Omega)} \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega), \qquad b(\tilde{\mathbf{y}}, q) = 0 \quad \forall q \in L^2(\Omega)/\mathbb{R}, \tag{60}$$

and

$$a(\mathbf{w}, \tilde{\mathbf{z}}) - b(\mathbf{w}, \tilde{r}) = \sum_{t \in \mathcal{D}} \langle (\tilde{\mathbf{y}} - \mathbf{y}_t)\delta_t, \mathbf{w} \rangle, \qquad b(\tilde{\mathbf{z}}, s) = 0 \tag{61}$$

for all $(\mathbf{w}, s) \in \mathbf{H}_0^1(\rho^{-1}, \Omega) \times L^2(\rho^{-1}, \Omega)/\mathbb{R}$, respectively, where $\tilde{\mathbf{u}} := \Pi(-\lambda^{-1}\bar{\mathbf{z}}_{\mathscr{T}})$.

Finally, we define $\mathbf{e_z} := \bar{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}}$, $e_r := \bar{r} - \bar{r}_{\mathscr{T}}$, $\mathbf{e_u} := \bar{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}$, and

$$\|\mathbf{e}\|_{\Omega}^2 := \|\mathbf{e_y}\|_{\mathbf{L}^\infty(\Omega)}^2 + \|e_p\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{e_z}\|_{\mathbf{L}^2(\rho,\Omega)}^2 + \|e_r\|_{L^2(\rho,\Omega)/\mathbb{R}}^2 + \|\mathbf{e_u}\|_{\mathbf{L}^2(\Omega)}^2, \tag{62}$$

where $\mathbf{e_y}$ and e_p are given as in the proof of Theorem 3.

5.3. A posteriori error estimator: Reliability. With all the previous ingredients at hand, we can establish the following result.

THEOREM 9 (global reliability of $\mathcal{E}_{ocp}$). *Let $(\bar{\mathbf{y}}, \bar{p}, \bar{\mathbf{z}}, \bar{r}, \bar{\mathbf{u}}) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R} \times \mathbf{H}_0^1(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R} \times \mathbb{U}_{ad}$ be the solution to the optimality system* (28), (35), *and* (36) *and $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ its numerical approximation given by* (42)–(44). *If $\alpha \in (d-2, 2)$, then*

$$\|\mathbf{e}\|_{\Omega}^2 \lesssim \ell_{\mathscr{T}}^{2\beta_d} \mathcal{E}_{st}^2 + \mathcal{E}_{ad}^2 + \mathcal{E}_{ct}^2 + E_{st}^2 \lesssim (1 + \ell_{\mathscr{T}}^{2\beta_d}) \mathcal{E}_{ocp}^2. \tag{63}$$

The term $\ell_{\mathscr{T}}$ is defined in (15), *β_d is given as in Theorem* 2, *and the hidden constants are independent of the continuous and discrete solutions, the size of the elements in the mesh $\mathscr{T}$ and $\#\mathscr{T}$. The constants, however, blow up as $\lambda \downarrow 0$.*

Proof. We proceed in six steps.

Step 1. We bound the error $\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)}$. To accomplish this task, we recall that $\tilde{\mathbf{u}} = \Pi(-\lambda^{-1}\bar{\mathbf{z}}_{\mathscr{T}})$ and notice that it can be equivalently characterized by

$$(\bar{\mathbf{z}}_{\mathscr{T}} + \lambda\tilde{\mathbf{u}}, \mathbf{u} - \tilde{\mathbf{u}})_{\mathbf{L}^2(\Omega)} \geq 0 \quad \forall\, \mathbf{u} \in \mathbb{U}_{ad} \tag{64}$$

[30, Lemma 2.26]. With the auxiliary control variable $\tilde{\mathbf{u}}$ at hand, a simple application of the triangle inequality yields

$$\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)} \leq \|\bar{\mathbf{u}} - \tilde{\mathbf{u}}\|_{\mathbf{L}^2(\Omega)} + \|\tilde{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)}. \tag{65}$$

In view of the definition of $\tilde{\mathbf{u}}$, the second term on the right-hand side of (65) corresponds to the global error estimator $\mathcal{E}_{ct}$ which is defined in (57). It thus suffices to control the term $\|\bar{\mathbf{u}} - \tilde{\mathbf{u}}\|_{\mathbf{L}^2(\Omega)}$. We thus begin by setting $\mathbf{u} = \tilde{\mathbf{u}}$ in (36) and $\mathbf{u} = \bar{\mathbf{u}}$ in (64). Adding the obtained inequalities we arrive at

$$\lambda\|\bar{\mathbf{u}} - \tilde{\mathbf{u}}\|^2_{\mathbf{L}^2(\Omega)} \leq (\bar{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)}. \tag{66}$$

We now invoke the auxiliary adjoint states $\hat{\mathbf{z}}$ and $\tilde{\mathbf{z}}$, defined as the solutions to problems (46) and (61), respectively, to write the previous inequality as follows:

$$\lambda\|\bar{\mathbf{u}} - \tilde{\mathbf{u}}\|^2_{\mathbf{L}^2(\Omega)} \leq (\bar{\mathbf{z}} - \tilde{\mathbf{z}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)} + (\tilde{\mathbf{z}} - \hat{\mathbf{z}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)} + (\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)}.$$

We bound the term $(\bar{\mathbf{z}} - \tilde{\mathbf{z}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)}$. Notice that $(\tilde{\mathbf{y}} - \bar{\mathbf{y}}, \tilde{p} - \bar{p}) \in \mathbf{H}^1_0(\Omega) \times L^2(\Omega)/\mathbb{R}$ and $(\bar{\mathbf{z}} - \tilde{\mathbf{z}}, \bar{r} - \tilde{r}) \in \mathbf{H}^1_0(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R}$ solve

$$a(\tilde{\mathbf{y}} - \bar{\mathbf{y}}, \mathbf{v}) + b(\mathbf{v}, \tilde{p} - \bar{p}) = (\tilde{\mathbf{u}} - \bar{\mathbf{u}}, \mathbf{v})_{\mathbf{L}^2(\Omega)}, \quad b(\tilde{\mathbf{y}} - \bar{\mathbf{y}}, q) = 0 \tag{67}$$

for all $\mathbf{v} \in \mathbf{H}^1_0(\Omega)$ and $q \in L^2(\Omega)/\mathbb{R}$, and

$$a(\mathbf{w}, \bar{\mathbf{z}} - \tilde{\mathbf{z}}) - b(\mathbf{w}, \bar{r} - \tilde{r}) = \sum_{t \in \mathcal{D}} \langle (\bar{\mathbf{y}} - \tilde{\mathbf{y}})\delta_t, \mathbf{w} \rangle, \quad b(\bar{\mathbf{z}} - \tilde{\mathbf{z}}, s) = 0 \tag{68}$$

for all $\mathbf{w} \in \mathbf{H}^1_0(\rho^{-1}, \Omega)$ and $s \in L^2(\rho^{-1}, \Omega)/\mathbb{R}$, respectively. We thus set $\mathbf{w} = \tilde{\mathbf{y}} - \bar{\mathbf{y}} \in \mathbf{H}^1_0(\Omega) \cap \mathbf{H}^1_0(\rho^{-1}, \Omega)$ in (68). Similar density arguments to the ones developed in the proof of Theorem 8 reveal that (67) holds with $\mathbf{v} = \bar{\mathbf{z}} - \tilde{\mathbf{z}}$. Consequently,

$$(\bar{\mathbf{z}} - \tilde{\mathbf{z}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)} = -\sum_{t \in \mathcal{D}} |\bar{\mathbf{y}}(t) - \tilde{\mathbf{y}}(t)|^2 \leq 0.$$

This estimate allows us to conclude that

$$\lambda\|\bar{\mathbf{u}} - \tilde{\mathbf{u}}\|^2_{\mathbf{L}^2(\Omega)} \leq (\tilde{\mathbf{z}} - \hat{\mathbf{z}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)} + (\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}}, \tilde{\mathbf{u}} - \bar{\mathbf{u}})_{\mathbf{L}^2(\Omega)}.$$

Standard estimates combined with the weighted Poincaré inequality of Theorem 6 allow us to arrive at

$$\|\bar{\mathbf{u}} - \tilde{\mathbf{u}}\|^2_{\mathbf{L}^2(\Omega)} \lesssim \|\tilde{\mathbf{z}} - \hat{\mathbf{z}}\|^2_{\mathbf{L}^2(\Omega)} + \|\nabla(\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|^2_{\mathbf{L}^2(\rho,\Omega)} \lesssim \|\tilde{\mathbf{z}} - \hat{\mathbf{z}}\|^2_{\mathbf{L}^2(\Omega)} + \mathcal{E}^2_{ad}, \tag{69}$$

where, in the last inequality, we have used the a posteriori error estimate (56).

We now bound the term $\|\tilde{\mathbf{z}} - \hat{\mathbf{z}}\|_{\mathbf{L}^2(\Omega)}$. Notice that the pair $(\tilde{\mathbf{z}} - \hat{\mathbf{z}}, \tilde{r} - \hat{r}) \in \mathbf{H}^1_0(\rho, \Omega) \times L^2(\rho, \Omega)$ solves

$$a(\mathbf{w}, \tilde{\mathbf{z}} - \hat{\mathbf{z}}) - b(\mathbf{w}, \tilde{r} - \hat{r}) = \sum_{t \in \mathcal{D}} \langle (\tilde{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}})\delta_t, \mathbf{w} \rangle, \qquad b(\tilde{\mathbf{z}} - \hat{\mathbf{z}}, s) = 0$$

for all $\mathbf{w} \in \mathbf{H}^1_0(\rho^{-1}, \Omega)$ and $s \in L^2(\rho^{-1}, \Omega)/\mathbb{R}$. We thus first apply the estimate of Theorem 6 and then the stability estimate of [29, Theorem 14] to conclude that

$$\|\tilde{\mathbf{z}} - \hat{\mathbf{z}}\|_{\mathbf{L}^2(\Omega)} \lesssim \|\nabla(\tilde{\mathbf{z}} - \hat{\mathbf{z}})\|_{\mathbf{L}^2(\rho,\Omega)} \lesssim \|\tilde{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)}. \tag{70}$$

To control the right-hand side of the previous expression, we use the triangle inequality to obtain that

$$\|\tilde{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \leq \|\tilde{\mathbf{y}} - \hat{\mathbf{y}}\|_{\mathbf{L}^\infty(\Omega)} + \|\hat{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)};$$

the pair $(\hat{\mathbf{y}}, \hat{p}) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R}$ solves (45). The results of [12, Theorem 2.9] (see also [24, Theorem 1.1] and [14, Lemma 12]) guarantee the existence of $l > d$ such that $\|\tilde{\mathbf{y}} - \hat{\mathbf{y}}\|_{\mathbf{W}^{1,l}(\Omega)} \lesssim \|\tilde{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{W}^{-1,l}(\Omega)}$. Thus,

$$\|\tilde{\mathbf{y}} - \hat{\mathbf{y}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim \|\tilde{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{W}^{-1,l}(\Omega)} \lesssim \|\tilde{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)} = \mathcal{E}_{ct}. \tag{71}$$

Now, since $\bar{\mathbf{y}}_{\mathscr{T}}$ is the Galerkin approximation of $\hat{\mathbf{y}}$, the term $\|\hat{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)}$ is bounded in view of the global reliability of the a posteriori error estimator $\mathcal{E}_{st}$ defined in (51): $\|\hat{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim \ell_{\mathscr{T}}^{\beta_d} \mathcal{E}_{st}$. Replacing the obtained estimates into (70), we obtain that $\|\tilde{\mathbf{z}} - \hat{\mathbf{z}}\|_{\mathbf{L}^2(\Omega)} \lesssim \mathcal{E}_{ct} + \ell_{\mathscr{T}}^{\beta_d} \mathcal{E}_{st}$. This, in light of (69) and (65), allows us to conclude

$$\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)}^2 \lesssim (1 + \ell_{\mathscr{T}}^{2\beta_d}) \mathcal{E}_{ocp}^2. \tag{72}$$

Step 2. The goal of this step is to bound the error $\|\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)}$. To accomplish this task, we write $\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}} = (\bar{\mathbf{y}} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}})$, and estimate each term separately. To control the first term we invoke a similar argument to the one that yields (71): $\|\bar{\mathbf{y}} - \hat{\mathbf{y}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim \|\bar{\mathbf{y}} - \hat{\mathbf{y}}\|_{\mathbf{W}^{1,l}(\Omega)} \lesssim \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)}$, which can be bounded with the use of (72). On the other hand, by using the global reliability of the error estimator $\mathcal{E}_{st}$ we arrive at $\|\hat{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim \ell_{\mathscr{T}}^{\beta_d} \mathcal{E}_{st}$. The collection of the previous results yield

$$\|\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)}^2 \lesssim (1 + \ell_{\mathscr{T}}^{2\beta_d}) \mathcal{E}_{ocp}^2. \tag{73}$$

Step 3. We bound the error $\nabla(\bar{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})$ in the $\mathbf{L}^2(\rho, \Omega)$-norm. A simple application of the triangle inequality yields

$$\|\nabla(\bar{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)} \leq \|\nabla(\bar{\mathbf{z}} - \hat{\mathbf{z}})\|_{\mathbf{L}^2(\rho,\Omega)} + \|\nabla(\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)}. \tag{74}$$

The first term on the right-hand side of the previous expression can be bounded in view of the stability estimate of [29, Theorem 14] and (73). In fact,

$$\|\nabla(\bar{\mathbf{z}} - \hat{\mathbf{z}})\|_{\mathbf{L}^2(\rho,\Omega)} \lesssim \|\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim (1 + \ell_{\mathscr{T}}^{\beta_d}) \mathcal{E}_{ocp}. \tag{75}$$

To control $\|\nabla(\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)}$, we resort to the global reliability of the error estimator $\mathcal{E}_{ad}$: $\|\nabla(\hat{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)} \lesssim \mathcal{E}_{ad}$. With this estimate at hand, we thus replace (75) into (74) to obtain that

$$\|\nabla(\bar{\mathbf{z}} - \bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)}^2 \lesssim (1 + \ell_{\mathscr{T}}^{2\beta_d}) \mathcal{E}_{ocp}^2. \tag{76}$$

Step 4. The goal of this step is to bound the term $\|\bar{r} - \bar{r}_{\mathscr{T}}\|_{L^2(\rho,\Omega)/\mathbb{R}}$. We write $\bar{r} - \bar{r}_{\mathscr{T}} = (\bar{r} - \hat{r}) + (\hat{r} - \bar{r}_{\mathscr{T}})$ and immediately notice that [29, Theorem 14] and (73) yields $\|\bar{r} - \hat{r}\|_{L^2(\rho,\Omega)/\mathbb{R}} \lesssim \|\bar{\mathbf{y}} - \bar{\mathbf{y}}_{\mathscr{T}}\|_{\mathbf{L}^\infty(\Omega)} \lesssim (1 + \ell_{\mathscr{T}}^{\beta_d}) \mathcal{E}_{ocp}$. We now invoke the global reliability of the error estimator $\mathcal{E}_{ad}$: $\|\hat{r} - \bar{r}_{\mathscr{T}}\|_{L^2(\rho,\Omega)/\mathbb{R}} \lesssim \mathcal{E}_{ad}$. The collection of our derived results allow us to arrive at

$$\|\bar{r} - \bar{r}_{\mathscr{T}}\|_{L^2(\rho,\Omega)/\mathbb{R}} \lesssim (1 + \ell_{\mathscr{T}}^{\beta_d}) \mathcal{E}_{ocp}. \tag{77}$$

Step 5. To obtain the estimate (63), we must estimate the term $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)}$. To accomplish this task, we write $\bar{p} - \bar{p}_{\mathscr{T}} = (\bar{p} - \hat{p}) + (\hat{p} - \bar{p}_{\mathscr{T}})$ and estimate each term separately. To estimate the first term, we use (29) and (72) to obtain that

$$\|\bar{p} - \hat{p}\|_{L^2(\Omega)} \lesssim \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_{\mathscr{T}}\|_{\mathbf{L}^2(\Omega)} \lesssim (1 + \ell_{\mathscr{T}}^{\beta_d})\mathcal{E}_{ocp}. \tag{78}$$

To estimate $(\hat{p} - \bar{p}_{\mathscr{T}})$, we invoke the global reliability property (54) of the error estimator E_{st} to obtain that $\|\hat{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim E_{st}$. We thus collect the derived estimates to obtain that

$$\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim (1 + \ell_{\mathscr{T}}^{\beta_d})\mathcal{E}_{ocp}. \tag{79}$$

Step 6. The collection of the estimates (72), (73), (76), (77), and (79) yields the desired estimate (63). ∎

5.4. A posteriori error estimator: Efficiency. In what follows we examine the efficiency properties of the a posteriori error estimator $\mathcal{E}_{ocp}$, which is defined as in (59). To accomplish this task, we analyze each of its contributions separately.

5.4.1. Efficiency properties of $\mathcal{E}_{st}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{\boldsymbol{p}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}})$. We begin by introducing the error equation associated to the state equations (28). Let us consider $\mathbf{v} \in \mathbf{H}_0^1(\Omega)$, which is such that $\mathbf{v}|_T \in \mathbf{C}^2(T)$ for all $T \in \mathscr{T}$. Following similar arguments to the ones that yield (20) we obtain, from the *momentum equation* in (28), that

$$\begin{aligned} &-\sum_{T\in\mathscr{T}} \Big((\mathbf{e}_{\mathbf{y}}, \Delta \mathbf{v})_{\mathbf{L}^2(T)} + (e_p, \operatorname{div} \mathbf{v})_{L^2(T)} \Big) - \sum_{S\in\mathscr{S}} (\mathbf{e}_{\mathbf{y}}, [\![\nabla \mathbf{v} \cdot \boldsymbol{\nu}]\!])_{\mathbf{L}^2(S)} \\ &\quad = \sum_{T\in\mathscr{T}} \Big((\bar{\mathbf{u}}_{\mathscr{T}} + \Delta \bar{\mathbf{y}}_{\mathscr{T}} - \nabla \bar{p}_{\mathscr{T}}, \mathbf{v})_{\mathbf{L}^2(T)} + (\mathbf{e}_{\mathbf{u}}, \mathbf{v})_{\mathbf{L}^2(T)} \Big) + \sum_{S\in\mathscr{S}} ([\![\nabla \bar{\mathbf{y}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], \mathbf{v})_{\mathbf{L}^2(S)}. \end{aligned} \tag{80}$$

On the basis of this error equation, the following efficiency result can be derived.

THEOREM 10 (local efficiency of $\mathcal{E}_{st}$). *Let $(\bar{\mathbf{y}}, \bar{p}, \bar{\mathbf{z}}, \bar{r}, \bar{\mathbf{u}}) \in \mathbf{H}_0^1(\Omega) \times L^2(\Omega)/\mathbb{R} \times \mathbf{H}_0^1(\rho, \Omega) \times L^2(\rho, \Omega)/\mathbb{R} \times \mathbb{U}_{ad}$ be the solution to the optimality system* (28), (35), *and* (36), *and $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ its numerical approximation given by* (42)–(44). *Then, for $T \in \mathscr{T}$, the local error indicator $\mathcal{E}_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}})$, defined in* (48), *satisfies that*

$$\mathcal{E}_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \lesssim \|\mathbf{e}_{\mathbf{y}}\|_{\mathbf{L}^\infty(\mathcal{N}_T)} + h_T^{1-\frac{d}{2}} \|e_p\|_{L^2(\mathcal{N}_T)} + h_T^{2-\frac{d}{2}} \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{L}^2(\mathcal{N}_T)}, \tag{81}$$

where $\mathcal{N}_T$ is defined as in (9). *The hidden constant is independent of the continuous and discrete solutions, the size of the elements in the mesh $\mathscr{T}$ and $\#\mathscr{T}$.*

Proof. The proof follows the arguments developed in the proof of Theorem 3. In fact, on the basis of (80), we proceed in three steps.

Step 1. Let $T \in \mathscr{T}$. We estimate the term $h_T^2 \|\bar{\mathbf{u}}_{\mathscr{T}} + \Delta \bar{\mathbf{y}}_{\mathscr{T}} - \nabla \bar{p}_{\mathscr{T}}\|_{\mathbf{L}^\infty(T)}$ in (48). To accomplish this task, and in order to simplify the presentation of the material, we define $\mathfrak{R}_T := (\bar{\mathbf{u}}_{\mathscr{T}} + \Delta \bar{\mathbf{y}}_{\mathscr{T}} - \nabla \bar{p}_{\mathscr{T}})|_T$. To derive the desired bound, we set $\mathbf{v} = \varphi_T^2 \mathfrak{R}_T$ in (80) and invoke Hölder's inequality. These arguments yield

$$\begin{aligned} \|\varphi_T \mathfrak{R}_T\|_{\mathbf{L}^2(T)}^2 &\lesssim \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{L}^2(T)} \|\varphi_T^2 \mathfrak{R}_T\|_{\mathbf{L}^2(T)} \\ &\quad + \|\mathbf{e}_{\mathbf{y}}\|_{\mathbf{L}^\infty(T)} \|\Delta(\varphi_T^2 \mathfrak{R}_T)\|_{\mathbf{L}^1(T)} + \|e_p\|_{L^2(T)} \|\operatorname{div}(\varphi_T^2 \mathfrak{R}_T)\|_{L^2(T)}. \end{aligned} \tag{82}$$

We now use properties of the function φ_T and similar arguments to the ones that yield (22) and (23) to arrive at

$$\|\mathfrak{R}_T\|_{\mathbf{L}^2(T)} \lesssim \|\mathbf{e_u}\|_{\mathbf{L}^2(T)} + h_T^{\frac{d}{2}-2}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T)} + h_T^{-1}\|e_p\|_{L^2(T)}. \tag{83}$$

Finally, we use the inverse estimate $\|\mathfrak{R}_T\|_{\mathbf{L}^\infty(T)} \lesssim h_T^{-\frac{d}{2}}\|\mathfrak{R}_T\|_{\mathbf{L}^2(T)}$ to conclude that

$$h_T^2\|\mathfrak{R}_T\|_{\mathbf{L}^\infty(T)} \lesssim h_T^{2-\frac{d}{2}}\|\mathbf{e_u}\|_{\mathbf{L}^2(T)} + \|\mathbf{e_y}\|_{\mathbf{L}^\infty(T)} + h_T^{1-\frac{d}{2}}\|e_p\|_{L^2(T)}. \tag{84}$$

Step 2. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We bound $h_T\|[\![\nabla \bar{\mathbf{y}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(\partial T \setminus \partial\Omega)}$ in (48). To accomplish this task, we invoke the vector-valued bubble function $\boldsymbol{\varphi}_S$ described in the proof of Theorem 3, set $\mathbf{v} = \boldsymbol{\varphi}_S$ in (80), and utilize the properties (24) to conclude that

$$\int_S [\![\nabla \bar{\mathbf{y}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\boldsymbol{\varphi}_S \lesssim \sum_{T' \in \mathcal{N}_S} \Big(\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T')}\|\Delta \boldsymbol{\varphi}_S\|_{\mathbf{L}^1(T')} + \|e_p\|_{L^2(T')}\|\operatorname{div} \boldsymbol{\varphi}_S\|_{L^2(T')} + \|\mathfrak{R}_T\|_{\mathbf{L}^\infty(T')}\|\boldsymbol{\varphi}_S\|_{\mathbf{L}^1(T')} + \|\mathbf{e_u}\|_{\mathbf{L}^2(T')}\|\boldsymbol{\varphi}_S\|_{\mathbf{L}^2(T')} \Big).$$

With this estimate at hand, we invoke standard arguments and the derived estimate for $\|\mathfrak{R}_T\|_{\mathbf{L}^\infty(T)}$ to arrive at

$$\int_S [\![\nabla \mathbf{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\boldsymbol{\varphi}_S \lesssim \sum_{T' \in \mathcal{N}_S} \Big(h_T^{d-2}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T')} + h_T^{\frac{d}{2}-1}\|e_p\|_{L^2(T')} + h_T^{\frac{d}{2}}\|\mathbf{e_u}\|_{\mathbf{L}^2(T')} \Big).$$

We thus replace the previous estimate into (25) and use that $|T|/|S| \approx h_T$ to obtain

$$h_T\|[\![\nabla \mathbf{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{\mathbf{L}^\infty(S)} \lesssim \sum_{T' \in \mathcal{N}_S} \Big(\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T')} + h_T^{1-\frac{d}{2}}\|e_p\|_{L^2(T')} + h_T^{2-\frac{d}{2}}\|\mathbf{e_u}\|_{\mathbf{L}^2(T')} \Big).$$

Step 3. Let $T \in \mathscr{T}$. The goal of this step is to control the term $h_T\|\operatorname{div} \bar{\mathbf{y}}_{\mathscr{T}}\|_{L^\infty(T)}$ in (14). This estimate follows from (26):

$$\|\operatorname{div} \bar{\mathbf{y}}_{\mathscr{T}}\|^2_{L^2(T)} \lesssim h_T^{\frac{d}{2}-1}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T)}\|\operatorname{div} \bar{\mathbf{y}}_{\mathscr{T}}\|_{L^2(T)}. \tag{85}$$

Consequently, the use of an inverse estimate yields

$$h_T\|\operatorname{div} \bar{\mathbf{y}}_{\mathscr{T}}\|_{L^\infty(T)} \lesssim \|\mathbf{e_y}\|_{\mathbf{L}^\infty(T)}. \tag{86}$$

The collection of the estimates derived in Steps 1, 2, and 3 yields (81). This concludes the proof. ∎

Remark 11 (efficiency for $d = 2$). Due to the presence of the term $h_T^{1-d/2}\|e_p\|_{L^2(\mathcal{N}_T)}$ in (81), we observe that the local indicator $\mathcal{E}_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}})$ is only efficient when $d = 2$.

We now derive efficiency properties for the local a posteriori error indicator $E_{st,T}$ introduced in (47).

THEOREM 12 (local efficiency of E_{st}). *Let $(\bar{\mathbf{y}}, \bar{p}, \bar{\mathbf{z}}, \bar{r}, \bar{\mathbf{u}}) \in \mathbf{H}^1_0(\Omega) \times L^2(\Omega)/\mathbb{R} \times \mathbf{H}^1_0(\rho,\Omega) \times L^2(\rho,\Omega)/\mathbb{R} \times \mathbb{U}_{ad}$ be the solution to the optimality system* (28), (35), *and* (36) *and $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ its numerical approximation given by* (42)–(44). *Then, for $T \in \mathscr{T}$, the local error indicator $E_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}})$, defined in* (47), *satisfies that*

$$E^2_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \lesssim h_T^{d-2}\|\mathbf{e_y}\|^2_{\mathbf{L}^\infty(\mathcal{N}_T)} + \|e_p\|^2_{L^2(\mathcal{N}_T)} + h_T^2\|\mathbf{e_u}\|^2_{\mathbf{L}^2(\mathcal{N}_T)}, \tag{87}$$

where $\mathcal{N}_T$ is defined as in (9). *The hidden constant is independent of the continuous and discrete solutions, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.*

Proof. The control of the terms $h_T^2\|\bar{\mathbf{u}}_{\mathscr{T}} + \Delta\bar{\mathbf{y}}_{\mathscr{T}} - \nabla\bar{p}_{\mathscr{T}}\|^2_{\mathbf{L}^2(T)}$ and $\|\operatorname{div}\bar{\mathbf{y}}_{\mathscr{T}}\|^2_{L^2(T)}$ in (47) follow directly from the estimates (83) and (85), respectively.

We now proceed to estimate the remaining term $h_T\|[\![\nabla\bar{\mathbf{y}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(\partial T\setminus\partial\Omega)}$ in (47). To accomplish this task, we set $\mathbf{v} = [\![\nabla\bar{\mathbf{y}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\varphi_S$ in (80), invoke (24), and utilize standard bubble functions arguments to conclude that

$$\begin{aligned}\|[\![\nabla\bar{\mathbf{y}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\varphi_S^{\frac12}\|^2_{\mathbf{L}^2(S)} \lesssim \sum_{T'\in\mathcal{N}_S}\Big(&h_T^{\frac{d}{2}-2}\|\mathbf{e_y}\|_{\mathbf{L}^\infty(T')} + h_T^{-1}\|e_p\|_{L^2(T')}\\ &+ \|\mathfrak{R}_T\|_{\mathbf{L}^2(T')} + \|\mathbf{e_u}\|_{\mathbf{L}^2(T')}\Big)h_T^{\frac12}\|[\![\nabla\bar{\mathbf{y}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^2(S)},\end{aligned}$$

where $\mathfrak{R}_T = (\bar{\mathbf{u}}_{\mathscr{T}} + \Delta\bar{\mathbf{y}}_{\mathscr{T}} - \nabla\bar{p}_{\mathscr{T}})|_T$. With this estimate at hand, we invoke standard arguments and the derived estimate for $\|\mathfrak{R}_T\|_{\mathbf{L}^2(T)}$ to arrive at

$$h_T\|[\![\nabla\bar{\mathbf{y}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(S)} \lesssim \sum_{T'\in\mathcal{N}_S}\Big(h_T^{d-2}\|\mathbf{e_y}\|^2_{\mathbf{L}^\infty(T')} + \|e_p\|^2_{L^2(T')} + h_T^2\|\mathbf{e_u}\|^2_{\mathbf{L}^2(T')}\Big). \tag{88}$$

This concludes the proof. ∎

5.4.2. Efficiency properties of $\boldsymbol{\mathcal{E}_{ad}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}})}$. To derive efficiency properties for the local error indicator $\mathcal{E}_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}})$, defined in (49), we utilize standard residual estimation techniques but on the basis of suitable bubble functions whose construction we owe to [1, section 5.2]; see also [4, section 5.1.2].

Given $T \in \mathscr{T}$, we consider a bubble function ψ_T which is such that

$$\psi_T(t) = 0 \ \ \forall\, t \in \mathcal{D}, \quad 0 \le \psi_T \le 1, \quad |T| \lesssim \int_T \psi_T, \quad \|\nabla\psi_T\|_{L^\infty(R_T)} \lesssim h_T^{-1}, \tag{89}$$

and there exists a simplex $T_* \subset T$ such that $R_T := \operatorname{supp}(\psi_T) \subset T_*$. Notice that, in light of (53), there is at most one $t \in \mathcal{D}$ for each element T. As a consequence of (89), we have, for every $g \in \mathbb{P}_2(R_T)$, that

$$\|g\|_{L^2(R_T)} \lesssim \|\psi_T^{\frac12} g\|_{L^2(R_T)}. \tag{90}$$

Given $S \in \mathscr{S}$, we also introduce a bubble function ψ_S such that $0 \le \psi_S \le 1$,

$$\psi_S(t) = 0 \ \ \forall\, t \in \mathcal{D}, \quad |S| \lesssim \int_S \psi_S, \quad \|\nabla\psi_S\|_{L^\infty(R_S)} \lesssim h_S^{-1} \tag{91}$$

where $R_S := \operatorname{supp}(\psi_S)$ is such that, if $\mathcal{N}_S = \{T, T'\}$, there exist two simplices $T_* \subset T$ and $T'_* \subset T'$ such that $R_S \subset T_* \cup T'_* \subset \mathcal{N}_S$; see Figure 1.

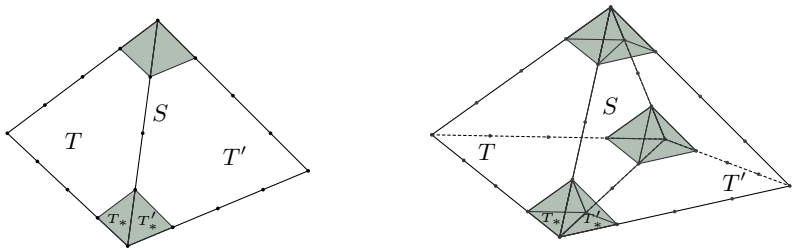

FIG. 1. *Simplices T and T', which share the side S, with their subsimplices T_* and T'_*, respectively, obtained after dividing their edges into four equal segments, in two dimensions (left) and three dimensions (right).*

An important ingredient in the analysis that we will provide below is the so-called *residual*. To define it, we introduce $\mathcal{Z} := \mathbf{H}^1_0(\rho^{-1},\Omega) \times L^2(\rho^{-1},\Omega)/\mathbb{R}$. We define the residual $\mathcal{R} = \mathcal{R}(\mathbf{z}_{\mathscr{T}}, r_{\mathscr{T}}) \in \mathcal{Z}'$ by

$$\langle \mathcal{R}, (\mathbf{w}, s)\rangle_{\mathcal{Z}',\mathcal{Z}} = \sum_{t\in\mathcal{D}} \langle (\mathbf{y}-\mathbf{y}_t)\delta_t, \mathbf{w}\rangle - (a(\mathbf{z},\mathbf{w}) - b(\mathbf{w},r) - b(\mathbf{z},s)), \tag{92}$$

where $\langle\cdot,\cdot\rangle_{\mathcal{Z}',\mathcal{Z}}$ denotes the duality pairing between $\mathcal{Z}'$ and $\mathcal{Z}$. We apply a standard integration by parts argument to conclude

$$\begin{aligned}\langle \mathcal{R}, (\mathbf{w}, s)\rangle_{\mathcal{Z}',\mathcal{Z}} &= \sum_{t\in\mathcal{D}} \langle (\mathbf{y}-\mathbf{y}_t)\delta_t, \mathbf{w}\rangle + \sum_{T\in\mathscr{T}} (\Delta \mathbf{z}_{\mathscr{T}} + \nabla r_{\mathscr{T}}, \mathbf{w})_{\mathbf{L}^2(T)} \\ &\quad + \sum_{S\in\mathscr{S}} (\llbracket \nabla \mathbf{z}_{\mathscr{T}}\cdot\boldsymbol{\nu}\rrbracket, \mathbf{w})_{\mathbf{L}^2(S)} - \sum_{T\in\mathscr{T}} (\operatorname{div}\mathbf{z}_{\mathscr{T}}, s)_{L^2(T)}\end{aligned} \tag{93}$$

for all $(\mathbf{w}, s) \in \mathbf{H}^1_0(\rho^{-1},\Omega) \times L^2(\rho^{-1},\Omega)/\mathbb{R}$.

With all these ingredients at hand, we derive local efficiency properties for the local error indicator $\mathcal{E}_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}})$.

THEOREM 13 (local efficiency of $\mathcal{E}_{ad}$). *Let $(\bar{\mathbf{y}}, \bar{p}, \bar{\mathbf{z}}, \bar{r}, \bar{\mathbf{u}}) \in \mathbf{H}^1_0(\Omega) \times L^2(\Omega)/\mathbb{R} \times \mathbf{H}^1_0(\rho,\Omega) \times L^2(\rho,\Omega)/\mathbb{R} \times \mathbb{U}_{ad}$ be the solution to the optimality system* (28), (35), *and* (36) *and $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ its numerical approximation given by* (42)–(44). *If $\alpha \in (d-2, d)$, then, for $T \in \mathscr{T}$, the local error indicator $\mathcal{E}_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}})$ defined in* (49) *satisfies that*

$$\begin{aligned}&\mathcal{E}^2_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}) \\ &\qquad \lesssim \|\nabla \mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,\mathcal{N}^*_T)} + \|e_r\|^2_{L^2(\rho,\mathcal{N}^*_T)} + \#(T\cap\mathcal{D}) h_T^{\alpha+2-d} \|\mathbf{e}_{\mathbf{y}}\|^2_{\mathbf{L}^\infty(T)},\end{aligned} \tag{94}$$

where $\mathcal{N}^*_T$ is defined as in (9). *The hidden constant is independent of the continuous and discrete solutions, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.*

Proof. We estimate each contribution in (49) separately.

Step 1. Let $T \in \mathscr{T}$. We bound $h_T^2 D_T^\alpha \|\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}\|^2_{\mathbf{L}^2(T)}$ in (49). To accomplish this task, we define $\boldsymbol{\Psi}_T := \psi_T(\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}})$ and use (90) to obtain that

$$\|\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}\|^2_{\mathbf{L}^2(T)} \lesssim \int_{R_T} |\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}|^2 \psi_T \lesssim (\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}, \boldsymbol{\Psi}_T)_{\mathbf{L}^2(T)}; \tag{95}$$

ψ_T denotes the bubble function that satisfies (89). Now, notice that, for $t \in \mathcal{D}$, we have that $\boldsymbol{\Psi}_T(t) = \psi_T(t)(\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}})(t) = 0$. Thus, by setting $(\mathbf{w}, s) = (\boldsymbol{\Psi}_T, 0)$ in

(93) we obtain that

$$
\begin{aligned}
(\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}, \boldsymbol{\Psi}_T)_{\mathbf{L}^2(T)} &= \langle \mathcal{R}, (\boldsymbol{\Psi}_T, 0)\rangle_{\mathcal{Z}',\mathcal{Z}} = a(\mathbf{e}_{\mathbf{z}}, \boldsymbol{\Psi}_T) - b(\boldsymbol{\Psi}_T, e_r) \\
&\lesssim \left(\|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,T)} + \|e_r\|^2_{L^2(\rho,T)}\right)^{\frac{1}{2}} \|\nabla\boldsymbol{\Psi}_T\|_{\mathbf{L}^2(\rho^{-1},T)}.
\end{aligned} \tag{96}
$$

In view of [1, Lemma 5.2] we conclude that $\|\nabla\boldsymbol{\Psi}_T\|_{\mathbf{L}^2(\rho^{-1},T)} \lesssim h_T^{-1} D_T^{-\frac{\alpha}{2}} \|\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}\|_{\mathbf{L}^2(T)}$; recall that $\boldsymbol{\Psi}_T := \psi_T(\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}})$. Replacing this estimate into (96), and the obtained one in (95), we conclude that

$$
h_T^2 D_T^{\alpha} \|\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla\bar{r}_{\mathscr{T}}\|^2_{\mathbf{L}^2(T)} \lesssim \|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,T)} + \|e_r\|^2_{L^2(\rho,T)}. \tag{97}
$$

Step 2. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We bound $h_T D_T^{\alpha} \|[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(\partial T\setminus\partial\Omega)}$ in (49). To accomplish this task, we first define $\boldsymbol{\Psi}_S := \psi_S[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]$. The use of (91) yields

$$
\|[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(S)} \lesssim \int_{R_S} |[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]|^2 \psi_S = ([\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], \boldsymbol{\Psi}_S)_{\mathbf{L}^2(S)}. \tag{98}
$$

We now set $(\mathbf{w}, r) = (\boldsymbol{\Psi}_S, 0)$ in (93) and recall that $\psi_S(t) = 0$ for every $t \in \mathcal{D}$ and that $R_S \subset T_* \cup T_*' \subset \mathcal{N}_S$, where $R_S = \operatorname{supp}(\psi_S)$. This yields

$$
\begin{aligned}
([\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], \boldsymbol{\Psi}_S)_{\mathbf{L}^2(S)} &= \sum_{T'\in\mathcal{N}_S} (\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla r_{\mathscr{T}}, \boldsymbol{\Psi}_S)_{\mathbf{L}^2(T')} - \langle \mathcal{R}, (\boldsymbol{\Psi}_S, 0)\rangle_{\mathcal{Z}',\mathcal{Z}} \\
&= \sum_{T'\in\mathcal{N}_S} (\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla r_{\mathscr{T}}, \boldsymbol{\Psi}_S)_{\mathbf{L}^2(T')} - a(\mathbf{e}_{\mathbf{z}}, \boldsymbol{\Psi}_S) + b(\boldsymbol{\Psi}_S, e_r) \\
&\lesssim \sum_{T'\in\mathcal{N}_S} \|\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla r_{\mathscr{T}}\|_{\mathbf{L}^2(T')} \|\boldsymbol{\Psi}_S\|_{\mathbf{L}^2(T')} \\
&\quad + \sum_{T'\in\mathcal{N}_S} \left(\|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,T')} + \|e_r\|^2_{L^2(\rho,T')}\right)^{\frac{1}{2}} \|\nabla\boldsymbol{\Psi}_S\|_{\mathbf{L}^2(\rho^{-1},T')}.
\end{aligned}
$$

We use that $\|\boldsymbol{\Psi}_S\|_{\mathbf{L}^2(T')} \approx |T'|^{\frac{1}{2}} |S|^{-\frac{1}{2}} \|\boldsymbol{\Psi}_S\|_{\mathbf{L}^2(S)}$ and apply [1, Lemma 5.2] to conclude

$$
\begin{aligned}
([\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], \boldsymbol{\Psi}_S)_{\mathbf{L}^2(S)} &\lesssim \sum_{T'\in\mathcal{N}_S} \|\Delta\bar{\mathbf{z}}_{\mathscr{T}} + \nabla r_{\mathscr{T}}\|_{\mathbf{L}^2(T')} |T'|^{\frac{1}{2}} |S|^{-\frac{1}{2}} \|\boldsymbol{\Psi}_S\|_{\mathbf{L}^2(S)} \\
&\quad + \sum_{T'\in\mathcal{N}_S} \left(\|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,T')} + \|e_r\|^2_{L^2(\rho,T')}\right)^{\frac{1}{2}} D_{T'}^{-\frac{\alpha}{2}} h_{T'}^{-\frac{1}{2}} \|\boldsymbol{\Psi}_S\|_{\mathbf{L}^2(S)}.
\end{aligned} \tag{99}
$$

We thus replace (99) into (98) to arrive at

$$
h_T D_T^{\alpha} \|[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|^2_{\mathbf{L}^2(S)} \lesssim \sum_{T'\subset\mathcal{N}_S} \left(\|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,T')} + \|e_r\|^2_{L^2(\rho,T')}\right), \tag{100}
$$

where we have also used that $|T|/|S| \approx h_T$.

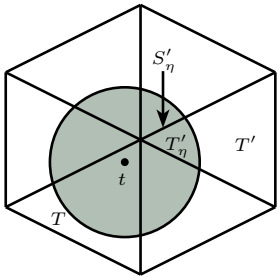

FIG. 2. *Support R_η of the function η (shaded area) on the patch $\mathcal{N}_T^*$.*

Step 3. Let $T \in \mathscr{T}$. We bound the term $\|\text{div } \bar{\mathbf{z}}_{\mathscr{T}}\|^2_{L^2(\rho,T)}$ in (49). Since, in view of (35), div $\bar{\mathbf{z}} = 0$, we immediately conclude that

$$\|\text{div } \bar{\mathbf{z}}_{\mathscr{T}}\|^2_{L^2(\rho,T)} = \|\text{div } \mathbf{e}_{\mathbf{z}}\|^2_{L^2(\rho,T)} \lesssim \|\nabla \mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,T)}. \tag{101}$$

Step 4. Let $T \in \mathscr{T}$ and $t \in \mathcal{D}$. In this step we estimate the term $h_T^{\alpha+2-d}|\bar{\mathbf{y}}_{\mathscr{T}}(t) - \mathbf{y}_t|^2\chi(\{t \in T\})$ in (49). We begin by noticing that, if $T \cap \{t\} = \emptyset$, then the estimate (94) follows directly from the previous three steps. If, instead, $T \cap \{t\} = \{t\}$, then the element indicator $\mathcal{E}_{ad,T}$ defined in (49) contains the term $h_T^{\alpha+2-d}|\bar{\mathbf{y}}_{\mathscr{T}}(t) - \mathbf{y}_t|^2\chi(\{t \in T\})$. If this is the case, a simple application of the triangle inequality yields

$$h_T^{\alpha+2-d}|\bar{\mathbf{y}}_{\mathscr{T}}(t) - \mathbf{y}_t|^2 \lesssim h_T^{\alpha+2-d}|\mathbf{e}_{\mathbf{y}}(t)|^2 + h_T^{\alpha+2-d}|\bar{\mathbf{y}}(t) - \mathbf{y}_t|^2. \tag{102}$$

The term $h_T^{\alpha+2-d}|\mathbf{e}_{\mathbf{y}}(t)|^2$ is trivially bounded by $h_T^{\alpha+2-d}\|\mathbf{e}_{\mathbf{y}}\|^2_{\mathbf{L}^\infty(T)}$. To control the second term on the right-hand side of (102), we follow the ideas developed in the proof of [1, Theorem 5.3] that yield the existence of a smooth function η such that

$$\eta(t) = 1, \quad \|\eta\|_{L^\infty(\Omega)} = 1, \quad \|\nabla\eta\|_{L^\infty(\Omega)} = h_T^{-1}, \quad R_\eta := \text{supp}(\eta) \subset \mathcal{N}_T^*. \tag{103}$$

We now define, given $T' \in \mathcal{N}_T^*$ and $S' \in \mathscr{S}_{T'}$, $T'_\eta := R_\eta \cap T'$ and $S'_\eta := R_\eta \cap S'$; see Figure 2. We also define $\mathbf{w}_\eta := (\bar{\mathbf{y}}(t) - \mathbf{y}_t)\eta \in \mathbf{H}^1_0(\rho^{-1}, \Omega)$. Since the pair $(\bar{\mathbf{z}}, \bar{r})$ solves (35) with $\mathbf{y} = \bar{\mathbf{y}}$, we thus have $|\bar{\mathbf{y}}(t) - \mathbf{y}_t|^2 = \langle(\bar{\mathbf{y}}(t) - \mathbf{y}_t)\delta_t, \mathbf{w}_\eta\rangle = a(\bar{\mathbf{z}}, \mathbf{w}_\eta) - b(\mathbf{w}_\eta, \bar{r}) = a(\mathbf{e}_{\mathbf{z}}, \mathbf{w}_\eta) - b(\mathbf{w}_\eta, e_r) + a(\bar{\mathbf{z}}_{\mathscr{T}}, \mathbf{w}_\eta) - b(\mathbf{w}_\eta, \bar{r}_{\mathscr{T}})$ and

$$\begin{aligned}
|\bar{\mathbf{y}}(t) - \mathbf{y}_t|^2 \lesssim{}& \left(\|\nabla \mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,R_\eta)} + \|e_r\|^2_{L^2(\rho,R_\eta)}\right)^{\frac{1}{2}} \|\nabla \mathbf{w}_\eta\|_{\mathbf{L}^2(\rho^{-1},R_\eta)} \\
&+ \sum_{T' \in \mathcal{N}_T^* : T'_\eta \subset R_\eta} \|\Delta \bar{\mathbf{z}}_{\mathscr{T}} + \nabla \bar{r}_{\mathscr{T}}\|_{\mathbf{L}^2(T'_\eta)} \|\mathbf{w}_\eta\|_{\mathbf{L}^2(T'_\eta)} \\
&+ \sum_{T' \in \mathcal{N}_T^* : T'_\eta \subset R_\eta} \sum_{S'_\eta \subset \partial T'_\eta : S'_\eta \not\subset \partial R_\eta} \|[\![\nabla \bar{\mathbf{z}}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{\mathbf{L}^2(S'_\eta)} \|\mathbf{w}_\eta\|_{L^2(S'_\eta)}.
\end{aligned}$$

Finally, the regularity of the mesh, in conjunction with the fact that, since $t \in T$, $h_T \approx D_T$, and the estimates $\|\nabla\eta\|_{L^2(\rho^{-1},R_\eta)} \lesssim h_T^{(d-2)/2-\alpha/2}$, $\|\eta\|_{L^2(R_\eta)} \lesssim h_T^{d/2}$, and

$\|\eta\|_{L^2(S'_\eta)} \lesssim h_T^{(d-1)/2}$, allow us to conclude that

(104)
$$
\begin{aligned}
&|\bar{\mathbf{y}}(t)-\mathbf{y}_t|^2\\
&\quad\lesssim h_T^{\frac{d-2}{2}-\frac{\alpha}{2}}|\bar{\mathbf{y}}(t)-\mathbf{y}_t|\left(\|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,R_\eta)}+\|e_r\|^2_{L^2(\rho,R_\eta)}\right)^{\frac{1}{2}}\\
&\qquad+h_T^{\frac{d-2}{2}-\frac{\alpha}{2}}|\bar{\mathbf{y}}(t)-\mathbf{y}_t|\left(\sum_{T'\in\mathcal{N}_T^*:T'_\eta\subset R_\eta} h_T D_T^{\frac{\alpha}{2}}\|\Delta\bar{\mathbf{z}}_{\mathscr{T}}+\nabla\bar{r}_{\mathscr{T}}\|_{\mathbf{L}^2(T'_\eta)}\right.\\
&\qquad\qquad\left.+\sum_{T'\in\mathcal{N}_T^*:T'_\eta\subset R_\eta}\ \sum_{S'_\eta\subset\partial T'_\eta:S'_\eta\not\subset\partial R_\eta} h_T^{\frac{1}{2}}D_T^{\frac{\alpha}{2}}\|[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^2(S'_\eta)}\right).
\end{aligned}
$$

Notice that $\|\Delta\bar{\mathbf{z}}_{\mathscr{T}}+\nabla\bar{r}_{\mathscr{T}}\|_{\mathbf{L}^2(T'_\eta)} \lesssim \|\Delta\bar{\mathbf{z}}_{\mathscr{T}}+\nabla\bar{r}_{\mathscr{T}}\|_{\mathbf{L}^2(T')}$ and $\|[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^2(S'_\eta)} \lesssim \|[\![\nabla\bar{\mathbf{z}}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{\mathbf{L}^2(S')}$. All these ingredients yield an estimate for $h_T^{\alpha+2-d}|\bar{\mathbf{y}}_{\mathscr{T}}(t)-\mathbf{y}_t|^2\chi(\{t\in T\})$.

A collection of the estimates (97), (100), (101), and (104) yield the desired result. ∎

We conclude with the global efficiency of the error estimator $\mathcal{E}_{ocp}$.

THEOREM 14 (global efficiency property of $\mathcal{E}_{ocp}$). *Let* $(\bar{\mathbf{y}},\bar{p},\bar{\mathbf{z}},\bar{r},\bar{\mathbf{u}})\in\mathbf{H}_0^1(\Omega)\times L^2(\Omega)/\mathbb{R}\times\mathbf{H}_0^1(\rho,\Omega)\times L^2(\rho,\Omega)/\mathbb{R}\times\mathbb{U}_{ad}$ *be the solution to the optimality system* (28), (35), *and* (36) *and* $(\bar{\mathbf{y}}_{\mathscr{T}},\bar{p}_{\mathscr{T}},\bar{\mathbf{z}}_{\mathscr{T}},\bar{r}_{\mathscr{T}},\bar{\mathbf{u}}_{\mathscr{T}})\in\mathbf{V}(\mathscr{T})\times Q(\mathscr{T})\times\mathbf{V}(\mathscr{T})\times Q(\mathscr{T})\times\mathbb{U}_{ad}(\mathscr{T})$ *its numerical approximation given by* (42)–(44). *If* $\Omega\subset\mathbb{R}^2$ *and* $\alpha\in(d-2,2)$, *then*

$$
\mathcal{E}^2_{ocp}(\bar{\mathbf{z}}_{\mathscr{T}},\bar{r}_{\mathscr{T}},\bar{\mathbf{y}}_{\mathscr{T}},\bar{p}_{\mathscr{T}},\bar{\mathbf{u}}_{\mathscr{T}})\lesssim\|\mathbf{e}\|^2_\Omega, \tag{105}
$$

where the hidden constant is independent of the size of the elements in the mesh $\mathscr{T}$ *and* $\#\mathscr{T}$ *but depends linearly on* $\#\mathcal{D}$ *and* $\operatorname{diam}(\Omega)^{\alpha+2-d}$.

Proof. We invoke the local efficiency estimates (81) and (87) to arrive at

$$
\mathcal{E}^2_{st}(\bar{\mathbf{y}}_{\mathscr{T}},\bar{p}_{\mathscr{T}},\bar{\mathbf{u}}_{\mathscr{T}})\lesssim\|\mathbf{e}_{\mathbf{y}}\|^2_{\mathbf{L}^\infty(\Omega)}+\operatorname{diam}(\Omega)^{2-d}\|e_p\|^2_{L^2(\Omega)}+\operatorname{diam}(\Omega)^{4-d}\|\mathbf{e}_{\mathbf{u}}\|^2_{\mathbf{L}^2(\Omega)}
$$

and $E^2_{st}(\bar{\mathbf{y}}_{\mathscr{T}},\bar{p}_{\mathscr{T}},\bar{\mathbf{u}}_{\mathscr{T}})\lesssim\operatorname{diam}(\Omega)^{d-2}\|\mathbf{e}_{\mathbf{y}}\|^2_{\mathbf{L}^\infty(\Omega)}+\|e_p\|^2_{L^2(\Omega)}+\operatorname{diam}(\Omega)^2\|\mathbf{e}_{\mathbf{u}}\|^2_{\mathbf{L}^2(\Omega)}$, respectively. On the other hand, in view of (52), the efficiency estimate (94) implies

$$
\mathcal{E}^2_{ad}(\bar{\mathbf{z}}_{\mathscr{T}},\bar{r}_{\mathscr{T}},\bar{\mathbf{y}}_{\mathscr{T}})\lesssim\|\nabla\mathbf{e}_{\mathbf{z}}\|^2_{\mathbf{L}^2(\rho,\Omega)}+\|e_r\|^2_{L^2(\rho,\Omega)}+\left(\sum_{T\in\mathscr{T}:T\cap\mathcal{D}\neq\emptyset}h_T^{\alpha+2-d}\right)\|\mathbf{e}_{\mathbf{y}}\|^2_{\mathbf{L}^\infty(\Omega)}.
$$

Now, since $\alpha\in(d-2,2)$ and $\#\mathcal{D}<\infty$, we can conclude that $\sum_{T\in\mathscr{T}:T\cap\mathcal{D}\neq\emptyset}h_T^{\alpha+2-d}\leq\#\mathcal{D}\operatorname{diam}(\Omega)^{\alpha+2-d}$. We notice that this estimate, which is where the linear dependence on $\#\mathcal{D}$ and $\operatorname{diam}(\Omega)^{\alpha+2-d}$ comes from, is independent of $\#\mathscr{T}$

Finally, an application of the triangle inequality yields

$$
\mathcal{E}_{ct}(\bar{\mathbf{z}}_{\mathscr{T}},\bar{\mathbf{u}}_{\mathscr{T}})\leq\|\bar{\mathbf{u}}_{\mathscr{T}}-\Pi(-\lambda^{-1}\bar{\mathbf{z}})\|_{\mathbf{L}^2(\Omega)}+\|\Pi(-\lambda^{-1}\bar{\mathbf{z}})-\Pi(-\lambda^{-1}\bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\Omega)}.
$$

This, in conjunction with the Lipschitz continuity of the projection operator Π and Theorem 6, implies that $\mathcal{E}_{ct}(\bar{\mathbf{z}}_{\mathscr{T}},\bar{\mathbf{u}}_{\mathscr{T}})\lesssim\|\bar{\mathbf{u}}_{\mathscr{T}}-\bar{\mathbf{u}}\|_{\mathbf{L}^2(\Omega)}+\lambda^{-1}\|\nabla(\bar{\mathbf{z}}-\bar{\mathbf{z}}_{\mathscr{T}})\|_{\mathbf{L}^2(\rho,\Omega)}$. The proof concludes by gathering all the obtained estimates. ∎

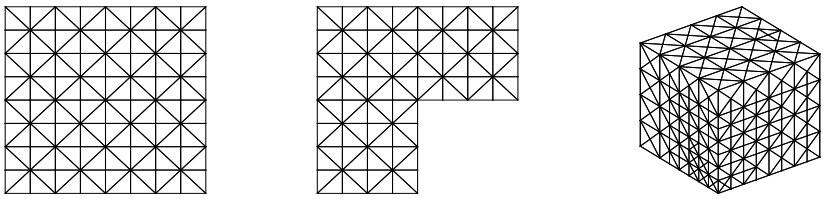

FIG. 3. *The initial meshes used when the domain Ω is a square (Examples 1–2), two-dimensional L-shape (Example 3), and cube (Examples 4–5).*

6. Numerical examples. In this section we conduct a series of numerical examples that illustrate the performance of the devised a posteriori error estimator. These experiments have been carried out with the help of a code that we implemented using C++. All matrices have been assembled exactly. The right-hand sides as well as the approximation errors are computed by a quadrature formula that is exact for polynomials of degree 19 for two-dimensional domains and degree 14 for three-dimensional domains. The global linear systems were solved using the multifrontal massively parallel sparse direct solver (MUMPS) [6, 7].

For a given partition $\mathscr{T}$ we seek $(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \in \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbf{V}(\mathscr{T}) \times Q(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ that solves the discrete optimality system (42)–(44). The underlying nonlinear system is solved by using the Newton-type primal-dual active set strategy of [30, section 2.12.4]. Once the discrete solution is obtained, we use the local error indicator $\mathcal{E}_{ocp,T}$, defined as

$$\begin{aligned}\mathcal{E}^2_{ocp,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) &:= E^2_{st}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) \\ &+ \mathcal{E}^2_{st,T}(\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}) + \mathcal{E}^2_{ad,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}}) + \mathcal{E}^2_{ct,T}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}),\end{aligned} \tag{106}$$

to drive the adaptive procedure described in Algorithm 1 and compute the global error estimator $\mathcal{E}_{ocp}$ in order to assess the accuracy of the approximation. A sequence of adaptively refined meshes is thus generated from the initial meshes shown in Figure 3. The total number of degrees of freedom is $\mathsf{Ndof} = 2\dim(\mathbf{V}(\mathscr{T})) + 2\dim(Q(\mathscr{T})) + \dim(\mathbf{U}(\mathscr{T}))$. The error is measured in the norm $\|\mathbf{e}\|_\Omega$, which is defined in (62). Finally, we define the effectivity index $\Upsilon := \mathcal{E}_{ocp}/\|\mathbf{e}\|_\Omega$.

We consider problems with homogeneous boundary conditions whose exact solutions are not known. We also consider problems with inhomogeneous Dirichlet boundary conditions whose exact solutions are known. Notice that this violates the assumption of homogeneous Dirichlet boundary conditions which is needed for the analysis that we have performed. In this case, we write the optimal adjoint pair $(\bar{\mathbf{z}}, \bar{r})$ in terms of fundamental solutions of the Stokes equations [18, section IV.2]:

$$\bar{\mathbf{z}}(\mathbf{x}) := \sum_{t\in\mathcal{D}} \vartheta_t \sum_{i=1}^{d} \widetilde{\mathbf{T}}_t(\mathbf{x}) \cdot \mathbf{e}_i + \mathfrak{C}, \qquad \bar{r}(\mathbf{x}) := \sum_{t\in\mathcal{D}} \vartheta_t \sum_{i=1}^{d} \mathbf{T}_t(\mathbf{x}) \cdot \mathbf{e}_i, \tag{107}$$

where $\mathfrak{C} \in \mathbb{R}^d$ and, if $\mathbf{r}_t = \mathbf{x} - t$ and $\mathbb{I}_d$ is the identity matrix in $\mathbb{R}^{d\times d}$, then

$$\widetilde{\mathbf{T}}_t(\mathbf{x}) = \begin{cases} -\dfrac{1}{4\pi}\left(\log|\mathbf{r}_t|\mathbb{I}_2 - \dfrac{\mathbf{r}_t\mathbf{r}_t^T}{|\mathbf{r}_t|^2}\right) & \text{if } d=2, \\ \dfrac{1}{8\pi}\left(\dfrac{1}{|\mathbf{r}_t|}\mathbb{I}_3 + \dfrac{\mathbf{r}_t\mathbf{r}_t^T}{|\mathbf{r}_t|^3}\right) & \text{if } d=3; \end{cases} \qquad \mathbf{T}_t(\mathbf{x}) = \begin{cases} -\dfrac{\mathbf{r}_t}{2\pi|\mathbf{r}_t|^2} & \text{if } d=2, \\ -\dfrac{\mathbf{r}_t}{4\pi|\mathbf{r}_t|^3} & \text{if } d=3; \end{cases}$$

Algorithm 1 Adaptive Algorithm.

Input: Initial mesh $\mathscr{T}_0$, set of observable points $\mathcal{D}$, set of desired states $\{\mathbf{y}_t\}_{t\in\mathcal{D}}$, Muckenhoupt parameter α, vector constraints $\mathbf{a}$ and $\mathbf{b}$, and regularization parameter λ.
Set: $i = 0$.
Active set strategy:
1 : Choose initial discrete guesses $\mathbf{u}^0_{\mathscr{T}}, \boldsymbol{\mu}^0_{\mathscr{T}} \in \mathbf{U}(\mathscr{T}_i)$ ($\mathbf{u}^0_{\mathscr{T}}$ is not necessarily admissible).
2 : Compute $[\bar{\mathbf{y}}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{u}}_{\mathscr{T}}] = \textbf{Active-Set}[\mathscr{T}_i, \mathbf{u}^0_{\mathscr{T}}, \boldsymbol{\mu}^0_{\mathscr{T}}, \lambda, \alpha, \mathbf{a}, \mathbf{b}, \mathcal{D}, \{\mathbf{y}_t\}_{t\in\mathcal{D}}]$.
Active-Set implements the active set strategy of [30, section 2.12.4]
Adaptive loop:
3 : For each $T \in \mathscr{T}_i$ compute the local error indicator $\mathcal{E}_{ocp,T}$ defined in (106).
4 : Mark an element T for refinement if $\mathcal{E}^2_{ocp,T} > \frac{1}{2}\max_{T'\in\mathscr{T}_i} \mathcal{E}^2_{ocp,T'}$.
5 : From step 4, construct a new mesh, using a longest edge bisection algorithm. Set $i \leftarrow i+1$ and go to step 1.

$\{\mathbf{e}_i\}_{i=1}^d$ denotes the canonical basis of $\mathbb{R}^d$ and $\vartheta_t \in \mathbb{R}$ for all $t \in \mathcal{D}$. We write the optimal velocity field $\bar{\mathbf{y}}$ in terms of the curl operator as described in [19, section 2.3]. The sequence of vectors $\{\mathbf{y}_t\}_{t\in\mathcal{D}}$ is computed from the constructed solutions in such a way that the adjoint equations (35) hold. A straightforward computation reveals that, for $t \in \mathcal{D}$, $\mathbf{y}_t = \bar{\mathbf{y}}(t) - \vartheta_t(\mathbf{e}_1 + \cdots + \mathbf{e}_d)$. We finally mention that in order to simplify the construction of exact solutions, we have incorporated, in the *momentum equation* of (28), an extra forcing term $\mathbf{f} \in \mathbf{L}^\infty(\Omega)$. With such a modification, the right-hand side of the *momentum equation* reads as follows: $(\mathbf{f} + \mathbf{u}, \mathbf{v})_{\mathbf{L}^2(\Omega)}$.

In the examples that we will present we have considered, as initial guesses in Algorithm 1, $\mathbf{u}^0_{\mathscr{T}} = \boldsymbol{\mu}^0_{\mathscr{T}} = \mathbf{0}$, where $\mathbf{0}$ denotes the null vector in $\mathbb{R}^d$.

6.1. Two-dimensional examples. We perform two-dimensional examples on convex and nonconvex domains, and with a different number of source points. The first two examples involve homogeneous Dirichlet boundary conditions in the state equations, but inhomogeneous Dirichlet boundary conditions in the adjoint equations. In the third example we consider homogeneous Dirichlet boundary conditions in both equations.

Example 1. We let $\Omega = (0,1)^2$, $\lambda = 1$, $\mathbf{a} = (-0.7,-0.7)^T$, $\mathbf{b} = (-0.3,-0.3)^T$, and $\mathcal{D} = \{(0.25,0.25),(0.25,0.75),(0.75,0.25),(0.75,0.75)\}$. The optimal state is $\bar{\mathbf{y}}(x_1,x_2) = 0.5\mathbf{curl}\left[(x_1x_2(1-x_1)(1-x_2))^2\right]$, $\bar{p}(x_1,x_2) = x_1x_2(1-x_1)(1-x_2) - \frac{1}{36}$, while the exact optimal adjoint state is taken to be as in (107) with $\vartheta_t = 1$, for all $t \in \mathcal{D}$ and $\mathfrak{C} = \mathbf{0}$.

In this example we investigate the effect of varying the exponent α in the Muckenhoupt weight ρ. We consider $\alpha \in \{0.5, 0.8, 1.2, 1.6, 1.9\}$.

Example 2. We let $\Omega = (0,1)^2$, $\mathbf{a} = (-120,-120)^T$, $\mathbf{b} = (-1,-1)^T$, $\mathcal{D} = \{(0.25,0.5),(0.5,0.5),(0.75,0.5)\}$, and $\alpha = 1.5$. The optimal state is given by

$$\bar{\mathbf{y}}(x_1,x_2) = \frac{5}{2\pi}\mathbf{curl}\left[(\sin(2\pi x_1)\sin(2\pi x_2))^2\right],$$
$$\bar{p}(x_1,x_2) = \frac{25}{2}\left(x_1 - 1 + \frac{(e^{-x_1}-1)}{(e^{-1}-1)}\right)\left(x_2 - 1 + \frac{(e^{-x_2}-1)}{(e^{-1}-1)}\right) - \frac{25}{2}\left(\frac{e-3}{e-1}\right)^2,$$

while the exact optimal adjoint state is taken to be as in (107) with $\vartheta_t = 3$ for all $t \in \mathcal{D}$ and $\mathfrak{C} = (-1/4,-1/4)$.

In this example we investigate the effect of varying the regularization paramater λ. We consider $\lambda \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$.

Example 3. We let $\Omega = (0,1)^2 \setminus [0.5,1) \times (0,0.5]$, $\mathbf{a} = (-0.3,-0.3)^T$, $\mathbf{b} = (0.4,0.4)^T$, $\lambda = 1$, $\mathcal{D} = \{(0.25,0.25),(0.25,0.75),(0.75,0.75)\}$, $\mathbf{y}_{(0.25,0.25)} = (3,3)^T$, $\mathbf{y}_{(0.25,0.75)} = (-1,-1)^T$, and $\mathbf{y}_{(0.75,0.75)} = (3,3)^T$.

In Figures 4 and 5 we present the results obtained for Example 1. In Figure 4 we present experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ together with their individual contributions for uniform and adaptive refinement ($\alpha = 1.9$). We observe that the devised adaptive loop outperforms uniform refinement. We also observe optimal experimental rates of convergence for all the individual contributions of $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ but with the exception of the ones associated to the optimal control $\bar{\mathbf{u}}$, i.e., $\mathcal{E}_{ct}$ and $\|\mathbf{e_u}\|_{\mathbf{L}^2(\Omega)}$. In Figure 5 we present, for different values of the exponent $\alpha \in \{0.5, 0.8, 1.2, 1.6\}$, experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$, effectivity indices Υ, and adaptively refined meshes. We observe optimal experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ in subfigures (A.1)–(D.1). In subfigures (A.3)–(D.3), the effect of varying the parameter α on the adaptively refined meshes can be appreciated. Notice that this behavior is consistent with (49) and (52). We also observe that the effectivity indices decrease as the exponent α increases; see subfigures (A.2)–(D.2). Notice that the well-posedness of our problem requires $\alpha \in (d-2, d)$ while the application of the results of Theorem 6 requires $\alpha \in (d-2, 2)$. Furthermore, we observe that the values of $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ decrease as the parameter α increases. This suggests that better accuracy properties can be expected for bigger values of α. Finally, all the effectivity indices $\Upsilon = \mathcal{E}_{ocp}/\|\mathbf{e}\|_\Omega$ are stabilized around values between 3 and 8. This shows the accuracy of the proposed a posteriori error estimator $\mathcal{E}_{ocp}$ when used in the adaptive loop described in Algorithm 1.

In Figures 6 we present the results obtained within the setting of Example 2. For $\lambda \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$, we observe optimal experimental rates of convergence for all the individual contributions of $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ but with the exception, again, of $\mathcal{E}_{ct}$ and $\|\mathbf{e_u}\|_{\mathbf{L}^2(\Omega)}$. Notice that, for the considered values of the parameter λ, the dominant contribution does not remain the same. In spite of this fact, optimal experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ are attained. Finally, we comment that the effectivity indices Υ increase as the regularization parameter λ increases. This behavior is in agreement with the results of Theorem 9.

In Figure 7 we present the results obtained for Example 3. We present the finite element solutions obtained for the adjoint and control variables, together with the 40th adaptively refined mesh and experimental rates of convergence for $\mathcal{E}_{ocp}$ and its individual contributions. We also observe that the refinement is being concentrated around the reentrant corner and the source points. Finally, it can be appreciated that an optimal experimental rate of convergence for the total error estimator $\mathcal{E}_{ocp}$ is attained.

Remark 15 (on the influence of α on Ndof). In Figure 5, we present experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$ for different values of the parameter α. Notice that each plot involves a different range of numbers of degrees of freedom. This is due to the fact that the adaptive refinement depends on the value of α: if $\alpha < 1$, the refinement is mostly performed on the elements that are close to the observation points; see Figure 5(A.3)–(D.3). When α is small, after a certain number of adaptive iterations, there are elements $T \in \mathscr{T}$, around the singular points, such that $|T| \approx 10^{-16}$. As a consequence, the assembly calculations reach machine precision numbers

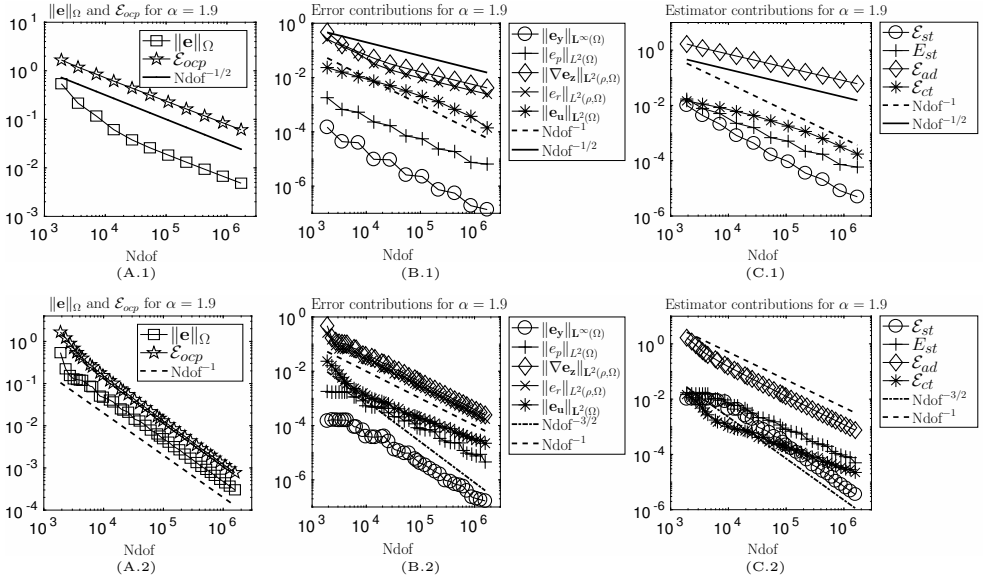

FIG. 4. *Example* 1. *Experimental rates of convergence for the total error and error estimator, as well as their individual contributions, for uniform refinement* (A.1)–(C.1) *and adaptive refinement* (A.2)–(C.2)*, respectively* ($\alpha = 1.9$).

and thus make impossible more computations within the adaptive procedure. This issue has also been observed in [1, section 6].

6.2. Three-dimensional examples. We now present three-dimensional examples with homogeneous and inhomogeneous Dirichlet boundary conditions and a different number of source points.

Example 4. We set $\Omega = (0,1)^3$, $\mathbf{a} = (-2,-2,-2)^T$, $\mathbf{b} = (-0.1,-0.1,-0.1)^T$, $\lambda = 10^{-2}$, $\alpha = 1.99$, and $\mathcal{D} = \{(0.5,0.5,0.5)\}$. The exact optimal state is given by

$$\bar{\mathbf{y}}(x_1,x_2,x_3) = 100\mathbf{curl}((x_1x_2x_3(1-x_1)(1-x_2)(1-x_3)\arctan(x-0.5))^2\mathbf{e}_1),$$
$$\bar{p}(x_1,x_2,x_3) = x_1x_2x_3 - \tfrac{1}{8}.$$

The optimal adjoint state is as in (107) with $\vartheta_t = 1$, for $t \in \mathcal{D}$, and $\mathfrak{C} = (-1,-1,-1)$.

Example 5. We set $\Omega = (0,1)^3$, $\mathbf{a} = (-2,-2,-2)^T$, $\mathbf{b} = (-1,-1,-1)^T$, $\lambda = 1$, $\alpha = 1.99$, and $\mathcal{D} = \{(0.25,0.25,0.25),(0.75,0.25,0.25),(0.25,0.75,0.75),(0.75,0.75,0.75)\}$. The observable points are $\mathbf{y}_{(0.25,0.25,0.25)} = (-5,-5,-5)^T$, $\mathbf{y}_{(0.75,0.25,0.25)} = (1,1,1)^T$, $\mathbf{y}_{(0.25,0.75,0.75)} = (5,5,5)^T$, and $\mathbf{y}_{(0.75,0.75,0.75)} = (-1,-1,-1)^T$.

In Figure 8 we present the results obtained for Example 4. We present experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$, as well as their individual contributions, the effectivity index Υ, and the 40th adaptively refined mesh. We observe optimal experimental rates of convergence for $\mathcal{E}_{ocp}$ and $\|\mathbf{e}\|_\Omega$. The effectivity index Υ is stabilized around the value of 11. This shows the accuracy of the proposed a posteriori error estimator $\mathcal{E}_{ocp}$ when used in the adaptive loop of Algorithm 1 in a three-dimensional setting. We also observe that most of the refinement occurs near the observation point which attests to the efficiency of the devised estimator.

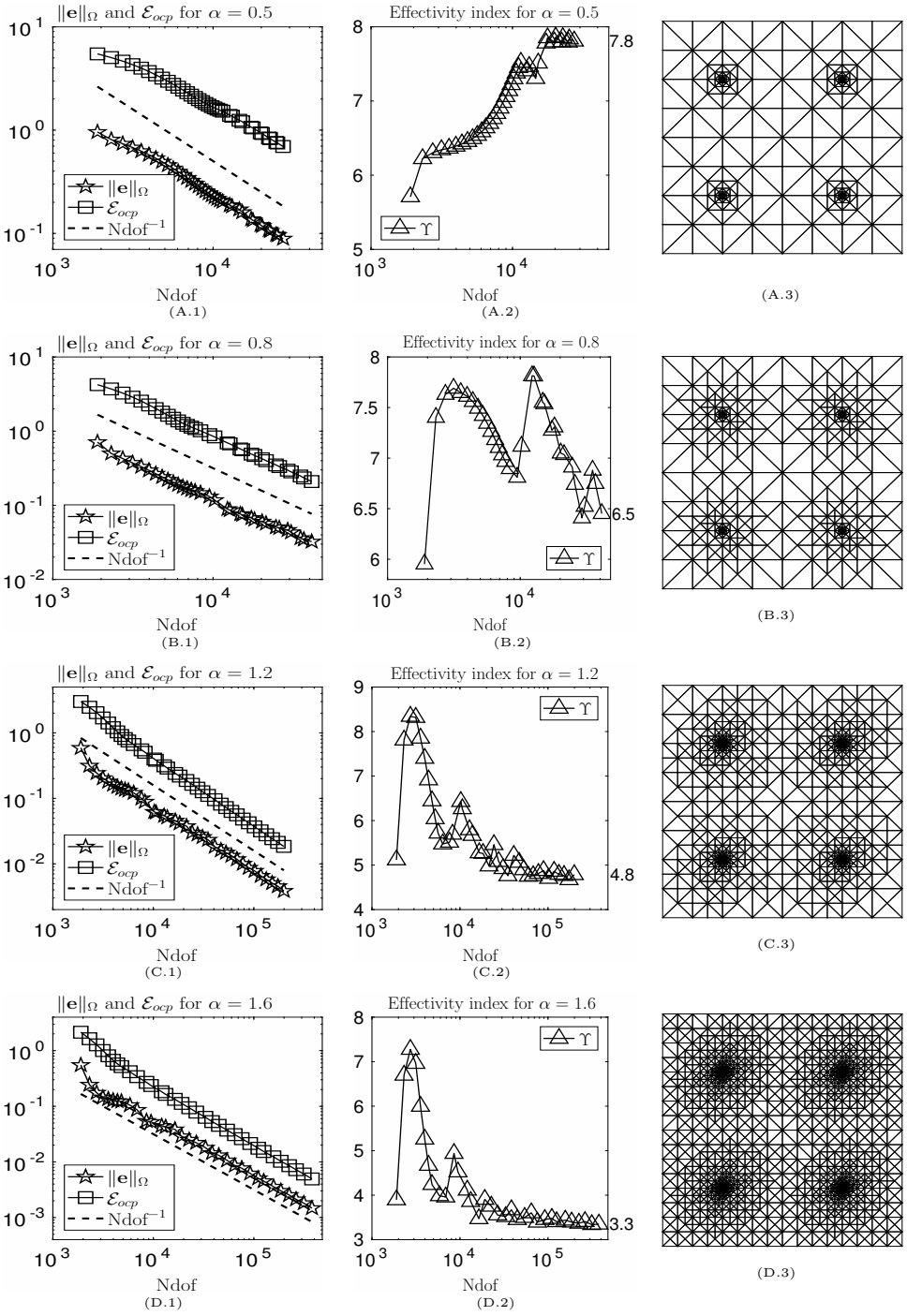

FIG. 5. *Example* 1. *Experimental rates of convergence for the total error* $\|\mathbf{e}\|_\Omega$ *and error estimator* $\mathcal{E}_{ocp}$ (A.1)–(D.1)*, effectivity indices* (A.2)–(D.2) *and the* 20*th adaptively refined mesh* (A.3)–(D.3) *for* $\alpha \in \{0.5, 0.8, 1.2, 1.6\}$ *and* $\lambda = 1$.

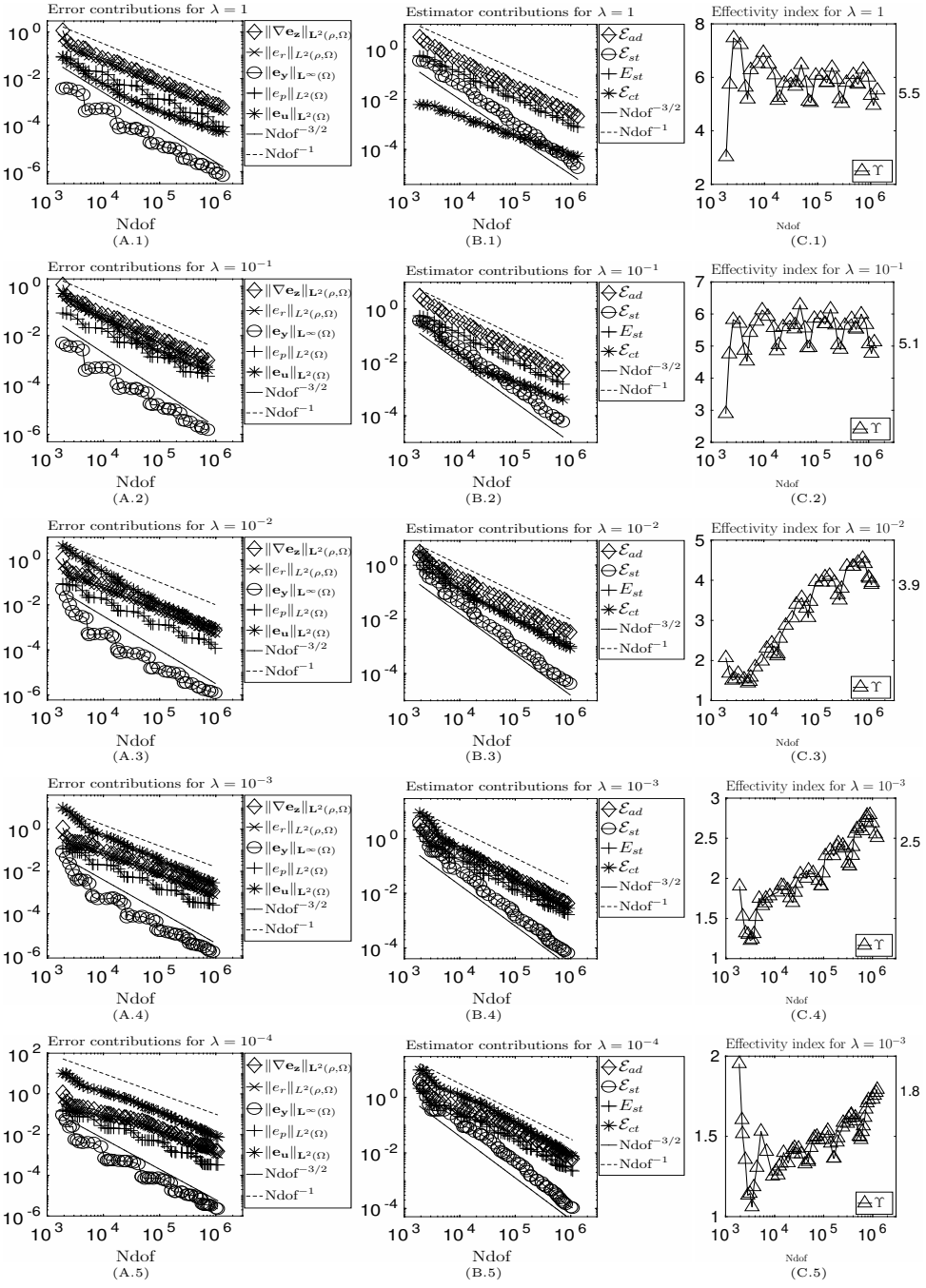

FIG. 6. *Example* 2. *Experimental rates of convergence for each contribution of the total error* $\|\mathbf{e}\|_\Omega$ (A.1)–(A.5) *and error estimator* $\mathcal{E}_{ocp}$ (B.1)–(B.5) *and effectivity indices* (C.1)–(C.5). *We have considered* $\lambda \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ *and* $\alpha = 1.5$.

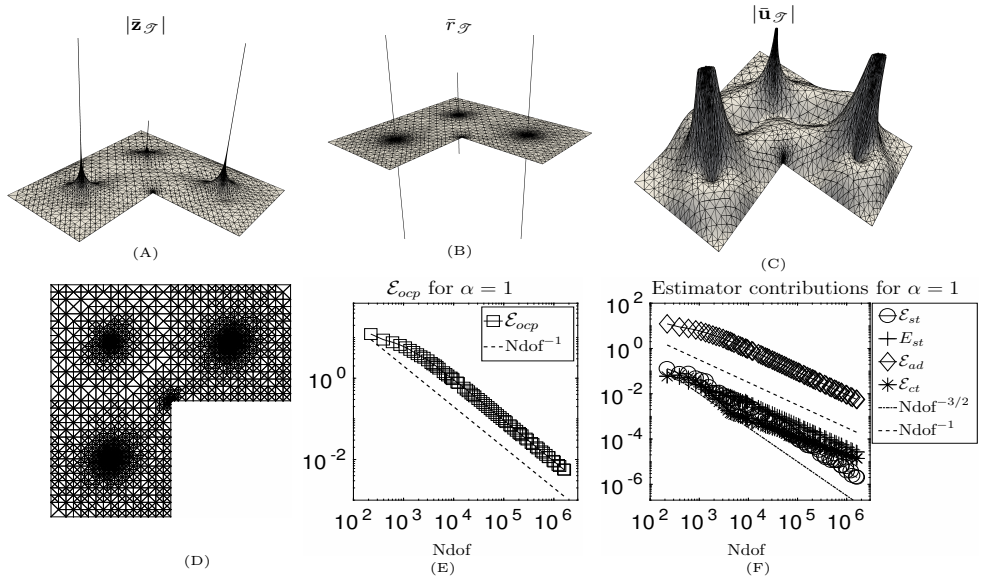

FIG. 7. *Example* 3. *Finite element approximations of* $|\bar{\mathbf{z}}_{\mathscr{T}}|$, $\bar{r}_{\mathscr{T}}$, *and* $|\bar{\mathbf{u}}_{\mathscr{T}}|$ (A)–(C) *obtained on the* 40*th adaptively refined mesh* (D), *and experimental rates of convergence for* $\mathcal{E}_{ocp}$ (E) *and its individual contributions* (F) ($\alpha = 1$).

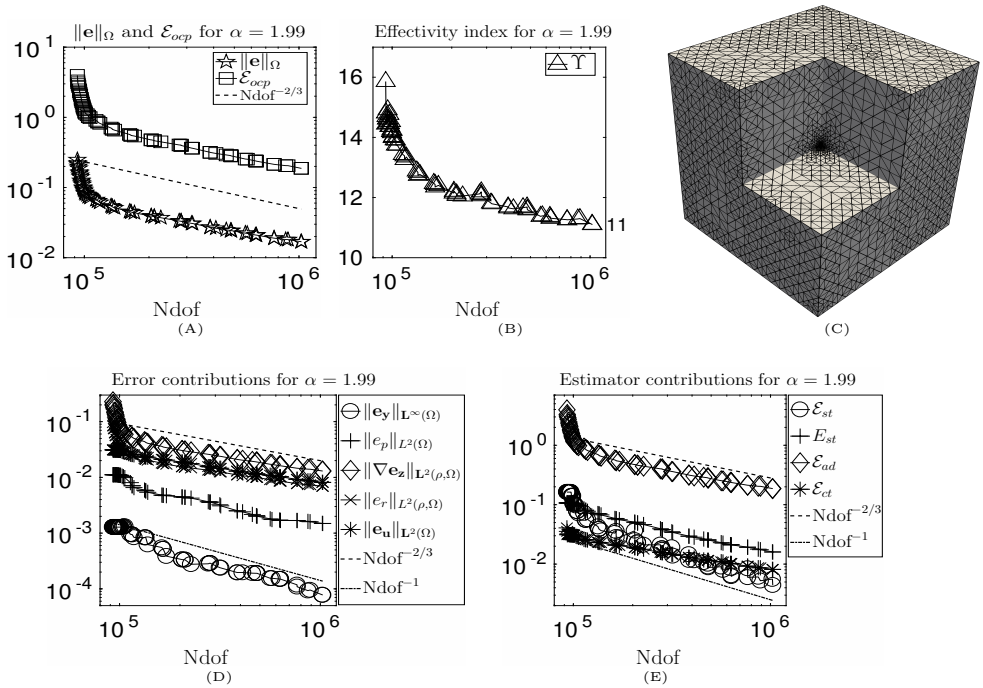

FIG. 8. *Example* 4. *Experimental rates of convergence for the error* $\|\mathbf{e}\|_\Omega$ *and error estimator* $\mathcal{E}_{ocp}$ (A), *the effectivity index* (B), *and the* 40*th adaptively refined mesh* (C). *Experimental rates of convergence for the individual contributions of* $\|\mathbf{e}\|_\Omega$ (D) *and* $\mathcal{E}_{ocp}$ (E) ($\alpha = 1.99$).

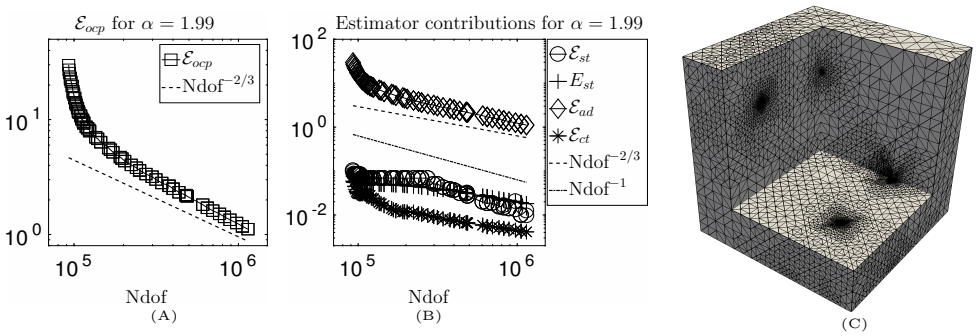

FIG. 9. *Example* 5. *Experimental rates of convergence for the error estimator* $\mathcal{E}_{ocp}$ (A)*, its individual contributions* (B)*, and the* 87*th adaptively refined mesh* (C) *by fixing* $\alpha = 1.99$.

In Figure 9 we present the results obtained for Example 5. We observe an optimal experimental rate of convergence for $\mathcal{E}_{ocp}$ and notice that the refinement occurs around the observation points.

6.3. Conclusions.

- Most of the refinement occurs near the observation points. This attests to the efficiency of the devised estimators. When the domain involves geometric singularities, refinement is also being performed in regions that are close to them.
- A larger value of α delivers the best results. Notice that if $h_T < 1$, the larger the value of α the smaller the value of $h_T^{\alpha+2-d}$.
- We observe that the contributions $\|\mathbf{e_u}\|_{\mathbf{L}^2(\Omega)}$ and $\mathcal{E}_{ct}$ do not exhibit an optimal decay. This might be due to the fact that the contributions of $\mathcal{E}_{ct}$ are not balanced in terms of error approximation, the very singular nature of the problem, and the reduced regularity properties of the optimal control variable.
- The contribution $\mathcal{E}_{ad}(\bar{\mathbf{z}}_{\mathscr{T}}, \bar{r}_{\mathscr{T}}, \bar{\mathbf{y}}_{\mathscr{T}})$ of the global error estimator $\mathcal{E}_{ocp}$ is, most of the time, the dominating one.
- When the regularization parameter λ increases, the effectivity index Υ increases as well. Notice that this is consistent with Theorem 9.
- In spite of the very singular nature of the problem that defines the adjoint variable, our proposed estimator is able to deliver optimal experimental rates of convergence, within an adaptive loop, for the contributions related to the discretization of the state and adjoint equations.

REFERENCES

[1] J. P. AGNELLI, E. M. GARAU, AND P. MORIN, *A posteriori error estimates for elliptic problems with Dirac measure terms in weighted spaces*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 1557–1581, https://doi.org/10.1051/m2an/2014010.

[2] H. AIMAR, M. CARENA, R. DURÁN, AND M. TOSCHI, *Powers of distances to lower dimensional sets as Muckenhoupt weights*, Acta Math. Hungar., 143 (2014), pp. 119–137, https://doi.org/10.1007/s10474-014-0389-1.

[3] M. AINSWORTH AND J. T. ODEN, *A Posteriori Error Estimation in Finite Element Analysis*, Pure Appl. Math. (New York), Wiley-Interscience, New York, 2000, https://doi.org/10.1002/9781118032824.

[4] 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 (2017), pp. 159–197, https://doi.org/10.1007/s00211-017-0867-9.

[5] A. Allendes, E. Otárola, and A. J. Salgado, *A posteriori error estimates for the Stokes problem with singular sources*, Comput. Methods Appl. Mech. Engrg., 345 (2019), pp. 1007–1032, https://doi.org/10.1016/j.cma.2018.11.004.

[6] P. R. Amestoy, I. S. Duff, and J.-Y. L'Excellent, *Multifrontal parallel distributed symmetric and unsymmetric solvers*, Comput. Methods Appl. Mech. Engrg., 184 (2000), pp. 501–520, https://doi.org/10.1016/S0045-7825(99)00242-X.

[7] P. R. Amestoy, I. S. Duff, J.-Y. L'Excellent, and J. Koster, *A fully asynchronous multifrontal solver using distributed dynamic scheduling*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 15–41, https://doi.org/10.1137/S0895479899358194.

[8] H. Antil, E. Otárola, and A. J. Salgado, *Some applications of weighted norm inequalities to the error analysis of PDE-constrained optimization problems*, IMA J. Numer. Anal., 38 (2018), pp. 852–883, https://doi.org/10.1093/imanum/drx018.

[9] N. Behringer, D. Meidner, and B. Vexler, *Finite Element Error Estimates for Optimal Control Problems with Pointwise Tracking*, preprint, arXiv:1802.02918, 2018.

[10] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008, https://doi.org/10.1007/978-0-387-75934-0.

[11] C. Brett, A. Dedner, and C. Elliott, *Optimal control of elliptic PDEs at points*, IMA J. Numer. Anal., 36 (2016), pp. 1015–1050, https://doi.org/10.1093/imanum/drv040.

[12] R. M. Brown and Z. Shen, *Estimates for the Stokes operator in Lipschitz domains*, Indiana Univ. Math. J., 44 (1995), pp. 1183–1206, https://doi.org/10.1512/iumj.1995.44.2025.

[13] L. Chang, W. Gong, and N. Yan, *Numerical analysis for the approximation of optimal control problems with pointwise observations*, Math. Methods Appl. Sci., 38 (2015), pp. 4502–4520, https://doi.org/10.1002/mma.2861.

[14] A. Demlow and S. Larsson, *Local pointwise a posteriori gradient error bounds for the Stokes equations*, Math. Comp., 82 (2013), pp. 625–649, https://doi.org/10.1090/S0025-5718-2012-02647-0.

[15] J. Duoandikoetxea, *Fourier Analysis*, Grad. Stud. Math. 29, American Mathematical Society, Providence, RI, 2001.

[16] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Appl. Math. Sci. 159, Springer, New York, 2004, https://doi.org/10.1007/978-1-4757-4355-5.

[17] E. B. Fabes, C. E. Kenig, and R. P. Serapioni, *The local regularity of solutions of degenerate elliptic equations*, Comm. Partial Differential Equations, 7 (1982), pp. 77–116, https://doi.org/10.1080/03605308208820218.

[18] G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations: Steady-State Problems*, Springer, New York, 2011, https://doi.org/10.1007/978-0-387-09620-9.

[19] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer, New York, 1986, https://doi.org/10.1007/978-3-642-61623-5.

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

[21] M. Hintermüller, R. 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 (2008), pp. 540–560, https://doi.org/10.1051/cocv:2007057.

[22] K. Kohls, A. Rösch, and K. Siebert, *A posteriori error analysis of optimal control problems with control constraints*, SIAM J. Control Optim., 52 (2014), pp. 1832–1861, https://doi.org/10.1137/130909251.

[23] V. A. Kozlov, V. G. Maz'ya, and J. Rossmann, *Elliptic Boundary Value Problems in Domains with Point Singularities*, Math. Surveys Monogr. 52, American Mathematical Society, Providence, RI, 1997.

[24] S. Larsson and E. D. Svensson, *Pointwise a Posteriori Error Estimates for the Stokes Equations in Polyhedral Domains*, preprint, 2006.

[25] W. Liu and N. Yan, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math., 15 (2001), pp. 285–309, https://doi.org/10.1023/A:1014239012739.

[26] R. H. Nochetto, E. Otárola, and A. J. Salgado, *Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications*, Numer. Math., 132 (2016), pp. 85–130, https://doi.org/10.1007/s00211-015-0709-6.

[27] R. H. Nochetto, K. G. Siebert, and A. Veeser, *Theory of adaptive finite element methods: An introduction*, in Multiscale, Nonlinear and Adaptive Approximation, Springer, New York, 2009, pp. 409–542, https://doi.org/10.1007/978-3-642-03413-8_12.

[28] R. H. Nochetto and A. Veeser, *Primer of adaptive finite element methods*, in Multiscale and Adaptivity: Modeling, Numerics and Applications, Lecture Notes in Math. 2040, Springer, New York, 2012, pp. 125–225, https://doi.org/10.1007/978-3-642-24079-9.

[29] E. Otárola and A. J. Salgado, *The Poisson and Stokes problems on weighted spaces in Lipschitz domains and under singular forcing*, J. Math. Anal. Appl., 471 (2019), pp. 599–612, https://doi.org/10.1016/j.jmaa.2018.10.094.

[30] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, Grad. Stud. Math. 112, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/gsm/112.

[31] B. O. Turesson, *Nonlinear Potential Theory and Weighted Sobolev Spaces*, Lecture Notes in Math. 1736, Springer, New York, 2000, https://doi.org/10.1007/BFb0103908.

[32] R. Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 2013, https://doi.org/10.1093/acprof:oso/9780199679423.001.0001.