ESAIM: M2AN 55 (2021) 2293–2322
https://doi.org/10.1051/m2an/2021033

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A POSTERIORI ERROR ESTIMATES FOR SEMILINEAR OPTIMAL CONTROL PROBLEMS

Alejandro Allendes, Francisco Fuica, Enrique Otárola* and Daniel Quero

Abstract. In two and three dimensional Lipschitz, but not necessarily convex, polytopal domains, we devise and analyze a reliable and efficient *a posteriori* error estimator for a semilinear optimal control problem; control constraints are also considered. We consider a fully discrete scheme that discretizes the state and adjoint equations with piecewise linear functions and the control variable with piecewise constant functions. The devised error estimator can be decomposed as the sum of three contributions which are associated to the discretization of the state and adjoint equations and the control variable. We extend our results to a scheme that approximates the control variable with piecewise linear functions and also to a scheme that approximates the solution to a nondifferentiable optimal control problem. We illustrate the theory with two and three-dimensional numerical examples.

Mathematics Subject Classification. 35J61, 49J20, 49M25, 65N15, 65N30.

Received July 29, 2020. Accepted July 8, 2021.

1. Introduction

In this work we will be interested in the design and analysis of *a posteriori* error estimates for finite element approximations of semilinear control–constrained optimal control problems; the state equation corresponds to a Dirichlet problem for a monotone, semilinear, and elliptic partial differential equation (PDE). 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 $\partial\Omega$. Notice that we do not assume that Ω is convex. Given a regularization parameter $\nu > 0$ and a desired state $y_\Omega \in L^2(\Omega)$, let us introduce the cost functional

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

We shall thus be concerned with the following *semilinear elliptic optimal control problem*: Find $\min J(y, u)$ subject to the *monotone, semilinear, and elliptic PDE*

$$-\Delta y + a(\cdot, y) = u \ \text{ in } \ \Omega, \qquad y = 0 \text{ on } \partial\Omega, \tag{1.2}$$

Keywords and phrases. optimal control problems, semilinear equations, finite element approximations, *a posteriori* error estimates.

Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile.
*Corresponding author: enrique.otarola@usm.cl

© The authors. Published by EDP Sciences, SMAI 2021

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

and the *control constraints*

$$u \in \mathbb{U}_{ad}, \quad \mathbb{U}_{ad} := \{v \in L^2(\Omega) : \mathtt{a} \leq v(x) \leq \mathtt{b} \text{ a.e. } x \in \Omega\}; \tag{1.3}$$

the control bounds $\mathtt{a}, \mathtt{b} \in \mathbb{R}$ are such that $\mathtt{a} < \mathtt{b}$. Assumptions on the function a will be deferred until Section 2.2.

The analysis of *a priori* error estimates for finite element approximations of semilinear optimal control problems has previously been considered in a number of works. The article [5] appears to be the first to provide error estimates for the distributed optimal control problem (1.1)–(1.3); notice that control constraints are considered. The authors of this work propose a fully discrete scheme on quasi–uniform meshes that discretizes the control variable with piecewise constant functions; piecewise linear functions are used for the discretization of the state and adjoint variables. In two and three dimensions and under the assumptions that Ω is convex, $\partial\Omega$ is of class $C^{1,1}$, and that the mesh–size is sufficiently small, the authors derive *a priori* error estimates for the approximation of the optimal control variable in the $L^2(\Omega)$-norm ([5], Thm. 5.1) and the $L^\infty(\Omega)$-norm Theorem 5.2 of [5]; the ones derived in the $L^2(\Omega)$-norm being optimal in terms of approximation. The analysis performed in [5] was later extended in [11] to a scheme that approximates the control variable with piecewise linear functions. The main result of this work reads as follows: $h_{\mathscr{T}}^{-1}\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \to 0$ as $h_{\mathscr{T}} \downarrow 0$ Theorem 4.1 of [11], where $\bar{u}_{\mathscr{T}}$ denotes the corresponding finite element approximation of $\bar{u}$. Under a suitable assumption, this result was later improved to Section 10 of [14]

$$\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim h_{\mathscr{T}}^{3/2}.$$

The following is a non-exhaustive list of extensions available in the literature: Neumann boundary optimal control [15], sparse optimal control [12], Dirichlet boundary optimal control [16], and state constrained optimal control [13].

While it is fair to say that the study of *a priori* error estimates for finite element solution techniques of semilinear optimal control problems is matured and well understood, the analysis of *a posteriori* error estimates is far from complete. An *a posteriori* error estimator is a computable quantity that depends on the discrete solution and data and is of primary importance in computational practice because of its ability to provide computable information about errors and drive the so-called adaptive finite element methods (AFEMs). The *a posteriori* error analysis for linear second–order elliptic boundary value problems have attained a mature understanding, including also the construction of AFEMs and their convergence and optimal complexity [1,27,32]. To the best of our knowledge, the first work that provided an advance regarding *a posteriori* error estimates for linear and distributed optimal control problems is [25]: the devised residual–type *a posteriori* error estimator is proven to yield an upper bound for the error Theorem 3.1 of [25]. These results were later improved in [21] where the authors explore a slight modification of the estimator of [25] and prove upper and lower error bounds which include oscillation terms Theorems 5.1 and 6.1 of [21]. Recently, these ideas were unified in [24]. In contrast to these advances the *a posteriori* error analysis for nonlinear optimal control problems is not as developed. To the best of our knowledge, the first work that provides an advance on this matter is [26]. In this work the authors derive *a posteriori* error estimates for distributed optimal control problems on Lipschitz domains and for nonlinear terms a which are such that

$$a(\cdot, y) \in W^{1,\infty}(-R, R), \; R > 0, \quad \frac{\partial a}{\partial y}(\cdot, y) \in L^2(\Omega), \quad \frac{\partial a}{\partial y}(\cdot, y) \geq 0.$$

Under the assumption that estimate (4.18) holds, with $\tilde{u} = \bar{u}_{\mathscr{T}}$, the authors devise an error estimator that yields an upper bound for the corresponding error on the $H^1(\Omega) \times H^1(\Omega) \times L^2(\Omega)$–norm Theorem 3.1 of [26]. We notice that no efficiency analysis is provided in [26]. We conclude this paragraph by mentioning the approach introduced in [7] for estimating the error in terms of the cost functional for linear/semilinear optimal control problems. This approach was later extended to problems with control constraints in [20,33] and state constraints in [8].

In this work, we propose an *a posteriori* error estimator for the optimal control problem (1.1)–(1.3) that can be decomposed as the sum of three contributions: one related to the discretization of the state equation, one

associated to the discretization of the adjoint equation, and another one that accounts for the discretization of the control variable. This error estimator is different to the one provided in [26]. In two and three dimensional Lipschitz polytopes, we obtain global reliability and efficiency properties. We must immediately mention that our global reliability result can be obtained under one of the either following assumptions:

- the underlying family of meshes is such that the conditions in (4.17) hold; see Theorems 4.5 and 5.2. These assumptions are needed in order to invoke second order sufficient optimality conditions. In particular, it is assumed that $\tilde{u} - \bar{u} \in C^{\tau}_{\tilde{u}}$, where $\tilde{u}$ is defined in (4.15). In Section 8.2, we provide numerical evidence that supports the claim that, after a suitable number of initial adaptive refinements, the condition $\tilde{u}-\bar{u} \in C^{\tau}_{\tilde{u}}$ holds true uniformly over subsequent adaptive refinements. In Remark 4.6, we show that a necessary condition for $\tilde{u} - \bar{u} \in C^{\tau}_{\tilde{u}}$ is that the estimate $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)} < \tau/2$ holds uniformly within adaptive refinement.
- the data of the problem is such the right-hand side estimate in (5.27) holds. In addition, it is also needed that $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq (2\mathcal{C})^{-1}\nu^2$, where $\mathcal{C}$ is defined in (5.34). In Section 8.2, we provide numerical evidence that supports the claim that, after a suitable number of initial adaptive refinements, the condition $\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq (2\mathcal{C})^{-1}\nu^2$ holds true uniformly over subsequent adaptive refinements.

On the basis of the devised error estimator, we also design a simple adaptive strategy that exhibits, for the examples that we perform, optimal experimental rates of convergence for all the optimal variables. We also provide numerical evidence that support the claim that our estimator outperforms the one in [26]; see Section 8. It is important to mention that, as opposed to the case when the state equation is linear, where, in general, only first order optimality conditions are needed to obtain *a posteriori* error estimates, here one of our strategies strongly relies on second order optimality conditions and on the particular structure of the associated critical cone. In addition, we mention that such a strategy is flexible enough to treat other PDE-constrained optimization problems. In particular, the following extensions of our theory are briefly discussed: piecewise linear approximation of the optimal control and sparse PDE-constrained optimization, where a nondifferentiable cost functional is considered.

The outline of this paper is as follows. In Section 2 we set notation and assumptions employed in the rest of the work. In Section 3 we review preliminary results about solutions to (1.2). Basic results for the optimal control problem (1.1)–(1.3) as well as first and second order optimality conditions are reviewed in Section 4. The core of our work are Sections 5 and 6, where we design an *a posteriori* error estimator for a suitable finite element discretization and show, in Sections 5 and 6, its reliability and efficiency, respectively. In Section 7 we present extensions of the theory developed in previous sections. Finally, numerical examples presented in Section 8 illustrate the theory and reveal a competitive performance of the devised error estimator.

2. Notation and assumptions

Let us set notation and describe the setting we shall operate with.

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 $\partial\Omega$. Notice that we do not assume that Ω is convex. If $\mathscr{X}$ and $\mathscr{Y}$ are Banach function spaces, $\mathscr{X} \hookrightarrow \mathscr{Y}$ means that $\mathscr{X}$ is continuously embedded in $\mathscr{Y}$. We denote by $\mathscr{X}'$ and $\|\cdot\|_{\mathscr{X}}$ the dual and norm, respectively, of $\mathscr{X}$. The relation $\mathfrak{a} \lesssim \mathfrak{b}$ indicates that $\mathfrak{a} \leq C\mathfrak{b}$, with a positive constant that depends neither on $\mathfrak{a}$, $\mathfrak{b}$ nor the discretization parameters. The value of C might change at each occurrence.

2.2. Assumptions

We assume that the nonlinear function a satisfies the following assumptions:

(A.1) $a : \Omega \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable and $a(\cdot, 0) \in L^2(\Omega)$.

(A.2) $\frac{\partial a}{\partial y}(x,y) \geq 0$ for a.e. $x \in \Omega$ and for all $y \in \mathbb{R}$.
(A.3) For all $\mathrm{M} > 0$, there exists a positive constant C_{M} such that

$$\sum_{i=1}^{2} \left| \frac{\partial^i a}{\partial y^i}(x,y) \right| \leq C_{\mathrm{M}}, \qquad \left| \frac{\partial^2 a}{\partial y^2}(x,v) - \frac{\partial^2 a}{\partial y^2}(x,w) \right| \leq C_{\mathrm{M}} |v-w|,$$

for a.e. $x \in \Omega$ and $|y|, |v|, |w| \leq \mathrm{M}$.

The following properties follow immediately from (A.1)–(A.3). First, a is monotone increasing in y for a.e. $x \in \Omega$. In particular, we have

$$(a(\cdot,v) - a(\cdot,w), v-w)_{L^2(\Omega)} \geq 0 \quad \forall v,w \in L^2(\Omega). \tag{2.1}$$

Second, a and $\frac{\partial a}{\partial y}$ are locally Lipschitz with respect to the second variable:

$$|a(x,v) - a(x,w)| \leq C_{\mathrm{M}} |v-w|, \qquad \left| \frac{\partial a}{\partial y}(x,v) - \frac{\partial a}{\partial y}(x,w) \right| \leq C_{\mathrm{M}} |v-w|, \tag{2.2}$$

for a.e. $x \in \Omega$ and $v,w \in \mathbb{R}$ such that $|v|, |w| \leq \mathrm{M}$.

3. Semilinear problems

In this section, we review some of the main results related to the existence and uniqueness of solutions for problem (1.2). We also review *a posteriori* error estimates for a particular finite element setting.

3.1. Weak formulation

Given $f \in L^q(\Omega)$ with $q > d/2$, we consider the following weak formulation of problem (1.2): Find $y \in H_0^1(\Omega)$ such that

$$(\nabla y, \nabla v)_{L^2(\Omega)} + (a(\cdot,y), v)_{L^2(\Omega)} = (f,v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega). \tag{3.1}$$

Invoking the main theorem on monotone operators ([35], Thm. 26.A; [28], Thm. 2.18) and an argument due to Stampacchia [30], ([23], Thm. B.2), the following result can be derived; see Section 2 of [14] and Theorem 4.8 of [31].

Theorem 3.1 (well–posedness)**.** *Let $q > d/2$ and $f \in L^q(\Omega)$. Let $a = a(x,y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function that is monotone increasing in y. Let $\psi_{\mathrm{M}} \in L^q(\Omega)$ be such that $|a(x,y)| \leq \psi_M(x)$ for a.e. $x \in \Omega$ and $|y| \leq \mathrm{M}$. Then problem (3.1) has a unique solution $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$. In addition, we have*

$$\|\nabla y\|_{L^2(\Omega)} + \|y\|_{L^\infty(\Omega)} \lesssim \|f - a(\cdot,0)\|_{L^q(\Omega)},$$

with a hidden constant that is independent of y, a, and f.

3.2. Finite element discretization

We begin by introducing some preliminary terminology. We denote by $\mathscr{T} = \{T\}$ a conforming partition of $\overline{\Omega}$ into simplices T with size $h_T := \mathrm{diam}(T)$. We denote by $\mathbb{T}$ the collection of conforming and shape regular meshes that are refinements of an initial mesh $\mathscr{T}_0$. We denote by $\mathscr{S}$ the set of internal $(d-1)$-dimensional interelement boundaries S of $\mathscr{T}$. If $T \in \mathscr{T}$, we define $\mathscr{S}_T$ as the subset of $\mathscr{S}$ that contains the sides of T. For $S \in \mathscr{S}$, we set $\mathcal{N}_S = \{T^+, T^-\}$, where $T^+, T^- \in \mathscr{T}$ are such that $S = T^+ \cap T^-$. We define the *star* or *patch* associated to the element $T \in \mathscr{T}$ as

$$\mathcal{N}_T = \bigcup \{T' \in \mathscr{T} : \mathscr{S}_T \cap \mathscr{S}_{T'} \neq \emptyset\}. \tag{3.2}$$

Given a mesh $\mathscr{T} \in \mathbb{T}$, we define the finite element space of continuous piecewise polynomials of degree one as

$$\mathbb{V}(\mathscr{T}) := \{v_{\mathscr{T}} \in C(\overline{\Omega}) : v_{\mathscr{T}}|_T \in \mathbb{P}_1(T)\ \forall T \in \mathscr{T}\} \cap H_0^1(\Omega). \tag{3.3}$$

Given a discrete function $v_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$, we define, for any internal side $S \in \mathscr{S}$, the jump or interelement residual $[\![\nabla v_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]$ by

$$[\![\nabla v_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!] := \boldsymbol{\nu}^+ \cdot \nabla v_{\mathscr{T}}|_{T^+} + \boldsymbol{\nu}^- \cdot \nabla v_{\mathscr{T}}|_{T^-},$$

where $\boldsymbol{\nu}^+, \boldsymbol{\nu}^-$ denote the unit normals to S pointing towards T^+, $T^- \in \mathscr{T}$, respectively; T^+, $T^- \in \mathscr{T}$ are such that $T^+ \neq T^-$ and $\partial T^+ \cap \partial T^- = S$.

We define the Galerkin approximation to problem (3.1) by

$$y_{\mathscr{T}} \in \mathbb{V}(\mathscr{T}) : \quad (\nabla y_{\mathscr{T}}, \nabla v_{\mathscr{T}})_{L^2(\Omega)} + (a(\cdot, y_{\mathscr{T}}), v_{\mathscr{T}})_{L^2(\Omega)} = (f, v_{\mathscr{T}})_{L^2(\Omega)} \tag{3.4}$$

for all $v_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$. The monotonicity of the nonlinear term a combined with an application of the Brouwer fixed point theorem yield the well–posedness of the discrete problem (3.4); see Section 7 of [14] for details.

3.3. *A posteriori* error analysis for the semilinear equation

Let $f \in L^2(\Omega)$ and let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be as in the statement of Theorem 3.1. Let us assume, in addition, that a is locally Lipschitz with respect to the second variable and $a(\cdot, 0) \in L^2(\Omega)$. With the notation introduced in Section 3.2 at hand, we define the following *a posteriori* local error indicators and error estimator

$$\mathcal{E}_T^2 := h_T^2 \|f - a(\cdot, y_{\mathscr{T}})\|_{L^2(T)}^2 + h_T \|[\![\nabla y_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T \setminus \partial\Omega)}^2, \quad \mathcal{E}_{\mathscr{T}}^2 := \sum_{T \in \mathscr{T}} \mathcal{E}_T^2,$$

respectively. Observe that, since a is locally Lipschitz with respect to the second variable and $a(\cdot, 0) \in L^2(\Omega)$, the residual term $h_T^2 \|f - a(\cdot, y_{\mathscr{T}})\|_{L^2(T)}^2$ is well–defined.

We present the following reliability result and, for the sake of readability, a proof.

Theorem 3.2 (global reliability of $\mathcal{E}_{\mathscr{T}}$)**.** *Let $f \in L^2(\Omega)$ and let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be as in the statement of Theorem 3.1. Let us assume, in addition, that $a(\cdot, 0) \in L^2(\Omega)$ and $a = a(x, y)$ is locally Lipschitz with respect to y. Let $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$ be the unique solution to problem (3.1) and $y_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ its finite element approximation obtained as the solution to (3.4). Then*

$$\|\nabla(y - y_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{\mathscr{T}},$$

with a hidden constant that is independent of y, $y_{\mathscr{T}}$, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. Let $v \in H_0^1(\Omega)$. Since y solves (3.1), we invoke Galerkin orthogonality and an elementwise integration by parts formula to arrive at

$$\begin{aligned}(\nabla(y - y_{\mathscr{T}}), \nabla v)_{L^2(\Omega)} &+ (a(\cdot, y) - a(\cdot, y_{\mathscr{T}}), v)_{L^2(\Omega)} \\ &= \sum_{T \in \mathscr{T}} \int_T (f - a(x, y_{\mathscr{T}}))(v - I_{\mathscr{T}} v) \mathrm{d}x + \sum_{S \in \mathscr{S}} \int_S [\![\nabla y_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!](v - I_{\mathscr{T}} v) \mathrm{d}x.\end{aligned}$$

Here, $I_{\mathscr{T}} : L^1(\Omega) \to \mathbb{V}(\mathscr{T})$ denotes the Clément interpolation operator [10, 18]. Standard approximation properties for $I_{\mathscr{T}}$ and the finite overlapping property of stars allow us to conclude that

$$\begin{aligned}(\nabla(y - y_{\mathscr{T}}), \nabla v)_{L^2(\Omega)} &+ (a(\cdot, y) - a(\cdot, y_{\mathscr{T}}), v)_{L^2(\Omega)} \\ &\lesssim \left[\sum_{T \in \mathscr{T}} h_T^2 \|f - a(\cdot, y_{\mathscr{T}})\|_{L^2(T)}^2 + h_T \|[\![\nabla y_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T \setminus \partial\Omega)}^2\right]^{\frac{1}{2}} \|\nabla v\|_{L^2(\Omega)}.\end{aligned}$$

Set $v = y - y_{\mathscr{T}} \in H_0^1(\Omega)$ and invoke property (2.1) to conclude. $\square$

4. A semilinear optimal control problem

In what follows, we assume that (A.1)–(A.3) hold and $q > d/2$. Let us introduce the following weak version of the optimal control problem (1.1)–(1.3): Find

$$\min\{J(y,u) : (y,u) \in H_0^1(\Omega) \times \mathbb{U}_{ad}\} \tag{4.1}$$

subject to the *monotone, semilinear, and elliptic state equation*

$$(\nabla y, \nabla v)_{L^2(\Omega)} + (a(\cdot, y), v)_{L^2(\Omega)} = (u, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega). \tag{4.2}$$

The existence of an optimal state-control pair is as follows; see Theorem 6.16 of [9], Theorem 4.15 of [31], and Theorem 6 of [14].

Theorem 4.1 (existence of the solution)**.** *The optimal control problem* (4.1)–(4.2) *admits at least one solution* $(\bar{y}, \bar{u}) \in H_0^1(\Omega) \cap L^\infty(\Omega) \times \mathbb{U}_{ad}$.

4.1. First order necessary optimality conditions

To formulate first order optimality conditions for problem (4.1)–(4.2), we introduce the so-called control-to-state map $\mathcal{S} : L^q(\Omega) \to H_0^1(\Omega) \cap L^\infty(\Omega)$ $(q > d/2)$, which, given a control $u \in \mathbb{U}_{ad} \subset L^q(\Omega)$, associates to it the unique state $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$ that solves (4.2). With this operator at hand, we introduce the reduced cost functional

$$j(u) := J(\mathcal{S}u, u) = \frac{1}{2}\|\mathcal{S}u - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}.$$

Theorem 4.17 in [31] guarantees that $\mathcal{S}$ is Fréchet differentiable from $L^q(\Omega)$ into $H_0^1(\Omega) \cap L^\infty(\Omega)$ $(q > d/2)$. We can thus formulate first order optimality conditions: if $\bar{u}$ is locally optimal for problem (4.1)–(4.2), then Lemma 4.18 of [31]

$$j'(\bar{u})(u - \bar{u}) \geq 0 \quad \forall u \in \mathbb{U}_{ad}. \tag{4.3}$$

Here, $j'(\bar{u})$ denotes the Gateâux derivative of the functional j at $\bar{u}$. To explore (4.3) we introduce the adjoint variable $p \in H_0^1(\Omega) \cap L^\infty(\Omega)$ as the unique solution to the *adjoint equation*

$$(\nabla w, \nabla p)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y)p, w\right)_{L^2(\Omega)} = (y - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega), \tag{4.4}$$

where $y = \mathcal{S}u$ solves (4.2). Observe that problem (4.4) is well–posed.

With these ingredients at hand, we present the following first order optimality conditions; see Theorem 4.20 of [31] and Theorem 3.2 of [5].

Theorem 4.2 (first order necessary optimality conditions)**.** *Every locally optimal control* $\bar{u} \in \mathbb{U}_{ad}$ *for problem* (4.1)–(4.2) *satisfies, together with the adjoint state* $\bar{p} \in H_0^1(\Omega) \cap L^\infty(\Omega)$*, the variational inequality*

$$(\bar{p} + \nu\bar{u}, u - \bar{u})_{L^2(\Omega)} \geq 0 \quad \forall u \in \mathbb{U}_{ad}. \tag{4.5}$$

Here, $\bar{p}$ *denotes the solution to* (4.4) *with* y *replaced by* $\bar{y} = \mathcal{S}\bar{u}$.

We now introduce the projection operator $\Pi_{[\mathtt{a},\mathtt{b}]} : L^1(\Omega) \to \mathbb{U}_{ad}$ as

$$\Pi_{[\mathtt{a},\mathtt{b}]}(v) := \min\{\mathtt{b}, \max\{v, \mathtt{a}\}\} \text{ a.e. in } \Omega. \tag{4.6}$$

With this projector at hand, we present the following result: If $\bar{u}$ denotes a locally optimal control for problem (4.1)–(4.2), then

$$\bar{u}(x) := \Pi_{[\mathtt{a},\mathtt{b}]}(-\nu^{-1}\bar{p}(x)) \text{ a.e. } x \in \Omega. \tag{4.7}$$

This formula implies that $\bar{u} \in H^1(\Omega) \cap L^\infty(\Omega)$; see Theorem A.1 of [23].

4.2. Second order sufficient optimality conditions

We follow [14, 17] and present necessary and sufficient second order optimality conditions.

Let $(\bar{y}, \bar{p}, \bar{u}) \in H_0^1(\Omega) \times H_0^1(\Omega) \times \mathbb{U}_{ad}$ satisfy the first order optimality conditions (4.2), (4.4), and (4.5). Define $\bar{\mathfrak{p}} := \bar{p} + \nu\bar{u}$. In view of (4.5), it follows that

$$\bar{\mathfrak{p}}(x) \begin{cases} = 0 & \text{a.e. } x \in \Omega \text{ if } \mathtt{a} < \bar{u} < \mathtt{b}, \\ \geq 0 & \text{a.e. } x \in \Omega \text{ if } \bar{u} = \mathtt{a}, \\ \leq 0 & \text{a.e. } x \in \Omega \text{ if } \bar{u} = \mathtt{b}. \end{cases}$$

Define the *cone of critical directions*

$$C_{\bar{u}} := \{v \in L^2(\Omega) \text{ satisfying (4.8) and } v(x) = 0 \text{ if } \bar{\mathfrak{p}}(x) \neq 0\},$$

with

$$v(x) \begin{cases} \geq 0 & \text{a.e. } x \in \Omega \text{ if } \bar{u}(x) = \mathtt{a}, \\ \leq 0 & \text{a.e. } x \in \Omega \text{ if } \bar{u}(x) = \mathtt{b}. \end{cases} \tag{4.8}$$

With these ingredients, we are in conditions to present second order necessary and sufficient optimality conditions; see Theorem 23 of [14].

Theorem 4.3 (second order necessary and sufficient optimality conditions)**.** *Suppose that assumptions (A.1)–(A.3) hold. If $\bar{u} \in \mathbb{U}_{ad}$ is a local minimum for problem (4.1)–(4.2), then*

$$j''(\bar{u})v^2 \geq 0 \quad \forall v \in C_{\bar{u}}.$$

Conversely, if $(\bar{y}, \bar{p}, \bar{u}) \in H_0^1(\Omega) \times H_0^1(\Omega) \times \mathbb{U}_{ad}$ satisfies the first order optimality conditions (4.2), (4.4), and (4.5), and

$$j''(\bar{u})v^2 > 0 \quad \forall v \in C_{\bar{u}} \setminus \{0\},$$

then, there exist $\mu > 0$ and $\varepsilon > 0$ such that

$$j(u) \geq j(\bar{u}) + \frac{\mu}{2}\|u - \bar{u}\|^2_{L^2(\Omega)} \quad \forall u \in \mathbb{U}_{ad} \cap \bar{B}_\varepsilon(\bar{u}),$$

where $\bar{B}_\varepsilon(\bar{u})$ denotes the closed ball in $L^2(\Omega)$ with center at $\bar{u}$ and radius ε.

Define

$$C^\tau_{\bar{u}} := \{v \in L^2(\Omega) \text{ satisfying (4.8) and } v(x) = 0 \text{ if } |\bar{\mathfrak{p}}(x)| > \tau\}. \tag{4.9}$$

The next result will be of importance for deriving *a posteriori* error estimates for suitable numerical discretizations of (4.1)–(4.2); see Theorem 25 of [14].

Theorem 4.4 (equivalent optimality condition)**.** *Suppose that assumptions (A.1)–(A.3) hold. If $\bar{u} \in \mathbb{U}_{ad}$ satisfies (4.5) then, the following statements are equivalent:*

$$j''(\bar{u})v^2 > 0 \quad \forall v \in C_{\bar{u}} \setminus \{0\}, \tag{4.10}$$

and

$$\exists \mu, \tau > 0 : \quad j''(\bar{u})v^2 \geq \mu\|v\|^2_{L^2(\Omega)} \quad \forall v \in C^\tau_{\bar{u}}. \tag{4.11}$$

We close this section with the following well–known estimate: Let $u, h, v \in L^\infty(\Omega)$ and $\mathsf{M} > 0$ be such that $\max\{\|u\|_{L^\infty(\Omega)}, \|h\|_{L^\infty(\Omega)}\} \leq \mathsf{M}$. Then, there exists $C_\mathsf{M} > 0$ such that Lemma 4.26 of [31]

$$|j''(u + h)v^2 - j''(u)v^2| \leq C_\mathsf{M}\|h\|_{L^2(\Omega)}\|v\|^2_{L^2(\Omega)}. \tag{4.12}$$

4.3. Finite element discretization

We present a finite element discretization for our optimal control problem. The approximation of the optimal control $\bar{u}$ is done by piecewise constant functions: $\bar{u}_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T})$, where

$$\mathbb{U}_{ad}(\mathscr{T}) := \mathbb{U}(\mathscr{T}) \cap \mathbb{U}_{ad}, \quad \mathbb{U}(\mathscr{T}) := \{u_{\mathscr{T}} \in L^\infty(\Omega) : u_{\mathscr{T}}|_T \in \mathbb{P}_0(T)\ \forall T \in \mathscr{T}\}.$$

The optimal state and adjoint state are discretized using the finite element space $\mathbb{V}(\mathscr{T})$ defined in (3.3). In this setting, the discrete counterpart of (4.1)–(4.2) reads as follows: Find min $J(y_{\mathscr{T}}, u_{\mathscr{T}})$ subject to the discrete state equation

$$y_{\mathscr{T}} \in \mathbb{V}(\mathscr{T}): \quad (\nabla y_{\mathscr{T}}, \nabla v_{\mathscr{T}})_{L^2(\Omega)} + (a(\cdot, y_{\mathscr{T}}), v_{\mathscr{T}})_{L^2(\Omega)} = (u_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)} \tag{4.13}$$

for all $v_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ and the discrete constraints $u_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T})$. This problem admits at least a solution Section 7 of [14]. In addition, if $\bar{u}_{\mathscr{T}}$ denotes a local solution, then

$$(\bar{p}_{\mathscr{T}} + \nu \bar{u}_{\mathscr{T}}, u_{\mathscr{T}} - \bar{u}_{\mathscr{T}})_{L^2(\Omega)} \geq 0 \quad \forall u_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T}),$$

where $\bar{p}_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ is such that

$$(\nabla w_{\mathscr{T}}, \nabla \bar{p}_{\mathscr{T}})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}, w_{\mathscr{T}}\right)_{L^2(\Omega)} = (\bar{y}_{\mathscr{T}} - y_\Omega, w_{\mathscr{T}})_{L^2(\Omega)} \tag{4.14}$$

for all $w_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$.

Define, on the basis of the projection operator (4.6), the auxiliary variable

$$\tilde{u} := \Pi_{[\mathsf{a},\mathsf{b}]}(-\nu^{-1}\bar{p}_{\mathscr{T}}). \tag{4.15}$$

Notice that $\tilde{u} \in \mathbb{U}_{ad}$ satisfies the following variational inequality Lemma 2.26 of [31]

$$(\bar{p}_{\mathscr{T}} + \nu\tilde{u}, u - \tilde{u})_{L^2(\Omega)} \geq 0 \quad \forall u \in \mathbb{U}_{ad}. \tag{4.16}$$

The following result is instrumental for our *a posteriori* error analysis.

Theorem 4.5 (auxiliary estimate)**.** *Suppose that assumptions (A.1)–(A.3) hold. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2) satisfying the sufficient second order optimality condition (4.10), or equivalently (4.11). If $\bar{u}_{\mathscr{T}}$ denotes a local minimum of the discrete optimal control problem and $\mathscr{T}$ is a mesh such that*

$$\tilde{u} - \bar{u} \in C^\tau_{\bar{u}}, \qquad \|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \nu\mu(2C_{\mathsf{M}})^{-1}, \tag{4.17}$$

then

$$\frac{\mu}{2}\|\bar{u} - \tilde{u}\|^2_{L^2(\Omega)} \leq (j'(\tilde{u}) - j'(\bar{u}))(\tilde{u} - \bar{u}). \tag{4.18}$$

The constant C_{M} is given by (4.12) and the auxiliary variable $\tilde{u}$ is defined in (4.15).

Proof. Since $\tilde{u} - \bar{u} \in C^\tau_{\bar{u}}$, with $C^\tau_{\bar{u}}$ defined in (4.9), and $\bar{u}$ satisfies the sufficient second order optimality condition (4.11), we are thus allow to set $v = \tilde{u} - \bar{u}$ there. This yields

$$\mu\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)} \leq j''(\bar{u})(\tilde{u} - \bar{u})^2. \tag{4.19}$$

On the other hand, in view of the mean value theorem, we obtain

$$(j'(\tilde{u}) - j'(\bar{u}))(\tilde{u} - \bar{u}) = j''(\zeta)(\tilde{u} - \bar{u})^2,$$

with $\zeta = \bar{u} + \theta_{\mathscr{T}}(\tilde{u} - \bar{u})$ and $\theta_{\mathscr{T}} \in (0,1)$. Thus, in view of (4.19), we arrive at

$$\mu\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)} \leq (j'(\tilde{u}) - j'(\bar{u}))(\tilde{u} - \bar{u}) + (j''(\bar{u}) - j''(\zeta))(\tilde{u} - \bar{u})^2. \tag{4.20}$$

Since there exists $\mathsf{M} > 0$ such that $\max\{\|\bar{u} + \theta_{\mathscr{T}}(\tilde{u} - \bar{u})\|_{L^\infty(\Omega)}, \|\tilde{u} - \bar{u}\|_{L^\infty(\Omega)}\} \leq \mathsf{M}$ and j is of class C^2 in $L^2(\Omega)$, we can thus apply (4.12) to derive

$$(j''(\bar{u}) - j''(\zeta))(\tilde{u} - \bar{u})^2 \leq C_{\mathsf{M}}\|\tilde{u} - \bar{u}\|_{L^2(\Omega)}\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)},$$

where we have also used that $\theta_{\mathscr{T}} \in (0,1)$. Invoke (4.7) and (4.15), the Lipschitz property of the projection operator $\Pi_{[\mathtt{a},\mathtt{b}]}$, defined in (4.6), and Assumption (4.17), to arrive at

$$(j''(\bar{u}) - j''(\zeta))(\tilde{u} - \bar{u})^2 \leq C_{\mathsf{M}}\nu^{-1}\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)}\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)} \leq \frac{\mu}{2}\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)}.$$

Replacing this inequality into (4.20) yields the desired inequality (4.18). □

Remark 4.6 (a sufficient condition for $\tilde{u} - \bar{u} \in C^\tau_{\bar{u}}$)**.** In what follows, we show that

$$\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)} < \tau/2 \implies \tilde{u} - \bar{u} \in C^\tau_{\bar{u}}.$$

In fact, since $\tilde{u} \in \mathbb{U}_{ad}$, we can immediately conclude that $\tilde{u} - \bar{u} \geq 0$ if $\bar{u} = \mathtt{a}$ and that $\tilde{u} - \bar{u} \leq 0$ if $\bar{u} = \mathtt{b}$. These arguments reveal that $v = \tilde{u} - \bar{u}$ satisfies (4.8). It thus suffices to verify the remaining condition in (4.9). To accomplish this task, we first use the triangle inequality and invoke the Lipschitz property of $\Pi_{[\mathtt{a},\mathtt{b}]}$, in conjunction with the assumption $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)} < \tau/2$, to obtain

$$\|\bar{p} + \nu\bar{u} - (\bar{p}_{\mathscr{T}} + \nu\tilde{u})\|_{L^\infty(\Omega)} \leq 2\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)} < \tau. \tag{4.21}$$

Now, let $\xi \in \Omega$ be such that $\bar{\mathfrak{p}}(\xi) = (\bar{p} + \nu\bar{u})(\xi) > \tau$. Since $\tau > 0$, this implies that $\bar{u}(\xi) > -\nu^{-1}\bar{p}(\xi)$. Therefore, from the projection formula (4.7), we conclude that $\bar{u}(\xi) = \mathtt{a}$. On the other hand, since $\xi \in \Omega$ is such that $(\bar{p} + \nu\bar{u})(\xi) > \tau$, from (4.21) we can conclude that

$$(\bar{p}_{\mathscr{T}} + \nu\tilde{u})(\xi) = \bar{p}_{\mathscr{T}}(\xi) + \nu\tilde{u}(\xi) > 0,$$

and thus that $\tilde{u}(\xi) > -\nu^{-1}\bar{p}_{\mathscr{T}}(\xi)$. This, on the basis of the definition of the auxiliary variable $\tilde{u}$, given in (4.15), yields that $\tilde{u}(\xi) = \mathtt{a}$. Consequently, $\bar{u}(\xi) = \tilde{u}(\xi) = \mathtt{a}$, and thus $(\tilde{u} - \bar{u})(\xi) = 0$. Similar arguments allow us to conclude that, if $\bar{\mathfrak{p}}(\xi) = (\bar{p} + \nu\bar{u})(\xi) < -\tau$, then $(\tilde{u} - \bar{u})(\xi) = 0$. Therefore, we have proved that $\tilde{u} - \bar{u} \in C^\tau_{\bar{u}}$.

5. *A POSTERIORI* ERROR ANALYSIS: RELIABILITY ESTIMATES

In this section, we devise and analyze an *a posteriori* error estimator for the discretization (4.13)–(4.14) of the optimal control problem (4.1)–(4.2). To simplify the exposition of the material, we define, for $(v, w, z) \in H^1_0(\Omega) \times H^1_0(\Omega) \times L^2(\Omega)$, the norm

$$|||(v, w, z)|||_\Omega := \|\nabla v\|_{L^2(\Omega)} + \|\nabla w\|_{L^2(\Omega)} + \|z\|_{L^2(\Omega)}. \tag{5.1}$$

The goal of this section is to obtain an upper bound for the error in the norm $|||\cdot|||_\Omega$. This will be obtained on the basis of estimates on the error between the solution to the discretization (4.13)–(4.14) and auxiliary variables that we define in what follows. Let $\hat{y} \in H^1_0(\Omega)$ be the solution to

$$(\nabla\hat{y}, \nabla v)_{L^2(\Omega)} + (a(\cdot, \hat{y}), v)_{L^2(\Omega)} = (\bar{u}_{\mathscr{T}}, v)_{L^2(\Omega)} \quad \forall v \in H^1_0(\Omega). \tag{5.2}$$

Define

$$\mathcal{E}^2_{st,T} := h^2_T\|\bar{u}_{\mathscr{T}} - a(\cdot, \bar{y}_{\mathscr{T}})\|^2_{L^2(T)} + h_T\|[\![\nabla\bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|^2_{L^2(\partial T\setminus\partial\Omega)}, \quad \mathcal{E}^2_{st} := \sum_{T\in\mathscr{T}} \mathcal{E}^2_{st,T}. \tag{5.3}$$

An application of Theorem 3.2 immediately yields the *a posteriori* error bound

$$\|\nabla(\hat{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st}. \tag{5.4}$$

Let $\hat{p} \in H_0^1(\Omega)$ be the solution to

$$(\nabla w, \nabla \hat{p})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\hat{p}, w\right)_{L^2(\Omega)} = (\bar{y}_{\mathscr{T}} - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega). \tag{5.5}$$

Define, for $T \in \mathscr{T}$, the local error indicators

$$\mathcal{E}_{ad,T}^2 := h_T^2 \|\bar{y}_{\mathscr{T}} - y_\Omega - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\|_{L^2(T)}^2 + h_T \|[\![\nabla \bar{p}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T \setminus \partial\Omega)}^2, \tag{5.6}$$

and the *a posteriori* error estimator

$$\mathcal{E}_{ad} := \left(\sum_{T \in \mathscr{T}} \mathcal{E}_{ad,T}^2\right)^{\frac{1}{2}}. \tag{5.7}$$

The following result yields an upper bound for the error $\|\nabla(\hat{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)}$ in terms of the computable quantity $\mathcal{E}_{ad}$.

Lemma 5.1 (global reliability of $\mathcal{E}_{ad}$)**.** *Suppose that assumptions (A.1)–(A.3) hold. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the discretization (4.13)–(4.14) with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the associated state and adjoint state, respectively. Then, the auxiliary variable $\hat{p}$, defined in (5.5), satisfies*

$$\|\nabla(\hat{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ad}. \tag{5.8}$$

The hidden constant is independent of the solution to (4.1)–(4.2), its finite element approximation, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. We proceed as in the proof of Theorem 3.2. Let $w \in H_0^1(\Omega)$. Since $\hat{p}$ solves (5.5), we invoke Galerkin orthogonality and an elementwise integration by parts formula to conclude that

$$\begin{aligned}(\nabla w, \nabla(\hat{p} - \bar{p}_{\mathscr{T}}))_{L^2(\Omega)} &+ \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})(\hat{p} - \bar{p}_{\mathscr{T}}), w\right)_{L^2(\Omega)} \\ &= \sum_{T \in \mathscr{T}} \left(\bar{y}_{\mathscr{T}} - y_\Omega - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}, w - I_{\mathscr{T}} w\right)_{L^2(T)} + \sum_{S \in \mathscr{S}} ([\![\nabla \bar{p}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], w - I_{\mathscr{T}} w)_{L^2(S)}.\end{aligned}$$

Standard approximation properties for $I_{\mathscr{T}}$ and the finite overlapping property of stars allow us to conclude that

$$\begin{aligned}(\nabla w, \nabla(\hat{p} - \bar{p}_{\mathscr{T}}))_{L^2(\Omega)} &+ \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})(\hat{p} - \bar{p}_{\mathscr{T}}), w\right)_{L^2(\Omega)} \\ &\lesssim \left[\sum_{T \in \mathscr{T}} h_T^2 \|\bar{y}_{\mathscr{T}} - y_\Omega - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\|_{L^2(T)}^2 + h_T \|[\![\nabla \bar{p}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T \setminus \partial\Omega)}^2\right]^{\frac{1}{2}} \|\nabla w\|_{L^2(\Omega)}.\end{aligned}$$

Set $w = \hat{p} - \bar{p}_{\mathscr{T}}$ and invoke assumption (A.2) to conclude. □

We define a global error estimator associated to the discretization of the optimal control variable as follows:

$$\mathcal{E}_{ct,T}^2 := \|\tilde{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)}^2, \qquad \mathcal{E}_{ct} := \left(\sum_{T \in \mathscr{T}} \mathcal{E}_{ct,T}^2\right)^{\frac{1}{2}}. \tag{5.9}$$

We recall that the auxiliary variable $\tilde{u}$ is defined as in (4.15).

The following two auxiliary variables, related to $\tilde{u} \in \mathbb{U}_{ad} \subset L^2(\Omega)$, will be of particular importance for our analysis. The variable $\tilde{y} \in H_0^1(\Omega)$, which solves

$$(\nabla \tilde{y}, \nabla v)_{L^2(\Omega)} + (a(\cdot, \tilde{y}), v)_{L^2(\Omega)} = (\tilde{u}, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega),$$

and $\tilde{p} \in H_0^1(\Omega)$, which is defined as the solution to

$$(\nabla w, \nabla \tilde{p})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \tilde{y})\tilde{p}, w\right)_{L^2(\Omega)} = (\tilde{y} - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega). \tag{5.10}$$

After all these definitions and preparations, we define an *a posteriori* error estimator for the optimal control problem (4.1)–(4.2), which can be decomposed as the sum of three contributions:

$$\mathcal{E}_{ocp}^2 := \mathcal{E}_{st}^2 + \mathcal{E}_{ad}^2 + \mathcal{E}_{ct}^2. \tag{5.11}$$

The estimators $\mathcal{E}_{st}$, $\mathcal{E}_{ad}$, and $\mathcal{E}_{ct}$, are defined in (5.3), (5.7), and (5.9), respectively.

We are now ready to state and prove one of the main results of this section.

Theorem 5.2 (global reliability I)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x, y)$ is globally Lipschitz with respect to $y \in \mathbb{R}$. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2) satisfying the sufficient second order condition (4.10), or equivalently (4.11). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the associated discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. Let $\mathscr{T}$ be a mesh such that (4.17) holds, then*

$$|||(\bar{y} - \bar{y}_{\mathscr{T}}, \bar{p} - \bar{p}_{\mathscr{T}}, \bar{u} - \bar{u}_{\mathscr{T}})|||_\Omega \lesssim \mathcal{E}_{ocp}. \tag{5.12}$$

The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. We proceed in four steps.

Step 1. The goal of this first step is to control the term $\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)}$. We begin with a simple application of the triangle inequality and write

$$\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \le \|\bar{u} - \tilde{u}\|_{L^2(\Omega)} + \mathcal{E}_{ct}, \tag{5.13}$$

where $\tilde{u} := \Pi_{[\mathsf{a},\mathsf{b}]}\left(-\nu^{-1}\bar{p}_{\mathscr{T}}\right)$ and $\mathcal{E}_{ct}$ is defined as in (5.9). Let us now bound the first term on the right hand side of (5.13). To accomplish this task, we set $u = \tilde{u}$ in (4.5) and $u = \bar{u}$ in (4.16) to obtain

$$-j'(\bar{u})(\tilde{u} - \bar{u}) = -(\bar{p} + \nu\bar{u}, \tilde{u} - \bar{u})_{L^2(\Omega)} \le 0, \qquad -(\bar{p}_{\mathscr{T}} + \nu\tilde{u}, \tilde{u} - \bar{u})_{L^2(\Omega)} \ge 0.$$

In light of these estimates, we invoke (4.18) to obtain

$$\tfrac{\mu}{2}\|\bar{u} - \tilde{u}\|_{L^2(\Omega)}^2 \le j'(\tilde{u})(\tilde{u} - \bar{u}) - j'(\bar{u})(\tilde{u} - \bar{u}) \le j'(\tilde{u})(\tilde{u} - \bar{u}) = (\tilde{p} + \nu\tilde{u}, \tilde{u} - \bar{u})_{L^2(\Omega)} \le (\tilde{p} - \bar{p}_{\mathscr{T}}, \tilde{u} - \bar{u})_{L^2(\Omega)}.$$

Adding and subtracting the auxiliary variable $\hat{p}$, defined as the solution to (5.5), and utilizing basic inequalities we arrive at

$$\|\bar{u} - \tilde{u}\|_{L^2(\Omega)}^2 \lesssim (\|\tilde{p} - \hat{p}\|_{L^2(\Omega)} + \|\hat{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)})\|\tilde{u} - \bar{u}\|_{L^2(\Omega)}. \tag{5.14}$$

We now invoke a Poincaré inequality and the *a posteriori* error estimate $\|\nabla(\hat{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ad}$, which follows from (5.8), to obtain

$$\|\bar{u} - \tilde{u}\|_{L^2(\Omega)} \lesssim \|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)} + \mathcal{E}_{ad}. \tag{5.15}$$

The rest of this step is dedicated to estimate the term $\|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)}$. To accomplish this task, we first notice that, for every $w \in H_0^1(\Omega)$, $\tilde{p} - \hat{p} \in H_0^1(\Omega)$ solves

$$(\nabla w, \nabla(\tilde{p} - \hat{p}))_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \tilde{y})\tilde{p} - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\hat{p}, w\right)_{L^2(\Omega)} = (\tilde{y} - \bar{y}_{\mathscr{T}}, w)_{L^2(\Omega)}.$$

Set $w = \tilde{p} - \hat{p}$ and invoke a generalized Hölder's inequality to obtain

$$\begin{aligned}
&\|\nabla(\tilde{p}-\hat{p})\|^2_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot,\tilde{y})(\tilde{p}-\hat{p}), \tilde{p}-\hat{p}\right)_{L^2(\Omega)} \\
&= (\tilde{y}-\bar{y}_{\mathscr{T}}, \tilde{p}-\hat{p})_{L^2(\Omega)} + \left(\left[\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}}) - \tfrac{\partial a}{\partial y}(\cdot,\tilde{y})\right]\hat{p}, \tilde{p}-\hat{p}\right)_{L^2(\Omega)} \\
&\leq \|\tilde{y}-\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}\|\tilde{p}-\hat{p}\|_{L^2(\Omega)} + \left\|\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}}) - \tfrac{\partial a}{\partial y}(\cdot,\tilde{y})\right\|_{L^2(\Omega)}\|\hat{p}\|_{L^4(\Omega)}\|\tilde{p}-\hat{p}\|_{L^4(\Omega)}.
\end{aligned}$$

Since $\frac{\partial a}{\partial y}$ is globally Lipschitz with respect to y, we thus obtain

$$\|\nabla(\tilde{p}-\hat{p})\|^2_{L^2(\Omega)} \lesssim \|\tilde{y}-\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}\left(\|\tilde{p}-\hat{p}\|_{L^2(\Omega)} + \|\hat{p}\|_{L^4(\Omega)}\|\tilde{p}-\hat{p}\|_{L^4(\Omega)}\right).$$

We thus use a Poincaré inequality and the embedding $H^1_0(\Omega) \hookrightarrow L^4(\Omega)$ to arrive at

$$\|\nabla(\tilde{p}-\hat{p})\|_{L^2(\Omega)} \lesssim \|\tilde{y}-\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}(1+\|\nabla\hat{p}\|_{L^2(\Omega)}). \tag{5.16}$$

Stability estimates for the problems that $\hat{p}$ and $\bar{y}_{\mathscr{T}}$ solve yield the estimate

$$\|\nabla\hat{p}\|_{L^2(\Omega)} \lesssim \|y_\Omega\|_{L^2(\Omega)} + \|y_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|y_\Omega\|_{L^2(\Omega)} + \rho|\Omega|^{\frac{1}{2}} + \|a(\cdot,0)\|_{L^2(\Omega)},$$

where $\rho = \max\{|\mathsf{a}|, |\mathsf{b}|\}$. Replacing this estimate into (5.16), and invoking, again, a Poincaré inequality, we obtain

$$\|\nabla(\tilde{p}-\hat{p})\|_{L^2(\Omega)} \lesssim \|\tilde{y}-\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|\nabla(\tilde{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)}, \tag{5.17}$$

with a hidden constant that is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$ but depends on the continuous problem data.

We now turn our attention to bounding the term $\|\nabla(\tilde{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)}$ in (5.17). To accomplish this task, we invoke the auxiliary variable $\hat{y}$, defined as the solution to (5.2), and use the triangle inequality to obtain

$$\|\nabla(\tilde{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \|\nabla(\tilde{y}-\hat{y})\|_{L^2(\Omega)} + \mathcal{E}_{st}, \tag{5.18}$$

where we have also used the *a posteriori* error estimate (5.4). It thus suffices to bound $\|\nabla(\tilde{y}-\hat{y})\|_{L^2(\Omega)}$. To do this, we first notice that $\tilde{y}-\hat{y} \in H^1_0(\Omega)$ solves the problem:

$$(\nabla(\tilde{y}-\hat{y}), \nabla v)_{L^2(\Omega)} + (a(\cdot,\tilde{y}) - a(\cdot,\hat{y}), v)_{L^2(\Omega)} = (\tilde{u}-\bar{u}_{\mathscr{T}}, v)_{L^2(\Omega)} \tag{5.19}$$

for all $v \in H^1_0(\Omega)$. Set $v = \tilde{y}-\hat{y}$. Since a is monotone increasing in y (2.1) we obtain that $\|\nabla(\tilde{y}-\hat{y})\|_{L^2(\Omega)} \lesssim \|\tilde{u}-\bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} = \mathcal{E}_{ct}$. Replacing this estimate into (5.18) and the obtained one into (5.17) yield

$$\|\nabla(\tilde{p}-\hat{p})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st} + \mathcal{E}_{ct}. \tag{5.20}$$

On the basis of (5.13), (5.15) and (5.20), we conclude the *a posteriori* error estimate

$$\|\bar{u}-\bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ad} + \mathcal{E}_{st} + \mathcal{E}_{ct}. \tag{5.21}$$

Step 2. The goal of this step is to bound $\|\nabla(\bar{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)}$. To accomplish this task, we invoke the auxiliary state $\hat{y}$, defined as the solution to (5.2) and apply the triangle inequality. In fact, we have

$$\|\nabla(\bar{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{y}-\hat{y})\|_{L^2(\Omega)} + \mathcal{E}_{st}, \tag{5.22}$$

where we have also used the *a posteriori* error estimate (5.4). It thus suffices to estimate $\|\nabla(\bar{y}-\hat{y})\|_{L^2(\Omega)}$. To achieve this goal, we invoke the state equation (4.2), with u replaced by $\bar{u}$, problem (5.2), and the monotonicity of the nonlinear term a (2.1). These arguments reveal that

$$\begin{aligned}
\|\nabla(\bar{y}-\hat{y})\|^2_{L^2(\Omega)} &\leq (\nabla(\bar{y}-\hat{y}), \nabla(\bar{y}-\hat{y}))_{L^2(\Omega)} + (a(\cdot,\bar{y}) - a(\cdot,\hat{y}), \bar{y}-\hat{y})_{L^2(\Omega)} \\
&= (\bar{u}-\bar{u}_{\mathscr{T}}, \bar{y}-\hat{y})_{L^2(\Omega)} \lesssim \|\bar{u}-\bar{u}_{\mathscr{T}}\|_{L^2(\Omega)}\|\nabla(\bar{y}-\hat{y})\|_{L^2(\Omega)}.
\end{aligned}$$

Consequently, $\|\nabla(\bar{y}-\hat{y})\|_{L^2(\Omega)} \lesssim \|\bar{u}-\bar{u}_{\mathscr{T}}\|_{L^2(\Omega)}$. Replacing this estimate into (5.22) and utilizing (5.21) allow us to conclude that

$$\|\nabla(\bar{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ad}+\mathcal{E}_{st}+\mathcal{E}_{ct}. \tag{5.23}$$

Step 3. We now bound the term $\|\nabla(\bar{p}-\bar{p}_{\mathscr{T}})\|_{L^2(\Omega)}$. To accomplish this task, we add and subtract $\hat{p}$, defined as the solution to (5.5), and use, again, the triangle inequality to obtain that

$$\|\nabla(\bar{p}-\bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{p}-\hat{p})\|_{L^2(\Omega)}+\mathcal{E}_{ad}, \tag{5.24}$$

where we have also used the *a posteriori* error estimate (5.8). It thus suffices to bound $\|\nabla(\bar{p}-\hat{p})\|_{L^2(\Omega)}$. Set $w=\bar{p}-\hat{p}$ in the weak problem that $\bar{p}-\hat{p}$ solves. This yields

$$\begin{aligned}\|\nabla(\bar{p}-\hat{p})\|^2_{L^2(\Omega)} &+ \left(\tfrac{\partial a}{\partial y}(\cdot,\bar{y})(\bar{p}-\hat{p}),\bar{p}-\hat{p}\right)_{L^2(\Omega)}\\ &= (\bar{y}-\bar{y}_{\mathscr{T}},\bar{p}-\hat{p})_{L^2(\Omega)} + \left(\left[\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})-\tfrac{\partial a}{\partial y}(\cdot,\bar{y})\right]\hat{p},\bar{p}-\hat{p}\right)_{L^2(\Omega)}.\end{aligned}$$

This identity, in view of a generalized Hölder's inequality, the Lipschitz property of $\frac{\partial a}{\partial y}$ with respect to y, and assumption (A.2), allows us to arrive at

$$\|\nabla(\bar{p}-\hat{p})\|^2_{L^2(\Omega)} \lesssim \|\bar{y}-\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}(\|\bar{p}-\hat{p}\|_{L^2(\Omega)}+\|\hat{p}\|_{L^4(\Omega)}\|\bar{p}-\hat{p}\|_{L^4(\Omega)}).$$

Using similar ideas to the ones that lead to (5.16) and (5.17), we can conclude that

$$\|\nabla(\bar{p}-\hat{p})\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{y}-\bar{y}_{\mathscr{T}})\|_{L^2(\Omega)}. \tag{5.25}$$

Replacing (5.23) into (5.25), and the obtained one into (5.24), we obtain

$$\|\nabla(\bar{p}-\bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ad}+\mathcal{E}_{st}+\mathcal{E}_{ct}. \tag{5.26}$$

Step 4. Combining (5.21), (5.23), and (5.26) allows us to arrive at (5.12). This concludes the proof. □

In Theorem 5.2, we obtained a global reliability result for our *a posteriori* error estimator $\mathcal{E}_{ocp}$ under the assumptions that the underlying adaptive loop driven by such an estimator and thus the family of adaptive meshes $\{\mathscr{T}\}$ are such that

$$\tilde{u}-\bar{u}\in C^{\tau}_{\bar{u}}, \qquad \|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \le \nu\mu(2C_{\mathsf{M}})^{-1}.$$

Numerical evidence supports the claim that, after a suitable number of initial adaptive refinements, the condition $\tilde{u}-\bar{u}\in C^{\tau}_{\bar{u}}$ holds true uniformly over subsequent adaptive refinements; see Figure 5 in Section 8.2. It is also observed that a quadratic experimental rate of convergence for $\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)}$ is attained so the condition $\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \le \nu\mu(2C_{\mathsf{M}})^{-1}$ is also computationally satisfied; see Figure 6 in Section 8.2.

As an alternative, in the next result, we obtain a global reliability result for $\mathcal{E}_{ocp}$ under the the following assumptions:

$$\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \le \frac{\nu^2}{2\mathcal{C}}, \qquad \max\{|\mathsf{a}|,|\mathsf{b}|,\|a(\cdot,0)\|_{L^2(\Omega)},\|y_\Omega\|_{L^2(\Omega)}\} \le C_{\max}, \tag{5.27}$$

where $\mathcal{C}$ and $C_{\max}>0$ are defined as in (5.34).

Theorem 5.3 (global reliability II)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that $\frac{\partial a}{\partial y}=\frac{\partial a}{\partial y}(x,y)$ is globally Lipschitz with respect to $y\in\mathbb{R}$. Let $\bar{u}\in\mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the associated discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. Assume that (5.27) holds. Then the a posteriori error estimate (5.12) holds, with a hidden constant independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.*

Proof. We begin with a simple application of the triangle inequality to obtain

$$\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \|\bar{u} - \tilde{u}\|_{L^2(\Omega)} + \mathcal{E}_{ct}, \tag{5.28}$$

where $\tilde{u} := \Pi_{[\mathsf{a},\mathsf{b}]}\left(-\nu^{-1}\bar{p}_{\mathscr{T}}\right)$ and $\mathcal{E}_{ct}$ is defined as in (5.9). Set $u = \tilde{u}$ in (4.5) and $u = \bar{u}$ in (4.16). Subtract the obtained inequalities to arrive at

$$\nu\|\bar{u} - \tilde{u}\|^2_{L^2(\Omega)} \leq (\bar{p} - \bar{p}_{\mathscr{T}}, \tilde{u} - \bar{u})_{L^2(\Omega)} = (\bar{p} - \tilde{p}, \tilde{u} - \bar{u})_{L^2(\Omega)} + (\tilde{p} - \bar{p}_{\mathscr{T}}, \tilde{u} - \bar{u})_{L^2(\Omega)},$$

where $\tilde{p}$ denotes the solution to (5.10). Young's inequality yields

$$\tfrac{3\nu}{4}\|\bar{u} - \tilde{u}\|^2_{L^2(\Omega)} \leq (\bar{p} - \tilde{p}, \tilde{u} - \bar{u})_{L^2(\Omega)} + \nu^{-1}\|\tilde{p} - \bar{p}_{\mathscr{T}}\|^2_{L^2(\Omega)}. \tag{5.29}$$

In what follows, we estimate the term $(\bar{p}-\tilde{p}, \tilde{u}-\bar{u})_{L^2(\Omega)}$. To accomplish this task, we notice that $\tilde{y}-\bar{y} \in H^1_0(\Omega)$ and $\bar{p} - \tilde{p} \in H^1_0(\Omega)$ solve

$$(\nabla(\tilde{y} - \bar{y}), \nabla v)_{L^2(\Omega)} + (a(\cdot, \tilde{y}) - a(\cdot, \bar{y}), v)_{L^2(\Omega)} = (\tilde{u} - \bar{u}, v)_{L^2(\Omega)} \quad \forall v \in H^1_0(\Omega), \tag{5.30}$$

and

$$\begin{aligned}(\nabla w, \nabla(\bar{p} - \tilde{p}))_{L^2(\Omega)} &+ \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y})(\bar{p} - \tilde{p}), w\right)_{L^2(\Omega)} \\ &= (\bar{y} - \tilde{y}, w)_{L^2(\Omega)} - \left(\tilde{p}\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}) - \tfrac{\partial a}{\partial y}(\cdot, \tilde{y})\right], w\right)_{L^2(\Omega)} \quad \forall w \in H^1_0(\Omega),\end{aligned} \tag{5.31}$$

respectively. Set $v = \bar{p} - \tilde{p}$ in (5.30) and $w = \tilde{y} - \bar{y}$ in (5.31) to obtain

$$\begin{aligned}(\bar{p} - \tilde{p}, \tilde{u} - \bar{u})_{L^2(\Omega)} = (a(\cdot, \tilde{y}) - a(\cdot, \bar{y}), \bar{p} - \tilde{p})_{L^2(\Omega)} &+ \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y})(\bar{p} - \tilde{p}), \bar{y} - \tilde{y}\right)_{L^2(\Omega)} \\ &+ \left(\tilde{p}\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}) - \tfrac{\partial a}{\partial y}(\cdot, \tilde{y})\right], \bar{y} - \tilde{y}\right)_{L^2(\Omega)} - \|\tilde{y} - \bar{y}\|^2_{L^2(\Omega)}.\end{aligned}$$

In view of Taylor's theorem, we conclude the existence of $\theta \in (0, 1)$ such that

$$a(\cdot, \tilde{y}) = a(\cdot, \bar{y}) + \tfrac{\partial a}{\partial y}(\cdot, \bar{y})(\tilde{y} - \bar{y}) + \tfrac{\partial^2 a}{\partial y^2}(\cdot, \mathsf{y})(\tilde{y} - \bar{y})^2,$$

with $\mathsf{y} = \bar{y} + \theta(\tilde{y} - \bar{y})$. This allows us to conclude that

$$(\bar{p}-\tilde{p}, \tilde{u}-\bar{u})_{L^2(\Omega)} = \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, \mathsf{y})(\tilde{y} - \bar{y})^2, \bar{p} - \tilde{p}\right)_{L^2(\Omega)} + \left(\tilde{p}\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}) - \tfrac{\partial a}{\partial y}(\cdot, \tilde{y})\right], \bar{y} - \tilde{y}\right)_{L^2(\Omega)} - \|\tilde{y} - \bar{y}\|^2_{L^2(\Omega)}. \tag{5.32}$$

Now, in view of the results of [19,22,29] and assumptions (A.1)–(A.3), we obtain that $\tilde{p} \in W^{1,r}_0(\Omega)$ for some $r > d$; recall that $\tilde{p}$ solves (5.10). This, in conjunction with the Sobolev embedding $W^{1,r}_0(\Omega) \hookrightarrow L^\infty(\Omega)$, for $r > d$, implies that

$$\|\tilde{p}\|_{L^\infty(\Omega)} \leq C_{\text{stab}}\|\tilde{y} - y_\Omega\|_{L^2(\Omega)}, \tag{5.33}$$

with $C_{\text{stab}} > 0$ being the stability constant. Invoke assumption (A.3), a generalized Hölder's inequality, the standard Sobolev embedding $H^1_0(\Omega) \hookrightarrow L^4(\Omega)$, and the stability estimate (5.33), to arrive at

$$(\bar{p} - \tilde{p}, \tilde{u} - \bar{u})_{L^2(\Omega)} \leq C_{\mathsf{M}}C_4^2\|\nabla(\bar{y} - \tilde{y})\|^2_{L^2(\Omega)}\|\bar{p} - \tilde{p}\|_{L^2(\Omega)} + (C_{\mathsf{M}}C_{\text{stab}}\|\tilde{y} - y_\Omega\|_{L^2(\Omega)} - 1)\|\bar{y} - \tilde{y}\|^2_{L^2(\Omega)}.$$

Given $q \in \{2, 4\}$, we denote by C_q the best constant associated to the embedding $H^1_0(\Omega) \hookrightarrow L^q(\Omega)$: C_q is such that $\|v\|_{L^q(\Omega)} \leq C_q\|\nabla v\|_{L^2(\Omega)}$ for all $v \in H^1_0(\Omega)$.

On the other hand, the well–posedness of problems (5.30) and (5.31) yields

$$\|\nabla(\tilde{y}-\bar{y})\|_{L^2(\Omega)} \leq C_2\|\bar{u}-\tilde{u}\|_{L^2(\Omega)},$$

and

$$\|\nabla(\bar{p}-\tilde{p})\|_{L^2(\Omega)} \leq C_2^2(1+C_{\mathsf{M}}\|\tilde{p}\|_{L^\infty(\Omega)})\|\nabla(\tilde{y}-\bar{y})\|_{L^2(\Omega)} \leq C_2^3(1+C_{\mathsf{M}}C_{\mathrm{stab}}\|\tilde{y}-y_\Omega\|_{L^2(\Omega)})\|\bar{u}-\tilde{u}\|_{L^2(\Omega)},$$

respectively. Hence, the combination of these estimates with the bound

$$\|\tilde{y}-y_\Omega\|_{L^2(\Omega)} \leq C_2\|\nabla\tilde{y}\|_{L^2(\Omega)}+\|y_\Omega\|_{L^2(\Omega)} \leq C_2^2\|\tilde{u}\|_{L^2(\Omega)}+C_2^2\|a(\cdot,0)\|_{L^2(\Omega)}+\|y_\Omega\|_{L^2(\Omega)} \leq C_{\max}(C_2^2|\Omega|^{\frac{1}{2}}+C_2^2+1),$$

which follows from assumption (5.27), yield

$$\begin{aligned}(\bar{p}-\tilde{p},\tilde{u}-\bar{u})_{L^2(\Omega)} \leq C_{\mathsf{M}}C_4^2C_2^6(1+C_{\mathsf{M}}C_{\mathrm{stab}}C_{\max}(C_2^2|\Omega|^{\frac{1}{2}}+C_2^2+1))\|\bar{u}-\tilde{u}\|^3_{L^2(\Omega)}\\ +(C_{\mathsf{M}}C_{\mathrm{stab}}C_{\max}(C_2^2|\Omega|^{\frac{1}{2}}+C_2^2+1)-1)\|\bar{y}-\tilde{y}\|^2_{L^2(\Omega)}.\end{aligned}$$

Define

$$C_{\max} := [C_{\mathsf{M}}C_{\mathrm{stab}}(C_2^2|\Omega|^{\frac{1}{2}}+C_2^2+1)]^{-1}, \qquad \mathcal{C} := 2C_{\mathsf{M}}C_4^2C_2^6, \tag{5.34}$$

and replace these constants in the previous inequality to conclude that

$$(\bar{p}-\tilde{p},\tilde{u}-\bar{u})_{L^2(\Omega)} \leq \mathcal{C}\|\bar{u}-\tilde{u}\|^3_{L^2(\Omega)}.$$

We now invoke the Lipschitz property of $\Pi_{[\mathsf{a},\mathsf{b}]}$ to obtain $\|\bar{u}-\tilde{u}\|_{L^2(\Omega)} \leq \nu^{-1}\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)}$. This estimate in combination with assumption (5.27) yield

$$(\bar{p}-\tilde{p},\tilde{u}-\bar{u})_{L^2(\Omega)} \leq \frac{\mathcal{C}}{\nu}\|\bar{u}-\tilde{u}\|^2_{L^2(\Omega)}\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \frac{\nu}{2}\|\bar{u}-\tilde{u}\|^2_{L^2(\Omega)}.$$

Replacing this estimate into (5.29) and invoking the auxiliary variable $\hat{p}$, solution to (5.5), allow us to obtain

$$\tfrac{\nu^2}{4}\|\bar{u}-\tilde{u}\|^2_{L^2(\Omega)} \leq \|\tilde{p}-\bar{p}_{\mathscr{T}}\|^2_{L^2(\Omega)} \lesssim (\|\nabla(\tilde{p}-\hat{p})\|^2_{L^2(\Omega)}+\|\nabla(\hat{p}-\bar{p}_{\mathscr{T}})\|^2_{L^2(\Omega)}).$$

This, in view of (5.8) and (5.20) (which do not rely on having second order sufficient optimality conditions), yields the estimate

$$\|\bar{u}-\tilde{u}\|^2_{L^2(\Omega)} \lesssim \mathcal{E}_{ad}^2+\mathcal{E}_{st}^2+\mathcal{E}_{ct}^2.$$

Replace this estimate into (5.28) to conclude $\|\bar{u}-\bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ad}+\mathcal{E}_{st}+\mathcal{E}_{ct}$.

The control of $\|\bar{y}-\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}$ and $\|\bar{p}-\tilde{p}_{\mathscr{T}}\|_{L^2(\Omega)}$ follow similar arguments to the ones developed in the proof of Theorem 5.2. This concludes the proof. □

6. *A POSTERIORI* ERROR ANALYSIS: EFFICIENCY ESTIMATES

In this section, we prove the local efficiency of the *a posteriori* error indicators $\mathcal{E}_{st,T}$ and $\mathcal{E}_{ad,T}$ and the global efficiency of the *a posteriori* error estimator $\mathcal{E}_{ocp}$. To accomplish this task, we will proceed on the basis of standard residual estimation techniques [1, 32]. Let us begin by introducing the following notation: for an edge/face or triangle/tetrahedron G, let $\mathcal{V}(G)$ be the set of vertices of G. With this notation at hand, we recall, for $T \in \mathscr{T}$ and $S \in \mathscr{S}$, the definition of the standard element and edge bubble functions [1, 32]

$$\varphi_T = (d+1)^{(d+1)}\prod_{\mathrm{v}\in\mathcal{V}(T)}\lambda_{\mathrm{v}}, \qquad \varphi_S = d^d\prod_{\mathrm{v}\in\mathcal{V}(S)}\lambda_{\mathrm{v}}|_{T'},$$

respectively, where $T' \subset \mathcal{N}_S$ and λ_{v} are the barycentric coordinates of T. Recall that $\mathcal{N}_S$ denotes the patch composed of the two elements of $\mathscr{T}$ that share S.

The following identities are essential to perform an efficiency analysis. First, since $\bar{y} \in H_0^1(\Omega)$ solves (4.2) with $u = \bar{u}$, an elementwise integration by parts formula implies that

$$(\nabla(\bar{y} - \bar{y}_{\mathscr{T}}), \nabla v)_{L^2(\Omega)} + (a(\cdot, \bar{y}) - a(\cdot, \bar{y}_{\mathscr{T}}), v)_{L^2(\Omega)} = (\bar{u} - \bar{u}_{\mathscr{T}}, v)_{L^2(\Omega)} + \sum_{T \in \mathscr{T}} \big[(\bar{u}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}}), v)_{L^2(T)} + (\mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}}) - a(\cdot, \bar{y}_{\mathscr{T}}), v)_{L^2(T)}\big] + \sum_{S \in \mathscr{S}} ([\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], v)_{L^2(S)} \tag{6.1}$$

for all $v \in H_0^1(\Omega)$. Second, since $\bar{p}$ solves (4.4) with $y = \bar{y}$, similar arguments yield

$$\begin{aligned}(\nabla w, \nabla(\bar{p} - \bar{p}_{\mathscr{T}}))_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y})(\bar{p} - \bar{p}_{\mathscr{T}}), w\right)_{L^2(\Omega)} &= (\bar{y} - \bar{y}_{\mathscr{T}}, w)_{L^2(\Omega)} + \left(\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}}) - \tfrac{\partial a}{\partial y}(\cdot, \bar{y})\right] \bar{p}_{\mathscr{T}}, w\right)_{L^2(\Omega)} \\ &+ \sum_{S \in \mathscr{S}} ([\![\nabla \bar{p}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!], w)_{L^2(S)} + \sum_{T \in \mathscr{T}} \Bigg[\left(\bar{y}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} y_\Omega - \mathscr{P}_{\mathscr{T}} \left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}}) \bar{p}_{\mathscr{T}}\right], w\right)_{L^2(T)} \\ &+ \left(\mathscr{P}_{\mathscr{T}} \left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}}) \bar{p}_{\mathscr{T}}\right] - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}}) \bar{p}_{\mathscr{T}}, w\right)_{L^2(T)} + (\mathscr{P}_{\mathscr{T}} y_\Omega - y_\Omega, w)_{L^2(T)}\Bigg]\end{aligned} \tag{6.2}$$

for all $w \in H_0^1(\Omega)$. In (6.1) and (6.2), $\mathscr{P}_{\mathscr{T}}$ denotes the L^2-projection onto piecewise constant functions over $\mathscr{T}$.

To derive local efficiency results, we introduce, for $w \in L^2(\Omega)$ and $\mathcal{M} \subset \mathscr{T}$,

$$\operatorname{osc}_{\mathscr{T}}(w; \mathcal{M}) := \left(\sum_{T \in \mathcal{M}} h_T^2 \|w - \mathscr{P}_{\mathscr{T}} w\|_{L^2(T)}^2\right)^{\frac{1}{2}}.$$

We are now ready to prove the local efficiency of the indicator $\mathcal{E}_{st}$ defined in (5.3).

Theorem 6.1 (local efficiency of $\mathcal{E}_{st}$)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that $a = a(x, y)$ is globally Lipschitz with respect to $y \in \mathbb{R}$. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the discretization (4.13)–(4.14) with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the associated state and adjoint state, respectively. Then, for $T \in \mathscr{T}$, the local error indicator $\mathcal{E}_{st,T}$ satisfies*

$$\mathcal{E}_{st,T} \lesssim \|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\mathcal{N}_T)} + h_T \|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\mathcal{N}_T)} + h_T \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\mathcal{N}_T)} + \operatorname{osc}_{\mathscr{T}}(a(\cdot, \bar{y}_{\mathscr{T}}), \mathcal{N}_T), \tag{6.3}$$

where $\mathcal{N}_T$ is defined as in (3.2). The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. We estimate each term in the definition of the local error indicator $\mathcal{E}_{st,T}$, given in (5.3), separately.

Step 1. Let $T \in \mathscr{T}$. We first bound the element term $h_T^2 \|\bar{u}_{\mathscr{T}} - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)}^2$. We begin with an application of the triangle inequality to obtain

$$h_T \|\bar{u}_{\mathscr{T}} - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)} \le h_T \|\bar{u}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)} + \operatorname{osc}_{\mathscr{T}}(a(\cdot, \bar{y}_{\mathscr{T}}); T).$$

It thus suffices to estimate the first term in the right hand side of the previous inequality. To accomplish this task, we set $v = \varphi_T(\bar{u}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}}))$ in (6.1) and invoke standard properties of the bubble function φ_T combined with basic inequalities to obtain

$$\begin{aligned}\|\bar{u}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)}^2 \lesssim \big(&h_T^{-1} \|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(T)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)} + \\ &\|a(\cdot, \bar{y}) - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)} + \|\mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}}) - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)}\big) \|\bar{u}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)}.\end{aligned}$$

This, in view of the global Lipschitz property of $a = a(x, y)$ with respect to y, yields

$$h_T^2\|\bar{u}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T)}^2 \lesssim \|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(T)}^2 + h_T^2\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)}^2 + h_T^2\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(T)}^2 + \operatorname{osc}_{\mathscr{T}}(a(\cdot, \bar{y}_{\mathscr{T}}); T)^2.$$

Step 2. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We bound $h_T\|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}^2$ in (5.3). To accomplish this task, we set $v = \varphi_S[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]$ in (6.1) and utilize standard bubble functions arguments to obtain

$$\begin{aligned}\|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}^2 \lesssim \sum_{T' \in \mathcal{N}_S} \Big(& h_{T'}^{-1}\|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(T')} + \|a(\cdot, \bar{y}) - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T')} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T')} \\ & + \|\bar{u}_{\mathscr{T}} - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T')} + \|\mathscr{P}_{\mathscr{T}} a(\cdot, \bar{y}_{\mathscr{T}}) - a(\cdot, \bar{y}_{\mathscr{T}})\|_{L^2(T')}\Big) h_T^{\frac{1}{2}}\|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}.\end{aligned}$$

Using again the global Lipschitz property of $a = a(x, y)$ with respect to y we arrive at

$$\begin{aligned}h_T\|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}^2 \lesssim \sum_{T' \in \mathcal{N}_S} \Big(& \|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(T')}^2 + h_T^2\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(T')}^2 \\ & + h_T^2\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T')}^2 + \operatorname{osc}_{\mathscr{T}}(a(\cdot, \bar{y}_{\mathscr{T}}); T')^2\Big).\end{aligned}$$

The collection of the estimates derived in Steps 1 and 2 concludes the proof. □

We now continue with the study of the local efficiency properties of the estimator $\mathcal{E}_{ad}$ defined in (5.7).

Theorem 6.2 (local efficiency of $\mathcal{E}_{ad}$)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x, y)$ is globally Lipschitz with respect to $y \in \mathbb{R}$. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the discretization (4.13)–(4.14) with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the associated state and adjoint state, respectively. Then, for $T \in \mathscr{T}$, the local error indicator $\mathcal{E}_{ad,T}$ satisfies*

$$\begin{aligned}\mathcal{E}_{ad,T} \lesssim & \|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\mathcal{N}_T)} + (1 + h_T)\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\mathcal{N}_T)} \\ & + h_T\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\mathcal{N}_T)} + \operatorname{osc}_{\mathscr{T}}(y_\Omega, \mathcal{N}_T) + \operatorname{osc}_{\mathscr{T}}\left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}, \mathcal{N}_T\right), \quad (6.4)\end{aligned}$$

where $\mathcal{N}_T$ is defined as in (3.2). The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. We estimate each term in the definition of the local error indicator $\mathcal{E}_{ad,T}$, given in (5.6), separately.

Step 1. Let $T \in \mathscr{T}$. A simple application of the triangle inequality yields

$$\begin{aligned}h_T\|\bar{y}_{\mathscr{T}} - y_\Omega - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\|_{L^2(T)} \le & \, h_T\left\|\bar{y}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} y_\Omega - \mathscr{P}_{\mathscr{T}}\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right]\right\|_{L^2(T)} \\ & + \operatorname{osc}_{\mathscr{T}}(y_\Omega; T) + \operatorname{osc}_{\mathscr{T}}\left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}; T\right).\end{aligned}$$

To estimate the first term on the right hand side of the previous estimate and also to simplify the presentation of the material, we define

$$\mathfrak{R}_T^{ad} := \bar{y}_{\mathscr{T}} - \mathscr{P}_{\mathscr{T}} y_\Omega - \mathscr{P}_{\mathscr{T}}\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right].$$

Now, set $w = \varphi_T \mathfrak{R}_T^{ad}$ in (6.2) and invoke basic inequalities to arrive at

$$\begin{aligned}\|\mathfrak{R}_T^{ad}\|_{L^2(T)}^2 \lesssim & \|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(T)}\|\nabla(\varphi_T \mathfrak{R}_T^{ad})\|_{L^2(T)} + \|\varphi_T \mathfrak{R}_T^{ad}\|_{L^2(T)}\Big(\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(T)} + \\ & \left\|\tfrac{\partial a}{\partial y}(\cdot, \bar{y})(\bar{p} - \bar{p}_{\mathscr{T}})\right\|_{L^2(T)} + \|\mathscr{P}_{\mathscr{T}} y_\Omega - y_\Omega\|_{L^2(T)} + \left\|\mathscr{P}_{\mathscr{T}}\left[\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right] - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right\|_{L^2(T)}\Big) \\ & + \left\|\tfrac{\partial a}{\partial y}(\cdot, \bar{y}) - \tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\right\|_{L^2(T)}\|\bar{p}_{\mathscr{T}}\|_{L^d(T)}\|\varphi_T \mathfrak{R}_T^{ad}\|_{L^\mathfrak{d}(T)}, \quad (6.5)\end{aligned}$$

where $\mathfrak{d} = \infty$ if $d = 2$ and $\mathfrak{d} = 6$ if $d = 3$. On the basis of (6.5), standard inverse inequalities ([10], Lem. 4.5.3) and bubble functions arguments yield

$$\|\mathfrak{R}_T^{ad}\|_{L^2(T)} \lesssim h_T^{-1}\|\nabla(\bar{p}-\bar{p}_{\mathscr{T}})\|_{L^2(T)} + \left\|\tfrac{\partial a}{\partial y}(\cdot,\bar{y})(\bar{p}-\bar{p}_{\mathscr{T}})\right\|_{L^2(T)} + \|\bar{y}-\bar{y}_{\mathscr{T}}\|_{L^2(T)} + \|\mathscr{P}_{\mathscr{T}} y_\Omega - y_\Omega\|_{L^2(T)}$$
$$+ \left\|\mathscr{P}_{\mathscr{T}}\left[\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right] - \tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right\|_{L^2(T)} + h_T^{-1}\|\bar{p}_{\mathscr{T}}\|_{L^d(T)}\left\|\tfrac{\partial a}{\partial y}(\cdot,\bar{y}) - \tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\right\|_{L^2(T)}. \quad (6.6)$$

Observe now that, since $d \in \{2,3\}$, $H_0^1(\Omega) \hookrightarrow L^d(\Omega)$. This, in combination with stability estimates for the problems that $\bar{p}_{\mathscr{T}}$ and $\bar{y}_{\mathscr{T}}$ solve yield the estimate

$$\|\bar{p}_{\mathscr{T}}\|_{L^d(T)} \le \|\bar{p}_{\mathscr{T}}\|_{L^d(\Omega)} \lesssim \|\bar{p}_{\mathscr{T}}\|_{H^1(\Omega)} \lesssim \|y_\Omega\|_{L^2(\Omega)} + \|y_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|y_\Omega\|_{L^2(\Omega)} + \rho|\Omega|^{\frac{1}{2}} + \|a(\cdot,0)\|_{L^2(\Omega)}, \quad (6.7)$$

where $\rho = \max\{|\mathsf{a}|,|\mathsf{b}|\}$. Replacing this estimate into (6.6), invoking the global Lipschitz property of $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x,y)$ with respect to y and assumption (A.3), we obtain

$$h_T\|\mathfrak{R}_T^{ad}\|_{L^2(T)} \lesssim \|\nabla(\bar{p}-\bar{p}_{\mathscr{T}})\|_{L^2(T)} + h_T\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(T)}$$
$$+ (1+h_T)\|\bar{y}-\bar{y}_{\mathscr{T}}\|_{L^2(T)} + \operatorname{osc}_{\mathscr{T}}(y_\Omega;T) + \operatorname{osc}_{\mathscr{T}}\left(\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}};T\right). \quad (6.8)$$

Notice that the hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$ but depends on the continuous problem data.

Step 2. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We now bound the jump term $\|[\![\nabla\bar{p}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{L^2(S)}$ in (5.6). To accomplish this task, we set $w = [\![\nabla\bar{p}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\varphi_S$ in (6.2) and proceed with similar arguments as the ones used in (6.5)–(6.6). We thus obtain

$$\|[\![\nabla\bar{p}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{L^2(S)}^2 \lesssim \sum_{T'\in\mathcal{N}_S}\Bigg(h_T^{-1}\|\nabla\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(T')} + \|\bar{y}-\bar{y}_{\mathscr{T}}\|_{L^2(T')} + \|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(T')}$$
$$+ \left\|\mathscr{P}_{T'}\left[\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right] - \tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}\right\|_{L^2(T')} + \|\mathscr{P}_{T'}y_\Omega - y_\Omega\|_{L^2(T')} + \|\mathfrak{R}_T^{ad}\|_{L^2(T')}$$
$$+ h_T^{-1}\|\bar{p}_{\mathscr{T}}\|_{L^d(T')}\left\|\tfrac{\partial a}{\partial y}(\cdot,\bar{y}) - \tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\right\|_{L^2(T')}\Bigg)h_T^{\frac{1}{2}}\|[\![\nabla\bar{p}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{L^2(S)}.$$

Finally, utilize the stability estimate (6.7), the global Lipschitz continuity of $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x,y)$ with respect to y, and estimate (6.8), to conclude

$$h_T^{\frac{1}{2}}\|[\![\nabla\bar{p}_{\mathscr{T}}\cdot\boldsymbol{\nu}]\!]\|_{L^2(S)} \lesssim \sum_{T'\in\mathcal{N}_S}\Bigg(\|\nabla(\bar{p}-\bar{p}_{\mathscr{T}})\|_{L^2(T)} + h_T\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^2(T)}$$
$$+ (1+h_T)\|\bar{y}-\bar{y}_{\mathscr{T}}\|_{L^2(T)} + \operatorname{osc}_{\mathscr{T}}(y_\Omega,T) + \operatorname{osc}_{\mathscr{T}}\left(\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}},T\right)\Bigg).$$

Combine the estimates derived in Steps 1 and 2 to arrive at the desired estimate (6.4). □

The results of Theorems 6.1 and 6.2 immediately yield the global efficiency of $\mathcal{E}_{ocp}$.

Theorem 6.3 (global efficiency of $\mathcal{E}_{ocp}$)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that $a = a(x,y)$ and $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x,y)$ are globally Lipschitz with respect to $y \in \mathbb{R}$. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the discretization (4.13)–(4.14) with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the associated state and adjoint state, respectively. Then,*

$$\mathcal{E}_{ocp} \lesssim \|\bar{p}-\bar{p}_{\mathscr{T}}\|_{H^1(\Omega)} + \|\bar{y}-\bar{y}_{\mathscr{T}}\|_{H^1(\Omega)} + \|\bar{u}-\bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} + \operatorname{osc}_{\mathscr{T}}(a(\cdot,\bar{y}_{\mathscr{T}});\mathscr{T}) + \operatorname{osc}_{\mathscr{T}}(y_\Omega;\mathscr{T}) + \operatorname{osc}_{\mathscr{T}}\left(\tfrac{\partial a}{\partial y}(\cdot,\bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}};\mathscr{T}\right).$$

The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. We begin by invoking the definition of the global indicator $\mathcal{E}_{st}$, given by (5.3), and the local efficiency estimate (6.3) to arrive at

$$\mathcal{E}_{st} \lesssim \|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} + \operatorname{diam}(\Omega)\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} + \operatorname{diam}(\Omega)\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} + \operatorname{osc}_{\mathscr{T}}(a(\cdot, \bar{y}_{\mathscr{T}}), \mathscr{T}). \tag{6.9}$$

On the other hand, in view of (5.7), the efficiency estimate (6.4) provides the bound

$$\begin{aligned}\mathcal{E}_{ad} \lesssim \|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} + (1 + \operatorname{diam}(\Omega))\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \\ + \operatorname{diam}(\Omega)\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} + \operatorname{osc}_{\mathscr{T}}(y_\Omega; \mathscr{T}) + \operatorname{osc}_{\mathscr{T}}\left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y}_{\mathscr{T}})\bar{p}_{\mathscr{T}}; \mathscr{T}\right).\end{aligned} \tag{6.10}$$

It thus suffices to control $\mathcal{E}_{ct}$. In view of (5.9), a trivial application of the triangle inequality yields

$$\mathcal{E}_{ct} \leq \|\tilde{u} - \bar{u}\|_{L^2(\Omega)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} = \|\Pi_{[\mathtt{a},\mathtt{b}]}(-\nu^{-1}\bar{p}_{\mathscr{T}}) - \Pi_{[\mathtt{a},\mathtt{b}]}(-\nu^{-1}\bar{p})\|_{L^2(\Omega)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)},$$

where $\Pi_{[\mathtt{a},\mathtt{b}]}$ is defined as in (4.6). This estimate, in conjunction with the Lipschitz property of $\Pi_{[\mathtt{a},\mathtt{b}]}$ and a Poincaré inequality, implies

$$\mathcal{E}_{ct} \lesssim \nu^{-1}\|\nabla(\bar{p}_{\mathscr{T}} - \bar{p})\|_{L^2(\Omega)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)}. \tag{6.11}$$

The proof concludes by gathering the estimates (6.9), (6.10), and (6.11). □

Remark 6.4 (local efficiency)**.** For $T \in \mathscr{T}$, let us introduce the local error indicator

$$\mathcal{E}_{ocp,T}^2 := \mathcal{E}_{st,T}^2 + \mathcal{E}_{ad,T}^2 + \mathcal{E}_{ct,T}^2. \tag{6.12}$$

A local efficiency result can be obtained for $\mathcal{E}_{ocp,T}$ if we measure the total error within a different norm. To be precise, define, for $(v, w, z) \in H^1(\Omega) \times H^1(\Omega) \times L^2(\Omega)$,

$$|||(v, w, z)|||_{1,\Omega} := \|v\|_{H^1(\Omega)} + \|w\|_{H^1(\Omega)} + \|z\|_{L^2(\Omega)}.$$

Let $T \in \mathscr{T}$. In view of (6.3) and (6.4), it suffices to bound $\mathcal{E}_{ct,T}$. Invoke (5.9), (4.7), (4.15), and the Lipschitz property of $\Pi_{[\mathtt{a},\mathtt{b}]}$ to arrive at

$$\begin{aligned}\mathcal{E}_{ct,T} &\leq \|\Pi_{[\mathtt{a},\mathtt{b}]}(-\nu^{-1}\bar{p}_{\mathscr{T}}) - \Pi_{[\mathtt{a},\mathtt{b}]}(-\nu^{-1}\bar{p})\|_{L^2(T)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)} \\ &\leq \nu^{-1}\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(T)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)} \leq \nu^{-1}\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{H^1(T)} + \|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)}.\end{aligned}$$

7. Extensions to the theory

We present a few extensions of the theory developed in the previous sections.

7.1. Piecewise linear approximation

In this section, we consider a similar finite element discretization as the one introduced in Section 4.3 with the difference that to approximate the optimal control variable $\bar{u}$ we employ piecewise linear functions, *i.e.*, we will seek for $\bar{u}_{\mathscr{T}}$ within the space $\mathbb{U}_{ad,1}(\mathscr{T})$, where

$$\mathbb{U}_{ad,1}(\mathscr{T}) := \mathbb{U}_1(\mathscr{T}) \cap \mathbb{U}_{ad}, \quad \mathbb{U}_1(\mathscr{T}) := \{u_{\mathscr{T}} \in C(\overline{\Omega}) : u_{\mathscr{T}}|_T \in \mathbb{P}_1(T) \forall\, T \in \mathscr{T}\}.$$

We propose the following discrete optimal control problem: Find $\min J(y_{\mathscr{T}}, u_{\mathscr{T}})$ subject to the discrete state equation

$$y_{\mathscr{T}} \in \mathbb{V}(\mathscr{T}): \quad (\nabla y_{\mathscr{T}}, \nabla v_{\mathscr{T}})_{L^2(\Omega)} + (a(\cdot, y_{\mathscr{T}}), v_{\mathscr{T}})_{L^2(\Omega)} = (u_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)} \tag{7.1}$$

for all $v_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ and the discrete control constraints $u_{\mathscr{T}} \in \mathbb{U}_{ad,1}(\mathscr{T})$. The well–posedness of this solution technique as well as first order optimality conditions follow from Theorem 3.3 of [11]. For *a priori* error estimates, we refer the reader to Theorem 4.1 of [11] and Section 10 of [14].

We propose an *a posteriori* error estimator that accounts for the discretization of the state, adjoint state, and control variables when the error, in each one of these variables, is measured in the $L^2(\Omega)$-norm. As it is customary when performing an *a posteriori* error analysis based on duality, we assume that Ω is convex.

Assume that we have at hand, *a posteriori* error estimators E_{st} and E_{ad} such that

$$\|\hat{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim E_{st}, \qquad \|\hat{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim E_{ad}. \tag{7.2}$$

We recall that the auxiliary variables $\hat{y} \in H_0^1(\Omega)$ and $\hat{p} \in H_0^1(\Omega)$ are defined as in (5.2) and (5.5), respectively.

Define, for $(v, w, z) \in L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)$, the norm

$$\|(v, w, z)\|_\Omega := \|v\|_{L^2(\Omega)} + \|w\|_{L^2(\Omega)} + \|z\|_{L^2(\Omega)}.$$

We present the following global reliability result.

Theorem 7.1 (global reliability)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x, y)$ is globally Lipschitz with respect to $y \in \mathbb{R}$. Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (4.1)–(4.2) satisfying the sufficient second order condition (4.10), or equivalently (4.11). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the associated discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. Let $\mathscr{T}$ be a mesh such that (4.17) holds, then*

$$\|(\bar{y} - \bar{y}_{\mathscr{T}}, \bar{p} - \bar{p}_{\mathscr{T}}, \bar{u} - \bar{u}_{\mathscr{T}})\|_\Omega \lesssim E_{st} + E_{ad} + \mathcal{E}_{ct}. \tag{7.3}$$

The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof. The proof of the estimate (7.3) follows closely the arguments developed in the proof of Theorem 5.2. In fact, with the estimate (5.14) at hand, we deduce

$$\|\bar{u} - \tilde{u}\|_{L^2(\Omega)} \lesssim \|\tilde{p} - \hat{p}\|_{L^2(\Omega)} + \|\hat{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|\tilde{p} - \hat{p}\|_{L^2(\Omega)} + E_{ad}, \tag{7.4}$$

where we have used (7.2). We thus use a Poincaré inequality in conjunction with the first estimate in (5.17) to obtain

$$\|\tilde{p} - \hat{p}\|_{L^2(\Omega)} \lesssim \|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)} \lesssim \|\tilde{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}. \tag{7.5}$$

The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$ but depends on the continuous problem data. To control $\|\tilde{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}$ we invoke the auxiliary state $\hat{y}$ defined as the solution to (5.2) and apply the triangle inequality. With these arguments we obtain

$$\|\tilde{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \|\tilde{y} - \hat{y}\|_{L^2(\Omega)} + \|\hat{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|\tilde{y} - \hat{y}\|_{L^2(\Omega)} + E_{st}, \tag{7.6}$$

where we have also used (7.2). To bound $\|\tilde{y} - \hat{y}\|_{L^2(\Omega)}$ we set $v = \tilde{y} - \hat{y}$ in problem (5.19). This, in view of the fact that a is monotone increasing with respect to y, yields

$$\|\tilde{y} - \hat{y}\|_{L^2(\Omega)} \lesssim \|\nabla(\tilde{y} - \hat{y})\|_{L^2(\Omega)} \lesssim \|\tilde{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} = \mathcal{E}_{ct}.$$

Replacing this estimate into (7.6), and the obtained one into (7.5), we obtain the estimate $\|\tilde{p} - \hat{p}\|_{L^2(\Omega)} \lesssim E_{st} + \mathcal{E}_{ct}$. This, in view of (7.4), reveals the *a posteriori* error estimate

$$\|\bar{u} - \bar{u}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim E_{st} + E_{ad} + \mathcal{E}_{ct}.$$

The control of $\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}$ and $\|\bar{p} - \tilde{p}_{\mathscr{T}}\|_{L^2(\Omega)}$ follow similar arguments as the ones elaborated in the proof of Theorem 5.2. For brevity, we skip details. □

7.2. Sparse PDE–constrained optimization

Throughout this section, we assume that the admissible set $\mathbb{U}_{ad}$ is such that $\mathtt{a} < 0 < \mathtt{b}$, with $\mathtt{a}, \mathtt{b} \in \mathbb{R}$. Define $\psi : L^1(\Omega) \to \mathbb{R}$ by $\psi(u) := \|u\|_{L^1(\Omega)}$. In this section, we present *a posteriori* error estimates for a semilinear optimal control problem that involves the cost functional

$$\mathfrak{J}(y,u) := J(y,u) + \vartheta\psi(u) = \frac{1}{2}\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \vartheta\|u\|_{L^1(\Omega)}.$$

Here, $\vartheta > 0$ denotes a sparsity parameter and $\nu > 0$ corresponds to the so-called regularization parameter. The cost functional involves the $L^1(\Omega)$-norm of the control variable, which is a natural measure of the control cost, and leads to sparsely supported optimal controls [12, 34]. Observe that $\mathfrak{J}$ is nondifferentiable; see Remark 2.1 [12]. We mention that the linear case, *i.e.*, $a \equiv 0$, has been investigated in [2].

We consider the following sparse PDE–constrained optimization problem: Find $\min\{\mathfrak{J}(y,u) : (y,u) \in H^1_0(\Omega) \times \mathbb{U}_{ad}\}$ subject to (4.2). This problem admits at least one optimal solution $(\bar{y}, \bar{u}) \in H^1_0(\Omega) \times \mathbb{U}_{ad}$. In addition, if $\bar{u}$ is a local minimum, then there exists $\bar{y} \in H^1_0(\Omega)$, $\bar{p} \in H^1_0(\Omega)$, and $\bar{\lambda} \in \partial\psi(\bar{u})$ such that (4.2) and (4.4) hold and

$$(\bar{p} + \nu\bar{u} + \vartheta\bar{\lambda}, u - \bar{u})_{L^2(\Omega)} \geq 0 \quad \forall\, u \in \mathbb{U}_{ad};$$

see Theorem 3.1 of [12]. The following characterizations for the optimal control $\bar{u}$ and its associated subgradient $\bar{\lambda}$ hold Corollary 3.2 of [12]:

$$\bar{\lambda}(x) := \Pi_{[-1,1]}\left(-\vartheta^{-1}\bar{p}(x)\right), \quad \bar{u}(x) = \Pi_{[\mathtt{a},\mathtt{b}]}\left(-\nu^{-1}\left[\bar{p}(x) + \vartheta\bar{\lambda}(x)\right]\right) \text{ a.e. } x \in \Omega.$$

We propose the following discrete optimal control problem: Find $\min \mathfrak{J}(y_{\mathscr{T}}, u_{\mathscr{T}})$ subject to (7.1) and the discrete control constraints $u_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T})$. The existence of solutions for this scheme as well as first order optimality conditions follow from Section 4 of [12].

Define the cones

$$\begin{aligned}\mathfrak{C}_{\bar{u}} &:= \{v \in L^2(\Omega) \text{ satisfying (4.8) and } j'(\bar{u})v + \vartheta\psi'(\bar{u};v) = 0\},\\ \mathfrak{C}^{\tau}_{\bar{u}} &:= \{v \in L^2(\Omega) \text{ satisfying (4.8) and } j'(\bar{u})v + \vartheta\psi'(\bar{u};v) \leq \tau\|v\|_{L^2(\Omega)}\}.\end{aligned}$$

Here, the term $\psi'(\bar{u};v)$ denotes the directional derivative of ψ at $\bar{u}$ in the direction v equation (3.2) of [12]. Necessary and sufficient second order optimality conditions follow from Theorems 3.7 and 3.9 of [12]: If $\bar{u}$ is a local minimum, then $j''(\bar{u})v^2 \geq 0$ for all $v \in \mathfrak{C}_{\bar{u}}$. Conversely, let $\bar{u} \in \mathbb{U}_{ad}$ and $\lambda \in \partial\psi(\bar{u})$ satisfy the associated first order optimality conditions. If $j''(\bar{u})v^2 > 0$ for all $v \in \mathfrak{C}_{\bar{u}} \setminus \{0\}$, then $\bar{u}$ is a local minimum. In addition, we have the equivalence Theorem 3.8 of [12]

$$j''(\bar{u})v^2 > 0 \;\forall v \in \mathfrak{C}_{\bar{u}} \setminus \{0\} \iff \exists \mu, \tau > 0 : j''(\bar{u})v^2 \geq \mu\|v\|^2_{L^2(\Omega)} \;\forall v \in \mathfrak{C}^{\tau}_{\bar{u}}. \tag{7.7}$$

Define, for a.e. $x \in \Omega$, the auxiliary variables

$$\tilde{\lambda}(x) := \Pi_{[-1,1]}\left(-\vartheta^{-1}\bar{p}_{\mathscr{T}}(x)\right), \qquad \tilde{u}(x) = \Pi_{[\mathtt{a},\mathtt{b}]}\left(-\nu^{-1}\left[\bar{p}_{\mathscr{T}}(x) + \vartheta\tilde{\lambda}(x)\right]\right). \tag{7.8}$$

To present *a posteriori* error estimates, we define the error indicators

$$\mathcal{E}^2_{sg,T} := \|\tilde{\lambda} - \bar{\lambda}_{\mathscr{T}}\|^2_{L^2(T)}, \qquad \mathcal{E}^2_{ct,T} := \|\tilde{u} - \bar{u}_{\mathscr{T}}\|^2_{L^2(T)},$$

and error estimators

$$\mathcal{E}_{sg} := \left(\sum_{T\in\mathscr{T}} \mathcal{E}^2_{sg,T}\right)^{\frac{1}{2}}, \qquad \mathcal{E}_{ct} := \left(\sum_{T\in\mathscr{T}} \mathcal{E}^2_{ct,T}\right)^{\frac{1}{2}}. \tag{7.9}$$

Theorem 7.2 (global reliability)**.** *Suppose that assumptions (A.1)–(A.3) hold. Assume, in addition, that* $\frac{\partial a}{\partial y} = \frac{\partial a}{\partial y}(x, y)$ *is globally Lipschitz with respect to* $y \in \mathbb{R}$*. Let* $\bar{u} \in \mathbb{U}_{ad}$ *be a local solution to the sparse PDE–constrained optimization problem satisfying the sufficient second order condition (7.7). Let* $\bar{u}_{\mathscr{T}}$ *be a local minimum of the associated discrete optimal control problem with* $\bar{y}_{\mathscr{T}}$*,* $\bar{p}_{\mathscr{T}}$*, and* $\bar{\lambda}_{\mathscr{T}}$ *being the corresponding state, adjoint state, and subgradient, respectively. Let* $\mathscr{T}$ *be a mesh such that (4.18) holds with* $\tilde{u}$ *as in (7.8), then*

$$|||(\bar{y} - \bar{y}_{\mathscr{T}}, \bar{p} - \bar{p}_{\mathscr{T}}, \bar{u} - \bar{u}_{\mathscr{T}})|||_{\Omega} + \|\bar{\lambda} - \bar{\lambda}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st} + \mathcal{E}_{ad} + \mathcal{E}_{ct} + \mathcal{E}_{sg}.$$

The hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh $\mathscr{T}$*, and* $\#\mathscr{T}$*.*

Proof. Since (4.18) is assumed to hold and it does not involve the nondifferentiable term ψ, the estimates of the errors associated to the state, adjoint state, and control variables are as presented in the proof of Theorem 5.2. It thus suffices to control the error associated to the approximation of the subgradient $\bar{\lambda}$. To accomplish this task, we invoke (7.9) and immediately conclude that

$$\|\bar{\lambda} - \bar{\lambda}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \|\bar{\lambda} - \tilde{\lambda}\|_{L^2(\Omega)} + \mathcal{E}_{sg}. \tag{7.10}$$

The Lipschitz property of $\Pi_{[-1,1]}$ and a Poincaré inequality yield

$$\|\bar{\lambda} - \tilde{\lambda}\|_{L^2(\Omega)} \leq \vartheta^{-1}\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)}.$$

Replace this estimate into (7.10) and invoke (5.26) to conclude. □

Remark 7.3 (feasibility of estimate (4.18))**.** Notice that $\tilde{u}$ coincides with the discrete approximation of $\bar{u}$ when the so–called variational discretization scheme is employed. For such an approximation scheme and within the framework of *a priori* error estimates, inequality (4.18) is proven in Section 5 of [12] and Lemma 4.6 of [12].

8. Numerical results

In this section, we conduct a series of numerical experiments that illustrate the performance of the devised *a posteriori* error estimator $\mathcal{E}_{ocp}$ defined in (5.11). In addition, in Section 8.2, we present a computational study related to the practical feasibility of the assumptions in (4.17).

All the experiments have been carried out with the help of a code that we implemented using `C++`. The involved matrices have been assembled exactly and global linear systems were solved using the multifrontal massively parallel sparse direct solver (MUMPS) [3,4]. The right hand sides and terms involving the functions $a(\cdot, y)$ and y_Ω, the approximation errors, and the error estimators are computed by quadrature formulas. In two dimensions, we use a quadrature formula that contains seventy-three (73) evaluation points on each element $T \in \mathscr{T}$; the associated numerical integration formula is exact for polynomials of degree nineteen (19). For three-dimensional domains, the employed quadrature formula contains two hundred and thirty six (236) evaluation points on each element $T \in \mathscr{T}$ and is exact for polynomials of degree fourteen (14).

8.1. Performance of the error estimator $\mathcal{E}_{ocp}$

For a given partition $\mathscr{T}$, we seek $(\bar{y}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{u}_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ that solves the discrete problem (4.13)–(4.14). This optimality system is solved by using a Newton–type primal–dual active set strategy as described in Algorithms 2 and 3. To be precise, Algorithm 2 presents a variant of the well–known primal–dual active set strategy that can be found, for instance, in Section 2.12.4 of [31]. On the other hand, Algorithm 3 describes the also well–known Newton method Section 4.4.1 of [6]. To present the latter, we define

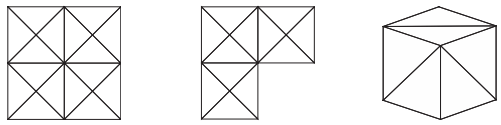

FIGURE 1. The initial meshes used when the domain Ω is a square, a L-shaped, and a cube.

$\mathcal{X}(\mathscr{T}) := \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}(\mathscr{T})$ and introduce, for $\Psi = (y_{\mathscr{T}}, p_{\mathscr{T}}, u_{\mathscr{T}})$ and $\Theta = (v_{\mathscr{T}}, w_{\mathscr{T}}, t_{\mathscr{T}})$ in $\mathcal{X}(\mathscr{T})$, the operator $F_{\mathscr{T}} : \mathcal{X}(\mathscr{T}) \to \mathcal{X}(\mathscr{T})'$ as

$$\langle F_{\mathscr{T}}(\Psi), \Theta\rangle := \begin{pmatrix} (\nabla y_{\mathscr{T}}, \nabla v_{\mathscr{T}})_{L^2(\Omega)} + (a(\cdot, y_{\mathscr{T}}) - u_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)} \\ (\nabla w_{\mathscr{T}}, \nabla p_{\mathscr{T}})_{L^2(\Omega)} + \left(\frac{\partial a}{\partial y}(\cdot, y_{\mathscr{T}}) p_{\mathscr{T}} - y_{\mathscr{T}} + y_\Omega, w_{\mathscr{T}}\right)_{L^2(\Omega)} \\ \left(\nu^{-1}\Pi_{\mathscr{T}} p_{\mathscr{T}}(\mathbf{1} - \boldsymbol{\chi}_{\mathtt{a}} - \boldsymbol{\chi}_{\mathtt{b}}) + u_{\mathscr{T}}\mathbf{1} - \mathtt{a}\boldsymbol{\chi}_{\mathtt{a}} - \mathtt{b}\boldsymbol{\chi}_{\mathtt{b}}, t_{\mathscr{T}}\right)_{L^2(\Omega)} \end{pmatrix}.$$

Here, $\Pi_{\mathscr{T}}$ denotes L^2–projection operator onto piecewise constant functions over $\mathscr{T}$ and $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathcal{X}(\mathscr{T})'$ and $\mathcal{X}(\mathscr{T})$. In addition,

$$\boldsymbol{\chi}_{\mathtt{a}}, \boldsymbol{\chi}_{\mathtt{b}} \in \mathbb{R}^{\#\mathscr{T}}, \quad \mathbf{1} = (1, \ldots, 1)^{\intercal} \in \mathbb{R}^{\#\mathscr{T}}.$$

Given an initial guess $\Psi_0 = (y^0_{\mathscr{T}}, p^0_{\mathscr{T}}, u^0_{\mathscr{T}}) \in \mathcal{X}(\mathscr{T})$ and $k \in \mathbb{N}_0$, we consider the following Newton iteration:

$$\Psi_{k+1} = \Psi_k + \eta,$$

where the incremental term $\eta = (\delta y_{\mathscr{T}}, \delta p_{\mathscr{T}}, \delta u_{\mathscr{T}}) \in \mathcal{X}(\mathscr{T})$ solves

$$\langle F'_{\mathscr{T}}(\Psi_k)(\eta), \Theta\rangle = -\langle F_{\mathscr{T}}(\Psi_k), \Theta\rangle \qquad \forall \Theta = (v_{\mathscr{T}}, w_{\mathscr{T}}, t_{\mathscr{T}}) \in \mathcal{X}(\mathscr{T}). \tag{8.1}$$

Here, $F'_{\mathscr{T}}(\Psi_k)(\eta)$ denotes the Gâteaux derivate of $F_{\mathscr{T}}$ at $\Psi_k = (y^k_{\mathscr{T}}, p^k_{\mathscr{T}}, u^k_{\mathscr{T}})$ in the direction η.

Once the discrete solution is obtained, we use the local error indicator $\mathcal{E}_{ocp,T}$, defined as in (6.12), to drive the adaptive procedure described in Algorithm 1. A sequence of adaptively refined meshes is thus generated from the initial meshes shown in Figure 1. The total number of degrees of freedom is $\mathsf{Ndof} = 2\dim(\mathbb{V}(\mathscr{T})) + \dim(\mathbb{U}(\mathscr{T}))$.

Finally, we define $e_y := \bar{y} - \bar{y}_{\mathscr{T}}$, $e_p := \bar{p} - \bar{p}_{\mathscr{T}}$, $e_u := \bar{u} - \bar{u}_{\mathscr{T}}$, and the total error $e := (e_y, e_p, e_u)$. To measure the total error we use $|||e|||_\Omega = |||(e_y, e_p, e_u)|||_\Omega$, where $|||\cdot|||_\Omega$ is defined as in (5.1).

Algorithm 1: Adaptive algorithm

Input: Initial mesh $\mathscr{T}_0$, constraints $\mathtt{a}$ and $\mathtt{b}$, and regularization parameter ν;
Set: $i = 0$.
Active set strategy:
1: Choose an initial discrete guess $(y^0_{\mathscr{T}_i}, p^0_{\mathscr{T}_i}, u^0_{\mathscr{T}_i}) \in \mathbb{V}(\mathscr{T}_i) \times \mathbb{V}(\mathscr{T}_i) \times \mathbb{U}(\mathscr{T}_i)$;
2: Compute $[\bar{y}_{\mathscr{T}_i}, \bar{p}_{\mathscr{T}_i}, \bar{u}_{\mathscr{T}_i}] = \textbf{Active-Set}[\mathscr{T}_i, \mathtt{a}, \mathtt{b}, \nu, y^0_{\mathscr{T}_i}, p^0_{\mathscr{T}_i}, u^0_{\mathscr{T}_i}]$ by using Algorithm 2;
Adaptive loop:
3: For each $T \in \mathscr{T}_i$ compute the local error indicator $\mathcal{E}_{ocp,T}$ defined in (6.12);
4: Mark an element $T \in \mathscr{T}_i$ 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**.

In order to simplify the construction of exact solutions, we incorporate an extra source term $f \in L^\infty(\Omega)$ in the state equation (4.2). With such a modification, the right hand side of (4.2) now reads $(f + u, v)_{L^2(\Omega)}$.

Algorithm 2: Active set algorithm

Input: Mesh $\mathscr{T}$, constraints $\mathtt{a}$ and $\mathtt{b}$, regularization parameter ν and initial guess $(y^0_{\mathscr{T}}, p^0_{\mathscr{T}}, u^0_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}(\mathscr{T})$;
1: Define $\boldsymbol{\chi}^{old}_{\mathtt{a}} = (\chi^{old}_{\mathtt{a},T})_{T\in\mathscr{T}}, \boldsymbol{\chi}^{old}_{\mathtt{b}} = (\chi^{old}_{\mathtt{b},T})_{T\in\mathscr{T}} \in \mathbb{R}^{\#\mathscr{T}}$ with $\chi^{old}_{\mathtt{a},T}, \chi^{old}_{\mathtt{b},T} \in \{0,1\}$.
Set: $j = 0$.
2: Compute $[y^{j+1}_{\mathscr{T}}, p^{j+1}_{\mathscr{T}}, u^{j+1}_{\mathscr{T}}] = \mathbf{Newton}[\mathscr{T}, \mathtt{a}, \mathtt{b}, \nu, \boldsymbol{\chi}^{old}_{\mathtt{a}}, \boldsymbol{\chi}^{old}_{\mathtt{b}}, y^j_{\mathscr{T}}, p^j_{\mathscr{T}}, u^j_{\mathscr{T}}]$ by using Algorithm 3.
3: For each $T \in \mathscr{T}$ compute

$$\chi^{new}_{\mathtt{a},T} = \begin{cases} 1 \text{ if } -\frac{1}{\nu}\Pi_T\left(p^{j+1}_{\mathscr{T}}\right) < \mathtt{a}, \\ 0 \text{ otherwise,} \end{cases} \qquad \chi^{new}_{\mathtt{b},T} = \begin{cases} 1 \text{ if } -\frac{1}{\nu}\Pi_T\left(p^{j+1}_{\mathscr{T}}\right) > \mathtt{b}, \\ 0 \text{ otherwise,} \end{cases}$$

where Π_T denotes the L^2–projection onto piecewise constant functions over T.
4: If $\sum_{T\in\mathscr{T}} \left(|\chi^{new}_{\mathtt{a},T} - \chi^{old}_{\mathtt{a},T}| + |\chi^{new}_{\mathtt{b},T} - \chi^{old}_{\mathtt{b},T}|\right) = 0$, set $(\bar{y}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{u}_{\mathscr{T}}) = (y^{j+1}_{\mathscr{T}}, p^{j+1}_{\mathscr{T}}, u^{j+1}_{\mathscr{T}})$. Otherwise, set $\boldsymbol{\chi}^{old}_{\mathtt{a}} := \boldsymbol{\chi}^{new}_{\mathtt{a}}$, $\boldsymbol{\chi}^{old}_{\mathtt{b}} := \boldsymbol{\chi}^{new}_{\mathtt{b}}$ and $j \leftarrow j+1$, and go to step **2**.

Algorithm 3: Newton method

Input: Mesh $\mathscr{T}$, constraints $\mathtt{a}$ and $\mathtt{b}$, regularization parameter ν, initial guess $(y^0_{\mathscr{T}}, p^0_{\mathscr{T}}, u^0_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}(\mathscr{T})$ and $\boldsymbol{\chi}_{\mathtt{a}}, \boldsymbol{\chi}_{\mathtt{b}} \in \mathbb{R}^{\#\mathscr{T}}$;
Set: $k = 0$.
1: Given $(y^k_{\mathscr{T}}, p^k_{\mathscr{T}}, u^k_{\mathscr{T}})$, compute the incremental $\eta = (\delta y_{\mathscr{T}}, \delta p_{\mathscr{T}}, \delta u_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}(\mathscr{T})$ as the solution to (8.1).
2: Set $(y^{k+1}_{\mathscr{T}}, p^{k+1}_{\mathscr{T}}, u^{k+1}_{\mathscr{T}}) = (y^k_{\mathscr{T}}, p^k_{\mathscr{T}}, u^k_{\mathscr{T}}) + (\delta y_{\mathscr{T}}, \delta p_{\mathscr{T}}, \delta u_{\mathscr{T}})$.
3: If $\max\{\|\delta y_{\mathscr{T}}\|_{L^\infty(\Omega)}, \|\delta p_{\mathscr{T}}\|_{L^\infty(\Omega)}, \|\delta u_{\mathscr{T}}\|_{L^\infty(\Omega)}\} < 10^{-8}$, set $(y_{\mathscr{T}}, p_{\mathscr{T}}, u_{\mathscr{T}}) = (y^{k+1}_{\mathscr{T}}, p^{k+1}_{\mathscr{T}}, u^{k+1}_{\mathscr{T}})$. Otherwise, set $k \leftarrow k+1$ and go to step **1**.

Example 8.1. We let $\Omega = (-1,1)^2 \setminus [0,1) \times (-1,0]$, $a(\cdot, y) = \arctan(y)$, $\mathtt{a} = -40$, $\mathtt{b} = -0.1$, and $\nu \in \{10^{-3}, 10^{-4}, 10^{-5}\}$. The exact optimal state and adjoint state are given, in polar coordinates (r,θ), with $\theta \in [0, 3\pi/2]$, by

$$\bar{y}(r,\theta) = \bar{p}(r,\theta) = \sin\left(\pi/2(r\sin\theta) + 1\right)\sin\left(\pi/2(r\cos\theta) + 1\right) r^{2/3}\sin(2\theta/3).$$

The purpose of this numerical example is threefold. First, we compare the performance of our adaptive FEM with uniform refinement. Second, we investigate the performance of the devised *a posteriori* error estimator when varying the parameter ν. Third, we compare the performance of our error estimator with the one presented in Section 3 of [26]. To present the error estimator of [26], we introduce

$$\mathfrak{E}_{st} := \mathcal{E}_{st}, \quad \mathfrak{E}_{ad} := \mathcal{E}_{ad}, \quad \mathfrak{E}_{ct,T} := h_T\|\nabla\bar{p}_{\mathscr{T}}\|_{L^2(T)}, \quad \mathfrak{E}_{ct} := \left(\sum_{T\in\mathscr{T}} \mathfrak{E}^2_{ct,T}\right)^{\frac{1}{2}},$$

where $\mathcal{E}_{st}$ and $\mathcal{E}_{ad}$ are defined as in (5.3) and (5.7), respectively. The total error indicator can thus be defined as follows Section 3 of [26]:

$$\mathfrak{E}^2_{ocp,T} = \mathfrak{E}^2_{st,T} + \mathfrak{E}^2_{ad,T} + \mathfrak{E}^2_{ct,T}. \tag{8.2}$$

This error indicator can be used to perform the adaptive FEM of Algorithm 1 with $\mathcal{E}_{ocp,T}$ replaced by $\mathfrak{E}_{ocp,T}$. We shall denote by $\mathfrak{e}_y$, $\mathfrak{e}_p$, and $\mathfrak{e}_u$ the approximation errors related to the state, adjoint state, and control variables, respectively, when the error indicator $\mathfrak{E}_{ocp,T}$ is considered in Algorithm 1. We measure the total error of the underlying AFEM with $|||\mathfrak{e}|||_\Omega = |||(\mathfrak{e}_y, \mathfrak{e}_p, \mathfrak{e}_u)|||_\Omega$, where $|||\cdot|||_\Omega$ is defined in (5.1). Finally, we introduce the effectivity indices $\Upsilon_{\mathcal{E}} := \mathcal{E}_{ocp}/|||e|||_\Omega$ and $\Upsilon_{\mathfrak{E}} := \mathfrak{E}_{ocp}/|||\mathfrak{e}|||_\Omega$.

In Figures 2 and 3 we present the results obtained for Example 8.1. In Figure 2 we present, for $\nu = 10^{-3}$, experimental rates of convergence for all the individual contributions of the total error $|||e|||_\Omega$ when uniform and

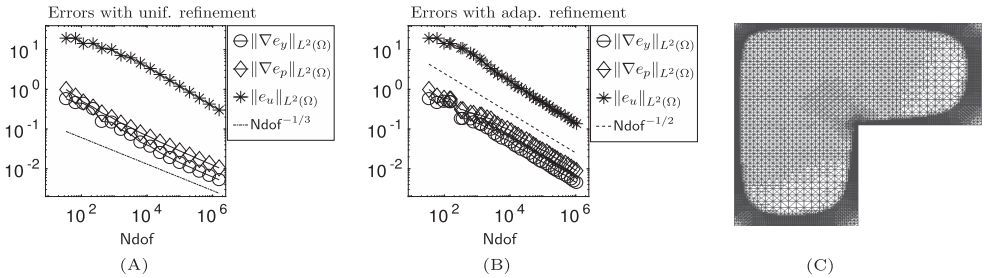

FIGURE 2. Example 8.1. Experimental rates of convergence for the individual contributions $\|\nabla e_y\|_{L^2(\Omega)}, \|\nabla e_p\|_{L^2(\Omega)}$, and $\|e_u\|_{L^2(\Omega)}$ for uniform (A) and adaptive refinement (B) and the 24th adaptively refined mesh (C) for $\nu = 10^{-3}$.

adaptive refinement are considered. We also present the adaptively refined mesh obtained after 24 adaptive loops. We observe that our adaptive loop *outperforms* uniform refinement. In addition, we observe optimal experimental rates of convergence for all the individual contributions of the total error $|||e|||_\Omega$. We also observe that most of the adaptive refinement occurs near to the interface of the control variable and the geometric singularity of the L–shaped domain, which attests to the efficiency of the devised estimator; see subfigure (C). In Figure 3, we present, for $\nu \in \{10^{-4}, 10^{-5}\}$, experimental rates of convergence for all the contributions of the total errors $|||e|||_\Omega$ and $|||\mathfrak{e}|||_\Omega$ and all the individual contributions of the *a posteriori* error estimators $\mathcal{E}_{ocp}$ and $\mathfrak{E}_{ocp}$ as well as the effectivity indices $\Upsilon_{\mathcal{E}}$ and $\Upsilon_{\mathfrak{E}}$. We observe that the behavior of the individual contributions of the total errors and error estimators associated to the state and adjoint variables are quite *similar* for both adaptive strategies. However, we observe an *important* difference when we compare the individual contributions associated to the control variable. In fact, as it can be observed from subfigures (B.3) and (D.3), the error norm $\|\mathfrak{e}_u\|_{L^2(\Omega)}$ *do not exhibit an optimal experimental rate of convergence*, while the error norm $\|e_u\|_{L^2(\Omega)}$ associated to our devised AFEM based on the error estimator $\mathcal{E}_{ocp}$ does. Finally, we observe, from subfigures (E) and (F), that the effectivity index $\Upsilon_{\mathcal{E}}$ is close to 1 for the two different values of ν that we consider. This shows the accuracy of the proposed *a posteriori* error estimator $\mathcal{E}_{ocp}$ when used in the adaptive loop described in Algorithm 1.

Example 8.2. We let $\Omega = (0,1)^3$, $\mathtt{a} = -80$, $\mathtt{b} = 100$, and $\nu = 10^{-3}$. We consider

$$f(x_1,x_2,x_3) = 10, \quad y_\Omega(x_1,x_2,x_3) = \begin{cases} 10^2 e^{\frac{1}{\xi}} \cos(4\pi\xi), & \text{if } \xi < 0, \\ 0, & \text{if } \xi \geq 0, \end{cases}$$

where $\xi = \xi(x_1,x_2,x_3) = 4(x_1-0.5)^2 + 4(x_2-0.5)^2 + 4(x_3-0.5)^2 - 1$.

The purpose of this numerical example is to investigate the performance of the devised error estimator when different choices of the nonlinear function a are considered. Let us, in particular, consider

$$a_1(\cdot,y) = 10y^3 - 2; \quad a_2(\cdot,y) = 10\arctan(80y) - 5; \quad a_3(\cdot,y) = 10\sinh(3y) - 2.$$

In Figure 4 we present the results obtained for Example 8.2. We show, for the considered three different nonlinear functions a, experimental rates of convergence for all the individual contributions of the error estimator $\mathcal{E}_{ocp}$ and the obtained 25th adaptively refined meshes. We observe optimal experimental rates of convergence for all the individual contributions of the error estimator $\mathcal{E}_{ocp}$.

8.2. Numerical verification of assumptions (4.17)

The main goal of this section is to show that, at a computational level, the assumptions in (4.17), namely,

$$\tilde{u} - \bar{u} \in C_{\bar{u}}^{\tau}, \qquad \|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \nu\mu(2C_{\mathsf{M}})^{-1},$$

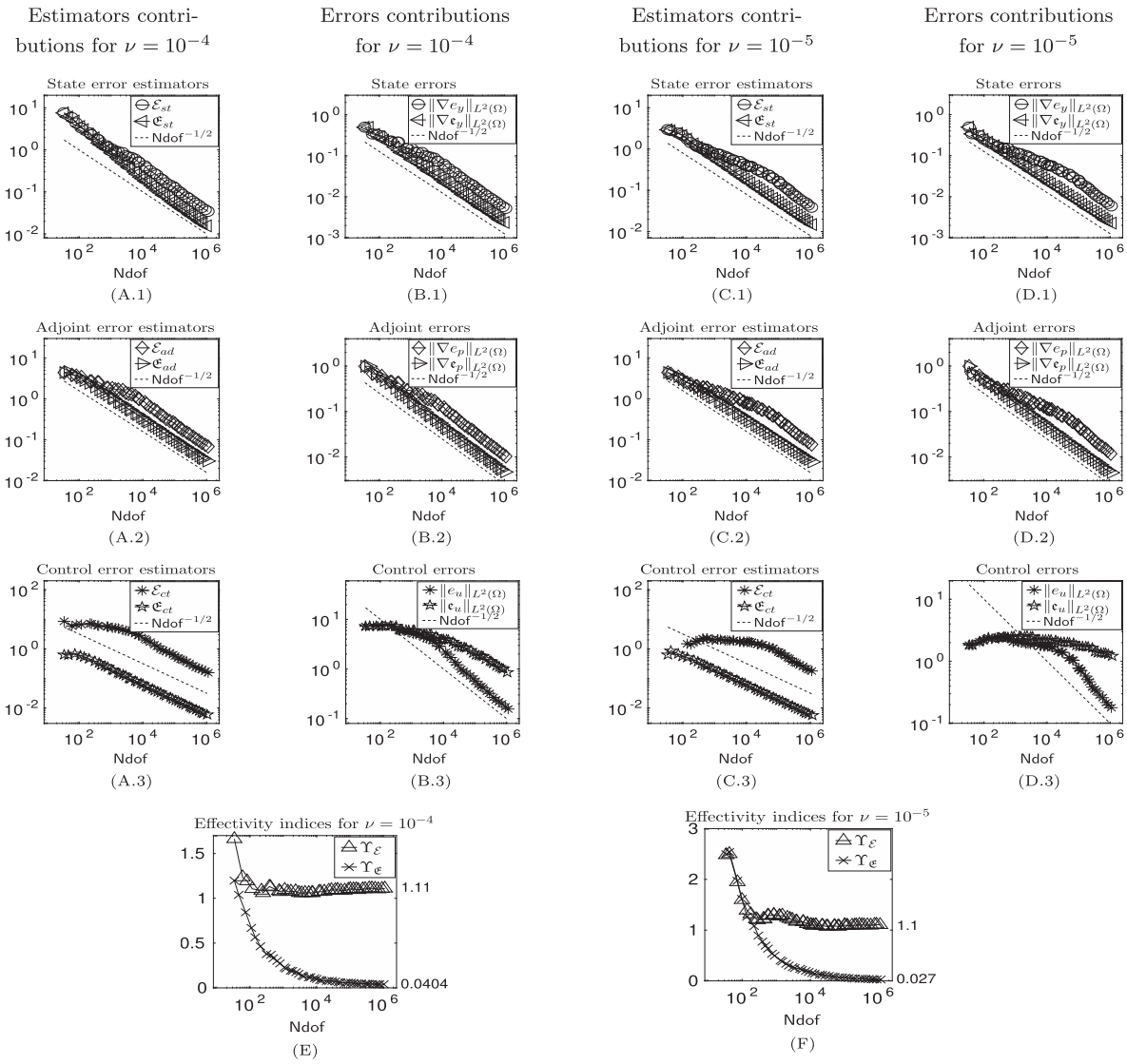

FIGURE 3. Example 8.1. Experimental rates of convergence for all the contributions of $\mathcal{E}_{ocp}$ and $\mathfrak{E}_{ocp}$ (A.1)–(A.3)–(C.1)–(C.3), experimental rates of convergence for all the contributions of the total errors $|||e|||_\Omega$ and $|||\mathfrak{e}|||_\Omega$ (B.1)–(B.3)–(D.1)–(D.3), and the effectivity indices $\Upsilon_{\mathcal{E}}$ and $\Upsilon_{\mathfrak{E}}$ (E)–(F) with $\nu = 10^{-4}$ and $\nu = 10^{-5}$.

are satisfied in practice. To do this, we will explore a few examples. We recall that the auxiliary variable $\tilde{u} \in \mathbb{U}_{ad}$ is defined by $\tilde{u} := \Pi_{[\mathsf{a},\mathsf{b}]}(-\nu^{-1}\bar{p}_{\mathscr{T}})$. Since, in practice, it is observed that $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)}$ decays as the number of degrees of freedom increases, the verification of assumption $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \nu\mu(2C_{\mathsf{M}})^{-1}$ is standard. We will thus particularly focus on the assumption $\tilde{u} - \bar{u} \in C_{\bar{u}}^{\tau}$ and describe, in what follows, a strategy that can be used to verify it. In light of definition (4.9), the term $v := \tilde{u} - \bar{u}$ belongs to $C_{\bar{u}}^{\tau}$ if v satisfies the sign condition (4.8) and $v(x) = 0$ if $|\bar{\mathfrak{p}}(x)| > \tau$ for a.e. $x \in \Omega$. Recall that $\bar{\mathfrak{p}} = \bar{p} + \nu\bar{u}$. We immediately notice that, since $\tilde{u} \in \mathbb{U}_{ad}$, then $v = \tilde{u} - \bar{u}$ satisfies the sign condition (4.8). Hence, in what follows we concentrate on the condition $v(x) = 0$

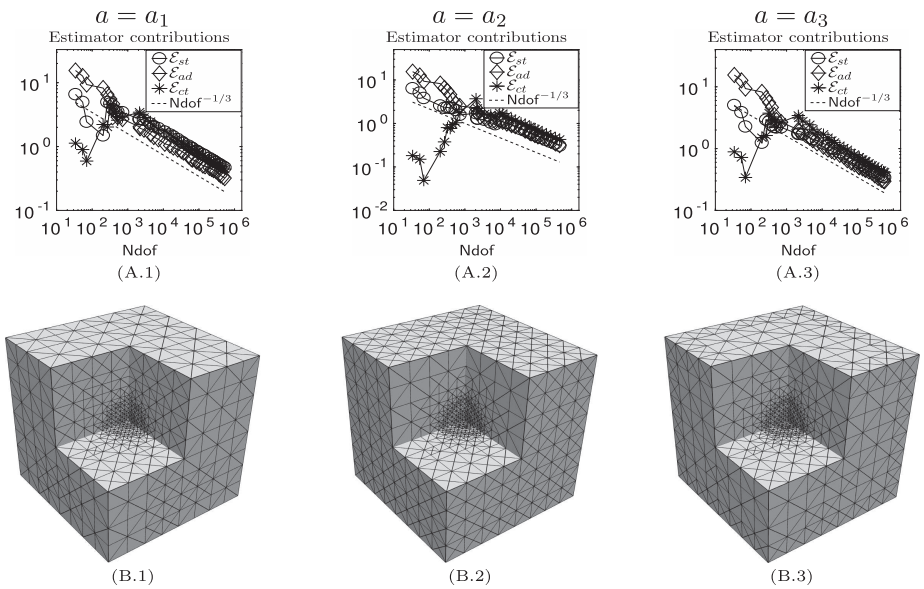

FIGURE 4. Example 8.2: Experimental rates of convergence for $\mathcal{E}_{st}$, $\mathcal{E}_{ad}$, and $\mathcal{E}_{ct}$ (A.1)–(A.3) and adaptively refined meshes obtained after 25 adaptive loops (B.1)–(B.3) with $\nu = 10^{-3}$.

if $|\bar{\mathfrak{p}}(x)| > \tau$ for a.e. $x \in \Omega$. Since this requires to make point evaluations, we will evaluate $v = \tilde{u} - \bar{u}$ at the quadrature points associated to the integration formulas mentioned at the beginning of Section 8.

Let $\mathscr{T}$ be a fixed mesh in $\mathbb{T}$. For each $T \in \mathscr{T}$, we define $Q_T \subset T$ as the discrete set that contains the corresponding quadrature points of the underlying integration formula. Given $\tau > 0$, we introduce the following sets:

$$C_{\bar{\mathfrak{p}},T} = \{x \in Q_T : |\bar{\mathfrak{p}}(x)| > \tau\}, \qquad C_{\bar{u},T} = \{x \in C_{\bar{\mathfrak{p}},T} : \tilde{u}(x) - \bar{u}(x) = 0\}.$$

With these ingredients at hand, we notice that the main goal now is reduced to verify that

$$v(x) = \tilde{u}(x) - \bar{u}(x) = 0 \text{ for every } x \in \bigcup_{T \in \mathscr{T}} C_{\bar{\mathfrak{p}},T},$$

which is, $C_{\bar{\mathfrak{p}},T} = C_{\bar{u},T}$ for every $T \in \mathscr{T}$. To verify this condition, we introduce the quantity

$$R_\tau = \frac{\sum_{T \in \mathscr{T}} \#C_{\bar{u},T}}{\sum_{T \in \mathscr{T}} \#C_{\bar{\mathfrak{p}},T}}.$$

Since, for $T \in \mathscr{T}$, $C_{\bar{u},T} \subseteq C_{\bar{\mathfrak{p}},T}$ we have that $0 \leq \#C_{\bar{u},T} \leq \#C_{\bar{\mathfrak{p}},T}$, which implies that $R_\tau \leq 1$. In particular, we observe that $R_\tau = 1$ if and only if $C_{\bar{\mathfrak{p}},T} = C_{\bar{u},T}$ for every $T \in \mathscr{T}$. Therefore, when $R_\tau = 1$ we will have $\tilde{u}(x) - \bar{u}(x) = 0$ if $|\bar{\mathfrak{p}}(x)| > \tau$ for every $x \in \cup_{T \in \mathscr{T}} Q_T$, and thus $\tilde{u} - \bar{u} \in C_{\bar{u}}^\tau$.

We now present numerical experiments where we verify the accomplishment of the condition $\tilde{u} - \bar{u} \in C_{\bar{u}}^\tau$ by utilizing R_τ. Since the exact value of $\tau > 0$ is unknown, we verify this condition for several values of τ. We also present experimental rates of convergence for $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)}$. In the following numerical examples, we have considered $\nu = 10^{-1}$ and $a(\cdot, y) = \arctan(y)$.

Example 1 (2D Convex domain). We let $\Omega = (0,1)^2$, $\mathtt{a} = -20$, and $\mathtt{b} = -0.1$. The exact optimal state and adjoint state are given by

$$\bar{y}(x_1, x_2) = \bar{p}(x_1, x_2) = \sin(2\pi x_1)\sin(2\pi x_2).$$

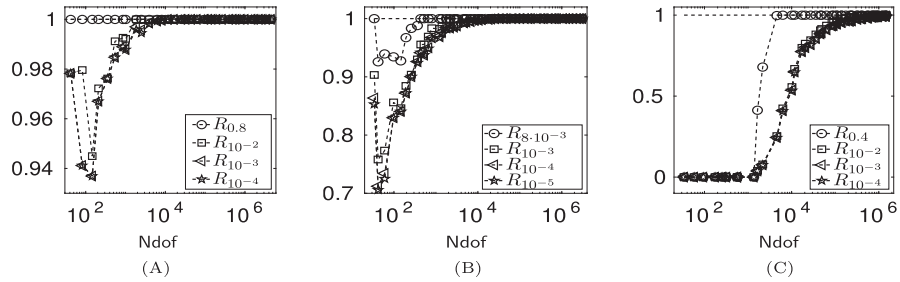

FIGURE 5. Experimental values of R_τ obtained for Example 1 (A), Example 2 (B), and Example 3 (C), with different values of τ.

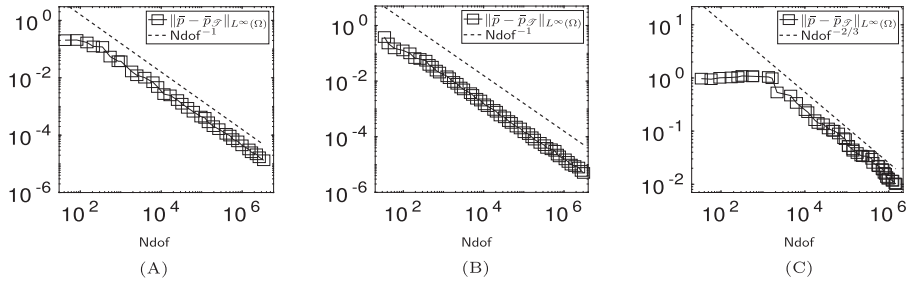

FIGURE 6. Experimental rates of convergences of $\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)}$ for Example 1 (A), Example 2 (B), and Example 3 (C).

Example 2 (2D Non-convex domain). We let $\Omega = (-1,1)^2 \setminus [0,1)\times(-1,0]$, $\mathtt{a} = -20$, and $\mathtt{b} = -0.1$. The exact optimal state and adjoint state are given, in polar coordinates (r,θ) with $\theta \in [0, 3\pi/2]$, by

$$\bar{y}(r,\theta) = \bar{p}(r,\theta) = \sin\left(\pi/2(r\sin\theta)+1\right)\sin\left(\pi/2(r\cos\theta)+1\right) r^{2/3}\sin(2\theta/3).$$

Example 3 (3D Convex domain). We let $\Omega = (0,1)^3$, $\mathtt{a} = -5$, and $\mathtt{b} = 5$. The exact optimal state and adjoint state are given by

$$\bar{y}(x_1,x_2,x_3) = \bar{p}(x_1,x_2,x_3) = -2\cos\left(2\pi x_2\right)\sin^2\left(2\pi x_1\right)\sin\left(2\pi x_2\right)\sin^2\left(2\pi x_3\right).$$

In Figure 5 we present the experimental values of R_τ obtained for each of the three numerical examples; we have considered different values of τ. We observe that, for all the considered values of τ, the quantity R_τ achieves the value 1 after a certain number of degrees of freedom; smaller values of τ requiring a bigger number of degrees of freedom. In Figure 6 we present experimental rates of convergence for $\|\bar{p}-\bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)}$: Ndof^{-1} for the two-dimensional examples and $\mathsf{Ndof}^{-2/3}$ for the three-dimensional example.

8.3. Conclusions

The following conclusions can be drawn:

– Most of the refinement occurs near to the interface of the control variable. When the domain involves geometric singularities, refinement is also being performed in regions that are close to them. These feature show a competitive performance of the devised *a posteriori* error estimator.

– All the individual contributions of the total error $|||e|||_\Omega$ exhibit optimal experimental rates of convergence for all the experiments and the nonlinear functions a considered in the experiments that we have performed.

– The devised *a posteriori* error estimator, defined in (5.11), is able to recognize the interface of $\bar{u}_{\mathscr{T}}$. This estimator also delivers, for all the numerical experiments that we have performed, optimal experimental rates of convergence for the involved individual contributions. This is not the case when the error estimator (8.2) is used in Algorithm 1.

– For some values of $\tau > 0$, we have provided consistent computational evidence that condition $\tilde{u} - \bar{u} \in C_{\bar{u}}^{\tau}$ is satisfied after achieving a suitable number of degrees of freedom and is uniform withing subsequently adaptive refinement. This suggests that the first assumption in (4.17) may hold.

– Numerical evidence shows that $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)}$ behaves as Ndof^{-1} and $\mathsf{Ndof}^{-2/3}$ in two and three dimensions, respectively. In particular, after a certain number of degrees of freedoms, we have that $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^\infty(\Omega)} < \tau/2$. This guarantees that $\tilde{u} - \bar{u} \in C_{\bar{u}}^{\tau}$; see Remark 4.6. In addition, this shows that $\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \nu\mu(2C_{\mathsf{M}})^{-1}$ uniformly after achieving a suitable number of degrees of freedom.

Acknowledgements. The first author is partially supported by ANID - Chile through FONDECYT project 1210729. The second author is supported by UTFSM through Beca de Mantención. The third author is partially supported by ANID - Chile through FONDECYT project 11180193.

References

[1] M. Ainsworth and J.T. Oden, A posteriori error estimation in finite element analysis, Pure and Applied Mathematics (New York), Wiley-Interscience, John Wiley & Sons, New York (2000).

[2] A. Allendes, F. Fuica and E. Otárola, Adaptive finite element methods for sparse PDE-constrained optimization. *IMA J. Numer. Anal.* **40** (2020) 2106–2142.

[3] P. Amestoy, I. Duff and J.-Y. L'Excellent, Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods Appl. Mech. Eng.* **184** (2000) 501–520.

[4] 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) 15–41.

[5] N. Arada, E. Casas and F. Tröltzsch, Error estimates for the numerical approximation of a semilinear elliptic control problem. *Comput. Optim. Appl.* **23** (2002) 201–229.

[6] K. Atkinson and W. Han, Theoretical Numerical Analysis, A functional analysis framework. Vol. 39 of *Texts in Applied Mathematics*, Springer-Verlag, New York (2001).

[7] R. Becker, H. Kapp and R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: basic concept. *SIAM J. Control Optim.* **39** (2000) 113–132.

[8] O. Benedix and B. Vexler, A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints. *Comput. Optim. Appl.* **44** (2009) 3–25.

[9] J.F. Bonnans and A. Shapiro, Perturbation Analysis of Optimization Problems, *Springer Series in Operations Research*, Springer-Verlag, New York, (2000).

[10] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, Vol. 15 of *Texts in Applied Mathematics*, Springer, New York, 3rd edition (2008).

[11] E. Casas, Using piecewise linear functions in the numerical approximation of semilinear elliptic control problems. *Adv. Comput. Math.* **26** (2007) 137–153.

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

[13] E. Casas and M. Mateos, Uniform convergence of the FEM. Applications to state constrained control problems, Vol. 21 (2002) 67–100. Special issue in memory of Jacques-Louis Lions.

[14] E. Casas and M. Mateos, Optimal Control of Partial Differential Equations, in Computational Mathematics, Numerical Analysis and Applications, Vol. 13 of *SEMA SIMAI Springer Ser.*, Springer, Cham (2017) 3–59.

[15] E. Casas, M. Mateos and F. Tröltzsch, Error estimates for the numerical approximation of boundary semilinear elliptic control problems. *Comput. Optim. Appl.* **31** (2005) 193–219.

[16] E. Casas and J.-P. Raymond, Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. *SIAM J. Control Optim.* **45** (2006) 1586–1611.

[17] E. Casas and F. Tröltzsch, Second order analysis for optimal control problems: improving results expected from abstract theory. *SIAM J. Optim.* **22** (2012) 261–279.

[18] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, Vol. 4 *Studies in Mathematics and its Applications*, North-Holland Publishing Co., Amsterdam-New York-Oxford (1978).

[19] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Vol. 24 of *Monographs and Studies in Mathematics*, Pitman (Advanced Publishing Program), Boston, MA (1985).

[20] M. Hintermüller and R.H.W. Hoppe, Goal-oriented adaptivity in control constrained optimal control of partial differential equations. *SIAM J. Control Optim.* **47** (2008) 1721–1743.

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

[22] D. Jerison and C.E. Kenig, The inhomogeneous Dirichlet problem in Lipschitz domains. *J. Funct. Anal.* **130** (1995) 161–219.

[23] D. Kinderlehrer and G. Stampacchia, An Introduction to Variational Inequalities and their Applications, Vol. 31 of *Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM)*, Philadelphia, PA (2000). Reprint of the 1980 original.

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

[25] W. Liu and N. Yan, A Posteriori Error Estimates for Distributed Convex Optimal Control Problems, *Adv. Comput. Math.* **15** (2001) 285–309.

[26] W. Liu and N. Yan, A Posteriori Error Estimates for Control Problems Governed by Nonlinear Elliptic Equations, Vol. 47, *2nd International Workshop on Numerical Linear Algebra, Numerical Methods for Partial Differential Equations and Optimization (Curitiba, 2001)* (2003) 173–187.

[27] R.H. Nochetto and A. Veeser, Primer of Adaptive Finite Element Methods, in Multiscale and Adaptivity: Modeling, Numerics and Applications, Vol. 2040 of *Lecture Notes in Math.*, Springer, Heidelberg (2012) 125–225.

[28] T. Roubíček, Nonlinear Partial Differential Equations with Applications, Vol. 153 of *International Series of Numerical Mathematics*, Birkhäuser/Springer Basel AG, Basel, 2nd edition (2013).

[29] G. Savaré, Regularity results for elliptic equations in Lipschitz domains. *J. Funct. Anal.* **152** (1998) 176–201.

[30] G. Stampacchia, Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier (Grenoble)* **15** (1965) 189–258.

[31] F. Tröltzsch, Optimal Control of Partial Differential Equations, Vol. 112 of *Graduate Studies in Mathematics*. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels. American Mathematical Society, Providence, RI (2010).

[32] R. Verfürth, A Posteriori Error Estimation Techniques for Finite Element Methods, *Numerical Mathematics and Scientific Computation*, Oxford University Press, Oxford (2013).

[33] B. Vexler and W. Wollner, Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Control Optim.* **47** (2008) 509–534.

[34] G. Wachsmuth and D. Wachsmuth, Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM: COCV* **17** (2011) 858–886.

[35] E. Zeidler, Nonlinear functional analysis and its applications. II/B. Nonlinear monotone operators, Translated from the German by the author and Leo F. Boron. Springer-Verlag, New York (1990).

Subscribe to Open (S2O)

A fair and sustainable open access model

This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: https://www.edpsciences.org/en/maths-s2o-programme