A Posteriori Error Estimates for an Optimal Control Problem with a Bilinear State Equation

Francisco Fuica[1] · **Enrique Otárola**[1]

Abstract We propose and analyze a posteriori error estimators for an optimal control problem that involves an elliptic partial differential equation as state equation and a control variable that enters the state equation as a coefficient; pointwise constraints on the control variable are considered as well. We consider two different strategies to approximate optimal variables: a fully discrete scheme in which the admissible control set is discretized with piecewise constant functions and a semi-discrete scheme where the admissible control set is not discretized; the latter scheme being based on the so-called variational discretization approach. We design, for each solution technique, an a posteriori error estimator and show, in two and three dimensional Lipschitz polygonal/polyhedral domains (not necessarily convex), that the proposed error estimator is reliable and efficient. We design, based on the devised estimators, adaptive strategies that deliver optimal experimental rates of convergence for the performed numerical examples.

✉ Enrique Otárola

enrique.otarola@usm.cl

[1] Departamento de Matemática, Universidad Técnica Federico Santa María, Av. España 1680, Valparaíso, Chile.

1 Introduction

The development and study of discretization techniques, based on finite elements, for distributed control–constrained linear–quadratic elliptic optimal control problems have been widely studied in the literature; see [16,15] for an extensive list of references. These discretization techniques are mainly divided into two categories, which rely on the discretization of the state and adjoint equations; they differ on whether or not the admissible control set is also discretized. In contrast to these advances, the study of solution techniques for optimal control problems where the control variable enters the state equation as a coefficient is not as developed. One of the main sources of difficulty within these type of problems is that the solution of the state equation depends nonlinearly on the control variable [18]. Consequently, uniqueness of solutions cannot be guaranteed.

In this work, we will focus on the development and analysis of efficient solution techniques for the optimal control problem (2)–(4), which incorporates the control variable as a coefficient in the state equation; the control variable is not a source term. We immediately mention that this problem can be interpreted as a particular instance of parameter estimation. To the best of our knowledge, the first work that provides an analysis for suitable finite element discretizations for problem (2)–(4) is [18]. In this work, the authors propose, on convex polygonal/polyhedral domains and quasi–uniform meshes, two fully discrete schemes that discretize the admissible control set with piecewise constant and piecewise linear functions; the state and adjoint equations are discretized with piecewise linear functions. Estimates for the error committed within the approximation of a control variable are derived in [18, Corollaries 5.6 and 5.10]. In addition, an error estimate for a post-processing strategy is obtained in [18, Theorem 5.18]. The results obtained in [18] were later extended to mixed and stabilized finite element methods in [8] and [11], respectively.

A particular class of numerical methods that has proven a competitive performance when are used to approximate solutions to PDE–constrained optimization problems, and the ones we shall consider in this work, are adaptive finite element methods (AFEMs). AFEMs are iterative methods recognizable by their capability to improve the quality of a discrete approximation to a corresponding PDE while keeping an efficient distribution of computational resources. A crucial component of an AFEM is an a posteriori error estimator, which is a computable quantity, depending on the problem data and discrete solution, that provides local information about the quality of the approximate solution. The a posteriori error analysis for control–constrained linear–quadratic optimal control problems has achieved several advances in recent years. We refer the interested reader to [20,5,13,17,12,2] for a discussion. As opposed to these advances, the analysis of AFEMs for optimal control problems involving nonlinear or bilinear equations is rather scarce. To the best of our knowledge, the work [19] appears to be the first that provides a posteriori error estimates for (2)–(4). In this work, the authors develop a posteriori error estimators for two fully discrete approximation schemes of (2)–(4) and obtain

global reliability estimates [19, Theorems 4.1 and 4.3] and *global* efficiency results [19, Lemmas 4.2 and 4.3]. We also mention the work [7], where a posteriori error estimates for a parabolic version of (2)–(4) have been analyzed; an efficiency analysis, however, was not provided. We conclude this paragraph by mentioning the work [24], where the authors provide, on the basis of a posteriori error estimators, upper bounds for discretization errors with respect to a cost functional and with respect to a given quantity of interest; the latter being an arbitrary functional depending on the control and the state variables. In our work, we derive upper and lower bounds for the approximation error when is measured in an *energy* norm; see below for a discussion.

In the present manuscript, we consider two different strategies to discretize the optimal control problem (2)–(4): a semi-discrete scheme, based on the so-called variational discretization approach [14], in which the admissible control set is not discretized, and a fully discrete scheme, where control variables are approximated by using piecewise constant functions. We devise, for each one of the aforementioned schemes, a residual–based a posteriori error estimator. For the fully discrete scheme the error estimator is formed by the sum of three contributions: two of them are related to the discretization of the state and adjoint equations while the remaining one is related to the discretization of the admissible control set. In contrast, the error estimator for the variational discretization approach is formed by only two contributions that are related to the discretization of the state and adjoint equations. In two and three dimensional Lipschitz polygonal/polyhedral domains (not necessarily convex), we obtain reliability and efficiency estimates.

In what follows we list what, we believe, are the main contributions of our work:

- For the fully and semi-discrete schemes that we consider, we devise a posteriori error estimators; both being different from the ones in [19] and [24].
- For the aforementioned solution techniques, we prove that the corresponding local error indicators associated to the discretization of the state and adjoint equations are *locally efficient*. This analysis improves the global one in [19, Lemmas 4.2 and 4.3]. We also prove that the total error indicator associated to the variational discretization approach is *locally efficient* (cf. Theorem 6.2); the one associated to the fully discrete scheme, as customary, being globally efficient (cf. Theorem 5.5).
- We design a simple adaptive loop that delivers optimal experimental rates of convergence for all the involved individual contributions of the corresponding error. The loop based on the a posteriori error indicators devised for the variational discretization approach delivers quadratic rates of convergence for the error approximation of a control variable. This substantially improves the approximation properties that can be achieved by the considered fully discrete scheme. The indicators devised for the latter scheme tend to refine the involved meshes in regions where the restrictions of the control variable become active. These DOFs seem not necessary for an accurate approximation of control variables. An scheme based on piece-

wise linear approximation of the admissible control set would suffer the same limitations in terms of avoidable refinement [2].

The rest of the paper is organized as follows. In section 2 we introduce the optimal control problem under consideration and set notation. Basic results for the state equation as well as basic a posteriori error estimates are reviewed in section 3. In section 4 we review the existence of solutions for the optimal control problem as well as first and second order optimality conditions. The crucial part of our work are sections 5 and 6, where we design and analyze a posteriori error estimators for the fully and semi-discrete schemes, respectively. Finally, in section 7 we present numerical examples in two and three dimensional domains that illustrate the theory and reveal a competitive performance of the devised AFEMs.

2 The Problem and Notation

Let us precisely introduce the optimal control problem that will be considered in our work, set notation, and describe the setting we shall operate with.

2.1 Presentation of the Problem

In this work we are interested in the design and analysis of a posteriori error estimates for an optimal control problem governed by an elliptic partial differential equation (PDE) as state equation. Our main source of difficulty here is that the control variable enters the state equation as a coefficient; control constraints are also considered. Let us make this discussion precise. Let $\Omega \subset \mathbb{R}^d$, with $d \in \{2, 3\}$, be an open and bounded polygonal/polyhedral domain with Lipschitz boundary $\partial\Omega$ [1, Chapter 4]. Given a desired state $y_\Omega \in L^2(\Omega)$ and a regularization parameter $\alpha > 0$, let us introduce the cost functional

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

We are thus interested in the following optimal control problem: Find

$$\min J(y, u) \tag{2}$$

subject to the elliptic PDE

$$-\Delta y + uy = f \text{ in } \Omega, \qquad y = 0 \text{ on } \partial\Omega, \tag{3}$$

where $f \in L^2(\Omega)$ denotes an external source, and the control constraints

$$u \in \mathbb{U}_{ad}, \qquad \mathbb{U}_{ad} := \{v \in L^2(\Omega) : 0 < \mathsf{a} \le v \le \mathsf{b} \text{ a.e. in } \Omega\}. \tag{4}$$

Here, $\mathsf{a}, \mathsf{b} \in \mathbb{R}^+$ satisfy $\mathsf{a} < \mathsf{b}$.

2.2 Notation

Let us set notation and describe the setting we shall operate with. Throughout this work $d \in \{2, 3\}$ and $\Omega \subset \mathbb{R}^d$ is an open and bounded polygonal/polyhedral domain with Lipschitz boundary $\partial\Omega$ [1, Chapter 4]. Notice that we do not assume that Ω is convex. If $\mathcal{X}$ and $\mathcal{Y}$ are normed vector spaces, we write $\mathcal{X} \hookrightarrow \mathcal{Y}$ to denote that $\mathcal{X}$ is continuously embedded in $\mathcal{Y}$. We denote by $\| \cdot \|_{\mathcal{X}}$ the norm of $\mathcal{X}$. The relation $a \lesssim b$ indicates that $a \leq Cb$, with a positive constant that depends neither on a, b nor on the involved discretization parameters. The value of C might change at each occurrence.

3 The State Equation

In this section, we briefly review some results related to the well-posedness of problem (3). Additionally, we present a posteriori error estimates for a specific finite element setting.

3.1 Weak Formulation

Let $\mathfrak{f}$ be a given forcing term in $L^2(\Omega)$ and $\mathfrak{u}$ be an arbitrary function in $\mathbb{U}_{ad}$. With this setting at hand, we introduce the following weak problem:

$$z \in H_0^1(\Omega): \quad (\nabla z, \nabla v)_{L^2(\Omega)} + (\mathfrak{u}z, v)_{L^2(\Omega)} = (\mathfrak{f}, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega). \quad (5)$$

Lax–Milgram Theorem immediately yields the well-posedness of problem (5). In particular, we have the following stability estimate $\|\nabla z\|_{L^2(\Omega)} \lesssim \|\mathfrak{f}\|_{L^2(\Omega)}$.

3.2 Finite Element Approximation

In this section, we introduce a basic finite element approximation for the weak problem (5) and review basic a posteriori error estimates. To accomplish this task, we first introduce some terminology and further basic ingredients.

We denote by $\mathscr{T} = \{T\}$ a conforming partition of $\overline{\Omega}$ into simplices T with size $h_T = \operatorname{diam}(T)$ and define $h_{\mathscr{T}} := \max_{T \in \mathscr{T}} h_T$. We denote by $\mathscr{S}$ the set of *internal* $(d-1)$-dimensional interelement boundaries S of $\mathscr{T}$. For $T \in \mathscr{T}$, we let $\mathscr{S}_T$ denote the subset of $\mathscr{S}$ which contains the sides of the element T. We denote by $\mathcal{N}_S \subset \mathscr{T}$ the subset that contains the two elements that have S as a side, namely, $\mathcal{N}_S = \{T^+, T^-\}$, where $T^+, T^- \in \mathscr{T}$ are such that $S = T^+ \cap T^-$. For $T \in \mathscr{T}$, we define the *star* associated with the element T as

$$\mathcal{N}_T := \{T' \in \mathscr{T} : \mathscr{S}_T \cap \mathscr{S}_{T'} \neq \emptyset\}. \quad (6)$$

In an abuse of notation, below we denote by $\mathcal{N}_T$ either the set itself or the union of its elements.

We define, for $T \in \mathscr{T}$, the *shape coefficient* σ_T of T as the ratio of the diameter, i.e., $\mathrm{diam}(T) = h_T$, and the inball diameter of T, i.e., $2\sup\{r > 0 : B_r(x) \subset T \text{ for } x \in T\}$. The *shape coefficient* of a triangulation $\mathscr{T}$ corresponds to the quantity $\sigma_\mathscr{T} := \max\{T : T \in \mathscr{T}\}$. A sequence of triangulations $\mathbb{T} = \{\mathscr{T}\}$, that is obtained by subsequent refinements of an initial mesh $\mathscr{T}_0$, is *shape regular* if $\sup\{\sigma_\mathscr{T} : \mathscr{T} \in \mathbb{T}\} \leq C$; see [21, section 3.2.1] for details.

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

$$\mathbb{V}(\mathscr{T}) := \{v_\mathscr{T} \in C(\overline{\Omega}) : v_\mathscr{T}|_T \in \mathbb{P}_1(T) \ \forall T \in \mathscr{T}\} \cap H_0^1(\Omega). \qquad (7)$$

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. Here, $T^+, T^- \in \mathscr{T}$ are such that $T^+ \neq T^-$ and $\partial T^+ \cap \partial T^- = S$.

With these ingredients at hand, we introduce a Galerkin approximation to problem (5) as follows:

$$z_\mathscr{T} \in \mathbb{V}(\mathscr{T}): \quad (\nabla z_\mathscr{T}, \nabla v_\mathscr{T})_{L^2(\Omega)} + (\mathfrak{u} z_\mathscr{T}, v_\mathscr{T})_{L^2(\Omega)} = (\mathfrak{f}, v_\mathscr{T})_{L^2(\Omega)} \qquad (8)$$

for all $v_\mathscr{T} \in \mathbb{V}(\mathscr{T})$. Here, $\mathfrak{f} \in L^2(\Omega)$ and $\mathfrak{u} \in \mathbb{U}_{ad}$. The existence and uniqueness of a solution $z_\mathscr{T} \in \mathbb{V}(\mathscr{T})$ of problem (8) is standard. In particular, we have the stability estimate $\|\nabla z_\mathscr{T}\|_{L^2(\Omega)} \lesssim \|\mathfrak{f}\|_{L^2(\Omega)}$.

3.3 An a Posteriori Error Estimate for the State Equation

We introduce the following local error indicators and a posteriori error estimator associated to the discretization (8) of problem (5):

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

We present the following global reliability result.

Theorem 3.1 (Global reliability of $\mathcal{E}$) *Let $\mathfrak{f} \in L^2(\Omega)$ and $\mathfrak{u} \in \mathbb{U}_{ad}$ be given. Let $z \in H_0^1(\Omega)$ be the unique solution to problem (5) and let $z_\mathscr{T} \in \mathbb{V}(\mathscr{T})$ be its finite element approximation obtained as the solution to (8). We thus have*

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

with a hidden constant that is independent of z, $z_\mathscr{T}$, the size of the elements in $\mathscr{T}$, and $\#\mathscr{T}$ but depends on the shape coefficient of the triangulation $\mathscr{T}$, i.e., $\sigma_\mathscr{T}$, and the dimension d.

Proof Since z solves (5), we invoke Galerkin orthogonality and an elementwise integration by parts formula to arrive at

$$(\nabla(z - z_{\mathscr{T}}), \nabla v)_{L^2(\Omega)} + (\mathfrak{u}(z - z_{\mathscr{T}}), v)_{L^2(\Omega)}$$
$$= \sum_{T \in \mathscr{T}} \int_T (\mathfrak{f} - \mathfrak{u} z_{\mathscr{T}})(v - I_{\mathscr{T}} v) \mathrm{d}x + \sum_{S \in \mathscr{S}} \int_S [\![\nabla z_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!] (v - I_{\mathscr{T}} v) \mathrm{d}x.$$

Here, $v \in H_0^1(\Omega)$ and $I_{\mathscr{T}} : L^1(\Omega) \to \mathbb{V}(\mathscr{T})$ denotes the Clément interpolation operator [6,9]. Standard approximation properties for $I_{\mathscr{T}}$ and the finite overlapping property of stars allow us to derive

$$(\nabla(z - z_{\mathscr{T}}), \nabla v)_{L^2(\Omega)} + (\mathfrak{u}(z - z_{\mathscr{T}}), v)_{L^2(\Omega)}$$
$$\lesssim \left[\sum_{T \in \mathscr{T}} h_T^2 \|\mathfrak{f} - \mathfrak{u} z_{\mathscr{T}}\|_{L^2(T)}^2 + h_T \|[\![\nabla z_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T \setminus \partial \Omega)}^2 \right]^{\frac{1}{2}} \|\nabla v\|_{L^2(\Omega)}.$$

Set $v = z - z_{\mathscr{T}} \in H_0^1(\Omega)$ and use the fact that $\mathfrak{u} > 0$ to conclude. $\qquad\square$

4 The Optimal Control Problem

In this section, we follow [18, section 2] and introduce a weak formulation for the optimal control problem (2)–(4). In addition, we review first and second order optimality conditions and introduce finite element discretization schemes.

4.1 Weak Formulation and Existence of a Solution

Let J be the cost functional defined in (1). We consider the following weak version of the optimization problem (2)–(4): Find

$$\min\{J(y, u) : (y, u) \in H_0^1(\Omega) \times \mathbb{U}_{ad}\} \tag{9}$$

subject to the state equation

$$(\nabla y, \nabla v)_{L^2(\Omega)} + (uy, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega). \tag{10}$$

The existence of an optimal solution $(\bar{y}, \bar{u}) \in H_0^1(\Omega) \times \mathbb{U}_{ad}$ for problem (9)–(10) follows standard arguments; see [18, Proposition 2.3].

4.2 Optimality Conditions

Due to the fact that the optimal control problem (9)–(10) is not convex, we discuss optimality conditions under the framework of local solutions in $L^2(\Omega)$. To be precise, a control $\bar{u} \in \mathbb{U}_{ad}$ is said to be locally optimal in $L^2(\Omega)$ for (9)–(10) if there exists a constant $\delta > 0$ such that $J(\bar{y}, \bar{u}) \le J(y, u)$ for all $u \in \mathbb{U}_{ad}$ such that $\|u - \bar{u}\|_{L^2(\Omega)} \le \delta$. Here, $\bar{y}$ and y denote the states associated to $\bar{u}$ and u, respectively.

Let us introduce the set $\mathcal{U} := \{u \in L^\infty(\Omega) : \exists c > 0 \text{ such that } u(x) > c > 0 \text{ a.e. } x \in \Omega\}$. We immediately notice that $\mathbb{U}_{ad} \subset \mathcal{U}$. Having defined $\mathcal{U}$, we introduce the control-to-state map $\mathcal{S}$ as follows: given a control $u \in \mathcal{U}$, $\mathcal{S}$ associates to it a unique state $y = \mathcal{S}u \in H_0^1(\Omega)$ solving (10). With these ingredients at hand, we define the reduced cost functional $j : \mathcal{U} \to \mathbb{R}_0^+$ by

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

We are now in position to formulate first order optimality conditions: if $\bar{u}$ is locally optimal for problem (9)–(10), then [18, Proposition 2.10]

$$j'(\bar{u})(u - \bar{u}) \ge 0 \quad \forall u \in \mathbb{U}_{ad}. \tag{11}$$

Here, $j'(\bar{u})$ denotes the Gateâux derivative of the functional j at $\bar{u}$ in the direction $u - \bar{u}$. We notice that, for $u, v \in \mathbb{U}_{ad}$, $j'(u)v = (\alpha u - yp, v)_{L^2(\Omega)}$ [18, equation (2.6)] and immediately comment that $\mathcal{S}$ and j are not Fréchet differentiable with respect to the $L^2(\Omega)$-topology [18, Remark 2.8]. To investigate the inequality (11), we introduce the *adjoint variable* $p \in H_0^1(\Omega)$ as the unique solution to the *adjoint equation*

$$(\nabla w, \nabla p)_{L^2(\Omega)} + (up, w)_{L^2(\Omega)} = (y - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega), \tag{12}$$

where $y = \mathcal{S}u$ solves (10). Observe that problem (12) is well-posed.

With the previous ingredients at hand, we reformulate first order optimality conditions as follows; see [18, Proposition 2.10 and equation (2.6)].

Theorem 4.1 (First order optimality conditions) *Every locally optimal control $\bar{u} \in \mathbb{U}_{ad}$ for problem (9)–(10) satisfies, together with the state $\bar{y} \in H_0^1(\Omega)$ and the adjoint state $\bar{p} \in H_0^1(\Omega)$, the variational inequality*

$$(\alpha\bar{u} - \bar{y}\bar{p}, u - \bar{u})_{L^2(\Omega)} \ge 0 \quad \forall u \in \mathbb{U}_{ad}. \tag{13}$$

Here, $\bar{p}$ denotes the solution to (12) *with y replaced by $\bar{y} = \mathcal{S}\bar{u}$.*

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

$$\Pi_{[\mathtt{a},\mathtt{b}]}(v) := \min\{\mathtt{b}, \max\{v, \mathtt{a}\}\} \text{ a.e. in } \Omega. \tag{14}$$

This operator allows us to present the following projection formula [18, equation (2.7)]: If $\bar{u}$ denotes a locally optimal control for (9)–(10), then

$$\bar{u}(x) := \Pi_{[\mathtt{a},\mathtt{b}]}(\alpha^{-1}\bar{y}(x)\bar{p}(x)) \text{ a.e. } x \in \Omega. \tag{15}$$

Let $\bar{u} \in \mathbb{U}_{ad}$ be a control that satisfies the first order necessary optimality condition (13). In what follows, we will assume that there exists a constant $\mu > 0$ such that

$$j''(\bar{u})v^2 \geq \mu \|v\|^2_{L^2(\Omega)} \quad \forall v \in L^\infty(\Omega); \tag{16}$$

see [18, Assumption 2.20]. Here, for each $v \in L^\infty(\Omega)$, we have that

$$j''(\bar{u})v^2 := \|\mathcal{S}'(\bar{u})v\|^2_{L^2(\Omega)} + (\bar{y} - y_\Omega, \mathcal{S}''(\bar{u})v^2)_{L^2(\Omega)} + \alpha\|v\|^2_{L^2(\Omega)},$$

where $\mathcal{S}'(\bar{u})v$ and $\mathcal{S}''(\bar{u})v^2$ are defined as in [18, Lemma 2.9]. We notice that assumption (16) is fulfilled if $\|\mathcal{S}\bar{u} - y_\Omega\|_{L^2(\Omega)}$ is sufficiently small or $\alpha > 0$ is sufficiently large; see [18, Remark 2.21] for details.

The following result states that every control $\bar{u}$ that satisfies (13) and (16) is a local solution for problem (9)–(10); see [18, Theorem 2.24].

Theorem 4.2 (Local optimality) *Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (9)–(10) satisfying the necessary and sufficient optimality conditions (13) and (16). Then, there exist positive constants $\delta, \sigma > 0$ such that*

$$j(u) \geq j(\bar{u}) + \sigma\|u - \bar{u}\|^2_{L^2(\Omega)} \quad \forall u \in \mathbb{U}_{ad} \cap B_\delta(\bar{u}),$$

where $B_\delta(\bar{u})$ denotes the closed ball in $L^2(\Omega)$ with center at $\bar{u}$ and radius δ.

We conclude this section with the following estimate [18, Proposition 2.22]: Let $u, v \in \mathbb{U}_{ad}$ and $w \in L^\infty(\Omega)$. Then, there exists $\mathfrak{C} > 0$, depending on $\|f\|_{L^2(\Omega)}$ and $\|y_\Omega\|_{L^2(\Omega)}$, such that

$$|j''(u)w^2 - j''(v)w^2| \leq \mathfrak{C}\|u - v\|_{L^2(\Omega)}\|w\|^2_{L^2(\Omega)}. \tag{17}$$

4.3 Finite Element Approximation

In this section, we introduce two finite element discretization schemes for our optimal control problem.

4.3.1 The Fully Discrete Scheme

To approximate a control variable, we introduce the space of piecewise constant functions

$$\mathbb{U}(\mathscr{T}) := \{u_\mathscr{T} \in L^\infty(\Omega) : u_\mathscr{T}|_T \in \mathbb{P}_0(T) \ \forall T \in \mathscr{T}\}$$

and define the discrete admissible set $\mathbb{U}_{ad}(\mathscr{T}) := \mathbb{U}(\mathscr{T}) \cap \mathbb{U}_{ad}$. The state and adjoint state variables, associated to a locally optimal control, are discretized by using the finite element space $\mathbb{V}(\mathscr{T})$ defined in (7). With this setting at hand, the fully discrete scheme reads as follows: Find $\min J(y_\mathscr{T}, u_\mathscr{T})$ subject to the discrete state equation

$$y_\mathscr{T} \in \mathbb{V}(\mathscr{T}) : \quad (\nabla y_\mathscr{T}, \nabla v_\mathscr{T})_{L^2(\Omega)} + (u_\mathscr{T} y_\mathscr{T}, v_\mathscr{T})_{L^2(\Omega)} = (f, v_\mathscr{T})_{L^2(\Omega)} \tag{18}$$

for all $v_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ and the discrete constraints $u_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T})$. The fully discrete scheme admits at least a solution; see [18, Section 3] for details. In addition, if $\bar{u}_{\mathscr{T}}$ denotes a discrete local solution, then

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

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

$$(\nabla w_{\mathscr{T}}, \nabla \bar{p}_{\mathscr{T}})_{L^2(\Omega)} + (\bar{u}_{\mathscr{T}} \bar{p}_{\mathscr{T}}, w_{\mathscr{T}})_{L^2(\Omega)} = (\bar{y}_{\mathscr{T}} - y_{\Omega}, w_{\mathscr{T}})_{L^2(\Omega)} \qquad (19)$$

for all $w_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$; see [18, equations (3.6) and (3.7)].

4.3.2 The Semi-discrete Scheme

In this section, we introduce the so-called variational discretization approach for (9)–(10). This scheme discretizes only the state space; the control space $\mathbb{U}_{ad}$ is not discretized. The scheme induces a discretization of an optimal control variable by projecting, in view of the operator introduced in (14), an optimal discrete adjoint state into $\mathbb{U}_{ad}$. The semi-discrete scheme is defined as follows: Find $\min J(y_{\mathscr{T}}, \mathsf{u})$ subject to the discrete state equation

$$y_{\mathscr{T}} \in \mathbb{V}(\mathscr{T}): \quad (\nabla y_{\mathscr{T}}, \nabla v_{\mathscr{T}})_{L^2(\Omega)} + (\mathsf{u} y_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)} = (f, v_{\mathscr{T}})_{L^2(\Omega)} \qquad (20)$$

for all $v_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ and the constraints $\mathsf{u} \in \mathbb{U}_{ad}$. As in the fully discrete case, this problem admits at least a solution and, if $\bar{\mathsf{u}}$ denotes a local solution, then

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

where $\bar{p}_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ solves

$$(\nabla w_{\mathscr{T}}, \nabla \bar{p}_{\mathscr{T}})_{L^2(\Omega)} + (\bar{\mathsf{u}} \bar{p}_{\mathscr{T}}, w_{\mathscr{T}})_{L^2(\Omega)} = (\bar{y}_{\mathscr{T}} - y_{\Omega}, w_{\mathscr{T}})_{L^2(\Omega)} \qquad (21)$$

for all $w_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$. Here, $\bar{y}_{\mathscr{T}} = \bar{y}_{\mathscr{T}}(\bar{\mathsf{u}})$ solves (20) with $\mathsf{u} = \bar{\mathsf{u}}$.

5 A Posteriori Error Analysis for the Fully Discrete Scheme

In this section, we devise and analyze an a posteriori error estimator for the fully discrete scheme. The error estimator will be formed by the sum of three contributions: two contributions related to the discretization of the state and adjoint equations and a one contribution associated to the discretization of the admissible control set $\mathbb{U}_{ad}$.

To begin with our studies we introduce, on the basis of the projection operator $\Pi_{[\mathsf{a},\mathsf{b}]}$ defined in (14), the auxiliary variable

$$\tilde{u} := \Pi_{[\mathsf{a},\mathsf{b}]} \left(\alpha^{-1} \bar{y}_{\mathscr{T}} \bar{p}_{\mathscr{T}} \right). \qquad (22)$$

A key property in favor of the definition of $\tilde{u}$ is that it satisfies the following variational inequality [22, Lemma 2.26]:

$$(\alpha \tilde{u} - \bar{y}_{\mathscr{T}} \bar{p}_{\mathscr{T}}, u - \tilde{u})_{L^2(\Omega)} \geq 0 \quad \forall u \in \mathbb{U}_{ad}. \qquad (23)$$

With the variable $\tilde{u}$ at hand, we present the following result which is instrumental for our a posteriori error analysis.

Theorem 5.1 (Auxiliary estimate) *Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (9)–(10) satisfying the sufficient second order optimality condition (16). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the fully discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. If $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ satisfy, on the mesh $\mathscr{T}$, the bound*

$$\|\bar{y}\bar{p} - \bar{y}_{\mathscr{T}}\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} \leq \alpha\mu(2\mathfrak{C})^{-1}, \tag{24}$$

then

$$\frac{\mu}{2}\|\bar{u} - \tilde{u}\|^2_{L^2(\Omega)} \leq (j'(\tilde{u}) - j'(\bar{u}))(\tilde{u} - \bar{u}). \tag{25}$$

The constants μ and $\mathfrak{C}$ are given as in (16) and (17), respectively.

Proof Since $\tilde{u} - \bar{u} \in L^\infty(\Omega)$, we are allowed to set $v = \tilde{u} - \bar{u}$ in the second order optimality condition (16). This yields

$$\mu\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)} \leq j''(\bar{u})(\tilde{u} - \bar{u})^2. \tag{26}$$

On the other hand, in view of the mean value theorem, we obtain $(j'(\tilde{u}) - j'(\bar{u}))(\tilde{u} - \bar{u}) = j''(\zeta)(\tilde{u} - \bar{u})^2$, where $\zeta = \bar{u} + \theta_{\mathscr{T}}(\tilde{u} - \bar{u})$ and $\theta_{\mathscr{T}} \in (0, 1)$. Inequality (26) thus yields

$$\mu\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)} \leq (j'(\tilde{u}) - j'(\bar{u}))(\tilde{u} - \bar{u}) + (j''(\bar{u}) - j''(\zeta))(\tilde{u} - \bar{u})^2. \tag{27}$$

Let us now concentrate on the second term on the right-hand side of inequality (27). To accomplish this task, let us first invoke (17) to arrive at

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

where we have also used that $\theta_{\mathscr{T}} \in (0, 1)$. Invoke (15) and (22), the Lipschitz property of the projection operator $\Pi_{[\mathbf{a},\mathbf{b}]}$, and assumption (24) to conclude

$$(j''(\bar{u}) - j''(\zeta))(\tilde{u} - \bar{u})^2 \leq \mathfrak{C}\alpha^{-1}\|\bar{y}\bar{p} - \bar{y}_{\mathscr{T}}\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)}\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)} \leq \frac{\mu}{2}\|\tilde{u} - \bar{u}\|^2_{L^2(\Omega)}.$$

Replacing this inequality into (27) yields the desired inequality (25). $\qquad\square$

Remark 5.1 (A sufficient condition for estimate (24)) An estimate that guarantees assumption (24) reads as follows:

$$\|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} + \|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \leq \alpha\mu(2\mathfrak{C}C_{cond})^{-1},$$

where C_{cond} is defined as in (28). In fact, let $q \in \{2, 4\}$ and let C_q be the best constant associated to the embedding $H_0^1(\Omega) \hookrightarrow L^q(\Omega)$, i.e., C_q is the best constant such that $\|v\|_{L^q(\Omega)} \leq C_q\|\nabla v\|_{L^2(\Omega)}$ for all $v \in H_0^1(\Omega)$. Hence, an application of the Cauchy–Schwarz inequality, the aforementioned Sobolev

embeddings, and the stability of the adjoint and discrete state equations allow us to obtain the bounds

$$
\begin{aligned}
\|\bar{y}\bar{p} - \bar{y}_{\mathscr{T}}\bar{p}_{\mathscr{T}}\|_{L^2(\Omega)} &\leq \|\bar{p}\|_{L^4(\Omega)}\|\bar{y} - \bar{y}_{\mathscr{T}}\|_{L^4(\Omega)} + \|\bar{y}_{\mathscr{T}}\|_{L^4(\Omega)}\|\bar{p} - \bar{p}_{\mathscr{T}}\|_{L^4(\Omega)} \\
&\leq C_4^2 \left(\|\nabla\bar{p}\|_{L^2(\Omega)}\|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} + \|\nabla\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}\|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \right) \\
&\leq C_4^2 C_2 \left[(C_2^2\|f\|_{L^2(\Omega)} + \|y_\Omega\|_{L^2(\Omega)})\|\nabla(\bar{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \right. \\
&\qquad\qquad\qquad\qquad \left. + \|f\|_{L^2(\Omega)}\|\nabla(\bar{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \right].
\end{aligned}
$$

The desired result can thus be concluded by setting

$$
C_{cond} := C_4^2 C_2 \max\{C_2^2\|f\|_{L^2(\Omega)} + \|y_\Omega\|_{L^2(\Omega)}, \|f\|_{L^2(\Omega)}\}. \tag{28}
$$

5.1 Global Reliability Analysis

The goal of this section is to derive an upper bound for the corresponding total error in terms of a devised a posteriori error estimator. The analysis relies on estimates on the error between a solution to the fully discrete optimal control problem (18)–(19) and auxiliary variables that we define in what follows.

We first define the variable $\hat{y} \in H_0^1(\Omega)$ as the solution to

$$
(\nabla\hat{y}, \nabla v)_{L^2(\Omega)} + (\bar{u}_{\mathscr{T}}\hat{y}, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega). \tag{29}
$$

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

$$
\mathcal{E}_{st,T}^2 := h_T^2 \|f - \bar{u}_{\mathscr{T}}\bar{y}_{\mathscr{T}}\|_{L^2(T)}^2 + h_T \|[\![\nabla\bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T\backslash\partial\Omega)}^2, \tag{30}
$$

and the global a posteriori error estimator associated to the finite element discretization of the state equation

$$
\mathcal{E}_{st,\mathscr{T}} := \left(\sum_{T\in\mathscr{T}} \mathcal{E}_{st,T}^2 \right)^{\frac{1}{2}}. \tag{31}
$$

An application of Theorem 3.1, with $\mathfrak{f} = f$ and $\mathfrak{u} = \bar{u}_{\mathscr{T}}$, immediately yields the a posteriori error estimate

$$
\|\nabla(\hat{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}}. \tag{32}
$$

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

$$
(\nabla w, \nabla\hat{p})_{L^2(\Omega)} + (\bar{u}_{\mathscr{T}}\hat{p}, w)_{L^2(\Omega)} = (\bar{y}_{\mathscr{T}} - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega). \tag{33}
$$

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

$$
\mathcal{E}_{adj,T}^2 := h_T^2 \|\bar{y}_{\mathscr{T}} - y_\Omega - \bar{u}_{\mathscr{T}}\bar{p}_{\mathscr{T}}\|_{L^2(T)}^2 + h_T \|[\![\nabla\bar{p}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T\backslash\partial\Omega)}^2, \tag{34}
$$

and the a posteriori error estimator associated to the discretization of the adjoint equation

$$
\mathcal{E}_{adj,\mathscr{T}} := \left(\sum_{T\in\mathscr{T}} \mathcal{E}_{adj,T}^2 \right)^{\frac{1}{2}}. \tag{35}
$$

An application of Theorem 3.1, again, with $\mathfrak{f} = \bar{y}_{\mathscr{T}} - y_\Omega$ and $\mathfrak{u} = \bar{u}_{\mathscr{T}}$, immediately yields the a posteriori error bound

$$\|\nabla(\hat{p} - \bar{p}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{adj,\mathscr{T}}. \tag{36}$$

Finally, we introduce local error indicators and an a posteriori error estimator associated to the discretization of a control variable. To be precise, we define, on the basis of the auxiliary variable $\tilde{u}$, defined in (22),

$$\mathcal{E}_{ct,T}^2 := \|\tilde{u} - \bar{u}_{\mathscr{T}}\|_{L^2(T)}^2, \qquad \mathcal{E}_{ct,\mathscr{T}} := \left(\sum_{T \in \mathscr{T}} \mathcal{E}_{ct,T}^2 \right)^{\frac{1}{2}}. \tag{37}$$

After having defined error estimators associated to the discretization of the state and adjoint equations and the admissible control set, we introduce an a posteriori error estimator for the fully discrete problem that approximates solutions to problem (9)–(10). The error estimator can be decomposed as the sum of three contributions:

$$\mathcal{E}_{ocp,\mathscr{T}}^2 := \mathcal{E}_{st,\mathscr{T}}^2 + \mathcal{E}_{adj,\mathscr{T}}^2 + \mathcal{E}_{ct,\mathscr{T}}^2. \tag{38}$$

The estimators $\mathcal{E}_{st,\mathscr{T}}$, $\mathcal{E}_{adj,\mathscr{T}}$, and $\mathcal{E}_{ct,\mathscr{T}}$ are defined as in (31), (35), and (37), respectively.

We are now ready to state and prove the main result of this section. As a final ingredient, we introduce $e_{\bar{y}} := \bar{y} - \bar{y}_{\mathscr{T}}$, $e_{\bar{p}} := \bar{p} - \bar{p}_{\mathscr{T}}$, and $e_{\bar{u}} := \bar{u} - \bar{u}_{\mathscr{T}}$.

Theorem 5.2 (Global reliability) *Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (9)–(10) satisfying the sufficient second order condition (16). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the fully discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. If $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ satisfy, on the mesh $\mathscr{T}$, the bound (24), then*

$$\|\nabla e_{\bar{y}}\|_{L^2(\Omega)}^2 + \|\nabla e_{\bar{p}}\|_{L^2(\Omega)}^2 + \|e_{\bar{u}}\|_{L^2(\Omega)}^2 \lesssim \mathcal{E}_{ocp,\mathscr{T}}^2, \tag{39}$$

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

Proof We proceed on the basis of four steps.

Step 1. The goal of this step is to control the term $\|e_{\bar{u}}\|_{L^2(\Omega)}$. To accomplish this task, we invoke the auxiliary variable $\tilde{u} = \Pi_{[\mathsf{a},\mathsf{b}]}(\alpha^{-1}\bar{y}_{\mathscr{T}}\bar{p}_{\mathscr{T}})$, a triangle inequality, and both definitions in (37) to arrive at

$$\|e_{\bar{u}}\|_{L^2(\Omega)} \leq \|\bar{u} - \tilde{u}\|_{L^2(\Omega)} + \mathcal{E}_{ct,\mathscr{T}}. \tag{40}$$

Let us now concentrate on the first term on the right-hand side of the previous estimate. Set $u = \tilde{u}$ in (11) and $u = \bar{u}$ in (23) to obtain, respectively,

$$j'(\bar{u})(\tilde{u} - \bar{u}) \geq 0, \qquad (\alpha\tilde{u} - \bar{p}_{\mathscr{T}}\bar{y}_{\mathscr{T}}, \bar{u} - \tilde{u})_{L^2(\Omega)} \geq 0.$$

With these estimates at hand, we invoke inequality (25) to arrive at

$$\frac{\mu}{2}\|\bar{u} - \tilde{u}\|_{L^2(\Omega)}^2 \leq j'(\tilde{u})(\tilde{u} - \bar{u}) - j'(\bar{u})(\tilde{u} - \bar{u}) \leq j'(\tilde{u})(\tilde{u} - \bar{u}) \qquad (41)$$

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

where the auxiliary variables $\tilde{y}, \tilde{p} \in H_0^1(\Omega)$ are defined as follows:

$$\tilde{y}: \quad (\nabla\tilde{y}, \nabla v)_{L^2(\Omega)} + (\tilde{u}\tilde{y}, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega), \qquad (42)$$

and

$$\tilde{p}: \quad (\nabla w, \nabla\tilde{p})_{L^2(\Omega)} + (\tilde{u}\tilde{p}, w)_{L^2(\Omega)} = (\tilde{y} - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega),$$

respectively.

Adding and subtracting the term $\tilde{p}\bar{y}_{\mathscr{T}}$ in inequality (41) and utilizing a generalized Hölder's inequality we arrive at

$$\|\bar{u} - \tilde{u}\|_{L^2(\Omega)}^2 \lesssim ([\bar{p}_{\mathscr{T}} - \tilde{p}]\bar{y}_{\mathscr{T}} + \tilde{p}[\bar{y}_{\mathscr{T}} - \tilde{y}], \tilde{u} - \bar{u})_{L^2(\Omega)}$$

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

In view of the Sobolev embedding $H_0^1(\Omega) \hookrightarrow L^4(\Omega)$, we thus conclude that

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

$$+ \|\nabla(\bar{y}_{\mathscr{T}} - \tilde{y})\|_{L^2(\Omega)}\|\nabla\tilde{p}\|_{L^2(\Omega)}. \quad (43)$$

We now control $\|\nabla\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)}$ and $\|\nabla\tilde{p}\|_{L^2(\Omega)}$. Set $v_{\mathscr{T}} = \bar{y}_{\mathscr{T}}$ in (18) and use that $\bar{u}_{\mathscr{T}} \in \mathbb{U}_{ad}$ to obtain $\|\nabla\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$. Similar arguments yield

$$\|\nabla\tilde{p}\|_{L^2(\Omega)} \lesssim \|\tilde{y} - y_\Omega\|_{L^2(\Omega)} \lesssim \|\nabla\tilde{y}\|_{L^2(\Omega)} + \|y_\Omega\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)} + \|y_\Omega\|_{L^2(\Omega)}.$$

These estimates, on the basis of (43), reveal that

$$\|\bar{u} - \tilde{u}\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{p}_{\mathscr{T}} - \tilde{p})\|_{L^2(\Omega)} + \|\nabla(\bar{y}_{\mathscr{T}} - \tilde{y})\|_{L^2(\Omega)}, \qquad (44)$$

with a hidden constant that is independent of continuous and discrete optimal variables but depends on the continuous problem data.

We now proceed to estimate $\|\nabla(\bar{y}_{\mathscr{T}} - \tilde{y})\|_{L^2(\Omega)}$ in (44). Invoke the auxiliary variable $\hat{y} \in H_0^1(\Omega)$, defined as the solution to (29), and the a posteriori error estimate (32), to immediately arrive at

$$\|\nabla(\bar{y}_{\mathscr{T}} - \tilde{y})\|_{L^2(\Omega)} \lesssim \|\nabla(\tilde{y} - \hat{y})\|_{L^2(\Omega)} + \mathcal{E}_{st,\mathscr{T}}. \qquad (45)$$

To control $\|\nabla(\tilde{y} - \hat{y})\|_{L^2(\Omega)}$, we first observe that $\tilde{y} - \hat{y} \in H_0^1(\Omega)$ solves

$$(\nabla(\tilde{y} - \hat{y}), \nabla v)_{L^2(\Omega)} + (\tilde{u}(\tilde{y} - \hat{y}), v)_{L^2(\Omega)} = (\hat{y}(\bar{u}_{\mathscr{T}} - \tilde{u}), v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega).$$

We can thus obtain, in view of a generalized Hölder's inequality, an stability bound for the problem that $\hat{y}$ solves, and definition (37) the following estimate:

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

Replacing this inequality into (45) and the obtained one into (44) yield

$$\|\bar{u} - \tilde{u}\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{p}_{\mathscr{T}} - \tilde{p})\|_{L^2(\Omega)} + \mathcal{E}_{ct,\mathscr{T}} + \mathcal{E}_{st,\mathscr{T}}. \tag{46}$$

The rest of this step is dedicated to bound the term $\|\nabla(\bar{p}_{\mathscr{T}} - \tilde{p})\|_{L^2(\Omega)}$ in (46). To accomplish this task, we first invoke the auxiliary variable $\hat{p} \in H_0^1(\Omega)$, defined as the solution to (33) and the a posteriori error estimate (36) to obtain

$$\|\nabla(\bar{p}_{\mathscr{T}} - \tilde{p})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{adj,\mathscr{T}} + \|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)}. \tag{47}$$

To bound $\|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)}$, we notice that $\tilde{p} - \hat{p} \in H_0^1(\Omega)$ solves

$$(\nabla w, \nabla(\tilde{p} - \hat{p}))_{L^2(\Omega)} + (\tilde{u}(\tilde{p} - \hat{p}), w)_{L^2(\Omega)}$$
$$= ((\bar{u}_{\mathscr{T}} - \tilde{u})\hat{p}, w)_{L^2(\Omega)} + (\tilde{y} - \bar{y}_{\mathscr{T}}, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega).$$

This and the application of basic inequalities reveal the estimate

$$\|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)} \lesssim \|\nabla\hat{p}\|_{L^2(\Omega)}\|\bar{u}_{\mathscr{T}} - \tilde{u}\|_{L^2(\Omega)} + \|\nabla(\tilde{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)}. \tag{48}$$

Observe that

$$\|\nabla\hat{p}\|_{L^2(\Omega)} \lesssim \|\nabla\bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} + \|y_{\Omega}\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)} + \|y_{\Omega}\|_{L^2(\Omega)}. \tag{49}$$

This estimate, combined with definition (37), and estimate $\|\nabla(\tilde{y} - \bar{y}_{\mathscr{T}})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}$ yield, on the basis of (48), the bound

$$\|\nabla(\tilde{p} - \hat{p})\|_{L^2(\Omega)} \lesssim \mathcal{E}_{ct,\mathscr{T}} + \mathcal{E}_{st,\mathscr{T}}. \tag{50}$$

Replace estimate (50) into (47) and the obtained one into (46) to arrive at $\|\bar{u} - \tilde{u}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{adj,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}$. Estimate (40) yields the desired bound

$$\|e_{\bar{u}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{adj,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}. \tag{51}$$

Step 2. The goal of this step is to control $\|\nabla e_{\bar{y}}\|_{L^2(\Omega)}$. Invoke the auxiliary variable $\hat{y} \in H_0^1(\Omega)$ and the a posteriori error estimate (32) to arrive at

$$\|\nabla e_{\bar{y}}\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{y} - \hat{y})\|_{L^2(\Omega)} + \mathcal{E}_{st,\mathscr{T}}. \tag{52}$$

To estimate $\|\nabla(\bar{y} - \hat{y})\|_{L^2(\Omega)}$, we first notice that $\bar{y} - \hat{y} \in H_0^1(\Omega)$ solves

$$(\nabla(\bar{y} - \hat{y}), \nabla v)_{L^2(\Omega)} + (\bar{u}(\bar{y} - \hat{y}), v)_{L^2(\Omega)} = ((\bar{u}_{\mathscr{T}} - \bar{u})\hat{y}, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega).$$

Consequently, estimate (51) yields

$$\|\nabla(\bar{y} - \hat{y})\|_{L^2(\Omega)} \lesssim \|\nabla\hat{y}\|_{L^2(\Omega)}\|\bar{u}_{\mathscr{T}} - \bar{u}\|_{L^2(\Omega)}$$
$$\lesssim \|f\|_{L^2(\Omega)}\|e_{\bar{u}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{adj,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}.$$

We finally replace this bound into (52) to conclude that

$$\|\nabla e_{\bar{y}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{adj,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}. \tag{53}$$

<u>Step 3.</u> The objective now is to bound the term $\|\nabla e_{\bar{p}}\|_{L^2(\Omega)}$. To accomplish this task, we invoke $\hat{p} \in H_0^1(\Omega)$ and the bound (36) to immediately arrive at

$$\|\nabla e_{\bar{p}}\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{p} - \hat{p})\|_{L^2(\Omega)} + \mathcal{E}_{adj,\mathscr{T}}. \tag{54}$$

To estimate $\|\nabla(\bar{p} - \hat{p})\|_{L^2(\Omega)}$, we observe that $\bar{p} - \hat{p} \in H_0^1(\Omega)$ solves

$$(\nabla w, \nabla(\bar{p} - \hat{p}))_{L^2(\Omega)} + (\bar{u}(\bar{p} - \hat{p}), w)_{L^2(\Omega)} = ((\bar{u}_{\mathscr{T}} - \bar{u})\hat{p}, w)_{L^2(\Omega)} + (\bar{y} - \bar{y}_{\mathscr{T}}, w)_{L^2(\Omega)}$$

for all $w \in H_0^1(\Omega)$. We thus invoke estimates (49), (51), and (53) to obtain

$$\|\nabla(\bar{p} - \hat{p})\|_{L^2(\Omega)} \lesssim \|\nabla\hat{p}\|_{L^2(\Omega)}\|e_{\bar{u}}\|_{L^2(\Omega)} + \|e_{\bar{y}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{adj,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}.$$

Replacing this bound into (54) we obtain

$$\|\nabla e_{\bar{p}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{st,\mathscr{T}} + \mathcal{E}_{adj,\mathscr{T}} + \mathcal{E}_{ct,\mathscr{T}}. \tag{55}$$

<u>Step 4.</u> The desired estimate (39) follows from collecting the estimates (51), (53), and (55). $\square$

5.2 Efficiency Analysis

In this section, we derive local efficiency estimates for the local a posteriori error indicators $\mathcal{E}_{st,T}$ and $\mathcal{E}_{ad,T}$ and a global efficiency estimate for the a posteriori error estimator $\mathcal{E}_{ocp}$. To accomplish this task, we will proceed on the basis of standard residual estimation techniques [23].

We begin our analysis by introducing the following notation: for an edge/face or triangle/tetrahedron G, we denote by $\mathcal{V}(G)$ the set of vertices of G. With this notation at hand, we introduce, for $T \in \mathscr{T}$ and $S \in \mathscr{S}$, the standard element and edge bubble functions [23], respectively, as

$$\varphi_T = (d+1)^{(d+1)} \prod_{\text{v} \in \mathcal{V}(T)} \lambda_{\text{v}}, \qquad \varphi_S = d^d \prod_{\text{v} \in \mathcal{V}(S)} \lambda_{\text{v}}|_{T'}, \quad T' \subset \mathcal{N}_S. \tag{56}$$

In these formulas, by λ_{v} we denote the barycentric coordinates of T. We 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 (10) with $u = \bar{u}$, an elementwise integration by parts formula yields, for $v \in H_0^1(\Omega)$, the identity

$$(\nabla e_{\bar{y}}, \nabla v)_{L^2(\Omega)} + (e_{\bar{u}}\bar{y}, v)_{L^2(\Omega)} = -(\bar{u}_{\mathscr{T}} e_{\bar{y}}, v)_{L^2(\Omega)} + \sum_{S \in \mathscr{S}} (\llbracket \nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu} \rrbracket, v)_{L^2(S)}$$

$$+ \sum_{T \in \mathscr{T}} \left[(\mathscr{P}_T f - \bar{y}_{\mathscr{T}} \bar{u}_{\mathscr{T}}, v)_{L^2(T)} + (f - \mathscr{P}_T f, v)_{L^2(T)} \right]. \tag{57}$$

Second, since $\bar{p}$ solves (12) with $y = \bar{y}$ and $u = \bar{u}$, similar arguments yield

$$(\nabla w, \nabla e_{\bar{p}})_{L^2(\Omega)} + (e_{\bar{u}}\bar{p}, w)_{L^2(\Omega)} = -(\bar{u}_{\mathscr{T}}e_{\bar{p}}, w)_{L^2(\Omega)} + (e_{\bar{y}}, w)_{L^2(\Omega)}$$
$$+ \sum_{S \in \mathscr{S}} (\llbracket \nabla \bar{p}_{\mathscr{T}} \cdot \boldsymbol{\nu} \rrbracket, w)_{L^2(S)} + \sum_{T \in \mathscr{T}} (\bar{y}_{\mathscr{T}} - \mathscr{P}_T y_\Omega - \bar{u}_{\mathscr{T}}\bar{p}_{\mathscr{T}}, w)_{L^2(T)}$$
$$+ \sum_{T \in \mathscr{T}} (\mathscr{P}_T y_\Omega - y_\Omega, w)_{L^2(T)} \quad \forall w \in H_0^1(\Omega). \quad (58)$$

In (57) and (58), $\mathscr{P}_T$ denotes the L^2-projection operator onto piecewise constant functions over $T \in \mathscr{T}$.

As a final ingredient, we introduce, for $v \in L^2(\Omega)$ and $\mathcal{M} \subset \mathscr{T}$,

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

We are now ready to prove the local efficiency of $\mathcal{E}_{st,T}$, defined in (30).

Theorem 5.3 (Local efficiency of $\mathcal{E}_{st,T}$) *Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (9)–(10). Let $\bar{u}_{\mathscr{T}}$ be a local minimum of the fully discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. Then, for $T \in \mathscr{T}$, the local error indicator $\mathcal{E}_{st,T}$ satisfies the bound*

$$\mathcal{E}_{st,T} \lesssim \|\nabla e_{\bar{y}}\|_{L^2(\mathcal{N}_T)} + \|e_{\bar{u}}\|_{L^2(\mathcal{N}_T)} + h_T\|e_{\bar{y}}\|_{L^2(\mathcal{N}_T)} + \mathrm{osc}_{\mathscr{T}}(f; \mathcal{N}_T), \quad (59)$$

where $\mathcal{N}_T$ is defined in (6). The hidden constant is independent of continuous and discrete optimal variables, the size of the elements in $\mathscr{T}$, and $\#\mathscr{T}$.

Proof We proceed in two steps and estimate each term in the definition of the local error indicator $\mathcal{E}_{st,T}$, given in (30), separately.

Step 1. Let $T \in \mathscr{T}$. We bound $h_T^2\|f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)}^2$ in (30). To accomplish this task, we begin with a simple application of a triangle inequality to write

$$h_T^2\|f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)}^2 \leq 2h_T^2\|\mathscr{P}_T f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)}^2 + 2\mathrm{osc}_{\mathscr{T}}(f; T)^2. \quad (60)$$

It thus suffices to bound the term $h_T\|\mathscr{P}_T f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)}$. To do this, we set $v = \varphi_T(\mathscr{P}_T f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}})$ in (57), where φ_T denotes the element bubble function introduced in (56). Standard properties of φ_T and inverse inequalities yield

$$\|\mathscr{P}_T f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)}^2 \lesssim \left(h_T^{-1}\|\nabla e_{\bar{y}}\|_{L^2(T)} + h_T^{-1}\|\bar{y}\|_{L^d(T)}\|e_{\bar{u}}\|_{L^2(T)} \right.$$
$$\left. + \|\bar{u}_{\mathscr{T}}\|_{L^\infty(T)}\|e_{\bar{y}}\|_{L^2(T)} + \|f - \mathscr{P}_T f\|_{L^2(T)} \right) \|\mathscr{P}_T f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)}.$$

We notice that, since $d \in \{2, 3\}$, $H_0^1(\Omega) \hookrightarrow L^d(\Omega)$. Consequently,

$$\|\bar{y}\|_{L^d(T)} \leq \|\bar{y}\|_{L^d(\Omega)} \lesssim \|\nabla \bar{y}\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}. \quad (61)$$

This bound combined with the fact that $\bar{u}_{\mathscr{T}} \in \mathbb{U}_{ad}$ yield

$$h_T\|\mathscr{P}_T f - \bar{y}_{\mathscr{T}}\bar{u}_{\mathscr{T}}\|_{L^2(T)} \lesssim \|\nabla e_{\bar{y}}\|_{L^2(T)} + \|e_{\bar{u}}\|_{L^2(T)} + h_T\|e_{\bar{y}}\|_{L^2(T)} + \mathrm{osc}_{\mathscr{T}}(f; T).$$

Replace this bound into (60) to obtain the desired one for $h_T^2 \|f - \bar{y}_{\mathscr{T}} \bar{u}_{\mathscr{T}}\|_{L^2(T)}^2$.

Step 2. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We now bound $h_T^{\frac{1}{2}} \|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}$ in (30). As a first step, we set $v = \varphi_S [\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]$ in identity (57). Here, φ_S denotes the edge bubble function introduced in (56). We thus invoke standard bubble functions arguments and inverse inequalities to arrive at

$$\|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}^2 \lesssim \sum_{T' \in \mathcal{N}_T} \left(h_{T'}^{-1} \|\nabla e_{\bar{y}}\|_{L^2(T')} + h_{T'}^{-1} \|\bar{y}\|_{L^d(T')} \|e_{\bar{u}}\|_{L^2(T')} \right.$$
$$+ \|\bar{u}_{\mathscr{T}}\|_{L^\infty(T')} \|e_{\bar{y}}\|_{L^2(T')} + \|f - \mathscr{P}_{T'} f\|_{L^2(T')}$$
$$\left. + \|\mathscr{P}_T f - \bar{y}_{\mathscr{T}} \bar{u}_{\mathscr{T}}\|_{L^2(T')} \right) h_T^{\frac{1}{2}} \|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)}.$$

In view of the derived estimate for $h_T \|\mathscr{P}_T f - \bar{y}_{\mathscr{T}} \bar{u}_{\mathscr{T}}\|_{L^2(T)}$, (61), and the fact that $\bar{u}_{\mathscr{T}} \in \mathbb{U}_{ad}$, we immediately conclude that

$$h_T^{\frac{1}{2}} \|[\![\nabla \bar{y}_{\mathscr{T}} \cdot \boldsymbol{\nu}]\!]\|_{L^2(S)} \lesssim \sum_{T' \in \mathcal{N}_T} \left(\|\nabla e_{\bar{y}}\|_{L^2(T')} + \|e_{\bar{u}}\|_{L^2(T')} \right.$$
$$\left. + h_T \|e_{\bar{y}}\|_{L^2(T')} + \mathrm{osc}_{\mathscr{T}}(f; T') \right).$$

A collection of the bounds derived in Steps 1 and 2 yield (59). This concludes the proof. □

We now continue with the study of local efficiency properties for the indicator $\mathcal{E}_{adj,T}$, which is defined in (34).

Theorem 5.4 (Local efficiency of $\mathcal{E}_{adj,T}$) *In the framework of Theorem 5.3, we have, for $T \in \mathscr{T}$, the local estimate*

$$\mathcal{E}_{adj,T} \lesssim \|\nabla e_{\bar{p}}\|_{L^2(\mathcal{N}_T)} + \|e_{\bar{u}}\|_{L^2(\mathcal{N}_T)}$$
$$+ h_T \|e_{\bar{p}}\|_{L^2(\mathcal{N}_T)} + h_T \|e_{\bar{y}}\|_{L^2(\mathcal{N}_T)} + \mathrm{osc}_{\mathscr{T}}(y_\Omega; \mathcal{N}_T), \quad (62)$$

where $\mathcal{N}_T$ is defined in (6). The hidden constant is independent of continuous and discrete optimal variables, the size of the elements in $\mathscr{T}$, and $\#\mathscr{T}$.

Proof The proof relies on utilizing identity (58) and similar arguments to the ones elaborated within the proof of Theorem 5.3. For brevity, we skip details. □

The results of Theorems 5.3 and 5.4 yield the following global efficiency estimate for $\mathcal{E}_{ocp,\mathscr{T}}$.

Theorem 5.5 (Global efficiency of $\mathcal{E}_{ocp,\mathscr{T}}$) *In the framework of Theorem 5.3, we have the global estimate*

$$\mathcal{E}_{ocp,\mathscr{T}} \lesssim \|\nabla e_{\bar{p}}\|_{L^2(\Omega)} + \|\nabla e_{\bar{y}}\|_{L^2(\Omega)} + \|e_{\bar{u}}\|_{L^2(\Omega)} + \mathrm{osc}_{\mathscr{T}}(f; \mathscr{T}) + \mathrm{osc}_{\mathscr{T}}(y_\Omega; \mathscr{T}).$$

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

Proof We begin by invoking the definition of the error estimator $\mathcal{E}_{st,\mathscr{T}}$, given in (31), and the local efficiency estimate (59) to arrive at

$$\mathcal{E}_{st,\mathscr{T}} \lesssim \|\nabla e_{\bar{y}}\|_{L^2(\Omega)} + \|e_{\bar{u}}\|_{L^2(\Omega)} + \mathrm{diam}(\Omega)\|e_{\bar{y}}\|_{L^2(\Omega)} + \mathrm{osc}_{\mathscr{T}}(f,\mathscr{T}). \quad (63)$$

On the other hand, the definition of the error estimator $\mathcal{E}_{adj,\mathscr{T}}$, given in (35), and the efficiency estimate (62) yield the bound

$$\mathcal{E}_{adj,\mathscr{T}} \lesssim \|\nabla e_{\bar{p}}\|_{L^2(\Omega)} + \|e_{\bar{u}}\|_{L^2(\Omega)} + \mathrm{diam}(\Omega)\|e_{\bar{p}}\|_{L^2(\Omega)}$$
$$+ \mathrm{diam}(\Omega)\|e_{\bar{y}}\|_{L^2(\Omega)} + \mathrm{osc}_{\mathscr{T}}(y_\Omega,\mathscr{T}). \quad (64)$$

It thus suffices to bound the estimator $\mathcal{E}_{ct,\mathscr{T}}$. In view of (37), a trivial application of a triangle inequality yields

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

where $\Pi_{[\mathsf{a},\mathsf{b}]}$ is defined in (14). This bound, the Lipschitz property of $\Pi_{[\mathsf{a},\mathsf{b}]}$, the Cauchy–Schwarz inequality, and the embedding $H_0^1(\Omega) \hookrightarrow L^4(\Omega)$ yield

$$\mathcal{E}_{ct,\mathscr{T}} \leq \alpha^{-1}\|\nabla e_{\bar{p}}\|_{L^2(\Omega)}\|\nabla \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} + \alpha^{-1}\|\nabla \bar{p}\|_{L^2(\Omega)}\|\nabla e_{\bar{y}}\|_{L^2(\Omega)} + \|e_{\bar{u}}\|_{L^2(\Omega)}.$$

Observe that $\|\nabla \bar{y}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$ and $\|\nabla \bar{p}\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)} + \|y_\Omega\|_{L^2(\Omega)}$. Consequently,

$$\mathcal{E}_{ct,\mathscr{T}} \lesssim \|\nabla e_{\bar{p}}\|_{L^2(\Omega)} + \|\nabla e_{\bar{y}}\|_{L^2(\Omega)} + \|e_{\bar{u}}\|_{L^2(\Omega)}, \quad (65)$$

with a hidden constant that is independent of continuous and discrete optimal variables but depends on the continuous problem data.

The proof concludes by gathering estimates (63), (64), and (65), upon utilizing a Poincaré inequality. $\qquad\square$

6 A Posteriori Error Analysis: the Semi-discrete Scheme

In this section, we design and analyze an a posteriori error estimator for the semi-discrete scheme of section 4.3.2. In contrast to the estimator devised in section 5, the estimator is now formed by only two contributions: one related to the discretization of the state equation and another one associated to the discretization of the adjoint equation.

6.1 Global Reliability Analysis

The goal of this section is to design an a posteriori error estimator and derive an upper bound for the corresponding total error in terms of the devised error estimator. As in section 5.1, the aforementioned upper bound will be obtained on the basis of estimates on the error between solutions to the semi-discrete optimal control problem (20)–(21) and suitable auxiliary variables.

The first auxiliary variable is $\hat{y}$ and is defined as follows:

$$\hat{y} \in H_0^1(\Omega): \quad (\nabla\hat{y}, \nabla v)_{L^2(\Omega)} + (\bar{\mathsf{u}}\hat{y}, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega). \tag{66}$$

With this variable at hand, we define, for $T \in \mathscr{T}$, the local error indicators and the corresponding a posteriori error estimator, respectively, by

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

We notice that, since $\bar{y}_\mathscr{T} \in \mathbb{V}(\mathscr{T})$, solution to (20) with $\mathsf{u} = \bar{\mathsf{u}}$, can be seen as a finite element approximation of $\hat{y}$, an application of Theorem 3.1 yields

$$\|\nabla(\hat{y} - \bar{y}_\mathscr{T})\|_{L^2(\Omega)} \lesssim \mathsf{E}_{st,\mathscr{T}}. \tag{67}$$

The second variable is $\hat{p} \in H_0^1(\Omega)$ and is defined as the solution to

$$(\nabla w, \nabla\hat{p})_{L^2(\Omega)} + (\bar{\mathsf{u}}\hat{p}, w)_{L^2(\Omega)} = (\bar{y}_\mathscr{T} - y_\Omega, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega). \tag{68}$$

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

$$\mathsf{E}_{adj,T}^2 := h_T^2 \|\bar{y}_\mathscr{T} - y_\Omega - \bar{\mathsf{u}}\bar{p}_\mathscr{T}\|_{L^2(T)}^2 + h_T \|[\![\nabla\bar{p}_\mathscr{T} \cdot \boldsymbol{\nu}]\!]\|_{L^2(\partial T \setminus \partial\Omega)}^2,$$

and the a posteriori error estimator

$$\mathsf{E}_{adj,\mathscr{T}}^2 := \sum_{T \in \mathscr{T}} \mathsf{E}_{adj,T}^2.$$

Since $\bar{p}_\mathscr{T} \in \mathbb{V}(\mathscr{T})$ can be seen as the finite element approximation of $\hat{p}$ within $\mathbb{V}(\mathscr{T})$, Theorem 3.1 yields

$$\|\nabla(\hat{p} - \bar{p}_\mathscr{T})\|_{L^2(\Omega)} \lesssim \mathsf{E}_{adj,\mathscr{T}}. \tag{69}$$

In order to present the following reliability result, we introduce the error $\mathsf{e}_{\bar{u}} := \bar{u} - \bar{\mathsf{u}}$ and the a posteriori error estimator $\mathsf{E}_{ocp,\mathscr{T}}^2 := \mathsf{E}_{st,\mathscr{T}}^2 + \mathsf{E}_{adj,\mathscr{T}}^2$.

Theorem 6.1 (Global reliability) *Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (9)–(10) satisfying the sufficient second order condition (16). Let $\bar{\mathsf{u}}$ be a local minimum of the semi-discrete optimal control problem with $\bar{y}_\mathscr{T}$ and $\bar{p}_\mathscr{T}$ being the corresponding state and adjoint state, respectively. If $\bar{y}_\mathscr{T}$ and $\bar{p}_\mathscr{T}$ satisfy, on the mesh $\mathscr{T}$, the bound (24), then*

$$\|\nabla e_{\bar{y}}\|_{L^2(\Omega)}^2 + \|\nabla e_{\bar{p}}\|_{L^2(\Omega)}^2 + \|\mathsf{e}_{\bar{u}}\|_{L^2(\Omega)}^2 \lesssim \mathsf{E}_{ocp,\mathscr{T}}^2, \tag{70}$$

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

Proof We immediately notice that, under the particular setting inherited by the semi-discrete scheme, the auxiliary variables that we have devised to perform our analysis satisfy that $\bar{\mathsf{u}} = \tilde{u}$, where $\tilde{u}$ is defined in (22). We can thus immediately conclude that $\hat{y} = \tilde{y}$, with $\hat{y}$ and $\tilde{y}$ being defined as the unique solutions to (66) and (42), respectively. Therefore, invoking (44), (45), (47), and (48) we conclude that

$$\|\mathsf{e}_{\bar{u}}\|_{L^2(\Omega)} = \|\bar{u} - \tilde{u}\|_{L^2(\Omega)} \lesssim \|\nabla(\bar{p}_{\mathscr{T}} - \tilde{p})\|_{L^2(\Omega)} + \mathsf{E}_{st,\mathscr{T}} \lesssim \mathsf{E}_{st,\mathscr{T}} + \mathsf{E}_{adj,\mathscr{T}}.$$

The estimation of the terms $\|\nabla e_{\bar{y}}\|_{L^2(\Omega)}$ and $\|\nabla e_{\bar{p}}\|_{L^2(\Omega)}$ follow by utilizing the bound $\|\mathsf{e}_{\bar{u}}\|_{L^2(\Omega)} \lesssim \mathsf{E}_{ocp,\mathscr{T}}$ and similar arguments to the ones developed in the proof of Theorem 5.2. For brevity, we skip the details. □

6.2 Efficiency Analysis

We begin the section by defining, for $T \in \mathscr{T}$, the local indicator

$$\mathsf{E}_{ocp,T}^2 := \mathsf{E}_{st,T}^2 + \mathsf{E}_{adj,T}^2. \tag{71}$$

The estimates obtained in Theorems 5.3 and 5.4 can also be obtained within the setting of the variational discretization approach.

Theorem 6.2 (Local estimates for $\mathsf{E}_{ocp,T}$) *Let $\bar{u} \in \mathbb{U}_{ad}$ be a local solution to (9)–(10). Let $\bar{\mathsf{u}}$ be a local minimum of the semi-discrete optimal control problem with $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$ being the corresponding state and adjoint state, respectively. Then, for $T \in \mathscr{T}$, the local error indicator $\mathsf{E}_{ocp,T}$ satisfies*

$$\begin{aligned}
\mathsf{E}_{ocp,T} \lesssim{} & \|\nabla e_{\bar{y}}\|_{L^2(\mathcal{N}_T)} + \|\nabla e_{\bar{p}}\|_{L^2(\mathcal{N}_T)} + \|\mathsf{e}_{\bar{u}}\|_{L^2(\mathcal{N}_T)} + h_T \|e_{\bar{p}}\|_{L^2(\mathcal{N}_T)} \\
& + h_T \|e_{\bar{y}}\|_{L^2(\mathcal{N}_T)} + \mathrm{osc}_{\mathscr{T}}(f; \mathcal{N}_T) + \mathrm{osc}_{\mathscr{T}}(y_{\Omega}; \mathcal{N}_T),
\end{aligned}$$

where $\mathcal{N}_T$ is defined as in (6). The hidden constant is independent of continuous and discrete optimal variables, the size of the elements in $\mathscr{T}$, and $\#\mathscr{T}$.

Proof The desired result follows by utilizing similar arguments to the ones that yield estimates (59) and (62). For brevity, we skip details. □

7 Numerical Examples

In this section, we conduct a series of numerical experiments that illustrate the performance of the devised a posteriori error estimators $\mathcal{E}_{ocp,\mathscr{T}}$ and $\mathsf{E}_{ocp,\mathscr{T}}$ when used to drive suitable AFEMs schemes based on the fully and semi-discrete schemes proposed in sections 4.3.1 and 4.3.2, respectively.

7.1 Implementation Details

The numerical examples that we shall present in what follows have been carried out with the help of a code that we implemented using `C++`. Global linear systems were solved using the multifrontal massively parallel sparse direct solver (MUMPS) [3,4]. We have used a quadrature formula to compute the right-hand sides, the approximation errors, and the error indicators; the quadrature formula being exact for polynomials of degree nineteen (19) for two dimensional domains and degree fourteen (14) for three dimensional domains.

In what follows, we discuss some pertinent implementation details that are particular for each discretization technique.

The Fully Discrete Scheme: For a given partition $\mathscr{T}$, we seek a discrete solution $(\bar{y}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}, \bar{u}_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}_{ad}(\mathscr{T})$ that solves the corresponding optimality system. This nonlinear system is solved on the basis of an adaptation of the semi–smooth Newton method described in [10, Appendix A.1]. We have exactly assembled all matrices involved in the left-hand side of the resulting linear system. The total number of degrees of freedom (DOFs) is $\mathsf{Ndof} = 2 \dim(\mathbb{V}(\mathscr{T})) + \dim(\mathbb{U}(\mathscr{T}))$. We measure the error within the norm $\|e\|_{\Omega} := [\|\nabla e_{\bar{y}}\|^2_{L^2(\Omega)} + \|\nabla e_{\bar{p}}\|^2_{L^2(\Omega)} + \|e_{\bar{u}}\|^2_{L^2(\Omega)}]^{\frac{1}{2}}$. Finally, we introduce the effectivity index $\mathcal{I}_{eff} := \mathcal{E}_{ocp,\mathscr{T}}/\|e\|_{\Omega}$.

The Semi-discrete Scheme: For a given partition $\mathscr{T}$, we seek a solution $(\bar{y}_{\mathscr{T}}, \bar{p}_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T})$ that solves the corresponding optimality system. This system is also solved by using an adaptation of the semi–smooth Newton method described in [10, Appendix A.1].

The following comments regarding the implementation of the variational discretization approach are of importance. In order to properly implement such a scheme, the assembling and exact computation of $(\bar{\mathsf{u}}\bar{y}_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)}$ and $(\bar{\mathsf{u}}\bar{p}_{\mathscr{T}}, w_{\mathscr{T}})_{L^2(\Omega)}$ are required. In particular, the exact integration of such terms on the simplices $T \in \mathscr{T}$ where the control $\bar{\mathsf{u}}$ exhibits kinks is necessary. Let us now describe an alternative to perform such a computation: First, recognize the simplices $T \in \mathscr{T}$ which are such that the control $\bar{\mathsf{u}}$ have kinks. Second, recognize the regions of such simplices where the control variable is inactive/active. The following difficulty thus appears: since we are using continuous piecewise polynomials of degree one to approximate $\bar{y}_{\mathscr{T}}$ and $\bar{p}_{\mathscr{T}}$, these regions have, in general, curved boundaries; see [18, Remark 5.19]. The third step is the computation of $(\bar{\mathsf{u}}\bar{y}_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)}$ and $(\bar{\mathsf{u}}\bar{p}_{\mathscr{T}}, w_{\mathscr{T}})_{L^2(\Omega)}$ by partitioning the integrals on the regions where the control is inactive/active. Since the computational implementation of the second and third steps is far from being simple, our implementation relies on computing the terms $(\bar{\mathsf{u}}\bar{y}_{\mathscr{T}}, v_{\mathscr{T}})_{L^2(\Omega)}$ and $(\bar{\mathsf{u}}\bar{p}_{\mathscr{T}}, w_{\mathscr{T}})_{L^2(\Omega)}$ with the help of a quadrature formula. We emphasize that this numerical implementation leads to an *approximated version* of the variational discretization approach which is capable of delivering, within an adaptive loop, optimal experimental rates of convergence for all the involved variables.

The total number of DOFs for the semi-discrete scheme corresponds to $\mathsf{Ndof} = 2 \dim(\mathbb{V}(\mathscr{T}))$. To measure the corresponding approximation error, we

use $\|e\|_{\Omega} = [\|\nabla e_{\bar{y}}\|^2_{L^2(\Omega)} + \|\nabla e_{\bar{p}}\|^2_{L^2(\Omega)} + \|e_{\bar{u}}\|^2_{L^2(\Omega)}]^{\frac{1}{2}}$. As a final ingredient, we introduce the effectivity index $I_{eff} := E_{ocp,\mathscr{T}}/\|e\|_{\Omega}$.

Once the discrete solution is obtained, we compute, for $T \in \mathscr{T}$, the error indicator $\mathcal{E}_{ocp,T}$, defined by

$$\mathcal{E}^2_{ocp,T} := \mathcal{E}^2_{st,T} + \mathcal{E}^2_{adj,T} + \mathcal{E}^2_{ct,T}, \tag{72}$$

or the indicator $E_{ocp,T}$, defined in (71), to drive the adaptive mesh refinement procedure described in Algorithm 1.

Algorithm 1 Adaptive algorithm

Input: Initial mesh $\mathscr{T}_0$, desired state y_{Ω}, external source f, constraints $\mathtt{a}$ and $\mathtt{b}$, and regularization parameter α;

Set: $i = 0$.

Newton strategy:

1: Choose initial guesses $y^0_{\mathscr{T}_i}, p^0_{\mathscr{T}_i} \in \mathbb{V}(\mathscr{T}_i)$ (and $u^0_{\mathscr{T}_i} \in \mathbb{U}(\mathscr{T}_i)$) when the fully discrete scheme is considered);

2 (Fully discrete solution technique): Compute $[\bar{y}_{\mathscr{T}_i}, \bar{p}_{\mathscr{T}_i}, \bar{u}_{\mathscr{T}_i}] = $ Semi-Smooth$[\mathscr{T}_i, y^0_{\mathscr{T}_i}, p^0_{\mathscr{T}_i}, u^0_{\mathscr{T}_i}, y_{\Omega}, f, \mathtt{a}, \mathtt{b}, \alpha]$, which implements an adaptation of the semi–smooth Newton method described in [10, Appendix A.1];

2 (Semi-discrete solution technique): Compute $[\bar{y}_{\mathscr{T}_i}, \bar{p}_{\mathscr{T}_i}] = $ Semi-Smooth$[\mathscr{T}_i, y^0_{\mathscr{T}_i}, p^0_{\mathscr{T}_i}, y_{\Omega}, f, \mathtt{a}, \mathtt{b}, \alpha]$, which implements an adaptation of the semi–smooth Newton method described in [10, Appendix A.1];

Adaptive loop:

3: For each $T \in \mathscr{T}_i$, compute the local error indicator $\mathcal{E}_{ocp,T}$ ($E_{ocp,T}$) defined in (72) ((71));

4: Mark an element $T \in \mathscr{T}_i$ for refinement if $\mathcal{E}^2_{ocp,T} > \frac{1}{2}\max_{T' \in \mathscr{T}_i} \mathcal{E}^2_{ocp,T'}$ ($E^2_{ocp,T} > \frac{1}{2}\max_{T' \in \mathscr{T}_i} E^2_{ocp,T'}$);

5: From step **4**, construct a new mesh $\mathscr{T}_{i+1}$ using a longest edge bisection algorithm. Set $i \leftarrow i+1$ and go to step **1**.

7.2 Numerical Experiments

We now provide two numerical experiments. In both examples, we consider problems where an exact solution can be obtained: succinctly, we fix the optimal state and adjoint state variables and compute the exact optimal control, the desired state y_{Ω}, and the source term f.

Example 1 (L-shaped domain). We set $\Omega = (-1,1)^2 \setminus [0,1) \times (-1,0]$, $\mathtt{a} = 0.01$, $\mathtt{b} = 5$, and $\alpha = 0.1$. The exact optimal state and adjoint state are given, in polar coordinates (ρ, ω) with $\omega \in [0, 3\pi/2]$, by

$$\bar{y} = 3\sin(\pi(\rho\sin(\omega)+1)/2)\sin(\pi(\rho\cos(\omega)+1)/2)\rho^{2/3}\sin(2\omega/3),$$

$$\bar{p} = 2\cos(\pi\rho\sin(\omega)/2)\sin(\pi(\rho\cos(\omega)+1)/2)\rho^{2/3}\sin(2\omega/3).$$

The purpose of this example is to investigate the performance of the devised a posteriori error estimators in a non–convex domain.

In Fig. 1 we present the results obtained for Example 1. We present, in subfigures (A.1)–(A.3), experimental rates of convergence for all the individual contributions of the total errors $\|e\|_\Omega$ and $\|\mathsf{e}\|_\Omega$ when uniform and adaptive refinements are considered within both discretization schemes, i.e., the fully discrete scheme (Fully) and the semi-discrete scheme (Semi). We observe that our adaptive loops *outperform* uniform refinement. In addition, we observe that our adaptive loops yield optimal experimental rates of convergence for all the individual contributions of the total errors $\|e\|_\Omega$ (Fully) and $\|\mathsf{e}\|_\Omega$ (Semi). We also observe, in subfigures (A.4) and (A.5), that the error estimators $\mathcal{E}_{ocp}$ (Fully) and E_{ocp} (Semi) exhibit optimal rates of convergence. Moreover, in subfigure (A.6), it can be observed that when Ndof increases, the effectivity indices $\mathcal{I}_{eff}$ (Fully) and I_{eff} (Semi) are stabilized around the values 4.7 and 6.4, respectively.
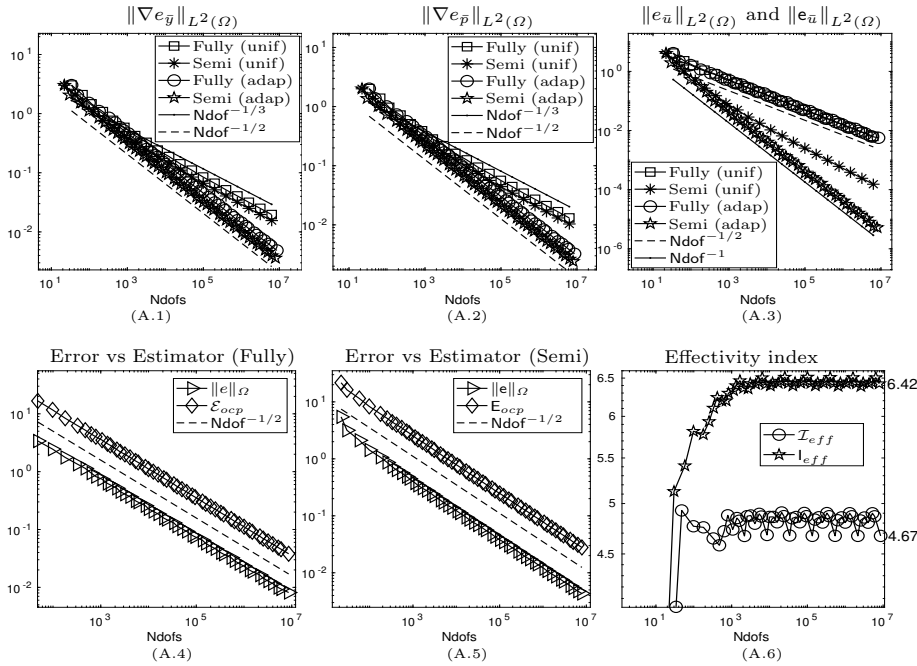

Fig. 1 Example 1: Experimental rates of convergence for the individual errors $\|\nabla e_{\bar{y}}\|_{L^2(\Omega)}, \|\nabla e_{\bar{p}}\|_{L^2(\Omega)}, \|e_{\bar{u}}\|_{L^2(\Omega)}$, and $\|\mathsf{e}_{\bar{u}}\|_{L^2(\Omega)}$, for uniform and adaptive refinement (A.1)–(A.3), the total errors $\|e\|_\Omega$ and $\|\mathsf{e}\|_\Omega$ and the error estimators $\mathcal{E}_{ocp,\mathscr{T}}$ and $\mathsf{E}_{ocp,\mathscr{T}}$ for adaptive refinement (A.4) and (A.5), and the effectivity indices (A.6).

Example 2 (Convex domain). We consider $\Omega = (0,1)^3$, $\alpha = 10^{-3}$, $\mathtt{a} = 0.2$, and $\mathtt{b} = 3$. The exact optimal state and adjoint state are given by $\bar{y} = x_1 x_2 x_3 (1-x_1)(1-x_2)(1-x_3)$ and $\bar{p} = \bar{y} \cdot \arctan(100(x_1 - 0.5))$, respectively.

In Fig. 2 we present the results obtained for Example 2. Similar conclusions to the ones presented for Example 1 can be derived. In particular, we observe optimal experimental rates of convergence for all the individual contributions of the total errors $\|e\|_\Omega$ (Fully) and $\|\mathsf{e}\|_\Omega$ (Semi) when adaptive refinement is considered.
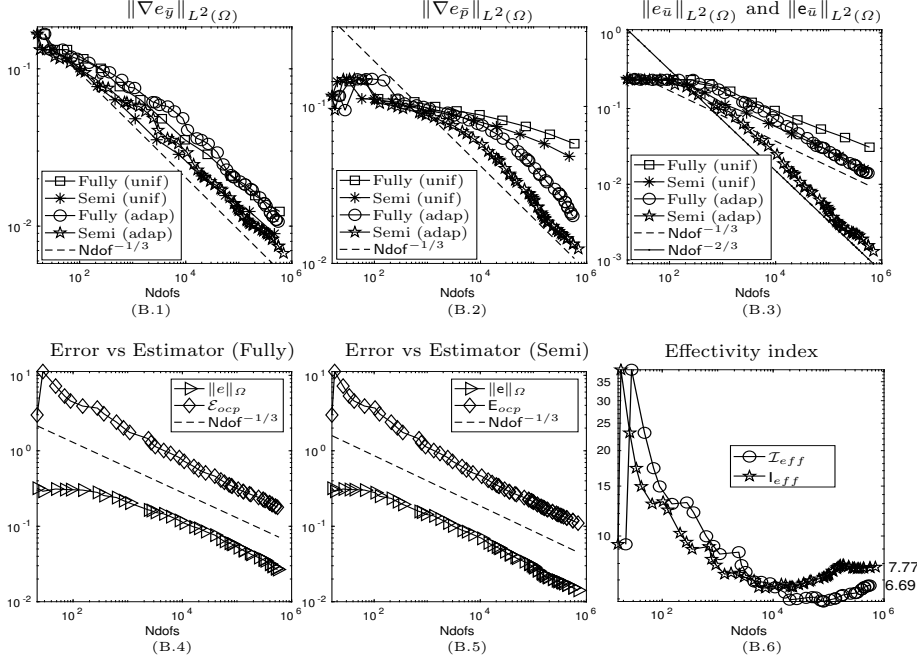

Fig. 2 Example 2: Experimental rates of convergence for the individual errors $\|\nabla e_{\bar{y}}\|_{L^2(\Omega)}, \|\nabla e_{\bar{p}}\|_{L^2(\Omega)}, \|e_{\bar{u}}\|_{L^2(\Omega)}$, and $\|\mathsf{e}_{\bar{u}}\|_{L^2(\Omega)}$, for uniform and adaptive refinement (B.1)–(B.3), the total errors $\|e\|_\Omega$ and $\|\mathsf{e}\|_\Omega$ and error estimators $\mathcal{E}_{ocp,\mathscr{T}}$ and $\mathsf{E}_{ocp,\mathscr{T}}$ for adaptive refinement (B.4) and (B.5), and the effectivity indices (B.6).

8 Conclusions

In this work, we designed AFEMs for an optimal control problem with a bilinear state equation; the control variable enters the state equation as a coefficient. Two different discretization schemes were considered to approximate an optimal solution: a fully discrete scheme and a semi-discrete one based on the so-called variation discretization approach. We constructed, for each one of these schemes, residual–type a posteriori error estimators that are formed by the sum of contributions related to the discretization of the state and adjoint equations and, additionally, the discretization of the control variable for when the fully discrete scheme is considered. We derived global reliability and efficiency estimates for the error estimator associated to the fully discrete scheme

whereas global reliability and *local* efficiency estimates were derived for the error estimator associated to the semi-discrete scheme. Finally, we performed numerical tests that showed competitive performances of both error estimators when used to drive adaptive procedures.

References

1. ADAMS, R. A., AND FOURNIER, J. J. F. *Sobolev Spaces*, Second ed., vol. 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, 2003.
2. ALLENDES, A., FUICA, F., AND OTÁROLA, E. Adaptive finite element methods for sparse PDE-constrained optimization. *IMA J. Numer. Anal. 40*, 3 (2020), 2106–2142.
3. AMESTOY, P., DUFF, I., AND L'EXCELLENT, J.-Y. Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering 184*, 2 (2000), 501 – 520.
4. AMESTOY, P. R., DUFF, I. S., L'EXCELLENT, J.-Y., AND KOSTER, J. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM J. Matrix Anal. Appl. 23*, 1 (2001), 15–41 (electronic).
5. BECKER, R., KAPP, H., AND RANNACHER, R. Adaptive finite element methods for optimal control of partial differential equations: basic concept. *SIAM J. Control Optim. 39*, 1 (2000), 113–132.
6. BRENNER, S. C., AND SCOTT, L. R. *The Mathematical Theory of Finite Element Methods*, Third ed., vol. 15 of *Texts in Applied Mathematics*. Springer, New York, 2008.
7. CHANG, Y., YANG, D., AND ZHANG, Z. Adaptive finite element approximation for a class of parameter estimation problems. *Appl. Math. Comput. 231* (2014), 284–298.
8. CHEN, Y., LU, Z., AND HUANG, Y. Superconvergence of triangular Raviart-Thomas mixed finite element methods for a bilinear constrained optimal control problem. *Comput. Math. Appl. 66*, 8 (2013), 1498–1513.
9. CIARLET, P. G. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.
10. CLASON, C., AND JIN, B. A semismooth Newton method for nonlinear parameter identification problems with impulsive noise. *SIAM J. Imaging Sci. 5*, 2 (2012), 505–538.
11. FU, H., GUO, H., HOU, J., AND ZHANG, J. A stabilized mixed finite element approximation of bilinear state optimal control problems. *Comput. Math. Appl. 74*, 6 (2017), 1246–1261.
12. GONG, W., AND YAN, N. Adaptive finite element method for elliptic optimal control problems: convergence and optimality. *Numer. Math. 135*, 4 (2017), 1121–1170.
13. HINTERMÜLLER, M., HOPPE, R. H. W., ILIASH, Y., AND KIEWEG, M. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM Control Optim. Calc. Var. 14*, 3 (2008), 540–560.
14. HINZE, M. A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl. 30*, 1 (2005), 45–61.
15. HINZE, M., PINNAU, R., ULBRICH, M., AND ULBRICH, S. *Optimization with PDE Constraints*, vol. 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.
16. HINZE, M., AND TRÖLTZSCH, F. Discrete concepts versus error analysis in PDE-constrained optimization. *GAMM-Mitt. 33*, 2 (2010), 148–162.
17. KOHLS, K., RÖSCH, A., AND SIEBERT, K. G. A posteriori error analysis of optimal control problems with control constraints. *SIAM J. Control Optim. 52*, 3 (2014), 1832–1861.
18. KRÖNER, A., AND VEXLER, B. A priori error estimates for elliptic optimal control problems with a bilinear state equation. *J. Comput. Appl. Math. 230*, 2 (2009), 781–802.

19. Kunisch, K., Liu, W., Chang, Y., Yan, N., and Li, R.   Adaptive finite element approximation for a class of parameter estimation problems. *J. Comput. Math. 28*, 5 (2010), 645–675.

20. Liu, W., and Yan, N.   A posteriori error estimates for distributed convex optimal control problems. *Adv. Comput. Math. 15*, 1-4 (2001), 285–309.

21. Nochetto, R. H., Siebert, K. G., and Veeser, A. Theory of Adaptive Finite Element Methods: an Introduction. In *Multiscale, Nonlinear and Adaptive Approximation*, R. A. DeVore and A. Kunoth, Eds. Springer, Berlin, 2009, pp. 409–542.

22. Tröltzsch, F. *Optimal Control of Partial Differential Equations*, vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, Methods and Applications, Translated from the 2005 German original by Jürgen Sprekels.

23. Verfürth, R. *A Posteriori Error Estimation Techniques for Finite Element Methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

24. Vexler, B., and Wollner, W. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Control Optim. 47*, 1 (2008), 509–534.