Maximum–norm a posteriori error estimates for an optimal control problem

Enrique Otárola
Richard Rankin
Abner J. Salgado

Abstract We analyze a reliable and efficient max-norm a posteriori error estimator for a control-constrained, linear–quadratic optimal control problem. The estimator yields optimal experimental rates of convergence within an adaptive loop.

Keywords Linear–quadratic optimal control problem · Finite element methods · A posteriori error analysis · Maximum–norm

Mathematics Subject Classification (2000) 49J20 · 49M25 · 65K10 · 65N15 · 65N30 · 65N50 · 65Y20

1 Introduction

Let Ω be an open and bounded polytope in $\mathbb{R}^d$, $d \in \{2,3\}$, with Lipschitz boundary $\partial\Omega$. Given $\mathsf{y}_\Omega \in L^2(\Omega)$ and $\lambda > 0$ we define the cost functional

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

In this article we devise max-norm a posteriori error estimators for the following optimal control problem: Find

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

subject to, for a given $\mathsf{f} \in L^2(\Omega)$, the linear elliptic partial differential equation (PDE)

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

Enrique Otárola
Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile
E-mail: enrique.otarola@usm.cl

Richard Rankin
School of Mathematical Sciences, University of Nottingham Ningbo China, Ningbo, China
E-mail: richard.rankin@nottingham.edu.cn

Abner J. Salgado
Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA
E-mail: asalgad1@utk.edu

and, for $\mathsf{a},\mathsf{b} \in \mathbb{R}$ with $\mathsf{a} \leq \mathsf{b}$, the control constraints

$$\mathsf{u} \in \mathbb{U}_{\text{ad}}, \qquad \mathbb{U}_{\text{ad}} := \{\mathsf{v} \in L^2(\Omega) : \ \mathsf{a} \leq \mathsf{v}(x) \leq \mathsf{b} \text{ for almost every } x \text{ in } \Omega\}. \tag{4}$$

Within the framework of finite element approximation, error estimates in the maximum norm are of special interest in applied sciences and engineering: they control pointwise accuracy, as opposed to the averaged L^2-based error estimates that arise naturally in a finite element setting. Starting with the pioneering works of Nitsche [31,32], Natterer [30], Scott [45], and Frehse and Rannacher [15], the study of a priori error estimates for piecewise linear finite element approximations of linear and elliptic problems has undergone great development [12,17,23,38,41,42,43,44]. The analysis of residual–type a posteriori error estimates in the maximum norm has also been considered in a number of works. To the best of our knowledge, the earliest two works that study L^∞ a posteriori error estimators for a Poisson problem with a bounded forcing term on two–dimensional domains are [13] and [33]. The results of Nochetto [33] were later extended to three–dimensional domains in [7] and subsequently improved in [9,10,34]. The theory has also been extended to problems involving unbounded forcing terms, the Stokes equations, and obstacle, monotone semilinear, and geometric problems [3,6,10,11,34,35].

A posteriori error estimators are computable quantities that depend only on the approximate solution and known data. They can be used to extract information about the local quality of the approximate solution. This makes them an essential ingredient of adaptive finite element methods (AFEMs) which are iterative methods that improve the quality of the approximation while striving to keep an efficient distribution of computational resources measured in terms of degrees of freedom. The a posteriori error analysis for standard finite element approximations of linear second–order elliptic boundary value problems has a solid foundation. We refer the reader to [1,36,37,47] for discussions that include the construction of AFEMs, their convergence and optimal complexity. In spite of these advances, the a posteriori error analysis for finite element approximations of constrained optimal control problems has not been fully developed. The main source of difficulty is their inherent nonlinear feature, which appears due to the control constraints. To the best of our knowledge, the first work that provides an advance is [26]. In this work the authors design an error estimator and prove that it yields an upper bound for the error [26, Theorem 3.1]. These results were later complemented in [19] where the authors propose a slight modification of the estimator of [26] and provide an efficiency analysis [19, Theorems 5.1 and 6.1]. A unification of these ideas has been carried out in [22]: on the basis of an important error equivalence the analysis is simplified to provide reliable and efficient estimators for the state and adjoint equations. The analysis is based on the energy norm inherited by the state and adjoint equations. However, many problems do not fit into this framework, and the problem we consider in this work is an instance of this issue. Indeed, since we are interested in maximum norm estimates, the energy based arguments of [22] do not apply. For extensions and different approaches based on weighted residual and goal-oriented methods we refer the reader to [2, 3,5,18,24,28,39].

Before proceeding with the description and analysis of our method, let us provide an overview of those advocated in the literature regarding the approximation of (2)–(4) in the maximum norm. In [29], the authors derive a *linear* order of convergence for the approximation error of the control variable for a scheme based on piecewise linear functions [29, Theorem 2.3]. The analysis of [29] is performed under the following assumptions:

$$d = 2, \quad \Omega \text{ is convex and } \Omega \in C^{1,1}, \quad \mathsf{y}_\Omega \in L^r(\Omega), \quad r > 2. \tag{5}$$

Since the control variable is approximated with piecewise linear functions the a priori error estimate of [29, Theorem 2.3] is *suboptimal* in terms of approximation. In this work we propose and analyze an a posteriori error estimator that, when used to drive an AFEM, delivers a *quadratic* experimental rate of convergence in the numerical examples that we perform. Such a rate of convergence is optimal in terms of approximation and consequently our scheme outperforms the one proposed in [29]. In addition, our analysis is valid under the following assumptions:

$$d \in \{2,3\}, \quad \Omega \text{ is a polytope with Lipschitz boundary}, \quad \mathsf{f}, \mathsf{y}_\Omega \in L^2(\Omega). \tag{6}$$

We comment that an external forcing f is not considered in [29]. Later, the authors of [4] study, when $\Omega \subset \mathbb{R}^2$ is a bounded and polygonal domain and $\mathsf{y}_\Omega \in C^{0,\sigma}(\overline{\Omega})$ with $\sigma \in (0,1]$, two solution techniques for (2)–(4): the so–called variational approach and a fully discrete scheme that discretizes the control with piecewise constant functions; standard piecewise linear functions are used to approximate the solutions to the state and adjoint equations. The latter scheme is combined with a post–processing step that defines a new control variable as the projection of the discrete adjoint state into the admissible set; see [4, equation (3.43)]. The authors of [4] prove that, in both schemes, the approximation errors of all three variables in the maximum norm behave like $h^2|\log(h)|$; see [4, Theorems 3.8 and 3.15]. The work [4] hinges on the use of appropriately graded meshes to compensate for the fact that the domain, on which the optimal control problem is posed, exhibits corners. In contrast, by embedding our a posteriori error estimators in an AFEM, suitable meshes will be automatically constructed within the adaptive loop.

The outline of this paper is as follows. In section 2 we introduce some terminology used throughout this work. In section 3 we study the optimal control problem (2)–(3) and obtain the associated optimality system. The core of our work are sections 4 and 5: In section 4 we design an a posteriori error estimator for a finite element discretization of problem (2)–(3) and study its reliability properties, while in section 5 we prove the local efficiency of the devised error estimator. Finally, in section 6, we present numerical examples to illustrate the theory.

2 Notation

Let $\mathscr{T} = \{T\}$ be a conforming simplicial mesh of $\overline{\Omega}$ [14], $h_T = \mathrm{diam}(T)$ and

$$\ell_{\mathscr{T}} = \left| \log \left(\max_{T \in \mathscr{T}} h_T^{-1} \right) \right|. \tag{7}$$

We assume that $\mathscr{T}$ is a member of a shape regular family of meshes. Define

$$\mathbb{W}(\mathscr{T}) := \{ w \in C^0(\overline{\Omega}) : \ w_{|T} \in \mathbb{P}_1(T) \ \forall \, T \in \mathscr{T} \}, \qquad \mathbb{V}(\mathscr{T}) := \mathbb{W}(\mathscr{T}) \cap H_0^1(\Omega),$$

and $\mathbb{U}_{\mathrm{ad}}(\mathscr{T}) := \mathbb{U}_{\mathrm{ad}} \cap \mathbb{W}(\mathscr{T})$. We denote by $\mathscr{S} = \{S\}$ the set of internal $(d-1)$-dimensional interelement boundaries of $\mathscr{T}$ and $h_S = \mathrm{diam}(S)$. If $T \in \mathscr{T}$, $\mathscr{S}_T \subset \mathscr{S}$ is the set of sides of T. For $S \in \mathscr{S}$ we set $\mathscr{N}_S = \{T^+, T^-\}$ such that $S = T^+ \cap T^-$. For $T \in \mathscr{T}$, we define

$$\mathscr{N}_T := \left\{ T' \in \mathscr{T} : \mathscr{S}_T \cap \mathscr{S}_{T'} \neq \emptyset \right\}. \tag{8}$$

For $w_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ and $S \in \mathscr{S}$ with $\mathscr{N}_S = \{T^+, T^-\}$, the jump or interelement residual is $[\![\nabla w_{\mathscr{T}} \cdot \nu]\!] = \nu^+ \cdot \nabla w_{\mathscr{T}|T^+} + \nu^- \cdot \nabla w_{\mathscr{T}|T^-}$, where ν^+, ν^- are the unit normals to S pointing towards T^+, $T^- \in \mathscr{T}$. The $L^2(\Omega)$ inner product is denoted by $(\cdot,\cdot)_{L^2(\Omega)}$. By $A \lesssim B$ we mean that $A \leq cB$ for a nonessential constant c that might change at each occurrence but is independent of the size of the elements in the mesh, the approximate solution, and the control cost parameter λ.

3 Optimal control problem

The analysis of the optimal control problem (2)–(4) follows standard arguments [25,46] which we now briefly present. Let us introduce the so–called control to state map $S : L^2(\Omega) \to H_0^1(\Omega)$, which to a given control $u \in L^2(\Omega)$, associates the unique solution to

$$(\nabla \mathsf{y}, \nabla \mathsf{v})_{L^2(\Omega)} = (\mathsf{f} + \mathsf{u}, \mathsf{v})_{L^2(\Omega)} \quad \forall\, \mathsf{v} \in H_0^1(\Omega). \tag{9}$$

With this map at hand, we define the reduced cost functional

$$F : L^2(\Omega) \to \mathbb{R}, \quad F(u) := \tfrac{1}{2}\|S(\mathsf{u}) - \mathsf{y}_\Omega\|^2_{L^2(\Omega)} + \tfrac{\lambda}{2}\|\mathsf{u}\|^2_{L^2(\Omega)}.$$

The functional F is weakly lower semicontinuous and strictly convex ($\lambda > 0$). In addition, the set $\mathbb{U}_{\text{ad}}$ is a bounded, convex, closed, and nonempty subset of $L^2(\Omega)$. Thus, it is weakly sequentially compact. We then are able to conclude the existence and uniqueness of an optimal control $\bar{\mathsf{u}}$ and an optimal state $\bar{\mathsf{y}}$ that satisfy (9) [46, Theorem 2.14]. In addition, $\bar{\mathsf{u}}$ satisfies the first–order optimality condition $F'(\bar{\mathsf{u}})(\mathsf{u} - \bar{\mathsf{u}}) \geq 0$ for all $\mathsf{u} \in \mathbb{U}_{\text{ad}}$ [46, Lemma 2.21]. To explore this variational inequality, we define the adjoint state $\mathsf{p} \in H_0^1(\Omega)$ as the solution to

$$(\nabla \mathsf{w}, \nabla \mathsf{p})_{L^2(\Omega)} = (\mathsf{y} - \mathsf{y}_\Omega, \mathsf{w})_{L^2(\Omega)} \quad \forall\, \mathsf{w} \in H_0^1(\Omega). \tag{10}$$

With the adjoint state at hand, the aforementioned variational inequality reads:

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

where by $\bar{\mathsf{p}}$ we denoted the adjoint state associated to $\bar{\mathsf{y}}$.

We have thus arrived at the following optimality system for (2)–(4): Find $(\bar{\mathsf{y}}, \bar{\mathsf{p}}, \bar{\mathsf{u}}) \in H_0^1(\Omega) \times H_0^1(\Omega) \times \mathbb{U}_{\text{ad}}$ such that

$$\begin{cases} (\nabla \bar{\mathsf{y}}, \nabla \mathsf{v})_{L^2(\Omega)} = (\mathsf{f} + \bar{\mathsf{u}}, \mathsf{v})_{L^2(\Omega)} & \forall\, \mathsf{v} \in H_0^1(\Omega), \\ (\nabla \mathsf{w}, \nabla \bar{\mathsf{p}})_{L^2(\Omega)} = (\bar{\mathsf{y}} - \mathsf{y}_\Omega, \mathsf{w})_{L^2(\Omega)} & \forall\, \mathsf{w} \in H_0^1(\Omega), \\ (\bar{\mathsf{p}} + \lambda \bar{\mathsf{u}}, \mathsf{u} - \bar{\mathsf{u}})_{L^2(\Omega)} \geq 0 & \forall\, \mathsf{u} \in \mathbb{U}_{ad}. \end{cases} \tag{11}$$

The following result will be instrumental in the analysis that we will perform [8,9,16, 20,21,27,40].

Proposition 1 (higher integrability) *Let* $\mathsf{y} \in H_0^1(\Omega)$ *and* $\mathsf{p} \in H_0^1(\Omega)$ *denote the solutions to* (9) *and* (10)*, respectively. If* $\mathsf{f}, \mathsf{y}_\Omega \in L^2(\Omega)$*, then there exists* $r > d$ *such that* $\mathsf{y}, \mathsf{p} \in W^{1,r}(\Omega)$*. In addition,*

$$\|\mathsf{y}\|_{W^{1,r}(\Omega)} \lesssim \|\mathsf{f} + \mathsf{u}\|_{L^2(\Omega)}, \quad \|\mathsf{p}\|_{W^{1,r}(\Omega)} \lesssim \|\mathsf{y} - \mathsf{y}_\Omega\|_{L^2(\Omega)}, \tag{12}$$

where the hidden constants are independent of u, f, y, *and* y_Ω*. In particular, we have, for* $\sigma = 1 - d/r > 0$*, that* $\mathsf{y}, \mathsf{p} \in C^{0,\sigma}(\overline{\Omega})$ *with similar estimates.*

As a consequence of the results of Proposition 1, we have that $\bar{\mathsf{u}} \in C(\overline{\Omega}) \cap H^1(\Omega)$ since

$$\bar{\mathsf{u}} = \Pi\left(-\lambda^{-1}\bar{\mathsf{p}}\right), \qquad \Pi(w)(x) := \min\{\mathsf{b}, \max\{\mathsf{a}, w(x)\}\} \text{ for all } x \text{ in } \overline{\Omega}. \tag{13}$$

We recall that the operator Π, defined in (13) is continuous in $H^1(\Omega)$ and, for $\mathscr{G} \subset \Omega$, it is nonexpansive in $L^\infty(\mathscr{G})$, i.e.,

$$\|\Pi(w_1) - \Pi(w_2)\|_{L^\infty(\mathscr{G})} \leq \|w_1 - w_2\|_{L^\infty(\mathscr{G})} \quad \forall\, w_1, w_2 \in L^\infty(\mathscr{G}). \tag{14}$$

We approximate the solution of (11) by finding $(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T}) \times \mathbb{V}(\mathscr{T}) \times \mathbb{U}_{\text{ad}}(\mathscr{T})$ such that

$$\begin{cases} (\nabla \bar{\mathsf{y}}_{\mathscr{T}}, \nabla \mathsf{v}_{\mathscr{T}})_{L^2(\Omega)} = (\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}, \mathsf{v}_{\mathscr{T}})_{L^2(\Omega)} & \forall\, \mathsf{v}_{\mathscr{T}} \in \mathbb{V}(\mathscr{T}), \\ (\nabla \mathsf{w}_{\mathscr{T}}, \nabla \bar{\mathsf{p}}_{\mathscr{T}})_{L^2(\Omega)} = (\bar{\mathsf{y}}_{\mathscr{T}} - \mathsf{y}_{\Omega}, \mathsf{w}_{\mathscr{T}})_{L^2(\Omega)} & \forall\, \mathsf{w}_{\mathscr{T}} \in \mathbb{V}(\mathscr{T}), \\ (\bar{\mathsf{p}}_{\mathscr{T}} + \lambda \bar{\mathsf{u}}_{\mathscr{T}}, \mathsf{u}_{\mathscr{T}} - \bar{\mathsf{u}}_{\mathscr{T}})_{L^2(\Omega)} \geq 0 & \forall\, \mathsf{u}_{\mathscr{T}} \in \mathbb{U}_{ad}(\mathscr{T}). \end{cases} \tag{15}$$

We conclude by noticing that, in view of (12) and (13), it is appropriate to study a posteriori error estimation in $L^\infty(\Omega)$.

4 A posteriori error analysis: reliability

We now begin the a posteriori error estimation in the maximum norm. To accomplish this task, we define the local error indicators

$$\mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; T) = h_T^{2-d/2} \|\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} + h_T \|[\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!]\|_{L^\infty(\partial T \setminus \partial\Omega)}, \tag{16}$$

$$\mathscr{E}_{\mathsf{p}}(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; T) = h_T^{2-d/2} \|\bar{\mathsf{y}}_{\mathscr{T}} - \mathsf{y}_{\Omega}\|_{L^2(T)} + h_T \|[\![\nabla \bar{\mathsf{p}}_{\mathscr{T}} \cdot \nu]\!]\|_{L^\infty(\partial T \setminus \partial\Omega)}, \tag{17}$$

$$\mathscr{E}_{\mathsf{u}}(\bar{\mathsf{u}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}; T) = \|\Pi(-\bar{\mathsf{p}}_{\mathscr{T}}/\lambda) - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T)}, \tag{18}$$

$$\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; T) = \left(\mathscr{E}_{\mathsf{y}}^2(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; T) + \mathscr{E}_{\mathsf{p}}^2(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; T) + \mathscr{E}_{\mathsf{u}}^2(\bar{\mathsf{u}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}; T)\right)^{\frac{1}{2}}, \tag{19}$$

and the global a posteriori error estimators

$$\mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T}) = \max_{T \in \mathscr{T}} \mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; T), \tag{20}$$

$$\mathscr{E}_{\mathsf{p}}(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; \mathscr{T}) = \max_{T \in \mathscr{T}} \mathscr{E}_{\mathsf{p}}(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; T), \tag{21}$$

$$\mathscr{E}_{\mathsf{u}}(\bar{\mathsf{u}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}; \mathscr{T}) = \max_{T \in \mathscr{T}} \mathscr{E}_{\mathsf{u}}(\bar{\mathsf{u}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}; T), \tag{22}$$

$$\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T}) = \left(\mathscr{E}_{\mathsf{y}}^2(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T}) + \mathscr{E}_{\mathsf{p}}^2(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; \mathscr{T}) + \mathscr{E}_{\mathsf{u}}^2(\bar{\mathsf{u}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}; \mathscr{T})\right)^{\frac{1}{2}}. \tag{23}$$

We also introduce two auxiliary variables: $\widehat{\mathsf{y}}, \widehat{\mathsf{p}} \in H_0^1(\Omega)$ which solve, respectively,

$$(\nabla \widehat{\mathsf{y}}, \nabla \mathsf{v}) = (\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}, \mathsf{v})_{L^2(\Omega)} \quad \forall\, \mathsf{v} \in H_0^1(\Omega) \tag{24}$$

and

$$(\nabla \mathsf{w}, \nabla \widehat{\mathsf{p}}) = (\bar{\mathsf{y}}_{\mathscr{T}} - \mathsf{y}_{\Omega}, \mathsf{w})_{L^2(\Omega)} \quad \forall\, \mathsf{w} \in H_0^1(\Omega). \tag{25}$$

We note that, as a consequence of Proposition 1, $\widehat{\mathsf{y}}, \widehat{\mathsf{p}} \in H_0^1(\Omega) \cap C(\overline{\Omega})$ and thus we can invoke [3, Lemma 4.2] to conclude that

$$\|\widehat{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\Omega)} \lesssim \ell_{\mathscr{T}} \mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T}), \qquad \|\widehat{\mathsf{p}} - \bar{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(\Omega)} \lesssim \ell_{\mathscr{T}} \mathscr{E}_{\mathsf{p}}(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; \mathscr{T}). \tag{26}$$

Finally, for $e_{\bar{\mathsf{y}}} = \bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}$, $e_{\bar{\mathsf{p}}} = \bar{\mathsf{p}} - \bar{\mathsf{p}}_{\mathscr{T}}$, and $e_{\bar{\mathsf{u}}} = \bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}$, we define

$$\|(e_{\bar{\mathsf{y}}}, e_{\bar{\mathsf{p}}}, e_{\bar{\mathsf{u}}})\|_{\Omega}^2 := \|e_{\bar{\mathsf{y}}}\|_{L^\infty(\Omega)}^2 + \|e_{\bar{\mathsf{p}}}\|_{L^\infty(\Omega)}^2 + \|e_{\bar{\mathsf{u}}}\|_{L^\infty(\Omega)}^2. \tag{27}$$

Theorem 1 (global reliability) *Let* $(\bar{\mathsf{y}},\bar{\mathsf{p}},\bar{\mathsf{u}}) \in H_0^1(\Omega)\times H_0^1(\Omega)\times L^2(\Omega)$ *be the solution to* (11) *and* $(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}}) \in \mathbb{V}(\mathscr{T})\times\mathbb{V}(\mathscr{T})\times\mathbb{U}_{\mathrm{ad}}(\mathscr{T})$ *its numerical approximation obtained as the solution to* (15). *Then*

$$\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_\Omega \lesssim \max\{1,\lambda^{-2}\}\max\{1,\ell_{\mathscr{T}}\}\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};\mathscr{T}), \tag{28}$$

where the hidden constant is independent of the control cost parameter λ*, the size of the elements in the mesh* $\mathscr{T}$ *and* $\#\mathscr{T}$.

Proof We proceed in five steps.
Step 1. First we control the error $\|\bar{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(\Omega)}$. Define $\widetilde{\mathsf{u}} = \Pi(-\lambda^{-1}\bar{\mathsf{p}}_{\mathscr{T}})$. Notice that $\widetilde{\mathsf{u}}$ can be equivalently characterized by [46, Lemma 2.26]

$$(\bar{\mathsf{p}}_{\mathscr{T}}+\lambda\widetilde{\mathsf{u}},\mathsf{u}-\widetilde{\mathsf{u}})_{L^2(\Omega)} \geq 0 \quad \forall\, \mathsf{u}\in\mathbb{U}_{\mathrm{ad}}. \tag{29}$$

We bound $\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|_{L^2(\Omega)}$ by setting $\mathsf{u}=\widetilde{\mathsf{u}}$ in (11), $\mathsf{u}=\bar{\mathsf{u}}$ in (29) and adding the results to obtain

$$\lambda\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} \leq (\bar{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}},\widetilde{\mathsf{u}}-\bar{\mathsf{u}})_{L^2(\Omega)}.$$

To bound the right hand side of the previous expression, we let $(\widetilde{\mathsf{y}},\widetilde{\mathsf{p}})\in H_0^1(\Omega)\times H_0^1(\Omega)$ be such that

$$(\nabla\widetilde{\mathsf{y}},\nabla\mathsf{v})_{L^2(\Omega)} = (\mathsf{f}+\widetilde{\mathsf{u}},\mathsf{v})_{L^2(\Omega)}\ \forall\,\mathsf{v}\in H_0^1(\Omega)$$

and

$$(\nabla\mathsf{w},\nabla\widetilde{\mathsf{p}})_{L^2(\Omega)} = (\widetilde{\mathsf{y}}-\mathsf{y}_\Omega,\mathsf{w})_{L^2(\Omega)}\ \forall\,\mathsf{w}\in H_0^1(\Omega).$$

With the auxiliary adjoint state $\widetilde{\mathsf{p}}$ at hand, we thus arrive at

$$\lambda\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} \leq (\bar{\mathsf{p}}-\widetilde{\mathsf{p}},\widetilde{\mathsf{u}}-\bar{\mathsf{u}})_{L^2(\Omega)} + (\widetilde{\mathsf{p}}-\widehat{\mathsf{p}},\widetilde{\mathsf{u}}-\bar{\mathsf{u}})_{L^2(\Omega)} + (\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}},\widetilde{\mathsf{u}}-\bar{\mathsf{u}})_{L^2(\Omega)}. \tag{30}$$

We now observe that $(\nabla(\widetilde{\mathsf{y}}-\bar{\mathsf{y}}),\nabla\mathsf{v})_{L^2(\Omega)} = (\widetilde{\mathsf{u}}-\bar{\mathsf{u}},\mathsf{v})_{L^2(\Omega)}$ for all $\mathsf{v}\in H_0^1(\Omega)$ and $(\nabla\mathsf{w},\nabla(\bar{\mathsf{p}}-\widetilde{\mathsf{p}}))_{L^2(\Omega)} = (\bar{\mathsf{y}}-\widetilde{\mathsf{y}},\mathsf{w})_{L^2(\Omega)}$ for all $\mathsf{w}\in H_0^1(\Omega)$. Hence,

$$(\bar{\mathsf{p}}-\widetilde{\mathsf{p}},\widetilde{\mathsf{u}}-\bar{\mathsf{u}})_{L^2(\Omega)} = (\nabla(\widetilde{\mathsf{y}}-\bar{\mathsf{y}}),\nabla(\bar{\mathsf{p}}-\widetilde{\mathsf{p}}))_{L^2(\Omega)} = -\|\bar{\mathsf{y}}-\widetilde{\mathsf{y}}\|^2_{L^2(\Omega)} \leq 0.$$

In view of this, an application of the Cauchy–Schwarz inequality yields

$$\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} \lesssim \lambda^{-1}\left(\|\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}\|_{L^2(\Omega)}+\|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|_{L^2(\Omega)}\right)\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|_{L^2(\Omega)}$$

from which it follows that

$$\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} \lesssim \lambda^{-2}\left(\|\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}\|^2_{L^2(\Omega)}+\|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^2(\Omega)}\right). \tag{31}$$

We now control $\|\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}\|^2_{L^2(\Omega)}$. Since $(\nabla\mathsf{w},\nabla(\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}))_{L^2(\Omega)} = (\widetilde{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}},\mathsf{w})_{L^2(\Omega)}$ for all $\mathsf{w}\in H_0^1(\Omega)$, we have that

$$\begin{aligned}\|\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}\|^2_{L^2(\Omega)} &\lesssim (\nabla(\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}),\nabla(\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}))_{L^2(\Omega)} \\ &= (\widetilde{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}},\widetilde{\mathsf{p}}-\widehat{\mathsf{p}})_{L^2(\Omega)} \lesssim \left(\|\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}\|_{L^2(\Omega)}+\|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|_{L^2(\Omega)}\right)\|\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}\|_{L^2(\Omega)}.\end{aligned}$$

Consequently, $\|\widetilde{\mathsf{p}}-\widehat{\mathsf{p}}\|_{L^2(\Omega)} \lesssim \|\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}\|_{L^2(\Omega)}+\|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|_{L^2(\Omega)}$. This, in conjunction with (31), yields

$$\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} \lesssim \lambda^{-2}\left(\|\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}\|^2_{L^2(\Omega)}+\|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^2(\Omega)}+\|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^2(\Omega)}\right). \tag{32}$$

Upon observing that $(\nabla(\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}),\nabla\mathsf{v})_{L^2(\Omega)} = (\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}},\mathsf{v})_{L^2(\Omega)}$ for all $\mathsf{v}\in H_0^1(\Omega)$, we obtain

$$\|\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}\|^2_{L^2(\Omega)} \lesssim (\nabla(\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}),\nabla(\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}))_{L^2(\Omega)} = (\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}},\widetilde{\mathsf{y}}-\widehat{\mathsf{y}})_{L^2(\Omega)} \leq \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(\Omega)}\|\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}\|_{L^2(\Omega)}$$

and hence, $\|\widetilde{\mathsf{y}}-\widehat{\mathsf{y}}\|_{L^2(\Omega)} \lesssim \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(\Omega)}$. Combining this with (32) implies that

$$\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} \lesssim \lambda^{-2}\left(\|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^2(\Omega)} + \|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^2(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^2(\Omega)}\right).$$

The triangle inequality then allows us to conclude that

$$\begin{aligned}\|e_{\bar{\mathsf{u}}}\|^2_{L^2(\Omega)} &\lesssim \|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^2(\Omega)} + \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^2(\Omega)}\\ &\lesssim \max\left\{1,\lambda^{-2}\right\}\left(\|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^2(\Omega)} + \|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^2(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^2(\Omega)}\right)\\ &\lesssim \max\left\{1,\lambda^{-2}\right\}\left(\|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}\right). \quad (33)\end{aligned}$$

Step 2. In this step we control $\|\bar{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\Omega)}$. In view of Proposition 1 there exists $r>d$ such that

$$\|\bar{\mathsf{y}}-\widehat{\mathsf{y}}\|_{L^\infty(\Omega)} \lesssim \|\bar{\mathsf{y}}-\widehat{\mathsf{y}}\|_{W^{1,r}(\Omega)} \lesssim \|\bar{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(\Omega)}.$$

Consequently, the triangle inequality and (33) give us that

$$\begin{aligned}\|e_{\bar{\mathsf{y}}}\|^2_{L^\infty(\Omega)} &\lesssim \|\bar{\mathsf{y}}-\widehat{\mathsf{y}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}\\ &\lesssim \max\left\{1,\lambda^{-2}\right\}\left(\|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}\right). \quad (34)\end{aligned}$$

Step 3. To bound $\|\bar{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(\Omega)}$ we again use Proposition 1 to conclude that there exists $r>d$ such that

$$\|\bar{\mathsf{p}}-\widehat{\mathsf{p}}\|_{L^\infty(\Omega)} \lesssim \|\bar{\mathsf{p}}-\widehat{\mathsf{p}}\|_{W^{1,r}(\Omega)} \lesssim \|\bar{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|_{L^2(\Omega)} \lesssim \|\bar{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\Omega)}.$$

Thus, this estimate and (34) imply that

$$\begin{aligned}\|e_{\bar{\mathsf{p}}}\|^2_{L^\infty(\Omega)} &\lesssim \|\bar{\mathsf{p}}-\widehat{\mathsf{p}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}\\ &\lesssim \max\left\{1,\lambda^{-2}\right\}\left(\|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}\right). \quad (35)\end{aligned}$$

Step 4. The goal of this step is to control the error $\|\bar{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(\Omega)}$. We begin with the basic estimate

$$\|\bar{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} \lesssim \|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^\infty(\Omega)} + \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}. \quad (36)$$

Using (14) we have that

$$\begin{aligned}\|\bar{\mathsf{u}}-\widetilde{\mathsf{u}}\|^2_{L^\infty(\Omega)} = \|\Pi(-\lambda^{-1}\bar{\mathsf{p}}) - \Pi(-\lambda^{-1}\bar{\mathsf{p}}_{\mathscr{T}})\|^2_{L^\infty(\Omega)} &\leq \|-\lambda^{-1}\bar{\mathsf{p}} - (-\lambda^{-1}\bar{\mathsf{p}}_{\mathscr{T}})\|^2_{L^\infty(\Omega)}\\ &= \lambda^{-2}\|\bar{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}.\end{aligned}$$

Therefore, by using (35), we can conclude that

$$\|e_{\bar{\mathsf{u}}}\|^2_{L^\infty(\Omega)} \lesssim \max\left\{1,\lambda^{-4}\right\}\left(\|\widehat{\mathsf{y}}-\bar{\mathsf{y}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widehat{\mathsf{p}}-\bar{\mathsf{p}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)} + \|\widetilde{\mathsf{u}}-\bar{\mathsf{u}}_{\mathscr{T}}\|^2_{L^\infty(\Omega)}\right) \quad (37)$$

upon using the fact that $1+\lambda^{-2}\max\left\{1,\lambda^{-2}\right\} \lesssim \max\left\{1,\lambda^{-2},\lambda^{-4}\right\} \leq \max\left\{1,\lambda^{-4}\right\}$ since

$$\lambda^{-2} \leq \frac{1}{2}\left(1+\lambda^{-4}\right) \leq \max\{1,\lambda^{-4}\}. \quad (38)$$

Step 5. The claimed result follows upon gathering (34), (35) and (37), and using (22), (26) and (38). □

5 A posteriori error analysis: efficiency

Let $\mathscr{P}_{\mathscr{T}}$ denote the L^2-projection onto piecewise linear functions over $\mathscr{T}$. For $g \in L^2(\Omega)$ and $\mathscr{M} \subset \mathscr{T}$ we define

$$\operatorname{osc}_{\mathscr{T}}(g;\mathscr{M}) = \left(\sum_{T\in\mathscr{M}} h_T^{4-d} \|g - \mathscr{P}_{\mathscr{T}} g\|_{L^2(T)}^2 \right)^{\frac{1}{2}}. \tag{39}$$

Lemma 1 (local efficiency of $\mathscr{E}_{\mathsf{y}}$) *In the setting of Theorem 1 we have that*

$$\mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; T) \lesssim \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} + h_T^2 \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} + \operatorname{osc}_{\mathscr{T}}(\mathsf{f}; \mathscr{N}_T), \tag{40}$$

where the hidden constant is independent of the control cost parameter λ, the size of the elements in the mesh $\mathscr{T}$, and $\#\mathscr{T}$.

Proof Let $\mathsf{v} \in H_0^1(\Omega)$ be such that $\mathsf{v}_{|T} \in C^2(T)$ for all $T \in \mathscr{T}$. Using (11) and integrating by parts yields

$$\int_\Omega \nabla(\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) \cdot \nabla \mathsf{v} = \sum_{T\in\mathscr{T}} \int_T (\mathsf{f} + \bar{\mathsf{u}})\mathsf{v} + \sum_{S\in\mathscr{S}} \int_S [\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!] \mathsf{v}.$$

Since on each $T \in \mathscr{T}$ we have that $\mathsf{v} \in C^2(T)$, we again apply integration by parts to conclude that

$$\int_\Omega \nabla(\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) \cdot \nabla \mathsf{v} = -\sum_{T\in\mathscr{T}} \int_T \Delta \mathsf{v} (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) - \sum_{S\in\mathscr{S}} \int_S [\![\nabla \mathsf{v} \cdot \nu]\!] (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}).$$

In conclusion, since the left hand sides of the previous expressions coincide, we arrive at the identity

$$\begin{aligned} \sum_{T\in\mathscr{T}} \int_T (\mathsf{f} + \bar{\mathsf{u}})\mathsf{v} + \sum_{S\in\mathscr{S}} \int_S [\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!] \mathsf{v} = &-\sum_{T\in\mathscr{T}} \int_T \Delta \mathsf{v} (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) \\ &- \sum_{S\in\mathscr{S}} \int_S [\![\nabla \mathsf{v} \cdot \nu]\!] (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}), \end{aligned} \tag{41}$$

for every $\mathsf{v} \in H_0^1(\Omega)$ such that $\mathsf{v}_{|T} \in C^2(T)$ for all $T \in \mathscr{T}$. We now proceed, on the basis of (16), in two steps.

Step 1. Let $T \in \mathscr{T}$. We begin with the basic estimate

$$h_T^{2-d/2} \|\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \le h_T^{2-d/2} \|\mathscr{P}_{\mathscr{T}} \mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} + h_T^{2-d/2} \|\mathsf{f} - \mathscr{P}_{\mathscr{T}} \mathsf{f}\|_{L^2(T)}. \tag{42}$$

By letting $\mathsf{v} = \beta_T = (\mathscr{P}_{\mathscr{T}} \mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}) \varphi_T^2$ in (41), where φ_T is the standard bubble function over T [1, 47], we obtain that

$$\int_T (\mathscr{P}_{\mathscr{T}} \mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}) \beta_T = -\int_T \Delta \beta_T (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) - \int_T (\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}) \beta_T - \int_T (\mathsf{f} - \mathscr{P}_{\mathscr{T}} \mathsf{f}) \beta_T, \tag{43}$$

since $\int_S [\![\nabla \beta_T \cdot \nu]\!] (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) = 0$ for all $S \in \mathscr{S}$. We now bound each term on the right–hand side of (43) separately. Since $\Delta(\mathscr{P}_{\mathscr{T}} \mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}) = 0$ on T, we have that $\Delta \beta_T = 4\nabla(\mathscr{P}_{\mathscr{T}} \mathsf{f} +$

$\bar{\mathsf{u}}_{\mathscr{T}}) \cdot \nabla \varphi_T \varphi_T + 2(\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}})(\varphi_T \Delta \varphi_T + \nabla \varphi_T \cdot \nabla \varphi_T)$. This equality, the properties of the bubble function φ_T and an inverse inequality allow us to conclude that

$$\begin{aligned}
&\left| \int_T \Delta \beta_T (\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}) \right| \\
&\lesssim \left(h_T^{d/2-1} \|\nabla(\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}})\|_{L^2(T)} + h_T^{d/2-2} \|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \right) \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(T)} \\
&\lesssim h_T^{d/2-2} \|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(T)}.
\end{aligned}$$

In addition, we have that

$$\begin{aligned}
\left| \int_T (\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}) \beta_T \right| &\lesssim \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \\
&\lesssim h_T^{d/2} \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T)} \|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)}
\end{aligned}$$

and

$$\left| \int_T (\mathsf{f} - \mathscr{P}_{\mathscr{T}}\mathsf{f}) \beta_T \right| \lesssim \|\mathsf{f} - \mathscr{P}_{\mathscr{T}}\mathsf{f}\|_{L^2(T)} \|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)}.$$

In view of the fact that $\|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)}^2 \lesssim \int_T (\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}) \beta_T$, the previous findings allow us to state that

$$h_T^{2-d/2} \|\mathscr{P}_{\mathscr{T}}\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \lesssim \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(T)} + h_T^2 \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T)} + h_T^{2-d/2} \|\mathsf{f} - \mathscr{P}_{\mathscr{T}}\mathsf{f}\|_{L^2(T)}.$$

Consequently, using (42) we conclude that

$$h_T^{2-d/2} \|\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T)} \lesssim \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(T)} + h_T^2 \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T)} + \operatorname{osc}_{\mathscr{T}}(\mathsf{f}; T). \qquad (44)$$

Step 2. Let $T \in \mathscr{T}$ and $S \in \mathscr{S}_T$. We proceed to control $h_T \|[\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!]\|_{L^\infty(S)}$. To accomplish this task, we use the property

$$|S| \|[\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!]\|_{L^\infty(S)} \lesssim \left| \int_S [\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!] \varphi_S \right|, \qquad (45)$$

of φ_S, the standard bubble function over S [1,47]. We now let $\mathsf{v} = \varphi_S$ in (41) and arrive at

$$\begin{aligned}
&\left| \int_S [\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!] \varphi_S \right| \\
&\leq \sum_{T' \in \mathscr{N}_S} \int_{T'} |\mathsf{f} + \bar{\mathsf{u}}| \varphi_S + \sum_{T' \in \mathscr{N}_S} \int_{T'} |\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}| |\Delta \varphi_S| + \sum_{T' \in \mathscr{N}_S} \sum_{S' \in \mathscr{S}_{T'}} \int_{S'} |\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}| |[\![\nabla \varphi_S \cdot \nu]\!]| \\
&\lesssim \sum_{T' \in \mathscr{N}_S} |T'|^{1/2} \left(\|\mathsf{f} + \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^2(T')} + |T'| \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T')} \right) \\
&+ \sum_{T' \in \mathscr{N}_S} \left(h_S^{-2} |T'| + h_S^{-1} \sum_{S' \in \mathscr{S}_{T'}} |S'| \right) \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(T')}.
\end{aligned}$$

In view of the fact that $h_T |S|^{-1} \approx h_T^{2-d}$, the previous estimate combined with (45) and (44) yields the bound

$$h_T \|[\![\nabla \bar{\mathsf{y}}_{\mathscr{T}} \cdot \nu]\!]\|_{L^\infty(S)} \lesssim h_T^2 \|\bar{\mathsf{u}} - \bar{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_S)} + \|\bar{\mathsf{y}} - \bar{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_S)} + \operatorname{osc}_{\mathscr{T}}(\mathsf{f}; \mathscr{N}_S).$$

We finally combine the results of Steps 1 and 2 and arrive at the desired estimate (40). This concludes the proof. □

Similar arguments to the ones elaborated in the proof of Lemma 1 allow us to conclude the following result.

Lemma 2 (local efficiency of $\mathscr{E}_{\mathsf{p}}$) *In the setting of Theorem 1 we have that*

$$\mathscr{E}_{\mathsf{p}}(\overline{\mathsf{p}}_{\mathscr{T}},\overline{\mathsf{y}}_{\mathscr{T}};T) \lesssim \|\overline{\mathsf{p}}-\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} + h_T^2\|\overline{\mathsf{y}}-\overline{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} + \operatorname{osc}_{\mathscr{T}}(\mathsf{y}_\Omega;\mathscr{N}_T), \tag{46}$$

where the hidden constant is independent of the control cost parameter λ, the size of the elements in the mesh $\mathscr{T}$ and $\#\mathscr{T}$.

Lemma 3 (local efficiency of $\mathscr{E}_{\mathsf{u}}$) *In the setting of Theorem 1 we have that*

$$\mathscr{E}_{\mathsf{u}}(\overline{\mathsf{u}}_{\mathscr{T}},\overline{\mathsf{p}}_{\mathscr{T}};T) \lesssim \|\overline{\mathsf{u}}-\overline{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T)} + \lambda^{-1}\|\overline{\mathsf{p}}-\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(T)}, \tag{47}$$

where the hidden constant is independent of the control cost parameter λ, the size of the elements in the mesh $\mathscr{T}$ and $\#\mathscr{T}$.

Proof Since $\widetilde{\mathsf{u}} = \Pi(-\lambda^{-1}\overline{\mathsf{p}}_{\mathscr{T}})$, definitions (18) and (13) reveal that

$$\mathscr{E}_{\mathsf{u}}(\overline{\mathsf{u}}_{\mathscr{T}},\overline{\mathsf{p}}_{\mathscr{T}};T) \leq \|\Pi(-\lambda^{-1}\overline{\mathsf{p}}_{\mathscr{T}}) - \Pi(-\lambda^{-1}\overline{\mathsf{p}})\|_{L^\infty(T)} + \|\overline{\mathsf{u}}-\overline{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(T)}.$$

The Lipschitz property (14) allows us to conclude that

$$\|\Pi(-\lambda^{-1}\overline{\mathsf{p}}_{\mathscr{T}}) - \Pi(-\lambda^{-1}\overline{\mathsf{p}})\|_{L^\infty(T)} \lesssim \|\lambda^{-1}\overline{\mathsf{p}} - \lambda^{-1}\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(T)} = \lambda^{-1}\|\overline{\mathsf{p}}-\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(T)}$$

from which the lemma follows. □

The results of Lemmas 1, 2, and 3 immediately yield the following result upon observing that Ω is bounded.

Theorem 2 (local and global efficiency of $\mathscr{E}$) *In the setting of Theorem 1 we have that*

$$\begin{aligned}\mathscr{E}(\overline{\mathsf{y}}_{\mathscr{T}},\overline{\mathsf{p}}_{\mathscr{T}},\overline{\mathsf{u}}_{\mathscr{T}};T) \lesssim & \|\overline{\mathsf{y}}-\overline{\mathsf{y}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} + \|\overline{\mathsf{p}}-\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} + \|\overline{\mathsf{u}}-\overline{\mathsf{u}}_{\mathscr{T}}\|_{L^\infty(\mathscr{N}_T)} \\ & + \lambda^{-1}\|\overline{\mathsf{p}}-\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(T)} + \operatorname{osc}_{\mathscr{T}}(\mathsf{f};\mathscr{N}_T) + \operatorname{osc}_{\mathscr{T}}(\mathsf{y}_\Omega;\mathscr{N}_T),\end{aligned} \tag{48}$$

and

$$\begin{aligned}\mathscr{E}(\overline{\mathsf{y}}_{\mathscr{T}},\overline{\mathsf{p}}_{\mathscr{T}},\overline{\mathsf{u}}_{\mathscr{T}};\mathscr{T}) \lesssim & \|(e_{\overline{\mathsf{y}}},e_{\overline{\mathsf{p}}},e_{\overline{\mathsf{u}}})\|_\Omega + \lambda^{-1}\|\overline{\mathsf{p}}-\overline{\mathsf{p}}_{\mathscr{T}}\|_{L^\infty(\Omega)} \\ & + \max_{T\in\mathscr{T}} \operatorname{osc}_{\mathscr{T}}(\mathsf{f};\mathscr{N}_T) + \max_{T\in\mathscr{T}} \operatorname{osc}_{\mathscr{T}}(\mathsf{y}_\Omega;\mathscr{N}_T),\end{aligned} \tag{49}$$

where the hidden constants are independent of the control cost parameter λ, the size of the elements in the mesh $\mathscr{T}$ and $\#\mathscr{T}$.

Remark 1 (weighted norm estimates) Motivated by the fact that in the definition of the cost functional (1) the weights of the involved norms do not have the same dependence with respect to λ, we could also measure the error using a norm that includes a weight $\alpha = \alpha(\lambda)$:

$$\|(e_{\overline{\mathsf{y}}},e_{\overline{\mathsf{p}}},e_{\overline{\mathsf{u}}})\|_{\alpha,\Omega}^2 := \|e_{\overline{\mathsf{y}}}\|_{L^\infty(\Omega)}^2 + \|e_{\overline{\mathsf{p}}}\|_{L^\infty(\Omega)}^2 + \alpha\|e_{\overline{\mathsf{u}}}\|_{L^\infty(\Omega)}^2. \tag{50}$$

If this norm is used to measure the error, then we can obtain the global reliability estimate

$$\|(e_{\overline{\mathsf{y}}},e_{\overline{\mathsf{p}}},e_{\overline{\mathsf{u}}})\|_{\alpha,\Omega} \lesssim \max\{1,\lambda^{-1},\alpha^{1/2},\alpha^{1/2}\lambda^{-2}\}\max\{1,\ell_{\mathscr{T}}\}\mathscr{E}(\overline{\mathsf{y}}_{\mathscr{T}},\overline{\mathsf{p}}_{\mathscr{T}},\overline{\mathsf{u}}_{\mathscr{T}};\mathscr{T}) \tag{51}$$

using (34), (35) and (37). We can also obtain the global efficiency estimate

$$\begin{aligned}\mathscr{E}(\overline{\mathsf{y}}_{\mathscr{T}},\overline{\mathsf{p}}_{\mathscr{T}},\overline{\mathsf{u}}_{\mathscr{T}};\mathscr{T}) \lesssim & \max\{1,\lambda^{-1},\alpha^{-1/2}\}\|(e_{\overline{\mathsf{y}}},e_{\overline{\mathsf{p}}},e_{\overline{\mathsf{u}}})\|_{\alpha,\Omega} \\ & + \max_{T\in\mathscr{T}} \operatorname{osc}_{\mathscr{T}}(\mathsf{f};\mathscr{N}_T) + \max_{T\in\mathscr{T}} \operatorname{osc}_{\mathscr{T}}(\mathsf{y}_\Omega;\mathscr{N}_T)\end{aligned} \tag{52}$$

using (40), (46) and (47). We remark that for certain choices of the weight α, such as $\alpha = \lambda$ or $\alpha = \lambda^2$, the dependence of the factor in the reliability bound on the control cost parameter λ is better than that in the result in Theorem 1 when λ is small.

6 Numerical examples

We illustrate the performance of the devised a posteriori error estimator with numerical examples. For each example, a sequence of adaptively refined meshes was generated from an initial mesh by using a maximum strategy: we mark elements T for refinement if

$$\mathscr{E}^2(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};T) > \frac{1}{2}\max_{T'\in\mathscr{T}} \mathscr{E}^2(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};T').$$

The number of degrees of freedom Ndof on a particular mesh is three times the number of vertices in the mesh. The initial meshes are shown in Figure 1.

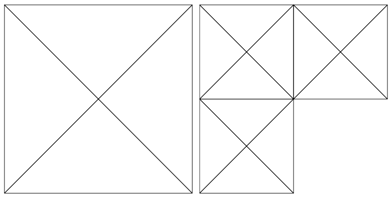

Fig. 1 The initial meshes for Example 1 (left) and Example 2 (right).

6.1 Example 1

We set $d=2$, $\Omega=(0,1)^2$, $\mathsf{a}=0$, $\mathsf{b}=1000000$, and the data f and y_Ω to be such that

$$\bar{\mathsf{y}}(x_1,x_2) = x_1x_2(1-x_1)(1-x_2)$$

and

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

Note that $\bar{\mathsf{u}}$ is given by (13). In order to investigate how the value of the control cost parameter affects the behavior of the a posteriori error estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};\mathscr{T})$, we performed this example for $\lambda\in\{1,0.1,0.01,0.001\}$.

The results are shown in Figures 2 and 3. Since we are approximating $\bar{\mathsf{y}}$, $\bar{\mathsf{p}}$, and $\bar{\mathsf{u}}$ using piecewise linear functions, the optimal behavior of the error $\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_\Omega$ and the estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};\mathscr{T})$ is $C\mathrm{Ndof}^{-1}$. We hence also plot $C\mathrm{Ndof}^{-1}$ for constants C that have been chosen so that this quantity is close to the dominating quantity on each plot. We can observe that, once the mesh has been sufficiently refined, the error $\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_\Omega$ and estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};\mathscr{T})$ are decreasing at the optimal rate. Note that we have proved that $\mathscr{E}$ is a reliable and efficient a posteriori error estimator for $\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_\Omega$. However, we have not proved that $\mathscr{E}_{\mathsf{y}}$, $\mathscr{E}_{\mathsf{p}}$, or $\mathscr{E}_{\mathsf{u}}$ is reliable and efficient for $\|e_{\bar{\mathsf{y}}}\|_{L^\infty(\Omega)}$, $\|e_{\bar{\mathsf{p}}}\|_{L^\infty(\Omega)}$, or $\|e_{\bar{\mathsf{u}}}\|_{L^\infty(\Omega)}$, respectively. Nevertheless, in Figure 2, we also show the errors $\|e_{\bar{\mathsf{y}}}\|_{L^\infty(\Omega)}$, $\|e_{\bar{\mathsf{p}}}\|_{L^\infty(\Omega)}$, and $\|e_{\bar{\mathsf{u}}}\|_{L^\infty(\Omega)}$ and the contributions $\mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};\mathscr{T})$, $\mathscr{E}_{\mathsf{p}}(\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{y}}_{\mathscr{T}};\mathscr{T})$, and $\mathscr{E}_{\mathsf{u}}(\bar{\mathsf{u}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}};\mathscr{T})$; it can be observed that optimal experimental rates of convergence are achieved. We also show the weighted error $\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_{\lambda,\Omega}$ which is defined by (50) with $\alpha=\lambda$. We observe that, for the smaller values of λ, the designed estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}},\bar{\mathsf{p}}_{\mathscr{T}},\bar{\mathsf{u}}_{\mathscr{T}};\mathscr{T})$ provides a better estimate of the weighted error $\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_{\lambda,\Omega}$ than the error $\|(e_{\bar{\mathsf{y}}},e_{\bar{\mathsf{p}}},e_{\bar{\mathsf{u}}})\|_\Omega$.

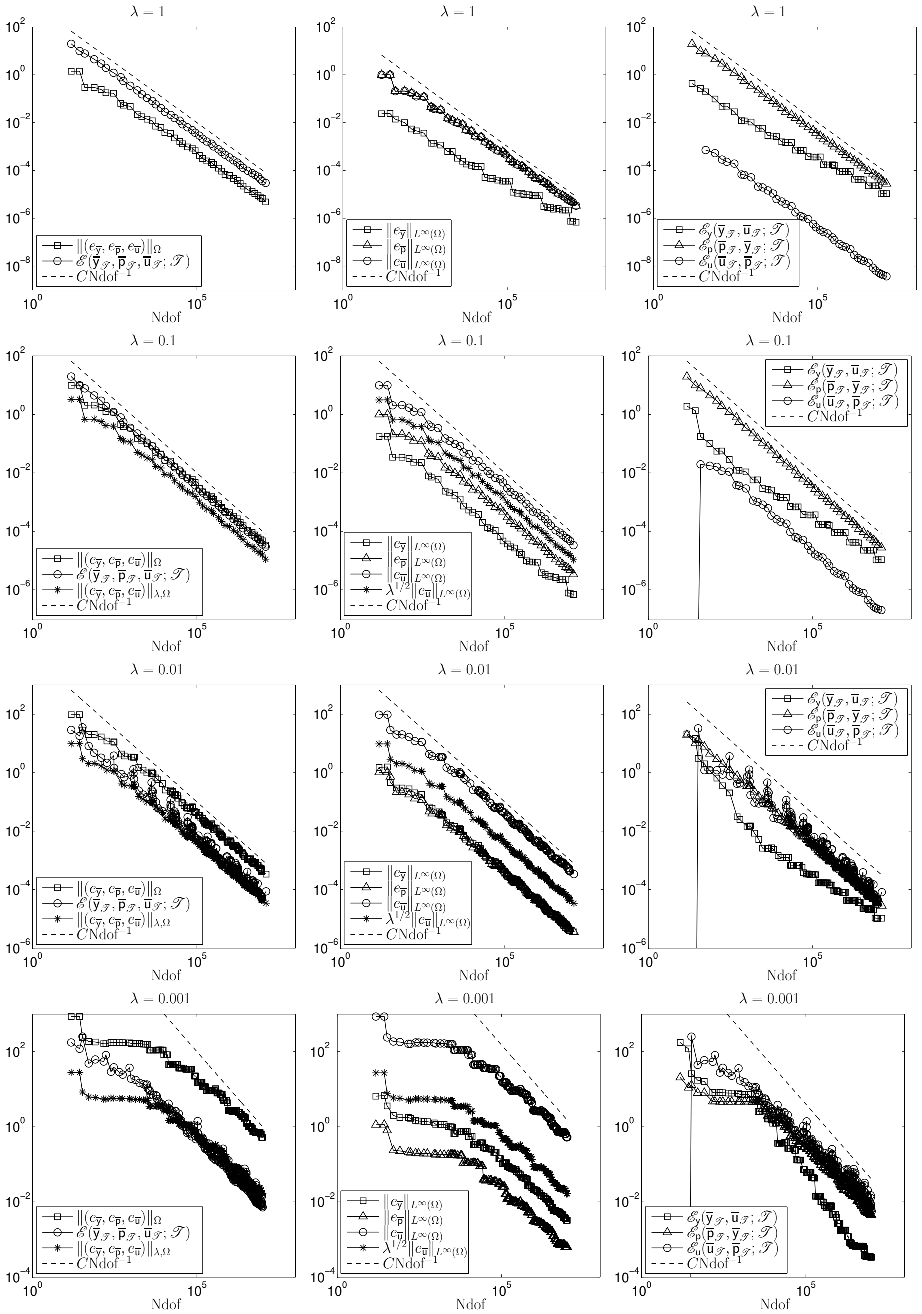

Fig. 2 Example 1: For $\lambda \in \{1, 0.1, 0.01, 0.001\}$, the error $\|(e_{\bar{\mathsf y}}, e_{\bar{\mathsf p}}, e_{\bar{\mathsf u}})\|_\Omega$, the estimator $\mathscr{E}(\bar{\mathsf y}_{\mathscr T}, \bar{\mathsf p}_{\mathscr T}, \bar{\mathsf u}_{\mathscr T}; \mathscr T)$, and the weighted error $\|(e_{\bar{\mathsf y}}, e_{\bar{\mathsf p}}, e_{\bar{\mathsf u}})\|_{\lambda,\Omega}$ (left), the errors $\|e_{\bar{\mathsf y}}\|_{L^\infty(\Omega)}$, $\|e_{\bar{\mathsf p}}\|_{L^\infty(\Omega)}$, and $\|e_{\bar{\mathsf u}}\|_{L^\infty(\Omega)}$ and the weighted error $\lambda^{1/2}\|e_{\bar{\mathsf u}}\|_{L^\infty(\Omega)}$ (center), and the contributions to the estimator $\mathscr{E}_{\mathsf y}(\bar{\mathsf y}_{\mathscr T}, \bar{\mathsf u}_{\mathscr T}; \mathscr T)$, $\mathscr{E}_{\mathsf p}(\bar{\mathsf p}_{\mathscr T}, \bar{\mathsf y}_{\mathscr T}; \mathscr T)$, and $\mathscr{E}_{\mathsf u}(\bar{\mathsf u}_{\mathscr T}, \bar{\mathsf p}_{\mathscr T}; \mathscr T)$ (right).

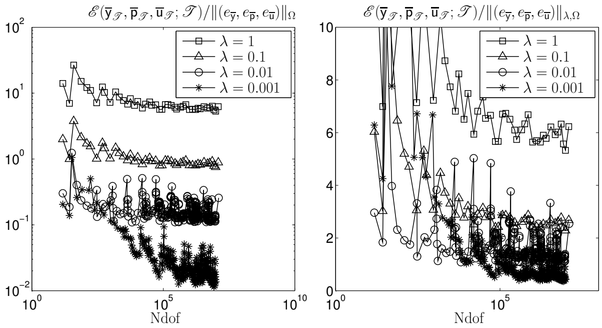

Fig. 3 Example 1: For $\lambda \in \{1, 0.1, 0.01, 0.001\}$, the effectivity indices $\mathscr{E}(\bar{y}_\mathscr{T},\bar{p}_\mathscr{T},\bar{u}_\mathscr{T};\mathscr{T})/\|(e_{\bar{y}},e_{\bar{p}},e_{\bar{u}})\|_\Omega$ (left) and $\mathscr{E}(\bar{y}_\mathscr{T},\bar{p}_\mathscr{T},\bar{u}_\mathscr{T};\mathscr{T})/\|(e_{\bar{y}},e_{\bar{p}},e_{\bar{u}})\|_{\lambda,\Omega}$ (right).

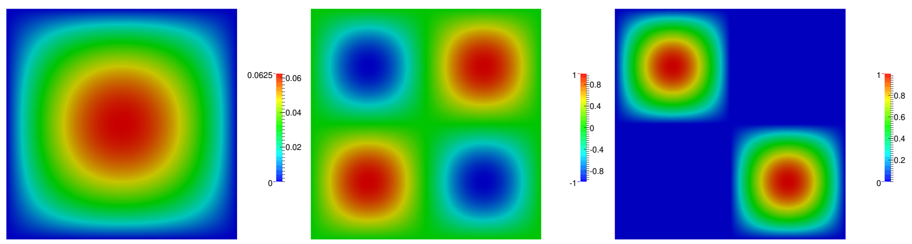

Fig. 4 Example 1: For $\lambda = 1$, the approximate solutions $\bar{y}_\mathscr{T}$ (left), $\bar{p}_\mathscr{T}$ (center), and $\bar{u}_\mathscr{T}$ (right) obtained on an adaptively refined mesh containing 10081 vertices.

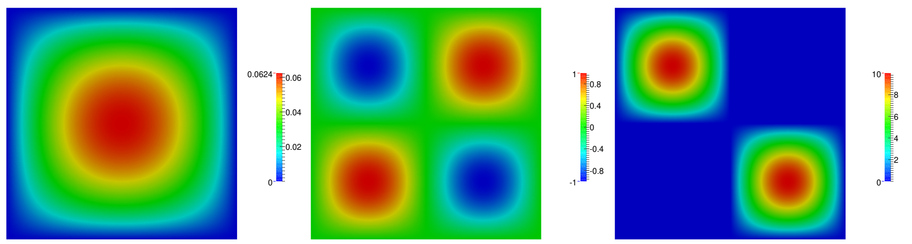

Fig. 5 Example 1: For $\lambda = 0.1$, the approximate solutions $\bar{y}_\mathscr{T}$ (left), $\bar{p}_\mathscr{T}$ (center), and $\bar{u}_\mathscr{T}$ (right) obtained on an adaptively refined mesh containing 10081 vertices.

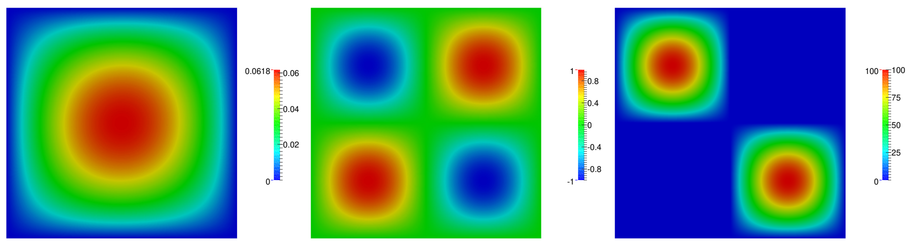

Fig. 6 Example 1: For $\lambda = 0.01$, the approximate solutions $\bar{y}_\mathscr{T}$ (left), $\bar{p}_\mathscr{T}$ (center), and $\bar{u}_\mathscr{T}$ (right) obtained on an adaptively refined mesh containing 10609 vertices.

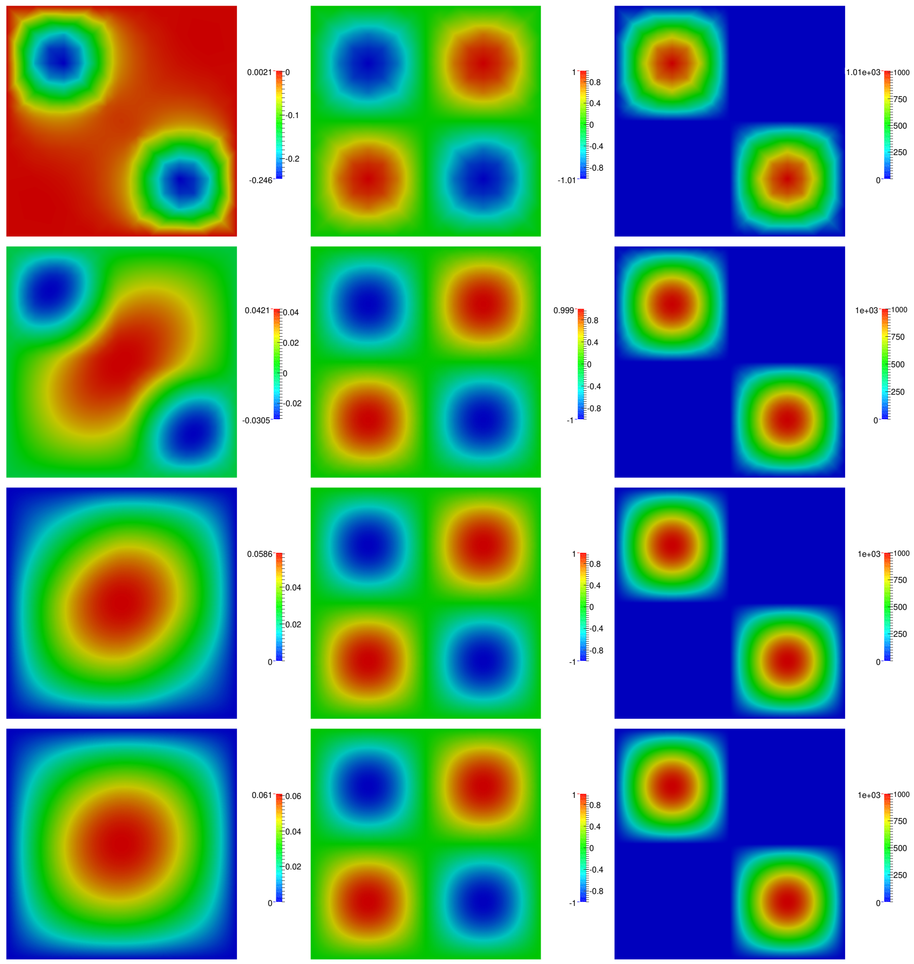

Fig. 7 Example 1: For $\lambda = 0.001$, the approximate solutions $\bar{\mathsf{y}}_{\mathscr{T}}$ (left), $\bar{\mathsf{p}}_{\mathscr{T}}$ (center), and $\bar{\mathsf{u}}_{\mathscr{T}}$ (right) obtained on adaptively refined meshes containing, from top to bottom, 10353, 110545, 1003629, and 3336999 vertices.

In Figures 4, 5, 6, and 7 we show some of the approximate solutions $\bar{\mathsf{y}}_{\mathscr{T}}$, $\bar{\mathsf{p}}_{\mathscr{T}}$, and $\bar{\mathsf{u}}_{\mathscr{T}}$ obtained and we observe that one of the control constraints is always active. This example has been constructed so that $\bar{\mathsf{y}}$ and $\bar{\mathsf{p}}$ do not depend on λ but $\bar{\mathsf{u}}$ does. We note that for all of the values of λ, the approximate solutions $\bar{\mathsf{p}}_{\mathscr{T}}$ and $\bar{\mathsf{u}}_{\mathscr{T}}$ appear similar on all of the meshes for which they are shown. However, for $\lambda = 0.001$, a lot of adaptive refinement had to be performed in order for $\bar{\mathsf{y}}_{\mathscr{T}}$ to reach the level of accuracy that was reached for the other values of λ with far less degrees of freedom. For $\lambda = 0.001$, the error in $\bar{\mathsf{u}}_{\mathscr{T}}$ is extremely dominant and, as previously noted, the estimator that we are using is an estimator for $\|(e_{\bar{\mathsf{y}}}, e_{\bar{\mathsf{p}}}, e_{\bar{\mathsf{u}}})\|_{\Omega}$.

We conclude that the a posteriori error estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T})$ performs well when the control cost parameter is not too small but that, in agreement with what we have proved, the estimator is not robust with respect to the control cost parameter λ.

6.2 Example 2

We set $d = 2$, $\Omega = (-1,1)^2 \setminus [0,1) \times (-1,0]$, $\mathsf{a} = 0$, $\mathsf{b} = 1$ and the data f and y_Ω to be such that, in polar coordinates (r,θ) with $\theta \in [0, 3\pi/2]$,

$$\bar{\mathsf{y}} = (1 - r^2\cos^2(\theta))(1 - r^2\sin^2(\theta))r^{2/3}\sin(2\theta/3)$$

and

$$\bar{\mathsf{p}} = \sin(2\pi r\cos(\theta))\sin(2\pi r\sin(\theta))r^{2/3}\sin(2\theta/3).$$

As in the previous example, $\bar{\mathsf{u}}$ is given by (13). For this example we took $\lambda = 1$.

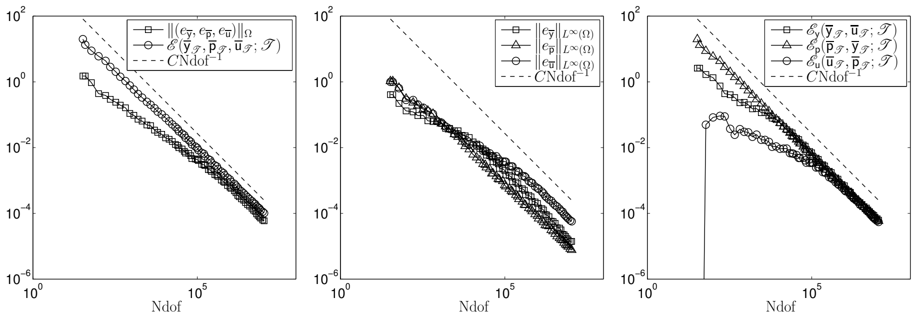

Fig. 8 Example 2: The error $\|(e_{\bar{\mathsf{y}}}, e_{\bar{\mathsf{p}}}, e_{\bar{\mathsf{u}}})\|_\Omega$ and the estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T})$ (left), the errors $\|e_{\bar{\mathsf{y}}}\|_{L^\infty(\Omega)}$, $\|e_{\bar{\mathsf{p}}}\|_{L^\infty(\Omega)}$, and $\|e_{\bar{\mathsf{u}}}\|_{L^\infty(\Omega)}$(center), and the contributions to the estimator $\mathscr{E}_{\mathsf{y}}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T})$, $\mathscr{E}_{\mathsf{p}}(\bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{y}}_{\mathscr{T}}; \mathscr{T})$, and $\mathscr{E}_{\mathsf{u}}(\bar{\mathsf{u}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}; \mathscr{T})$ (right).

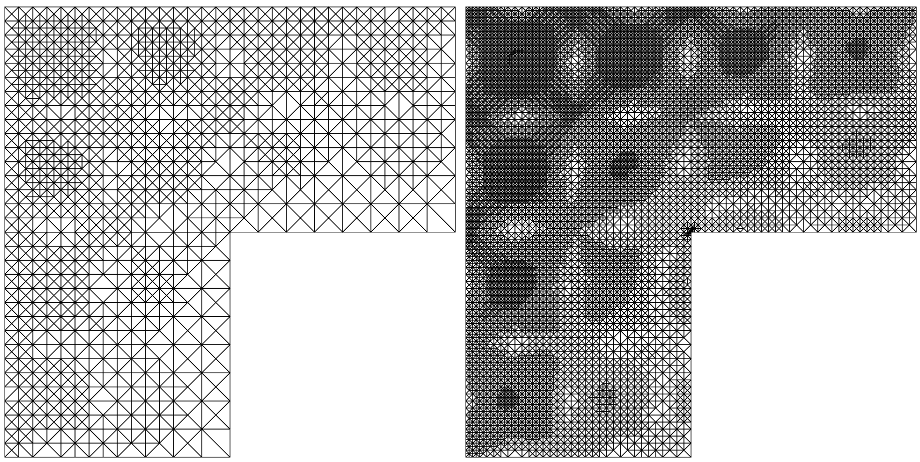

Fig. 9 Example 2: Adaptively refined meshes containing 1264 (left) and 11545 (right) vertices.

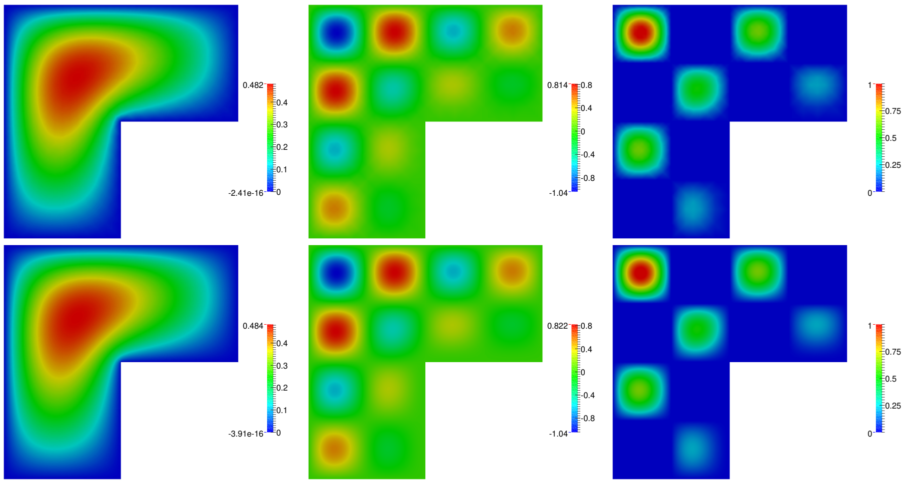

Fig. 10 Example 2: The approximate solutions $\bar{\mathsf{y}}_{\mathscr{T}}$ (left), $\bar{\mathsf{p}}_{\mathscr{T}}$ (center), and $\bar{\mathsf{u}}_{\mathscr{T}}$ (right) obtained on adaptively refined meshes containing 1264 (top) and 11545 (bottom) vertices.

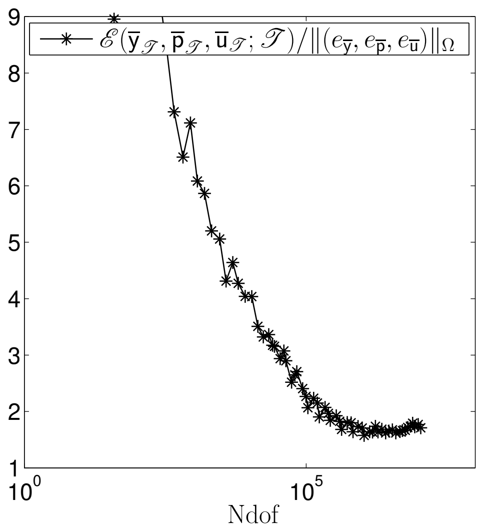

Fig. 11 Example 2: The effectivity indices $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T})/\|(e_{\bar{\mathsf{y}}}, e_{\bar{\mathsf{p}}}, e_{\bar{\mathsf{u}}})\|_\Omega$.

Figure 8 shows the results and we can observe that, once the mesh has been sufficiently refined, the error $\|(e_{\bar{\mathsf{y}}}, e_{\bar{\mathsf{p}}}, e_{\bar{\mathsf{u}}})\|_\Omega$ and estimator $\mathscr{E}(\bar{\mathsf{y}}_{\mathscr{T}}, \bar{\mathsf{p}}_{\mathscr{T}}, \bar{\mathsf{u}}_{\mathscr{T}}; \mathscr{T})$ are decreasing at the optimal rate. Adaptively refined meshes are shown in Figure 9 and the approximate solutions obtained on these meshes are shown in Figure 10 where we can see that both control constraints are active. The estimator performs well for this example and we observe from Figure 11 that the effectivity index is above and, once the mesh has been sufficiently refined, close to 1.

Acknowledgements E. Otárola was supported in part by CONICYT through FONDECYT project 11180193. A. J. Salgado was supported in part by NSF grant DMS-1418784. R. Rankin was supported in part by Universidad de Chile through BASAL PFB03 CMM project. The authors would like to thank Alejandro Allendes.

References

1. Ainsworth, M., Oden, J.T.: A posteriori error estimation in finite element analysis. Wiley-Interscience, New York (2000)
2. Allendes, A., Otárola, E., Rankin, R.: A posteriori error estimation for a PDE-constrained optimization problem involving the generalized Oseen equations. SIAM J. Sci. Comput. **40**(4), A2200–A2233 (2018)
3. Allendes, A., Otárola, E., Rankin, R., Salgado, A.J.: An *a posteriori* error analysis for an optimal control problem with point sources. ESAIM Math. Model. Numer. Anal. **52**(5), 1617–1650 (2018)
4. Apel, T., Rösch, A.A., Sirch, D.: L^{∞}-error estimates on graded meshes with application to optimal control. SIAM J. Control Optim. **48**(3), 1771–1796 (2009)
5. Becker, R., Kapp, H., Rannacher, R.: Adaptive finite element methods for optimal control of partial differential equations: basic concept. SIAM J. Control Optim. **39**(1), 113–132 (electronic) (2000)
6. Camacho, F., Demlow, A.: L_2 and pointwise a posteriori error estimates for FEM for elliptic PDEs on surfaces. IMA J. Numer. Anal. **35**(3), 1199–1227 (2015)
7. Dari, E., Durán, R.G., Padra, C.: Maximum norm error estimators for three-dimensional elliptic problems. SIAM J. Numer. Anal. **37**(2), 683–700 (2000)
8. Dauge, M.: Neumann and mixed problems on curvilinear polyhedra. Integral Equations Operator Theory **15**(2), 227–261 (1992)
9. Demlow, A., Georgoulis, E.H.: Pointwise a posteriori error control for discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **50**(5), 2159–2181 (2012)
10. Demlow, A., Kopteva, N.: Maximum-norm a posteriori error estimates for singularly perturbed elliptic reaction-diffusion problems. Numer. Math. **133**(4), 707–742 (2016)
11. Demlow, A., Larsson, S.: Local pointwise a posteriori gradient error bounds for the Stokes equations. Math. Comp. **82**(282), 625–649 (2013)
12. Demlow, A., Leykekhman, D., Schatz, A.H., Wahlbin, L.B.: Best approximation property in the W^1_{∞} norm for finite element methods on graded meshes. Math. Comp. **81**(278), 743–764 (2012)
13. Eriksson, K.: An adaptive finite element method with efficient maximum norm error control for elliptic problems. Math. Models Methods Appl. Sci. **4**(3), 313–329 (1994)
14. Ern, A., Guermond, J.L.: Theory and practice of finite elements. Springer, New York (2004)
15. Frehse, J., Rannacher, R.: Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente pp. 92–114. Bonn. Math. Schrift., No. 89 (1976)
16. Grisvard, P.: Elliptic problems in nonsmooth domains, *Classics in Applied Mathematics*, vol. 69. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2011). Reprint of the 1985 original [MR0775683], With a foreword by Susanne C. Brenner
17. Guzmán, J., Leykekhman, D., Rossmann, J., Schatz, A.: Hölder estimates for Green's functions on convex polyhedral domains and their applications to finite element methods. Numer. Math. **112**(2), 221–243 (2009)
18. Hintermüller, M., Hoppe, R.: Goal-oriented adaptivity in control constrained optimal control of partial differential equations. SIAM J. Control Optim. **47**(4), 1721–1743 (2008)
19. Hintermüller, M., Hoppe, R., Iliash, Y., Kieweg, M.: An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. ESAIM: Control Optim. Calc. of Var. **14**, 540–560 (2008)
20. Jerison, D., Kenig, C.E.: The inhomogeneous Dirichlet problem in Lipschitz domains. J. Funct. Anal. **130**(1), 161–219 (1995)
21. Jerison, D.S., Kenig, C.E.: The Neumann problem on Lipschitz domains. Bull. Amer. Math. Soc. (N.S.) **4**(2), 203–207 (1981)
22. Kohls, K., Rösch, A., Siebert, K.: A posteriori error analysis of optimal control problems with control constraints. SIAM J. Control Optim. **52**(3), 1832–1861 (2014)
23. Leykekhman, D., Vexler, B.: Finite element pointwise results on convex polyhedral domains. SIAM J. Numer. Anal. **54**(2), 561–587 (2016)
24. Li, R., Liu, W., Yan, N.: A posteriori error estimates of recovery type for distributed convex optimal control problems. J. Sci. Comput. **33**(2), 155–182 (2007)
25. Lions, J.L.: Optimal control of systems governed by partial differential equations. Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170. Springer-Verlag, New York-Berlin (1971)

26. Liu, W., Yan, N.: A posteriori error estimates for distributed convex optimal control problems. Adv. in Comput. Math. **15**(1-4), 285–309 (2001)
27. Maz'ya, V., Rossmann, J.: Elliptic equations in polyhedral domains, *Mathematical Surveys and Monographs*, vol. 162. American Mathematical Society, Providence, RI (2010)
28. Meyer, C., Rademacher, A., Wollner, W.: Adaptive optimal control of the obstacle problem. SIAM J. Sci. Comput. **37**(2), 918–945 (2015)
29. Meyer, C., Rösch, A.: L^∞-estimates for approximated optimal control problems. SIAM J. Control Optim. **44**(5), 1636–1649 (2005)
30. Natterer, F.: über die punktweise Konvergenz finiter Elemente. Numer. Math. **25**(1), 67–77 (1975/76)
31. Nitsche, J.: Lineare Spline-Funktionen und die Methoden von Ritz für elliptische Randwertprobleme. Arch. Rational Mech. Anal. **36**, 348–355 (1970)
32. Nitsche, J.: L_∞-convergence of finite element approximation. In: Journées "Éléments Finis" (Rennes, 1975), p. 18. Univ. Rennes, Rennes (1975)
33. Nochetto, R.: Pointwise a posteriori error estimates for elliptic problems on highly graded meshes. Math. Comp. **64**(209), 1–22 (1995)
34. Nochetto, R.H., Schmidt, A., Siebert, K.G., Veeser, A.: Pointwise a posteriori error estimates for monotone semi-linear equations. Numer. Math. **104**(4), 515–538 (2006)
35. Nochetto, R.H., Siebert, K.G., Veeser, A.: Pointwise a posteriori error control for elliptic obstacle problems. Numer. Math. **95**(1), 163–195 (2003)
36. Nochetto, R.H., Siebert, K.G., Veeser, A.: Theory of adaptive finite element methods: an introduction. In: Multiscale, nonlinear and adaptive approximation, pp. 409–542. Springer, Berlin (2009)
37. Nochetto, R.H., Veeser, A.: Primer of adaptive finite element methods. In: Multiscale and adaptivity: modeling, numerics and applications, *Lecture Notes in Math.*, vol. 2040, pp. 125–225. Springer, Heidelberg (2012)
38. Rannacher, R., Scott, R.: Some optimal error estimates for piecewise linear finite element approximations. Math. Comp. **38**(158), 437–445 (1982)
39. Rösch, A., Siebert, K.G., Steinig, S.: Reliable a posteriori error estimation for state-constrained optimal control. Computational Optimization and Applications **68**(1), 121–162 (2017)
40. Savaré, G.: Regularity results for elliptic equations in Lipschitz domains. J. Funct. Anal. **152**(1), 176–201 (1998)
41. Schatz, A., Wahlbin, L.: Interior maximum norm estimates for finite element methods. Math. Comp. **31**(138), 414–442 (1977)
42. Schatz, A., Wahlbin, L.: Maximum norm estimates in the finite element method on plane polygonal domains. I. Math. Comp. **32**(141), 73–109 (1978)
43. Schatz, A., Wahlbin, L.: Maximum norm estimates in the finite element method on plane polygonal domains. II. Refinements. Math. Comp. **33**(146), 465–492 (1979)
44. Schatz, A.H., Wahlbin, L.B.: On the quasi-optimality in L_∞ of the $\dot{H}^1$-projection into finite element spaces. Math. Comp. **38**(157), 1–22 (1982)
45. Scott, R.: Optimal L^∞ estimates for the finite element method on irregular meshes. Math. Comp. **30**(136), 681–697 (1976)
46. Tröltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods, and Applications. Graduate Studies in Mathematics. American Mathematical Society (2010)
47. Verfürth, R.: A posteriori error estimation techniques for finite element methods. Oxford University Press, Oxford (2013)