ESAIM: COCV 14 (2008) 540–560
DOI: 10.1051/cocv:2007057

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

AN *A POSTERIORI* ERROR ANALYSIS OF ADAPTIVE FINITE ELEMENT METHODS FOR DISTRIBUTED ELLIPTIC CONTROL PROBLEMS WITH CONTROL CONSTRAINTS

Michael Hintermüller[1], Ronald H.W. Hoppe[2,3], Yuri Iliash[3] and Michael Kieweg[3]

Abstract. We present an *a posteriori* error analysis of adaptive finite element approximations of distributed control problems for second order elliptic boundary value problems under bound constraints on the control. The error analysis is based on a residual-type *a posteriori* error estimator that consists of edge and element residuals. Since we do not assume any regularity of the data of the problem, the error analysis further invokes data oscillations. We prove reliability and efficiency of the error estimator and provide a bulk criterion for mesh refinement that also takes into account data oscillations and is realized by a greedy algorithm. A detailed documentation of numerical results for selected test problems illustrates the convergence of the adaptive finite element method.

Mathematics Subject Classification. 65N30, 65N50, 49K20, 65K10

Received October 6, 2005. Revised March 8, 2006.
Published online November 21, 2007.

1. Introduction

Adaptive finite element methods have been widely and successfully used for the efficient numerical solution of boundary and initial-boundary value problems for partial differential equations and systems thereof (*cf.*, *e.g.*, the monographs [1,3,4,14,26,27] and the references therein).

Several error concepts have been developed over the past three decades including residual-type estimators [2,3,27] that rely on the appropriate evaluation of the residual in a dual norm, hierarchical type estimators [5,18,19] where the error equation is solved locally using higher order elements, error estimators that are based on local averaging [9,28], the so-called goal oriented dual weighted approach [4,14] where information about the error is extracted from the solution of the dual problem, and functional type error majorants [26] that provide guaranteed sharp upper bounds for the error.

As far as the *a posteriori* error analysis of adaptive finite element schemes for optimal control problems is concerned, there is not much work available. The unconstrained case has been addressed in [4,6], whereas

Keywords and phrases. A posteriori error analysis, distributed optimal control problems, control constraints, adaptive finite element methods, residual-type *a posteriori* error estimators, data oscillations.

[1] Institute of Mathematics, University of Graz, 8010 Graz, Austria.

[2] Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA.

[3] Institute of Mathematics, University of Augsburg, 86159 Augsburg, Germany; `Ronald.Hoppe@math.uni-augsburg.de`

residual-type *a posteriori* error estimators in the control constrained case have been derived and analyzed in [20,23,24]. In contrast to the approach used in [20,23,24], the error analysis in this paper pertains to the error in the state, the adjoint state, the control, and the adjoint control and incorporates oscillations in terms of the data of the problem. The data oscillations may significantly contribute to the error and thus have to be considered in the adaptive refinement process. The paper is organized as follows:

In Section 2, as a model problem we consider a distributed optimal control problem for a two-dimensional, second order elliptic PDE with a quadratic objective functional and unilateral constraints on the control variable. The optimality conditions are stated in terms of the state, the adjoint state, the control, and the Lagrangian multiplier for the control which will be referred to as the adjoint control.

In Section 3, the control problem is discretized with respect to a shape regular simplicial triangulation of the computational domain using continuous, piecewise linear finite elements for the state and the adjoint state and elementwise constant approximations of the control and the adjoint control.

The residual-type *a posteriori* error estimator for the global discretization errors in the state, the adjoint state, the control, and the adjoint control consists of edge and element residuals. In contrast to [20], we include the error in the adjoint control. Moreover, we do not assume any regularity of the data. Consequently, the *a posteriori* error analysis also has to take into account data oscillations. Both the *a posteriori* error estimator and the data oscillations are presented in Section 4.

In Section 5, we prove reliability of the error estimator, *i.e.*, up to data oscillations, it provides an upper bound for the global discretization errors. Section 6 is devoted to the efficiency of the estimator. Here, it is shown that, modulo data oscillations, the error estimator also gives rise to a lower bound for the discretization errors.

In Section 7, we address the issue of adaptive mesh refinement on the basis of the local components of the error estimator and the data oscillations. This is done by means of a bulk criterion where edges and elements of the triangulation are selected for refinement in such a way that the sum of the associated error terms/data oscillations exceeds the total sum by a certain margin. The bulk criterion is realized by a greedy algorithm.

Finally, Section 8 contains a detailed documentation of numerical results for selected test examples in terms of the convergence history of the adaptive finite element method including visualizations of the adaptively generated simplicial triangulations.

2. The distributed elliptic control problem

We consider the following optimal control problem for a linear second order elliptic boundary value problem with constrained distributed controls

$$\text{minimize } J(y,u) \;:=\; \frac{1}{2}\|y-y^d\|_{0,\Omega}^2 \;+\; \frac{\alpha}{2}\|u-u^d\|_{0,\Omega}^2 \tag{2.1a}$$

$$\text{over } (y,u) \in H_0^1(\Omega)\times L^2(\Omega),$$

$$\text{subject to } -\Delta y \;=\; f \;+\; u, \tag{2.1b}$$

$$u \in K \;:=\; \{v \in L^2(\Omega) \mid v \le \psi \;\; a.e. \;\text{ in } \Omega\}. \tag{2.1c}$$

Here, $\Omega \subset \mathbb{R}^2$ is a bounded, polygonal domain with boundary $\Gamma := \partial\Omega$. Moreover, we suppose that

$$u^d, y^d \in L^2(\Omega),\;\; f \in L^2(\Omega),\;\; \psi \in L^\infty(\Omega),\;\; \alpha \in \mathbb{R}_+. \tag{2.2}$$

It is well-known that under the assumption (2.2) the distributed optimal control problem (2.1a)–(2.1c) admits a unique solution $(y,u) \in H_0^1(\Omega)\times L^2(\Omega)$ (*cf.*, *e.g.*, [15,21–23]) which is characterized by the existence of a co-state (adjoint state) $p \in H_0^1(\Omega)$ and a Lagrange multiplier for the inequality constraint (adjoint control)

$\sigma \in L^2(\Omega)$ such that

$$a(y,v) = (f+u,v)_{0,\Omega}, \quad v \in H_0^1(\Omega), \tag{2.3a}$$
$$a(p,v) = -(y-y^d,v)_{0,\Omega}, \quad v \in H_0^1(\Omega), \tag{2.3b}$$
$$u = u^d + \frac{1}{\alpha}(p-\sigma), \tag{2.3c}$$
$$\sigma \in \partial I_K(u). \tag{2.3d}$$

Here, $a(\cdot,\cdot)$ stands for the bilinear form

$$a(w,z) := \int\limits_\Omega \nabla w \cdot \nabla z \, \mathrm{d}x, \quad w,z \in H_0^1(\Omega),$$

and $\partial I_K : L^2(\Omega) \to 2^{L^2(\Omega)}$ denotes the subdifferential of the indicator function I_K of the constraint set K (*cf.*, *e.g.*, [17]).

We note that (2.3d) can be equivalently written as the variational inequality

$$(\sigma, u-v)_{0,\Omega} \geq 0, \quad v \in K, \tag{2.4}$$

and the complementarity problem

$$\begin{aligned} \sigma \in L_+^2(\Omega),\ \psi - u \in L_+^2(\Omega), \\ (\sigma, \psi - u)_{0,\Omega} = 0, \end{aligned} \tag{2.5}$$

where $(\cdot,\cdot)_{0,\Omega}$ stands for the L^2-inner product and $L_+^2(\Omega)$ refers to the nonnegative cone in $L^2(\Omega)$.

We define the active control set $\mathcal{A}(u)$ as the maximal open set $A \subset \Omega$ such that $u(x) = \psi(x)$ f.a.a. $x \in A$ and the inactive control set $\mathcal{I}(u)$ according to $\mathcal{I}(u) := \bigcup_{\varepsilon>0} B_\varepsilon$, where B_ε is the maximal open set $B \subset \Omega$ such that $u(x) \leq \psi(x) - \varepsilon$ for almost all $x \in B$. Then, the complementarity conditions (2.5) can be equivalently stated as:

$$\sigma(x) \geq 0 \quad \text{f.a.a. } x \in \Omega, \tag{2.6a}$$
$$\sigma(x) = 0 \quad \text{f.a.a. } x \in \mathcal{I}(u), \tag{2.6b}$$
$$\sigma(x) = \alpha(u^d(x) - \psi(x)) + p(x) \quad \text{f.a.a. } x \in \mathcal{A}(u). \tag{2.6c}$$

3. Finite element approximation

We assume that $\{\mathcal{T}_h(\Omega)\}$ is a family of shape-regular simplicial triangulations of Ω. We refer to $\mathcal{N}_h(D)$ and $\mathcal{E}_h(D),\ D \subseteq \overline{\Omega}$, as the sets of vertices and edges of $\mathcal{T}_h(\Omega)$ in $D \subseteq \overline{\Omega}$. We denote by h_T and $|T|$ the diameter and area of an element $T \in \mathcal{T}_h(\Omega)$ and by h_E the length of an edge $E \in \mathcal{E}_h(D)$.

The distributed optimal control problem (2.1a)–(2.1c) is discretized by continuous piecewise linear finite elements with respect to the triangulation $\mathcal{T}_h(\Omega)$. In particular, we refer to

$$V_h := \{v_h \in C_0(\Omega)\ |v_h|_T \in P_1(T),\ T \in \mathcal{T}_h(\Omega)\},$$

as the finite element space spanned by the canonical nodal basis functions $\varphi_h^a,\ a \in \mathcal{N}_h(\Omega)$, associated with the nodal points in $\overline{\Omega}$. Moreover, we denote by

$$W_h := \{w_h \in L^2(\Omega)\ |w_h|_T \in P_0(T),\ T \in \mathcal{T}_h(\Omega)\}$$

the linear space of elementwise constant functions on Ω. We refer to $y_h \in V_h$ and $u_h \in W_h$ as finite element approximations of the state y and the control u, respectively. We approximate the upper obstacle ψ by $\psi_h \in W_h$ with $\psi_h|_T := |T|^{-1} \int_T \psi \, \mathrm{d}x,\ T \in \mathcal{T}_h(\Omega)$.

The finite element approximation of the distributed optimal control problem (2.1a)–(2.1c) reads as follows:

$$\text{minimize } J_h(y_h, u_h) := \frac{1}{2}\|y_h - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2}\|u_h - u^d\|_{0,\Omega}^2, \tag{3.1a}$$

$$\text{over } (y_h, u_h) \in V_h \times W_h, \tag{3.1b}$$

$$\text{subject to } a(y_h, v_h) = (f + u_h, v_h)_{0,\Omega},\ v_h \in V_h, \tag{3.1c}$$

$$u_h \in K_h := \{w_h \in W_h \ |w_h|_T \le \psi_h|_T,\ T \in \mathcal{T}_h(\Omega)\}. \tag{3.1d}$$

As in the continuous regime, the necessary and sufficient optimality conditions for (3.1a)–(3.1d) involve the existence of an adjoint state $p_h \in V_h$ and an adjoint control $\sigma_h \in W_h$ such that

$$a(y_h, v_h) \;=\; (f + u_h, v_h)_{0,\Omega},\ v_h \in V_h, \tag{3.2a}$$

$$a(p_h, v_h) \;=\; -(y_h - y^d, v_h)_{0,\Omega},\ v_h \in V_h, \tag{3.2b}$$

$$u_h \;=\; u_h^d \;+\; \frac{1}{\alpha}(M_h p_h - \sigma_h), \tag{3.2c}$$

$$\sigma_h \in \partial I_{K_h}(u_h), \tag{3.2d}$$

where $u_h^d \in W_h$ with $u_h^d|_T := |T|^{-1} \int_T u^d \, \mathrm{d}x,\ T \in \mathcal{T}_h(\Omega)$, and $M_h : V_h \to W_h$ is the operator given by

$$(M_h v_h)_T \;:=\; |T|^{-1} \int\limits_T v_h(x)\, \mathrm{d}x,\ T \in \mathcal{T}_h(\Omega). \tag{3.3}$$

Again, (3.2d) can be stated as the complementarity problem

$$\sigma_h \;\ge\; 0, \quad \psi_h - u_h \;\ge\; 0, \qquad (\sigma_h, \psi_h - u_h)_{0,\Omega} \;=\; 0. \tag{3.4}$$

We define $\mathcal{A}(u_h)$ and $\mathcal{I}(u_h)$ as the discrete active and inactive control sets according to

$$\mathcal{A}(u_h) \;:=\; \bigcup \{T \in \mathcal{T}_h(\Omega) \ |u_h|_T = \psi_h|_T\}, \tag{3.5a}$$

$$\mathcal{I}(u_h) \;:=\; \bigcup \{T \in \mathcal{T}_h(\Omega) \ |u_h|_T < \psi_h|_T\}. \tag{3.5b}$$

The complementarity conditions (3.4) readily imply

$$\sigma_h|_T \;\ge\; 0, \quad T \in \mathcal{T}_h(\Omega), \tag{3.6a}$$

$$\sigma_h|_T \;=\; 0, \quad T \in \mathcal{I}(u_h), \tag{3.6b}$$

$$\sigma_h|_T \;=\; \alpha(u_h^d - \psi_h)|_T \;+\; (M_h p_h)|_T, \quad T \in \mathcal{A}(u_h). \tag{3.6c}$$

We note that the discrete state and co-state $y_h, p_h \in V_h$ may also be considered as finite element approximations of the coupled elliptic system: given $u_h \in W_h$, find $y(u_h), p(u_h) \in H_0^1(\Omega)$ such that

$$a(y(u_h), v) \;=\; (f + u_h, v)_{0,\Omega},\ v \in H_0^1(\Omega), \tag{3.7a}$$

$$a(p(u_h), v) \;=\; -(y(u_h) - y^d, v)_{0,\Omega},\ v \in H_0^1(\Omega). \tag{3.7b}$$

Obviously, we have

$$|y(u_h) - y|_{1,\Omega} \ \leq \ c(\Omega) \ \|u - u_h\|_{0,\Omega}, \tag{3.8a}$$

$$|p(u_h) - p|_{1,\Omega} \ \leq \ c(\Omega) \ \|y - y(u_h)\|_{0,\Omega}, \tag{3.8b}$$

where $c(\Omega) > 0$ is the constant in the Poincaré-Friedrichs inequality

$$\|v\|_{0,\Omega} \ \leq \ c(\Omega) \ |v|_{1,\Omega}, \ \ v \in H_0^1(\Omega). \tag{3.9}$$

Moreover, choosing $v = p(u_h) - p$ in (3.7a) and $v = y(u_h) - y$ in (3.7b), we find

$$(p - p(u_h), u - u_h)_{0,\Omega} \ = \ - \ \|y - y(u_h)\|^2_{0,\Omega} \leq 0. \tag{3.10}$$

4. The residual type error estimator

The residual type error estimator consists of easily computable element and edge residuals with respect to the finite element approximations $y_h \in V_h$ and $p_h \in V_h$ of the state $y \in H_0^1(\Omega)$ and the co-state $p \in H_0^1(\Omega)$ as well as of data oscillations.

In particular, we define

$$\eta_y \ := \ \left(\sum_{T \in \mathcal{T}_h(\Omega)} \eta^2_{y,T} \ + \ \sum_{E \in \mathcal{E}_h(\Omega)} \eta^2_{y,E} \right)^{1/2}, \tag{4.1a}$$

$$\eta_p \ := \ \left(\sum_{T \in \mathcal{T}_h(\Omega)} \sum_{i=1}^{2} (\eta^{(i)}_{p,T})^2 \ + \ \sum_{E \in \mathcal{E}_h(\Omega)} \eta^2_{p,E} \right)^{1/2}. \tag{4.1b}$$

Here, the element residuals $\eta_{y,T}, \eta^{(i)}_{p,T}$, $1 \leq i \leq 2$, and the edge residuals $\eta_{y,E}, \eta_{p,E}$ are given by

$$\eta_{y,T} \ := \ h_T \|f + u_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \tag{4.2a}$$

$$\eta^{(1)}_{p,T} \ := \ h_T \|y^d - y_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \tag{4.2b}$$

$$\eta^{(2)}_{p,T} \ := \ \|M_h p_h - p_h\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \tag{4.2c}$$

$$\eta_{y,E} \ := h_E^{1/2} \ \|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \tag{4.2d}$$

$$\eta_{p,E} \ := h_E^{1/2} \ \|\boldsymbol{\nu}_E \cdot [\nabla p_h]\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \tag{4.2e}$$

where $E = T_1 \cap T_2, \ T_\nu \in \mathcal{T}_h(\Omega), \ 1 \leq \nu \leq 2$, and $\boldsymbol{\nu}_E$ is the exterior unit normal vector on E directed towards T_2, whereas $[\nabla y_h]$ and $[\nabla p_h]$ denote the jumps of $\nabla y_h, \nabla p_h$ across E.

The residual type error estimator η for the finite element approximation of the distributed control problem (2.1a)–(2.1c) is then given by

$$\eta \ := \ \left(\eta_y^2 \ + \ \eta_p^2 \right)^{1/2}. \tag{4.3}$$

Moreover, we define the low order data oscillations

$$\mu_h(u^d) := \Big(\sum_{T\in\mathcal{T}_h(\Omega)} \mu_T(u^d)^2\Big)^{1/2}, \qquad (4.4a)$$
$$\mu_T(u^d) := \|u^d - u_h^d\|_{0,T},$$
$$\mu_h(\psi) := \Big(\sum_{T\in\mathcal{T}_h(\Omega)} \mu_T(\psi)^2\Big)^{1/2}, \qquad (4.4b)$$
$$\mu_T(\psi) := \|\psi - \psi_h\|_{0,T},$$

as well as the data oscillations

$$\mathrm{osc}_h(y^d) := \Big(\sum_{T\in\mathcal{T}_h(\Omega)} \mathrm{osc}_T(y^d)^2\Big)^{1/2}, \qquad (4.5a)$$
$$\mathrm{osc}_T(y^d) := h_T\,\|y^d - y_h^d\|_{0,T},$$
$$\mathrm{osc}_h(f) := \Big(\sum_{T\in\mathcal{T}_h(\Omega)} \mathrm{osc}_T(f)^2\Big)^{1/2}, \qquad (4.5b)$$
$$\mathrm{osc}_T(f) := h_T\|f - f_h\|_{0,T},$$

where $y_h^d \in W_h$ and $f_h \in W_h$ with $y_h^d|_T := |T|^{-1}\int_T y^d\,\mathrm{d}x$, $f_h|_T := |T|^{-1}\int_T f\,\mathrm{d}x$, $T\in\mathcal{T}_h(\Omega)$.

Compared to the element residuals $\eta_{y,T},\eta_{p,T}$ and the edge residuals $\eta_{y,E},\eta_{p,E}$, the data oscillations $\mathrm{osc}_h(y^d)$, $\mathrm{osc}_h(f)$ are of the same order for non smooth y^d, f and of higher order for smooth y^d, f, *e.g.*, $y^d, f\in H^1(\Omega)$.

Remark 4.1. The element residuals $\eta_{y,T}$ and $\eta_{p,T}^{(1)}$ in (4.2a), (4.2b) may be replaced by $\hat\eta_{y,T} := h_T\|f_h + u_h\|_{0,T}$ and $\hat\eta_{p,T}^{(1)} := h_T\|y_h^d - y_h\|_{0,T}$. Obviously, $\eta_{y,T},\eta_{p,T}^{(1)}$ and $\hat\eta_{y,T},\hat\eta_{p,T}^{(1)}$ are equivalent up to the data oscillations $\mathrm{osc}_T(f)$ and $\mathrm{osc}_T(y^d)$, respectively. The following results remain valid up to these (additional) data oscillations.

5. Reliability of the error estimator

Theorem 5.1. *Let (y,p,u,σ) and (y_h,p_h,u_h,σ_h) be the solutions of* (2.3a)–(2.3d) *and* (3.2a)–(3.2d)*, and let η and $\mu_h(u^d),\mu_h(\psi)$ be the residual error estimator and the data oscillations as given by* (4.3) *and* (4.4a), (4.4b)*, respectively. Then, there exist positive constants Λ and C, depending on α,Ω and the shape regularity of $\{\mathcal{T}_h(\Omega)\}$, such that*

$$|y-y_h|_{1,\Omega} + |p-p_h|_{1,\Omega} + \|u-u_h\|_{0,\Omega} + \|\sigma-\sigma_h\|_{0,\Omega} \le \Lambda\eta + C\Big(\mu_h(u^d) + \mu_h(\psi)\Big). \qquad (5.1)$$

The idea of the proof of Theorem 5.1 is to show that the discretization errors, making up the left-hand side in (5.1), can be bounded by the discretization errors in the finite element approximations of $y(u_h)$ and $p(u_h)$ by y_h and p_h and the data oscillation $\mu_h(u^d)$ and $\mu_h(\psi)$. An upper bound for the latter discretization errors can be obtained as in the case of the finite element approximations of standard second order elliptic boundary value problems. As a first step in this direction, we prove the following result.

Lemma 5.2. *Let (y,p,u,σ) and (y_h,p_h,u_h,σ_h) be the solutions of* (2.3a)–(2.3d) *and* (3.2a)–(3.2d)*, respectively, and let $\mu_h(u^d)$ be the data oscillation according to* (4.4a)*. Moreover, let $y(u_h)$ and $p(u_h)$ be the intermediate state* *and intermediate adjoint state as given by* (3.7a), (3.7b). *Then, there exists a positive constant C depending only on α and Ω such that*

$$|y-y_h|_{1,\Omega}+|p-p_h|_{1,\Omega}+\|u-u_h\|_{0,\Omega}+\|\sigma-\sigma_h\|_{0,\Omega}\le C\Bigg(|y_h-y(u_h)|_{1,\Omega}\;+\;|p_h-p(u_h)|_{1,\Omega}\\ +\Big(\sum_{T\in\mathcal{T}_h(\Omega)}(\eta_{p,T}^{(2)})^{1/2}\Big)^2\;+\;\mu_h(\psi)\;+\;\mu_h(u^d)\Bigg). \tag{5.2}$$

Proof. Using (3.8a), (3.8b) and (3.9), we find

$$|y-y_h|_{1,\Omega}\le|y_h-y(u_h)|_{1,\Omega}+c(\Omega)\|u-u_h\|_{0,\Omega}, \tag{5.3}$$

$$\begin{aligned}|p-p_h|_{1,\Omega}&\le|p_h-p(u_h)|_{1,\Omega}+c(\Omega)\|y-y_h\|_{0,\Omega}\\ &\le|p_h-p(u_h)|_{1,\Omega}+c(\Omega)^2|y_h-y(u_h)|_{1,\Omega}+c(\Omega)^3\|u-u_h\|_{0,\Omega},\end{aligned} \tag{5.4}$$

$$\begin{aligned}\|\sigma-\sigma_h\|_{0,\Omega}^2\le&\;\;2\alpha^2\Big(\|u-u_h\|_{0,\Omega}+\mu_h(u^d)\Big)^2+2\|p-M_hp_h\|_{0,\Omega}^2\\ &\le4\alpha^2\Big(\|u-u_h\|_{0,\Omega}^2+\mu_h^2(u^d)\Big)+4c(\Omega)^2|p-p_h|_{1,\Omega}^2\\ &\quad+4\|p_h-M_hp_h\|_{0,\Omega}^2\le4(\alpha^2+3c(\Omega)^8)\|u-u_h\|_{0,\Omega}^2\\ &\quad+12c(\Omega)^2|p_h-p(u_h)|_{1,\Omega}^2+12c(\Omega)^6|y_h-y(u_h)|_{1,\Omega}^2\\ &\quad+4\|p_h-M_hp_h\|_{0,\Omega}^2+4\alpha^2\mu_h^2(u^d).\end{aligned} \tag{5.5}$$

Moreover, in view of (2.3c) and (3.2c), using Young's inequality we get

$$\begin{aligned}\alpha\|u-u_h\|_{0,\Omega}^2&=(\sigma_h-\sigma,u-u_h)_{0,\Omega}+(p-p_h,u-u_h)_{0,\Omega}\\ &\quad+(p_h-M_hp_h,u-u_h)_{0,\Omega}+\alpha\left(u^d-u_h^d,u-u_h\right)_{0,\Omega}\\ &\le\;(\sigma_h-\sigma,u-u_h)_{0,\Omega}+(p-p_h,u-u_h)_{0,\Omega}\\ &\quad+\frac{\alpha}{4}\|u-u_h\|_{0,\Omega}^2+\frac{2}{\alpha}\|p_h-M_hp_h\|_{0,\Omega}^2+2\alpha\mu_h(u^d).\end{aligned} \tag{5.6}$$

Observing (2.5) and (3.4), for the first term on the right-hand side in (5.6) it follows that

$$\begin{aligned}(\sigma_h-\sigma,u-u_h)_{0,\Omega}=&\;\underbrace{(\sigma_h,u-\psi)_{0,\Omega}}_{\le0}+(\sigma_h-\sigma,\psi-\psi_h)_{0,\Omega}+\underbrace{(\sigma_h,\psi_h-u_h)_{0,\Omega}}_{=0}\\ &-\underbrace{(\sigma,u-\psi)_{0,\Omega}}_{=0}-\underbrace{(\sigma,\psi_h-u_h)_{0,\Omega}}_{\ge0}\;\le\;|(\sigma_h-\sigma,\psi-\psi_h)_{0,\Omega}|.\end{aligned}$$

An application of Young's inequality yields

$$(\sigma_h-\sigma,u-u_h)_{0,\Omega}\le\frac{\alpha}{16(\alpha^2+3c(\Omega)^8)}\|\sigma-\sigma_h\|_{0,\Omega}^2+4\frac{\alpha^2+3c(\Omega)^8}{\alpha}\mu_h^2(\psi). \tag{5.7}$$

On the other hand, in view of (3.10), for the second term on the right-hand side in (5.6) we obtain

$$(p-p_h,u-u_h)_{0,\Omega}\;\le\;(p(u_h)-p_h,u-u_h)_{0,\Omega}.$$

Using Young's inequality once more, the right-hand side can be further estimated according to

$$(p(u_h)-p_h,u-u_h)_{0,\Omega} \;\le\; \frac{\alpha}{4}\|u-u_h\|_{0,\Omega}^2+\frac{c(\Omega)^2}{\alpha}|p(u_h)-p_h|_{1,\Omega}^2. \tag{5.8}$$

Using (5.7) and (5.8) in (5.6), we end up with

$$\begin{aligned}\|u-u_h\|_{0,\Omega}^2 \le{}& \frac{1}{8(\alpha^2+3c(\Omega)^8)}\;\|\sigma-\sigma_h\|_{0,\Omega}^2\\ &+2\left(\frac{c(\Omega)}{\alpha}\right)^2\;|p_h-p(u_h)|_{1,\Omega}^2+\frac{4}{\alpha^2}\|p_h-M_hp_h\|_{0,\Omega}^2\\ &+4\mu_h^2(u^d)+8\frac{\alpha^2+3c(\Omega)^8}{\alpha^2}\mu_h^2(\psi).\end{aligned} \tag{5.9}$$

Hence, taking advantage of (5.9) in (5.5), we obtain

$$\begin{aligned}\|\sigma-\sigma_h\|_{0,\Omega}^2 \le{}& 8c(\Omega)^2\left(3+\frac{2(\alpha^2+3c(\Omega)^8)}{\alpha^2}\right)|p_h-p(u_h)|_{1,\Omega}^2\\ &+24c(\Omega)^6|y_h-y(u_h)|_{1,\Omega}^2+8\left(1+4\frac{\alpha^2+3c(\Omega)^8}{\alpha^2}\right)\|p_h-M_hp_h\|_{0,\Omega}^2\\ &+8\big(5\alpha^2+12c(\Omega)^8\big)\mu_h^2(u^d)+64\frac{(\alpha^2+3c(\Omega)^8)^2}{\alpha^2}\mu_h^2(\psi).\end{aligned} \tag{5.10}$$

On the other hand, using (5.10) in (5.9) readily gives

$$\begin{aligned}\|u-u_h\|_{0,\Omega}^2 \le{}& \frac{c(\Omega)^2}{\alpha^2}\frac{5\alpha^2+6c(\Omega)^8}{\alpha^2+3c(\Omega)^8}|p_h-p(u_h)|_{1,\Omega}^2\\ &+\frac{3c(\Omega)^6}{\alpha^2+3c(\Omega)^8}|y_h-y(u_h)|_{1,\Omega}^2+\frac{9\alpha^2+24c(\Omega)^8}{\alpha^2+3c(\Omega)^8}\mu_h^2(u^d)\\ &+\left(\frac{4}{\alpha^2}+\frac{5\alpha^2+12c(\Omega)^8}{\alpha^2(\alpha^2+3c(\Omega)^8)}\right)\|p_h-M_hp_h\|_{0,\Omega}^2\\ &+16\frac{\alpha^2+3c(\Omega)^8}{\alpha^2}\mu_h^2(\psi).\end{aligned} \tag{5.11}$$

Combining (5.3), (5.4), (5.10) and (5.11), gives the assertion. □

Lemma 5.3. *Let (y_h,p_h,u_h,σ_h) be the solution of (3.2a)–(3.2d) and let $y(u_h),p(u_h)$ be the solutions of (3.7a), (3.7b), respectively. Further, let η_y and $\eta_{p,T}^{(1)},\eta_{p,E}$ be the parts of the residual error estimator η as given by (4.1a) and (4.2b), (4.2e). Then, there exist positive constants C_ν, $4\le\nu\le5$, depending only on the shape regularity of $\{\mathcal{T}_h(\Omega)\}$, such that*

$$|y(u_h)-y_h|_{1,\Omega}^2 \;\le\; C_4\;\eta_y^2, \tag{5.12a}$$

$$|p(u_h)-p_h|_{1,\Omega}^2 \;\le\; C_5\left(\eta_y^2+\sum_{T\in\mathcal{T}_h(\Omega)}(\eta_{p,T}^{(1)})^2+\sum_{E\in\mathcal{E}_h(\Omega}\eta_{p,E}^2\right). \tag{5.12b}$$

Proof. Using standard techniques based on Clément's interpolation operator (*cf.*, *e.g.*, [27]), for the discretization error $|y(u_h) - y_h|_{1,\Omega}$ we obtain

$$|y(u_h) - y_h|^2_{1,\Omega} \le C\Big(\sum_{T\in\mathcal{T}_h(\Omega)} \underbrace{h_T^2\|f + u_h\|^2_{0,T}}_{=\,\eta^2_{y,T}} + \sum_{E\in\mathcal{E}_h(\Omega)} \underbrace{h_E\|\boldsymbol{\nu}_E\cdot[\nabla y_h]\|^2_{0,E}}_{=\,\eta^2_{y,E}}\Big),$$

which is (5.12a).

Applying the same techniques to the discretization error $|p(u_h) - p_h|_{1,\Omega}$, we obtain

$$|p(u_h) - p_h|^2_{1,\Omega} \le C\Big(\sum_{T\in\mathcal{T}_h(\Omega)} h_T^2\|y^d - y(u_h)\|^2_{0,T} + \sum_{E\in\mathcal{E}_h(\Omega)} \underbrace{h_E\|\boldsymbol{\nu}_E\cdot[\nabla p_h]\|^2_{0,E}}_{=\,\eta^2_{p,E}}\Big). \tag{5.13}$$

For the first term on the right-hand side in (5.13), taking advantage of (5.12a) it follows that

$$\begin{aligned}\sum_{T\in\mathcal{T}_h(\Omega)} h_T^2\|y^d - y(u_h)\|^2_{0,T} &\le 2\Big(\sum_{T\in\mathcal{T}_h(\Omega)} \underbrace{h_T^2\|y^d - y_h\|^2_{0,T}}_{=\,(\eta^{(1)}_{p,T})^2} + \sum_{T\in\mathcal{T}_h(\Omega)} h_T^2\|y(u_h) - y_h\|^2_{0,T}\Big) \\ &\le 2\sum_{T\in\mathcal{T}_h(\Omega)} (\eta^{(1)}_{T,p})^2 + 2h^2c(\Omega)^2|y(u_h) - y_h|^2_{1,\Omega} \\ &\le 2\sum_{T\in\mathcal{T}_h(\Omega)} (\eta^{(1)}_{p,T})^2 + 2h^2c(\Omega)^2C^2\eta_y^2.\end{aligned} \tag{5.14}$$

Combining (5.13) and (5.14) results in (5.12b). □

6. Local efficiency of the error estimator

Theorem 6.1. *Let* (y,p,u,σ) *and* (y_h,p_h,u_h,σ_h) *be the solutions of* (2.3a)–(2.3d) *and* (3.2a)–(3.2d) *and let* $\eta, \mu_h(u^d), \mu_h(\psi)$ *and* $\mathrm{osc}_h(y^d), \mathrm{osc}_h(f)$ *be given by* (4.3), (4.4a), (4.4b) *and* (4.5a), (4.5b), *respectively. Then, there exist positive constants* λ *and* c *depending only on* Ω *and the shape regularity of* $\{\mathcal{T}_h(\Omega)\}$ *such that*

$$|y - y_h|^2_{1,\Omega} + |p - p_h|^2_{1,\Omega} + \|u - u_h\|^2_{0,\Omega} + \|\sigma - \sigma_h\|^2_{0,\Omega} \ \ge \lambda\eta^2 - c\Big(\mu_h^2(u^d) + \mathrm{osc}_h^2(y^d) + \mathrm{osc}_h^2(f)\Big). \tag{6.1}$$

The proof of Theorem 6.1 will be given by a series of lemmas.

We denote by $\lambda_i^T, 1 \le i \le 3$, the barycentric coordinates of $T \in \mathcal{T}_h(\Omega)$ and refer to $\vartheta_T := 27\prod_{i=1}^3 \lambda_i^T$ as the associated element bubble function. Likewise, $\lambda_i^E,\ 1 \le i \le 2$, stand for the barycentric coordinates of $E \in \mathcal{E}_h(\Omega)$ and $\vartheta_E := 4\prod_{i=1}^2 \lambda_i^E$ denotes the associated edge bubble function. We recall from [27] that there exist constants $c_i,\ 1 \le i \le 3$, depending only on the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$ such that for $\zeta_T \in P_k(T),\ k \in \mathbb{N}_0$, and $\zeta_E \in P_k(E),\ k \in \mathbb{N}_0$, there holds

$$\|\zeta_T\|^2_{0,T} \ \le\ c_1(\zeta_T, \zeta_T\vartheta_T)_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \tag{6.2a}$$

$$\|\zeta_T\vartheta_T\|_{0,T} \ \le\ \|\zeta_T\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \tag{6.2b}$$

$$|\zeta_T\vartheta_T|_{1,T} \ \le\ c_2h_T^{-1}\|\zeta_T\|_{0,T}, \quad T \in \mathcal{T}_h(\Omega), \tag{6.2c}$$

$$\|\zeta_E\|^2_{0,E} \ \le\ c_3(\zeta_E, \zeta_E\vartheta_E)_{0,E}, \quad E \in \mathcal{E}_h(\Omega), \tag{6.2d}$$

$$\|\zeta_E\vartheta_E\|_{0,E} \ \le\ \|\zeta_E\|_{0,E}, \quad E \in \mathcal{E}_h(\Omega). \tag{6.2e}$$

For $E \in \mathcal{E}_h(T)$ and $\zeta_E \in P_k(E),\ k \in \mathbb{N}_0$, we further refer to $\tilde{\zeta}_E$ as the extension of ζ_E to the patch

$$\omega_E := T_1 \cup T_2,\ E = T_1 \cap T_2,\ T_\nu \in \mathcal{T}_h(\Omega),\ 1 \le \nu \le 2, \tag{6.3}$$

in the sense that for fixed $E'_\nu \in \mathcal{E}_h(T_\nu) \setminus \{E\}$, for $x \in T_\nu$ we have $\tilde{\zeta}_E(x) := \zeta_E(x_E)$ where $x_E \in E$ is such that $x - x_E$ is parallel to E'_ν. Again, referring to [27], there exist positive constants $c_i,\ 4 \le i \le 5$, which only depend on the shape regularity of $\{\mathcal{T}_h(\Omega)\}$ such that

$$\|\tilde{\zeta}_E \vartheta_E\|_{0,\omega_E} \ \le\ c_4 h_E^{1/2} \|\zeta_E\|_{0,E}, \tag{6.4a}$$

$$|\tilde{\zeta}_E \vartheta_E|_{1,\omega_E} \ \le\ c_5 h_E^{-1/2} \|\zeta_E\|_{0,E}. \tag{6.4b}$$

Lemma 6.2. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3a)–(2.3d) and (3.2a)–(3.2d) and let $\eta_{y,T}, \mathrm{osc}_T(f)$ be given by (4.2a) and (4.5b), respectively. Then, there exists a positive constant c depending only on the shape regularity of $\{\mathcal{T}_h(\Omega)\}$ such that for $T \in \mathcal{T}_h(\Omega)$*

$$\eta_{y,T}^2 \ \le\ c\Big(|y - y_h|_{1,T}^2 + h_T^2\ \|u - u_h\|_{0,T}^2 + \mathrm{osc}_T^2(f) \Big). \tag{6.5}$$

Proof. We have

$$\eta_{y,T}^2 \ =\ h_T^2 \|f + u_h\|_{0,T}^2 \ \le\ 2\ h_T^2 \|f_h + u_h\|_{0,T}^2 + 2\,\mathrm{osc}_T^2(f). \tag{6.6}$$

Setting $z_h := (f_h + u_h)|_T \vartheta_T$, applying (6.2a) and observing that $\Delta y_h|_T = 0$, Green's formula and the fact that z_h is an admissible test function in (2.3a) imply

$$\begin{aligned} h_T^2 \|f_h + u_h\|_{0,T}^2 \ \le\ c_1 h_T^2 (f_h + u_h + \Delta y_h, z_h)_{0,T} &= c_1 h_T^2 \Big(- a(y_h, z_h) + (f + u, z_h)_{0,T} \\ &\quad + (f_h - f, z_h)_{0,T} + (u_h - u, z_h)_{0,T} \Big) \\ &= c_1 h_T^2 \Big(a(y - y_h, z_h) + ((f_h - f) + (u_h - u), z_h)_{0,T} \Big) \\ &\le c_1 \Big(h_T^2\ |y - y_h|_{1,T} |z_h|_{1,T} + \big(h_T^2\ \|u - u_h\|_{0,T} \\ &\quad +\ h_T\, \mathrm{osc}_T(f)\big) \|z_h\|_{0,T} \Big). \end{aligned} \tag{6.7}$$

Now, by (6.2b), (6.2c) and Young's inequality, (6.7) gives rise to

$$h_T^2 \|f_h + u_h\|_{0,T}^2 \le 2c_1^2 \Big(c_2 |y - y_h|_{1,T}^2 + h_T^2 \|u - u_h\|_{0,T}^2 + \mathrm{osc}_T^2(f) \Big) + \frac{1}{2} h_T^2 \|f_h + u_h\|_{0,T}^2. \tag{6.8}$$

Combining (6.6) and (6.8), readily gives the assertion. □

Lemma 6.3. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3a)–(2.3d) and (3.2a)–(3.2d) and let $\eta_{p,T}^{(1)}, \mathrm{osc}_T(y^d)$ be given by (4.2b) and (4.5a), respectively. Then, there exists a positive constant c depending only on the shape regularity of $\{\mathcal{T}_h(\Omega)\}$ such that for $T \in \mathcal{T}_h(\Omega)$*

$$(\eta_{p,T}^{(1)})^2 \ \le\ c\Big(|p - p_h|_{1,T}^2 + h_T^2\ \|y - y_h\|_{0,T}^2 + \mathrm{osc}_T^2(y^d) \Big). \tag{6.9}$$

Proof. The assertion (6.9) follows using the same arguments as in the proof of the previous lemma. □

Lemma 6.4. *Let (y, p, u, σ) and $(y_h, p_h, u_h, \sigma_h)$ be the solutions of (2.3a)–(2.3d) and (3.2a)–(3.2d) and let $\eta_{p,T}^{(2)}$ and $\mu_T(u^d)$ be given by (4.2c) and (4.4a), respectively. Then, for $T \in \mathcal{T}_h(\Omega)$ there holds*

$$\eta_{p,T}^{(2)} \ \le\ \|p - p_h\|_{0,T} + \|\sigma - \sigma_h\|_{0,T} + \alpha\Big(\|u - u_h\|_{0,T} + \mu_T(u^d) \Big). \tag{6.10}$$

Proof. We have

$$\|M_h p_h - p_h\|_{0,T} \;\leq\; \|p - p_h\|_{0,T} + \|M_h p_h - p\|_{0,T}.$$

Observing (2.3c) and (3.2c), for the second term on the right-hand side we find

$$\|M_h p_h - p\|_{0,T} \;\leq\; \|\sigma - \sigma_h\|_{0,T} + \alpha\Big(\|u - u_h\|_{0,T} + \mu_T(u^d)\Big),$$

which readily gives (6.10). □

Lemma 6.5. *Let (y,p,u,σ) and (y_h,p_h,u_h,σ_h) be the solutions of (2.3a)–(2.3d) and (3.2a)–(3.2d) and let $\eta_{y,T}$ and $\eta_{y,E}$ be given by (4.2a) and (4.2d), respectively. Then, there exists a positive constant c depending only on the shape regularity of $\{\mathcal{T}_h(\Omega)\}$ such that for $E \in \mathcal{E}_h(\Omega)$*

$$\eta_{y,E}^2 \;\leq\; c\left(|y - y_h|_{1,\omega_E}^2 + h_E^2\|u - u_h\|_{0,\omega_E}^2 + \sum_{\nu=1}^{2}\eta_{y,T_\nu}^2\right), \tag{6.11}$$

where ω_E is the patch as given by (6.3).

Proof. We set $\zeta_E := (\boldsymbol{\nu}_E \cdot [\nabla y_h])|_E$ and $z_h := \tilde{\zeta}_E \vartheta_E$. Then, using (6.2d), applying Green's formula, observing that z_h is an admissible test function in (2.3a), and taking advantage of (6.4a), (6.4b), we find

$$\begin{aligned}
\eta_{y,E}^2 &= h_E\|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}^2 \leq\; c_3 h_E\,(\boldsymbol{\nu}_E \cdot [\nabla y_h], \zeta_E\vartheta_E)_{0,E}\\
&= c_3 h_E \sum_{\nu=1}^{2} (\boldsymbol{\nu}_{\partial T_\nu} \cdot [\nabla y_h], z_h)_{0,\partial T_\nu}\\
&= c_3 h_E\Big(a(y_h - y, z_h) + (u - u_h, z_h)_{0,\omega_E} + (f + u_h, z_h)_{0,\omega_E}\Big)\\
&\leq c_3 h_E^{1/2}\|\boldsymbol{\nu}_E \cdot [\nabla y_h]\|_{0,E}\left(c_5|y - y_h|_{1,\omega_E} + c_4\big(h_E\|u - u_h\|_{0,\omega_E} + \big(\sum_{\nu=1}^{2}\eta_{y,T_\nu}^2\big)^{1/2}\big)\right).
\end{aligned}$$

An application of Young's inequality results in (6.11). □

Lemma 6.6. *Let (y,p,u,σ) and (y_h,p_h,u_h,σ_h) be the solutions of (2.3a)–(2.3d) and (3.2a)–(3.2d) and let $\eta_{p,T}^{(1)}$ and $\eta_{p,E}$ be given by (4.2b) and (4.2e), respectively. Then, there exists a positive constant c depending only on the shape regularity of $\{\mathcal{T}_h(\Omega)\}$ such that for $E \in \mathcal{E}_h(T),\ T \in \mathcal{T}_h(\Omega)$*

$$\eta_{p,E}^2 \;\leq\; c\left(|p - p_h|_{1,\omega_E}^2 + h_E^2\|y - y_h\|_{0,\omega_E}^2 + \sum_{\nu=1}^{2}(\eta_{p,T_\nu}^{(1)})^2\right), \tag{6.12}$$

where ω_E is the patch as given by (6.3).

Proof. The assertion (6.12) can be verified along the same lines of proof as in Lemma 6.5. □

Remark 6.7. The lower estimates provided by Lemmas 6.2 to 6.6 show that the magnitude of the element and edge residuals can be used for the purpose of mesh adaptivity as will be described in detail in the subsequent section.

7. The adaptive refinement process

The refinement of the triangulation $\mathcal{T}_h(\Omega)$ is based on a bulk criterion that has been previously used in the convergence analysis of adaptive finite element for nodal finite element methods [8,13,25] and for nonconforming, mixed and edge element methods [10–12]. Here, we adopt the bulk criterion for the finite element approximation of the distributed optimal control problem under consideration: Given the universal constants Θ_i, $1 \le i \le 4$ with $0 < \Theta_i < 1$, the outcome is a set of edges $\mathcal{M}^E \subset \mathcal{E}_h(\Omega)$ and sets of elements $\mathcal{M}^{\eta,T}, \mathcal{M}^{\mu,T}, \mathcal{M}^{\mathrm{osc},T} \subset \mathcal{E}_h(\Omega)$ such that

$$\Theta_1\Big(\sum_{E\in\mathcal{E}_h(\Omega)} (\eta_{y,E}^2 + \eta_{p,E}^2)\Big) \le \sum_{E\in\mathcal{M}^E} (\eta_{y,E}^2 + \eta_{p,E}^2), \tag{7.1}$$

$$\Theta_2\Big(\sum_{T\in\mathcal{T}_h(\Omega)} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2)\Big) \le \sum_{T\in\mathcal{M}^{\eta,T}} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2), \tag{7.2}$$

$$\Theta_3\Big(\sum_{T\in\mathcal{T}_h(\Omega)} \mu_T^2(u^d) + \sum_{T\in\mathcal{A}_{u_h}} \mu_T^2(\psi)\Big) \le \sum_{T\in\mathcal{M}^{\mu,T}} (\mu_T^2(u^d) + \mu_T^2(\psi)), \tag{7.3}$$

$$\Theta_4\Big(\sum_{T\in\mathcal{T}_h(\Omega)} \mathrm{osc}_T^2(y^d) + \mathrm{osc}_T^2(f)\Big) \le \sum_{T\in\mathcal{M}^{\mathrm{osc},T}} (\mathrm{osc}_T^2(y^d) + \mathrm{osc}_T^2(f)). \tag{7.4}$$

We set

$$\mathcal{M}^T \;:=\; \mathcal{M}^{\eta,T} \cup \mathcal{M}^{\mu,T} \cup \mathcal{M}^{\mathrm{osc},T}$$

and refine an element $T \in \mathcal{T}_h(\Omega)$ regularly (*i.e.*, subdividing it into four congruent subtriangles by joining the midpoints of the edges), if

- $T \in \mathcal{M}^T$ or
- at least two edges $E \in \mathcal{E}_h(T)$ belong to $\mathcal{M}^E$.

Denoting by $\mathcal{N}_T := \{T' \in \mathcal{T}_h(\Omega) | T' \cap T \ne \emptyset\}$ the set of all neighboring triangles of $T \in \mathcal{T}_h(\Omega)$, we define the set

$$\mathcal{F}_h(u_h) \;:=\; \partial\mathcal{A}(u_h) \cup \partial\mathcal{I}(u_h),$$

where

$$\partial\mathcal{A}(u_h) := \bigcup \{T \subset \mathcal{A}(u_h) \mid \mathcal{N}_T \cap \mathcal{I}(u_h) \ne \emptyset\},$$
$$\partial\mathcal{I}(u_h) := \bigcup \{T \subset \mathcal{I}(u_h) \mid \mathcal{N}_T \cap \mathcal{A}(u_h) \ne \emptyset\}.$$

The set $\mathcal{F}_h(u_h)$ represents a neighborhood of the discrete free boundary between the discrete active and inactive sets $\mathcal{A}(u_h)$ and $\mathcal{I}(u_h)$. In order to guarantee a sufficient resolution of the continuous free boundary, at each refinement step, the elements $T \in \mathcal{F}_h(u_h)$ are refined regularly.

Further irregular refinements by bisection are only performed in order to guarantee that the refined triangulation is geometrically conforming.

The bulk criterion (7.1)–(7.4) is realized by the following greedy algorithm:

Algorithm (bulk criterion):

Step 1. Initialization:
Set

$$\mathcal{M}_0^E \;:=\; \emptyset, \quad \mathcal{M}_0^{T,\eta} \;:=\; \mathcal{F}_h(u_h) \quad \text{and} \quad k = 0.$$

Step 2. Iteration loop:

Step 2a. Check edge residuals:
If

$$\Theta_1\left(\sum_{E\in\mathcal{E}_h(\Omega)} (\eta_{y,E}^2 + \eta_{p,E}^2)\right) \le \sum_{E\in\mathcal{M}^E} (\eta_{y,E}^2 + \eta_{p,E}^2),$$

go to Step 2b, else select some

$$F \in \mathcal{E}_h(\Omega) \setminus \mathcal{M}_k^E$$

such that

$$\eta_{E,F} \;=\; \max_{G \in \mathcal{E}_h(\Omega) \setminus \mathcal{M}_k^E} \Big(\eta_{y,G}, \eta_{p,G} \Big)$$

and set

$$\mathcal{M}_{k+1}^E \;:=\; \mathcal{M}_k^E \cup \{F\}, \quad k \;:=\; k+1.$$

Step 2b. Check element residuals:
Set

$$\mathcal{M}_k^{T,\eta} \;:=\; \bigcup \Big\{ T \in \mathcal{T}_h(\Omega) \mid \text{card}\,\Big(\mathcal{E}_h(T) \cap \mathcal{M}_k^E \Big) \;\geq\; 2 \Big\}.$$

If

$$\Theta_2 \,\Big(\sum_{T \in \mathcal{T}_h(\Omega)} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2) \Big) \leq \sum_{T \in \mathcal{M}^{\eta,T}} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2),$$

go to Step 2c, else select some

$$\eta_{T,S} \;:=\; \max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_k^{\eta,T}} \Big(\eta_{y,T}, \eta_{p,T}^{(1)}, \eta_{p,T}^{(2)} \Big)$$

and set

$$\mathcal{M}_{k+1}^{\eta,T} \;:=\; \mathcal{M}_k^{\eta,T} \cup \{S\}, \quad k \;:=\; k+1.$$

Step 2c. Check low order data residuals:
Set

$$\mathcal{M}_k^{\mu,T} \;:=\; \mathcal{M}_k^{\eta,T}.$$

If

$$\Theta_3 \,\Big(\sum_{T \in \mathcal{T}_h(\Omega)} \mu_T^2(u^d) + \sum_{T \in \mathcal{A}_{u_h}} \mu_T^2(\psi) \Big) \;\leq\; \sum_{T \in \mathcal{M}^{\mu,T}} (\mu_T^2(u^d) + \mu_T^2(\psi)),$$

go to Step 2d, else select some

$$\eta_{T,S} \;:=\; \max(\max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_k^{\mu,T}} \mu_T(u^d), \max_{T \in \mathcal{A}(u_h) \setminus \mathcal{M}_k^{\mu,T}} \mu_T(\psi))$$

and set

$$\mathcal{M}_{k+1}^{T,\mu} \;:=\; \mathcal{M}_k^{T,\mu} \cup \{S\}, \quad k \;:=\; k+1.$$

Step 2d. Check remaining data residuals:
Set

$$\mathcal{M}_k^{\text{osc},T} \;:=\; \mathcal{M}_k^{\mu,T}.$$

If

$$\Theta_4 \,\Big(\sum_{T \in \mathcal{T}_h(\Omega)} \text{osc}_T^2(y^d) + \text{osc}_T^2(f)) \Big) \leq \sum_{T \in \mathcal{M}^{\text{osc},T}} (\text{osc}_T^2(y^d) + \text{osc}_T^2(f)),$$

go to Step 3, else select some

$$\eta_{T,S} \;:=\; \max_{T \in \mathcal{T}_h(\Omega) \setminus \mathcal{M}_k^{\text{osc},T}} \Big(\text{osc}_T(y^d), \text{osc}_T(f) \Big)$$

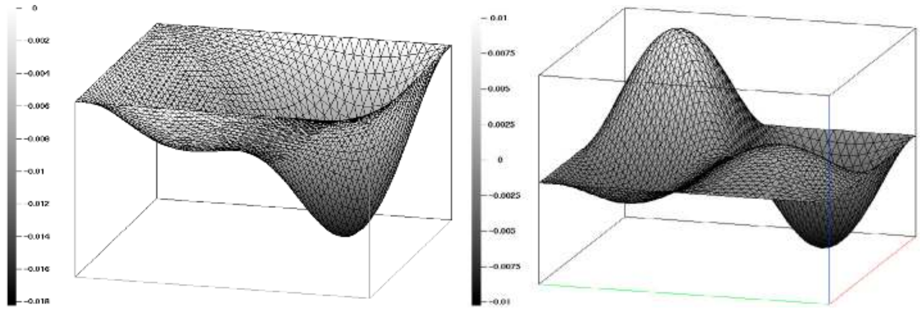

FIGURE 1. Example 1: visualization of the optimal state y (left) and the optimal adjoint state p (right).

and set

$$\mathcal{M}_{k+1}^{\mathrm{osc},T} \ := \ \mathcal{M}_k^{\mathrm{osc},T} \cup \{S\}, \quad k \ := \ k+1.$$

Step 3. Final output:
Output the set of marked edges and elements

$$\mathcal{M}^E \ := \ \mathcal{M}_k^E, \quad \mathcal{M}^T \ := \ \mathcal{M}_k^{\mathrm{osc},T}.$$

8. Numerical results

We provide a documentation of numerical results illustrating the performance of the adaptive finite element approximation for two representative distributed optimal control problems that have been considered in [7] in the framework of primal-dual active set strategies as iterative solvers for such kind of control problems (*cf.* also [16]). In particular, the second example considers a variable obstacle and exhibits a lack of strict complementarity. It thus features particular cases that have not been included in the numerical examples presented in [20]. Moreover, the numerical results clearly demonstrate that, at least at the beginning of the refinement process, the data oscillations have to be taken into account.

Example 1. Constant obstacle.
The data in the optimal distributed control problem (2.1a)–(2.1c) are chosen as follows:

$$\begin{aligned} \Omega &:= (0,1)^2, \quad y^d \ := \ \sin(2\pi x_1)\sin(2\pi x_2)\frac{\exp(2x_1)}{6}, \\ u^d &:= 0, \quad \alpha \ := \ 0.01, \quad \psi \ := \ 0, \quad f \ := \ 0. \end{aligned}$$

Figures 1 and 2 show a visualization of the optimal state, the optimal adjoint state, the optimal control, and the optimal adjoint control, respectively.

The initial simplicial triangulation $\mathcal{T}_{h_0}$ was chosen according to a subdivision of Ω by joining the four vertices resulting in one interior nodal point and four congruent triangles. Since the obstacle ψ is zero, we have $\psi_h = 0$ as well. Moreover, since also $u^d = 0$ and $f = 0$, for the data oscillations we have $\mu_h(u^d) = 0$ and $\mathrm{osc}_h(f) = \mathrm{osc}_h(\psi) = 0$.

Figure 3 displays the adaptively generated triangulations after six (left) and eight (right) refinement steps with $\Theta_i = 0.6, \ 1 \leq i \leq 4$, in the bulk criterion. The two areas at the upper left and the bottom right corner represent the discrete inactive set $\mathcal{I}(u_h)$, whereas the simply-connected region in between is the discrete active set $\mathcal{A}(u_h)$. The continuous free boundary between the continuous active and inactive sets is indicated by the black curves. We see that the continuous free boundary is accurately resolved by the adaptive refinement process. Moreover, there are local areas of pronounced refinement within the discrete active and inactive sets.

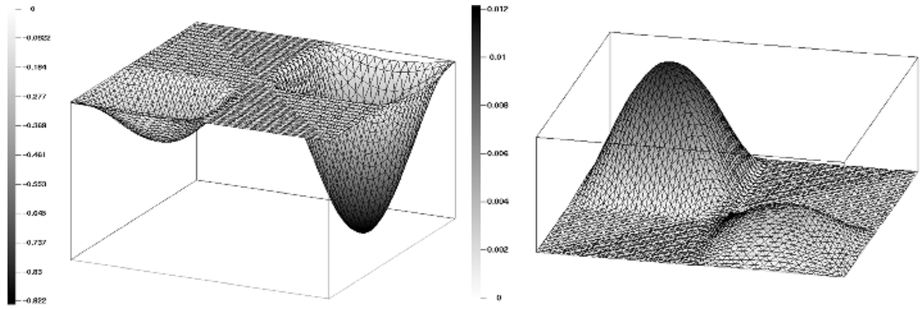

FIGURE 2. Example 1: visualization of the optimal control u (left) and the optimal adjoint control σ (right).

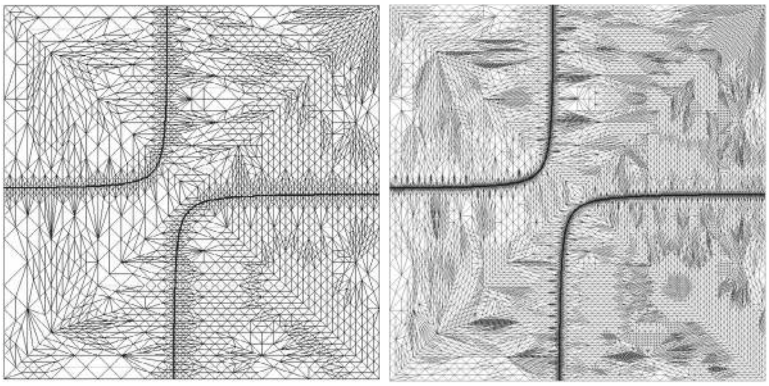

FIGURE 3. Example 1: adaptively generated grid after 6 (left) and 8 (right) refinement steps, $\Theta_i = 0.6$.

TABLE 1. Example 1: convergence history of the adaptive FEM, part I: total discretization error and discretization errors in the state, adjoint state, control, and adjoint control.

l	$N_{\rm dof}$	$\vert\vert\vert z - z_h\vert\vert\vert$	$\vert y - y_h\vert_1$	$\vert p - p_h\vert_1$	$\Vert u - u_h\Vert_0$	$\Vert \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\Vert_0$
1	13	2.27 e–01	1.92 e–02	1.48 e–02	1.91 e–01	2.11 e–03
2	41	1.24 e–01	1.34 e–02	1.36 e–02	9.59 e–02	1.06 e–03
3	126	6.19 e–02	6.83 e–03	7.86 e–03	4.67 e–02	5.48 e–04
4	374	3.57 e–02	3.93 e–03	4.89 e–03	2.65 e–02	3.67 e–04
5	968	2.50 e–02	2.63 e–03	3.34 e–03	1.88 e–02	2.50 e–04
6	2553	1.77 e–02	1.92 e–03	2.33 e–03	1.33 e–02	1.57 e–04
7	5396	1.25 e–02	1.31 e–03	1.67 e–03	9.39 e–03	1.17 e–04
8	12 318	8.71 e–03	9.34 e–04	1.17 e–03	6.53 e–03	7.58 e–05
9	26 887	6.17 e–03	6.52 e–04	8.38 e–04	4.62 e–03	5.66 e–05

It should be emphasized that we are working with only one grid for all variables (state, adjoint state, control, and adjoint control). Hence, the grid reflects regions of substantial change in all these variables (*cf.* Figs. 1 and 2).

More detailed information is given in Tables 1–4. In particular, Table 1 displays the error reduction in the total error

$$|||z - z_h||| \;:=\; (|y - y_h|^2_{1,\Omega} + |p - p_h|^2_{1,\Omega} + \|u - u_h\|^2_{0,\Omega} + \|\sigma - \sigma_h\|^2_{0,\Omega})^{1/2}$$

TABLE 2. Example 1: convergence history of the adaptive FEM, part II: components of the error estimator and data oscillations.

l	$N_{\rm dof}$	$\boldsymbol{\eta}_y$	$\boldsymbol{\eta}_p$	$\mathrm{osc}_h(y^d)$
1	13	7.73 e–02	1.56 e–01	1.12 e–01
2	41	5.79 e–02	8.29 e–02	2.58 e–02
3	126	3.72 e–02	4.63 e–02	8.06 e–03
4	374	2.26 e–02	2.92 e–02	3.79 e–03
5	968	1.53 e–02	1.98 e–02	1.86 e–03
6	2553	1.11 e–02	1.35 e–02	8.74 e–04
7	5396	7.51 e–03	9.38 e–03	4.39 e–04
8	12 318	5.35 e–03	6.61 e–03	2.24 e–04
9	26 887	3.68 e–03	4.77 e–03	1.13 e–04

TABLE 3. Example 1: convergence history of the adaptive FEM, part III: average values of the local estimators.

l	$N_{\rm dof}$	$\boldsymbol{\eta}_{y,T}$	$\boldsymbol{\eta}_{y,E}$	$\boldsymbol{\eta}_{p,T}^{(1)}$	$\boldsymbol{\eta}_{p,T}^{(2)}$	$\boldsymbol{\eta}_{p,E}$
1	13	9.72 e–03	6.07 e–03	3.27 e–02	5.33 e–04	1.17 e–02
2	41	3.15 e–03	2.15 e–03	7.16 e–03	1.31 e–04	3.04 e–03
3	126	1.07 e–03	8.33 e–04	2.01 e–03	3.89 e–05	1.10 e–03
4	374	3.69 e–04	2.95 e–04	6.58 e–04	1.33 e–05	3.86 e–04
5	968	1.44 e–04	1.19 e–04	2.52 e–04	5.22 e–06	1.59 e–04
6	2553	6.12 e–05	5.42 e–05	1.06 e–04	2.21 e–06	7.30 e–05
7	5396	2.71 e–05	2.55 e–05	4.70 e–05	9.82 e–07	3.45 e–05
8	12 318	1.21 e–05	1.22 e–05	2.09 e–05	4.36 e–07	1.68 e–05
9	26 887	5.61 e–06	5.92 e–06	9.65 e–06	2.02 e–07	8.21 e–06

TABLE 4. Example 1: convergence history of the adaptive FEM, part IV: average values of the data oscillations, bulk criterion.

l	$N_{\rm dof}$	$\mathrm{osc}(y^d)$	$M_{fb,T}$	$M_{\boldsymbol{\eta},E}$	$M_{\boldsymbol{\eta},T}$	$M_{\mathrm{osc},T}$
1	13	2.53 e–02	68.8	35.0	31.2	37.5
2	41	2.83 e–03	42.2	13.6	15.6	25.0
3	126	4.47 e–04	25.9	16.7	16.5	20.1
4	374	9.47 e–05	16.3	16.8	14.4	6.1
5	968	2.53 e–05	12.0	14.8	11.5	5.1
6	2553	7.18 e–06	9.9	12.9	12.2	3.3
7	5396	2.28 e–06	8.6	12.5	13.5	3.8
8	12 318	6.90 e–07	7.5	11.7	11.8	1.5
9	26 887	2.31 e–07	7.5	10.8	13.6	2.1

and the errors in the state, the adjoint state, the control, and the adjoint control, respectively. On the other hand, the actual components of the residual type *a posteriori* error estimator are given in Table 2, whereas Table 3 contains the average values of the local element and edge contributions of the error estimator. Finally, Table 4 lists the average values of the local data oscillation $\mathrm{osc}_T(y^d)$, $T \in \mathcal{T}_h(\Omega)$ and the percentages of elements and edges that have been marked for refinement according to the bulk criterion. Here, $M_{fb,T}$, $M_{\eta,T}$ and $M_{\mathrm{osc},T}$

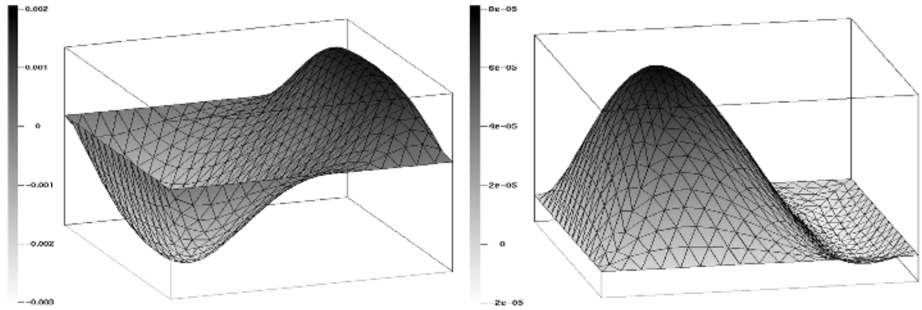

FIGURE 4. Example 2: visualization of the optimal state y (left) and the optimal adjoint state p (right).

stand for the level l elements marked for refinement due to the resolution of the free boundary, the element residuals, and the data oscillations, respectively, whereas $M_{\eta,E}$ refers to the edges marked for refinement with regard to the edge residuals. On the coarsest grid, the sum of the percentages exceeds 100%, since an element $T \in \mathcal{T}_h(\Omega)$ may satisfy more than one criterion in the adaptive refinement process. The refinement is initially dominated by the resolution of the free boundary, whereas at a later stage edge and element residuals dominate.

The second example features a variable obstacle and is such that no strict complementarity holds at the optimal solution.

Example 2. Variable obstacle.

The data in (2.1a)–(2.1c) have been chosen as follows:

$$\Omega := (0,1)^2, \quad y^d := 0, \quad u^d := \hat{u} + \alpha^{-1} (\hat{\sigma} + \Delta^{-2}\hat{u}),$$
$$\psi := \begin{cases} (x_1 - 0.5)^8, & (x_1,x_2) \in \Omega_1 \\ (x_1 - 0.5)^2, & \text{otherwise} \end{cases}, \quad \alpha := 0.1, \quad f := 0.$$

Here, $\hat{u}$ and $\hat{\sigma}$ are given by

$$\hat{u} := \begin{cases} \psi(x_1,x_2), & (x_1,x_2) \in \Omega_1 \cup \Omega_2 \\ -1.01\ \psi(x_1,x_2), & \text{otherwise} \end{cases}$$
$$\hat{\sigma} := \begin{cases} 2.25\ (x_1 - 0.75) \times 10^{-4}, & (x_1,x_2) \in \Omega_2 \\ 0, & \text{otherwise} \end{cases}$$

with Ω_1 and Ω_2 specified as follows

$$\Omega_1 := \{(x_1,x_2) \in \Omega \mid ((x_1 - 0.5)^2 + (x_2 - 0.5)^2)^{1/2} \leq 0.15\},$$
$$\Omega_2 := \{(x_1,x_2) \in \Omega \mid x_1 \geq 0.75\}.$$

Figures 4 and 5 display the optimal state y, the optimal adjoint state p, the optimal control $u = \hat{u}$, and the optimal adjoint control $\sigma = \hat{\sigma}$, respectively.

The initial simplicial triangulation $\mathcal{T}_{h_0}$ has been chosen as in Example 1, whereas the parameters Θ_i in the bulk criterion have been specified according to $\Theta_i = 0.7,\ 1 \leq i \leq 4$. Figure 6 shows the adaptively generated triangulations after six (left) and eight (right) refinement steps. As in Example 1, we see that the continuous free boundary $\mathcal{F} := \{(x_1,x_2) \in \Omega \mid x_1 = 0.75\}$ and the boundary layer at the left vertical boundary of the computational domain (*cf.* Fig. 4) are well resolved by the adaptive solution process.

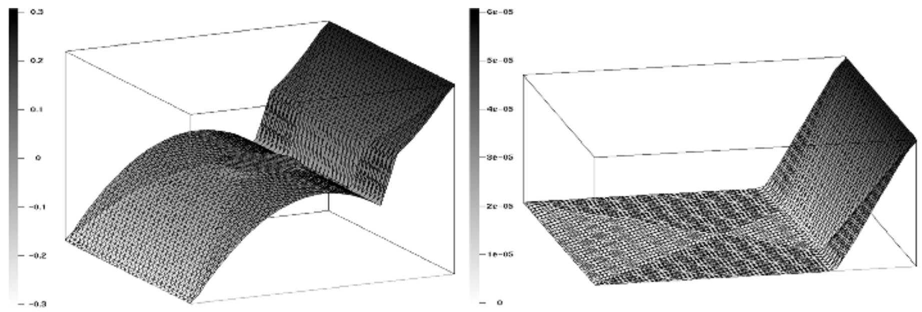

FIGURE 5. Example 2: visualization of the optimal control u (left) and the optimal adjoint control σ (right).

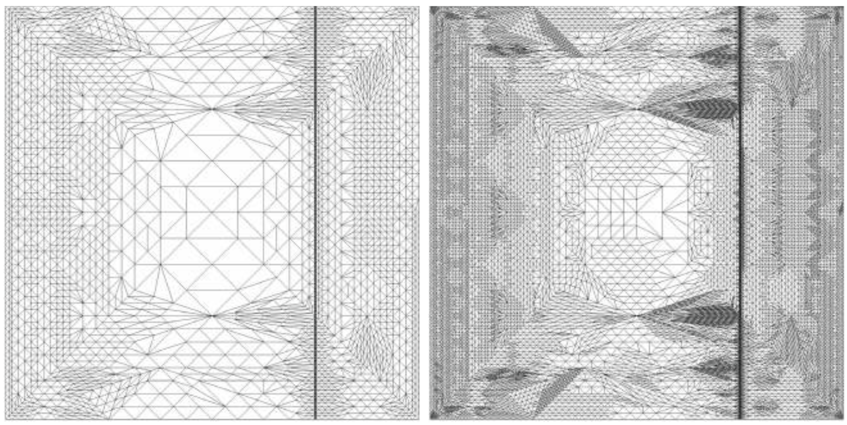

FIGURE 6. Example 2: adaptively generated grid after 6 (left) and 8 (right) refinement steps, $\Theta_i = 0.7$.

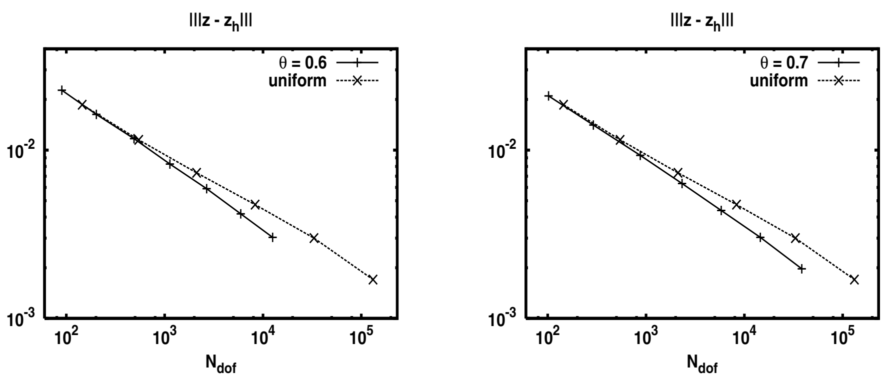

FIGURE 7. Example 2: adaptive *versus* uniform refinement, $\Theta_i = 0.6$ (left) and $\Theta_i = 0.7$ (right).

Figure 7 displays the benefit of adaptive *versus* uniform refinement. In particular, the total discretization error in the state, adjoint state, control, and adjoint control is shown as a function of the total number of degrees of freedom.

Table 5. Example 2: convergence history of the adaptive FEM, part I: total discretization error and discretization errors in the state, adjoint state, control, and adjoint control.

l	N_{dof}	$\|\|\|z - z_h\|\|\|$	$\|y - y_h\|_1$	$\|p - p_h\|_1$	$\\|u - u_h\\|_0$	$\\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\\|_0$
1	13	5.36 e–02	6.85 e–03	1.04 e–04	4.66 e–02	8.86 e–06
2	41	3.12 e–02	3.83 e–03	5.99 e–05	2.74 e–02	4.63 e–06
3	102	2.10 e–02	2.39 e–03	4.10 e–05	1.85 e–02	2.29 e–06
4	291	1.41 e–02	1.58 e–03	2.94 e–05	1.24 e–02	1.39 e–06
5	873	9.31 e–03	9.73 e–04	1.93 e–05	8.32 e–03	8.41 e–07
6	2325	6.33 e–03	6.17 e–04	1.22 e–05	5.70 e–03	5.60 e–07
7	5816	4.38 e–03	4.02 e–04	7.62 e–06	3.97 e–03	3.76 e–07
8	14 524	3.03 e–03	2.66 e–04	5.26 e–06	2.76 e–03	2.42 e–07
9	38 364	1.97 e–03	1.71 e–04	3.42 e–06	1.80 e–03	1.54 e–07

Table 6. Example 2: convergence history of the adaptive FEM, part II: components of the error estimator and data oscillations.

l	N_{dof}	$\boldsymbol{\eta}_y$	$\boldsymbol{\eta}_p$	$\boldsymbol{\mu}_h(u^d)$	$\boldsymbol{\mu}_h(\boldsymbol{\psi})$
1	13	5.45 e–02	5.76 e–04	4.77 e–02	3.93 e–02
2	41	2.99 e–02	3.30 e–04	2.64 e–02	2.06 e–02
3	102	1.72 e–02	2.71 e–04	1.83 e–02	1.34 e–02
4	291	1.01 e–02	1.80 e–04	1.21 e–02	8.62 e–03
5	873	6.10 e–03	1.21 e–04	8.33 e–03	5.49 e–03
6	2325	3.93 e–03	7.50 e–05	5.74 e–03	3.73 e–03
7	5816	2.54 e–03	4.84 e–05	4.22 e–03	2.34 e–03
8	14 524	1.65 e–03	3.23 e–05	3.08 e–03	1.55 e–03
9	38 364	1.07 e–03	2.17 e–05	2.34 e–03	1.02 e–03

Table 7. Example 2: convergence history of the adaptive FEM, part III: average values of the local estimators.

l	N_{dof}	$\boldsymbol{\eta}_{y,T}$	$\boldsymbol{\eta}_{y,E}$	$\boldsymbol{\eta}^{(1)}_{p,T}$	$\boldsymbol{\eta}^{(2)}_{p,T}$	$\boldsymbol{\eta}_{p,E}$
1	13	1.01 e–02	2.50 e–03	1.22 e–04	2.19 e–06	4.21 e–05
2	41	2.52 e–03	8.10 e–04	3.05 e–05	6.38 e–07	1.43 e–05
3	102	9.61 e–04	3.46 e–04	1.17 e–05	2.69 e–07	5.94 e–06
4	291	3.24 e–04	1.33 e–04	3.95 e–06	9.53 e–08	2.13 e–06
5	873	1.06 e–04	4.73 e–05	1.27 e–06	3.14 e–08	7.42 e–07
6	2325	3.94 e–05	1.94 e–05	4.70 e–07	1.19 e–08	3.02 e–07
7	5816	1.57 e–05	8.19 e–06	1.87 e–07	4.75 e–09	1.28 e–07
8	14 524	6.35 e–06	3.55 e–06	7.55 e–08	1.94 e–09	5.37 e–08
9	38 364	2.42 e–06	1.44 e–06	2.88 e–08	7.45 e–10	2.20 e–08

Tables 5–8 contain the same information as Tables 1–4 for Example 1. Since in Example 2, the obstacle ψ is not constant, the data oscillation $\mu_h(\psi)$ has been considered. As far as the selection step MARK is concerned,we again observe a pronounced refinement for the resolution of the free boundary at the beginning of the refinement process, whereas edge and element residuals dominate at a later stage.

TABLE 8. Example 2: convergence history of the adaptive FEM, part IV: average values of the data oscillations, bulk criterion.

l	$\boldsymbol{\mu}_T(u^d)$	$\boldsymbol{\mu}_T(\boldsymbol{\psi})$	$M_{fb,T}$	$M_{\boldsymbol{\eta},E}$	$M_{\boldsymbol{\eta},T}$	$M_{\boldsymbol{\mu},T}$
1	1.04 e–02	8.88 e–03	37.5	25.0	25.0	43.8
2	2.70 e–03	2.26 e–03	18.8	9.1	21.9	34.4
3	1.09 e–03	8.86 e–04	14.0	12.8	31.4	23.3
4	3.83 e–04	3.10 e–04	9.1	14.0	35.7	16.7
5	1.29 e–04	1.03 e–04	5.8	14.6	32.5	9.6
6	4.98 e–05	3.97 e–05	4.3	12.8	28.7	7.4
7	2.04 e–05	1.61 e–05	3.4	13.6	29.1	3.4
8	8.33 e–06	6.51 e–06	2.7	16.1	32.0	1.5
9	3.26 e–06	2.51 e–06	2.0	14.6	28.6	0.9

Acknowledgements. The first two authors acknowledge the support by the NSF under Grant No. DMS-0411403. The fourth author expresses his thanks for support by the Bavarian Government within the elite curriculum TOPMATH.

References

[1] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis.* Wiley, Chichester (2000).

[2] I. Babuska and W. Rheinboldt, Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.* **15** (1978) 736–754.

[3] I. Babuska and T. Strouboulis, *The Finite Element Method and its Reliability.* Clarendon Press, Oxford (2001).

[4] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations. Lectures in Mathematics.* ETH-Zürich, Birkhäuser, Basel (2003).

[5] R.E. Bank and A. Weiser, Some *a posteriori* error estimators for elliptic partial differential equations. *Math. Comput.* **44** (1985) 283–301.

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

[7] M. Bergounioux, M. Haddou, M. Hintermüller and K. Kunisch, A comparison of a Moreau-Yosida based active set strategy and interior point methods for constrained optimal control problems. *SIAM J. Optim.* **11** (2000) 495–521.

[8] P. Binev, W. Dahmen and R. DeVore, Adaptive finite element methods with convergence rates. *Numer. Math.* **97** (2004) 219–268.

[9] C. Carstensen and S. Bartels, Each averaging technique yields reliable *a posteriori* error control in FEM on unstructured grids. Part I: Low order conforming, nonconforming, and mixed FEM. *Math. Comput.* **71** (2002) 945–969.

[10] C. Carstensen and R.H.W. Hoppe, Convergence analysis of an adaptive edge finite element method for the 2d eddy current equations. *J. Numer. Math.* **13** (2005) 19–32.

[11] C. Carstensen and R.H.W. Hoppe, Error reduction and convergence for an adaptive mixed finite element method. *Math. Comp.* **75** (2006) 1033–1042.

[12] C. Carstensen and R.H.W. Hoppe, Convergence analysis of an adaptive nonconforming finite element method. *Numer. Math.* **103** (2006) 251–266.

[13] W. Dörfler, A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* **33** (1996) 1106–1124.

[14] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational Differential Equations.* Cambridge University Press, Cambridge (1995).

[15] H.O. Fattorini, *Infinite Dimensional Optimization and Control Theory.* Cambridge University Press, Cambridge (1999).

[16] M. Hintermüller, A primal-dual active set algorithm for bilaterally control constrained optimal control problems. *Quart. Appl. Math.* **LXI** (2003) 131–161.

[17] J.B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms.* Springer, Berlin-Heidelberg-New York (1993).

[18] R.H.W. Hoppe and B. Wohlmuth, Element-oriented and edge-oriented local error estimators for nonconforming finite element methods. *RAIRO Modél. Math. Anal. Numér.* **30** (1996) 237–263.

[19] R.H.W. Hoppe and B. Wohlmuth, Hierarchical basis error estimators for Raviart-Thomas discretizations of arbitrary order, in *Finite Element Methods: Superconvergence, Post-Processing, and A Posteriori Error Estimates,* M. Krizek, P. Neittaanmäki and R. Steinberg Eds., Marcel Dekker, New York (1998) 155–167.

[20] R. Li, W. Liu H. Ma and T. Tang, Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM J. Control Optim.* **41** (2002) 1321–1349.

[21] X.J. Li and J. Yong, *Optimal Control Theory for Infinite Dimensional Systems.* Birkhäuser, Boston-Basel-Berlin (1995).

[22] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations.* Springer, Berlin-Heidelberg-New York (1971).

[23] W. Liu and N. Yan, A posteriori error estimates for distributed optimal control problems. *Adv. Comp. Math.* **15** (2001) 285–309.

[24] W. Liu and N. Yan, *A posteriori error estimates for convex boundary control problems.* Preprint, Institute of Mathematics and Statistics, University of Kent, Canterbury (2003).

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

[26] P. Neittaanmäki and S. Repin, *Reliable methods for mathematical modelling. Error control and a posteriori estimates.* Elsevier, New York (2004).

[27] R. Verfürth, *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques.* Wiley-Teubner, New York, Stuttgart (1996).

[28] O. Zienkiewicz and J. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis. *J. Numer. Meth. Eng.* **28** (1987) 28–39.