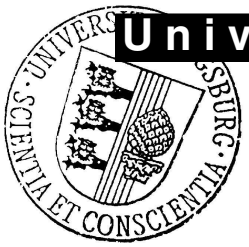

Universität Augsburg

Institut für
Mathematik

M. Hintermüller, Ronald H.W. Hoppe

Adaptive Finite Element Methods for Control Constrained Distributed and Boundary Optimal Control Problems

Preprint Nr. 17/2008 — 14. April 2008

Institut für Mathematik, Universitätsstraße, D-86 135 Augsburg http://www.math.uni-augsburg.de/

Impressum:

Herausgeber:

Institut für Mathematik
Universität Augsburg
86135 Augsburg
http://www.math.uni-augsburg.de/forschung/preprint/

ViSdP:

Ronald H.W. Hoppe
Institut für Mathematik
Universität Augsburg
86135 Augsburg

Adaptive finite element methods for control constrained distributed and boundary optimal control problems

M. Hintermüller[1] and R.H.W. Hoppe[2,3]

[1]Institute of Mathematics and Scientific Computing, University of Graz, A-8010 Graz, Austria
[2]Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany
[3]Department of Mathematics, University of Houston, Houston TX 77204-3008, U.S.A.

Abstract - This contribution is concerned with the development, analysis and implementation of Adaptive Finite Element Methods (AFEMs) for distributed and boundary control problems with control constraints. AFEMs consist of successive loops of the cycle 'SOLVE', 'ESTIMATE', 'MARK', and 'REFINE'. Emphasis will be on the steps 'SOLVE' and 'ESTIMATE'. In this context, 'SOLVE' stands for the efficient solution of the finite element discretized problems and the following step 'ESTIMATE' is devoted to a residual-type a posteriori error estimation of the global discretization errors in the state, the co-state, the control and the co-control. A bulk criterion is the core of the step 'MARK' to indicate selected edges and elements for refinement. The final step 'REFINE' deals with the technical realization of the refinement process itself. The efficient solution of the underlying constrained minimization problems is achieved by employing a primal-dual active set strategy. This method is equivalent to a class of semismooth Newton algorithms and converges locally at a superlinear rate. The a posteriori error analysis features convergence in the states, co-states, controls, and co-controls and addresses the issue of a guaranteed error reduction. Finally, numerical results for distributed as well as boundary control problems will be discussed. The test problems under consideration also include the case of lack of strict complementarity.

Keywords: a posteriori error analysis, boundary control, distributed control, control constraints, adaptive finite element methods, residual-type a posteriori error estimators, data oscillations

1 Introduction

Adaptive finite element methods have proved to be significant tools for the efficient numerical solution of boundary and initial-boundary value problems for partial differential equations as documented by a series of monographs on this subject [2, 4, 5, 14, 33, 34]. These methods are based on appropriately chosen a posteriori error estimators. The concepts for the a posteriori error analysis include residual-type estimators [3, 4, 34], hierarchical-type estimators [6, 25, 26], the so-called goal oriented dual weighted approach [5, 14], and functional type error majorants [33].

However, significantly less work has been performed in the framework of an a posteriori error analysis of adaptive finite element schemes for optimal control problems associated with partial differential equations. In the unconstrained case, we refer to [5, 7]. Residual-type a posteriori error estimators for control constrained distributed and boundary control problems have been derived and analyzed in [21, 24, 28, 30, 31]. In contrast to the approach used in [28, 30, 31], the error analysis in [21, 24] includes the error in the state, the co-state, the control, and the co-control and also takes data oscillations into account which may sig-

[1]The first author acknowledges the support by the Austrian Science Fund (FWF) within the START project Y305 "Interfaces and Free Boundaries". The second author has been partially supported by the NSF under Grant No. DMS-0411403 and Grant No. DMS-0511611.

nificantly contribute to the error and thus have to be considered in the adaptive refinement process. This paper contains a survey of the results from [17, 21] and [24]. It is organized as follows:
In section 2, we introduce classes of distributed and boundary control problems for a two-dimensional, second order elliptic equation with a quadratic objective functional and bilateral constraints on the control variable. The optimality conditions are given in terms of the state, the co-state, the control, and the Lagrangian multiplier for the control which will be referred to as the co-control. The control problems are discretized with respect to a family of shape regular, simplicial triangulations of the computational domain using continuous, piecewise linear finite elements for the state and the co-state and elementwise constant approximations of the control and the co-control. Section 3 is devoted to the description of an efficient solver for the underlying constrained minimization problems. By means of a nonlinear complementarity problem function the complementarity system involved in first order optimality conditions for inequality constrained problems is reformulated as a single not necessarily Fréchet-differentiable equality. Employing a generalized Newton methodology, the solution procedure is shown to be equivalent to a primal-dual active set strategy. The method converges locally at a superlinear rate in the original function space setting as well as after appropriate discretization in finite dimensional space. In Section 4, we are concerned with an a posteriori error analysis of residual-type error estimators for the global discretization errors in the state, the co-state, the control, and the co-control. We address the important issue of data oscillations and the so-called bulk criteria for an appropriate selection of elements and edges for refinement. We provide the reliability and the discrete local efficiency of the error estimators which are the basic tools to establish convergence results and to derive error reduction properties. Finally, section 5 contains numerical results for selected test examples illustrating the performance of the adaptive approach.

2 Distributed and boundary control problems

We suppose $\Omega \subset \mathbb{R}^2$ to be a bounded, polygonal domain with boundary $\Gamma := \partial\Omega$. We adopt standard notation from Lebesgue and Sobolev space theory (see, e.g., [1]) and refer to $(\cdot,\cdot)_{k,D}$ and $|\cdot|_{k,\Omega}, \|\cdot\|_{k,D}, k \in \mathbb{N}_0, D \subseteq \Omega$, as the $H^k(D)$-inner product and associated semi-norm and norm, respectively. For given $u^d, y^d \in L^2(\Omega), f \in L^2(\Omega), \psi^{(\nu)} \in L^\infty(\Omega), 1 \le \nu \le 2$, and $\alpha, c \in \mathbb{R}_+$, we consider the following distributed optimal control problem with constraints on the controls

$$\begin{aligned}
&\text{minimize } J(y,u) \;:=\; \frac{1}{2}\,\|y-y^d\|^2_{0,\Omega} \;+\; \frac{\alpha}{2}\,\|u-u^d\|^2_{0,\Omega} && \text{(2.1a)}\\
&\text{over } (y,u) \in V \times W := H^1_0(\Omega) \times L^2(\Omega),\\
&\text{subject to } -\Delta y + cy \;=\; f \;+\; u\;, && \text{(2.1b)}\\
&\qquad u \in K \;:=\; \{v \in L^2(\Omega) \mid \psi^{(1)} \le v \le \psi^{(2)} \text{ a.e.}\}\;. && \text{(2.1c)}
\end{aligned}$$

We also consider the constrained boundary control problem

$$\begin{aligned}
&\text{minimize } J(y,u) \;:=\; \frac{1}{2}\,\|y-y^d\|^2_{0,\Omega} \;+\; \frac{\alpha}{2}\,\|u-u^d\|^2_{0,\Gamma_1} && \text{(2.2a)}\\
&\text{over } (y,u) \in V \times W := H^1_{0,\Gamma_2}(\Omega) \times L^2(\Gamma_1),\\
&\text{subject to } -\Delta y + cy \;=\; f \quad \text{in } \Omega\;, && \text{(2.2b)}\\
&\qquad\qquad n \cdot \nabla y \;=\; u \quad \text{on } \Gamma_1\;,\\
&\qquad u \in K \;:=\; \{v \in L^2(\Gamma_1) \mid \psi^{(1)} \le v \le \psi^{(2)} \text{ a.e.}\}\;. && \text{(2.2c)}
\end{aligned}$$

Here, $\Gamma = \Gamma_1 \cup \Gamma_2, \Gamma_1 \cap \Gamma_2 = \emptyset$ and $u^d \in L^2(\Gamma_1), \psi^{(\nu)} \in L^\infty(\Gamma_1), 1 \le \nu \le 2$, whereas y^d, c, f and α are given as in (2.1a)-(2.1c).
It is well-known that (2.1a)-(2.1c) and (2.2a)-(2.2c) admit a unique solution $(y,u) \in V \times K$ (cf., e.g., [16, 29, 30]), respectively. The optimality conditions involve the existence of a co-state $p \in H_0^1(\Omega)$ and and a co-control $\sigma \in L_+^2(\Omega)$ resp. $\sigma \in L_+^2(\Gamma_1)$ such that y, p, u, σ satisfy

$$a(y,v) \;=\; (f,v)_{0,\Omega} + \ell_1(u,v) \quad , \quad v \in V \;, \tag{2.3a}$$

$$a(p,v) \;=\; (y^d,v)_{0,\Omega} - \ell_2(y,v) \quad , \quad v \in H_0^1(\Omega) \;, \tag{2.3b}$$

$$u - u^d \;= \frac{1}{\alpha}\,(p-\sigma) \;, \tag{2.3c}$$

$$\sigma \in \partial I_K(u) \;. \tag{2.3d}$$

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

$$a(w,z) \;:=\; (\nabla w, \nabla z)_{0,\Omega} \;+\; (cw,z)_{0,\Omega} \quad , \quad w,z \in V \;.$$

The functionals $\ell_1 : W \times V \to \mathbb{R}$ and $\ell_2 : V \times V \to \mathbb{R}$ are given by

$$\ell_1(u,v) \;:=\; (u,v)_{0,\Omega} \quad , \quad v \in V$$

for the distributed control problem and

$$\ell_1(u,v) \;:=\; (u,v)_{0,\Gamma_1} \quad , \quad v \in V$$

for the boundary control problem as well as

$$\ell_2(y,v) \;:=\; (y,v)_{0,\Omega} \quad , \quad v \in V$$

for both problems. Moreover, ∂I_K denotes the subdifferential of the indicator function I_K of the constraint set K (cf., e.g., [23]).
We define the active control sets $\mathcal{A}^{(\nu)}(u), 1 \le \nu \le 2$, as the maximal open sets $A^{(\nu)} \subset \Omega$ resp. $A^{(\nu)} \subset \Gamma_1$ such that $u(x) = \psi^{(\nu)}(x)$ a.e. 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 $B \subset \Gamma_1$ such that $\psi^{(1)}(x) + \varepsilon \le u(x) \le \psi^{(2)}(x) - \varepsilon$ f.a.a. x. Then, the inclusion (2.3d) can be equivalently stated by means of the complementarity conditions:

$$\sigma(x) \;\begin{cases} \le 0 \quad , \quad \text{f.a.a. } x \in \mathcal{A}^{(1)}(u) \;, \\ = 0 \quad , \quad \text{f.a.a. } x \in \mathcal{I}(u) \;, \\ \ge 0 \quad , \quad \text{f.a.a. } x \in \mathcal{A}^{(2)}(u) \end{cases} \;. \tag{2.4}$$

In particular, we may split the co-control σ according to

$$\sigma \;=\; \sigma^{(1)} \;+\; \sigma^{(2)} \;, \tag{2.5}$$

where $\sigma^{(1)} \in L_-^2(\Omega)$ resp. $\sigma^{(1)} \in L_-^2(\Gamma_1)$ with $\mathrm{supp}(\sigma^{(1)}) = \mathrm{cl}(\mathcal{A}^{(1)}(u))$ and $\sigma^{(2)} \in L_+^2(\Omega)$ resp. $\sigma^{(2)} \in L_+^2(\Gamma_1)$ with $\mathrm{supp}(\sigma^{(2)}) = \mathrm{cl}(\mathcal{A}^{(2)}(u))$. It follows from (2.4) that $\sigma^{(\nu)}, 1 \le \nu \le 2$, satisfy the complementarity conditions

$$(\sigma^{(1)}, u - \psi^{(1)})_{0,\Lambda} = 0 \quad , \quad (\sigma^{(2)}, \psi^{(2)} - u)_{0,\Lambda} = 0 \;, \tag{2.6}$$

where $\Lambda = \Omega$ for the distributed control problem and $\Lambda = \Gamma_1$ in case of the boundary control problem.
We assume that $\{\mathcal{T}_\ell\}_{\ell \in \mathbb{N}_0}$ is a family of shape-regular simplicial triangulations of Ω such that in case of the boundary control problem the subsets $\Gamma_\nu \subset \Gamma, 1 \le \nu \le 2$, inherit geometrically

conforming triangulations $\mathcal{T}_\ell(\Gamma_\nu)$. For $D \subseteq \Omega$, we denote by $P_k(D), k \in \mathbb{N}_0$, the linear space of polynomials of degree k on D and we refer to $\mathcal{N}_\ell(D), \mathcal{E}_\ell(D)$ and $\mathcal{T}_\ell(D)$, $D \subseteq \overline{\Omega}$, as the sets of vertices, edges and elements of $\mathcal{T}_\ell$ in D. We set $h := \max\{h_T | T \in \mathcal{T}_\ell\}$, where h_T denotes the diameter of an element and we further denote by h_E the length of an edge $E \in \mathcal{E}_\ell(\overline{\Omega})$. For an interior edge $E \in \mathcal{E}_\ell(\Omega)$ with $E = T_1 \cap T_2, T_\nu \in \mathcal{T}_\ell, 1 \leq \nu \leq 2$, we refer to $\omega_E = T_1 \cup T_2$ as the patch formed by the union of the two elements sharing E as a common edge.
We refer to $V_\ell \subset V$ as the finite element space of continuous, piecewise linear finite elements with respect to $\mathcal{T}_\ell$

$$V_\ell \ := \ \{v_\ell \in C(\overline{\Omega}) \mid v_\ell|_T \in P_1(T) \ , \ v_\ell|_{\partial T \cap \Gamma^*} = 0 \ , \ T \in \mathcal{T}_\ell\} \ ,$$

where $\Gamma^* = \Gamma$ for the distributed control problem and $\Gamma^* = \Gamma_2$ for the boundary control problem. We further define $W_\ell \subset L^2(\Omega)$ resp. $W_\ell \subset L^2(\Gamma_1)$ as the linear space of piecewise constants with respect to $\mathcal{T}_\ell$ resp. $\mathcal{E}_\ell(\Gamma_1)$

$$\begin{aligned} W_\ell &:= \{w_\ell \in L^2(\Omega) \mid w_\ell|_T \in P_0(T) \ , \ T \in \mathcal{T}_\ell\} \quad \text{resp.} \\ W_\ell &:= \{w_\ell \in L^2(\Gamma_1) \mid w_\ell|_E \in P_0(E) \ , \ E \in \mathcal{E}_\ell(\Gamma_1)\} \ . \end{aligned}$$

We define $f_\ell, y_\ell^d \in W_\ell$ by $f_\ell|_T := f_T, y_\ell^d|_T := y_T^d \in P_0(T), T \in \mathcal{T}_\ell$, where f_T is given by the integral mean with respect to T, i.e., $f_T := |T|^{-1} \int_T f \, dx$, and $y_T^d \in P_0(T)$ is given analogously. We further define $u_\ell^d \in W_\ell$ in case of distributed control by $u_\ell^d|_T = u_T^d, T \in \mathcal{T}_\ell$, and by $u_\ell^d|_E = u_E^d, E \in \mathcal{E}_\ell(\Gamma_1)$, in the boundary control case, where $u_E^d \in P_0(E)$ is the integral mean of u^d with respect to E, i.e., $u_E^d := h_E^{-1} \int_E u^d \, ds$.
The lower bound $\psi^{(1)}$ and the upper bound $\psi^{(2)}$ for the boundary controls are approximated by $\psi_\ell^{(\nu)} \in W_\ell, 1 \leq \nu \leq 2$ such that $\psi_T^{(1)} := \psi_\ell^{(1)}|_T < \psi_\ell^{(2)}|_T =: \psi_T^{(2)}, T \in \mathcal{T}_\ell$, for distributed control and $\psi_E^{(1)} := \psi_\ell^{(1)}|_E < \psi_\ell^{(2)}|_E =: \psi_E^{(2)}, E \in \mathcal{E}_\ell(\Gamma_1)$, in case of boundary control.
Then, the finite element approximations of the control problems (2.1a)-(2.1c) resp. (2.2a)-(2.2c) amount to the computation of $(y_\ell, u_\ell) \in V_\ell \times K_\ell$ as the solution of the constrained finite dimensional minimization problem

$$\text{minimize } J_\ell(y_\ell, u_\ell) := \frac{1}{2} \ \|y_\ell - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u_\ell - u^d\|_{0,\Lambda}^2 \ , \tag{2.7a}$$

$$\text{over } (y_\ell, u_\ell) \in V_\ell \times K_\ell \ , \tag{2.7b}$$

$$\text{subject to } a(y_\ell, v_\ell) = \ell_{\ell,1}(u_\ell, v_\ell) \ . \ v_\ell \in V_\ell \ . \tag{2.7c}$$

Here, $\Lambda := \Omega$ for distributed control and $\Lambda := \Gamma_1$ for boundary control, whereas K_ℓ denotes the discrete constraint set for the controls

$$K_\ell \ := \ \{ \ w_\ell \in W_\ell \mid \psi_\ell^{(1)} \leq w_\ell \leq \psi_\ell^{(2)} \ \} \ . \tag{2.8}$$

Moreover, the functional $\ell_{\ell,1} : W_\ell \times V_\ell \to \mathbb{R}$ is given by

$$\begin{aligned} \ell_{\ell,1}(u_\ell, v_\ell) &:= (u_\ell, v_\ell)_{0,\Omega} \quad , \quad v_\ell \in V_\ell, \quad \text{resp.} \\ \ell_{\ell,1}(u_\ell, v_\ell) &:= (u_\ell, v_\ell)_{0,\Gamma_1} \quad , \quad v_\ell \in V_\ell \ . \end{aligned}$$

The optimality conditions for (2.7a)-(2.7c) involve the existence of a discrete co-state $p_\ell \in V_\ell$ and of a discrete co-control $\sigma_\ell \in W_\ell$ such that

$$a(y_\ell, v_\ell) \ = \ (f, v_\ell)_{0,\Omega} + \ell_{\ell,1}(u_\ell, v_\ell) \ , \ v_\ell \in V_\ell \ , \tag{2.9a}$$

$$a(p_\ell, v_\ell) \ = \ (y^d, v_\ell)_{0,\Omega} - \ell_{\ell,2}(y_\ell, v_\ell) \ , \ v_\ell \in V_\ell \ , \tag{2.9b}$$

$$u_\ell \ = \ u_\ell^d \ + \ \frac{1}{\alpha} \ (M_\ell p_\ell - \sigma_\ell) \ , \tag{2.9c}$$

$$\sigma_\ell \in \partial I_{K_\ell}(u_\ell) \ . \tag{2.9d}$$

Here, $\ell_{\ell,2} : V_\ell \times V_\ell \to \mathbb{R}$ stands for the functional

$$\ell_{\ell,2}(y_\ell, v_\ell) \;:=\; \int_\Omega y_\ell v_\ell \; dx \quad , \quad v_\ell \in V_\ell$$

and $M_\ell : V_\ell \to W_\ell$ is the averaging operator given by

$$\begin{aligned} M_\ell v_\ell|_T \;&:=\; |T|^{-1} \int_T v_\ell \; dx \;,\; T \in \mathcal{T}_\ell \quad \text{resp.} \\ M_\ell v_\ell|_E \;&:=\; h_E^{-1} \int_E v_\ell \; ds \;,\; E \in \mathcal{E}_\ell(\Gamma_1) \;. \end{aligned} \tag{2.10}$$

We define $\mathcal{A}^{(\nu)}(u_\ell), 1 \le \nu \le 2$, and $\mathcal{I}(u_\ell)$ as the discrete active and inactive control sets according to

$$\mathcal{A}^{(1)}(u_\ell) \;:=\; \bigcup \{\; T \in \mathcal{T}_\ell \mid u_\ell|_T = \psi_T^{(1)} \;\} \;, \tag{2.11a}$$

$$\mathcal{A}^{(2)}(u_\ell) \;:=\; \bigcup \{\; T \in \mathcal{T}_\ell \mid u_\ell|_T = \psi_T^{(2)} \;\} \;, \tag{2.11b}$$

$$\mathcal{I}(u_\ell) \;:=\; \bigcup \{\; T \in \mathcal{T}_\ell \mid \psi_T^{(1)} < u_\ell|_T < \psi_T^{(2)} \;\} \tag{2.11c}$$

in case of distributed control and analogously for boundary control.
As in the continuous regime, we split the adjoint control σ_h by means of

$$\sigma_\ell \;=\; \sigma_\ell^{(1)} \;+\; \sigma_\ell^{(2)} \;, \tag{2.12}$$

with $\mathrm{supp}(\sigma_\ell^{(1)}) = \mathcal{A}^{(1)}(u_\ell)$ and $\mathrm{supp}(\sigma_\ell^{(2)}) = \mathcal{A}^{(2)}(u_\ell)$ such that

$$\sigma_\ell^{(1)}|_D \le 0 \;,\; D \in \mathcal{A}^{(1)}(u_\ell) \quad , \quad \sigma_\ell^{(2)}|_D \ge 0 \;,\; D \in \mathcal{A}^{(2)}(u_\ell)) \;, \tag{2.13}$$

where $D = T$ for distributed control and $D = E$ in case of boundary control.
The inclusion (2.9d) implies that $\sigma_\ell^{(\nu)}, 1 \le \nu \le 2$, satisfy the complementarity conditions

$$(\sigma_\ell^{(1)}, u_\ell - \psi_\ell^{(1)})_{0,\Lambda} = 0 \quad , \quad (\sigma_\ell^{(2)}, \psi_\ell^{(2)} - u_\ell)_{0,\Lambda} = 0 \;, \tag{2.14}$$

where $\Lambda := \Omega$ for distributed control and $\Lambda := \Gamma_1$ in case of boundary control.

3 Prima-dual active set strategies

For the computation of the solution of (2.9a)-(2.9d) we use the primal-dual active set strategies as described in [8] and [20], respectively. These methods are based on an alternative representation of the complementarity system (2.14) and

$$\sigma_\ell^{(1)} \ge 0, \quad u_h \ge \psi_\ell^{(1)}, \tag{3.1a}$$

$$\sigma_\ell^{(2)} \ge 0, \quad u_h \le \psi_\ell^{(2)}. \tag{3.1b}$$

In fact, a simple calculation shows that (2.14) and (3.1) are equivalent to

$$\sigma_\ell = \min(0, \sigma_\ell + c(u_\ell - \psi_\ell^{(1)})) + \max(0, \sigma_\ell + c(u_\ell - \psi_\ell^{(2)})), \tag{3.2}$$

with some arbitrarily fixed real $c > 0$. From nonsmooth analysis (see, e.g., [23]) it is known that the max- and min-operations in (3.2) are generalized differentiable in the sense

of Clarke. Particular generalized derivatives of $(u_\ell, \sigma_\ell) \mapsto \min(0, \sigma_\ell + c(u_\ell - \psi_\ell^{(1)}))$ and $(u_\ell, \sigma_\ell) \mapsto \max(0, \sigma_\ell + c(u_\ell - \psi_\ell^{(2)}))$ are given by

$$G_\ell^{\min}(u_\ell, \sigma_\ell)(x) = \begin{cases} (c,1)^\top & \text{if } (\sigma_\ell + c(u_\ell - \psi_\ell^{(1)}))(x) < 0, \\ (0,0)^\top & \text{else,} \end{cases} \tag{3.3}$$

and

$$G_h^{\max}(u_\ell, \sigma_\ell)(x) = \begin{cases} (c,1)^\top & \text{if } (\sigma_\ell + c(u_\ell - \psi_\ell^{(2)}))(x) > 0, \\ (0,0)^\top & \text{else,} \end{cases} \tag{3.4}$$

respectively.

The primal-dual active set strategy is obtained by performing a generalized Newton step with respect to (2.9a)–(2.9c) and (3.2). This amounts to setting the linearization of (2.9a)–(2.9c) and (3.2) at a current estimate (or iterate in an algorithmic process) $(y_\ell^k, u_\ell^k, p_\ell^k, \sigma_\ell^k)$ of $(y_\ell, u_\ell, p_\ell, \sigma_\ell)$ in direction $(\delta y_\ell, \delta u_\ell, \delta p_\ell, \delta \sigma_\ell)$ equal to zero. This gives the following system

$$a(\delta y_\ell, v_\ell) - \ell_{\ell,1}(\delta u_\ell, v_\ell) = (f, v_\ell)_{0,\Omega} - a(y_\ell^k, v_\ell) + \ell_{\ell,1}(u_\ell^k, v_\ell) \quad \forall v_\ell \in V_\ell, \tag{3.5a}$$

$$a(\delta p_\ell, v_\ell) + \ell_{\ell,2}(\delta y_\ell, v_\ell) = (y^d, v_\ell)_{0,\Omega} - \ell_{\ell,2}(y_\ell^k, v_\ell) \quad \forall v_\ell \in V_\ell, \tag{3.5b}$$

$$\delta u_\ell - \frac{1}{\alpha} M_\ell \delta p_\ell + \delta \sigma_\ell = u_\ell^d - u_\ell^k + \frac{1}{\alpha}(M_\ell p_\ell^k - \sigma_\ell^k) \tag{3.5c}$$

$$\begin{aligned} &\delta \sigma_\ell - G_\ell^{\min}(u_\ell^k, \sigma_\ell^k)(\delta u_\ell, \delta \sigma_\ell) - G_\ell^{\max}(u_\ell^k, \sigma_\ell^k)(\delta u_\ell, \delta \sigma_\ell) = \\ &\quad = -\sigma_\ell^k + \min(0, \sigma_\ell^k + c(u_\ell^k - \psi_\ell^{(1)})) + \max(0, \sigma_\ell^k + c(u_\ell^k - \psi_\ell^{(2)})). \end{aligned} \tag{3.5d}$$

Let us next closer analyze (3.5d). For this purpose we define the sets

$$\mathcal{A}_\ell^{(1),k} := \bigcup \{T \in \mathcal{T}_\ell \mid \sigma_\ell^k|_T + c(u_\ell^k|_T - \psi_T^{(1)}) < 0\}, \tag{3.6a}$$

$$\mathcal{A}_\ell^{(2),k} := \bigcup \{T \in \mathcal{T}_\ell \mid \sigma_\ell^k|_T + c(u_\ell^k|_T - \psi_T^{(2)}) > 0\}, \tag{3.6b}$$

$$\mathcal{I}_\ell^k := \mathcal{T}_h \setminus (\mathcal{A}_\ell^{(1),k} \cup \mathcal{A}_\ell^{(2),k}), \tag{3.6c}$$

and analogously for the boundary control case. Based on the current iterate $(y_\ell^k, u_\ell^k, p_\ell^k, \sigma_\ell^k)$, these sets provide approximations of the true active and inactive sets, respectively, defined in (2.11). From (3.5d) we infer

$$(u_\ell^k + \delta u_\ell)|_T = \psi_T^{(1)} \quad \forall T \in \mathcal{A}_\ell^{(1),k}, \tag{3.7a}$$

$$(u_\ell^k + \delta u_\ell)|_T = \psi_T^{(2)} \quad \forall T \in \mathcal{A}_\ell^{(2),k}, \tag{3.7b}$$

$$(\sigma_\ell^k + \delta \sigma_\ell)|_T = 0 \quad \forall T \in \mathcal{I}_\ell^k. \tag{3.7c}$$

These observations imply the following algorithmic procedure.

Algorithm 1. (Primal-dual active set strategy)

(i) Choose $(y_\ell^0, u_\ell^0, p_\ell^0, \sigma_\ell^0)$, $c > 0$, and set $k := 0$.

(ii) Unless some stopping rule is satisfied, determine the active set estimates $\mathcal{A}_\ell^{(\nu),k}$, $1 \le \nu \le 2$, and the inactive set estimate $\mathcal{I}_\ell^k$ by (3.6).

(iii) Solve for $(\delta y_\ell, \delta u_\ell, \delta p_\ell, \delta\sigma_\ell)$*:*

$$
\begin{aligned}
&a(\delta y_\ell, v_\ell) - \ell_{\ell,1}(\delta u_\ell, v_\ell) = (f, v_\ell)_{0,\Omega} - a(y_\ell^k, v_\ell) + \ell_{\ell,1}(u_\ell^k, v_\ell) \quad \forall v_\ell \in V_\ell,\\
&a(\delta p_\ell, v_\ell) + \ell_{\ell,2}(\delta y_\ell, v_\ell) = (y^d, v_\ell)_{0,\Omega} - \ell_{\ell,2}(y_\ell^k, v_\ell) \quad \forall v_\ell \in V_\ell,\\
&\delta u_\ell - \frac{1}{\alpha} M_\ell \delta p_\ell + \delta\sigma_\ell = u_\ell^d - u_\ell^k + \frac{1}{\alpha}(M_\ell p_\ell^k - \sigma_\ell^k)\\
&(u_\ell^k + \delta u_\ell)|_T = \psi_T^{(1)} \quad \forall T \in \mathcal{A}_\ell^{(1),k},\\
&(u_\ell^k + \delta u_\ell)|_T = \psi_T^{(2)} \quad \forall T \in \mathcal{A}_\ell^{(2),k},\\
&(\sigma_\ell^k + \delta\sigma_\ell)|_T = 0 \quad \forall T \in \mathcal{I}_\ell^k.
\end{aligned}
$$

(iv) Set $(y_\ell^{k+1}, u_\ell^{k+1}, p_\ell^{k+1}, \sigma_\ell^{k+1}) := (y_\ell^k, u_\ell^k, p_\ell^k, \sigma_\ell^k) + (\delta y_\ell, \delta u_\ell, \delta p_\ell, \delta\sigma_\ell)$*,* $k := k+1$ *and return to (ii).*

A few comments on the algorithm are in order. The primal-dual active set strategy is rather stable with respect to its initialization. A practical choice is given by

$$
\begin{aligned}
&u_\ell^0 = \psi_\ell^{(1)} \quad (\text{or } u_\ell^0 = \psi_\ell^{(2)}),\\
&a(y_\ell^0, v_\ell) - \ell_{\ell,1}(u_\ell^0, v_\ell) = (f, v_\ell)_{0,\Omega} \quad \forall v_\ell \in V_\ell,\\
&a(p_\ell^0, v_\ell) + \ell_{\ell,2}(y_\ell^0, v_\ell) = (y^d, v_\ell)_{0,\Omega} \quad \forall v_\ell \in V_\ell,\\
&\sigma_\ell^0 = \alpha(u_\ell^d - u_\ell^0) + M_\ell p_\ell^0.
\end{aligned}
$$

But other choices are possible. Concerning the selection of c we argue below that the best choice is given by $c = \alpha > 0$. It is related to the function space specific convergence analysis of the algorithm. As far as the stopping rule is concerned, it can be shown [8, 20] that if

$$
\mathcal{A}_\ell^{(1),k} = \mathcal{A}_\ell^{(1),k-1} \quad \text{and} \quad \mathcal{A}_\ell^{(2),k} = \mathcal{A}_\ell^{(2),k-1} \quad \text{for } k \geq 1, \tag{3.8}
$$

then $(y_\ell^k, u_\ell^k, p_\ell^k, \sigma_\ell^k) = (y_\ell, u_\ell, p_\ell, \sigma_\ell)$, the solution of the discrete optimal control problem. This suggests to use (3.8) as our stopping criterion. Alternatively, assuming that the primal and the adjoint equations are satisfied to sufficient accuracy, then one may use

$$
\|\sigma_\ell^k - \min(0, \sigma_\ell^k + \alpha(u_\ell^k - \psi_\ell^{(1)})) - \max(0, \sigma_\ell^k + \alpha(u_\ell^k - \psi_\ell^{(2)}))\|_{0,\Omega} \leq \epsilon_c
$$

instead of (3.8). Here, $\epsilon_c > 0$ represents a user-specified stopping tolerance which needs to be aligned with stopping tolerances for iterative solvers of the discrete state and adjoint equations. Also notice that due to the settings of δu_h on $\mathcal{A}_\ell^{(1),k} \cup \mathcal{A}_\ell^{(2),k}$ the system in step (iii) can be reduced to a system on $\mathcal{I}_\ell^k$, only. The multiplier update $\delta\sigma_\ell$ can be easily obtained from the third equation in (iii) upon solving the linear system on $\mathcal{I}_\ell^k$. Especially in the distributed control case this may result in a significant reduction of the computational effort per iteration.

Finally let us spend a few words on the (local) convergence theory for the primal-dual active set strategy. As we can see from the derivation of the method, it is tightly related to generalized Newton methods. Moreover, it is known that the max- and min-operations involved in (3.5d) are semismooth, i.e., they satisfy

$$
\begin{aligned}
\|\min(0, \sigma_\ell + \delta\sigma_\ell + \alpha(u_\ell + \delta u_\ell - \psi_\ell^{(1)})) - \min(0, \sigma_\ell + \alpha(u_\ell - \psi_\ell^{(1)}))\\
- G_\ell^{\min}(u_\ell + \delta u_\ell, \sigma_\ell + \delta\sigma_\ell)(\delta u_\ell, \delta\sigma_\ell)\| = o(\|(\delta u_\ell, \delta\sigma_\ell)\|)
\end{aligned} \tag{3.9}
$$

as $\|(\delta u_\ell, \delta\sigma_\ell)\| \to 0$ and analogously for $\max(0, \sigma_\ell + \alpha(u_\ell - \psi_\ell^{(2)}))$. As a consequence the primal-dual active set strategy, which is—as we have just argued—equivalent to a semi-smooth Newton method, converges locally at a q-superlinear rate. For a proof and further

details in the context of constrained optimal control problems we refer to [22]. For general problems in finite dimensional space see [15, 27].

The function space counterpart of Algorithm 1 uses the active and inactive set prediction according to

$$\begin{aligned} \mathcal{A}^{(1),k} &:= \{x \in \Omega \mid (\sigma^k + \alpha(u^k - \psi^{(1)})) < 0\}, && \text{(3.10a)} \\ \mathcal{A}^{(2),k} &:= \{x \in \Omega \mid (\sigma^k + \alpha(u^k - \psi^{(2)})) > 0\}, && \text{(3.10b)} \\ \mathcal{I}^k &:= \Omega \setminus (\mathcal{A}^{(1),k} \cup \mathcal{A}^{(2),k}). && \text{(3.10c)} \end{aligned}$$

The derivation of these settings is similar to the one in finite dimensions. In fact, first observe that the function space analogue of (3.5c) yields

$$\sigma^k = \alpha(u^d - u^k) + p^k,$$

where p^k solves the adjoint equation. Using this expression in $\sigma^k + \alpha(u^k - \psi^{(\nu)})$, $1 \le \nu \le 2$, we obtain

$$\sigma^k = \alpha(u^d - u^k) + p^k = \alpha(u^d - \psi^{(\nu)}) + p(u^k),$$

where we emphasize that the adjoint state $p^k = p(u^k)$ depends on u^k via $y^k = y(u^k)$, which solves the state equation. Due to our choice $c = \alpha$, the above expression only consists of (fixed) data terms and the adjoint state, which, considered as a function of the control, satisfies $p : L^2(\Omega) \to H_0^1(\Omega)$. Thus, it is a smoothing operation. Further note that the reformulation of the complementarity system (i.e., the function space counterpart of (3.5d)) is considered in $L^2(\Omega)$. In [22] it was observed that the smoothing property of the control-to-adjoint-state mapping and the particular choice of c are prerequisites for the min- and max-operations to be generalized differentiable as mappings from $L^2(\Omega)$ to $L^2(\Omega)$ and the satisfaction of the analogue of (3.9) in $L^2(\Omega)$, respectively. As a consequence, using the derivatives analogous to (3.3) and (3.4) the function space version of Algorithm 1, with (3.10) in step (ii), again is equivalent to a semismooth Newton method and, hence, converges locally at a q-superlinear rate. For proof details we refer to [22].

Finally we mention that in [20] a global convergence result for Algorithm 1 is established. The techniques used in this work for bilaterally constrained boundary control problems also apply to distributed control. The key aspect is the fact that, regardless of the initial choice, a modified augmented Lagrangian merit function is shown to be strictly decreasing along the iteration sequence produced by the primal-dual active set strategy. This yields the global convergence of the method.

4 A posteriori error analysis

This section is devoted to the steps 'ESTIMATE', 'MARK', and 'REFINE' of the adaptive loop. Furthermore, it contains a convergence analysis of the AFEMs including convergence of the states, co-states, controls, and co-controls and - under some non-degeneracy assumptions on the continuous and discrete problems - a guaranteed reduction of the weighted discretization errors as given by

$$|||z - z_\ell|||_\kappa^2 := \kappa_1 \|y - y_\ell\|_{1,\Omega}^2 + \kappa_2 \|p - p_\ell\|_{1,\Omega}^2 + \kappa_3 \|u - u_\ell\|_{0,\Omega}^2 + \kappa_4 \|\sigma - \sigma_\ell\|_{0,\Omega}^2 \qquad (4.1)$$

for distributed control problems and by

$$|||z - z_\ell|||_\kappa^2 := \kappa_1 \|y - y_\ell\|_{1,\Omega}^2 + \kappa_2 \|p - p_\ell\|_{1,\Omega}^2 + \kappa_3 \|u - u_\ell\|_{0,\Gamma_1}^2 + \kappa_4 \|\sigma - \sigma_\ell\|_{0,\Gamma_2}^2 \qquad (4.2)$$

in case of boundary control, where $\kappa_i > 0, 1 \le i \le 4$, in (4.1) (resp. (4.2)) are weighting factors that only depend on the data of the problems, the shape regularity of the triangulations, and on some universal constants in the selection of edges and elements for refinement in the step 'MARK' of the adaptive loop. We write $|||\cdot|||_1$ for $\kappa_i = 1, 1 \le i \le 4$.

We use residual type a posteriori error estimators consisting of edge and element residuals. For distributed control problems, the estimator is of the form

$$\eta \;:=\; \Big(\eta_y^2 \;+\; \eta_p^2\Big)^{1/2} \;, \tag{4.3}$$

$$\eta_y \;:=\; \Big(\sum_{T\in\mathcal{T}_\ell} \eta_{y,T}^2 \;+\; \sum_{E\in\mathcal{E}_\ell} \eta_{y,E}^2\Big)^{1/2} \;, \tag{4.4}$$

$$\eta_p \;:=\; \Big(\sum_{T\in\mathcal{T}_\ell} \sum_{i=1}^{2} (\eta_{p,T}^{(i)})^2 \;+\; \sum_{E\in\mathcal{E}_\ell} \eta_{p,E}^2\Big)^{1/2} \;, \tag{4.5}$$

where for $T \in \mathcal{T}_\ell$ the element residuals $\eta_{y,T}$ and $\eta_{p,T}^{(i)}, 1 \le i \le 2$, are given by

$$\eta_{y,T} \;:=\; h_T \;\|f + u_h - cy_\ell\|_{0,T} \;, \tag{4.6}$$

$$\eta_{p,T}^{(1)} \;:=\; h_T \;\|y^d - y_\ell - cp_\ell\|_{0,T} \quad , \quad \eta_{p,T}^{(2)} \;:=\; \|M_\ell p_\ell - p_\ell\|_{0,T} \;. \tag{4.7}$$

On the other hand, for an interior edge $E = T_1 \cap T_2, T_\nu \in \mathcal{T}_\ell, 1 \le \nu \le 2$, and $\boldsymbol{\nu}_E$ denoting the exterior unit normal vector on E directed towards T_2 the edge residuals $\eta_{y,E}, \eta_{p,E}$ are as follows

$$\eta_{y,E} \;:=\; h_E^{1/2} \;\|\boldsymbol{\nu}_E \cdot [\nabla y_\ell]\|_{0,E} \quad , \quad \eta_{p,E} \;:=\; h_E^{1/2} \;\|\boldsymbol{\nu}_E \cdot [\nabla p_\ell]\|_{0,E} \;, \tag{4.8}$$

where $[\nabla y_\ell]$ and $[\nabla p_\ell]$ denote the jumps of $\nabla y_\ell, \nabla p_\ell$ across E.
The error analysis further has to take into account oscillations

$$osc_\ell \;:=\; \big(\mu_\ell^2(u^d) \;+\; \sum_{\nu=1}^{2} \mu_\ell^2(\psi^{(\nu)}) \;+\; osc_\ell^2(y^d) \;+\; osc_\ell^2(f)\big)^{1/2} \tag{4.9}$$

in the data $u^d, \psi^{(\nu)}, 1 \le \nu \le 2$, and y^d, f, where

$$\mu_\ell(u^d) := \Big(\sum_{T\in\mathcal{T}_\ell} \mu_T(u^d)^2\Big)^{1/2} \;,\; \mu_T(u^d) := \|u^d - u_T^d\|_{0,T} \;, \tag{4.10a}$$

$$\mu_\ell(\psi^{(\nu)}) := \Big(\sum_{T\in\mathcal{T}_\ell} \mu_T(\psi^{(\nu)})^2\Big)^{1/2} \;,\; \mu_T(\psi^{(\nu)}) := \|\psi^{(\nu)} - \psi_T^{(\nu)}\|_{0,T} \;, \tag{4.10b}$$

$$osc_\ell(y^d) := \Big(\sum_{T\in\mathcal{T}_\ell} osc_T(y^d)^2\Big)^{1/2} \;,\; osc_T(y^d) := h_T\|y^d - y_T^d\|_{0,T} \;, \tag{4.10c}$$

$$osc_\ell(f) := \Big(\sum_{T\in\mathcal{T}_\ell} osc_T(f)^2\Big)^{1/2} \;,\; osc_T(f) := h_T\|f - f_T\|_{0,T} \;. \tag{4.10d}$$

The residual type error estimator η for the finite element approximation of the boundary control problem (2.2a)-(2.2c) is defined according to

$$\eta \;:=\; \Big(\eta_y^2 \;+\; \eta_p^2\Big)^{1/2} \;, \tag{4.11a}$$

$$\eta_y \;:=\; \Big(\sum_{T\in\mathcal{T}_\ell} \eta_{y,T}^2 + \sum_{E\in\mathcal{E}_\ell(\Omega)} (\eta_{y,E}^{int})^2 + \sum_{E\in\mathcal{E}_\ell(\Gamma_1)} (\eta_{y,E}^{\Gamma_1})^2\Big)^{1/2} \;, \tag{4.11b}$$

$$\eta_p \;:=\; \Big(\sum_{T\in\mathcal{T}_\ell} \eta_{p,T}^2 + \sum_{E\in\mathcal{E}_\ell(\Omega)} (\eta_{p,E}^{int})^2 + \sum_{E\in\mathcal{E}_\ell(\Gamma_1)} \sum_{\nu=1}^{2} (\eta_{p,E}^{\Gamma_1,\nu})^2\Big)^{1/2} \;, \tag{4.11c}$$

where the element residuals $\eta_{y,T}, \eta_{p,T}, T \in \mathcal{T}_\ell$, and the edge residuals $\eta_{y,E}^{int}, \eta_{p,E}^{int}$ associated with the interior edges $E \in \mathcal{E}_h(\Omega)$ as well as the edge residuals $\eta_{y,E}^{\Gamma_1}, \eta_{p,E}^{\Gamma_1,\nu}, 1 \leq \nu \leq 2$, with respect to the boundary edges $E \in \mathcal{E}_\ell(\Gamma_1)$ are given as follows

$$\eta_{y,T} := h_T \, \|f - cy_\ell\|_{0,T} \quad , \quad T \in \mathcal{T}_\ell \, , \tag{4.12a}$$

$$\eta_{p,T} := h_T \, \|y^d - y_\ell - cp_\ell\|_{0,T} \quad , \quad T \in \mathcal{T}_\ell \, , \tag{4.12b}$$

$$\eta_{y,E}^{int} := h_E^{1/2} \, \|\boldsymbol{\nu}_E \cdot [\nabla y_\ell]\|_{0,E} \quad , \quad E \in \mathcal{E}_\ell(\Omega) \, , \tag{4.12c}$$

$$\eta_{p,E}^{int} := h_E^{1/2} \, \|\boldsymbol{\nu}_E \cdot [\nabla p_\ell]\|_{0,E} \quad , \quad E \in \mathcal{E}_\ell(\Omega) \, , \tag{4.12d}$$

$$\eta_{y,E}^{\Gamma_1} := h_E^{1/2} \, \|u_\ell - \boldsymbol{\nu}_E \cdot \nabla y_\ell\|_{0,E} \quad , \quad E \in \mathcal{E}_\ell(\Gamma_1) \, , \tag{4.12e}$$

$$\eta_{p,E}^{\Gamma_1,1} := h_E^{1/2} \, \|\boldsymbol{\nu}_E \cdot \nabla p_\ell\|_{0,E} \quad , \quad E \in \mathcal{E}_\ell(\Gamma_1) \, , \tag{4.12f}$$

$$\eta_{p,E}^{\Gamma_1,2} := \|M_\ell p_\ell - p_\ell\|_{0,E} \quad , \quad E \in \mathcal{E}_\ell(\Gamma_1) \, . \tag{4.12g}$$

The data oscillations read as follows

$$\mu_\ell(u^d) := \Big(\sum_{E \in \mathcal{E}_\ell(\Gamma_1)} \mu_E(u^d)^2 \Big)^{1/2} \, , \quad \mu_E(u^d) := \|u^d - u_E^d\|_{0,E} \, , \tag{4.13a}$$

$$\mu_\ell(\psi^{(\nu)}) := \Big(\sum_{E \in \mathcal{E}_\ell(\Gamma_1)} \sum_{\nu=1}^{2} \mu_E(\psi^{(\nu)})^2 \Big)^{1/2} \, , \quad \mu_E(\psi^{(\nu)}) := \|\psi^{(\nu)} - \psi_E^{(\nu)}\|_{0,E} \, , \tag{4.13b}$$

$$osc_\ell(y^d) := \Big(\sum_{T \in \mathcal{T}_\ell} osc_T(y^d)^2 \Big)^{1/2} \, , \quad osc_T(y^d) := h_T \, \|y^d - y_T^d\|_{0,T} \, , \tag{4.13c}$$

$$osc_\ell(f) := \Big(\sum_{T \in \mathcal{T}_\ell} osc_T(f)^2 \Big)^{1/2} \, , \quad osc_T(f) := h_T \, \|f - f_T\|_{0,T} \, . \tag{4.13d}$$

In the step 'MARK' of the adaptive loop, the selection of edges and elements is realized by means of bulk criteria originally suggested in [13]. In particular, given universal constants $\Theta_i, 1 \leq i \leq 2$ with $0 < \Theta_i < 1$, in case of distributed control problems we select a set of edges $\mathcal{M}_\ell^E \subset \mathcal{E}_\ell(\Omega)$ and a set of elements $\mathcal{M}_\ell^T \subset \mathcal{T}_\ell$ such that

$$\Theta_1 \sum_{E \in \mathcal{E}_\ell(\Omega)} (\eta_{y,E}^2 + \eta_{p,E}^2) \leq \sum_{E \in \mathcal{M}_\ell^E} (\eta_{y,E}^2 + \eta_{p,E}^2) \, , \tag{4.14}$$

$$\Theta_2 \Big(\sum_{T \in \mathcal{T}_\ell} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2) \Big) \leq \sum_{T \in \mathcal{M}_\ell^T} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2) \, . \tag{4.15}$$

Obvious modifications apply for boundary control problems.
In practice, we also include the data oscillations in the bulk criteria which are realized by a greedy algorithm (cf., e.g., [24]). For the implementation of the refinement process in the step 'REFINE' of the adaptive loop, we proceed as follows: If an element T or at least one of its edges E has been selected in the bulk criterion, we refine T in such a way that T contains at least one interior nodal point. Bisection is used in case of newly created nodes at midpoint of edges not contained in $\mathcal{M}_\ell^E$ in order to provide a geometrically conforming refined triangulation.

The a posteriori error analysis will be performed exemplarily for distributed control. The case of boundary control can be treated analogously. Basic tools are the reliability of the estimator and its discrete local efficiency. The reliability states that up to oscillations of the lower and upper bounds $\psi^{(\nu)}, 1 \leq \nu \leq 2$, for the control the total error $|||z - z_\ell|||_1$ can be bounded from above by the residual-type error estimator η.

Lemma 1. Let (y, p, u, σ) and $(y_\ell, p_\ell, u_\ell, \sigma_\ell)$ be the solutions of (2.3a)-(2.3d), and let η and $\mu_\ell(u^d), \mu_\ell(\psi^\nu), 1 \le \nu \le 2$, be the residual error estimator and the data oscillations as given by (4.3) and (4.10a),(4.10b), respectively. Then, there exists a constant $C_1 > 1$, depending only on α and on the shape regularity of the triangulations, such that

$$|||z - z_\ell|||_1^2 \le C_1 \left(\eta^2 + \mu_\ell^2(u^d) + \mu_\ell^2(\psi^{(1)}) + \mu_\ell^2(\psi^{(2)})\right) . \tag{4.16}$$

Proof. In case of an upper obstacle for the control, the proof has been established in [21]. It can be easily generalized to bilateral constraints. □

The discrete local efficiency of the estimator means that, again up to data oscillations, for refined elements T and edges E the local components of the estimator can be bounded from above by the norms of the differences of the fine and coarse mesh approximations on T and the patches ω_E, respectively. Then, forming the sum over all refined edges and elements, we end up with:

Lemma 2. Let (y, p, u, σ) and $(y_k, p_k, u_k, \sigma_k), k \in \{\ell, \ell+1\}$, be the solutions of (2.3a)-(2.3d) and (2.9a)-(2.9d), and let η and $osc_\ell(y^d), osc_\ell(f)$ be the residual error estimator and the data oscillations as given by (4.3) and (4.10c),(4.10d), respectively. Then, there exists a constant $C_2 > 1$, depending only on the constants $\Theta_i, 1 \le i \le 2$, in the bulk criteria (4.14),(4.15), and on the shape regularity of the triangulations, such that

$$\begin{aligned} &\sum_{T \in \mathcal{M}_\ell^T} \left(\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2\right) + \sum_{E \in \mathcal{M}_\ell^E} \left(\eta_{y,E}^2 + \eta_{p,E}^2\right) \le \\ &\le C_2 \left(\|y_\ell - y_{\ell+1}\|_{1,\Omega}^2 + \|p_\ell - p_{\ell+1}\|_{1,\Omega}^2 + \alpha^2 \|u_\ell - u_{\ell+1}\|_{0,\Omega}^2 + \right. \\ &\qquad \left. + \|\sigma_\ell - \sigma_{\ell+1}\|_{0,\Omega}^2 + osc_\ell^2(y^d) + osc_\ell^2(f)\right) . \end{aligned} \tag{4.17}$$

Proof. The idea of proof is to replace the element resp. edge bubble functions as used in the proof of the efficiency of the estimators [21, 24] by suitably scaled nodal basis functions associated with interior nodal points in $\mathcal{N}_{\ell+1}(T)$ resp. $\mathcal{N}_{\ell+1}(E)$ that are admissible test functions of the fine mesh finite element approximation (2.9a)-(2.9d). For details, we refer to [17]. □

Theorem 3. *Let (y, p, u, σ) and $(y_\ell, p_\ell, u_\ell, \sigma_\ell)$ be the solutions of* (2.3a)-(2.3d) *and* (2.9a)-(2.9d) *and let $osc_k, k \in \{\ell, \ell+1\}$ be given by* (4.9)*. Assume further that for some constant $0 \le \xi < 1$*

$$osc_{\ell+1}^2 \le \xi \, osc_\ell^2 . \tag{4.18}$$

Then, there holds

$$|||z - z_\ell|||_1 \to 0 \quad \text{for } \ell \to \infty . \tag{4.19}$$

Proof. Using the reliability of the estimator (Lemma 1), the discrete local efficiency (Lemma 2) and a perturbed Galerin orthogonality (cf. [17]; Lemma 5.1 and Lemma 5.2), it can be shown that there exists a constant $\Lambda > 0$, depending on the data of the problem, the constants $\Theta_i, 1 \le i \le 2$, in the bulk criteria (4.14),(4.15) and on the shape regularity of the triangulations, such that

$$\sum_{\ell=0}^{\infty} |||z - z_\ell|||^2 \le \Lambda , \tag{4.20}$$

which gives the assertion. □

As far as guaranteed error reductions are concerned, we observe an error reduction of the objective functional J in the sense that for some $0 \le \rho < 1$ and $C > 0$

$$\begin{pmatrix} \delta_{\ell+1} \\ osc_{\ell+1}^2 \end{pmatrix} \le \begin{pmatrix} \rho & C \\ 0 & \xi \end{pmatrix} \begin{pmatrix} \delta_\ell \\ osc_\ell^2 \end{pmatrix} , \tag{4.21}$$

where $\delta_k := J(y_k,u_k) - J(y,u), k \in \{\ell,\ell+1\}$. On the other hand, we can not expect a reduction of the discretization errors in the state, the co-state, the control, and the co-control, unless the continuous free boundary is sufficiently resolved by its discrete counterparts. To this end, we define the sets

$$\hat{\mathcal{A}}^{(\nu)}(u_\ell) := \mathrm{int}\big(\bigcup\{T \in \mathcal{T}_\ell \mid u_\ell|_{T'} = \psi_\ell^{(\nu)}|_{T'} \ , \ T' \in \mathcal{T}_\ell, T' \cap T \neq \emptyset\}\big) \ , \ 1 \leq \nu \leq 2, \tag{4.22}$$

$$\hat{\mathcal{I}}(u_\ell) := \mathrm{int}\big(\bigcup\{T \in \mathcal{T}_\ell \mid \psi_\ell^{(1)}|_T + \hat{\varepsilon} \leq u_\ell|_T \leq \psi_\ell^{(2)}|_T - \hat{\varepsilon}\}\big) \ , \tag{4.23}$$

$$\hat{\mathcal{F}}(u_\ell) := \Omega \setminus \big(\bigcup_{\nu=1}^{2} \hat{\mathcal{A}}^{(\nu)}(u_\ell) \cup \hat{\mathcal{I}}(u_\ell)\big) \tag{4.24}$$

for some $\hat{\varepsilon} > 0$ in (4.23) and enhance the resolution of the free boundary by the following extension of the bulk criteria (4.14)-(4.15):

(E) In the step 'MARK' of the adaptive loop, all edges $E \in \mathcal{E}_\ell(\hat{\mathcal{F}}(u_\ell))$ are marked for refinement.

We assume that there exists some $\ell^* \in \mathbb{N}$ such that for all $\ell \geq \ell^*$

$$\hat{\mathcal{I}}(u_\ell) \subseteq \mathcal{I}(u) \quad , \quad \hat{\mathcal{A}}^{(\nu)}(u_\ell) \subseteq \mathcal{A}(u) \ , \ 1 \leq \nu \leq 2 \ . \tag{4.25}$$

Theorem 4. *Let (y,p,u,σ) and $(y_\ell,p_\ell,u_\ell,\sigma_\ell)$ be the solutions of* (2.3a)-(2.3d) *and* (2.9a)-(2.9d) *and let $osc_k, k \in \{\ell,\ell+1\}$, be the data oscillations given by* (4.9)*. Assume that the extension* **(E)** *of the bulk criteria* (4.14),(4.15) *is implemented and that* (4.18) *and* (4.25) *hold true. Then, there exist constants $0 \leq \rho < 1$ and $C > 0$, depending on the data of the problem, the constants $\Theta_i, 1 \leq i \leq 2$, in the bulk criteria and on the shape regularity of the triangulations, such that*

$$\begin{pmatrix} |||z - z_{\ell+1}|||_\kappa^2 \\ osc_{\ell+1}^2 \end{pmatrix} \leq \begin{pmatrix} \rho_1 & C \\ 0 & \xi \end{pmatrix} \begin{pmatrix} |||z - z_\ell|||_\kappa^2 \\ osc_\ell^2 \end{pmatrix} .$$

Proof. The assumption (4.25) allows to find constants $0 < \kappa_i < 1, 1 \leq i \leq 4$, and $0 \leq \rho < 1, C > 0$ such that (cf. [17]; Theorem 6.4)

$$|||z - z_{\ell+1}|||_\kappa^2 \ \leq \ \rho_1 \ |||z - z_\ell|||_\kappa^2 \ + \ C \ osc_\ell^2 \ ,$$

which together with (4.18) gives the assertion. □

Remark. As has been shown in [17], the condition (4.25) is satisfied under the assumption of strict complementarity $\sigma|_{\mathcal{I}(u)}\, 0$ of the continuous problem (2.3a)-(2.3d) and the following non-degeneracy assumptions for the discrete problems (2.9a)-(2.9d): There exist $\varepsilon_i^* > 0, \gamma_i, 1 \leq i \leq 3$, such that for all $0 < \varepsilon < \varepsilon_i^*$ and for all sufficiently large $\ell \in \mathbb{N}$

$$\begin{aligned} &\mathrm{meas}(\{x \in \mathcal{I}(u_\ell) \mid 0 < \psi_\ell(x) - u_\ell(x) < \varepsilon^2\}) < \gamma_1 \ \varepsilon \ , \\ &\{x \in \mathcal{I}(u_\ell) \mid 0 < \psi_\ell(x) - u_\ell(x) < \varepsilon^2\} \subseteq \{x \in \mathcal{I}(u_\ell) \mid \mathrm{dist}(x,\mathcal{F}_\ell) < \gamma_2 \varepsilon\} \ , \\ &\{x \in \mathcal{I}(u_\ell) \mid \mathrm{dist}(x,\mathcal{F}_\ell) < \varepsilon\} \subseteq \{x \in \mathcal{I}(u_\ell) \mid 0 < \psi_\ell(x) - u_\ell(x) < \gamma_3 \varepsilon^2\} \ . \end{aligned}$$

5 Numerical results

The first two examples illustrate the performance of the adaptive finite element approximation for distributed optimal control problems. The first one focuses on the dependence on the regularization parameter α in the objective functional, whereas the second one features lack of strict complementarity.

Example 1: Distributed control (dependence on α)
The data in (2.1a)-(2.1c) have been chosen as follows:

$$\Omega = (0,1)^2 \ , \ u^d = f = 0 \ , \ \psi = 1 \ , \ \alpha \ = \ 10^{-k} \ , \ 1 \leq \ k \leq 5 \ ,$$
$$y^d := \begin{cases} 200x_1x_2(x_1 - 1/2)^2(1 - x_2) & , \quad 0 \leq x_1 \leq 1/2 \\ 200(x_1 - 1)x_2(x_1 - 1/2)^2(1 - x_2) & , \quad 1/2 < x_1 \leq 1 \end{cases} .$$

Figure 1 shows a visualization of the optimal state and the optimal control for various values of α. The flat region in the visualization of the control corresponds to the active set.

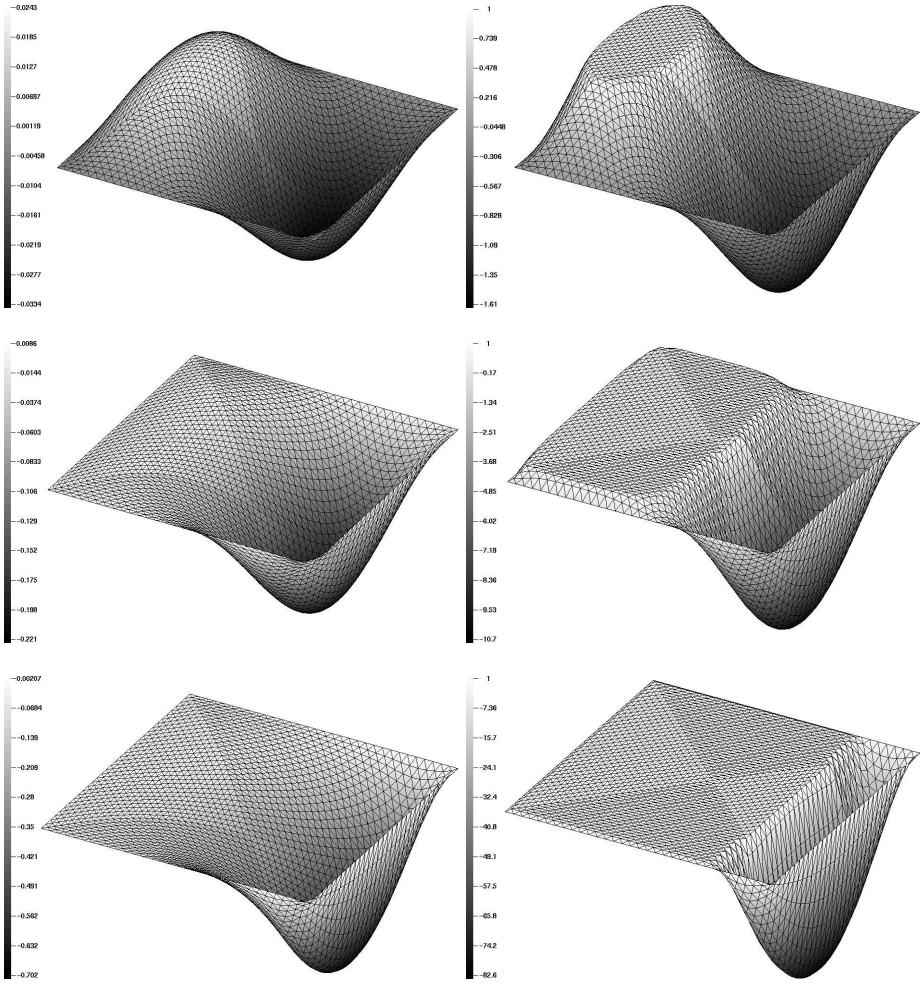

Figure 1: Visualization of the optimal state y and the optimal control u for $\alpha = 0.01$ (top), $\alpha = 0.001$ (middle) and $\alpha = 0.00001$ (bottom)

The initial simplicial triangulation $\mathcal{T}_0$ was chosen according to a subdivision of Ω by joining the four vertices resulting in one interior nodal point and four congruent triangles. Since u^d, f and ψ are constant, we have $osc_\ell(u^d) = osc_\ell(f) = osc_\ell(\psi) = 0, \ell \in \mathbb{N}_0$.

For various values of α, Figure 2 displays the adaptively generated triangulations after six refinement steps with $\Theta_i = 0.6, 1 \leq i \leq 2$, in the bulk criteria. In case $\alpha = 0.01$, the elliptically shaped area in the left part represents the active set. We observe that the active set is growing for decreasing α. The continuous free boundary between the active and inactive sets, displayed by a black curve, is well resolved by the adaptive refinement due to the extension **(E)** of the bulk criteria.

Table 1: Total discretization error, discretization errors in the state, co-state, control, and co-control (Example 1, $\alpha = 0.01$)

$\mathbf{l}$	$\mathbf{N_{dof}}$	$\|\|\|\mathbf{z}-\mathbf{z}_\ell\|\|\|_1$	$\|\mathbf{y}-\mathbf{y}_\ell\|_1$	$\|\mathbf{p}-\mathbf{p}_\ell\|_1$	$\|\mathbf{u}-\mathbf{u}_\ell\|_0$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_\ell\|_0$
1	**13**	**9.38e-02**	**5.69e-02**	**3.18e-02**	**4.17e-01**	**1.37e-03**
2	**41**	**5.37e-02**	**3.35e-02**	**1.76e-02**	**2.07e-01**	**6.63e-04**
3	**134**	**3.02e-02**	**1.89e-02**	**9.67e-03**	**1.30e-01**	**3.24e-04**
4	**319**	**2.24e-02**	**1.39e-02**	**7.47e-03**	**8.07e-02**	**1.98e-04**
5	**795**	**1.47e-02**	**9.14e-03**	**4.84e-03**	**5.92e-02**	**1.10e-04**
6	**1998**	**1.02e-02**	**6.35e-03**	**3.33e-03**	**3.87e-02**	**9.16e-05**
7	**4373**	**7.16e-03**	**4.46e-03**	**2.37e-03**	**2.69e-02**	**6.70e-05**
8	**10612**	**4.93e-03**	**3.08e-03**	**1.61e-03**	**1.83e-02**	**4.60e-05**
9	**23019**	**3.44e-03**	**2.14e-03**	**1.13e-04**	**1.32e-02**	**3.24e-05**

Table 2: Components of the error estimator and data oscillations (Example 1, $\alpha = 0.01$)

$\mathbf{l}$	$\mathbf{N_{dof}}$	$\boldsymbol{\eta}_{\mathbf{y},\mathbf{T},\ell}$	$\boldsymbol{\eta}_{\mathbf{p},\mathbf{T},\ell}$	$\boldsymbol{\eta}_{\mathbf{y},\mathbf{E},\ell}$	$\boldsymbol{\eta}_{\mathbf{p},\mathbf{E},\ell}$	$\mathbf{osc}_\ell(\mathbf{y^d})$
1	**13**	**2.54e-01**	**2.23e-01**	**1.56e-01**	**9.97e-02**	**9.76e-02**
2	**41**	**1.70e-01**	**1.10e-01**	**1.09e-01**	**6.50e-02**	**2.88e-02**
3	**134**	**1.03e-01**	**5.86e-02**	**6.63e-02**	**3.63e-02**	**1.03e-02**
4	**319**	**6.43e-02**	**3.83e-02**	**4.74e-02**	**2.63e-02**	**5.09e-03**
5	**795**	**4.18e-02**	**2.48e-02**	**3.25e-02**	**1.78e-02**	**2.21e-03**
6	**1998**	**2.80e-02**	**1.66e-02**	**2.30e-02**	**1.24e-02**	**1.02e-03**
7	**4373**	**1.90e-02**	**1.15e-02**	**1.64e-02**	**8.95e-03**	**5.01e-04**
8	**10612**	**1.28e-02**	**7.63e-03**	**1.15e-02**	**6.14e-03**	**2.53e-04**
9	**23019**	**8.75e-03**	**5.30e-03**	**8.35e-03**	**4.41e-03**	**1.30e-04**

Table 3: Percentages of elements/edges selected for refinement by the bulk criteria and its extension (Example 1, $\alpha = 0.01$)

$\mathbf{l}$	$\mathbf{N_{dof}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{T}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{E}}$	$\mathbf{M}_{\mathbf{osc},\mathbf{E}}$	$\mathbf{M}_{\mathbf{fb},\mathbf{E}}$
0	**5**	**50.0**	**75.0**	**75.0**	**0.0**
1	**13**	**25.0**	**20.0**	**43.8**	**0.0**
2	**41**	**23.4**	**20.5**	**29.7**	**21.9**
3	**134**	**18.8**	**20.6**	**10.3**	**13.2**
4	**319**	**17.5**	**13.2**	**8.7**	**10.4**
5	**795**	**16.0**	**13.6**	**6.6**	**8.2**
6	**1998**	**15.4**	**11.8**	**5.8**	**6.4**
7	**4373**	**16.3**	**13.0**	**5.0**	**5.8**
8	**10612**	**15.7**	**12.5**	**2.6**	**4.7**
9	**23019**	**15.2**	**11.8**	**1.8**	**4.4**

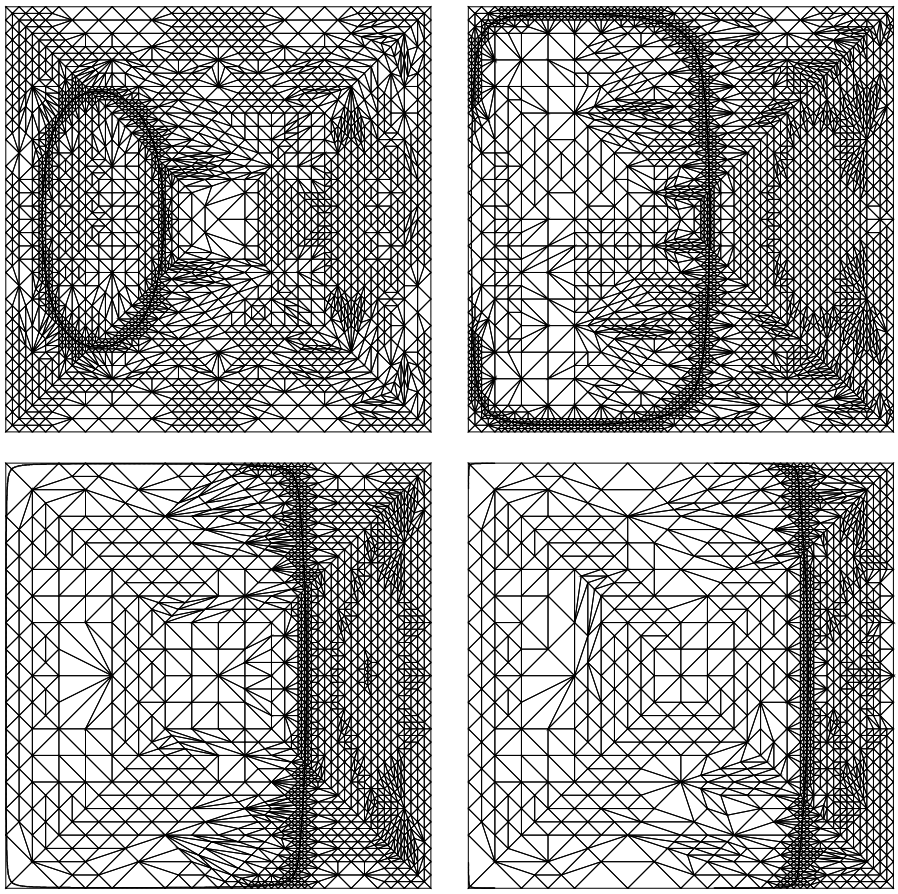

Figure 2: Adaptively generated grids after 6 refinement steps ($\alpha = 0.01$ (top left), $\alpha = 0.001$ (top right), $\alpha = 0.0001$ (bottom left) and $\alpha = 0.00001$ (bottom right))

More detailed information is given in Table 1 - Table 3. In particular, Table 1 displays the error reduction in the total error

$$|||z - z_\ell||| \;:=\; (|y - y_\ell|^2_{1,\Omega} + |p - p_\ell|^2_{1,\Omega} \|u - u_\ell\|^2_{0,\Omega} + \|\sigma - \sigma_\ell\|^2_{0,\Omega})^{1/2}$$

and the errors in the state, the co-state, the control, and the co-control, whereas the actual element and edge related components of the residual type a posteriori error estimator are given in Table 2. Table 3 contains the percentages of elements and edges that have been marked for refinement according to the bulk criteria and their extension **(E)**. Here, $M_{\eta,T}$ stands for the level l elements marked for refinement due to the element residuals and the data oscillations. On the other hand, $M_{\eta,E}, M_{osc,E}$and $M_{fb,E}$ refer to the edges marked for refinement with regard to the edge residuals, data oscillations and the extension **(E)** of the bulk criteria (resolution of the free boundary). On the coarsest grid, the sum of the percentages exceeds 100 %, since an edge may satisfy more than one criterion in the adaptive refinement process. The refinement is initially dominated by the resolution of the free boundary and the data oscillations, whereas at a later stage edge and element residuals dominate.

The second example is a unilaterally control constrained distributed control problem which features a lack of strict complementarity in a vicinity of the free boundary between the active and the inactive sets.

Example 2: Distributed control (lack of strict complementarity)
The data in (2.1a)-(2.1c) have been chosen as follows:

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

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

$$\begin{aligned} \hat{u} &:= \begin{cases} \psi^{(2)}(x_1, x_2) & , \quad (x_1, x_2) \in \Omega_1 \cup \Omega_2 \\ -1.01\ \psi^{(2)}(x_1, x_2) & , \quad \text{otherwise} \end{cases} \,, \\ \hat{\sigma} &:= \begin{cases} 2.25\ (x_1 - 0.75) \cdot 10^{-4} & , \quad (x_1, x_2) \in \Omega_2 \\ 0 & , \quad \text{otherwise} \end{cases} \end{aligned}$$

with Ω_1 and Ω_2 specified as follows

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

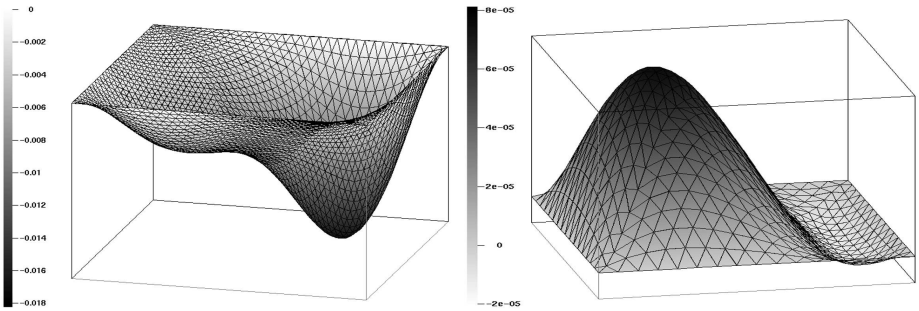

Figure 3: Example 2: Visualization of the optimal state y (left) and the optimal co-state p (right))

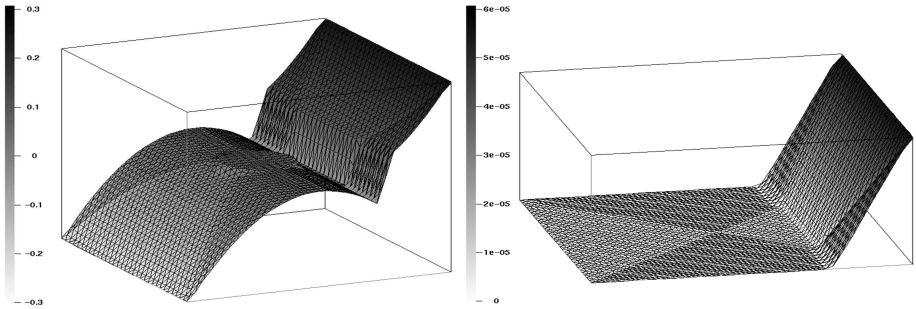

Figure 4: Example 2: Visualization of the optimal control u (left) and the optimal co-control σ (right))

Figures 3 and 4 display the optimal state y, the optimal co-state p, the optimal control $u = \hat{u}$, and the optimal co-control $\sigma = \hat{\sigma}$, respectively.

The initial simplicial triangulation $\mathcal{T}_{h_0}$ has been chosen according to a subdivision of Ω by joining the four vertices resulting in one interior nodal point and four congruent triangles.

The parameters Θ_i in the bulk criterion have been specified according to $\Theta_i = 0.7, 1 \leq i \leq 2$. Figure 5 shows the adaptively generated triangulations after six (left) and eight (right) refinement steps. We observe 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. Figure 3) are well resolved by the adaptive solution process.

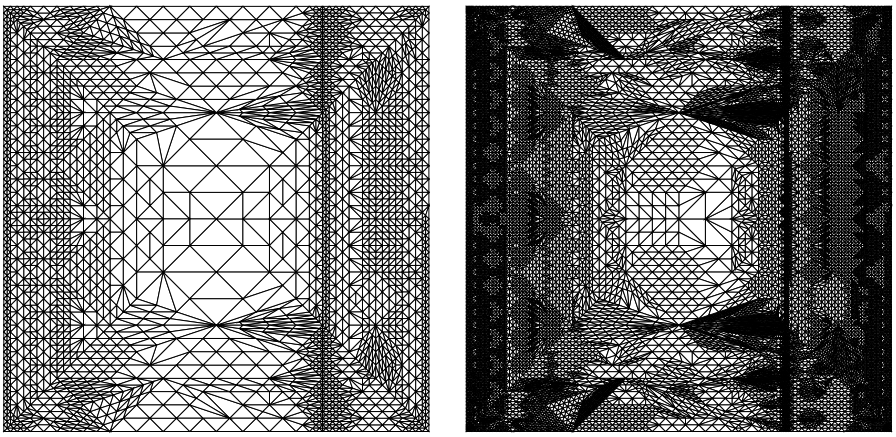

Figure 5: Example 2: Adaptively generated grid after 6 (left) and 8 (right) refinement steps, $\Theta_i = 0.7$

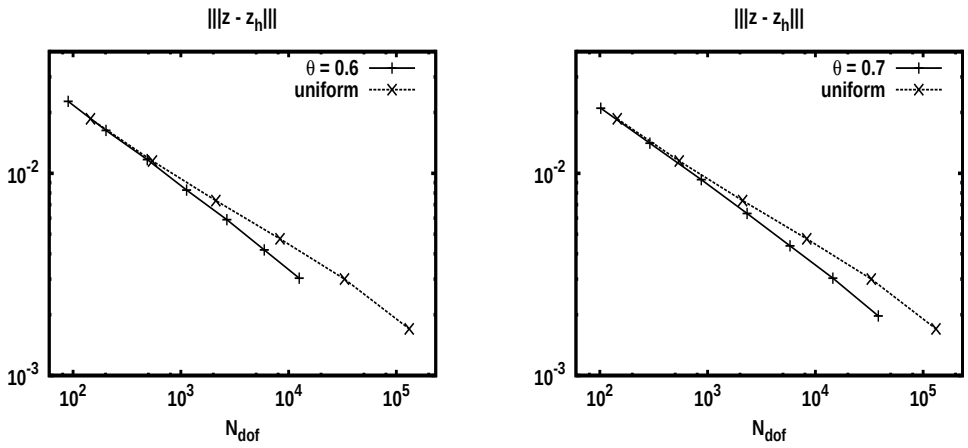

Figure 6: Example 2: Adaptive versus uniform refinement, $\Theta_i = 0.6$ (left) and $\Theta_i = 0.7$ (right)

Figure 6 displays the benefit of adaptive versus uniform refinement. In particular, the total discretization error in the state, co-state, control, and co-control is shown as a function of the total number of degrees of freedom.
More detailed information is given in Table 4 - Table 6. In particular, Table 4 displays the error reduction in the total error $|||z - z_h|||$ and the errors in the state, the co-state, the control, and the co-control, respectively. On the other hand, the actual components of the residual type a posteriori error estimator and data oscillations are given in Table 5. Finally, Table 6 lists 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 $M_{\mu,T}$ 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 $T \in \mathcal{T}_h$ may satisfy more than one criterion in the adaptive refinement process. We 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 4: Total error, errors in the state, co-state, control, and co-control (Example 2)

$\mathbf{l}$	$\mathbf{N_{dof}}$	$\|\|\|\mathbf{z}-\mathbf{z}_\ell\|\|\|_1$	$\|\mathbf{y}-\mathbf{y}_\ell\|_1$	$\|\mathbf{p}-\mathbf{p}_\ell\|_1$	$\|\mathbf{u}-\mathbf{u}_\ell\|_0$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_\ell\|_0$
1	**13**	**5.36e-02**	**6.85e-03**	**1.04e-04**	**4.66e-02**	**8.86e-06**
2	**41**	**3.12e-02**	**3.83e-03**	**5.99e-05**	**2.74e-02**	**4.63e-06**
3	**102**	**2.10e-02**	**2.39e-03**	**4.10e-05**	**1.85e-02**	**2.29e-06**
4	**291**	**1.41e-02**	**1.58e-03**	**2.94e-05**	**1.24e-02**	**1.39e-06**
5	**873**	**9.31e-03**	**9.73e-04**	**1.93e-05**	**8.32e-03**	**8.41e-07**
6	**2325**	**6.33e-03**	**6.17e-04**	**1.22e-05**	**5.70e-03**	**5.60e-07**
7	**5816**	**4.38e-03**	**4.02e-04**	**7.62e-06**	**3.97e-03**	**3.76e-07**
8	**14524**	**3.03e-03**	**2.66e-04**	**5.26e-06**	**2.76e-03**	**2.42e-07**
9	**38364**	**1.97e-03**	**1.71e-04**	**3.42e-06**	**1.80e-03**	**1.54e-07**

Table 5: Components of the error estimator and data oscillations (Example 2)

$\mathbf{l}$	$\mathbf{N_{dof}}$	$\boldsymbol{\eta_y}$	$\boldsymbol{\eta_p}$	$\boldsymbol{\mu_h(u^d)}$	$\boldsymbol{\mu_h(\psi)}$
1	**13**	**5.45e-02**	**5.76e-04**	**4.77e-02**	**3.93e-02**
2	**41**	**2.99e-02**	**3.30e-04**	**2.64e-02**	**2.06e-02**
3	**102**	**1.72e-02**	**2.71e-04**	**1.83e-02**	**1.34e-02**
4	**291**	**1.01e-02**	**1.80e-04**	**1.21e-02**	**8.62e-03**
5	**873**	**6.10e-03**	**1.21e-04**	**8.33e-03**	**5.49e-03**
6	**2325**	**3.93e-03**	**7.50e-05**	**5.74e-03**	**3.73e-03**
7	**5816**	**2.54e-03**	**4.84e-05**	**4.22e-03**	**2.34e-03**
8	**14524**	**1.65e-03**	**3.23e-05**	**3.08e-03**	**1.55e-03**
9	**38364**	**1.07e-03**	**2.17e-05**	**2.34e-03**	**1.02e-03**

Table 6: Percentages of elements/edges marked for refinement (Example 2)

$\mathbf{l}$	$\mathbf{N_{dof}}$	$\mathbf{M_{fb,T}}$	$\mathbf{M_{\eta,E}}$	$\mathbf{M_{\eta,T}}$	$\mathbf{M_{\mu,T}}$
1	**13**	**37.5**	**25.0**	**25.0**	**43.8**
2	**41**	**18.8**	**9.1**	**21.9**	**34.4**
3	**102**	**14.0**	**12.8**	**31.4**	**23.3**
4	**291**	**9.1**	**14.0**	**35.7**	**16.7**
5	**873**	**5.8**	**14.6**	**32.5**	**9.6**
6	**2325**	**4.3**	**12.8**	**28.7**	**7.4**
7	**5816**	**3.4**	**13.6**	**29.1**	**3.4**
8	**14524**	**2.7**	**16.1**	**32.0**	**1.5**
9	**38364**	**2.0**	**14.6**	**28.6**	**0.9**

Example 3: Boundary control problem (constant constraints)

The data in (2.2a)-(2.2c) have been chosen as follows:

$$\begin{aligned} \Omega &= (0,1)^2 \quad , \quad \Gamma_1 = (0,1)\times\{0\} \quad , \quad c = 1 \, , \\ y^d &= \sin(2\pi x_1)\sin(2\pi x_2)\exp(2x_1) \quad , \quad u^d = \cos(5\pi x_1^2) \quad , \quad \alpha = 10^{-5} \, , \\ f &:= 0 \quad , \quad \psi^{(1)} = -0.75 \quad , \quad \psi^{(2)} = 0.75 \, . \end{aligned}$$

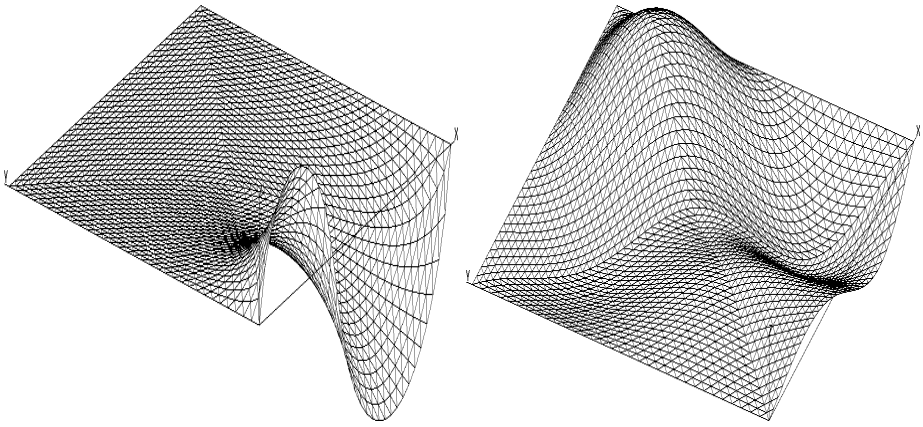

Figure 7: Example 3: Visualization of the optimal state y (left) and of the optimal co-state p (right))

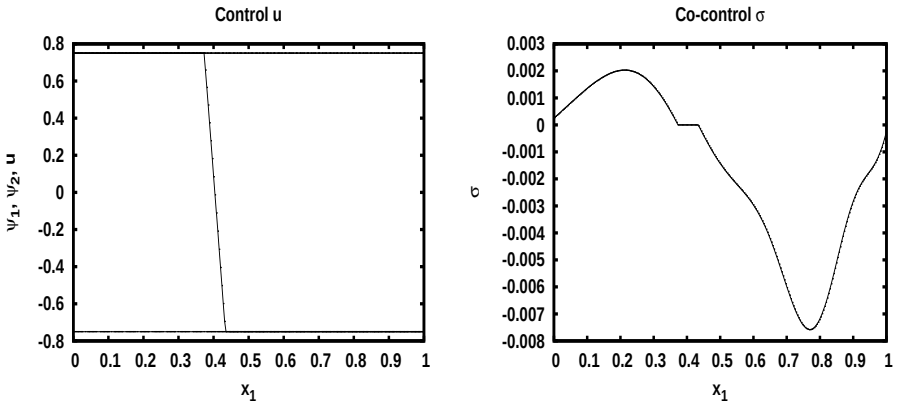

Figure 8: Example 3: Visualization of the optimal control u (left) and of the optimal co-control σ (right)). The lower and upper bounds on the control are shown as 'dashed' and 'dotted' lines,respectively.

Figures 7 and 8 show a visualization of the optimal state, the optimal co-state, the optimal control, and the optimal co-control, respectively. The optimal control switches from the lower to the upper bound in a very narrow region which corresponds to the inactive set associated with the optimal control.
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 $f = 0$ and the lower and upper bounds $\psi^{(\nu)}, 1 \le \nu \le 2$, are constant, we have $\mu_h(\psi^{(\nu)}) = 0$

and $osc_h(f) = 0$. Figure 9 displays the adaptively generated triangulations after four (left) and eight (right) refinement steps with $\Theta_i = 0.8$ in the bulk criteria. It should be emphasized that we are working with only one grid for both the state and the co-state. Consequently, the grid reflects regions of substantial change in these variables.
Figure 10 (left) displays the actual size of the state and co-state related components of the error estimator as well as the relevant data oscillations as functions of the number of levels in the hierarchy of adaptively generated grids for $\Theta_i = 0.7$ (left) and $\Theta_i = 0.8$ (right). As can be expected, $\mu_h(u^d)$ is dominant due to the highly oscillatory character of u^d.

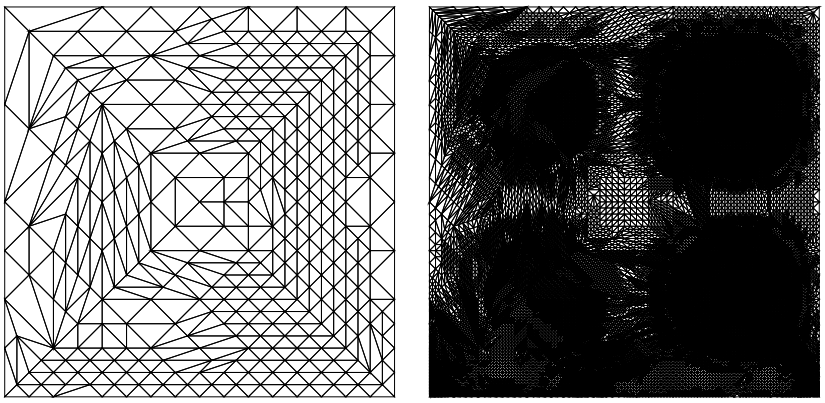

Figure 9: Example 3: Adaptively generated grid after 4 (left) and 8 (right) refinement steps ($\Theta_i = 0.8, 1 \leq i \leq 5$, in the bulk criteria)

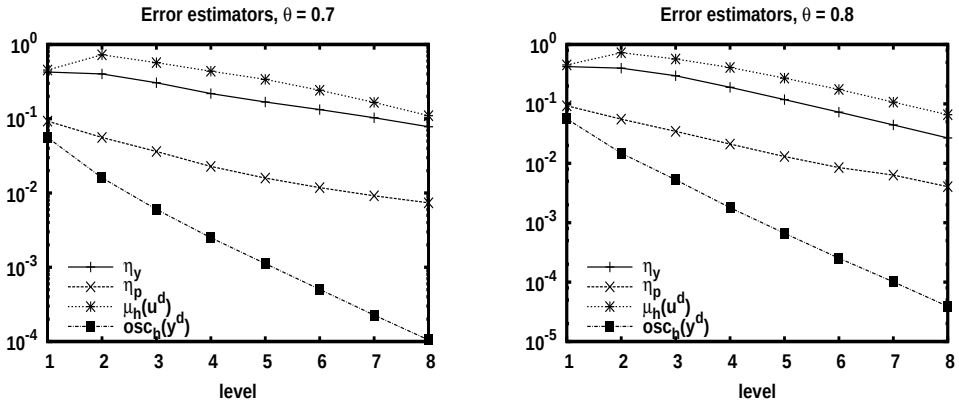

Figure 10: Example 3: Components of the error estimator and data oscillations for $\Theta_i = 0.8, 1 \leq i \leq 5$(left), norm of the total error as a function of the number of grid points for uniform and adaptive refinement (right)

Table 7 - Table 9 reflect the history of the refinement process. In particular, Table 7 displays the error reduction in the total error

$$|||z - z_h||| \ := \ (|y - y_h|^2_{1,\Omega} + |p - p_h|^2_{1,\Omega} + \|u - u_h\|^2_{0,\Gamma_1} + \|\sigma - \sigma_h\|^2_{0,\Gamma_1})^{1/2}$$

and the errors in the state, the co-state, the control, and the co-control, respectively. On the other hand, the values of the components of the residual type a posteriori error estimator and of the relevant data oscillations are given in Table 8. Table 9 lists the percentages of

elements and edges that have been marked for refinement according to the bulk criteria. Here, $M_{\eta,T}$ and $M_{osc,T}$ stand for the level l elements marked for refinement due to the element residuals and the data oscillations, whereas $M_{\eta,E}, M_{\eta,E,\Gamma_1}$ and M_{μ,E,Γ_1} refer to the edges marked for refinement with regard to the interior edge residuals. the edge residuals on Γ_1 and the data oscillations on Γ_1. We note that the sum of the percentages may exceed 100 %, since an element or edge may satisfy more than one of the bulk criteria in the adaptive refinement process. We see that the refinement is initially dominated by the data oscillation $\mu_h(u^d)$. whereas at a later stage the element residuals dominate.

Table 7: Total error, errors in the state, co-state, control, and co-control (Example 3)

l	$\mathbf{N_{dof}}$	$\|\|\|\mathbf{z}-\mathbf{z}_\ell\|\|\|_1$	$\|\mathbf{y}-\mathbf{y}_\ell\|_1$	$\|\mathbf{p}-\mathbf{p}_\ell\|_1$	$\|\mathbf{u}-\mathbf{u}_\ell\|_0$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_\ell\|_0$
0	**5**	**1.05e+00**	**2.80e-01**	**3.26e-02**	**7.29e-01**	**3.38e-03**
1	**13**	**8.21e-01**	**2.66e-01**	**1.71e-02**	**5.35e-01**	**3.14e-03**
2	**38**	**5.07e-01**	**1.58e-01**	**1.38e-02**	**3.33e-01**	**1.74e-03**
3	**106**	**2.79e-01**	**9.75e-02**	**8.81e-03**	**1.72e-01**	**8.45e-04**
4	**331**	**1.73e-01**	**5.59e-02**	**5.41e-03**	**1.11e-01**	**4.27e-04**
5	**1053**	**8.97e-02**	**3.14e-02**	**3.33e-03**	**5.46e-02**	**2.77e-04**
6	**3311**	**4.68e-02**	**1.83e-02**	**2.05e-03**	**2.64e-02**	**1.61e-04**
7	**9986**	**3.31e-02**	**1.07e-02**	**1.47e-03**	**2.09e-02**	**1.10e-04**
8	**29751**	**2.29e-02**	**5.86e-03**	**8.59e-04**	**1.61e-02**	**5.96e-05**

Table 8: Components of the error estimator and data oscillations (Example 3)

l	$\mathbf{N_{dof}}$	$\boldsymbol{\eta}_\mathbf{y}$	$\boldsymbol{\eta}_\mathbf{p}$	$\mu_\mathbf{h}(\mathbf{u^d})$	$\mathbf{osc_h}(\mathbf{y^d})$
0	**5**	**3.14e-02**	**1.44e-01**	**7.73e-01**	**1.42e-01**
1	**13**	**4.25e-01**	**9.28e-02**	**4.52e-01**	**5.60e-02**
2	**38**	**3.99e-01**	**5.54e-02**	**7.26e-01**	**1.48e-02**
3	**106**	**2.98e-01**	**3.45e-02**	**5.69e-01**	**5.28e-03**
4	**331**	**1.90e-01**	**2.11e-02**	**4.10e-01**	**1.79e-03**
5	**1053**	**1.18e-01**	**1.30e-02**	**2.71e-01**	**6.64e-04**
6	**3311**	**7.27e-02**	**8.43e-03**	**1.75e-01**	**2.53e-04**
7	**9986**	**4.43e-02**	**6.33e-03**	**1.07e-01**	**1.01e-04**
8	**29751**	**2.67e-02**	**4.05e-03**	**6.57e-02**	**3.90e-05**

Table 9: Percentages of elements/edges marked for refinement (Example 3)

l	$\mathbf{N_{dof}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{T}}$	$\mathbf{M_{osc,T}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{E}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{E},\boldsymbol{\Gamma}_1}$	$\mathbf{M}_{\boldsymbol{\mu},\mathbf{E},\boldsymbol{\Gamma}_1}$
0	**5**	**75.0**	**75.0**	**100.0**	**100.0**	**100.0**
1	**13**	**50.0**	**50.0**	**20.0**	**50.0**	**100.0**
2	**38**	**32.2**	**44.1**	**11.1**	**75.0**	**75.0**
3	**106**	**35.0**	**34.4**	**13.4**	**50.0**	**62.5**
4	**331**	**39.4**	**21.4**	**12.5**	**33.3**	**53.3**
5	**1053**	**44.7**	**17.6**	**10.7**	**36.0**	**40.0**
6	**3311**	**48.6**	**12.7**	**9.3**	**33.3**	**35.7**
7	**9986**	**49.6**	**8.2**	**9.6**	**23.2**	**33.3**
8	**29751**	**43.4**	**5.7**	**10.0**	**21.3**	**38.0**

Example 4: Boundary control problem (highly oscillating constraints)

In this example, the lower and upper bounds for the boundary control represent highly oscillatory functions on the control boundary. In particular, the data in (2.2a)-(2.2c) have

been chosen as follows:

$$\begin{aligned} \Omega &= (0,1)^2 \quad , \quad \Gamma_1 = (0,1)\times\{0\} \quad , \quad c = 1 \, , \\ y^d &= \begin{cases} 0 & , \ x_1 \le 0.5 \, , \\ 1 & , \ 0.5 < x_1 \le 0.75 \, , \\ -1 & , \ 0.75 < x_1 \end{cases} \quad , \quad u^d = 0 \quad , \quad \alpha = 10^{-3} \, , \\ f &:= 0 \quad , \quad \psi^{(1)} = \sin(9\pi x_1)| \quad , \quad \psi^{(2)} = 2 + \cos(\pi/2 + 8x_1) \, . \end{aligned}$$

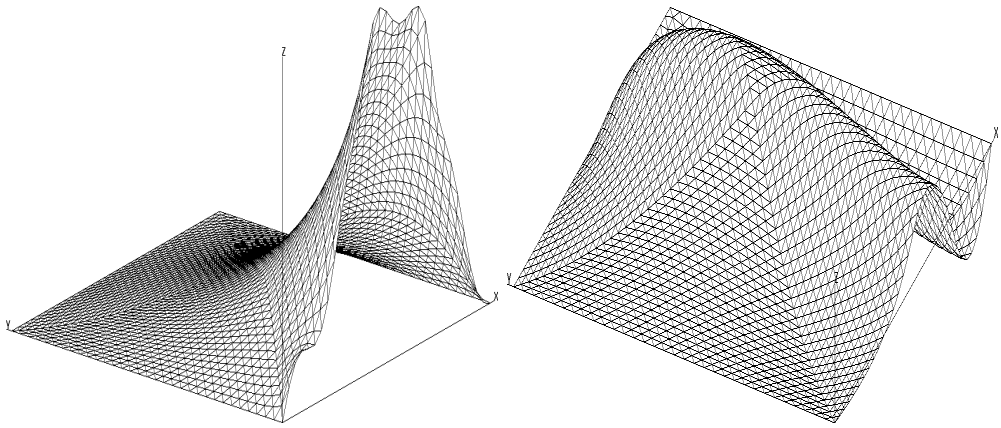

Figure 11: Example 4: Visualization of the optimal state y (left) and of the optimal co-state p (right))

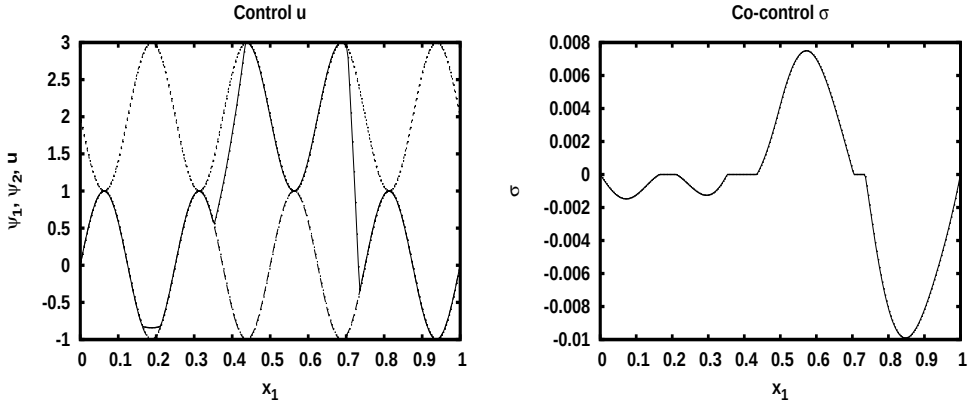

Figure 12: Example 4: Visualization of the optimal control u (left) and of the optimal co-control σ (right)). The lower and upper bounds on the control are shown as 'dashed' and 'dotted' lines,respectively.

Figures 11 and 12 display the optimal state y, the optimal co-state p, the optimal control u, and the optimal co-control σ, respectively. We observe an almost 'bang-bang' type of optimal control which switches from the lower to the upper bound and back again to the lower bound on Γ_1.
The initial simplicial triangulation $\mathcal{T}_{h_0}$ and the parameters Θ_i in the bulk criterion have been chosen as in Example 1. Figure 13 shows the adaptively generated triangulations after four (left) and eight (right) refinement steps. Figure 14 (left) shows the components of the error

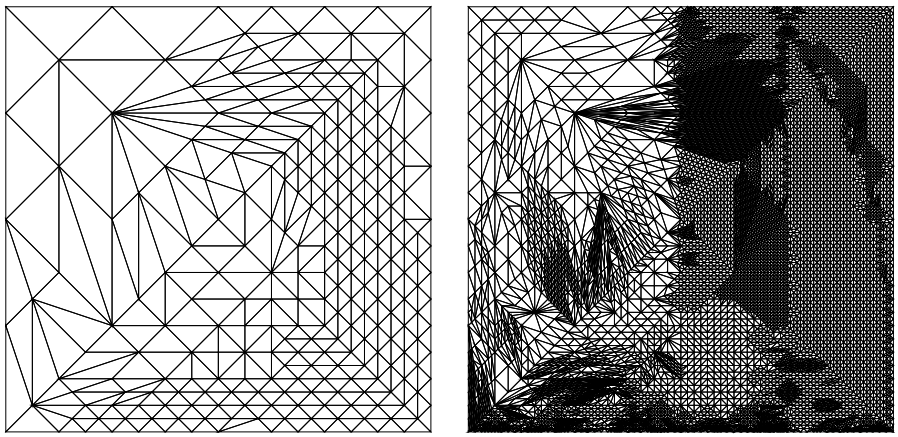

Figure 13: Example 4: Adaptively generated grid after 4 (left) and 8 (right) refinement steps ($\Theta_i = 0.7, 1 \leq i \leq 2$, in the bulk criteria)

estimator and the relevant data oscillations as functions of the number of levels of the grid hierarchy ($\Theta_i = 0.7, 1 \leq i \leq 2$, in the bulk criteria). Furthermore, Figure 14 (right) displays the total discretization error in the state, co-state, control, and co-control as a function of the total number of degrees of freedom. The benefits of adaptive versus uniform refinement set in, once the highly oscillatory bounds $\psi^{(\nu)}, 1 \leq \nu \leq 2$, have been sufficiently resolved. This is also reflected by the results in Table 10 - Table 12. The refinement process is at the very beginning clearly dominated by the edge residuals associated with $\psi^{(1)}$ and $\psi^{(2)}$.

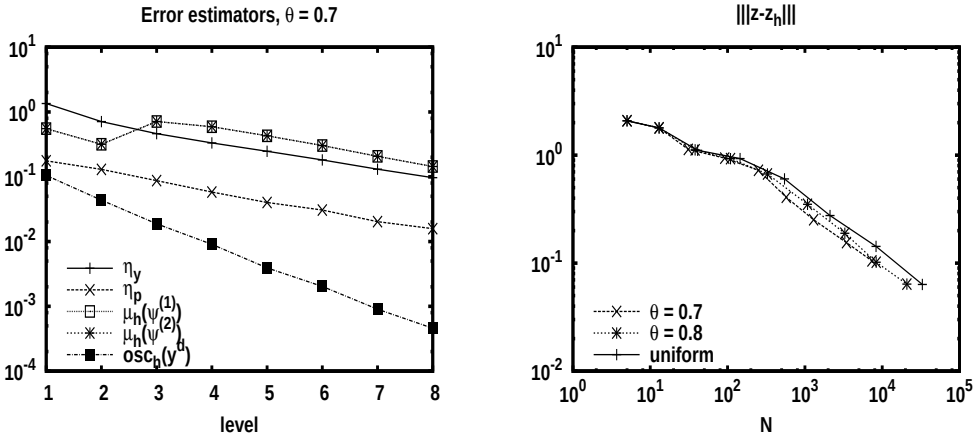

Figure 14: Example 4: Components of the error estimator and data oscillations for $\Theta_i = 0.8, 1 \leq i \leq 5$ (left), adaptive versus uniform refinement (right).

Table 10: Total error, errors in the state, co-state, control, and co-control (Example 4)

l	$\mathbf{N_{dof}}$	$\|\|\|\mathbf{z}-\mathbf{z}_\ell\|\|\|_1$	$\|\mathbf{y}-\mathbf{y}_\ell\|_1$	$\|\mathbf{p}-\mathbf{p}_\ell\|_1$	$\|\mathbf{u}-\mathbf{u}_\ell\|_0$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_\ell\|_0$
1	**13**	**1.79e+00**	**3.70e-01**	**5.34e-02**	**1.37e+00**	**4.47e-03**
2	**31**	**1.13e+00**	**2.75e-01**	**3.48e-02**	**8.12e-01**	**4.17e-03**
3	**92**	**9.39e-01**	**1.87e-01**	**2.28e-02**	**7.27e-01**	**2.10e-03**
4	**251**	**7.24e-01**	**1.37e-01**	**1.42e-02**	**5.71e-01**	**1.17e-03**
5	**574**	**4.07e-01**	**8.52e-02**	**9.28e-03**	**3.12e-01**	**7.32e-04**
6	**1296**	**2.52e-01**	**5.35e-02**	**7.31e-03**	**1.91e-01**	**5.75e-04**
7	**3452**	**1.55e-01**	**3.39e-02**	**4.77e-03**	**1.16e-01**	**3.23e-04**
8	**7372**	**1.05e-01**	**2.29e-02**	**3.44e-03**	**7.81e-02**	**2.14e-04**

Table 11: Components of the error estimator and data oscillations (Example 4)

l	$\mathbf{N_{dof}}$	$\boldsymbol{\eta}_\mathbf{y}$	$\boldsymbol{\eta}_\mathbf{p}$	$\boldsymbol{\mu}_\mathbf{h}(\boldsymbol{\psi}^{(1)})$	$\boldsymbol{\mu}_\mathbf{h}(\boldsymbol{\psi}^{(2)})$	$\mathbf{osc_h}(\mathbf{y^d})$
1	**13**	**1.35e+00**	**1.76e-01**	**5.53e-01**	**5.53e-01**	**1.05e-01**
2	**31**	**7.11e-01**	**1.30e-01**	**3.17e-01**	**3.17e-01**	**4.36e-02**
3	**92**	**4.63e-01**	**8.72e-02**	**7.13e-01**	**7.13e-01**	**1.88e-02**
4	**251**	**3.33e-01**	**5.80e-02**	**5.92e-01**	**5.92e-01**	**9.05e-03**
5	**574**	**2.47e-01**	**4.01e-02**	**4.28e-01**	**4.28e-01**	**3.92e-03**
6	**1296**	**1.82e-01**	**3.06e-02**	**3.03e-01**	**3.03e-01**	**2.03e-03**
7	**3452**	**1.31e-01**	**2.03e-02**	**2.07e-01**	**2.07e-01**	**9.04e-04**
8	**7372**	**9.70e-02**	**1.58e-02**	**1.44e-01**	**1.44e-01**	**4.58e-04**

Table 12: Percentages of elements/edges marked for refinement (Example 4)

l	$\mathbf{N_{dof}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{T}}$	$\mathbf{M}_{\mathbf{osc},\mathbf{T}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{E}}$	$\mathbf{M}_{\boldsymbol{\eta},\mathbf{E},\boldsymbol{\Gamma}_1}$	$\mathbf{M}_{\boldsymbol{\mu},\mathbf{E},\boldsymbol{\Gamma}_1}$
1	**13**	**37.5**	**18.8**	**15.0**	**100.0**	**100.0**
2	**31**	**42.6**	**14.9**	**18.8**	**50.0**	**75.0**
3	**92**	**37.8**	**5.8**	**12.7**	**37.5**	**75.0**
4	**251**	**24.4**	**1.7**	**10.0**	**35.7**	**64.3**
5	**574**	**24.9**	**1.6**	**9.5**	**24.0**	**28.0**
6	**1296**	**39.1**	**1.6**	**10.3**	**32.5**	**27.5**
7	**3452**	**24.8**	**0.8**	**8.1**	**25.8**	**18.2**
8	**7372**	**29.9**	**0.7**	**7.7**	**25.5**	**29.9**

References

[1] R.A. Adams (1975), *Sobolev spaces*, Pure and Applied Mathematics, Vol. 65, Academic Press, New York-London.

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

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

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

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

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

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

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

[9] C. Carstensen and S. Bartels (2002). *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**, 945–969.

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

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

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

[13] W. Dörfler (1996). *A convergent adaptive algorithm for Poisson's equation.* SIAM J. Numer. Anal., **33**, 3, 1106–1124.

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

[15] F. Facchinei and J.-S. Pang (2003), *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vol. II, Springer Series in Operations Research and Financial Engineering, Springer-Verlag, Berlin-New York.

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

[17] A. Gaevskaya, R.H.W. Hoppe, Y. Iliash, and M. Kieweg (2007) *Convergence analysis of an adaptive finite element method for distributed control problems with control constraints.* In Proc. Conf. Optimal Control for PDEs, Oberwolfach, Germany (G. Leugering et al.; eds.), pp. 47–68, Birkhäuser, Basel.

[18] A. Gaevskaya, R.H.W. Hoppe, Y. Iliash, and M. Kieweg (2006). *A posteriori error analysis of control constrained distributed and boundary control problems.* In: Proc. Conf. Scientific Computing, September 16/17, 2006, Moscow, Russia (W. Fitzgibbon et al.; eds.), pp. 85–108, Russian Academy of Sciences, Moscow.

[19] A. Gaevskaya, R.H.W. Hoppe, and S. Repin (2006). *A Posteriori Estimates for Cost Functionals of Optimal Control Problems.* In: Numerical Methods and Advanced Applications (A. Bermudez et al.; eds.), pp. 308–316, Spriner, Berlin-Heidelberg, New York.

[20] M. Hintermüller (2003). *A primal-dual active set algorithm for bilaterally control constrained optimal control problems.* Quarterly of Applied Mathematics, LXI: 131–161.

[21] M. Hintermüller, R.H.W. Hoppe, Y. Iliash, and M. Kieweg (2008). *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints.* ESAIM, Calculus of Variations and Optimal Control (in press).

[22] M. Hintermüller, K. Ito, and K. Kunisch (2003), *The primal-dual active set strategy as a semismooth Newton method.* SIAM J. Optim., **13**, 865–888.

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

[24] R.H.W. Hoppe, Y. Iliash, C. Iyyunni, and N. Sweilam (2006). *A posteriori error estimates for adaptive finite element discretizations of boundary control problems.* J. Numer. Anal. **14**, 57–82.

[25] R.H.W. Hoppe and B. Wohlmuth (1996). *Element-oriented and edge-oriented local error estimators for nonconforming finite element methods.* M^2AN, Modélisation Math. Anal. Numér. **30**, 237–263.

[26] R.H.W. Hoppe and B. Wohlmuth (1998). *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.), pp. 155-167, Marcel Dekker, New York.

[27] D. Klatte and B. Kummer (2002), *Nonsmooth Equations in Optimization: Regularity, Claculus, Methods and Applications*, Kluwer, Dordrecht-Boston-London.

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

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

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

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

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

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

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