

Convergence Analysis of an Adaptive Finite Element Method for Distributed Control Problems with Control Constraints

A. Gaevskaya, R.H.W. Hoppe, Y. Iliash and M. Kieweg

Abstract. We develop an adaptive finite element method for a class of distributed optimal control problems with control constraints. The method is based on a residual-type a posteriori error estimator and incorporates data oscillations. The analysis is carried out for conforming P1 approximations of the state and the co-state and elementwise constant approximations of the control and the co-control. We prove convergence of the error in the state, the co-state, the control, and the co-control. Under some additional non-degeneracy assumptions on the continuous and the discrete problems, we then show that an error reduction property holds true at least asymptotically. The analysis uses the reliability and the discrete local efficiency of the a posteriori estimator as well as quasi-orthogonality properties as essential tools. Numerical results illustrate the performance of the adaptive algorithm.

Mathematics Subject Classification (2000). Primary 65K10; Secondary 49M15.

Keywords. Distributed optimal control, control constraints, adaptive finite elements, residual-type a posteriori error estimators, convergence analysis.

1. Introduction

We present a convergence analysis of adaptive finite element approximations of a distributed optimal control problem with control constraints. In particular, assuming $\Omega \subset \mathbb{R}^2$ to be a bounded, polygonal domain with boundary $\Gamma := \partial\Omega$ and given data $y^d \in L^2(\Omega)$ and $f \in L^2(\Omega)$, $\psi \in H^1(\Omega) \cap L^\infty(\Omega)$ as well as a parameter $0 < \alpha \leq 1$, we consider the following distributed optimal control problems with

The second author has been partially supported by the NSF under Grant No. DMS-0411403 and Grant No. DMS-0511611. The fourth author acknowledges the support by the elite graduate school TopMath.

bound constrained controls

$$\text{minimize } J(y, u) := \frac{1}{2} \|y - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u\|_{0,\Omega}^2 \quad (1.1a)$$

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

$$\text{subject to } -\Delta y = f + u, \quad (1.1b)$$

$$u \in K := \{v \in L^2(\Omega) \mid v \leq \psi \text{ a.e. in } \Omega\}. \quad (1.1c)$$

It is well known (cf., e.g., [15, 20, 21]) that (1.1a)–(1.1c) admits a unique solution $(y, u) \in H_0^1(\Omega) \times L^2(\Omega)$. The optimality conditions involve the existence of a co-state $p \in H_0^1(\Omega)$ and a co-control $\sigma \in L_+^2(\Omega)$ such that y, p, u, σ satisfy

$$a(y, v) = (f + u, v)_{0,\Omega}, \quad v \in H_0^1(\Omega), \quad (1.2a)$$

$$a(p, v) = -(y - y^d, v)_{0,\Omega}, \quad v \in H_0^1(\Omega), \quad (1.2b)$$

$$u = \frac{1}{\alpha} (p - \sigma) \in K, \quad (1.2c)$$

$$(\sigma, u - v)_{0,\Omega} \geq 0, \quad v \in K. \quad (1.2d)$$

Here, $(\cdot, \cdot)_{0,\Omega}$ refers to the standard L^2 inner product and $a(\cdot, \cdot)$ stands for the bilinear form

$$a(w, z) := \int_{\Omega} \nabla w \cdot \nabla z \, dx, \quad w, z \in H_0^1(\Omega).$$

We note that the variational inequality (1.2d) can be equivalently stated as the complementarity condition

$$\begin{aligned} \sigma &\in L_+^2(\Omega), \quad \psi - u \in L_+^2(\Omega), \\ (\sigma, \psi - u)_{0,\Omega} &= 0. \end{aligned} \quad (1.3)$$

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 $x \in B$. Further, we refer to $\mathcal{F}(u) := \partial\mathcal{A}(u)$ as the free boundary between the active and inactive sets.

The control problem (1.1a)–(1.1c) will be approximated by Lagrangian type finite elements with respect to an adaptively generated hierarchy of simplicial triangulations of the computational domain. We note that adaptive finite element methods (AFEM) are efficient and reliable algorithmic tools in the numerical solution of partial differential equations. AFEMs typically consist of successive loops of the sequence

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \quad (1.4)$$

Here, SOLVE stands for the numerical solution of the finite element discretized problem, ESTIMATE requires the a posteriori estimation of the global discretization error in some appropriate norm or with respect to a goal oriented error functional. The step MARK is devoted to the selection of elements and edges for

refinement, and the final step REFINE takes care of the technical realization of the refinement process.

The development, analysis and implementation of efficient and reliable a posteriori error estimators has been the subject of intensive research in the past two decades and has actually reached some level of maturity (see, e.g., the monographs [1, 3, 4, 14, 25, 26] and the references therein). On the other hand, a rigorous convergence analysis of (1.4) relying on appropriate error reduction properties has so far only been done for conforming AFEMs [8, 13, 24] and, very recently, by Carstensen and the second author for mixed and nonconforming finite element methods in [10, 11] as well as for edge element methods for eddy current equations in [12].

As far as the a posteriori error analysis of adaptive finite element schemes for optimal control problems is concerned, the unconstrained case has been considered in [4, 6], whereas residual-type a posteriori error estimators in the control constrained case have been derived and analyzed in [17, 19, 21, 22]. No convergence analysis has been addressed so far.

This contribution aims to provide a convergence analysis of AFEM for (1.1a)–(1.1c). The paper is organized as follows:

In Section 2, we consider the finite element approximation of (1.1a)–(1.1c) and present details of the adaptive loop (1.4) focusing on a residual-type a posteriori error estimator in the step ESTIMATE and a bulk criterion for the selection of edges and elements for refinement in the step MARK. The reliability of the estimator and its discrete local efficiency are shown in Section 3. In Section 4, we prove convergence of the discrete states and co-states in $H_0^1(\Omega)$ and of the discrete controls and co-controls in $L^2(\Omega)$. Under the assumptions of strict complementarity and non-degeneracy, in Section 5 we show that an error reduction property holds true at least asymptotically. Finally, Section 6 illustrates the performance of the estimator by an illustrative numerical example.

2. Finite element discretization

For the finite element discretization of (1.1a)–(1.1c) we assume that \mathcal{T}_ℓ is a shape-regular simplicial triangulation of Ω . 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}_ℓ in $D \subseteq \overline{\Omega}$. We set $h_\ell := \max\{h_T | T \in \mathcal{T}_\ell\}$ where h_T stands for the diameter of an element $T \in \mathcal{T}_\ell$ and we denote by h_E the length of an edge $E \in \mathcal{E}_\ell$. Further, we refer to g_T as the integral mean of $g \in L^2(\Omega)$ on $T \in \mathcal{T}_\ell$, i.e., $g_T = |T|^{-1} \int_T g \, dx$. We denote by

$$V_\ell := \{ v_\ell \in C_0(\Omega) \mid v_\ell|_T \in P_1(T), T \in \mathcal{T}_\ell \},$$

the standard conforming P1 finite element space and by

$$W_\ell := \{ w_\ell \in L^2(\Omega) \mid w_\ell|_T \in P_0(T), T \in \mathcal{T}_\ell \}$$

the linear space of elementwise constant functions on Ω . We refer to $y_\ell \in V_\ell$ and $u_\ell \in W_\ell$ as finite element approximations of the state y and the control u , respectively. The upper obstacle ψ is approximated by the elementwise constant function $\psi_\ell \in W_\ell$ with $\psi_\ell|_T := \psi_T, T \in \mathcal{T}_\ell$.

Then, the finite element approximation of the distributed optimal control problem (1.1a)–(1.1c) reads as follows:

$$\text{minimize } J_\ell(y_\ell, u_\ell) := \frac{1}{2} \|y_\ell - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u_\ell\|_{0,\Omega}^2, \quad (2.1a)$$

$$\text{over } (y_\ell, u_\ell) \in V_\ell \times W_\ell,$$

$$\text{subject to } a(y_\ell, v_\ell) = (f + u_\ell, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell, \quad (2.1b)$$

$$u_\ell \in K_\ell := \{w_\ell \in W_\ell \mid w_\ell|_T \leq \psi_T, T \in \mathcal{T}_\ell\}. \quad (2.1c)$$

The optimality conditions for (2.1a)–(2.1c) again give rise to the existence of a co-state $p_\ell \in V_\ell$ and a co-control $\sigma_\ell \in W_\ell$ such that

$$a(y_\ell, v_\ell) = (f + u_\ell, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell, \quad (2.2a)$$

$$a(p_\ell, v_\ell) = -(y_\ell - y^d, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell, \quad (2.2b)$$

$$u_\ell = \frac{1}{\alpha} (M_\ell p_\ell - \sigma_\ell) \in K_\ell, \quad (2.2c)$$

$$(\sigma_\ell, v_\ell - u_\ell)_{0,\Omega} = 0 \quad v_\ell \in K_\ell. \quad (2.2d)$$

Here, $M_\ell : H_0^1(\Omega) \rightarrow W_\ell$ stands for the operator given by

$$(M_\ell v)_T := v_T = |T|^{-1} \int_T v(x) dx, \quad T \in \mathcal{T}_\ell. \quad (2.3)$$

As in the continuous case, (2.2d) can be stated as the complementarity condition

$$\begin{aligned} \sigma_\ell &\geq 0, \quad \psi_\ell - u_\ell \geq 0, \\ (\sigma_\ell, \psi_\ell - u_\ell)_{0,\Omega} &= 0. \end{aligned} \quad (2.4)$$

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

$$\mathcal{A}(u_\ell) := \bigcup \{ T \in \mathcal{T}_\ell \mid u_\ell|_T = \psi_\ell|_T \}, \quad (2.5a)$$

$$\mathcal{I}(u_\ell) := \bigcup \{ T \in \mathcal{T}_\ell \mid u_\ell|_T < \psi_\ell|_T \} \quad (2.5b)$$

and refer to $\mathcal{F}(u_\ell) := \partial\mathcal{A}(u_\ell)$ as the discrete free boundary between the discrete active and inactive sets.

We note that the discrete state and co-state $y_\ell, p_\ell \in V_\ell$ may also be considered as finite element approximations of an auxiliary state $y(u_\ell) \in H_0^1(\Omega)$ and an auxiliary co-state $p(u_\ell) \in H_0^1(\Omega)$ as given by the coupled elliptic system

$$a(y(u_\ell), v) = (f + u_\ell, v)_{0,\Omega}, \quad v \in H_0^1(\Omega), \quad (2.6a)$$

$$a(p(u_\ell), v) = -(y(u_\ell) - y^d, v)_{0,\Omega}, \quad v \in H_0^1(\Omega). \quad (2.6b)$$

Obviously, we have the Galerkin orthogonality

$$a(y_\ell - y(u_\ell), v_\ell) = 0 \quad , \quad v_\ell \in V_\ell . \quad (2.7)$$

Furthermore, there holds

$$\|y(u_\ell) - y\|_{1,\Omega} \leq c_F(\Omega) \|u - u_\ell\|_{0,\Omega} , \quad (2.8a)$$

$$\|p(u_\ell) - p\|_{1,\Omega} \leq c_F(\Omega) \|y - y(u_\ell)\|_{0,\Omega} , \quad (2.8b)$$

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

$$\|v\|_{0,\Omega} \leq c_F(\Omega) |v|_{1,\Omega} , \quad v \in H_0^1(\Omega) . \quad (2.9)$$

Throughout the rest of this paper, we assume that the coupled system (2.6a), (2.6b) is $H^{1+\gamma}$ -regular for some $\gamma > 0$ which implies the existence of a constant $C_r > 0$, depending only on the shape regularity of the triangulations, such that

$$\|y_\ell - y(u_\ell)\|_{0,\Omega} \leq C_r h_\ell^\gamma |y_\ell - y(u_\ell)|_{1,\Omega} . \quad (2.10)$$

3. The adaptive loop

In the step SOLVE of the adaptive loop, for the computation of the solution of (2.1a)–(2.1c) we use the primal-dual active set strategy as described in [7]. In the step ESTIMATE, we use the residual type error estimator

$$\eta := \left(\eta_y^2 + \eta_p^2 \right)^{1/2} , \quad (3.1)$$

$$\eta_y := \left(\sum_{T \in \mathcal{T}_\ell} \eta_{y,T}^2 + \sum_{E \in \mathcal{E}_\ell} \eta_{y,E}^2 \right)^{1/2} , \quad (3.2)$$

$$\eta_p := \left(\sum_{T \in \mathcal{T}_\ell} \sum_{i=1}^2 (\eta_{p,T}^{(i)})^2 + \sum_{E \in \mathcal{E}_\ell} \eta_{p,E}^2 \right)^{1/2} . \quad (3.3)$$

The estimator consists of easily computable element residuals and edge residuals. In particular, for $T \in \mathcal{T}_\ell$ the element residuals $\eta_{y,T}$ and $\eta_{p,T}^{(i)}$, $1 \leq i \leq 2$, are as follows

$$\eta_{y,T} := h_T \|f + u_\ell\|_{0,T} , \quad (3.4)$$

$$\eta_{p,T}^{(1)} := h_T \|y^d - y_\ell\|_{0,T} , \quad (3.5)$$

$$\eta_{p,T}^{(2)} := \|M_\ell p_\ell - p_\ell\|_{0,T} , \quad (3.6)$$

whereas for $E \in \mathcal{E}_\ell$ the edge residuals $\eta_{y,E}$, $\eta_{p,E}$ are given by

$$\eta_{y,E} := h_E^{1/2} \|\boldsymbol{\nu}_E \cdot [\nabla y_\ell]\|_{0,E} , \quad (3.7)$$

$$\eta_{p,E} := h_E^{1/2} \|\boldsymbol{\nu}_E \cdot [\nabla p_\ell]\|_{0,E} . \quad (3.8)$$

Here, $E = T_1 \cap T_2$, $T_\nu \in \mathcal{T}_\ell$, $1 \leq \nu \leq 2$, and $\boldsymbol{\nu}_E$ is the exterior unit normal vector on E directed towards T_2 , whereas $[\nabla y_\ell]$ and $[\nabla p_\ell]$ denote the jumps of ∇y_ℓ , ∇p_ℓ across E .

Moreover, the convergence analysis invokes data oscillations in y^d , f and ψ

$$\text{osc}_\ell(y^d) := \left(\sum_{T \in \mathcal{T}_\ell} \text{osc}_T(y^d)^2 \right)^{1/2}, \quad \text{osc}_T(y^d) := h_T \|y^d - y_\ell^d\|_{0,T}, \quad (3.9a)$$

$$\text{osc}_\ell(f) := \left(\sum_{T \in \mathcal{T}_\ell} \text{osc}_T(f)^2 \right)^{1/2}, \quad \text{osc}_T(f) := h_T \|f - f_\ell\|_{0,T}, \quad (3.9b)$$

$$\text{osc}_\ell(\psi) := \left(\sum_{T \in \mathcal{T}_\ell} \text{osc}_T(\psi)^2 \right)^{1/2}, \quad \text{osc}_T(\psi) := h_T \|\nabla \psi\|_{0,T}. \quad (3.9c)$$

where $y_\ell^d \in W_\ell$ and $f_\ell \in W_\ell$ with $y_\ell^d|_T := y_T^d$, $f_\ell|_T := f_T$, $T \in \mathcal{T}_\ell$.

Given universal constants Θ_i , $1 \leq i \leq 2$ with $0 < \Theta_i < 1$, in the bulk criterion of step MARK we select a set of edges $\mathcal{M}_\ell^E \subset \mathcal{E}_\ell$ and a set of elements $\mathcal{M}_\ell^T := \mathcal{M}_\ell^{\eta,T} \cup \mathcal{M}^{\text{osc},T} \subset \mathcal{T}_\ell$ such that

$$\Theta_1 \sum_{E \in \mathcal{E}_\ell} (\eta_{y,E}^2 + \eta_{p,E}^2) \leq \sum_{E \in \mathcal{M}^E} (\eta_{y,E}^2 + \eta_{p,E}^2), \quad (3.10)$$

$$\begin{aligned} & \Theta_2 \left(\sum_{T \in \mathcal{T}_\ell} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2) \right) \\ & \leq \sum_{T \in \mathcal{M}^{\eta,T}} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2). \end{aligned} \quad (3.11)$$

The bulk criteria are realized by a greedy algorithm (cf., e.g., [17]).

In the final step REFINE, an element T selected in the bulk criterion is refined by successive bisection such that at least one interior nodal point is generated ('interior node property'). If two or three edges of an element have been marked for refinement, the triangle is subdivided into four subtriangles by joining the midpoints of the edges, whereas simple bisection is used, if only one edge has been selected in the bulk criterion. Bisection is also used in case of newly created nodes at midpoint of edges not contained in \mathcal{M}^E in order to provide a geometrically conforming new triangulation $\mathcal{T}_{\ell+1}$. Setting

$$\text{osc}_\ell := \left(\text{osc}_\ell^2(y^d) + \text{osc}_\ell^2(f) + \text{osc}_\ell^2(\psi) \right)^{1/2}, \quad (3.12)$$

we assume that $\mathcal{T}_{\ell+1}$ is such that there exists $0 \leq \rho_2 < 1$ satisfying

$$\text{osc}_{\ell+1}^2 \leq \rho_2 \text{osc}_\ell^2. \quad (3.13)$$

In practice, the oscillation term osc_ℓ is included in the bulk criteria of step MARK (cf., e.g., [24] for a thorough discussion of the oscillation term and see [17] for details of the algorithmic realization).

4. Reliability and discrete local efficiency

The reliability of the estimator has been established in [17].

Theorem 4.1. *Let (y, p, u, σ) and $(y_\ell, p_\ell, u_\ell, \sigma_\ell)$ be the solutions of (1.2a)–(1.2d) and (2.2a)–(2.2d), and let η and $\mu_\ell(\psi)$ be the residual error estimator and the data oscillations as given by (3.1) and (3.9c), respectively. Then, there exists a constant $C_1 > 1$, depending only on α and on the shape regularity of the triangulations, such that*

$$\begin{aligned} |y - y_\ell|_{1,\Omega}^2 + |p - p_\ell|_{1,\Omega}^2 + \|u - u_\ell\|_{0,\Omega}^2 \\ + \|\sigma - \sigma_\ell\|_{0,\Omega}^2 \leq C_1 (\eta^2 + \text{osc}_\ell^2(\psi)) . \end{aligned} \quad (4.1)$$

For the discrete local efficiency of the estimator we have to show that 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.

Lemma 4.2. *For a refined element $T \in \mathcal{T}_\ell$ there holds*

$$\eta_{y,T}^2 \lesssim |y_\ell - y_{\ell+1}|_{1,T}^2 + h_T^2 \|u_\ell - u_{\ell+1}\|_{0,T}^2 + \text{osc}_T^2(f) , \quad (4.2)$$

$$(\eta_{p,T}^{(1)})^2 \lesssim |p_\ell - p_{\ell+1}|_{1,T}^2 + |y_\ell - y_{\ell+1}|_{1,T}^2 + \text{osc}_T^2(y^d) , \quad (4.3)$$

$$(\eta_{p,T}^{(2)})^2 \lesssim |p_\ell - p_{\ell+1}|_{0,T}^2 + \alpha^2 \|u_\ell - u_{\ell+1}\|_{0,T}^2 + \|\sigma_\ell - \sigma_{\ell+1}\|_{0,T}^2 . \quad (4.4)$$

Proof. Let $\varphi_{\ell+1}^a \in V_{\ell+1}$ be a nodal basis function associated with an interior point $a \in \mathcal{N}_{\ell+1}(T)$ and $D_a := \text{supp}(\varphi_{\ell+1}^a)$. Then, the function $z_{\ell+1} := (f_T + u_\ell)\varphi_{\ell+1}^a$ satisfies

$$\|f_T + u_\ell\|_{0,T}^2 \lesssim (f_T + u_\ell, z_{\ell+1})_{0,T} , \quad (4.5)$$

$$\|z_{\ell+1}\|_{0,T} \lesssim \|f_T + u_\ell\|_{0,T} \quad , \quad |z_{\ell+1}|_{1,T} \lesssim h_T^{-1} \|f_T + u_\ell\|_{0,T} . \quad (4.6)$$

Using (4.5) and (4.6) we find

$$\begin{aligned} \eta_{y,T}^2 &= h_T^2 \|f_T + u_\ell\|_{0,T}^2 \lesssim h_T^2 (f + u_{\ell+1}, z_{\ell+1})_{0,T} \\ &+ h_T^2 \left(\|u_\ell - u_{\ell+1}\|_{0,T} + \|f - f_T\|_{0,T} \right) \|z_{\ell+1}\|_{0,T} . \end{aligned} \quad (4.7)$$

Since $z_{\ell+1}$ is an admissible test function in (2.5a) (with ℓ replaced by $\ell + 1$), we have

$$a(y_{\ell+1}, z_{\ell+1}) = (f + u_{\ell+1}, z_{\ell+1})_{0,T} .$$

Observing $\Delta y_\ell = 0$ on T and $z_{\ell+1}|_{\partial D_a} = 0$, a simple integration by parts shows

$$\begin{aligned} a|_T(y_\ell, z_{\ell+1}) &= \sum_{T' \in \mathcal{T}_{\ell+1}(D_a)_{T'}} \int_{T'} \nabla y_\ell \cdot \nabla z_{\ell+1} \, dx \\ &= - \sum_{T' \in \mathcal{T}_{\ell+1}(D_a)_{T'}} \int_{T'} \Delta y_\ell z_{\ell+1} \, dx + \int_{\partial D_a} \boldsymbol{\nu}_{\partial D_a} \cdot \nabla y_\ell z_{\ell+1} \, ds = 0 . \end{aligned}$$

Consequently, we obtain

$$\begin{aligned} & h_T^2 (f + u_{\ell+1}, z_h)_{0,T} \\ &= h_T^2 a(y_{\ell+1} - y_\ell, z_{\ell+1}) \leq h_T^2 |y_{\ell+1} - y_\ell|_{1,T} |z_{\ell+1}|_{1,T}. \end{aligned} \quad (4.8)$$

Inserting (4.8) into (4.7) and using (4.6) as well as Young's inequality gives the assertion.

The proof of (4.3) follows by similar arguments, this time choosing $z_{\ell+1} = (y_\ell^d - \hat{y}_\ell)\varphi_{\ell+1}^a$, where \hat{y}_ℓ is the integral mean of y_ℓ on T .

For the proof of (4.4), the triangle inequality readily gives

$$\eta_{p,T}^{(2)} \leq \|M_\ell p_\ell - M_{\ell+1} p_{\ell+1}\|_{0,T} + \|p_\ell - p_{\ell+1}\|_{0,T} + \|M_{\ell+1} p_\ell - p_\ell\|_{0,T}. \quad (4.9)$$

Using the relationship (2.2c) both for the coarse and the fine mesh, for the first term on the right-hand side in (4.9) we obtain

$$\|M_\ell p_\ell - M_{\ell+1} p_{\ell+1}\|_{0,T} \leq \alpha \|u_\ell - u_{\ell+1}\|_{0,T} + \|\sigma_\ell - \sigma_{\ell+1}\|_{0,T}. \quad (4.10)$$

For the third term on the right-hand side in (4.9), there exists $0 \leq q < 1$ such that

$$\|M_{\ell+1} p_\ell - p_\ell\|_{0,T} \leq q \|M_\ell p_\ell - p_\ell\|_{0,T}. \quad (4.11)$$

Taking advantage of (4.10), (4.11) in (4.9) yields

$$\eta_{p,T}^{(2)} \leq \frac{1}{1-q} (\|p_\ell - p_{\ell+1}\|_{0,\Omega} + \alpha \|u_\ell - u_{\ell+1}\|_{0,\Omega} + \|\sigma_\ell - \sigma_{\ell+1}\|_{0,\Omega}).$$

□

Lemma 4.3. *For a refined edge $E \in \mathcal{E}_\ell$ there holds*

$$\eta_{y,E}^2 \lesssim |y_\ell - y_{\ell+1}|_{1,\omega_E}^2 + h_T^2 \|u_\ell - u_{\ell+1}\|_{0,\omega_E}^2 + \eta_{y,\omega_E}^2, \quad (4.12)$$

$$\eta_{p,E}^2 \lesssim |p_\ell - p_{\ell+1}|_{1,\omega_E}^2 + |y_\ell - y_{\ell+1}|_{1,\omega_E}^2 + \eta_{p,\omega_E}^2. \quad (4.13)$$

Proof. Let $\varphi_{\ell+1}^{mid_E} \in V_{\ell+1}$ be the nodal basis function associated with $\text{mid}(E) \in \mathcal{N}_{\ell+1}(\Omega)$. Then, the function $z_{\ell+1} := [\boldsymbol{\nu}_E \cdot \nabla y_\ell] \varphi_{\ell+1}^{mid_E}$ satisfies

$$\|[\boldsymbol{\nu}_E \cdot \nabla y_\ell]\|_{0,E}^2 \lesssim ([\boldsymbol{\nu}_E \cdot \nabla y_\ell], z_{\ell+1})_{0,E}, \quad (4.14)$$

$$\|z_{\ell+1}\|_{0,\omega_E} \lesssim h_E^{1/2} \|[\boldsymbol{\nu}_E \cdot \nabla y_\ell]\|_{0,E}, \quad (4.15)$$

$$|z_{\ell+1}|_{1,\omega_E} \lesssim h_E^{-1/2} \|[\boldsymbol{\nu}_E \cdot \nabla y_\ell]\|_{0,E}. \quad (4.16)$$

Using (4.14)–(4.16) and the fact that $z_{\ell+1}$ is an admissible test function in (2.5a) (with ℓ replaced by $\ell + 1$), we find

$$\begin{aligned} \eta_{y,E}^2 &= h_E \|[\boldsymbol{\nu}_E \cdot \nabla y_\ell]\|_{0,E}^2 \lesssim h_E ([\boldsymbol{\nu}_E \cdot \nabla y_\ell], z_{\ell+1})_{0,E} \\ &= h_E \left(a|_{\omega_E} (y_\ell - y_{\ell+1}, z_{\ell+1}) + (u_\ell - u_{\ell+1}, z_{\ell+1})_{0,\Omega_E} - (f + u_\ell, z_{\ell+1})_{0,\Omega_E} \right) \\ &\lesssim h_E^{1/2} \|[\boldsymbol{\nu}_E \cdot \nabla y_\ell]\|_{0,E} \left(|y_\ell - y_{\ell+1}|_{1,\omega_E} + h_T \|u_\ell - u_{\ell+1}\|_{0,\omega_E} + \eta_{y,\omega_E} \right), \end{aligned} \quad (4.17)$$

which immediately leads to (4.12). The estimate (4.13) is shown in exactly the same way. □

Summarizing the results of Lemma 4.2 and Lemma 4.3 and taking account that the union of the patches ω_E has a finite overlap, we obtain

Theorem 4.4. *Let (y, p, u, σ) and $(y_k, p_k, u_k, \sigma_k), k \in \{\ell, \ell + 1\}$, be the solutions of (1.2a)–(1.2d) and (2.2a)–(2.2d), and let η and $\text{osc}_\ell(y^d), \text{osc}_\ell(f)$ be the residual error estimator and the data oscillations as given by (3.1) and (3.9a), (3.9b), respectively. Then, there exists a constant $C_2 > 1$, depending only on the constants $\Theta_i, 1 \leq i \leq 2$, in the bulk criteria (3.10), (3.11), and on the shape regularity of the triangulations, such that*

$$\begin{aligned} & \sum_{T \in \mathcal{M}_T^\ell} (\eta_{y,T}^2 + (\eta_{p,T}^{(1)})^2 + (\eta_{p,T}^{(2)})^2) + \sum_{E \in \mathcal{M}_E^\ell} (\eta_{y,E}^2 + \eta_{p,E}^2) \\ & \leq C_2 (|y_\ell - y_{\ell+1}|_{1,\Omega}^2 + |p_\ell - p_{\ell+1}|_{1,\Omega}^2 + \alpha^2 \|u_\ell - u_{\ell+1}\|_{0,\Omega}^2 \\ & \quad + \|\sigma_\ell - \sigma_{\ell+1}\|_{0,\Omega}^2 + \text{osc}_\ell^2(y^d) + \text{osc}_\ell^2(f)) . \end{aligned} \quad (4.18)$$

5. Convergence result

In this section, we will prove convergence of the discrete states, co-states, controls, and co-controls to its continuous counterparts.

The reliability (4.1), the bulk criteria (3.10), (3.11), and the discrete local efficiency (4.18) imply

$$\begin{aligned} & |y - y_\ell|_{1,\Omega}^2 + |p - p_\ell|_{1,\Omega}^2 + \|u - u_\ell\|_{0,\Omega}^2 + \|\sigma - \sigma_\ell\|_{0,\Omega}^2 \\ & \leq C_3 (\alpha |y_\ell - y_{\ell+1}|_{1,\Omega}^2 + |p_\ell - p_{\ell+1}|_{1,\Omega}^2) + C_4 (\alpha^2 \|u_\ell - u_{\ell+1}\|_{0,\Omega}^2 \\ & \quad + \|\sigma_\ell - \sigma_{\ell+1}\|_{0,\Omega}^2) + C_5 (\text{osc}_\ell^2(y^d) + \text{osc}_\ell^2(f) + \text{osc}_\ell^2(\psi)) , \end{aligned} \quad (5.1)$$

where $C_3 := C_1 C_2 \alpha^{-1}$, $C_4 := C_1 C_2$ and $C_5 := \max(C_1, C_2)$.

Now, using the fundamental relationships

$$|s_\ell - s_{\ell+1}|_{1,\Omega}^2 \quad (5.2)$$

$$= |s - s_\ell|_{1,\Omega}^2 - |s - s_{\ell+1}|_{1,\Omega}^2 + 2a(s - s_{\ell+1}, s_\ell - s_{\ell+1}) ,$$

$$\|w_\ell - w_{\ell+1}\|_{0,\Omega}^2 \quad (5.3)$$

$$= \|w - w_\ell\|_{0,\Omega}^2 - \|w - w_{\ell+1}\|_{0,\Omega}^2 + 2(w - w_{\ell+1}, w_\ell - w_{\ell+1})_{0,\Omega}$$

for $s = y, s_k = y_k$ and $s = p, s_k = p_k$ and for $w = u, w_k = u_k$ and $w = \sigma, w_k = \sigma_k, k \in \{\ell, \ell + 1\}$, we would be able to deduce not only convergence, but even an error reduction property, if we had Galerkin orthogonality of the AFEM. However, Galerkin orthogonality does not apply here. Instead, we will establish some quasi-orthogonality properties which allow to prove convergence.

Lemma 5.1. *Let (y, p, u, σ) and $(y_k, p_k, u_k, \sigma_k), k \in \{\ell, \ell + 1\}$, be the solutions of (1.2a)–(1.2d) and (2.2a)–(2.2d), and let $y(u_{\ell+1}) \in V$ be the auxiliary state as given*

by (2.6). Then, there holds

$$\begin{aligned} & \alpha a(y - y_{\ell+1}, y_\ell - y_{\ell+1}) + a(p - p_{\ell+1}, p_\ell - p_{\ell+1}) \\ &= (y_{\ell+1} - y(u_{\ell+1}), p_\ell - p_{\ell+1})_{0,\Omega} + (y(u_{\ell+1}) - y, \sigma_\ell - \sigma_{\ell+1})_{0,\Omega} \\ & \quad + (y(u_{\ell+1}) - y, (I - M_\ell)p_\ell - (I - M_{\ell+1})p_{\ell+1})_{0,\Omega} \\ & \quad + \alpha (u_{\ell+1} - u, (y(u_\ell) - y_\ell) + (y_{\ell+1} - y(u_{\ell+1})))_{0,\Omega} , \end{aligned} \quad (5.4)$$

$$\begin{aligned} & \alpha^2 (u - u_{\ell+1}, u_\ell - u_{\ell+1})_{0,\Omega} + (\sigma - \sigma_{\ell+1}, \sigma_\ell - \sigma_{\ell+1}) \\ & \leq \alpha (\sigma, u_{\ell+1} - u_\ell)_{0,\Omega} + \alpha (\sigma_{\ell+1} - \sigma, \psi_\ell - \psi_{\ell+1})_{0,\Omega} \\ & \quad + \alpha (\sigma, \psi_\ell - \psi_{\ell+1})_{0,\Omega} + \alpha (\sigma_{\ell+1} - \sigma_\ell, u - \psi)_{0,\Omega} \\ & \quad + \alpha (\sigma_\ell - \sigma_{\ell+1}, \psi_{\ell+1} - \psi)_{0,\Omega} + (p - M_{\ell+1}p_{\ell+1}, M_\ell p_\ell - M_{\ell+1}p_{\ell+1})_{0,\Omega}. \end{aligned} \quad (5.5)$$

Proof. In view of (1.2a), (2.2a) and (2.6), (2.7), we readily get

$$\begin{aligned} & \alpha a(y - y_{\ell+1}, y_\ell - y_{\ell+1}) \\ &= \alpha a(y - y(u_{\ell+1}), y_\ell - y_{\ell+1}) = \alpha (u - u_{\ell+1}, y_\ell - y_{\ell+1})_{0,\Omega} . \end{aligned} \quad (5.6)$$

On the other hand, observing (1.2b), (1.2c) and (2.2b), (2.2c) as well as (2.8), we find

$$\begin{aligned} & a(p - p_{\ell+1}, p_\ell - p_{\ell+1}) = (y_{\ell+1} - y(u_{\ell+1}), p_\ell - p_{\ell+1})_{0,\Omega} \\ & \quad + (y(u_{\ell+1}) - y, \sigma_\ell - \sigma_{\ell+1} + (I - M_\ell)p_\ell - (I - M_{\ell+1})p_{\ell+1})_{0,\Omega} \\ & \quad + \alpha (y(u_{\ell+1}) - y, u_\ell - u_{\ell+1})_{0,\Omega} . \end{aligned} \quad (5.7)$$

Moreover, since $y(u_{\ell+1}) - y = (-\Delta)^{-1}(u_{\ell+1} - u)$ and $(-\Delta)^{-1}(u_\ell - u_{\ell+1}) = y(u_\ell) - y(u_{\ell+1})$, we obtain

$$\begin{aligned} & \alpha (y(u_{\ell+1}) - y, u_\ell - u_{\ell+1})_{0,\Omega} \\ &= \alpha (u_{\ell+1} - u, y(u_\ell) - y(u_{\ell+1}))_{0,\Omega} = \alpha (u_{\ell+1} - u, y_\ell - y_{\ell+1})_{0,\Omega} \\ & \quad + \alpha (u_{\ell+1} - u, (y(u_\ell) - y_\ell) + (y_{\ell+1} - y(u_{\ell+1})))_{0,\Omega} . \end{aligned} \quad (5.8)$$

Using (5.8) in (5.7) and combining (5.6) and (5.7) results in (5.4).

As far as the proof of (5.5) is concerned, using again (1.2c), (2.2c), we find

$$\begin{aligned} & \alpha^2 (u - u_{\ell+1}, u_\ell - u_{\ell+1})_{0,\Omega} + (\sigma - \sigma_{\ell+1}, \sigma_\ell - \sigma_{\ell+1}) \\ &= \alpha (\sigma_{\ell+1} - \sigma, u_\ell - u_{\ell+1})_{0,\Omega} + \alpha (u - u_{\ell+1}, \sigma_{\ell+1} - \sigma_\ell)_{0,\Omega} \\ & \quad + (p - M_{\ell+1}p_{\ell+1}, M_\ell p_\ell - M_{\ell+1}p_{\ell+1})_{0,\Omega} . \end{aligned} \quad (5.9)$$

For the first term on the right-hand side in (5.9), we obtain

$$\begin{aligned} & \alpha (\sigma_{\ell+1} - \sigma, u_\ell - u_{\ell+1})_{0,\Omega} \\ &= \alpha (\sigma_{\ell+1}, u_\ell - \psi_\ell)_{0,\Omega} + \alpha (\sigma_{\ell+1}, \psi_\ell - \psi_{\ell+1})_{0,\Omega} \\ & \quad + \alpha (\sigma_{\ell+1}, \psi_{\ell+1} - u_{\ell+1})_{0,\Omega} + \alpha (\sigma, u_{\ell+1} - u_\ell)_{0,\Omega} \\ & \leq \alpha (\sigma_{\ell+1}, \psi_\ell - \psi_{\ell+1})_{0,\Omega} + \alpha (\sigma, u_{\ell+1} - u_\ell)_{0,\Omega} , \end{aligned} \quad (5.10)$$

where we have used that due to (2.4)

$$(\sigma_{\ell+1}, u_\ell - \psi_\ell)_{0,\Omega} \leq 0 \quad \text{and} \quad (\sigma_{\ell+1}, \psi_{\ell+1} - u_{\ell+1})_{0,\Omega} = 0 .$$

For the second term, similar arguments yield

$$\begin{aligned} & \alpha(\sigma_{\ell+1} - \sigma_\ell, u - u_{\ell+1})_{0,\Omega} \\ &= \alpha(\sigma_{\ell+1} - \sigma_\ell, u - \psi)_{0,\Omega} + \alpha(\sigma_{\ell+1} - \sigma_\ell, \psi_{\ell+1} - \psi)_{0,\Omega} . \end{aligned} \quad (5.11)$$

Using (5.10), (5.11) in (5.9) gives the assertion. \square

For the terms in (5.4), (5.5) involving the averaging operators M_ℓ and $M_{\ell+1}$ we provide the following result.

Lemma 5.2. *Under the assumptions of Lemma 5.1 there holds*

$$(y_{\ell+1} - y, (I - M_\ell)p_\ell - (I - M_{\ell+1})p_{\ell+1})_{0,\Omega} \quad (5.12)$$

$$\lesssim h_\ell^2 (|y - y_{\ell+1}|_{1,\Omega}^2 + |p - p_\ell|_{1,\Omega}^2 + |p - p_{\ell+1}|_{1,\Omega}^2) + \mu_\ell^2(p) ,$$

$$(p - M_{\ell+1}p_{\ell+1}, M_\ell p_\ell - M_{\ell+1}p_{\ell+1})_{0,\Omega} \quad (5.13)$$

$$\lesssim h_\ell^2 (|p - p_\ell|_{1,\Omega}^2 + |p - p_{\ell+1}|_{1,\Omega}^2) + \mu_\ell^2(p) ,$$

where

$$\mu_\ell(p) := \left(\sum_{T \in \mathcal{T}_\ell} \mu_T^2(p) \right)^{1/2} , \quad \mu_T(p) := h_T |p|_{1,T} . \quad (5.14)$$

Proof. We split the left-hand side in (5.9) according to

$$\begin{aligned} & (y_{\ell+1} - y, (I - M_\ell)p_\ell - (I - M_{\ell+1})p_{\ell+1})_{0,\Omega} \\ &= (y_{\ell+1} - y, (I - M_{\ell+1})(p_\ell - p_{\ell+1}))_{0,\Omega} \\ & \quad + (y_{\ell+1} - y, (M_{\ell+1} - M_\ell)p_\ell)_{0,\Omega} . \end{aligned}$$

For $T \in \mathcal{T}_{\ell+1}$ we set $\hat{y}_{\ell+1} := |T|^{-1} \int_T (y_{\ell+1} - y) dx$. Since $M_{\ell+1}(p_\ell - p_{\ell+1})$ has zero integral mean on $T \in \mathcal{T}_{\ell+1}$, an elementwise application of Poincaré's inequality and of Young's inequality gives

$$\begin{aligned} & (y_{\ell+1} - y, (I - M_{\ell+1})(p_\ell - p_{\ell+1}))_{0,T} \\ &= (y_{\ell+1} - y - \hat{y}_{\ell+1}, (I - M_{\ell+1})(p_\ell - p_{\ell+1}))_{0,T} \\ &\lesssim h_T |y - y_{\ell+1}|_{1,T} h_T |p_\ell - p_{\ell+1}|_{1,T} \\ &\lesssim h_T^2 |y - y_{\ell+1}|_{1,T}^2 + h_T^2 |p_\ell - p_{\ell+1}|_{1,T}^2 . \end{aligned}$$

Summing over all $T \in \mathcal{T}_{\ell+1}$, we obtain

$$\begin{aligned} & (y_{\ell+1} - y, (I - M_\ell)p_\ell - (I - M_{\ell+1})p_{\ell+1})_{0,\Omega} \\ &\lesssim h_\ell^2 (|y - y_{\ell+1}|_{1,\Omega}^2 + |p_\ell - p_{\ell+1}|_{1,\Omega}^2) . \end{aligned}$$

Moreover, using similar arguments

$$\begin{aligned} & (y_{\ell+1} - y, (M_{\ell+1} - M_\ell)p_\ell)_{0,\Omega} \\ &\lesssim h_\ell^2 (|y - y_{\ell+1}|_{1,\Omega}^2 + |p - p_\ell|_{1,\Omega}^2) + \mu_\ell^2(p) . \end{aligned}$$

Combing both inequalities proves (5.12). The proof of (5.13) is along the same lines. \square

Since the $|\cdot|_{1,\Omega}$ -norm of the co-state p can be bounded from above by means of the given data of the problem (cf., e.g., [20]), we may interpret $\mu_\ell(p)$ as a data term. As far as the reduction of that data term is concerned, we may assume the existence of $0 \leq \rho_3 < 1$ such that

$$\mu_{\ell+1}^2(p) \leq \rho_3 \mu_\ell^2(p). \quad (5.15)$$

For the convergence proof, we set $z := (y, p, u, \sigma)$, $z_\ell := (y_\ell, p_\ell, u_\ell, \sigma_\ell)$, $\ell \in \mathbb{N}_0$, and introduce the norm

$$\| \|z - z_\ell \| \| := \left(|y - y_\ell|_{1,\Omega}^2 + |p - p_\ell|_{1,\Omega}^2 + \|u - u_\ell\|_{0,\Omega}^2 + \|\sigma - \sigma_\ell\|_{0,\Omega}^2 \right)^{1/2}. \quad (5.16)$$

We establish convergence with respect to $\| \cdot \|$ in the sense that the sequence $\{ \|z - z_\ell \| \}_{\mathbb{N}_0}$ belongs to ℓ^2 .

Theorem 5.3. *Let (y, p, u, σ) and $(y_\ell, p_\ell, u_\ell, \sigma_\ell)$ be the solutions of (1.2a)–(1.2d) and (2.2a)–(2.2d) and let $\text{osc}_\ell(y^d), \text{osc}_\ell(f), \text{osc}_\ell(\psi), \mu_\ell(p)$ be the data oscillations and data terms given by (3.9a)–(3.9c) and (5.14). Assume that (3.13) and (5.15) are satisfied. Then, there exists a constant $\Lambda > 0$, depending on the data of the problem, the constants Θ_i , $1 \leq i \leq 2$, in the bulk criteria (3.10), (3.11) and on the shape regularity of the triangulations, such that*

$$\sum_{\ell=0}^{\infty} \| \|z - z_\ell \| \|^2 \leq \Lambda. \quad (5.17)$$

Proof. In addition to Lemma (5.2), we provide further estimates for the remaining terms on the right-hand side in (5.4). In particular, by means of (2.8a), (2.9) and (2.10), setting $c_\Omega := \max(1, c_F(\Omega))$ we obtain

$$\begin{aligned} & (y_{\ell+1} - y(u_{\ell+1}), p_\ell - p_{\ell+1})_{0,\Omega} \leq C_r c_\Omega h_\ell^\gamma (|y - y_{\ell+1}|_{1,\Omega} \\ & + |y - y(u_{\ell+1})|_{1,\Omega}) (|p - p_\ell|_{1,\Omega} + |p - p_{\ell+1}|_{1,\Omega}) \\ & \leq \frac{1}{2} C_r c_\Omega h_\ell^\gamma (|y - y_{\ell+1}|_{1,\Omega}^2 + 4|p - p_\ell|_{1,\Omega}^2 + 4|p - p_{\ell+1}|_{1,\Omega}^2) \\ & + \frac{1}{2} C_r c_\Omega^2 h_\ell^\gamma \|u - u_{\ell+1}\|_{0,\Omega}^2, \end{aligned} \quad (5.18)$$

$$\begin{aligned} & \alpha(u_{\ell+1} - u, (y(u_\ell) - y_\ell) + (y_{\ell+1} - y(u_{\ell+1})))_{0,\Omega} \\ & \leq \alpha C_r h_\ell^\gamma \|u - u_{\ell+1}\|_{0,\Omega} (|y - y_\ell|_{1,\Omega} + |y - y_{\ell+1}|_{1,\Omega} \\ & + c_\Omega (\|u - u_\ell\|_{0,\Omega} + \|u - u_{\ell+1}\|_{0,\Omega})) \\ & \leq \alpha C_r h_\ell^\gamma (|y - y_\ell|_{1,\Omega}^2 + |y - y_{\ell+1}|_{1,\Omega}^2 \\ & + c_\Omega^2 \|u - u_\ell\|_{0,\Omega}^2 + (1 + c_\Omega^2) \|u - u_{\ell+1}\|_{0,\Omega}^2). \end{aligned} \quad (5.19)$$

Moreover, in view of (2.8a), (2.9) and Young's inequality

$$\begin{aligned} & (y(u_{\ell+1}) - y, \sigma_\ell - \sigma_{\ell+1})_{0,\Omega} \leq c_\Omega^2 \|u - u_{\ell+1}\|_{0,\Omega} \|\sigma_\ell - \sigma_{\ell+1}\|_{0,\Omega} \\ & \leq \varepsilon_1 \|u - u_{\ell+1}\|_{0,\Omega}^2 + \frac{c_\Omega^4}{4\varepsilon_1} \|\sigma_\ell - \sigma_{\ell+1}\|_{0,\Omega}^2, \end{aligned} \quad (5.20)$$

where $\varepsilon_1 > 0$ can be arbitrarily chosen. Likewise, for some arbitrary $\varepsilon_2 > 0$ we get

$$\begin{aligned} & \alpha(\sigma_{\ell+1} - \sigma, \psi_\ell - \psi_{\ell+1})_{0,\Omega} \\ & \leq \alpha \|\sigma - \sigma_{\ell+1}\|_{0,\Omega} (\|\psi - \psi_\ell\|_{0,\Omega} + \|\psi - \psi_\ell\|_{0,\Omega}) \\ & \leq \varepsilon_2 \|\sigma - \sigma_{\ell+1}\|_{0,\Omega}^2 + \frac{\alpha^2}{4\varepsilon_2} (1 + \rho_2) \operatorname{osc}_\ell^2(\psi) , \end{aligned} \quad (5.21)$$

$$\begin{aligned} & \alpha(\sigma_\ell - \sigma_{\ell+1}, \psi_{\ell+1} - \psi)_{0,\Omega} \\ & \leq \varepsilon_2 (\|\sigma - \sigma_\ell\|_{0,\Omega}^2 + \|\sigma - \sigma_{\ell+1}\|_{0,\Omega}^2) + \frac{\alpha^2}{2\varepsilon_2} \rho_2 \operatorname{osc}_\ell^2(\psi) , \end{aligned} \quad (5.22)$$

where we have used (3.13) in both estimates.

Now, we choose $\varepsilon_1 > 0, \varepsilon_2 > 0$ according to

$$\varepsilon_1 := \alpha/(16C_4) , \quad \varepsilon_2 := \alpha/(16C_4(\alpha + 4c_\Omega^4))$$

and $h^* \in \mathbb{R}_+$ by means of

$$h^* := (\alpha/(120C_4 \max(C_6, C_7, C_7)c_\Omega^4))^{1/\gamma} .$$

Then, there exists $\ell^* \in \mathbb{N}$ such that $h_\ell \leq h^*$ for $\ell \geq \ell^*$. If we take advantage of (5.18)–(5.22) as well as (5.12), (5.13) from Lemma 5.2 in (5.4), (5.5) and use the result in (5.1), setting $C_8 := 4C_3c_\Omega^4 + C_4\alpha^2$ and $C_9 := 4C_3c_\Omega^4 + C_4$, for $\ell \geq \ell^*$ we get

$$\begin{aligned} |||z - z_\ell|||^2 & \leq (C_4 + \frac{1}{12})|y - y_\ell|_{1,\Omega}^2 - (C_4 - \frac{1}{4})|y - y_{\ell+1}|_{1,\Omega}^2 \\ & + (C_3 + \frac{1}{4})|p - p_\ell|_{1,\Omega}^2 - (C_3 - \frac{1}{4})|p - p_{\ell+1}|_{1,\Omega}^2 \\ & + (C_8 + \frac{1}{4})\|u - u_\ell\|_{0,\Omega}^2 - (C_8 - \frac{1}{4})\|u - u_{\ell+1}\|_{0,\Omega}^2 \\ & + (C_9 + \frac{1}{2})\|\sigma - \sigma_\ell\|_{0,\Omega}^2 - (C_9 - \frac{1}{4})\|\sigma - \sigma_{\ell+1}\|_{0,\Omega}^2 \\ & + C_{10}(\operatorname{osc}_\ell^2 + \mu_\ell^2(p)) + 2\alpha(\sigma, u_{\ell+1} - u_\ell)_{0,\Omega} \\ & + 2\alpha(\sigma, \psi_\ell - \psi_{\ell+1})_{0,\Omega} + 2\alpha(\sigma_{\ell+1} - \sigma_\ell, u - \psi)_{0,\Omega} , \end{aligned} \quad (5.23)$$

where $C_{10} := C_5 + 2C_3C_6 + 2C_9(C_7 + 4(1 + \rho_2))$. We define constants $0 < \kappa_i \leq 1, 1 \leq i \leq 3$, and $0 < \rho_1 < 1$ according to

$$\begin{aligned} \kappa_1 & := \frac{C_4 - 1/4}{C_9 - 1/4} , \quad \kappa_2 := \frac{C_3 - 1/4}{C_9 - 1/4} , \\ \kappa_3 & := \frac{C_8 - 1/4}{C_9 - 1/4} , \quad \rho_1 := \frac{C_9 - 1/2}{C_9 - 1/4} \end{aligned} \quad (5.24)$$

(observe $C_4 < C_9, C_3 < C_9$, and $C_8 \leq C_9$). We further introduce the weighted norm

$$|||z - z_\ell|||_\kappa := (\kappa_1|y - y_\ell|_{1,\Omega}^2 + \kappa_2|p - p_\ell|_{1,\Omega}^2 + \kappa_3\|u - u_\ell\|_{0,\Omega}^2 + \|\sigma - \sigma_\ell\|_{0,\Omega}^2)^{1/2} . \quad (5.25)$$

Then, from (5.23) we deduce that

$$\begin{aligned} & \|||z - z_{\ell+1}\|||_{\kappa}^2 \\ & \leq \rho \|||z - z_{\ell}\|||_{\kappa}^2 + 2\alpha(\sigma, u_{\ell+1} - u_{\ell})_{0,\Omega} + 2\alpha(\sigma, \psi_{\ell} - \psi_{\ell+1})_{0,\Omega} \\ & \quad + 2\alpha(\sigma_{\ell+1} - \sigma_{\ell}, u - \psi)_{0,\Omega} + C_{11} (\text{osc}_{\ell}^2 + \mu_{\ell}^2(p)) , \end{aligned} \quad (5.26)$$

where $C_{11} := C_{10}/(C_9 - 1/4)$. Summing in (5.26) over ℓ from $\ell = \ell^*$ to $\ell = n > \ell^*$ results in

$$\begin{aligned} (1 - \rho) \sum_{\ell=\ell^*}^{n-1} & \|||z - z_{\ell+1}\|||_{\kappa}^2 + \|||z - z_{n+1}\|||_{\kappa}^2 \leq \rho \|||z - z_{\ell^*}\|||_{\kappa}^2 \\ & + 2\alpha(\sigma, u_{n+1} - u_{\ell^*})_{0,\Omega} + 2\alpha(\sigma, \psi_{\ell^*} - \psi_{n+1})_{0,\Omega} \\ & + 2\alpha(\sigma_{n+1} - \sigma_{\ell^*}, u - \psi)_{0,\Omega} + C_7 \sum_{\ell=\ell^*}^n (\text{osc}_{\ell}^2 + \mu_{\ell}^2(p)) . \end{aligned} \quad (5.27)$$

Now, taking (1.3) and (2.4) into account, we have

$$\begin{aligned} & (\sigma, u_{n+1} - u_{\ell^*})_{0,\Omega} + (\sigma, \psi_{\ell^*} - \psi_{n+1})_{0,\Omega} \leq (\sigma, \psi_{\ell^*} - u_{\ell^*})_{0,\Omega} , \\ & (\sigma_{n+1} - \sigma_{\ell^*}, u - \psi)_{0,\Omega} \leq (\sigma_{\ell^*}, \psi - u)_{0,\Omega} . \end{aligned}$$

Moreover, due to (3.13) and (5.15)

$$\sum_{\ell=\ell^*}^{\infty} (\text{osc}_{\ell}^2 + \mu_{\ell}^2(p)) \leq \frac{C_7}{1 - \rho_2} \text{osc}_{\ell^*}^2 + \frac{C_7}{1 - \rho_3} \mu_{\ell^*}^2(p) .$$

Using the preceding estimates in (5.27) implies the existence of a constant ϑ such that

$$\min_{1 \leq i \leq 3} \kappa_i \sum_{\ell=0}^{\infty} \|||z - z_{\ell}\|||^2 \leq \vartheta ,$$

which gives the assertion. \square

Corollary 5.4. *Under the assumptions of Theorem 5.3 there holds*

$$|y - y_{\ell}|_{1,\Omega}, |p - p_{\ell}|_{1,\Omega}, \|u - u_{\ell}\|_{0,\Omega}, \|\sigma - \sigma_{\ell}\|_{0,\Omega} \rightarrow 0 \quad \text{as } \ell \rightarrow \infty . \quad (5.28)$$

6. Error reduction property

An error reduction property of the adaptive finite element approximation of the obstacle problem in the weighted norm $\||| \cdot \|||_{\kappa}$ can be established under some additional assumptions. In particular, we suppose that the sequence $\{W_{\ell}\}_{\mathbb{N}_0}$ of spaces of elementwise constants is limit dense in $L^2(\Omega)$ in the sense

- (L) For each $w \in L^2(\Omega)$, there is a sequence $\{w_{\ell}\}_{\mathbb{N}}$, $w_{\ell} \in W_{\ell}$, $\ell \in \mathbb{N}$, such that $w_{\ell} \rightarrow w$ in $L^2(\Omega)$ as $\ell \rightarrow \infty$.

We further assume strict complementarity of the continuous problem

$$\text{(C)} \quad \sigma|_{\mathcal{I}(u)} > 0 ,$$

as well as the following non-degeneracy properties of the discrete control problems:

(N₁) There exist $\varepsilon_1^* > 0$ and $C_1 > 0$ such that for all $0 < \varepsilon < \varepsilon_1^*$ and for all sufficiently large $\ell \in \mathbb{N}$

$$\text{meas}(\{x \in \mathcal{I}(u_\ell) \mid 0 < \psi_\ell(x) - u_\ell(x) < \varepsilon^2\}) \leq C_1 \varepsilon .$$

(N₂) There exist $\varepsilon_2^* > 0$ and $C_2 > 0$ such that for all $0 < \varepsilon < \varepsilon_2^*$ and for all sufficiently large $\ell \in \mathbb{N}$

$$\{x \in \mathcal{I}(u_\ell) \mid 0 < \psi_\ell(x) - u_\ell(x) < \varepsilon^2\} \subseteq \{x \in \mathcal{I}(u_\ell) \mid \text{dist}(x, \mathcal{F}_\ell) < C_2 \varepsilon\} .$$

(N₃) There exist $\varepsilon_3^* > 0$ and $C_3 > 0$ such that for all $0 < \varepsilon < \varepsilon_3^*$ and for all sufficiently large $\ell \in \mathbb{N}$

$$\{x \in \mathcal{I}(u_\ell) \mid \text{dist}(x, \mathcal{F}_\ell) < \varepsilon\} \subseteq \{x \in \mathcal{I}(u_\ell) \mid 0 < \psi_\ell(x) - u_\ell(x) < C_3 \varepsilon^2\} .$$

The error reduction property holds asymptotically, i.e., once the continuous free boundary has been sufficiently resolved by its discrete counterpart. We enhance the resolution of the free boundary by an extension of the bulk criteria. To this end, we define the sets

$$\hat{\mathcal{A}}(u_\ell) := \text{int}\left(\bigcup\{T \in \mathcal{T}_\ell \mid u_\ell|_{T'} = \psi_\ell|_{T'} , T' \in \mathcal{T}_\ell, T' \cap T = \emptyset\}\right) , \quad (6.1)$$

$$\hat{\mathcal{I}}(u_\ell) := \text{int}\left(\bigcup\{T \in \mathcal{T}_\ell \mid \psi_\ell|_T - u_\ell|_T \geq \hat{\varepsilon} > 0\}\right) , \quad (6.2)$$

$$\hat{\mathcal{F}}(u_\ell) := \Omega \setminus (\hat{\mathcal{A}}(u_\ell) \cup \hat{\mathcal{I}}(u_\ell)) \quad (6.3)$$

for some $\hat{\varepsilon} > 0$ in (6.2). Then, the extension of the bulk criteria (3.10)–(3.11) is as follows:

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

Proposition 6.1. *Assume that the discrete problem (2.2a)–(2.2d) satisfies **(N₁)**, **(N₂)** and that the refinement is done based on the bulk criteria (3.10), (3.11) and its extension **(E)**. Then, there exists a subsequence $\mathbb{N}^* \subset \mathbb{N}$ such that for all $\ell \in \mathbb{N}^*$*

$$\hat{\mathcal{I}}(u_\ell) \subseteq \hat{\mathcal{I}}(u_{\ell+1}) \quad , \quad \hat{\mathcal{A}}(u_\ell) \subseteq \hat{\mathcal{A}}(u_{\ell+1}) \quad , \quad (6.4)$$

$$\hat{\mathcal{F}}(u_{\ell+1}) \subset \hat{\mathcal{F}}(u_\ell) . \quad (6.5)$$

Proof. If the assertion does not hold true, we have $\hat{\mathcal{I}}(u_{m+1}) \subset \hat{\mathcal{I}}(u_m)$ and $\hat{\mathcal{A}}(u_{m+1}) \subset \hat{\mathcal{A}}(u_m)$ for $m > \ell$ which implies $\hat{\mathcal{F}}(u_m) \subseteq \hat{\mathcal{F}}(u_{m+1})$. Hence, in view of **(N₁)** and **(E)**, there exists $T \in \hat{\mathcal{F}}(u_m)$ such that $\text{dist}(x, \mathcal{F}(u_m)) > \tau, x \in T$, where $\tau > C_2 h_m$ with C_2 from **(N₂)**, and $0 = \psi_m(a_\nu) - u_m(a_\nu) < \psi_m(a_\mu) - u_m(a_\mu)$ for some vertices $a_\nu, a_\mu, \nu = \mu$, of T . Then, we find $\mathcal{U}(a_\nu) := \{x \in T' \mid 0 < \psi_m(x) - u_m(x) < h_m^2\}$ and **(N₂)** implies $\text{dist}(x, \mathcal{F}(u_m)) < C_2 h_m < \tau$ contradicting $\text{dist}(x, \mathcal{F}(u_m)) > \tau, x \in \mathcal{U}(a_\nu) \subset T$. Note that (6.5) is a direct consequence of (6.4). \square

Proposition 6.2. *Assume that (1.2a)–(1.2d) and (2.2a)–(2.2d) satisfy **(S)** and **(N₁)**, **(N₂)**, **(N₃)**, respectively, and that **(L)** holds true. Then, there exists $\ell^* \in \mathbb{N}$ such that*

$$\hat{\mathcal{I}}(u_\ell) \subseteq \mathcal{I}(u) \quad \text{for all } \ell \geq \ell^* . \quad (6.6)$$

Proof. If (6.6) does not hold true, there is a subsequence $\mathbb{N}' \subset \mathbb{N}$ such that $\hat{\mathcal{I}}(u_\ell) \cap \mathcal{A}(u) = \emptyset$ for all $\ell \in \mathbb{N}'$. Hence, we find $D_\ell \subset \mathcal{A}$, $\text{meas}(D_\ell) = 0$ such that $D_\ell \subset \hat{\mathcal{I}}(u_\ell)$, $\ell \in \mathbb{N}'$. If $D := \bigcap_{\ell \in \mathbb{N}'} D_\ell$ is such that $\text{meas}(D) = 0$, for $w \in L^2(D)_+$, $w \equiv 0$, there is $\{w_\ell\}_{\mathbb{N}'}$ with $w_\ell \in L^2(\hat{\mathcal{I}}(u_\ell))W_\ell \cap L^2(\hat{\mathcal{I}}(u_\ell))$, $\ell \in \mathbb{N}'$, such that $w_\ell \rightarrow w$ in $L^2(\Omega)$ as $\ell \rightarrow \infty$ whence

$$\langle \sigma_\ell, w_\ell \rangle_{*, \hat{\mathcal{I}}(u_\ell)} \rightarrow \langle \sigma, w \rangle_{*, D} \quad \text{as } \ell \rightarrow \infty .$$

But $\langle \sigma_\ell, w_\ell \rangle_{*, \hat{\mathcal{I}}(u_\ell)} = 0$, $\ell \in \mathbb{N}'$, due to (2.4), and hence, $\langle \sigma, w \rangle_{*, D} = 0$ contradicting $\sigma|_D > 0$, $D \subset \mathcal{A}$ (cf. **(C)**). If $\text{meas}(D) = 0$, for the Hausdorff distance $d_H(\mathcal{F}, \partial \hat{\mathcal{N}}_\ell)$ we must have $d_H(\mathcal{F}, \partial \hat{\mathcal{N}}_\ell) \rightarrow 0$ as $\ell \rightarrow \infty$ whence $d_H(\mathcal{F}, \mathcal{F}_\ell) \rightarrow 0$ as $\ell \rightarrow \infty$, since otherwise we arrive at a contradiction to (6.5). Consequently, there exist $x_\ell \in \hat{\mathcal{N}}_\ell$, $\ell \in \mathbb{N}'$, such that $u_\ell(x_\ell) - \chi(x_\ell) \geq \varepsilon > 0$ and $\text{dist}(x_\ell, \mathcal{F}_\ell) \rightarrow 0$ as $\ell \rightarrow \infty$ contradicting **(N₃)**. \square

We are now in a position to prove an error reduction property. The essential ingredient is a refined quasi-orthogonality property that can be derived by a more subtle treatment of the terms $\alpha(\sigma_\ell - \sigma, u_\ell - u_{\ell+1})_{0, \Omega}$ and $\alpha(\sigma_{\ell+1} - \sigma, u - u_{\ell+1})_{0, \Omega}$ in (5.6) of the proof of Lemma 5.1.

Lemma 6.3. *Under the same assumptions as in Proposition 6.1, for any $\varepsilon > 0$ and $\ell \geq \ell^*$ there holds*

$$\alpha(\sigma_\ell - \sigma, u_\ell - u_{\ell+1})_{0, \Omega} \leq \alpha\varepsilon \left(\|u - u_\ell\|_{0, \Omega}^2 + \|\sigma - \sigma_{\ell+1}\|_{0, \Omega}^2 \right) \quad (6.7)$$

$$+ \alpha \left(\varepsilon + \frac{1}{4\varepsilon} \right) \|\sigma\|_{0, \hat{\mathcal{F}}(u_\ell)}^2 + \frac{1}{4\varepsilon} \left(\text{osc}_\ell^2(\psi) + \text{osc}_{\ell+1}^2(\psi) \right) ,$$

$$\alpha(\sigma_{\ell+1} - \sigma, u - u_{\ell+1})_{0, \Omega} \quad (6.8)$$

$$\leq \alpha\varepsilon \left(\|\sigma - \sigma_\ell\|_{0, \Omega}^2 + \|\sigma - \sigma_{\ell+1}\|_{0, \Omega}^2 \right) + \frac{1}{4\varepsilon} \left(\|\psi - u\|_{0, \hat{\mathcal{F}}(u_\ell)}^2 + \text{osc}_{\ell+1}^2(\psi) \right) .$$

Proof. Taking advantage of the complementarity conditions (1.3), (2.4), we obtain

$$\begin{aligned} \alpha(\sigma_{\ell+1} - \sigma, u_\ell - u_{\ell+1})_{0, \Omega} &= \alpha(\sigma_{\ell+1}, \psi_{\ell+1} - u_{\ell+1})_{0, \Omega} \quad (6.9) \\ &+ \alpha(\sigma_{\ell+1}, u_\ell - \psi_\ell)_{0, \Omega} + \alpha(\sigma, u_{\ell+1} - \psi_{\ell+1})_{0, \Omega} \\ &+ \alpha(\sigma_{\ell+1} - \sigma, \psi_\ell - \psi_{\ell+1})_{0, \Omega} + \alpha(\sigma, \psi_\ell - u_\ell)_{0, \Omega} \\ &\leq \alpha(\sigma, \psi_\ell - u_\ell)_{0, \Omega} + \alpha(\sigma_{\ell+1} - \sigma, \psi_\ell - \psi_{\ell+1})_{0, \Omega} , \end{aligned}$$

where we have used that the first term after the equality sign is zero, whereas the second and the term is non-positive.

The first term on the right-hand side is obviously zero in $\mathcal{I}(u)$ and in $\hat{\mathcal{A}}(u_\ell)$. Due to Proposition 6.1 we thus have

$$(\sigma, \psi_\ell - u_\ell)_{0, \Omega} = (\sigma, \psi_\ell - u_\ell)_{0, \hat{\mathcal{F}}(u_\ell)} .$$

We further get

$$(\sigma, \psi_H - u_H)_{0, \hat{\mathcal{F}}(u_\ell)} = (\sigma, u - u_H)_{0, \hat{\mathcal{F}}(u_\ell)} + (\sigma, \psi_H - \psi)_{0, \hat{\mathcal{F}}(u_\ell)} , \quad (6.10)$$

where we have used that $(\sigma, \psi - u)_{0, \hat{\mathcal{F}}(u_\ell)} = 0$ due to (1.3). Using (6.10) in (6.9) and applying Cauchy's and Young's inequality results in (6.7).

The proof of (6.8) is done by similar arguments. \square

In view of Lemma 6.3, we define

$$\mu_\ell(u, \sigma) := \left(\sum_{T \in \hat{\mathcal{F}}(u_\ell)} (\|\sigma\|_T^2 + \|\psi - u\|_T^2) \right)^{1/2} . \quad (6.11)$$

We assume that the sequence $\{\mathcal{T}_\ell\}_{\mathbb{N}}$ of triangulations, generated by (3.10), (3.11) and **(E)**, is such that there exists $0 \leq \rho_4 < 1$ satisfying

$$\mu_{\ell+1}^2(u, \sigma) \leq \rho_4 \mu_\ell^2(u, \sigma) \quad , \quad \ell \in \mathbb{N}^* . \quad (6.12)$$

Theorem 6.4. *Let (y, p, u, σ) and $(y_\ell, p_\ell, u_\ell, \sigma_\ell)$ be the solutions of (1.2a)–(1.2d) and (2.2a)–(2.2d) and let $\text{osc}_\ell, \mu_\ell(p), \mu_\ell(u, \sigma)$ be the data oscillations and data terms given by (3.12), (5.14) and (6.11). Assume that (2.10), **(L)**, **(C)**, **(N₁)** – **(N₃)** and (3.13), (5.15), (6.12) are satisfied. Then, there exist constants $0 \leq \rho_1 < 1$ and $\Lambda > 0$, depending on the data of the problem, the constants $\Theta_i, 1 \leq i \leq 2$, in the bulk criteria (3.10), (3.11) and on the shape regularity of the triangulations, such that*

$$\begin{pmatrix} \| \|z - z_{\ell+1}\| \|_{\kappa}^2 \\ \text{osc}_{\ell+1}^2 \\ \mu_{\ell+1}^2(p) \\ \mu_{\ell+1}^2(u, \sigma) \end{pmatrix} \leq \begin{pmatrix} \rho_1 & \Lambda & \Lambda & \Lambda \\ 0 & \rho_2 & 0 & 0 \\ 0 & 0 & \rho_3 & 0 \\ 0 & 0 & 0 & \rho_4 \end{pmatrix} \begin{pmatrix} \| \|z - z_\ell\| \|_{\kappa}^2 \\ \text{osc}_\ell^2 \\ \mu_\ell^2(p) \\ \mu_\ell^2(u, \sigma) \end{pmatrix} .$$

Proof. Using the results of Lemma 6.3, as in the proof of Theorem 5.3 we find constants $0 < \kappa_i < 1, 1 \leq i \leq 3$, and $0 \leq \rho_1 < 1$ such that for some $\Lambda > 0$

$$\| \|z - z_{\ell+1}\| \|_{\kappa}^2 \leq \rho_1 \| \|z - z_\ell\| \|_{\kappa}^2 + \Lambda (\text{osc}_\ell^2 + \mu_\ell^2(p) + \mu_\ell^2(u, \sigma)) . \quad \square$$

7. Numerical results

We provide numerical results that illustrate the performance of the adaptive finite element approximation for a distributed optimal control problems where the data are given as follows

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

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.

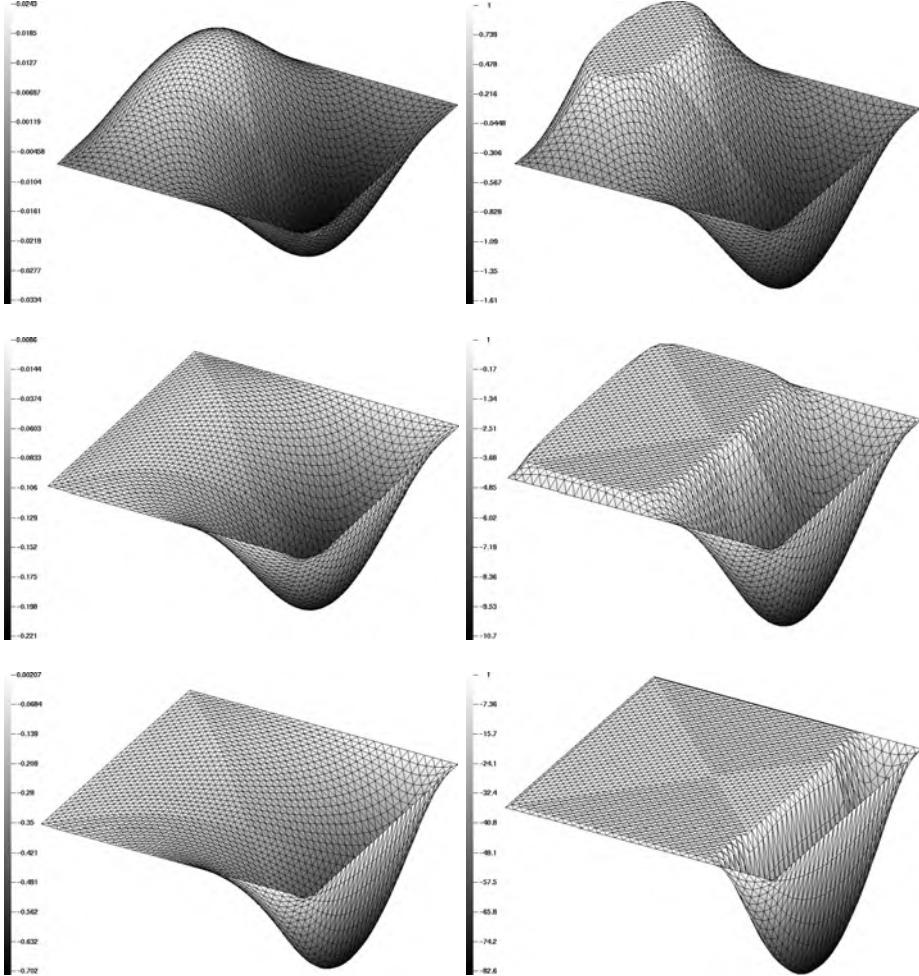


FIGURE 1. Visualization of the optimal state y and the optimal control u for $\alpha = 0.01$ (top), $\alpha = 0.001$ (middle) and $\alpha = 0.00001$ (bottom)

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 $\text{osc}_\ell(u^d) = \text{osc}_\ell(f) = \text{osc}_\ell(\psi) = 0, \ell \in \mathbb{N}_0$.

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

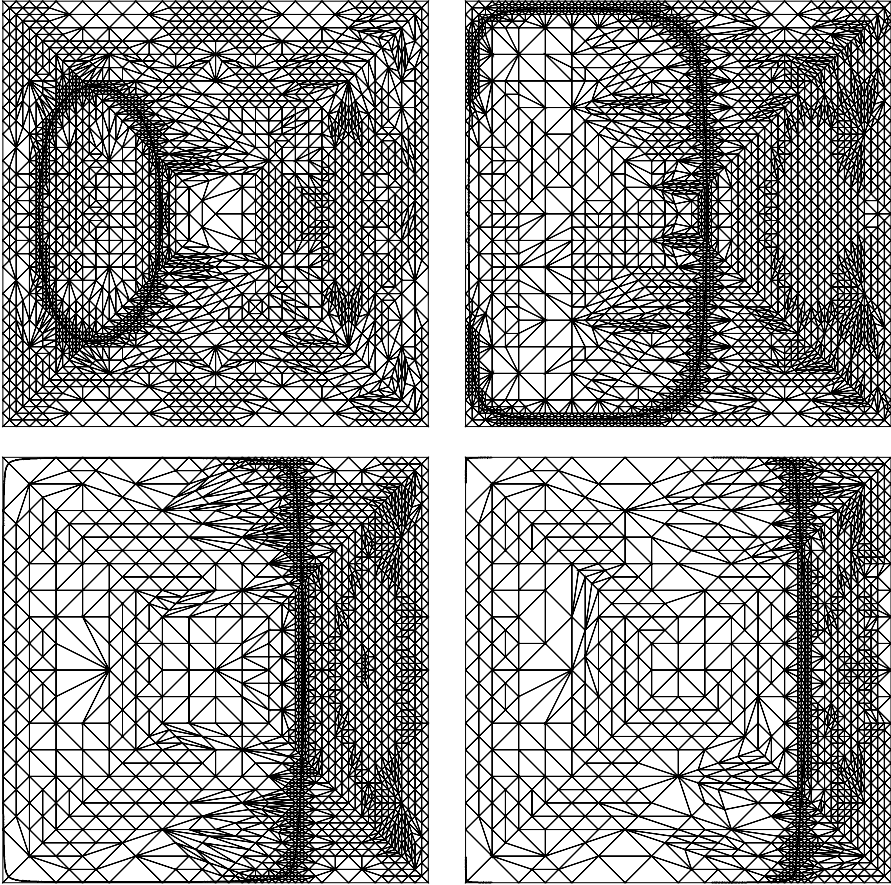


FIGURE 2. Adaptively generated grids after 6 refinement steps ($\alpha = 0.01$ (top left), $\alpha = 0.001$ (top right), $\alpha = 0.0001$ (bottom left) and $\alpha = 0.00001$ (bottom right))

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.

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

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

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

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

l	N_{dof}	$\ z - z_\ell\ $	$ y - y_\ell _1$	$ p - p_\ell _1$	$\ u - u_\ell\ _0$	$\ \sigma - \sigma_\ell\ _0$
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 ($\alpha = 0.01$)

l	N_{dof}	$\eta_{\text{v},\text{T},\ell}$	$\eta_{\text{p},\text{T},\ell}$	$\eta_{\text{v},\text{E},\ell}$	$\eta_{\text{p},\text{E},\ell}$	$\text{osc}_\ell(y^{\text{d}})$
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 ($\alpha = 0.01$)

l	N_{dof}	$M_{\eta,\text{T}}$	$M_{\eta,\text{E}}$	$M_{\text{osc},\text{E}}$	$M_{\text{fb},\text{E}}$
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

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, $M_{\eta,\text{T}}$ stands for the level l elements marked for refinement due to the element residuals and the data oscillations. On the other hand, $M_{\eta,\text{E}}$, $M_{\text{osc},\text{E}}$ and $M_{\text{fb},\text{E}}$ 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.

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, *Error reduction and convergence for an adaptive mixed finite element method*. Math. Comp. (2006) (in press).
- [11] C. Carstensen and R.H.W. Hoppe, *Convergence analysis of an adaptive nonconforming finite element method*. Numer. Math. (2006) (in press).
- [12] 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.
- [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*. Quarterly of Applied Mathematics **LXI** (2003), 131–161.

- [17] M. Hintermüller, R.H.W. Hoppe, Y. Iliash, and M. Kieweg, *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints*. to appear in ESAIM (2006).
- [18] J.B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*. Springer, Berlin-Heidelberg-New York, 1993.
- [19] 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.
- [20] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*. Springer, Berlin-Heidelberg-New York, 1971.
- [21] W. Liu and N. Yan, *A posteriori error estimates for distributed optimal control problems*. Adv. Comp. Math. **15** (2001), 285–309.
- [22] 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).
- [23] X.J. Li and J. Yong, *Optimal Control Theory for Infinite-Dimensional Systems*. Birkhäuser, Boston-Basel-Berlin, 1995.
- [24] P. Morin, R.H. Nochetto, and K.G. Siebert, *Data oscillation and convergence of adaptive FEM*. SIAM J. Numer. Anal. **38** (2000), 466–488.
- [25] P. Neittaanmäki and S. Repin, *Reliable methods for mathematical modelling. Error control and a posteriori estimates*. Elsevier, New York, 2004.
- [26] R. Verfürth, *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, New York, Stuttgart, 1996.

A. Gaevskaya, Y. Iliash and M. Kieweg
Institute of Mathematics
Universität Augsburg
D-86159 Augsburg, Germany
e-mail: gaevskaya@math.uni-augsburg.de
e-mail: iliash@math.uni-augsburg.de
e-mail: Michael.Kieweg@gmx.de

R.H.W. Hoppe
Department of Mathematics
University of Houston
Houston, TX 77204-3008, USA
e-mail: rohop@math.uh.edu