Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems

Ronald H. W. Hoppe, Michael Kieweg

Angaben zur Veröffentlichung / Publication details:

Hoppe, Ronald H. W., and Michael Kieweg. 2008. "Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems." *Computational Optimization and Applications* 46 (3): 511–33. https://doi.org/10.1007/s10589-008-9195-4.

Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems

R.H.W. Hoppe · M. Kieweg

Abstract Mixed control-state constraints are used as a relaxation of originally state constrained optimal control problems for partial differential equations to avoid the intrinsic difficulties arising from measure-valued multipliers in the case of pure state constraints. In particular, numerical solution techniques known from the pure control constrained case such as active set strategies and interior-point methods can be used in an appropriately modified way. However, the residual-type a posteriori error estimators developed for the pure control constrained case can not be applied directly. It is the essence of this paper to show that instead one has to resort to that type of estimators known from the pure state constrained case. Up to data oscillations and consistency error terms, they provide efficient and reliable estimates for the discretization errors in the state, a regularized adjoint state, and the control. A documentation of numerical results is given to illustrate the performance of the estimators.

1 Introduction

Adaptive finite element methods based on reliable a posteriori error estimators are powerful algorithmic tools for the efficient numerical solution of boundary and initial-boundary value problems for partial differential equations (PDEs) (cf. [1, 3, 4, 10, 25, 31] and the references therein). On the other hand, considerably less work

R.H.W. Hoppe (✉) · M. Kieweg
Institute of Mathematics, University of Augsburg, 86159 Augsburg, Germany
e-mail: ronald.hoppe@math.uni-augsburg.de

R.H.W. Hoppe · M. Kieweg
Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA

has been done with regard to optimal control problems for PDEs. The so-called goal oriented dual weighted approach has been applied in the unconstrained case in [4, 5], to control constrained problems in [14, 32] and to state constrained problems in [13], whereas residual-type a posteriori error estimators for control constrained problems have been derived and analyzed in [11, 12, 15, 17, 19–21]. Unlike the control constrained case, pointwise state constrained optimal control problems are much more difficult to handle due to the fact that the Lagrange multiplier for the state constraints lives in a measure space (see, e.g., [6, 7, 16, 29]). Therefore, it is a natural idea to regularize state constrained problems by means of mixed control-state constrained ones, since with regard to numerical solution techniques the regularized problems can be formally treated as in the case of pure control constraints (cf. e.g., [2, 8, 23, 26–30]). This method is called Lavrentiev regularization. However, so far an a posteriori error analysis of adaptive finite element approximations has not been provided for mixed control-state constrained control problems.

In this paper, we will be concerned with the development, analysis and implementation of a residual-type a posteriori error estimator for mixed control-state constrained distributed optimal control problems for linear second order elliptic boundary value problems. The paper is organized as follows: In Sect. 2, as a model problem we consider a distributed optimal control problem for a two-dimensional, second order elliptic PDE with a quadratic objective functional and mixed unilateral constraints on the state and on the control. The optimality conditions are stated in terms of the state, the adjoint state, the control, and a Lagrangian multiplier for the mixed constraints. However, the a posteriori error analysis known from pure control constraints can not be applied in a meaningful way to the mixed control-state constrained case, since the constants involved in the reliability estimates blow up as the regularization parameter tends to zero. Instead, one has to adopt the error analysis as developed for the pure state constrained case [18]. In this spirit, we further consider a regularized multiplier and a regularized adjoint state which will play an essential role in the error analysis. In Sect. 3, we describe the finite element discretization of the control problem with respect to a family of shape regular simplicial triangulations of the computational domain using continuous, piecewise linear finite elements for the state, the control, and for the adjoint and the regularized adjoint state. In Sect. 4, we present the residual-type a posteriori error estimator for the global discretization errors in the state, the regularized adjoint state, and the control. Data oscillations and a consistency error are considered as well, since they enter the subsequent error analysis. In Sect. 5, we prove reliability of the error estimator, i.e., we show that it gives rise to an upper bound for the global discretization errors up to data oscillations and a consistency error. A computable upper bound for the consistency error is derived based on suitable approximations of the continuous coincidence and non-coincidence sets. Section 6 is devoted to the efficiency of the estimator by showing that, modulo data oscillations, the error estimator also provides a lower bound for the discretization errors. Finally, Sect. 7 contains a documentation of numerical results for two representative test examples in terms of the convergence history of the adaptive finite element process.

2 The mixed control-state constrained distributed control problem

Let Ω be a bounded domain in $\mathbb{R}^2$ with boundary $\Gamma := \Gamma_D \cup \Gamma_N, \Gamma_D \cap \Gamma_N = \emptyset$. We use standard notation from Lebesgue and Sobolev space theory. In particular, we refer to $L^2(\Omega)$ as the Hilbert space of square integrable functions with inner product $(\cdot,\cdot)_{0,\Omega}$ and associated norm $\|\cdot\|_{0,\Omega}$. We denote by $L^2_+(\Omega)$ the non-negative cone of $L^2(\Omega)$, i.e., $L^2_+(\Omega) := \{v \in L^2(\Omega) | v(x) \geq 0 \text{ f.a.a. } x \in \Omega\}$. Moreover, $H^k(\Omega), k \in \mathbb{N}$, stands for the Sobolev space of square integrable functions whose weak derivatives up to order k are square integrable as well, equipped with the norm $\|\cdot\|_{k,\Omega}$. $H_0^k(\Omega)$ denotes its subspace $H_0^k(\Omega) := \{v \in H^k(\Omega) | D^\alpha v|_\Gamma = 0, |\alpha| \leq k-1\}$ and $H^{-k}(\Omega)$ is the dual of $H_0^k(\Omega)$.

For given $c \in \mathbb{R}_+$, we refer to $A : V \to H^{-1}(\Omega)$, $V := \{v \in H^1(\Omega) \mid v|_{\Gamma_D} = 0\}$, as the linear second order elliptic differential operator

$$Ay := -\Delta y + cy, \quad y \in V$$

and to $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ with $a(y,v) := \int_\Omega (\nabla y \cdot \nabla v + cyv)dx$ as the associated bilinear form. We assume $c > 0$ or meas$(\Gamma_D) \neq 0$. In particular, this assures that A is bounded and V-elliptic, i.e., there exist constants $C > 0$ and $\gamma > 0$ such that

$$|a(y,v)| \leq C\|y\|_{1,\Omega}\|v\|_{1,\Omega}, \qquad a(y,y) \geq \gamma\|y\|^2_{1,\Omega}. \tag{2.1}$$

Now, given a desired state $y^d \in L^2(\Omega)$, a shift control $u^d \in L^2(\Omega)$, regularization parameters $\alpha > 0$, $\varepsilon > 0$, and a function $\psi \in L^\infty(\Omega)$, we consider the objective functional

$$J(y,u) := \frac{1}{2}\|y - y^d\|^2_{0,\Omega} + \frac{\alpha}{2}\|u - u^d\|^2_{0,\Omega} \tag{2.2}$$

and the mixed control-state constrained distributed optimal control problem: Find $(y,u) \in V \times L^2(\Omega)$ such that

$$\inf_{y,u} J(y,u), \tag{2.3a}$$

subject to the constraints

$$Ay = u \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \Gamma_D, \qquad \boldsymbol{\nu} \cdot \nabla y = 0 \quad \text{on } \Gamma_N, \tag{2.3b}$$

$$\varepsilon u + y \in K := \{v \in L^2(\Omega) \mid v(x) \leq \psi(x) \text{ f.a.a. } x \in \Omega\}. \tag{2.3c}$$

The usual way to look at (2.3a)–(2.3c) is as a regularized state constrained problem, since the multiplier associated with the inequality constraint (2.3c), usually called the adjoint control for control constrained problems, lives in the non-negative cone of $L^2(\Omega)$ and not in a measure space as in the case of pure state constraints. Obviously, the latter case is much more difficult to handle.

We define $G : L^2(\Omega) \to V$ as the control-to-state map which assigns to $u \in L^2(\Omega)$ the unique solution $y = y(u)$ of (2.3b). We note that the control-to-state map G is a bounded linear operator. Substituting the state $y = y(u)$ by $y(u) = Gu$ leads to the reduced objective functional

$$J_{\mathrm{red}}(u) := \frac{1}{2}\|Gu - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2}\|u - u^d\|_{0,\Omega}^2, \tag{2.4}$$

and the mixed control-state constrained problem (2.3a)–(2.3c) can be restated as

$$\inf_{\varepsilon u + Gu \in K} J_{\mathrm{red}}(u). \tag{2.5}$$

Standard arguments from convex optimization reveal the existence and uniqueness of a solution. The optimality conditions for the optimal solution $(y, u) \in V \times L^2(\Omega)$ are as follows.

Theorem 2.1 *The optimal solution $(y, u) \in V \times L^2(\Omega)$ of (2.3) is characterized by the existence of an adjoint state $p \in V$, and a multiplier $\sigma \in L^2_+(\Omega)$ such that*

$$(\nabla y, \nabla v)_{0,\Omega} + (cy, v)_{0,\Omega} = (u, v)_{0,\Omega}, \quad v \in V, \tag{2.6a}$$

$$(\nabla p, \nabla w)_{0,\Omega} + (cp, w)_{0,\Omega} = (y - y^d, w)_{0,\Omega} + (\sigma, w)_{0,\Omega}, \quad w \in V, \tag{2.6b}$$

$$p + \alpha(u - u^d) + \varepsilon\sigma = 0, \tag{2.6c}$$

$$(\sigma, \varepsilon u + y - \psi)_{0,\Omega} = 0. \tag{2.6d}$$

The residual-type a posteriori error analysis known from the pure control constrained case [15] is not applicable to (2.6a)–(2.6d) uniformly in the regularization parameter ε, since the constants in the reliability and efficiency estimates for the associated error estimator depend on ε in the sense that they blow up as $\varepsilon \to 0$. Indeed, denoting by $(y_\ell, u_\ell, p_\ell, \sigma_\ell)$ P1 conforming finite element approximations of (y, u, p, σ), by $\eta_\ell, osc_\ell(u^d), osc_\ell(\psi)$ the residual-type a posteriori error estimator and the data oscillations from [15] and mimicking the proof of Theorem 5.1 in [15], we get the reliability estimate

$$\begin{aligned} &|y - y_\ell|_{1,\Omega} + |p - p_\ell|_{1,\Omega} + \|u - u_\ell\|_{0,\Omega} + \|\sigma - \sigma_\ell\|_{0,\Omega} \\ &\quad \le C(\varepsilon)\big(\eta_\ell + osc_\ell(u^d) + osc_\ell(\psi)\big) \end{aligned}$$

with a positive constant $C(\varepsilon) = O(1 + \varepsilon^{-2})$ (using the complementarity conditions in the mixed control-state constrained case we have to estimate $u - u_\ell$ via $u - u_\ell = \varepsilon^{-1}(\varepsilon u + y - (\varepsilon u_\ell + y_\ell)) - \varepsilon^{-1}(y - y_\ell)$). Therefore, the a posteriori error analysis has to be adopted to that what is known from the pure state constrained case. Following [18], we define a regularized multiplier $\overline{\sigma} \in V$ as the solution of

$$(\nabla\overline{\sigma}, \nabla v)_{0,\Omega} + (c\overline{\sigma}, v)_{0,\Omega} = (\sigma, v)_{0,\Omega}, \quad v \in V, \tag{2.7}$$

and further introduce a regularized adjoint state $\overline{p} \in V$ according to

$$(\nabla\overline{p}, \nabla v)_{0,\Omega} + (c\overline{p}, v)_{0,\Omega} = (y - y^d, v)_{0,\Omega}, \quad v \in V. \tag{2.8}$$

Obviously, p, $\overline{p}$, and $\overline{\sigma}$ are related by

$$p := \overline{p} + \overline{\sigma}. \tag{2.9}$$

3 Finite element approximation

We consider a family $\{\mathcal{T}_\ell(\Omega)\}$ of shape-regular simplicial triangulations of Ω which align with Γ_D, Γ_N on Γ. We denote by $\mathcal{N}_\ell(D)$ and $\mathcal{E}_\ell(D)$, $D \subseteq \overline{\Omega}$, the sets of vertices and edges of $\mathcal{T}_\ell(\Omega)$ in $D \subseteq \overline{\Omega}$, and we refer to h_T and $|T|$ as the diameter and the area of an element $T \in \mathcal{T}_\ell(\Omega)$, whereas h_E stands for the length of an edge $E \in \mathcal{E}_\ell(D)$. For $E \in \mathcal{E}_\ell(\Omega)$ such that $E = T_+ \cap T_-$, $T_\pm \in \mathcal{T}_\ell(\Omega)$, we define $\omega_E := T_+ \cup T_-$. For $T \in \mathcal{T}_\ell(\Omega)$ and $E \in \mathcal{E}_\ell(\Omega)$ we further denote by $\lambda_i(T)$, $1 \le i \le 3$, and $\lambda_i(E)$, $1 \le i \le 2$, the barycentric coordinates associated with the vertices of T and E, respectively. We will also use the following notation: If A and B are two quantities, then $A \lesssim B$ means that there exists a positive constant C such that $A \le CB$, where C only depends on the shape regularity of the triangulations, but not on their granularities.

In order to get a useful discrete optimality system, we use an approach which first optimizes and then discretizes, i.e., we discretize the continuous optimality conditions (2.6). We first equivalently reformulate these conditions by means of the active set $\mathcal{A} \subset \Omega$ and the inactive set $\mathcal{I}$. We define the active set $\mathcal{A}$ as the maximal open set $A \subset \Omega$ such that $\varepsilon u(x) + y(x) + \sigma(x) > \psi(x)$ f.a.a. $x \in A$ and the inactive set $\mathcal{I}$ according to $\mathcal{I} := \bigcup_{\rho > 0} B_\rho$, where B_ρ is the maximal open set $B \subset \Omega$ such that $\varepsilon u(x) + y(x) \le \psi(x) - \rho$ for almost all $x \in B$. One can easily verify (cf. [22]) that the optimality system (2.6) is then equivalent to

$$(\nabla y, \nabla v)_{0,\Omega} + c(y, v)_{0,\Omega} = (u, v)_{0,\Omega} \quad \forall v \in V, \tag{3.1a}$$

$$(\nabla p, \nabla v)_{0,\Omega} + c(p, v)_{0,\Omega} = (y + \sigma - y^d, v)_{0,\Omega} \quad \forall v \in V, \tag{3.1b}$$

$$p + \alpha(u - u^d) + \varepsilon\sigma = 0 \quad \text{a.e. in } \Omega, \tag{3.1c}$$

$$\sigma = 0 \quad \text{a.e. in } \mathcal{I}, \qquad \varepsilon u + y = \psi \quad \text{a.e. in } \mathcal{A}. \tag{3.1d}$$

Now, we discretize (3.1) by conforming P1 finite elements. In particular, we refer to $S_\ell := \{v_\ell \in C_0(\Omega) \mid v_\ell|_T \in P_1(T), T \in \mathcal{T}_\ell(\Omega)\}$ as the finite element space spanned by the canonical nodal basis functions φ_ℓ^p, $p \in \mathcal{N}_\ell(\Omega)$, associated with the nodal points in $\overline{\Omega}$ and to V_ℓ as its subspace $V_\ell := \{v_\ell \in S_\ell \mid v_\ell|_{\Gamma_D} = 0\}$. Assume that $\{n_1, \ldots, n_N\} = \mathcal{N}(\overline{\Omega})$ are the numbered nodes of the triangulation. Similar to [22], the discrete solution consists of $(y_\ell, u_\ell, p_\ell, \sigma_\ell) \in V_\ell \times S_\ell \times V_\ell \times S_\ell$ and two subsets of indexes $\mathcal{A}_\ell$ and $\mathcal{I}_\ell$ such that there holds

$$(\nabla y_\ell, \nabla v_\ell)_{0,\Omega} + c(y_\ell, v_\ell)_{0,\Omega} = (u_\ell, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell, \tag{3.2a}$$

$$(\nabla p_\ell, \nabla v_\ell)_{0,\Omega} + c(p_\ell, v_\ell)_{0,\Omega} = (y_\ell - y^d + \sigma_\ell, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell, \tag{3.2b}$$

$$p_\ell + \alpha(u_\ell - u_\ell^d) + \varepsilon\sigma_\ell = 0, \tag{3.2c}$$

$$\sigma_\ell(n_i) = 0 \quad \text{if } i \in \mathcal{I}_\ell, \tag{3.2d}$$

$$\varepsilon u_\ell(n_i) + y_\ell(n_i) = \psi_\ell(n_i) \quad \text{if } i \in \mathcal{A}_\ell,$$

where $\mathcal{A}_\ell$ and $\mathcal{I}_\ell$ are given by

$$\mathcal{A}_\ell := \big\{i \in \{1, \ldots, N\} \mid \varepsilon u_\ell(n_i) + y_\ell(n_i) + \sigma_\ell(n_i) > \psi_\ell(n_i)\big\}, \tag{3.3a}$$

$$\mathcal{I}_\ell := \{1, \ldots, N\} \setminus \mathcal{A}_\ell. \tag{3.3b}$$

Such a discrete solution can, for example, be achieved by a primal-dual active set strategy (cf. [22]).

As in Sect. 2 before, we introduce a regularized discrete multiplier $\overline{\sigma}_\ell \in V_\ell$ as the solution of

$$(\nabla\overline{\sigma}_\ell, \nabla v_\ell)_{0,\Omega} + (c\overline{\sigma}_\ell, v_\ell)_{0,\Omega} = (\sigma_\ell, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell, \tag{3.4}$$

and define $\overline{p}_\ell \in V_\ell$ as the solution of the discrete analogue of (2.8), i.e.,

$$(\nabla\overline{p}_\ell, \nabla v_\ell)_{0,\Omega} + (c\overline{p}_\ell, v_\ell)_{0,\Omega} = (y_\ell - y^d, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell. \tag{3.5}$$

As in the continuous setting, we obtain the fundamental relationship

$$p_\ell = \overline{p}_\ell + \overline{\sigma}_\ell. \tag{3.6}$$

We further define $\mathcal{A}_\ell^T$ and $\mathcal{I}_\ell^T$ as the discrete active and inactive sets according to

$$\mathcal{A}_\ell^T := \bigcup\big\{T \in \mathcal{T}_\ell(\Omega) \mid \varepsilon u_\ell(p) + y_\ell(p) = \psi_\ell(p),\, p \in \mathcal{N}_\ell(T)\big\},$$

$$\mathcal{I}_\ell^T := \bigcup\big\{T \in \mathcal{T}_\ell(\Omega) \mid \varepsilon u_\ell(p) + y_\ell(p) < \psi_\ell(p),\, p \in \mathcal{N}_\ell(T)\big\}.$$

The discrete free boundary is given by

$$\mathcal{F}_\ell^T := \mathcal{T}_\ell(\Omega) \setminus \{\mathcal{A}_\ell^T \cup \mathcal{I}_\ell^T\}.$$

With these definitions at hand and taking (3.2d) into account, it is easy to see that there holds

$$(\sigma_\ell, \varepsilon u_\ell + y_\ell - \psi_\ell)_{0,\Omega} = \sum_{T\in\mathcal{F}_\ell^T} (\sigma_\ell, \varepsilon u_\ell + y_\ell - \psi_\ell)_{0,T}. \tag{3.7}$$

4 The residual-type error estimator

The residual-type a posteriori error estimator involves estimators of the state y and of the regularized adjoint state $\bar{p}$ according to

$$\eta_\ell := \eta_\ell(y) + \eta_\ell(\overline{p}). \tag{4.1}$$

where $\eta_\ell(y)$ and $\eta_\ell(\overline{p})$ consist of element and edge residuals

$$\eta_\ell(y) := \bigg(\sum_{T\in\mathcal{T}_\ell(\Omega)} \eta_T^2(y) + \sum_{E\in\mathcal{E}_\ell(\Omega)} \eta_E^2(y)\bigg)^{1/2}, \tag{4.2a}$$

$$\eta_\ell(\bar{p}) := \Big(\sum_{T \in \mathcal{T}_\ell(\Omega)} \eta_T^2(\bar{p}) + \sum_{E \in \mathcal{E}_\ell(\Omega)} \eta_E^2(\bar{p}) \Big)^{1/2}. \tag{4.2b}$$

The element residuals $\eta_T(y)$ and $\eta_T(\overline{p}), T \in \mathcal{T}_\ell(\Omega)$, are weighted elementwise L^2-residuals with respect to the strong form of the state equation (2.3b) and the modified adjoint state equation (2.8), respectively:

$$\eta_T(y) := h_T \|c y_\ell - u_\ell\|_{0,T}, \quad T \in \mathcal{T}_\ell(\Omega), \tag{4.3a}$$

$$\eta_T(\overline{p}) := h_T \|c \overline{p}_\ell - (y_\ell - y^d)\|_{0,T}, \quad T \in \mathcal{T}_\ell(\Omega). \tag{4.3b}$$

The edge residuals $\eta_E(y)$ and $\eta_E(\bar{p}), E \in \mathcal{E}_\ell(\Omega)$, are weighted L^2-norms of the jumps $\nu_E \cdot [\nabla y_\ell]$ and $\nu_E \cdot [\nabla \overline{p}_\ell]$ of the normal derivatives across the interior edges

$$\eta_E(y) := h_E^{1/2} \|\nu_E \cdot [\nabla y_\ell]\|_{0,E}, \quad E \in \mathcal{E}_\ell(\Omega), \tag{4.4a}$$

$$\eta_E(\overline{p}) := h_E^{1/2} \|\nu_E \cdot [\nabla \overline{p}_\ell]\|_{0,E}, \quad E \in \mathcal{E}_\ell(\Omega). \tag{4.4b}$$

Denoting by $y_\ell^d \in S_\ell$ some approximation of the desired state y^d, we further have to take into account data oscillations with respect to the data u^d, y^d, ψ of the problem

$$osc_\ell := \big(osc_\ell^2(u^d) + osc_\ell^2(y^d) + osc_\ell^2(\psi)\big)^{1/2}, \tag{4.5}$$

where $osc_\ell(u^d)$, $osc_\ell(y^d)$, and $osc_\ell(\psi)$ are given by

$$osc_\ell(u^d) := \Big(\sum_{T \in \mathcal{T}_\ell(\Omega)} osc_T^2(u^d) \Big)^{1/2}, \tag{4.6a}$$

$$osc_T(u^d) := \|u^d - u_\ell^d\|_{0,T}, \quad T \in \mathcal{T}_\ell(\Omega),$$

$$osc_\ell(y^d) := \Big(\sum_{T \in \mathcal{T}_\ell(\Omega)} osc_T^2(y^d) \Big)^{1/2}, \tag{4.6b}$$

$$osc_T(y^d) := h_T \|y^d - y_\ell^d\|_{0,T}, \quad T \in \mathcal{T}_\ell(\Omega),$$

$$osc_\ell(\psi) := \Big(\sum_{T \in \mathcal{T}_\ell(\Omega)} osc_T^2(\psi) \Big)^{1/2}, \tag{4.6c}$$

$$osc_T(\psi) := \|\psi - \psi_\ell\|_{0,T}, \quad T \in \mathcal{T}_\ell(\Omega).$$

We will show that, up to data oscillations and consistency errors, the residual-type a posteriori error estimator (4.1) provides an upper and a lower bound for the discretization errors in the state, the regularized adjoint state, and the control which are given by

$$e_y := y - y_\ell, \qquad e_{\overline{p}} := \overline{p} - \overline{p}_\ell, \qquad e_u := u - u_\ell. \tag{4.7}$$

In much the same way as in case of adaptive finite element discretizations of pure state constrained elliptic boundary value problems (cf. [18]), the a posteriori error analysis requires an auxiliary state $y(u_\ell) \in V$ and an auxiliary adjoint state $\overline{p}(y_\ell) \in V$. These are defined as the solutions of the following variational equations

$$(\nabla y(u_\ell), \nabla v)_{0,\Omega} + (cy(u_\ell), v)_{0,\Omega} = (u_\ell, v)_{0,\Omega}, \quad v \in V, \tag{4.8a}$$

$$(\nabla \overline{p}(y_\ell), \nabla v)_{0,\Omega} + (c\overline{p}(y_\ell), v)_{0,\Omega} = (y_\ell - y^d, v)_{0,\Omega}, \quad v \in V. \tag{4.8b}$$

We further introduce an auxiliary discrete state $y_\ell(u) \in V_\ell$ as the solution of the finite dimensional variational problem

$$(\nabla y_\ell(u), \nabla v_\ell)_{0,\Omega} + (cy_\ell(u), v_\ell)_{0,\Omega} = (u, v_\ell)_{0,\Omega}, \quad v_\ell \in V_\ell. \tag{4.9}$$

The auxiliary states $y(u_\ell) \in V$ and $y_\ell(u) \in V_\ell$ do not necessarily satisfy the mixed control-state constraints, i.e., it may happen that $\varepsilon u_\ell + y(u_\ell) \notin K$ and $\varepsilon u + y_\ell(u) \notin K_\ell$. This gives rise to the consistency error (cf. Lemma 5.3)

$$\begin{aligned} e_{c,\ell} := \max\big((\sigma_\ell, \varepsilon(u - u_\ell) + y_\ell(u) - y_\ell)_{0,\Omega} \\ + (\sigma, \varepsilon(u_\ell - u) + y(u_\ell) - y)_{0,\Omega}, 0\big). \end{aligned} \tag{4.10}$$

The consistency error is not an a posteriori term, since neither u, $y_\ell(u)$ nor $y(u_\ell)$ are available. However, at the end of Sect. 5 we will derive a computable upper bound which shows that it can be controlled by the residual estimator $\eta_\ell(y)$ and the data oscillation $osc_\ell(\psi)$. Therefore, it will not be included in the refinement strategy except that we refine all triangles within the discrete free boundary $\mathcal{F}_\ell^T$ to reduce the mismatch between the continuous and discrete coincidence and non-coincidence sets.

The refinement of a triangulation $\mathcal{T}_\ell(\Omega)$ is based on bulk criteria that have been previously used in the convergence analysis of adaptive finite element for nodal finite element methods [9, 24]. For the mixed control-state constrained optimal control problem under consideration, the bulk criteria are as follows: Given universal constants $\Theta_i \in (0, 1)$, $1 \le i \le 5$, we create a set of edges $\mathcal{M}_E \subset \mathcal{E}_h(\Omega)$ and sets of elements $\mathcal{M}_{\eta,T}, \mathcal{M}_{u^d,T}, \mathcal{M}_{y^d,T}, \mathcal{M}_{\psi,T} \subset \mathcal{T}_\ell(\Omega)$ such that

$$\Theta_1 \sum_{E \in \mathcal{E}_\ell(\Omega)} \big(\eta_E^2(y) + \eta_E^2(\bar{p})\big) \le \sum_{E \in \mathcal{M}_E} \big(\eta_E^2(y) + \eta_E^2(\bar{p})\big),$$

$$\Theta_2 \sum_{T \in \mathcal{T}_\ell(\Omega)} \big(\eta_T^2(y) + \eta_T^2(\bar{p})\big) \le \sum_{T \in \mathcal{M}_{\eta,T}} \big(\eta_T^2(y) + \eta_T^2(\bar{p})\big),$$

$$\Theta_3 \sum_{T \in \mathcal{T}_\ell(\Omega)} osc_T^2(u^d) \le \sum_{T \in \mathcal{M}_{u^d,T}} osc_T^2(u^d),$$

$$\Theta_4 \sum_{T \in \mathcal{T}_\ell(\Omega)} osc_T^2(y^d) \le \sum_{T \in \mathcal{M}_{y^d,T}} osc_T^2(y^d),$$

$$\Theta_5 \sum_{T \in \mathcal{T}_\ell(\Omega)} osc_T^2(\psi) \le \sum_{T \in \mathcal{M}_{\psi,T}} osc_T^2(\psi).$$

The bulk criteria are realized by a greedy algorithm (cf., e.g., [15]). We set

$$\mathcal{M}_T := \mathcal{M}_{\eta,T} \cup \mathcal{M}_{u^d,T} \cup \mathcal{M}_{y^d,T} \cup \mathcal{M}_{\psi,T}$$

and refine an element $T \in \mathcal{T}_\ell(\Omega)$ by bisection (i.e., by joining the midpoint of the longest edge with the opposite vertex), if $T \in \mathcal{M}^T$ and an edge $E \in \mathcal{E}_\ell(T)$ by bisection (joining its midpoint with the opposite vertices of the adjacent elements), if $E \in \mathcal{M}^E$. We further extend the refinement strategy in so far as we refine all triangles within the discrete free boundary $\mathcal{F}_\ell^T$.

5 Reliability of the error estimator

We prove reliability of the residual-type error estimator (4.1) in the sense that it provides an upper bound for the discretization errors e_y, e_u, and $e_{\overline{p}}$ up to the data oscillations $osc_\ell(u^d)$ and $osc_\ell(\psi)$ and the consistency error $e_{c,\ell}$.

Theorem 5.1 *Let* (y, u, p, σ) *and* $(y_\ell, u_\ell, p_\ell, \sigma_\ell)$ *be the solutions of* (2.6a)–(2.6d) *and* (3.2a)–(3.2d) *and let* η_ℓ, $osc_\ell(u^d)$ *and* $e_{c,\ell}$ *be the error estimator, the data oscillation in the shift control, and the consistency error according to* (4.1), (4.6) *and* (4.10), *respectively. Further, let* $\overline{p}$ *and* $\overline{p}_\ell$ *be the regularized adjoint states as given by* (2.8), (3.5). *Then, there holds*

$$\|e_y\|_{1,\Omega}^2 + \|e_u\|_{0,\Omega}^2 + \|e_{\overline{p}}\|_{1,\Omega}^2 \lesssim \eta_\ell^2 + osc_\ell^2(u^d) + e_{c,\ell}. \tag{5.1}$$

The proof of Theorem 5.1 will be given by the following three Lemmas 5.2, 5.3 and 5.4.

Lemma 5.2 *In addition to the assumptions of Theorem* 5.1 *let* $y(u_\ell)$ *and* $\overline{p}(y_\ell)$ *be the auxiliary state and auxiliary adjoint state according to* (4.8a), (4.8b). *Then, there holds*

$$\|e_y\|_{1,\Omega}^2 + \|e_{\overline{p}}\|_{1,\Omega}^2 \lesssim \|y(u_\ell) - y_\ell\|_{1,\Omega}^2 + \|\overline{p}(y_\ell) - \overline{p}_\ell\|_{1,\Omega}^2 + \|e_u\|_{0,\Omega}^2. \tag{5.2}$$

Proof Obviously, e_y and $e_{\overline{p}}$ can be estimated from above by

$$\|e_y\|_{1,\Omega}^2 \le 2\|y - y(u_\ell)\|_{1,\Omega}^2 + 2\|y(u_\ell) - y_\ell\|_{1,\Omega}^2, \tag{5.3a}$$

$$\|e_{\overline{p}}\|_{1,\Omega}^2 \le 2\|\overline{p} - \overline{p}(y_\ell)\|_{1,\Omega}^2 + 2\|\overline{p}(y_\ell) - \overline{p}_\ell\|_{1,\Omega}^2. \tag{5.3b}$$

Setting $v = y - y(u_\ell)$ in (2.6a), (4.8a), and $M := 1/\gamma$ with γ from (2.1), for the first term on the right-hand side in (5.3a) we readily get

$$\|y - y(u_\ell)\|_{1,\Omega}^2 \le M\|e_u\|_{0,\Omega}\|y - y(u_\ell)\|_{0,\Omega} \le M\|e_u\|_{0,\Omega}\|y - y(u_\ell)\|_{1,\Omega},$$

and hence,

$$\|y - y(u_\ell)\|_{1,\Omega}^2 \le M^2\|e_u\|_{0,\Omega}^2. \tag{5.4}$$

Likewise, choosing $v = \overline{p} - \overline{p}(y_\ell)$ in (2.8) and (4.8b), for the first term on the right-hand side in (5.3b) it follows that

$$\|\overline{p} - \overline{p}(y_\ell)\|_{1,\Omega}^2 \leq M\|e_y\|_{0,\Omega}\|\overline{p} - \overline{p}(y_\ell)\|_{0,\Omega} \leq M\|e_y\|_{1,\Omega}\|\overline{p} - \overline{p}(y_\ell)\|_{1,\Omega}.$$

Consequently, in view of (5.3a) and (5.4) we obtain

$$\|\overline{p} - \overline{p}(y_\ell)\|_{1,\Omega}^2 \ \leq \ M^2\|e_y\|_{1,\Omega}^2 \leq 2M^4\|e_u\|_{0,\Omega}^2 + 2M^2\|y(u_\ell) - y_\ell\|_{1,\Omega}^2. \tag{5.5}$$

Using (5.4), (5.5) in (5.3a), (5.3b) gives (5.2). □

Lemma 5.3 *Under the same assumptions as in Lemma 5.2 there holds*

$$\|e_u\|_{0,\Omega}^2 \leq \frac{4}{\alpha}\left(\frac{M^2}{\alpha}\|y(u_\ell) - y_\ell\|_{1,\Omega}^2 + \frac{1}{\alpha}\|\overline{p}(y_\ell) - \overline{p}_\ell\|_{1,\Omega}^2 + e_{c,\ell}\right) + osc_\ell^2(u^d). \tag{5.6}$$

Proof Using (2.6c), (2.9) and (3.2c), (3.6), we find

$$\begin{aligned}\alpha\|e_u\|_{0,\Omega}^2 = (e_u, \overline{p}_\ell - \overline{p})_{0,\Omega} + \varepsilon(e_u, \sigma_\ell - \sigma)_{0,\Omega} \\ + (e_u, \overline{\sigma}_\ell - \overline{\sigma})_{0,\Omega} + \alpha(e_u, u^d - u_\ell^d)_{0,\Omega}.\end{aligned} \tag{5.7}$$

The first term on the right-hand side in (5.7) can be split according to

$$(e_u, \overline{p}_\ell - \overline{p})_{0,\Omega} = (e_u, \overline{p}_\ell - \overline{p}(y_\ell))_{0,\Omega} + (e_u, \overline{p}(y_\ell) - \overline{p})_{0,\Omega}. \tag{5.8}$$

An application of Young's inequality yields

$$(e_u, \overline{p}_\ell - \overline{p}(y_\ell))_{0,\Omega} \leq \frac{\alpha}{4}\|e_u\|_{0,\Omega}^2 + \frac{1}{\alpha}\|\overline{p}_\ell - \overline{p}(y_\ell)\|_{0,\Omega}^2. \tag{5.9}$$

On the other hand, choosing $v = \overline{p} - \overline{p}(y_\ell)$ in (2.6a), (4.8a) and $v = y - y(u_\ell)$ in (2.8), (4.8b), for the second term on the right-hand side in (5.8) we get

$$\begin{aligned}(e_u, \overline{p}(y_\ell) - \overline{p})_{0,\Omega} &= -(y - y_\ell, y - y(u_\ell))_{0,\Omega} \\ &= -\|y - y(u_\ell)\|_{0,\Omega}^2 + (y_\ell - y(u_\ell), y - y(u_\ell))_{0,\Omega} \\ &\leq \|y - y(u_\ell)\|_{1,\Omega}\|y(u_\ell) - y_\ell\|_{1,\Omega} \\ &\leq \frac{\alpha}{4}\|e_u\|_{0,\Omega}^2 + \frac{M^2}{\alpha}\|y(u_\ell) - y_\ell\|_{1,\Omega}^2,\end{aligned} \tag{5.10}$$

where we have further made use of (5.4) and of Young's inequality in the last estimate. Using (5.9) and (5.10) in (5.8) results in

$$(e_u, \overline{p}_\ell - \overline{p})_{0,\Omega} \leq \frac{\alpha}{2}\|e_u\|_{0,\Omega}^2 + \frac{1}{\alpha}\big(\|\overline{p}_\ell - \overline{p}(y_\ell)\|_{1,\Omega}^2 + M^2\|y(u_\ell) - y_\ell\|_{1,\Omega}^2\big). \tag{5.11}$$

As far as the third term on the right-hand side in (5.7) is concerned, in view of (2.6a), (2.7), (3.2a), (3.4), (4.8a) and (4.9) we obtain

$$(e_u, \overline{\sigma}_\ell - \overline{\sigma})_{0,\Omega} = (\nabla(y_\ell(u) - y_\ell), \nabla\overline{\sigma}_\ell)_{0,\Omega} + (c(y_\ell(u) - y_\ell), \overline{\sigma}_\ell)_{0,\Omega}$$

$$
\begin{aligned}
&- (\nabla(y - y(u_\ell)), \nabla\overline{\sigma})_{0,\Omega} - (c(y - y(u_\ell)), \overline{\sigma})_{0,\Omega} \\
&= (\sigma_\ell, y_\ell(u) - y_\ell)_{0,\Omega} + (\sigma, y(u_\ell) - y)_{0,\Omega}.
\end{aligned} \tag{5.12}
$$

Combining (5.12) with the second term on the right-hand side in (5.7) and using the complementarity conditions (2.6d), (3.2d) and (3.7) as well as the definition (4.10) of the consistency error, we find

$$
\begin{aligned}
&\varepsilon(e_u, \sigma_\ell - \sigma)_{0,\Omega} + (e_u, \overline{\sigma}_\ell - \overline{\sigma})_{0,\Omega} \\
&\quad = (\sigma_\ell, \varepsilon u + y_\ell(u) - (\varepsilon u_\ell + y_\ell))_{0,\Omega} - \big(\sigma, \varepsilon u + y - (\varepsilon u_\ell + y(u_\ell))\big)_{0,\Omega} \\
&\quad = (\sigma_\ell, \varepsilon u + y_\ell(u) - \psi_\ell)_{0,\Omega} + (\sigma, \varepsilon u_\ell + y(u_\ell) - \psi)_{0,\Omega} \\
&\qquad + (\sigma_\ell, \psi_\ell - (\varepsilon u_\ell + y_\ell))_{0,\Omega} - (\sigma, \varepsilon u + y - \psi)_{0,\Omega} \le e_{c,\ell}.
\end{aligned} \tag{5.13}
$$

Finally, another application of Young's inequality gives us the following upper bound for the fourth term on the right-hand side in (5.7)

$$
\alpha(e_u, u^d - u_\ell^d)_{0,\Omega} \le \frac{\alpha}{4}\|e_u\|_{0,\Omega}^2 + \alpha\, osc_\ell^2(u^d). \tag{5.14}
$$

Taking advantage of the estimates (5.11), (5.13), (5.14) in (5.7) allows to conclude. □

Lemma 5.4 *Under the same assumptions as in Lemma 5.2 there holds*

$$
\|y(u_\ell) - y_\ell\|_{1,\Omega}^2 \lesssim (\eta_\ell(y))^2, \tag{5.15a}
$$

$$
\|\overline{p}(y_\ell) - \overline{p}_\ell\|_{1,\Omega}^2 \lesssim (\eta_\ell(\overline{p}))^2. \tag{5.15b}
$$

Proof Due to Galerkin orthogonality, the assertion follows by standard arguments from the a posteriori error analysis of adaptive finite element methods (see, e.g., [31]). □

Proof of Theorem 5.1 Combining the estimates from Lemmas 5.2, 5.3, and 5.4 results in (5.1). □

Remark 5.5 It follows from the proofs of the previous results that the constants involved in the reliability estimates do not depend on the regularization parameter ε.

We conclude this section by deriving a computable upper bound for the consistency error $e_{c,\ell}$. Taking into account that $\sigma|_{\mathcal{I}} = 0, \sigma_\ell|_{\mathcal{I}_\ell} = 0$ and $(\varepsilon u + y)|_{\mathcal{A}} = \psi|_{\mathcal{A}}, (\varepsilon u_\ell + y_\ell)|_{\mathcal{A}_\ell} = \psi_\ell|_{\mathcal{A}_\ell}$, we find

$$
e_{c,\ell}|_{\mathcal{I}\cap\mathcal{I}_\ell} = 0, \tag{5.16a}
$$

$$
\begin{aligned}
e_{c,\ell}|_{\mathcal{A}\cap\mathcal{A}_\ell} &\le (\sigma, y(u_\ell) - y_\ell)_{0,\mathcal{A}\cap\mathcal{A}_\ell} + (\sigma_\ell, y_\ell(u) - y)_{0,\mathcal{A}\cap\mathcal{A}_\ell} \\
&\quad + (\sigma - \sigma_\ell, \psi_\ell - \psi)_{0,\mathcal{A}\cap\mathcal{A}_\ell},
\end{aligned} \tag{5.16b}
$$

$$
e_{c,\ell}|_{\mathcal{I}\cap\mathcal{A}_\ell} \le (\sigma_\ell, y_\ell(u) - y)_{0,\mathcal{I}\cap\mathcal{A}_\ell} + (\sigma_\ell, \psi - \psi_\ell)_{0,\mathcal{I}\cap\mathcal{A}_\ell}, \tag{5.16c}
$$

$$
e_{c,\ell}|_{\mathcal{A}\cap\mathcal{I}_\ell} \le (\sigma, y(u_\ell) - y_\ell)_{0,\mathcal{A}\cap\mathcal{I}_\ell} + (\sigma, \psi_\ell - \psi)_{0,\mathcal{A}\cap\mathcal{I}_\ell}. \tag{5.16d}
$$

This leads to the estimates

$$e_{c,\ell}|_{\mathcal{A}\cap\mathcal{A}_\ell} \le \|\sigma\|_{0,\mathcal{A}\cap\mathcal{A}_\ell}\|y(u_\ell)-y_\ell\|_{0,\mathcal{A}\cap\mathcal{A}_\ell} + \|\sigma_\ell\|_{0,\mathcal{A}\cap\mathcal{A}_\ell}\|y_\ell(u)-y\|_{0,\mathcal{A}\cap\mathcal{A}_\ell} + (\|\sigma\|_{0,\mathcal{A}\cap\mathcal{A}_\ell}+\|\sigma_\ell\|_{0,\mathcal{A}\cap\mathcal{A}_\ell})\|\psi-\psi_\ell\|_{0,\mathcal{A}\cap\mathcal{A}_\ell}, \tag{5.17a}$$

$$e_{c,\ell}|_{\mathcal{I}\cap\mathcal{A}_\ell} \le \|\sigma_\ell\|_{0,\mathcal{I}\cap\mathcal{A}_\ell}\|y_\ell(u)-y\|_{0,\mathcal{I}\cap\mathcal{A}_\ell} + \|\sigma_\ell\|_{0,\mathcal{I}\cap\mathcal{A}_\ell}\|\psi-\psi_\ell\|_{0,\mathcal{I}\cap\mathcal{A}_\ell}, \tag{5.17b}$$

$$e_{c,\ell}|_{\mathcal{A}\cap\mathcal{I}_\ell} \le \|\sigma\|_{0,\mathcal{A}\cap\mathcal{I}_\ell}\|y(u_\ell)-y_\ell\|_{0,\mathcal{A}\cap\mathcal{I}_\ell} + \|\sigma\|_{0,\mathcal{A}\cap\mathcal{I}_\ell}\|\psi-\psi_\ell\|_{0,\mathcal{A}\cap\mathcal{I}_\ell}. \tag{5.17c}$$

In view of (5.15a) in Lemma 5.4 we have $\|y(u_\ell)-y_\ell\|_{0,\Omega} \lesssim \eta_\ell(y)$. By the same arguments $\|y_\ell(u)-y\|_{0,\Omega} \lesssim \bar{\eta}_\ell(y)$, where $\bar{\eta}_\ell(y) := (\sum_{T\in\mathcal{T}_\ell(\Omega)} \bar{\eta}_T^2(y) + \sum_{E\in\mathcal{E}_\ell(\Omega)} \bar{\eta}_E^2(y)$ with $\bar{\eta}_T(y)$ and $\bar{\eta}_E(y)$ given as in (4.3a) and (4.4a) with y_ℓ replaced by $y_\ell(u)$. Of course, $y_\ell(u)$ is not available, but we may expect $\bar{\eta}_T(y) \approx \eta_T(y)$, $\bar{\eta}_E(y) \approx \eta_E(y)$ and hence, replace $\bar{\eta}_\ell(y)$ by $\eta_\ell(y)$. Moreover, due to the results in [15] for control constrained elliptic optimal control problems the sequence $(\|\sigma_\ell\|_{0,\Omega})_{\ell\in\mathbb{N}}$ can be expected to be uniformly bounded. Then, (5.17a)–(5.17c) already show that the consistency error $e_{c,\ell}$ can be controlled by the residual error estimator $\eta_\ell(y)$ and the data oscillation $osc_\ell(\psi)$. The upper bounds become computable as soon as we construct suitable approximations of $\|\sigma\|_{0,S}$, $S \in \{\mathcal{A}\cap\mathcal{A}_\ell, \mathcal{A}\cap\mathcal{I}_\ell\}$, in (5.17a), (5.17c) as well as estimates for the sets $\mathcal{A}\cap\mathcal{A}_\ell, \mathcal{I}\cap\mathcal{A}_\ell$ and $\mathcal{A}\cap\mathcal{I}_\ell$. To this end, we provide a modification of the approximation of the continuous coincidence set $\mathcal{A}$ from [19] (cf. also [14]) by using locally quadratic interpolations of the computed state, adjoint state and control in the spirit of the goal-oriented dual weighted approach (cf., e.g., [4]). In particular, referring to $P_2(T)$, $T \in \mathcal{T}_{\ell-1}(\Omega)$, as the set of quadratic polynomials on T, we define

$$i_{\ell-1}^y y_\ell, i_{\ell-1}^p p_\ell, i_{\ell-1}^u u_\ell \in \prod_{T\in\mathcal{T}_{\ell-1}(\Omega)} P_2(T)$$

such that $i_{\ell-1}^y y_\ell|_T, i_{\ell-1}^p p_\ell|_T, i_{\ell-1}^u u_\ell|_T$, $T \in \mathcal{T}_{\ell-1}(\Omega)$, are the quadratic interpolations of y_ℓ, p_ℓ and u_ℓ on T with respect to the values in the vertices and the midpoints of the edges of T. We approximate the characteristic function $\chi(\mathcal{A})$ of the continuous coincidence set $\mathcal{A}$ by

$$\chi_\ell^{\mathcal{A}} := 1 - \frac{\psi - (\varepsilon i_{\ell-1}^u u_\ell + i_{\ell-1}^y y_\ell)}{\gamma h_\ell^r + \psi - (\varepsilon i_{\ell-1}^u u_\ell + i_{\ell-1}^y y_\ell)}, \tag{5.18}$$

where $0 < \gamma \le 1$ and $r > 0$ are fixed. Indeed, for $T \subset \mathcal{A}$ we find

$$\|\chi(\mathcal{A}) - \chi_\ell^{\mathcal{A}}\|_{0,T} = \left\| \frac{\gamma^{-1}h_\ell^{-r}(\varepsilon u + y - (\varepsilon i_{\ell-1}^u u_\ell + i_{\ell-1}^y y_\ell))}{1+\gamma^{-1}h_\ell^{-r}(\varepsilon u + y - (\varepsilon i_{\ell-1}^u u_\ell + i_{\ell-1}^y y_\ell))} \right\|_{0,T},$$

which converges to zero whenever $\|\varepsilon u + y - (\varepsilon i_{\ell-1}^u u_\ell + i_{\ell-1}^y y_\ell)\|_{0,T} = O(h_\ell^q)$, $q > r$. By the same arguments, for $T \subset \mathcal{I}$ one can show as well that $\|\chi(\mathcal{A}) - \chi_\ell^{\mathcal{A}}\|_{0,T} \to 0$ as $h_\ell \to 0$. Now, for fixed $0 < \kappa \le 1$ and $0 < s \le r$ we provide approximations $\hat{\mathcal{A}}_\ell$ of $\mathcal{A}$ and $\hat{\mathcal{I}}_\ell$ of $\mathcal{I}$ according to

$$\hat{\mathcal{A}}_\ell := \bigcup \{T \in \mathcal{T}_\ell(\Omega) \mid \chi_\ell^{\mathcal{A}}(x) \ge 1 - \kappa h_\ell^s \text{ for all } x \in T\}, \tag{5.19a}$$

$$\hat{\mathcal{I}}_\ell := \bigcup \{T \in \mathcal{T}_\ell(\Omega) \mid \chi_\ell^{\mathcal{A}}(x) < 1 - \kappa h_\ell^s \text{ for some } x \in T\}. \tag{5.19b}$$

We define approximations $\mathcal{T}_{\mathcal{A}\cap\mathcal{A}_\ell}$, $\mathcal{T}_{\mathcal{I}\cap\mathcal{A}_\ell}$ and $\mathcal{T}_{\mathcal{A}\cap\mathcal{I}_\ell}$ of $\mathcal{A}\cap\mathcal{A}_\ell, \mathcal{I}\cap\mathcal{A}_\ell$ and $\mathcal{A}\cap\mathcal{I}_\ell$ by means of

$$\mathcal{T}_{\mathcal{A}\cap\mathcal{A}_\ell} := \hat{\mathcal{A}}_\ell \cap \mathcal{A}_\ell, \qquad \mathcal{T}_{\mathcal{I}\cap\mathcal{A}_\ell} := \hat{\mathcal{I}}_\ell \cap \mathcal{A}_\ell, \qquad \mathcal{T}_{\mathcal{A}\cap\mathcal{I}_\ell} := \hat{\mathcal{A}}_\ell \cap \mathcal{I}_\ell.$$

For the approximation of σ on $\mathcal{A}\cap\mathcal{A}_\ell$ we may assume $\|\sigma\|_{0,\mathcal{A}\cap\mathcal{A}_\ell} \approx \|\sigma_\ell\|_{0,\mathcal{A}\cap\mathcal{A}_\ell}$. On $\mathcal{A}\cap\mathcal{I}_\ell$ we take advantage of (2.6c), i.e., $\sigma = \varepsilon^{-1}(\alpha(u^d - u) - p)$, and approximate p and u by $i_{\ell-1}^p p_\ell$ and $i_{\ell-1}^u u_\ell$, respectively. This leads to $\|\sigma\|_{0,\mathcal{A}\cap\mathcal{I}_\ell} \approx \|z_\ell\|_{0,\mathcal{A}\cap\mathcal{I}_\ell}$, where $z_\ell := \varepsilon^{-1}(\alpha(u^d - i_{\ell-1}^u u_\ell) - i_{\ell-1}^p p_\ell)$. Altogether, we thus arrive at the estimates

$$e_{c,\ell}|_{\mathcal{A}\cap\mathcal{A}_\ell} \lesssim 2\|\sigma_\ell\|_{0,\mathcal{T}_{\mathcal{A}\cap\mathcal{A}_\ell}}\big(\eta_\ell(y) + \|\psi - \psi_\ell\|_{0,\mathcal{T}_{\mathcal{A}\cap\mathcal{A}_\ell}}\big), \tag{5.20a}$$

$$e_{c,\ell}|_{\mathcal{I}\cap\mathcal{A}_\ell} \lesssim \|\sigma_\ell\|_{0,\mathcal{T}_{\mathcal{I}\cap\mathcal{A}_\ell}}\big(\eta_\ell(y) + \|\psi - \psi_\ell\|_{0,\mathcal{T}_{\mathcal{I}\cap\mathcal{A}_\ell}}\big), \tag{5.20b}$$

$$e_{c,\ell}|_{\mathcal{A}\cap\mathcal{I}_\ell} \lesssim \|z_\ell\|_{0,\mathcal{T}_{\mathcal{A}\cap\mathcal{I}_\ell}}\big(\eta_\ell(y) + \|\psi - \psi_\ell\|_{0,\mathcal{T}_{\mathcal{A}\cap\mathcal{I}_\ell}}\big). \tag{5.20c}$$

The upper bounds in (5.20a)–(5.20c) are computable, but appear too pessimistic. Actually, in (4.10) we may expect $(\varepsilon u + y_\ell(u))|_{\mathcal{A}} \approx \psi$ and $(\varepsilon u_\ell + y(u_\ell)|_{\mathcal{A}_\ell} \approx \psi_\ell$ such that the consistency error will be essentially controlled by the data oscillations in ψ. This is confirmed by the numerical examples in Sect. 7.

Remark 5.6 The definition of $\chi_\ell^{\mathcal{A}}$ as in [19], where in (5.18) $i_{\ell-1}^y y_\ell$ and $i_{\ell-1}^u u_\ell$ are replaced by y_ℓ and u_ℓ, is not sufficient. For instance, in case $\psi_\ell = \psi$ we would have $\chi_\ell^{\mathcal{A}}(x) = 1, x \in \mathcal{A}_\ell$, and hence $\hat{\mathcal{I}}_\ell \cap \mathcal{A}_\ell = \emptyset$ due to $\mathcal{A}_\ell \subset \hat{\mathcal{A}}_\ell$, i.e., we would not get a reasonable approximation of $\mathcal{I}\cap\mathcal{A}_\ell$.

Remark 5.7 The limiting case $\varepsilon \to 0$ corresponds to the case of pure state constraints where the multipliers σ and σ_ℓ are Radon measures. Denoting for clarity of notation the states, the controls, and the multipliers in the mixed control-state constrained case by y^ε, y_ℓ^ε etc. and by $\langle\cdot,\cdot\rangle$ the dual pairing between the measure space $\mathcal{M}(\Omega)$ and $C(\overline{\Omega})$, we can only expect $\langle\sigma^\varepsilon, z\rangle \to \langle\sigma, z\rangle$ and $\langle\sigma_\ell^\varepsilon, z\rangle \to \langle\sigma_\ell, z\rangle, z \in C(\overline{\Omega})$, as $\varepsilon \to 0$. Consequently, for the consistency error $e_{c,\ell}^\varepsilon$ it follows that $e_{c,\ell}^\varepsilon \to e_{c,\ell} = \max(\langle\sigma_\ell, y_\ell(u) - \psi_\ell\rangle + \langle\sigma, y(u_\ell) - \psi\rangle, 0)$, if $y_\ell^\varepsilon(u^\varepsilon) \to y_\ell(u)$ and $y^\varepsilon(u_\ell^\varepsilon) \to y(u_\ell)$ in $C(\overline{\Omega})$ as $\varepsilon \to 0$. Indeed, in this case the limit $e_{c,\ell}$ corresponds to the consistency error in the purely state constrained case (cf. [18]).

6 Local efficiency of the error estimator

Efficiency of the estimator means that up to data oscillations it also provides a lower bound for the discretization errors in the state, the regularized adjoint state, and the control.

Theorem 6.1 *Let* (y, u, p, σ) *and* $(y_\ell, u_\ell, p_\ell, \sigma_\ell)$ *be the solutions of* (2.6a)–(2.6d) *and* (3.2a)–(3.2d) *and let* η_ℓ *and* $osc_\ell(y^d)$ *be the error estimator and the data oscillation as given by* (4.1) *and* (4.6b), *respectively. Further, let* $\overline{p}$ *and* $\overline{p}_\ell$ *be the modified adjoint states as given by* (2.8), (3.5). *Then, there holds*

$$\eta_\ell^2 - osc_\ell^2(y^d) \lesssim \|e_y\|_{1,\Omega}^2 + \|e_u\|_{0,\Omega}^2 + \|e_{\overline{p}}\|_{1,\Omega}^2. \tag{6.1}$$

The proof of the efficiency of the estimator is usually done by establishing local efficiency in the sense that the element and edge residuals can be bounded from above by norms of the discretization errors on the elements and associated patches, respectively. Local efficiency will be provided by the subsequent two lemmas.

Lemma 6.2 *Let* $\eta_T(y)$ *and* $\eta_T(\overline{p}), T \in \mathcal{T}_\ell(\Omega)$, *be the element residuals as given by* (4.3). *Then, there holds*

$$\eta_T^2(y) \lesssim \|e_y\|_{1,T}^2 + h_T^2 \|e_u\|_{0,T}^2, \tag{6.2}$$

$$\eta_T^2(\overline{p}) \lesssim \|e_{\overline{p}}\|_{1,T}^2 + h_T^2 \|e_y\|_{0,T}^2 + osc_T^2(y^d). \tag{6.3}$$

Proof We denote by $\varphi_T = \prod_{i=1}^3 \lambda_i(T)$ the element bubble function associated with $T \in \mathcal{T}_\ell(\Omega)$ (cf., e.g., [31]). Then, $\xi_\ell := (u_\ell - cy_\ell)\varphi_T$ is an admissible test function in (2.6a). Observing $\Delta y_\ell|_T = 0$, we obtain

$$\begin{aligned}\eta_T^2(y) &\lesssim h_T^2 (u_\ell - cy_\ell, \xi_\ell)_{0,T} \\ &= h_T^2 \big((u, \xi_\ell)_{0,T} + (\Delta y_\ell - cy_\ell, \xi_\ell)_{0,T} + (u_\ell - u, \xi_\ell)_{0,T}\big) \\ &= h_T^2 \big((\nabla(y - y_\ell), \nabla \xi_\ell)_{0,T} + (c(y - y_\ell), \xi_\ell)_{0,T} + (u_\ell - u, \xi_\ell)_{0,T}\big).\end{aligned}$$

Using standard estimates for $\|\nabla \xi_\ell\|_{0,T}$ and $\|\xi_\ell\|_{0,T}$ (cf., e.g., [31]) readily gives (6.2). The estimate (6.3) can be verified in the same way. □

Lemma 6.3 *Let* $\eta_T(y), \eta_T(\overline{p}), T \in \mathcal{T}_\ell(\Omega)$, *and* $\eta_E(y), \eta_E(\overline{p}), E \in \mathcal{E}_\ell(\Omega)$, *be the element and edge residuals as given by* (4.3), (4.4). *Further, let* $osc_T(y^d), T \in \mathcal{T}_\ell(\Omega)$, *be the element contribution to the data oscillation in* y^d *according to* (4.6b). *Then, there holds*

$$\eta_E^2(y) \lesssim \|e_y\|_{1,\omega_E}^2 + h_E^2 \|e_u\|_{0,\omega_E}^2 + \eta_{\omega_E}^2(y), \tag{6.4}$$

$$\eta_E^2(\overline{p}) \lesssim \|e_{\overline{p}}\|_{1,\omega_E}^2 + h_E^2 \|e_y\|_{0,\omega_E}^2 + \eta_{\omega_E}^2(\overline{p}) + osc_{\omega_E}^2(y^d), \tag{6.5}$$

where $\eta_{\omega_E}(y) := (\eta_{T_+}^2(y) + \eta_{T_-}^2(y))^{1/2}$ *and* $\eta_{\omega_E}(\overline{p}), osc_{\omega_E}(y^d)$ *are defined analogously.*

Proof We denote by $\varphi_E = \prod_{i=1}^2 \lambda_i(E)$, the edge bubble function associated with $E \in \mathcal{E}_\ell(\Omega)$ (cf., e.g., [31]). We set $\zeta_E := (\nu_E \cdot [\nabla y_\ell])|_E$ and $\xi_\ell := \tilde{\zeta}_E \varphi_E$, where $\tilde{\zeta}_E$ is the extension of ζ_E to ω_E as in [31]. Taking advantage of the fact that ξ_ℓ is an admissible test function in (2.6a) and $\Delta y_\ell|_T = 0$, it follows that

$$\eta_E^2(y) \lesssim h_E (\nu_E \cdot [\nabla y_\ell], \zeta_E \varphi_E)_{0,E}$$

$$
\begin{aligned}
&= h_E \sum_{T \subset \omega_E} \big((\nu_{\partial T} \cdot \nabla y_\ell, \xi_\ell)_{0,\partial T} - (\Delta y_\ell, \xi_\ell)_{0,T} \big) \\
&= h_E \big((\nabla(y_\ell - y), \nabla \xi_\ell)_{0,\omega_E} + (c(y_\ell - y), \xi_\ell)_{0,\omega_E} \\
&\quad + (u - u_\ell, \xi_\ell)_{0,\omega_E} + (u_\ell - c y_\ell, \xi_\ell)_{0,\omega_E} \big).
\end{aligned}
$$

Standard estimates for ξ_ℓ (cf., e.g., [31]) readily give (6.4). The estimate (6.5) can be proved along the same lines. □

Proof of Theorem 6.1 The efficiency estimate (6.1) follows by summing up the estimates (6.3)–(6.5) over all $T \in \mathcal{T}_\ell(\Omega)$ and $E \in \mathcal{E}_\ell(\Omega)$. Using the fact that the union of the patches ω_E has a finite overlap allows to conclude. □

7 Numerical results

In this section, we illustrate the approximation of state constrained optimal control problems by its Lavrentiev type regularizations using two numerical examples. The first example features a solution y that strongly oscillates around the origin where the coincidence set is a connected subdomain with smooth boundary. In contrast to that, the second example, which is taken from [22], features a multiplier in $\mathcal{M}_+(\Omega)$ where the coincidence set degenerates to a single point. In both cases the adaptive process generates finite element meshes that are close to those created when one uses the adaptive strategy for state constrained problems as suggested in [18].

Example 1 (Simply connected coincidence set with smooth boundary) The data of the problem are as follows:

$$
\begin{aligned}
&\Omega := (-2, 2)^2, \qquad \psi := 0, \qquad \alpha := 0.1, \qquad c = 0, \qquad \Gamma_D := \partial\Omega, \\
&y^d := y(r) + \Delta p(r) + \sigma(r), \qquad u^d := u(r) + \alpha^{-1}\, p(r).
\end{aligned}
$$

Here, $y = y(r), u = u(r), p = p(r)$ and $\sigma = \sigma(r), r := (x_1^2 + x_2^2)^{1/2}, (x_1, x_2)^T \in \Omega$, represent the exact optimal solution of the pure state constrained problem ($\varepsilon = 0$) according to

$$
\begin{aligned}
&y(r) := -r^{\frac{4}{3}} \gamma_1(r), \qquad u(r) := -\Delta y(r), \\
&p(r) := \gamma_2(r) \left(r^4 - \frac{3}{2} r^3 + \frac{9}{16} r^2 \right), \qquad \sigma(r) := \begin{cases} 0, & r < 0.75, \\ 0.1, & \text{otherwise}, \end{cases}
\end{aligned}
$$

where

$$
\gamma_1 := \begin{cases} 1, & r < 0.25, \\ -192(r - 0.25)^5 + 240(r - 0.25)^4 - 80(r - 0.25)^3 + 1, & 0.25 < r < 0.75, \\ 0, & \text{otherwise}, \end{cases}
$$

$$\gamma_2 := \begin{cases} 1, & r < 0.75, \\ 0, & \text{otherwise.} \end{cases}$$

Figure 1 shows the computed optimal state y_ℓ and optimal control u_ℓ in case of an adaptively generated simplicial triangulation with 9194 degrees of freedom and $\epsilon = 10^{-6}$, whereas Fig. 2 displays the adaptively generated meshes after 12 (left) and 14 (right) refinement steps for $\epsilon = 10^{-6}$.

Table 1 documents the convergence history of the adaptive refinement process with respect to the convergence of the solutions of the discrete mixed control-state problems ($\varepsilon = 10^{-6}$) to the exact solution of the pure state constrained problem. We remark that the impact of the regularization parameter ε has to be observed in the error estimates. In particular, Table 1 contains the H^1-error in the state, the L^2-errors in the control and in the adjoint state as well as the H^1-error in the regularized adjoint state. The adaptive refinement process has been terminated when the size of the regularization parameter $\varepsilon = 10^{-6}$ started to blur the results. The same effect could be observed for larger values of ε (see also Fig. 3 (right)). For a comparison, Table 2 shows the decrease of the discretization errors in the limiting case $\varepsilon = 0$ (pure state

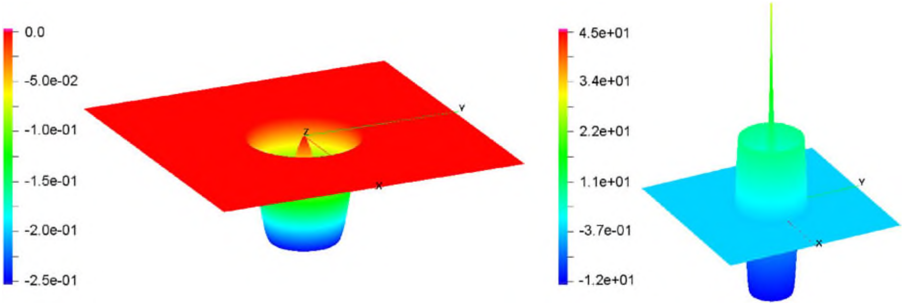

Fig. 1 Example 1. Visualization of the discrete optimal state y_l (*left*) and the discrete optimal control u_l (*right*) on a triangulation with 9194 nodes and with regularization parameter $\epsilon = 10^{-6}$

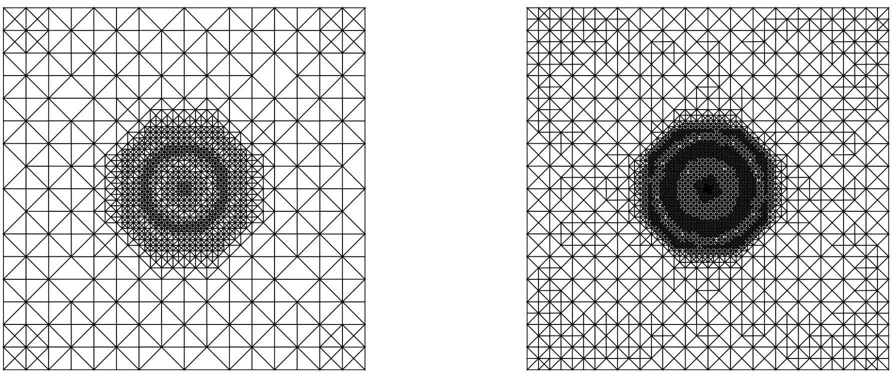

Fig. 2 Example 1. Adaptively generated grid after 12 (*left*) and 14 (*right*) refinement steps, $\Theta_i = 0.7, \epsilon = 10^{-6}$

Table 1 Example 1. Convergence history of the adaptive FEM. Errors in the state, the control, the adjoint state, and the regularized adjoint state ($\varepsilon = 10^{-6}$)

ℓ	N_{dof}	$\|u - u_\ell\|_0$	$\|y - y_\ell\|_1$	$\|\bar{p} - \bar{p}_\ell\|_1$	$\| p - p_\ell\|_0$
1	13	1.37e+01	1.03e+00	1.91e-01	9.67e-01
2	41	1.01e+01	1.58e+00	1.51e-01	8.61e-01
4	74	9.53e+00	1.31e+00	1.01e-01	9.48e-02
6	142	6.01e+00	6.56e-01	1.12e-01	3.06e-02
8	290	3.36e+00	3.72e-01	1.27e-01	1.94e-02
10	623	2.19e+00	2.34e-01	1.29e-01	8.47e-03
12	1412	1.47e+00	1.32e-01	1.31e-01	1.08e-02
14	3498	1.01e+00	7.92e-02	1.30e-01	1.29e-02

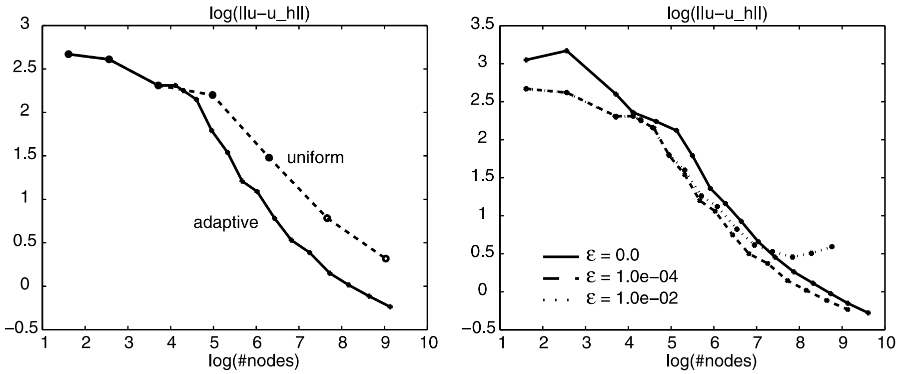

Fig. 3 Example 1. Adaptive refinement [*straight line*] versus uniform refinement [*dotted line*] (*left*) and discretization error in the control for different regularization parameters (*right*)

Table 2 Example 1. Convergence history of the adaptive FEM. Discretization errors in the pure state constrained case ($\varepsilon = 0$); from [18]

ℓ	N_{dof}	$\|u - u_\ell\|_0$	$\|y - y_\ell\|_1$	$\|\bar{p} - \bar{p}_\ell\|_1$	$\| p - p_\ell\|_0$
1	13	2.37e+01	1.51e+00	6.74e-01	2.06e+00
2	41	1.35e+01	1.02e+00	1.06e-01	1.28e-01
4	105	9.41e+00	7.34e-01	7.88e-02	9.54e-02
6	244	6.01e+00	5.41e-01	6.02e-02	4.78e-02
8	532	3.18e+00	2.80e-01	4.53e-02	3.92e-02
10	1147	1.91e+00	1.74e-01	3.44e-02	2.36e-02
12	2651	1.29e+00	1.03e-01	2.02e-02	1.81e-02
14	6340	9.74e-01	6.32e-02	1.17e-02	1.22e-02

constraints). We observe that in this example the most significant impact of the regularization parameter is on the errors in the adjoint state and the regularized adjoint state.

Table 3 Example 1. Convergence history of the adaptive FEM. Estimator in the state and regularized adjoint state, data oscillations, and consistency error ($\varepsilon = 10^{-6}$)

ℓ	N_{dof}	$\eta_\ell(y)$	$\eta_\ell(\bar{p})$	$osc_\ell(u^d)$	$osc_\ell(y^d)$	$e_{c,\ell}$
1	13	4.20e+00	1.04e+00	1.37e+01	5.42e-01	1.10e+00
2	41	4.25e+00	1.04e+00	1.36e+01	6.22e-01	0.87e+00
4	74	4.01e+00	4.71e-01	9.67e+00	3.32e-01	0.00e+00
6	142	1.77e+00	3.14e-01	6.03e+00	1.11e-01	0.00e+00
8	290	1.27e+00	2.49e-01	3.38e+00	5.36e-02	0.00e+00
10	623	8.70e-01	1.80e-01	2.19e+00	2.78e-02	0.00e+00
12	1412	5.50e-01	1.12e-01	1.47e+00	1.50e-02	0.00e+00
14	3498	3.42e-01	6.90e-02	1.01e+00	7.08e-03	0.00e+00

Table 4 Example 1. Convergence history of the adaptive FEM. Estimators and data oscillations in the pure state constrained case ($\varepsilon = 0$); from [18]

ℓ	N_{dof}	$\eta_\ell(y)$	$\eta_\ell(\bar{p})$	$osc_\ell(u^d)$	$osc_\ell(y^d)$
1	13	2.19e+01	2.04e+00	1.37e+01	5.42e-01
2	41	9.83e+00	8.10e-01	1.36e+01	6.22e-01
4	105	3.67e+00	4.35e-01	9.42e+00	3.32e-01
6	244	1.63e+00	2.60e-01	5.99e+00	1.11e-01
8	532	1.17e+00	1.69e-01	3.17e+00	4.47e-02
10	1147	7.72e-01	1.22e-01	1.90e+00	2.17e-02
12	2651	4.71e-01	7.37e-02	1.29e+00	9.27e-03
14	6340	2.93e-01	4.55e-02	9.74e-01	4.62e-03

Figure 3 (left) illustrates the benefit of adaptive versus uniform refinement by showing on a logarithmic scale the error in the control as a function of the degrees of freedom ($\epsilon = 10^{-6}$). Figure 3 (right) contains a comparison of the error in the control for the pure state constrained problem ($\epsilon = 0$) and its Lavrentiev regularizations ($\varepsilon = 10^{-2}, 10^{-4}, 10^{-6}$).

The residual-type estimators $\eta_\ell(y)$ in the state and $\eta_\ell(\bar{p})$ in the regularized adjoint state, the data oscillations $osc_\ell(u^d)$ in the shift control and $osc_\ell(y^d)$ in the desired state as well as the consistency error are given in Table 3. As expected, the consistency error vanishes, once the continuous free boundary has been properly resolved by its discrete counterpart. For a comparison with the pure state constrained case, Table 4 contains corresponding values in case $\varepsilon = 0$.

Example 2 (Degenerated coincidence set [22]) The data of the problem are as follows:

$$\Omega := B(0,1), \qquad \Gamma_D = \emptyset, \qquad \alpha := 1.0, \qquad c = 1.0,$$

$$y^d(r) := 4 + \frac{1}{\pi} - \frac{1}{4\pi} r^2 + \frac{1}{2\pi} \ln(r),$$

$$u^d(r) := 4 + \frac{1}{4\pi} r^2 - \frac{1}{2\pi} \ln(r), \qquad \psi(r) := r + 4.$$

The optimal solution in the pure state constrained case is given by

$$y(r) \equiv 4, \qquad p(r) = \frac{1}{4\pi} r^2 - \frac{1}{2\pi} \ln(r), \qquad u(r) \equiv 4, \qquad \sigma = \delta_0.$$

Figure 4 displays the computed optimal state y_ℓ and optimal control u_ℓ for a simplicial triangulation with 964 degrees of freedom. For the regularization parameter $\epsilon = 10^{-6}$, the adaptively generated grids after 12 and 14 refinement steps are shown in Fig. 5.

In case $\varepsilon = 10^{-6}$, Tables 5 and 7 reflect the convergence history of the adaptive refinement process with data analogous to those in Example 1. As before, in order to compare with the pure state constrained case ($\varepsilon = 0$), the associated data are given in Tables 6 and 8. We remark that there are no data oscillations in ψ in the pure state constrained case $\varepsilon = 0$, since $\langle \sigma - \sigma_\ell, \psi - \psi_\ell \rangle = 0$ where $\langle \cdot, \cdot \rangle$ stands for the dual pairing between the space of Radon measures and the space of continuous functions (cf. [18]). There is no consistency error, since coincidence with the obstacle only

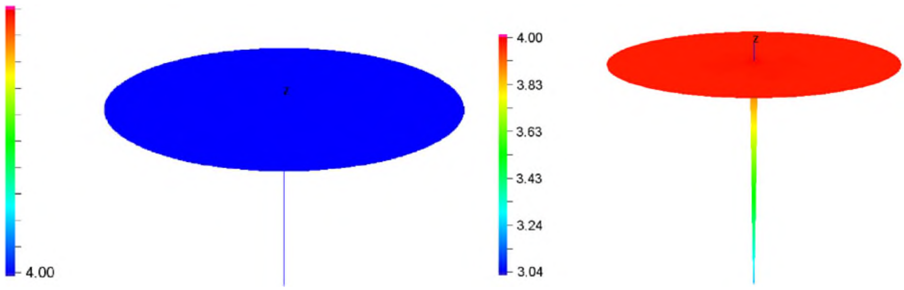

Fig. 4 Example 2. Visualization of the discrete state y_l (*left*) and the discrete control u_l (*right*) on an adaptive generated mesh with 964 nodes and with regularization parameter $\epsilon = 10^{-6}$

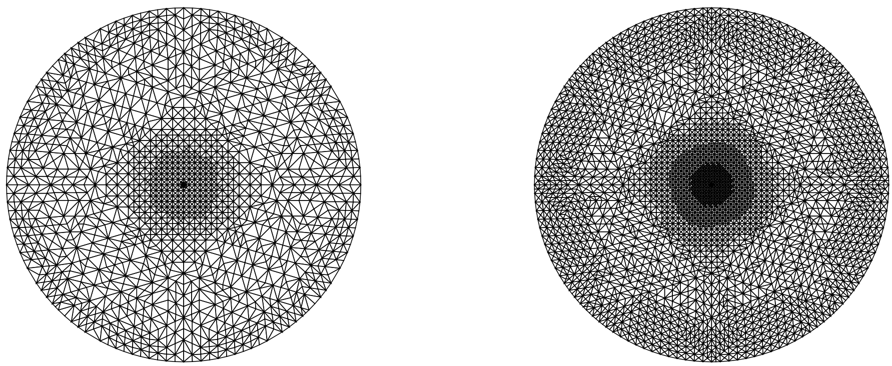

Fig. 5 Example 2. Adaptively generated grid after 12 (*left*) and 14 (*right*) refinement steps, $\Theta_i = 0.7$, $\epsilon = 10^{-6}$

Table 5 Example 2. Convergence history of the adaptive FEM. Errors in the state, the control, the adjoint state, and the regularized adjoint state ($\varepsilon = 10^{-6}$)

ℓ	N_{dof}	$\|u - u_\ell\|_0$	$\|y - y_\ell\|_1$	$\|\bar{p} - \bar{p}_\ell\|_1$	$\|p - p_\ell\|_0$
1	13	8.62e-02	1.75e-02	2.24e-02	6.73e-02
2	41	6.50e-02	1.01e-02	1.21e-02	2.93e-02
4	73	5.56e-02	6.89e-03	9.36e-03	1.54e-02
6	121	3.39e-02	2.34e-03	6.02e-03	8.30e-03
8	243	1.98e-02	6.96e-04	3.91e-03	4.35e-03
10	603	1.14e-02	2.25e-04	2.32e-03	2.00e-03
12	1618	6.39e-03	6.98e-05	1.46e-03	9.37e-04
14	3989	3.55e-03	2.58e-05	8.54e-04	4.57e-04
16	10656	1.95e-03	1.10e-05	4.76e-04	2.21e-04

Table 6 Example 2. Convergence history of the adaptive FEM. Discretization errors in the pure state constrained case ($\varepsilon = 0$); from [18]

ℓ	N_{dof}	$\|u - u_\ell\|_0$	$\|y - y_\ell\|_1$	$\|\bar{p} - \bar{p}_\ell\|_1$	$\|p - p_\ell\|_0$
1	13	1.04e-01	8.51e-03	1.74e-02	3.73e-02
2	41	6.95e-02	4.43e-03	9.01e-03	1.86e-02
4	73	5.73e-02	2.30e-03	7.36e-03	1.00e-02
6	121	3.42e-02	1.79e-03	6.11e-03	7.41e-03
8	243	1.99e-02	1.07e-03	4.02e-03	4.13e-03
10	604	1.14e-02	4.02e-04	2.43e-03	1.95e-03
12	1621	6.39e-03	1.60e-04	1.52e-03	9.26e-04
14	3991	3.55e-03	6.81e-05	8.79e-04	4.55e-04

Table 7 Example 2. Convergence history of the adaptive FEM. Estimators, data oscillations and consistency error ($\varepsilon = 10^{-6}$)

ℓ	N_{dof}	$\eta_\ell(y)$	$\eta_\ell(\bar{p})$	$osc_\ell(u^d)$	$osc_\ell(\psi)$	$osc_\ell(y^d)$	$e_{c,\ell}$
1	13	6.15e-02	7.38e-02	1.29e-01	1.11e-01	4.36e-02	0.00e+00
2	41	2.29e-02	3.76e-02	8.14e-02	3.25e-02	1.26e-02	0.00e+00
4	73	1.00e-02	2.52e-02	5.95e-02	2.13e-02	7.78e-03	0.00e+00
6	121	3.11e-03	2.01e-02	3.56e-02	1.23e-02	4.96e-03	0.00e+00
8	243	9.15e-04	1.32e-02	2.06e-02	5.27e-03	1.87e-03	0.00e+00
10	603	2.59e-04	8.12e-03	1.17e-02	2.25e-03	8.28e-04	0.00e+00
12	1618	7.23e-05	4.76e-03	6.54e-03	9.86e-04	3.17e-04	0.00e+00
14	3989	2.01e-05	2.89e-03	3.62e-03	3.95e-04	1.42e-04	0.00e+00
16	10656	5.53e-06	1.78e-03	1.98e-03	1.86e-04	5.89e-05	0.00e+00

Table 8 Example 2. Convergence history of the adaptive FEM. Estimators and data oscillations in the pure state constrained case ($\varepsilon = 0$); from [18]

ℓ	N_{dof}	$\eta_\ell(y)$	$\eta_\ell(\bar{p})$	$osc_\ell(u^d)$	$osc_\ell(y^d)$
1	13	7.32e-02	7.62e-02	1.29e-01	4.36e-02
2	41	2.45e-02	3.83e-02	8.14e-02	1.26e-02
4	73	1.02e-02	2.54e-02	5.95e-02	7.78e-03
6	121	3.11e-03	1.97e-02	3.56e-02	4.96e-03
8	243	9.10e-04	1.32e-02	2.06e-02	1.87e-03
10	604	2.59e-04	8.07e-03	1.17e-02	8.27e-04
12	1621	7.22e-05	4.75e-03	6.54e-03	3.16e-04
14	3991	2.01e-05	2.89e-03	3.62e-03	1.41e-04

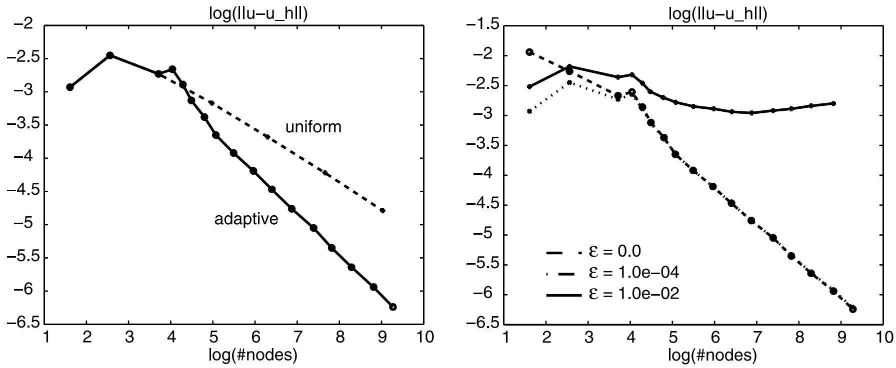

Fig. 6 Example 2. Adaptive versus uniform refinement (*left*) and discretization error in the control for different regularization parameters (*right*)

occurs at the origin. We see that here the impact of the regularization parameter is less pronounced than in the first example (provided ε is chosen sufficiently small; see also Fig. 6 (right)). The benefit of adaptive versus uniform refinement is addressed in Fig. 6 (left) which displays on a logarithmic scale the discretization error in the control as a function of the degrees of freedom (dotted line: adaptive refinement, straight line: uniform refinement, $\varepsilon = 10^{-6}$). Finally, Fig. 6 (right) contains a comparison of the error in the control for the pure state constrained problem ($\epsilon = 0$) and its Lavrentiev regularizations ($\varepsilon = 10^{-2}, 10^{-4}, 10^{-6}$).

Acknowledgements The authors acknowledge support by the NSF under Grant No. DMS-0411403 and Grant No. DMS-0511611 as well as by the DFG within the Priority Program SPP 1253 'PDE Constrained Optimization'. The authors also would like to express their sincere thanks to the referees for most valuable suggestions.

References

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

2. Arada, N., Raymond, J.P.: Optimal control problems with mixed control-state constraints. SIAM J. Control Optim. **39**, 1391–1407 (2000)
3. Babuska, I., Strouboulis, T.: The Finite Element Method and Its Reliability. Clarendon Press, Oxford (2001)
4. Bangerth, W., Rannacher, R.: Adaptive Finite Element Methods for Differential Equations. Lectures in Mathematics. ETH-Zürich. Birkhäuser, Basel (2003)
5. Becker, R., Kapp, H., Rannacher, R.: Adaptive finite element methods for optimal control of partial differential equations: Basic concepts. SIAM J. Control Optim. **39**, 113–132 (2000)
6. Bergounioux, M., Kunisch, K.: On the structure of the Lagrange multiplier for state-constrained optimal control problems. Syst. Control Lett. **48**, 169–176 (2002)
7. Casas, E.: Control of an elliptic problem with pointwise state controls. SIAM J. Control Optim. **24**, 1309–1318 (1986)
8. Casas, E., Raymond, J.-P., Zidani, H.: Pontryagin's principle for local solutions of control problems with mixed control-state constraints. SIAM J. Control Optim. **39**, 1182–1203 (2000)
9. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**, 1106–1124 (1996)
10. Eriksson, K., Estep, D., Hansbo, P., Johnson, C.: Computational Differential Equations. Cambridge University Press, Cambridge (1995)
11. Gaevskaya, A., Hoppe, R.H.W., Iliash, Y., Kieweg, M.: A posteriori error analysis of control constrained distributed and boundary control problems. In: Fitzgibbon, W., et al. (eds.) Proceedings of Conference on Advances in Scientific Computing, Moscow, Russia, pp. 85–108. Russian Academy of Sciences, Moscow (2006)
12. Gaevskaya, A., Hoppe, R.H.W., Iliash, Y., Kieweg, M.: Convergence analysis of an adaptive finite element method for distributed control problems with control constraints. In: Leugering, G., et al. (eds.) Proceedings of Conference on Optimal Control for PDEs, Oberwolfach, Germany, pp. 47–68. Birkhäuser, Basel (2007)
13. Günther, A., Hinze, M.: A posteriori error control of a state constrained elliptic control problem. J. Numer. Math. **16** (2008, in press)
14. Hintermüller, M., Hoppe, R.H.W.: Goal-oriented adaptivity in control constrained optimal control of partial differential equations. SIAM J. Control Optim. **47**, 1721–1743 (2008)
15. Hintermüller, M., Hoppe, R.H.W., Iliash, Y., Kieweg, M.: An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. ESAIM COCV **14**, 540–560 (2008). doi:10.1051/cocv:2007057
16. Hintermüller, M., Kunisch, K.: Stationary state constrained optimal control problems. Preprint IFB, Report No. 3, Karl-Franzens-University of Graz (2006)
17. Hoppe, R.H.W., Iliash, Y., Iyyunni, C., Sweilam, N.: A posteriori error estimates for adaptive finite element discretizations of boundary control problems. J. Numer. Anal. **14**, 57–82 (2006)
18. Hoppe, R.H.W., Kieweg, M.: A posteriori error estimation of finite element approximations of pointwise state constrained distributed control problems. Preprint No. 16, Inst. of Math., Univ. of Augsburg (2007)
19. Li, R., Liu, W., Ma, H., Tang, T.: Adaptive finite element approximation for distributed elliptic optimal control problems. SIAM J. Control Optim. **41**, 1321–1349 (2002)
20. Liu, W., Yan, N.: A posteriori error estimates for distributed optimal control problems. Adv. Comput. Math. **15**, 285–309 (2001)
21. Liu, W., Yan, N.: A posteriori error estimates for convex boundary control problems. Preprint, Institute of Mathematics and Statistics, University of Kent, Canterbury (2003)
22. Meyer, C., Prüfert, U., Tröltzsch, F.: On two numerical methods for state-constrained elliptic control problems. Preprint, Department of Mathematics, Berlin University of Technology (2005)
23. Meyer, C., Rösch, A., Tröltzsch, F.: Optimal control problems of PDEs with regularized pointwise state constraints. Comput. Optim. Appl. **33**, 209–228 (2006)
24. Morin, P., Nochetto, R.H., Siebert, K.G.: Data oscillation and convergence of adaptive FEM. SIAM J. Numer. Anal. **38**(2), 466–488 (2000)
25. Neittaanmäki, P., Repin, S.: Reliable Methods for Mathematical Modelling. Error Control and a Posteriori Estimates. Elsevier, New York (2004)
26. Prüfert, U., Tröltzsch, F., Weiser, M.: The convergence of an interior point method for an elliptic control problem with mixed control-state constraints. Comput. Optim. Appl. **39**, 369–383 (2008)
27. Rösch, A., Tröltzsch, F.: Sufficient second order optimality conditions for an elliptic optimal control problem with pointwise control-state constraints. SIAM J. Optim. **17**, 776–794 (2006)

28. Tröltzsch, F.: A minimum principle and a generalized bang-bang principle for a distributed optimal control problem with constraints on the control and the state. Z. Angew. Math. Mech. **59**, 737–739 (1979)
29. Tröltzsch, F.: Optimale Steuerung partieller Differentialgleichungen. Vieweg, Wiesbaden (2005)
30. Tröltzsch, F.: Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints. SIAM J. Optim. **15**, 616–634 (2005)
31. Verfürth, R.: A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques. Wiley/Teubner, New York/Stuttgart (1996)
32. Vexler, B., Wollner, W.: Adaptive finite element methods for elliptic optimization problems with control constraints. SIAM J. Control Optim. **47**, 509–534 (2008)