Adaptive finite elements for optimally controlled elliptic variational inequalities of obstacle type

A. Gaevskaya, M. Hintermüller, Ronald H. W. Hoppe, C. Löbhard

Adaptive Finite Elements for Optimally Controlled Elliptic Variational Inequalities of Obstacle Type

A. Gaevskaya, M. Hintermüller, R.H.W. Hoppe, and C. Löbhard

Abstract We are concerned with the numerical solution of distributed optimal control problems for second order elliptic variational inequalities by adaptive finite element methods. Both the continuous problem as well as its finite element approximations represent subclasses of Mathematical Programs with Equilibrium Constraints (MPECs) for which the optimality conditions are stated by means of stationarity concepts in function space (Hintermüller and Kopacka, SIAM J. Optim. 20:868–902, 2009) and in a discrete, finite dimensional setting (Scheel and Scholtes, Math. Oper. Res. 25:1–22, 2000) such as (ε-almost, almost) C- and S-stationarity. With regard to adaptive mesh refinement, in contrast to the work in (Hintermüller, ESAIM Control Optim. Calc. Var., 2012, submitted) which adopts a goal oriented dual weighted approach, we consider standard residual-type a posteriori error estimators.

The first main result states that for a sequence of discrete C-stationary points there exists a subsequence converging to an almost C-stationary point, provided the associated sequence of nested finite element spaces is limit dense in its continuous counterpart. As the second main result, we prove the reliability and efficiency of the residual-type a posteriori error estimators. Particular emphasis is put on the approximation of the reliability and efficiency related consistency errors by

A. Gaevskaya
Institute of Mathematics, University of Augsburg Universitätsstr. 14, D-86159 Augsburg, Germany
e-mail: gaevskaya@math.uni-augsburg.de

M. Hintermüller • C. Löbhard
Department of Mathematics, Humboldt-Universität zu Berlin Unter den Linden 6, D-10099 Berlin, Germany
e-mail: hint@math.hu-berlin.de; loebhard@math.hu-berlin.de

R.H.W. Hoppe (✉)
Institute of Mathematics, University of Augsburg Universitätsstr. 14, D-86159 Augsburg, Germany

Department of Mathematics, University of Houston 659 P.G. Hoffman, Houston, TX 77204-3008, USA
e-mail: hoppe@math.uni-augsburg.de; rohop@math.uh.edu

heuristically motivated computable quantities and on the approximation of the continuous active, strongly active, and inactive sets by their discrete counterparts.

A detailed documentation of numerical results for two representative test examples illustrates the performance of the adaptive approach.

1 Introduction

This paper is devoted to the study of adaptive finite element methods for the approximation of optimally controlled elliptic variational inequalities of obstacle type. Such problems can be formulated as Mathematical Programs with Complementarity Constraints (MPCCs) representing a subclass of Mathematical Programs with Equilibrium Constraints (MPECs) which have been investigated both in function space [4, 18, 27, 30–34] as well as in finite dimensions [10, 26, 29, 36–38]. Due to the inherent non-convexity and non-differentiability, MPECs are not amenable to classical approaches from optimal control/optimization theory and thus require tools from non-smooth analysis such as generalized derivatives. In particular, this leads to optimality systems in terms of various stationarity concepts such as C(larke)-stationarity and S(trong)-stationarity (cf., e.g., [18] for MPECs in function space). For the spatial discretization of the problems we use continuous, piecewise linear finite elements with respect to an adaptively generated hierarchy of geometrically conforming simplicial triangulations of the computational domain. Although adaptive mesh refinement relying on various a posteriori error estimators has been extensively studied for elliptic variational inequalities (cf., e.g., [2, 6–8,22,24,35,42,43]) as well as for unconstrained and control and/or state constrained elliptic optimal control problems (cf., e.g., [5, 12, 14–17, 19, 20, 23, 40, 45]), the only adaptive approach for optimally controlled elliptic variational inequalities we are aware of is the one in [21] based on goal oriented dual weighted residuals. Instead, here we study standard residual-type a posteriori error estimators in terms of element and edge residuals and prove both reliability and efficiency up to consistency errors and data oscillations.

The paper is organized as follows: After introducing basic notations and some preliminary results, in Sect. 2 we state the distributed optimal control problem for a second order elliptic variational inequality of obstacle type, specify the associated active and inactive sets including a possible set of biactivity in case of a lack of strict complementarity, and introduce the relevant stationarity concepts in function space. Section 3 is devoted to the finite element approximation of the problem under consideration giving rise to a discrete optimally controlled variational inequality,

the specification of the discrete active and inactive sets, and the discrete stationarity concepts. Particular emphasis is put on suitable extensions of the discrete Lagrange multipliers which will play a significant role both in the subsequent convergence analysis and in the a posteriori error analysis. In Sect. 4, we prove the first main result of this paper. Under the assumption that the sequence of nested finite element spaces is limit dense in the function space for the continuous state and adjoint state, we show that for a bounded sequence of discrete C-stationary points there exists a subsequence which converges to an almost C-stationary point (cf. Theorem 4.2). Section 5 is concerned with the a posteriori error analysis based on residual-type a posteriori error estimators.

As the second main result, we establish reliability and efficiency of the error estimator up to consistency errors due to a mismatch in complementarity and data oscillations (cf. Theorem 5.1 and Theorem 5.1). Since in the original formulation the consistency errors are not a posteriori, we provide heuristically motivated fully computable quantities in terms of approximations of the characteristic functions of the continuous active and inactive sets as well as of the continuous states and multipliers (cf. Sect. 5.4). The final Sect. 6 contains a documentation of numerical results for two representative test examples, one with strict complementarity and the other without. The numerical results exhibit experimental convergence rates that asymptotically approach the expected optimal convergence rates. Moreover, it is shown that at least some of the heuristically derived approximations of the consistency errors provide close upper bounds.

2 The Optimal Control Problem and Stationarity Concepts

2.1 Notations and Preliminaries

For a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$, we denote by $\mathcal{D}(\Omega)$ the space of infinitely often continuously differentiable functions with compact support in Ω, and we refer to $\mathcal{D}(\Omega)'$ as the dual space of distributions. Further, we adopt standard notation from Lebesgue and Sobolev space theory (cf., e.g., [1]). In particular, for $D \subseteq \Omega$, we denote by $L^2(D)$ the Hilbert space of square integrable functions on D with inner product $(\cdot, \cdot)_{0,D}$ and associated norm $\| \cdot \|_{0,D}$. $L^2(D)_+$ refers to the positive cone of $L^2(D)$ with respect to the partial order on $L^2(D)$, i.e., $L^2(D)_+ := \{v \in L^2(D) \mid v \geq 0 \text{ a.e. in } D\}$. For $k \in \mathbb{N}$, we denote by $H^k(D)$ the Sobolev space with inner product $(\cdot, \cdot)_{k,D}$, seminorm $|\cdot|_{k,D}$, and norm $\| \cdot \|_{k,D}$. We define $H_0^k(D)$ as the closure of $\mathcal{D}(D)$ in $H^k(D)$ and refer to $H^{-k}(D)$ as the dual space. In particular, we set $V := H_0^1(\Omega)$ so that $V^* = H^{-1}(\Omega)$, and we refer to $\langle \cdot, \cdot \rangle$ as the dual pairing between V^* and V. We define V_+ as the positive cone of V with respect to the partial ordering inherited from $L^2(\Omega)$, i.e., $V_+ := \{v \in V \mid v \geq 0 \text{ a.e. in } \Omega\}$ and we refer to V_+^* as the positive cone of V^*, i.e., $V_+^* := \{\lambda \in V^* \mid \langle \lambda, v \rangle \geq 0 \text{ for all } v \in V_+\}$.

As far as localizations of functionals $\lambda \in V^*$ are concerned, we note that for a distribution $T \in \mathcal{D}(\Omega)'$ and an open set $\omega \subseteq \Omega$ it is said that $T = 0$ on ω, if $T(v) = 0$ for all $v \in \mathcal{D}(\Omega)$ with $\mathrm{supp}(v) \subseteq \omega$ (cf., e.g., [41]). Further, denoting by $\mathcal{O}_T$ the maximal open set where $T = 0$, the support of T is defined by $\mathrm{supp}(T) := \Omega \setminus \mathcal{O}_T$. We set $V_\omega := \{v \in V \mid \mathrm{supp}(v) \subseteq \bar{\omega}\}$. Since a functional $\lambda \in V^*$ can be viewed as a distribution, we introduce the set

$$V_{\omega,0} := \{v \in V_\omega \mid v|_{\Omega \setminus \omega} = 0 \text{ a.e.}, v|_\omega \in H_0^1(\omega)\} \tag{2.1}$$

of test functions and say that $\lambda = 0$ on ω, if $\langle \lambda, v \rangle = 0$ for all $v \in V_{\omega,0}$ (for alternative definitions see [18]). Further, we say that $\lambda \geq 0$ ($\lambda \leq 0$) on ω, if $\langle \lambda, v \rangle \geq 0$ ($\langle \lambda, v \rangle \leq 0$) for all $v \in V_{\omega,0} \cap V_+$. The support of $\lambda \in V^*$ is defined by

$$\mathrm{supp}(\lambda) := \Omega \setminus \mathcal{O}_\lambda. \tag{2.2}$$

We note that $V_{\omega,0} \subseteq V_\omega$. If ω is Lipschitz, we have $V_{\omega,0} = V_\omega$ (cf., e.g., [27]).

In the sequel, we will need characterizations of functionals $\lambda \in V^*$ with restricted support. To this end, we first consider the question of extension by zero of $v|_\omega, v \in V$, for $\omega \subseteq \Omega$. If ω is Lipschitz, we denote by $\partial\omega^0(v)$ that part of the boundary $\partial\omega$ such that $v = 0$ a.e. on $\partial\omega^0(v)$ and $v \neq 0$ a.e. on $\partial\omega \setminus \partial\omega^0(v)$. Then, for $v \in V$ and an open Lipschitz domain $\omega \subseteq \Omega$ there exist an open Lipschitz set $\tilde{\omega}$ such that $\omega \subseteq \tilde{\omega} \subseteq \Omega$ and a function $v_\omega^{ext} \in V_{\tilde{\omega},0}$ with $v_\omega^{ext}|_\omega = v|_\omega$ a.e. in ω. If $\partial\omega^0(v) \neq \emptyset$, $\tilde{\omega}$ can be chosen so that $\partial\tilde{\omega} \cap \partial\omega = \partial\omega^0(v)$. If ω is non-Lipschitz, the previous property remains true, if ω is replaced by $\mathrm{Lip}(\omega)$ which is the minimal open Lipschitz set with $\omega \subseteq \mathrm{Lip}(\omega)$.

The following result allows to make use of the restricted support of functionals in V^* to describe their action on functions from V.

Proposition 2.1. *For $\lambda \in V^*$ set $\Lambda := int(supp(\lambda))$, if $supp(\lambda)$ is Lipschitz, and $\Lambda := Lip(int(supp(\lambda))$, otherwise. For any $v \in V$ there exist an open Lipschitz set $\tilde{\Lambda}$ with $\Lambda \subseteq \tilde{\Lambda} \subseteq \Omega, \partial\tilde{\Lambda} \cap \partial\Lambda = \partial\Lambda^0(v)$ and a function $v_\Lambda^{ext} \in V_{\tilde{\Lambda},0}$ such that $v_\Lambda^{ext}|_\Lambda = v|_\Lambda$ a.e. in Λ and*

$$\langle \lambda, v \rangle = \langle \lambda, v_\Lambda^{ext} \rangle. \tag{2.3}$$

Proof. Since Λ is an open Lipschitz domain, there exist $\tilde{\Lambda}$ with $\Lambda \subseteq \tilde{\Lambda} \subseteq \Omega, \partial\tilde{\Lambda} \cap \partial\Lambda = \partial\Lambda^0(v)$ and a function $v_\Lambda^{ext} \in V_{\tilde{\Lambda}}$ such that $v_\Lambda^{ext}|_\Lambda = v|_\Lambda$ a.e. in Λ. Hence, it suffices to prove (2.3). Let $\bar{v} \in V_{\Omega \setminus \Lambda, 0}$ be defined according to

$$\bar{v} = \begin{cases} 0 \text{ in } \Lambda, \\ v - v_\Lambda^{ext} \text{ in } int(\Omega \setminus \Lambda). \end{cases}$$

In view of the construction of Λ it holds $\mathrm{int}(\Omega \setminus \Lambda) \subseteq \mathcal{O}_\lambda$, where $\mathcal{O}_\lambda$ is the maximal open set where λ vanishes, and hence, $\langle \lambda, \bar{v} \rangle = 0$. It follows that $\langle \lambda, v \rangle = \langle \lambda, v_\Lambda^{ext} \rangle + \langle \lambda, \bar{v} \rangle = \langle \lambda, v_\Lambda^{ext} \rangle$. $\qquad\square$

Remark 2.2. We note that $\langle \lambda, v \rangle = \langle \lambda, v|_{\mathrm{supp}(\lambda)} \rangle$ only if $v \in V_{\mathrm{supp}(\lambda),0}$. Otherwise, λ 'reaches' the values of v slightly outside of $\mathrm{int}(\mathrm{supp}(\lambda))$.

2.2 The Optimal Control Problem

Given a domain $\Omega \subset \mathbb{R}^2$ with boundary $\Gamma = \partial\Omega$, a bilinear form $a(\cdot,\cdot): V \times V \to \mathbb{R}$, where $V := H_0^1(\Omega)$, a desired state y^d, a shift control u^d, a force density f, an upper obstacle ψ, and a regularization parameter α such that

$$\Omega \text{ is a bounded, polygonal Lipschitz domain,} \tag{2.4a}$$

$$a(\cdot,\cdot): V \times V \to \mathbb{R} \text{ is symmetric, bounded and V-elliptic, i.e.,}$$

$$|a(y,v)| \leq C \|y\|_{1,\Omega} \|v\|_{1,\Omega}, \quad \gamma \|y\|_{1,\Omega}^2 \leq a(y,y), \ \gamma, C > 0, \tag{2.4b}$$

$$y^d \in L^2(\Omega), \quad u^d \in L^2(\Omega), \quad f \in L^2(\Omega), \tag{2.4c}$$

$$\psi \in V, \quad \alpha > 0, \tag{2.4d}$$

we consider the following distributed optimal control problem with a variational inequality constraint:

$$\text{Minimize} \quad J(y,u) := \frac{1}{2} \|y - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \|u - u^d\|_{0,\Omega}^2 \tag{2.5a}$$

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

$$\text{subject to} \quad a(y, y-v) \leq (f+u, y-v)_{0,\Omega}, \ v \in K, \tag{2.5b}$$

$$K := \{v \in V \mid v \leq \psi \text{ a.e. in } \Omega\}.$$

Here, J is referred to as the objective functional, y and u stand for the state and the control, and K denotes the constraint set which makes (2.5b) to a variational inequality of obstacle type. We further denote by $A : V \to V^*$ the bounded linear operator associated with the bilinear form $a(\cdot,\cdot)$. Although the subsequent analysis can be carried out for a general second order elliptic differential operator in divergence form, in the sequel we will restrict ourselves to the case $A = -\Delta$.

The optimal control problem (2.5) can be equivalently written in the so-called control-reduced form by means of the control-to-state map $S : L^2(\Omega) \to V$ which assigns to a control $u \in L^2(\Omega)$ the unique solution of the variational inequality (2.5b):

Minimize $\qquad J^{red}(u) := \frac{1}{2}\|Su - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2}\|u - u^d\|_{0,\Omega}^2$ (2.6)

over $\qquad u \in L^2(\Omega)$.

The existence of minimizers for (2.5) is guaranteed by the following result:

Theorem 2.3. *Under the assumptions* (2.4) *on the data, the optimal control problem* (2.5) *admits an optimal solution.*

Proof. We refer to [4, 31]. $\qquad\square$

By introducing a slack variable $\sigma \in V^*$, the variational inequality constraint (2.5b) can be equivalently reformulated in terms of a complementarity system so that (2.5) reads:

Minimize $\qquad J(y, u) := \frac{1}{2}\|y - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2}\|u - u^d\|_{0,\Omega}^2$ (2.7a)

over $\qquad (y, \sigma, u) \in V \times V^* \times L^2(\Omega)$,

subject to $\qquad a(y, v) = (f + u, v)_{0,\Omega} - \langle\sigma, v\rangle,\ v \in V$, (2.7b)

$\qquad\qquad \psi - y \in V_+,\ \sigma \in V_+^*,\ \langle\sigma, \psi - y\rangle = 0$.

The problem (2.7) is commonly referred to as a Mathematical Program with Complementarity Constraints (MPCC).

2.3 Continuous Active and Inactive Sets

For given $u \in L^2(\Omega)$, (2.5b) represents an obstacle problem which, under the assumptions (2.4), admits a unique solution $(y, \sigma) \in V \times V^*$ (cf., e.g., [25]). The complementary behavior of y and σ according to (2.7b) gives rise to the following definitions:

Definition 2.4. We define the active set $\mathcal{A}$ as the maximal open subset $D \subseteq \Omega$ such that $\psi - y = 0$ a.e. in D. We denote by $\mathcal{I} := \bigcup_{\varepsilon>0} B_\varepsilon(\psi - y)$ the inactive set, where $B_\varepsilon(\psi - y)$ is the maximal open set $D \subseteq \Omega$ such that $\psi - y \geq \varepsilon$ a.e. in D. Finally, $\mathcal{F}(y) := \Omega \setminus (\mathcal{A} \cup \mathcal{I})$ is said to be the free boundary with respect to y.

Obviously, the sets $\mathcal{A}, \mathcal{I}$, and $\mathcal{F}(y)$ provide a partition of Ω, i.e., it holds $\Omega = \mathcal{A} \cup \mathcal{I} \cup \mathcal{F}(y)$. An alternative partition can be achieved in terms of properties of the multiplier σ:

Definition 2.5. The zero set $\mathcal{Z}$ is defined as the maximal open set D such that $\langle \sigma, v \rangle = 0$ for all $v \in V_{D,0}$, whereas the set $\mathcal{C} := \text{int}(\text{supp}(\sigma))$ is referred to as the strongly active set (for the definitions of $V_{D,0}$ and $\text{supp}(\sigma)$ see (2.1) and (2.2) in Sect. 2.1). The set $\mathcal{F}(\sigma) := \Omega \setminus (\mathcal{Z} \cup \mathcal{C})$ is called the free boundary with respect to σ.

Remark 2.6. If in addition to the assumptions (2.4) on the data of the problem we suppose

$$\Omega \subset \mathbb{R}^2 \text{ is convex or of class } \mathcal{C}^{1,1}, \tag{2.8a}$$

$$\psi \in V \cap H^2(\Omega), \tag{2.8b}$$

the solution of the obstacle problem satisfies $(y, \sigma) \in V \cap H^2(\Omega) \times L^2(\Omega)$. In this regular case, we define the active and the inactive set according to $\mathcal{A}_{reg} := \text{int}(\{x \in \Omega \mid \psi(x) - y(x) = 0\}), \mathcal{I}_{reg} := \text{int}(\Omega \setminus \mathcal{A}_{reg})$. Moreover, the zero set $\mathcal{Z}_{reg}$ is the maximal open set $D \subseteq \Omega$ such that $\sigma = 0$ a.e. in D, and the strongly active set is given by $\mathcal{C}_{reg} := \text{int}(\Omega \setminus \mathcal{Z}_{reg})$.

The special case where $\psi - y$ and the slack variable σ are simultaneously zero in some subset of Ω is taken care of by the definition of the so-called biactive set:

Definition 2.7. The set $\mathcal{B} := \text{int}(\mathcal{A} \setminus \mathcal{C})$ is called the biactive set. If $\text{meas}(\mathcal{B}) = 0$, the solution of the obstacle problem is said to satisfy the strict complementarity condition. Otherwise, it is said that the solution exhibits a lack of strict complementarity.

The following results which were proven in [11] provide characterizations of the active set, the inactive set, the zero set, and of the slack variable σ. They all refer to the complementarity conditions (2.7b).

Proposition 2.8. *For any $v \in V_+$ let the zero set $\Omega^0(v)$ be the maximal open set $D \subseteq \Omega$ such that $v = 0$ a.e. in D and let $\Omega^+(v) := \bigcup_{\varepsilon > 0} B_\varepsilon(v)$ be the positive set, where $B_\varepsilon(v)$ is the maximal open set $D \subseteq \Omega$ such that $v \geq \varepsilon$ a.e. in D. Then, it holds*

$$\mathcal{A} = \Omega^0(\psi - y), \quad \mathcal{I} = \Omega^+(\psi - y). \tag{2.9}$$

Moreover, for any $v \in V_+$ such that $\langle \sigma, v \rangle = 0$ it holds

$$\Omega^+(v) \subseteq \mathcal{Z}. \tag{2.10}$$

Corollary 2.9. *For any $v \in V$ such that $\langle \sigma, v^+ \rangle = 0$ and $\langle \sigma, v^- \rangle = 0$ it holds*

$$v = 0 \text{ in } \mathcal{C} \quad \text{and} \quad \langle \sigma, v \rangle = 0. \tag{2.11}$$

Proposition 2.10. *The slack variable σ satisfies*

$$\sigma = 0 \text{ in } \mathcal{I}, \text{ i.e., } \mathcal{C} \subseteq \mathcal{A}, \tag{2.12a}$$

$$\sigma = f + u - A\psi \text{ in } \mathcal{A}. \tag{2.12b}$$

Corollary 2.11. *A lack of strict complementarity of the solution of the obstacle problem occurs if and only if there exists a set $\mathcal{B} \subseteq \mathcal{A}$ such that $f + u - A\psi = 0$ in $\mathcal{B}$. Hence, there must hold $\langle A\psi, v \rangle = (f + u, v)_{0,\mathcal{B}}$, i.e., $A\psi|_{\mathcal{B}} \in L^2(\mathcal{B})$.*

2.4 Stationarity Concepts

In this subsection, we present various concepts of stationarity associated with the optimal control problem (2.5). We note that for MPCC in function space the concepts of C(larke)-stationarity and S(trong)-stationarity have been introduced in [18].

Definition 2.12. For $(y, \sigma, u) \in V \times V^* \times L^2(\Omega)$ assume that there exists a pair $(p, \mu) \in V \times V^*$ such that the following conditions hold true

$$a(y, v) = (f + u, v)_{0,\Omega} - \langle \sigma, v \rangle, \ v \in V, \tag{2.13a}$$

$$\psi - y \in V_+, \ \sigma \in V_+^*, \ \langle \sigma, \psi - y \rangle = 0, \tag{2.13b}$$

$$a(p, v) = (y^d - y, v)_{0,\Omega} - \langle \mu, v \rangle, \ v \in V, \tag{2.13c}$$

$$p = \alpha \, (u - u^d), \tag{2.13d}$$

$$p = 0 \text{ a.e. in } \mathcal{C}, \tag{2.13e}$$

$$\langle \mu, p \rangle \geq 0, \tag{2.13f}$$

$$\langle \mu, \psi - y \rangle = 0. \tag{2.13g}$$

A triple $(y, \sigma, u) \in V \times V^* \times L^2(\Omega)$ is called

(i) an ε-almost C-stationary point of (2.5), if (2.13a)–(2.13g) hold true and the pair $(p, \mu) \in V \times V^*$ satisfies:
For all $\varepsilon > 0$ there exists $U_\varepsilon \subseteq \mathcal{I}$ with $\text{meas}(\mathcal{I} \setminus U_\varepsilon) \leq \varepsilon$ such that

$$\langle \mu, v \rangle = 0, \ v \in V_{U_\varepsilon}, \tag{2.13h}$$

(ii) an almost C-stationary point of (2.5), if (2.13a)–(2.13g) hold true and the pair $(p, \mu) \in V \times V^*$ fulfills

$$\langle \mu, v \rangle = 0, \ v \in V_{\mathcal{I},0}, \tag{2.13i}$$

(iii) a C-stationary point of (2.5), if (2.13a)–(2.13g) hold true and the pair $(p, \mu) \in V \times V^*$ satisfies

$$\langle \mu, v \rangle = 0, \ v \in V_{\mathcal{I}}. \tag{2.13j}$$

Definition 2.13. Let $(y, \sigma, u) \in V \times V^* \times L^2(\Omega)$ be an ε-almost C-stationary point (almost C-stationary, C-stationary) point of (2.5). Then, the triple (y, σ, u) is said to be an ε-almost S-stationary (almost S-stationary, S-stationary) point of (2.5), if the pair $(p, \mu) \in V \times V^*$ additionally satisfies

$$\langle \mu, v \rangle \geq 0, \ v \in V_{\mathcal{B}} \cap V_+, \tag{2.14a}$$

$$p \geq 0 \text{ a.e. in } \mathcal{B}. \tag{2.14b}$$

Remark 2.14. In the Definitions 2.12 and 2.13, the function $p \in V$ is referred to as the adjoint state and Eq. (2.13c) is called the adjoint state equation. The functional $\mu \in V^*$ is said to be the Lagrange multiplier associated with the adjoint state equation.

Remark 2.15. In the previous Definitions 2.12 and 2.13, S-stationarity is the strongest and ε-almost C-stationarity is the weakest concept. The hierarchy of the above introduced stationarity concepts is displayed in the commuting diagram below:

$$\text{S-stationarity} \implies \text{almost S-stationarity} \implies \varepsilon\text{-almost S-stationarity}$$
$$\Downarrow \qquad\qquad\qquad \Downarrow \qquad\qquad\qquad \Downarrow$$
$$\text{C-stationarity} \implies \text{almost C-stationarity} \implies \varepsilon\text{-almost C-stationarity}$$

The following result reveals local properties of almost C-stationary points with respect to the sets $\mathcal{C}, \mathcal{B}$, and $\mathcal{I}$ defined in Sect. 2.3.

Proposition 2.16. *Let* $(y, \sigma, u) \in V \times V^* \times L^2(\Omega)$ *be an almost C-stationary point of* (2.5) *and let* $(p, \mu) \in V \times V^*$ *be the associated adjoint state and Lagrange multiplier. Then, with regard to the strongly active set* $\mathcal{C}$*, the biactive set* $\mathcal{B}$*, and the inactive set* $\mathcal{I}$ *it holds*

	$\mathcal{C}$	$\mathcal{B}$	$\mathcal{I}$
y	$= \psi$ a.e.	$= \psi$ a.e.	$-$
p	$= 0$ a.e.	$= -\alpha \, (\Delta\psi + f + u^d)$ a.e.	$-$
u	$= u^d$ a.e.	$= -\Delta\psi - f$ a.e.	$-$
σ	$= f + u^d + \Delta\psi$	$= 0$	$= 0$
μ	$= y^d - \psi$	$= y^d - \psi + \alpha \, \Delta(\Delta\psi + f + u^d)$	$= 0$

Proof. In view of the definitions of the sets $\mathcal{A}, \mathcal{C}$, and $\mathcal{B}$, we obviously have $y = \psi$ a.e. in $\mathcal{A} = \mathcal{C} \cup \mathcal{B}$. Taking $V_{\mathcal{B},0} \subseteq V_{\mathcal{Z},0}$ and $V_{\mathcal{I},0} \subseteq V_{\mathcal{Z},0}$ into account, it holds

$$\langle \sigma, v \rangle = 0, \quad v \in V_{\mathcal{B},0}, \qquad \langle \sigma, v \rangle = 0, \quad v \in V_{\mathcal{I},0}.$$

Further, due to (2.13d) and (2.13e)

$$p = 0 \text{ a.e. in } \mathcal{C}, \qquad u = u^d \text{ a.e. in } \mathcal{C}.$$

Hence, (2.13c) implies

$$\langle \mu, v \rangle = (y^d - \psi, v)_{0,\mathcal{C}}, \quad v \in V_{\mathcal{C},0},$$

i.e., $\mu|_{\mathcal{C}} = y^d - \psi \in L^2(\mathcal{C})$. By (2.13a) it holds

$$\langle \sigma, v \rangle = (f + u^d, v)_{0,\mathcal{C}} - a(\psi, v), \quad v \in V_{\mathcal{C},0},$$

whence $\sigma = f + u^d + \Delta\psi$ a.e. in $\mathcal{C}$. Moreover, in $\mathcal{B}$ we have

$$(f + u, v)_{0,\mathcal{B}} = (\nabla\psi, \nabla v)_{0,\mathcal{B}}, \quad v \in V_{\mathcal{B},0}.$$

Consequently, the weak divergence of $\nabla\psi$ in $\mathcal{B}$ exists and equals $-(f + u)|_{\mathcal{B}} \in L^2(\mathcal{B})$. It follows that $-\Delta\psi = f + u$ a.e. in $\mathcal{B}$. Hence,

$$u = -\Delta\psi - f \text{ a.e. in } \mathcal{B},$$

and, due to (2.13d)

$$p = -\alpha \left(\Delta\psi + f + u^d \right) \text{ a.e. in } \mathcal{B}.$$

The previous equation gives rise to $\Delta\psi + f + u^d \in H^1(\mathcal{B})$. Hence, (2.13c) implies

$$\langle \mu, v \rangle = (y^d - \psi, v)_{0,\mathcal{B}} + \alpha\, a(\Delta\psi + f + u^d, v), \quad v \in V_{\mathcal{B},0}.$$

$\square$

Stationarity in the Regular Case. If in addition to the assumptions (2.4) on the data of the problem we suppose

$$\Omega \text{ is either convex and polygonal or of class } C^{1,1}, \tag{2.15a}$$

$$\psi \in V \cap H^2(\Omega), \tag{2.15b}$$

for fixed $u \in L^2(\Omega)$ the solution (y, σ) of the obstacle problem belongs to $V \cap H^2(\Omega) \times L^2(\Omega)$. In this regular case, the optimal control problem (2.5) can be

rewritten according to:

$$\text{Minimize} \qquad J(y,u) := \frac{1}{2}\,\|y - y^d\|_{0,\Omega}^2 + \frac{\alpha}{2}\,\|u - u^d\|_{0,\Omega}^2 \qquad (2.16a)$$

$$\text{over} \qquad (y, \sigma, u) \in V \times L^2(\Omega) \times L^2(\Omega),$$

$$\text{subject to} \qquad a(y, y - v) = (f + u - \sigma, v)_{0,\Omega}, \ v \in V, \qquad (2.16b)$$

$$\psi - y \geq 0 \text{ a.e. in } \Omega, \ \sigma \geq 0 \text{ a.e. in } \Omega, \ (\sigma, \psi - y)_{0,\Omega} = 0.$$

The stationarity concepts can be formulated as in Definitions 2.12 and 2.13.

3 Finite Element Approximation

For a null sequence $\mathcal{H}$ of positive real numbers we assume $\{\mathcal{T}_h(\Omega)\}_{h\in\mathcal{H}}$ to be a shape regular family of geometrically conforming simplicial triangulations of the computational domain Ω. For $D \subset \bar{\Omega}$, we denote by $\mathcal{N}_h(D), \mathcal{E}_h(D)$, and $\mathcal{T}_h(D)$ the sets of nodal points, edges, and triangles of $\mathcal{T}_h(\Omega)$ in D. For $T \in \mathcal{T}_h(\Omega)$, we refer to h_T and $|T|$ as the diameter and the area of T, whereas for $E \in \mathcal{E}_h(\bar{\Omega})$ we denote by h_E the length of the edge E. We further introduce the following patches of triangles of $\mathcal{T}_h(\Omega)$:

$$\omega_a := \bigcup\{T \in \mathcal{T}_h(\Omega) \mid a \in N_h(T)\}, \qquad (3.1a)$$

$$\omega_E := \bigcup\{T \in \mathcal{T}_h(\Omega) \mid E \in \mathcal{E}_h(T)\}, \qquad (3.1b)$$

$$\omega_T := \bigcup\{T' \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T') \cap \mathcal{N}_h(T) \neq \emptyset\}, \qquad (3.1c)$$

and the following set of edges of $\mathcal{E}_h(\Omega)$:

$$\mathcal{E}_h^a := \bigcup\{E \in \mathcal{E}_h(\Omega) \mid a \in \mathcal{N}_h(E)\}. \qquad (3.2)$$

Moreover, for $T \in \mathcal{T}_h(\Omega)$ we refer to $P_k(T), k \in \mathbb{N}_0$, as the linear space of polynomials of degree $\leq k$ on T, and we define

$$S_h^{(1)} := \{v_h \in C(\bar{\Omega}) \mid v_h|_T \in P_1(T), \ T \in \mathcal{T}_h(\Omega)\} \qquad (3.3)$$

as the finite element space of continuous piecewise linear functions. We set

$$V_h := \{v_h \in S_h^{(1)} \mid v_h|_\Gamma = 0\} \qquad (3.4)$$

and denote by $\varphi_h^{(a)}$ the nodal basis function associated with $a \in \mathcal{N}_h(\Omega)$ such that $V_h = \text{span}(\{\varphi_h^{(a)} \mid a \in \mathcal{N}_h(\Omega)\})$ with $\dim V_h = N_h := \text{card}(\mathcal{N}_h(\Omega))$. As the dual space of V_h we consider linear combinations of the Dirac delta functionals δ_a associated with $a \in \mathcal{N}_h(\Omega)$, i.e.,

$$M_h := \{\lambda_h \in \mathcal{M}(\bar{\Omega}) \mid \lambda_h = \sum_{a \in \mathcal{N}_h(\Omega)} \lambda_h(a)\, \delta_a,\ \lambda_h(a) \in \mathbb{R}\}. \tag{3.5}$$

Here, $\mathcal{M}(\bar{\Omega})$ stands for the space of regular Borel measures.

3.1 The Discrete Optimal Control Problem

For the finite element approximation of the optimal control problem (2.5) we denote by $\psi_h \in V_h$ and $u_h^d \in S_h^{(1)}$ the interpolants of $\psi \in V$ and $u^d \in L^2(\Omega)$ in V_h and $S_h^{(1)}$ and refer to $y_h^d \in S_h^{(1)}$ and $f_h \in S_h^{(1)}$ as the L^2-projections of $y^d \in L^2(\Omega)$ and $f \in L^2(\Omega)$ onto $S_h^{(1)}$. Approximating the state $y \in V$ and the control $u \in L^2(\Omega)$ by finite element functions $y_h \in V_h$ and $u_h \in S_h^{(1)}$, the discrete optimal control problem is given as follows:

Minimize $\quad J_h(y_h, u_h) := \dfrac{1}{2}\, \|y_h - y_h^d\|_{0,\Omega}^2 + \dfrac{\alpha}{2}\, \|u_h - u_h^d\|_{0,\Omega}^2 \qquad (3.6\text{a})$

over $\quad (y_h, u_h) \in V_h \times S_h^{(1)},$

subject to $\quad a(y_h, y_h - v_h) \leq (f_h + u_h, y_h - v_h)_{0,\Omega},\ v_h \in K_h, \qquad (3.6\text{b})$

$\qquad\qquad K_h := \{v_h \in V_h \mid v_h \leq \psi_h \text{ in } \Omega\}.$

We refer to J_h and K_h as the discrete objective functional and the discrete constraint set and to y_h and u_h as the discrete state and the discrete control.

Denoting by $S_h : S_h^{(1)} \to V_h$ the discrete control-to-state map which assigns to a control $u_h \in S_h^{(1)}$ the unique solution $y_h \in V_h$ of the discrete variational inequality (3.6b), the control-reduced form of (3.6) reads:

Minimize $\quad J_h^{red}(u_h) := \dfrac{1}{2}\, \|S_h u_h - y_h^d\|_{0,\Omega}^2 + \dfrac{\alpha}{2}\, \|u_h - u_h^d\|_{0,\Omega}^2 \qquad (3.7)$

over $\quad u_h \in S_h^{(1)}.$

Theorem 3.1. *The discrete optimal control problem* (3.6) *admits an optimal solution* $(y_h, u_h) \in V_h \times S_h^{(1)}$.

Proof. The proof can be given in much the same way as that of Theorem 2.3. □

As in the continuous regime, by introducing a slack variable $\sigma_h \in M_h$, the discrete optimal control problem (3.6) can be equivalently reformulated as the discrete complementarity problem:

Minimize $\quad J_h(y_h, u_h) := \frac{1}{2} \, \|y_h - y_h^d\|_{0,\Omega}^2 + \frac{\alpha}{2} \, \|u_h - u_h^d\|_{0,\Omega}^2 \qquad$ (3.8a)

over $\qquad (y_h, \sigma_h, u_h) \in V_h \times M_h \times S_h^{(1)},$

subject to $\quad a(y_h, v_h) = (f_h + u_h, v_h)_{0,\Omega} - \langle\langle \sigma_h, v_h \rangle\rangle, \ v_h \in V_h, \qquad$ (3.8b)

$$y_h \in K_h, \ \sigma_h \in M_h \cap \mathcal{M}_+(\bar{\Omega}), \ \langle\langle \sigma_h, \psi_h - y_h \rangle\rangle = 0,$$

where $\langle\langle \cdot, \cdot \rangle\rangle$ refers to the dual pairing between $C(\bar{\Omega})$ and $\mathcal{M}(\bar{\Omega})$.

3.2 Discrete Active and Inactive Sets

For $v_h \in V_h$ we denote by

$$\mathcal{Z}_h(v_h) := \{a \in \mathcal{N}_h(\bar{\Omega}) \mid v_h(a) = 0\}, \quad \mathcal{C}_h(v_h) := \mathcal{N}_h(\bar{\Omega}) \setminus \mathcal{Z}_h(v_h) \qquad (3.9)$$

the sets of zero and non-zero nodal points with respect to $v_h \in V_h$, and we partition the triangulation $\mathcal{T}_h(\Omega)$ into the sets of zero, non-zero, and mixed triangles with respect to $v_h \in V_h$ according to

$$\mathcal{T}_h(\Omega) = \mathcal{T}_h^z(v_h) \cup \mathcal{T}_h^c(v_h) \cup \mathcal{T}_h^m(v_h), \qquad (3.10)$$

where

$$\mathcal{T}_h^z(v_h) := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subset \mathcal{Z}_h(v_h)\}, \qquad (3.11a)$$

$$\mathcal{T}_h^c(v_h) := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subset \mathcal{C}_h(v_h)\}, \qquad (3.11b)$$

$$\mathcal{T}_h^m(v_h) := \mathcal{T}_h(\Omega) \setminus (\mathcal{T}_h^z(v_h) \cup \mathcal{T}_h^c(v_h)). \qquad (3.11c)$$

Definition 3.2. For $y_h \in K_h$ we denote by $A_h := \mathcal{Z}_h(\psi_h - y_h) \cap \mathcal{N}_h(\Omega)$ and $I_h := \mathcal{C}_h(\psi_h - y_h) \cap \mathcal{N}_h(\Omega)$ the sets of active and inactive nodal points. A nodal point is said to be an isolated active (inactive) nodal point, if $\mathcal{N}_h(\omega_a) \setminus \{a\} \subset I_h \cup \mathcal{N}_h(\Gamma)$ $(\mathcal{N}_h(\omega_a) \setminus \{a\} \subset A_h \cup \mathcal{N}_h(\Gamma))$. Moreover, the sets

$$\mathcal{A}_h := \bigcup\{T \in \mathcal{T}_h^z(\psi_h - y_h)\}, \qquad (3.12a)$$

$$\overset{\circ}{\mathcal{I}}_h := \bigcup\{T \in \mathcal{T}_h^c(\psi_h - y_h)\}, \qquad (3.12b)$$

$$\mathcal{F}_h(y_h) := \bigcup\{T \in \mathcal{T}_h^m(\psi_h - y_h)\} \qquad (3.12c)$$

are referred to as the discrete active set, the discrete purely inactive set, and the discrete free boundary with respect to y_h. The set

$$\mathcal{I}_h := \overset{\circ}{\mathcal{I}}_h \cup \mathcal{F}_h(y_h) \tag{3.12d}$$

is said to be the discrete inactive set.

An edge $E \in \mathcal{E}_h(\bar{\Omega})$ is called active (purely inactive), if $\mathcal{N}_h(E) \subset \mathcal{A}_h$ ($\mathcal{N}_h(E) \subset \mathcal{I}_h$). The sets of active and purely inactive edges will be denoted by $\mathcal{E}_{\mathcal{A}_h}$ and $\overset{\circ}{\mathcal{E}}_{\mathcal{I}_h}$. We set $\mathcal{E}_{\mathcal{F}_h(y_h)} := \mathcal{E}_h(\bar{\Omega}) \setminus (\mathcal{E}_{\mathcal{A}_h} \cup \overset{\circ}{\mathcal{E}}_{\mathcal{I}_h})$ and $\mathcal{E}_{\mathcal{I}_h} := \overset{\circ}{\mathcal{E}}_{\mathcal{I}_h} \cup \mathcal{E}_{\mathcal{F}_h(y_h)}$. An active edge $E \in \mathcal{E}_{\mathcal{A}_h}$ is called isolated, if $E \in \mathcal{E}_{\mathcal{A}_h} \setminus \mathcal{E}_h(\mathcal{A}_h)$.

Likewise, for $\lambda_h \in M_h$ we denote by

$$\mathcal{Z}_h(\lambda_h) := \{a \in \mathcal{N}_h(\Omega) \mid \lambda_h(a) = 0\}, \quad \mathcal{C}_h(\lambda_h) := \mathcal{N}_h(\Omega) \setminus \mathcal{Z}_h(\lambda_h) \tag{3.13}$$

the sets of zero and non-zero nodal points with respect to λ_h and we partition $\mathcal{T}_h(\Omega)$ as follows

$$\mathcal{T}_h(\Omega) = \mathcal{T}_h^z(\lambda_h) \cup \mathcal{T}_h^c(\lambda) \cup \mathcal{T}_h^m(\lambda_h), \tag{3.14}$$

where

$$\mathcal{T}_h^z(\lambda_h) := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subset \mathcal{Z}_h(\lambda_h) \cup \mathcal{N}_h(\Gamma)\}, \tag{3.15a}$$

$$\mathcal{T}_h^c(\lambda_h) := \{T \in \mathcal{T}_h(\Omega) \mid T \cap \Gamma = \emptyset \text{ and } \mathcal{N}_h(T) \subset \mathcal{C}_h(\lambda_h)\} \cup \tag{3.15b}$$

$$\{T \in \mathcal{T}_h(\Omega) \mid T \cap \Gamma \neq \emptyset \wedge \mathcal{N}_h(T) \cap \mathcal{N}_h(\Omega) \subset \mathcal{C}_h(\lambda_h) \wedge T \subset \mathcal{A}_h\},$$

$$\mathcal{T}_h^m(\lambda_h) := \mathcal{T}_h(\Omega) \setminus (\mathcal{T}_h^z(\lambda_h) \cup \mathcal{T}_h^c(\lambda_h)). \tag{3.15c}$$

Definition 3.3. For $\sigma_h \in M_h \cap \mathcal{M}_+(\bar{\Omega})$ the sets $Z_h := \mathcal{Z}_h(\sigma_h)$ and $C_h := \mathcal{C}_h(\sigma_h)$ are said to be the sets of zero and strongly active nodal points. Isolated zero (strongly active) nodal points are defined analogously to Definition 3.2.

An edge $E \in \mathcal{E}_h(\bar{\Omega})$ is said to be strongly active (purely zero), if $\mathcal{N}_h(E) \subseteq C_h$ ($\mathcal{N}_h(E) \subseteq Z_h$). The sets of strongly active and purely zero edges are denoted by $\mathcal{E}_{C_h}$ and $\overset{\circ}{\mathcal{E}}_{Z_h}$. We set $\mathcal{E}_{\mathcal{F}_h(\sigma_h)} := \mathcal{E}_h(\bar{\Omega}) \setminus (\mathcal{E}_{C_h} \cup \overset{\circ}{\mathcal{E}}_{Z_h})$ and $\mathcal{E}_{Z_h} := \overset{\circ}{\mathcal{E}}_{Z_h} \cup \mathcal{E}_{\mathcal{F}_h(\sigma_h)}$.

Moreover, the sets

$$\overset{\circ}{\mathcal{Z}}_h := \bigcup \{T \in \mathcal{T}_h^z(\sigma_h)\}, \tag{3.16a}$$

$$\mathcal{C}_h := \bigcup \{T \in \mathcal{T}_h^c(\sigma_h)\}, \tag{3.16b}$$

$$\mathcal{F}_h(\sigma_h) := \bigcup \{T \in \mathcal{T}_h^m(\sigma_h)\} \tag{3.16c}$$

are referred to as the discrete purely zero set, the discrete strongly active set, and the discrete free boundary with respect to σ_h. The set

$$\mathcal{Z}_h := \overset{\circ}{\mathcal{Z}}_h \cup \mathcal{F}_h(\sigma_h) \tag{3.16d}$$

is said to be the discrete zero set and the set

$$\mathcal{B}_h := \mathrm{cl}(\mathcal{A}_h \setminus \mathcal{C}_h) \tag{3.16e}$$

is called the discrete biactive set. If $\mathcal{B}_h = \emptyset$, we say that discrete strict complementarity holds true. Otherwise, there is a lack of discrete strict complementarity.

Zero (strongly active) edges and isolated zero (isolated strongly active) edges are defined similarly to Definition 3.2.

3.3 Discrete Stationarity Concepts

The discrete (strongly) active sets $\mathcal{A}_h, \mathcal{C}_h$, the discrete biactive set $\mathcal{B}_h$ and the discrete inactive set $\mathcal{I}_h$ will be used to classify stationary points in the discrete regime.

Definition 3.4. For $(y_h, \sigma_h, u_h) \in V_h \times M_h \times S_h^{(1)}$ assume that there exist $(p_h, \mu_h) \in V_h \times M_h$ such that it holds

$$a(y_h, v_h) = (f + u_h, v_h)_{0,\Omega} - \langle\langle \sigma_h, v_h \rangle\rangle, \ v_h \in V_h, \tag{3.17a}$$

$$\psi_h - y_h \geq 0, \ \sigma_h \in M_h \cap \mathcal{M}_+(\bar{\Omega}), \ \langle\langle \sigma_h, \psi_h - y_h \rangle\rangle = 0, \tag{3.17b}$$

$$a(p_h, v_h) = (y^d - y_h, v_h)_{0,\Omega} - \langle\langle \mu_h, v_h \rangle\rangle, \ v_h \in V_h, \tag{3.17c}$$

$$p_h = \alpha \, (u_h - u_h^d), \tag{3.17d}$$

$$p_h(a) = 0, \ a \in \mathcal{C}_h, \tag{3.17e}$$

$$\mu_h(a) = 0, \ a \in \mathcal{I}_h. \tag{3.17f}$$

The triple $(y_h, \sigma_h, u_h) \in V_h \times M_h \times S_h^{(1)}$ is called

(i) a discrete C-stationary point of (3.6), if the pair $(p_h, \mu_h) \in V_h \times M_h$ satisfies

$$\mu_h(a) \, p_h(a) \geq 0, \ a \in \mathcal{B}_h, \tag{3.17g}$$

(ii) a discrete S-stationary point of (3.6), if the pair $(p_h, \mu_h) \in V_h \times M_h$ fulfills

$$\mu_h(a) \geq 0, \ p_h(a) \geq 0, \ a \in \mathcal{B}_h, \tag{3.17h}$$

(iii) a discrete stationary point of (3.6), if $\mathcal{B}_h = \emptyset$, i.e.,

$$\mathcal{C}_h = \mathcal{A}_h. \tag{3.17i}$$

Remark 3.5. In view of (3.17e) and (3.17f), condition (3.17g) implies

$$\langle\langle \mu_h, p_h \rangle\rangle \geq 0. \tag{3.18}$$

However, the reverse does not hold true. If $\langle\langle \mu_h, p_h \rangle\rangle = \sum_{a \in \mathcal{N}_h(\mathcal{B}_h)} \mu_h(a) p_h(a) \geq 0$, this does not imply that every summand is nonnegative. In other words, condition (3.18) is weaker than (3.17g).

3.4 Extensions of the Discrete Lagrange Multipliers

In this subsection, we will first derive an explicit representation of the operation of the discrete Lagrange multipliers σ_h and μ_h on functions $v_h \in V_h$ and then provide two extensions $\hat{\sigma}_h, \hat{\mu}_h$ and $\tilde{\sigma}_h, \tilde{\mu}_h$ to functionals on V. The extensions $\hat{\sigma}_h, \hat{\mu}_h$ will be used in the convergence analysis of the finite element approximations in Sect. 4, whereas the extensions $\tilde{\sigma}_h, \tilde{\mu}_h$ will play an essential role in the a posteriori error analysis in Sect. 5.

For notational convenience, we introduce the operator $I_{D_h} : V_h \to V_h, D_h \subset \mathcal{N}_h(\Omega)$, defined by means of

$$I_{D_h}(v_h)(a) := \begin{cases} v_h(a), \, a \in D_h \\ 0, \, a \in \mathcal{N}_h(\Omega) \setminus D_h \end{cases}, \quad v_h \in V_h. \tag{3.19}$$

It follows that $I_{\mathcal{C}_h}$ is the identity on $\mathcal{C}_h$, vanishes on $\overset{\circ}{\mathcal{Z}}_h$, whereas for $D = T \in \mathcal{T}_h(\mathcal{F}_h(\sigma_h))$ and $D = E \in \mathcal{E}_{\mathcal{F}_h(\sigma_h)}$:

$$I_{\mathcal{C}_h}(v_h)|_D = \sum_{a \in \mathcal{N}_h(D) \cap \mathcal{C}_h} v_h(a) \, \varphi_h^{(a)}.$$

Likewise, $I_{\mathcal{A}_h}$ is the identity on $\mathcal{A}_h$, vanishes on $\overset{\circ}{\mathcal{I}}_h$, whereas for $D = T \in \mathcal{T}_h(\mathcal{F}_h(y_h))$ and $D = E \in \mathcal{E}_{\mathcal{F}_h(y_h)}$:

$$I_{\mathcal{A}_h}(v_h)|_D = \sum_{a \in \mathcal{N}_h(D) \cap \mathcal{A}_h} v_h(a) \, \varphi_h^{(a)}.$$

Proposition 3.6. *Let* σ_h, μ_h *be the discrete Lagrange multipliers from* Defini-tion *3.4, let* $\mathcal{F}_h(y_h), \mathcal{F}_h(\sigma_h)$ *be the discrete free boundaries with respect to* y_h *and* σ_h *according to* (3.12d) *and* (3.16c)*, and let* I_{D_h} *be given by* (3.19)*. Then, for* $v_h \in V_h$ *it holds*

$$\langle\langle \sigma_h, v_h \rangle\rangle = \sum_{T \in \mathcal{T}_h(\mathcal{C}_h \cup \mathcal{F}_h(\sigma_h))} \left((f + u_h, I_{C_h}(v_h))_{0,T} - (\nabla y_h, \nabla I_{C_h}(v_h))_{0,T} \right) =$$
(3.20a)

$$\sum_{T \in \mathcal{T}_h(\mathcal{C}_h \cup \mathcal{F}_h(\sigma_h))} (f + u_h, I_{C_h}(v_h))_{0,T} - \sum_{E \in \mathcal{E}_{\mathcal{C}_h} \cup \mathcal{E}_{\mathcal{F}_h(\sigma_h)}} (\nu_E \cdot [\nabla y_h]_E, I_{C_h}(v_h))_{0,E},$$

$$\langle\langle \mu_h, v_h \rangle\rangle = \sum_{T \in \mathcal{T}_h(\mathcal{A}_h \cup \mathcal{F}_h(y_h))} \left((y^d - y_h, I_{A_h}(v_h))_{0,T} - (\nabla p_h, \nabla I_{A_h}(v_h))_{0,T} \right) =$$
(3.20b)

$$\sum_{T \in \mathcal{T}_h(\mathcal{A}_h \cup \mathcal{F}_h(y_h))} (y^d - y_h, I_{A_h}(v_h))_{0,T} - \sum_{E \in \mathcal{E}_{\mathcal{A}_h} \cup \mathcal{E}_{\mathcal{F}_h(y_h)}} (\nu_E \cdot [\nabla p_h]_E, I_{A_h}(v_h))_{0,E}.$$

Proof. In view of (3.17a) and (3.17c) we have

$$\langle\langle \sigma_h, \varphi_h^{(a)} \rangle\rangle = (f + u_h, \varphi_h^{(a)})_{0,\omega_a} - (\nabla y_h, \nabla \varphi_h^{(a)})_{0,\omega_a}, \quad a \in \mathcal{C}_h,$$

$$\langle\langle \mu_h, \varphi_h^{(a)} \rangle\rangle = (y^d - y_h, \varphi_h^{(a)})_{0,\omega_a} - (\nabla p_h, \nabla \varphi_h^{(a)})_{0,\omega_a}, \quad a \in \mathcal{A}_h.$$

Due to (3.16d) and (3.17f) $\sigma_h(a) = 0, a \in \mathcal{Z}_h$, and $\mu_h(a) = 0, a \in \mathcal{I}_h$, whence

$$\sigma_h(a) = \begin{cases} \sum\limits_{T \in \mathcal{T}_h(\omega_a)} \left((f + u_h, \varphi_h^{(a)})_{0,T} - (\nabla y_h, \nabla \varphi_h^{(a)})_{0,T} \right), & a \in \mathcal{C}_h \\ 0, & a \in \mathcal{Z}_h \end{cases}, \qquad (3.21)$$

and

$$\mu_h(a) = \begin{cases} \sum\limits_{T \in \mathcal{T}_h(\omega_a)} \left((y^d - y_h, \varphi_h^{(a)})_{0,T} - (\nabla p_h, \nabla \varphi_h^{(a)})_{0,T} \right), & a \in \mathcal{A}_h \\ 0, & a \in \mathcal{I}_h \end{cases}. \qquad (3.22)$$

Applying Green's formula elementwise to the second terms on the right-hand side in (3.21) and (3.22) yields

$$\sigma_h(a) = \begin{cases} \sum\limits_{T \in \mathcal{T}_h(\omega_a)} (f + u_h, \varphi_h^{(a)})_{0,T} - \sum\limits_{E \in \mathcal{E}_h^a} (\nu_E \cdot [\nabla y_h]_E, \varphi_h^{(a)})_{0,E}, & a \in \mathcal{C}_h \\ 0, & a \in \mathcal{Z}_h \end{cases},$$
(3.23)

and

$$\mu_h(a) = \begin{cases} \sum_{T \in \mathcal{T}_h(\omega_a)} (y^d - y_h, \varphi_h^{(a)})_{0,T} - \sum_{E \in \mathcal{E}_h^a} (\nu_E \cdot [\nabla p_h]_E, \varphi_h^{(a)})_{0,E} \ , a \in \mathcal{A}_h \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0 \ , a \in \mathcal{I}_h \end{cases}.$$

(3.24)

Taking $\langle\langle \sigma_h, v_h \rangle\rangle = \sum_{a \in \mathcal{N}_h(\mathcal{C}_h)} \sigma_h(a) v_h(a)$ into account, from (3.21) and (3.23) we deduce

$$\langle\langle \sigma_h, v_h \rangle\rangle = \sum_{a \in \mathcal{N}_h(\mathcal{C}_h)} \Big(\sum_{T \in \mathcal{T}_h(\omega_a)} \big((f + u_h, v_h(a)\, \varphi_h^{(a)})_{0,T} - (\nabla y_h, v_h(a)\nabla \varphi_h^{(a)})_{0,T} \big) \Big)$$

and

$$\langle\langle \sigma_h, v_h \rangle\rangle =$$
$$\sum_{a \in \mathcal{N}_h(\mathcal{C}_h)} \Big(\sum_{T \in \mathcal{T}_h(\omega_a)} (f + u_h, v_h(a)\varphi_h^{(a)})_{0,T} - \sum_{E \in \mathcal{E}_h(\mathcal{E}_h^a)} (\nu_E \cdot [\nabla y_h]_E, v_h(a)\varphi_h^{(a)})_{0,E} \Big).$$

Regrouping the summands in the above expressions gives (3.20a). The representation (3.20b) follows similarly. □

The first extensions $\hat{\sigma}_h, \hat{\mu}_h \in V^*$ of the discrete multipliers are defined in a similar way to the finite element analysis of variational inequalities of obstacle type (cf., e.g., [6]), whereas the second extensions $\tilde{\sigma}_h, \tilde{\mu}_h \in V^*$ are defined in view of Proposition 3.6.

Definition 3.7. Let $(y_h, \sigma_h, u_h, p_h, \mu_h) \in V_h \times M_h \times S_h^{(1)} \times V_h \times M_h$ satisfy (3.17a)–(3.17f). We define functionals $\hat{\sigma}_h, \hat{\mu}_h \in V^*$ by means of

$$\langle \hat{\sigma}_h, v \rangle := (f + u_h, v)_{0,\Omega} - a(y_h, v), \quad v \in V, \tag{3.25a}$$

$$\langle \hat{\mu}_h, v \rangle := (y^d - y_h, v)_{0,\Omega} - a(p_h, v), \quad v \in V, \tag{3.25b}$$

and functionals $\tilde{\sigma}_h, \tilde{\mu}_h \in V^*$ according to

$$\langle \tilde{\sigma}_h, v \rangle := \tag{3.26a}$$
$$\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} (f + u_h, v)_{0,T} - \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} (\nu_E \cdot [\nabla y_h]_E, v)_{0,E} + F_h^{(\sigma)}(P_h^{SZ}v), \quad v \in V,$$

$$\langle \tilde{\mu}_h, v \rangle := \tag{3.26b}$$
$$\sum_{T \in \mathcal{T}_h(\mathcal{A}_h)} (y^d - y_h, v)_{0,T} - \sum_{E \in \mathcal{E}_{\mathcal{A}_h}} (\nu_E \cdot [\nabla p_h]_E, v)_{0,E} + F_h^{(\mu)}(P_h^{SZ}v), \quad v \in V,$$

where P_h^{SZ} stands for the Scott–Zhang interpolation operator (see, e.g., [9,39]) and

$$F_h^{(\sigma)}(v_h) := \tag{3.26c}$$

$$\sum_{T \in \mathcal{T}_h(\mathcal{F}_h(\sigma_h))} (f + u_h, I_{C_h}(v_h))_{0,T} - \sum_{E \in \mathcal{E}_{\mathcal{F}_h(\sigma_h)}} (v_E \cdot [\nabla y_h]_E, I_{C_h}(v_h))_{0,E}, \tag{3.26d}$$

$$F_h^{(\mu)}(v_h) := \tag{3.26e}$$

$$\sum_{T \in \mathcal{T}_h(\mathcal{F}_h(y_h))} (y^d - y_h, I_{A_h}(v_h))_{0,T} - \sum_{E \in \mathcal{E}_{\mathcal{F}_h(y_h)}} (v_E \cdot [\nabla p_h]_E, I_{A_h}(v_h))_{0,E}. \tag{3.26f}$$

Remark 3.8. For later use in Sect. 5, we recall the definition of the Scott–Zhang interpolation operator: For each $a \in \mathcal{T}_h(\Omega)$ let $T \in \omega_a$ be an arbitrarily but fixed chosen element. Further, let $\{\Phi_T^{(a)} \mid a \in \mathcal{N}_h(T)\}$ be the $L^2(T)$-dual basis of $\{\varphi_h^{(a)} \mid a \in \mathcal{N}_h(T)\}$. Then, $P_h^{SZ} : L^2(\Omega) \to V_h$ is defined by means of

$$P_h^{SZ} v := \sum_{a \in \mathcal{N}_h(\Omega)} (P_h^{SZ} v)(a) \varphi_h^{(a)}, \tag{3.27}$$

where the nodal coefficients $(P_h^{SZ} v)(a)$ are given by

$$(P_h^{SZ} v)(a) := \int_T \Phi_T^{(a)}(x) v(x) \, dx. \tag{3.28}$$

Proposition 3.9. *The functionals $\hat{\sigma}_h, \hat{\mu}_h \in V^*$ and $\tilde{\sigma}_h, \tilde{\mu}_h \in V^*$ are extensions of $\sigma_h, \mu_h \in M_h$, i.e., for $v_h \in V_h$ it holds*

$$\langle \hat{\sigma}_h, v_h \rangle = \langle \tilde{\sigma}_h, v_h \rangle = \langle\langle \sigma_h, v_h \rangle\rangle,$$

$$\langle \hat{\mu}_h, v_h \rangle = \langle \tilde{\mu}_h, v_h \rangle = \langle\langle \mu_h, v_h \rangle\rangle.$$

Proof. The results are immediate consequences of (3.17) and Proposition 3.6. □

Remark 3.10. Fine properties of the extensions $\hat{\sigma}_h, \hat{\mu}_h \in V^*$ in terms of localizations involving the discrete active/inactive sets are difficult to obtain, whereas the extensions $\tilde{\sigma}_h, \tilde{\mu}_h \in V^*$ obviously satisfy

$$\mathcal{C}_h \subseteq \text{supp}(\tilde{\sigma}_h) \subseteq \mathcal{C}_h \cup \mathcal{F}_h(\sigma_h), \tag{3.29a}$$

$$\text{supp}(\tilde{\mu}_h) \subseteq \mathcal{A}_h \cup \mathcal{F}_h(y_h). \tag{3.29b}$$

The precise structure of $\tilde{\sigma}_h \in V^*$ depends on the definition of the Scott–Zhang interpolation operator P_h^{SZ}. In particular, under the condition

$$\text{For all } a \in \mathcal{C}_h \text{ there exists } T^{(a)} \subset \omega_a \text{ such that } T^{(a)} \subset \mathcal{C}_h, \tag{3.30}$$

we obtain supp($\tilde{\sigma}_h$) = $\mathcal{C}_h$, if the triangles satisfying (3.30) are used in the definition of P_h^{SZ}. We note that (3.30) excludes isolated strongly active nodal points and edges. However, utilizing a Scott–Zhang interpolation operator defined by averaging over edges instead of triangles (see [39]), allows to show supp($\tilde{\sigma}_h$) = $\mathcal{C}_h$, if we only exclude isolated strongly active nodal points. Similar remarks apply to $\tilde{\mu}_h$, i.e., it is possible to achieve supp($\tilde{\mu}_h$) $\subseteq$ $\mathcal{A}_h$ instead of (3.29b), if no isolated active nodal points occur and the modified P_h^{SZ} is used.

4   Convergence Analysis of the Finite Element Approximation

In this section, we prove that for a sequence of discrete C-stationary points there exists a subsequence converging to an almost C-stationary point. To this end, we assume:

(**A₁**) $\{(y_h, u_h, \sigma_h)\}_{\mathcal{H}}$ is a sequence of global minima of (3.7) or the sequences $\{y_h\}_{\mathcal{H}}$ and $\{u_h\}_{\mathcal{H}}$ are uniformly bounded in $L^2(\Omega)$.
(**A₂**) The obstacle ψ satisfies $\Delta\psi \in L^2(\Omega)$.

Remark 4.1. Under assumption (**A₂**) we may restrict ourselves to the case $\psi = 0$, since otherwise we can replace f by $f + \Delta\psi$ and y^d by $y^d - \psi$.

Theorem 4.2. *Let $\{(y_h, \sigma_h, u_h)\}_{\mathcal{H}}, (y_h, \sigma_h, u_h) \in V_h \times M_h \times S_h^{(1)}, h \in \mathcal{H}$,   be   a sequence of discrete C-stationary points of* (3.6). *Further, let* $\{(p_h, \mu_h)\}_{\mathcal{H}}, (p_h, \mu_h) \in V_h \times M_h, h \in \mathcal{H}$, *be the sequence of associated discrete adjoint states and multipliers computed with respect to a sequence* $\{V_h\}_{\mathcal{H}}$ *of nested finite element spaces. Finally, let* $\hat{\sigma}_h \in V^*$ *and* $\hat{\mu}_h \in V^*$ *be the extensions of the multipliers* σ_h *and* μ_h *as given by* (3.25).

If the assumptions (**A₁**) *and* (**A₂**) *are satisfied and the sequence* $\{V_h\}_{\mathcal{H}}$ *is limit dense in* V, *then there exist a subsequence* $\mathcal{H}' \subset \mathcal{H}$ *and an almost C-stationary point* $(y^*, \sigma^*, u^*) \in V \times V^* \times L^2(\Omega)$ *of* (2.5) *with associated adjoint state* $p^* \in V$ *and multiplier* $\mu^* \in V^*$ *such that for* $h \in \mathcal{H}', h \to 0$ *it holds*

$$y_h \to y^* \quad \text{in } V, \tag{4.1a}$$

$$y_h \to y^* \quad \text{in } L^2(\Omega), \tag{4.1b}$$

$$\hat{\sigma}_h \to \sigma^* \quad \text{in } V^*, \tag{4.1c}$$

$$u_h \to u^* \quad \text{in } L^2(\Omega), \tag{4.1d}$$

$$p_h \rightharpoonup p^* \quad \text{in } V, \tag{4.1e}$$

$$p_h \to p^* \quad \text{in } L^2(\Omega), \tag{4.1f}$$

$$\hat{\mu}_h \rightharpoonup^* \mu^* \quad \text{in } V^*. \tag{4.1g}$$

Moreover, if $\{S_h^{(1)}\}_{\mathcal{H}}$ is limit dense in $H^1(\Omega)$, we have

$$\langle \mu^*, y^* v \rangle = 0 \quad \text{for all } v \in C^1(\bar{\Omega}). \tag{4.1h}$$

Proof. Assume that $\{(y_h, \sigma_h, u_h)\}_{\mathcal{H}}$ is a sequence of global minima. The triple $(y_h, \sigma_h, u_h) = (0, -f_h, 0)$ is a feasible point for (3.6) and hence, $J_h(y_h, u_h) \leq J_h(0, -f_h)$. By the inverse triangle inequality and Young's inequality it follows that the sequences $\{y_h\}_{\mathcal{H}}$ and $\{u_h\}_{\mathcal{H}}$ are bounded in $L^2(\Omega)$.

If $\{(y_h, \sigma_h, u_h)\}_{\mathcal{H}}$ is a sequence of stationary points, the boundedness of $\{y_h\}_{\mathcal{H}}$ and $\{u_h\}_{\mathcal{H}}$ in $L^2(\Omega)$ follows from assumption (**A₁**).

Choosing $v_h = y_h$ in (3.17a) and $v_h = p_h$ in (3.17c) and taking (2.4b),(3.17b), and (3.18) into account, we obtain

$$\gamma \, \|y_h\|_{1,\Omega}^2 \leq a(y_h, y_h) = (f + u_h, y_h)_{0,\Omega} \leq \left(\|f\|_{0,\Omega} + \|u_h\|_{0,\Omega} \right) \|y_h\|_{1,\Omega},$$

$$\gamma \, \|p_h\|_{1,\Omega}^2 \leq a(p_h, p_h) = (y^d - y_h, p_h)_{0,\Omega} - \langle\langle \mu_h, p_h \rangle\rangle$$

$$\leq (y^d - y_h, p_h)_{0,\Omega} \leq \left(\|y^d\|_{0,\Omega} + \|y_h\|_{0,\Omega} \right) \|p_h\|_{1,\Omega}.$$

In view of the boundedness of $\{y_h\}_{\mathcal{H}}$ and $\{u_h\}_{\mathcal{H}}$ in $L^2(\Omega)$, the preceding two inequalities imply the boundedness of $\{y_h\}_{\mathcal{H}}$ and $\{p_h\}_{\mathcal{H}}$ in V. Moreover, observing (2.4b), for $v \in V$ we have

$$|\langle \hat{\sigma}_h, v \rangle| \leq \|f + u_h\|_{0,\Omega} \|v\|_{0,\Omega} + C \, \|y_h\|_{1,\Omega} \|v\|_{1,\Omega}$$

$$\leq (\|f + u_h\|_{0,\Omega} + C \, \|y_h\|_{1,\Omega}) \|v\|_{1,\Omega},$$

$$|\langle \hat{\mu}_h, v \rangle| \leq \|y^d - y_h\|_{0,\Omega} \|v\|_{0,\Omega} + C \, \|p_h\|_{1,\Omega} \|v\|_{1,\Omega}$$

$$\leq (\|y^d - y_h\|_{0,\Omega} + C \, \|p_h\|_{1,\Omega}) \|v\|_{1,\Omega},$$

whence

$$\|\hat{\sigma}_h\|_{V^*} \leq \|f + u_h\|_{0,\Omega} + C \, \|y_h\|_{1,\Omega}), \quad \|\hat{\mu}_h\|_{V^*} \leq \|y^d - y_h\|_{0,\Omega} + C \, \|y_h\|_{1,\Omega}).$$

This implies boundedness of the sequences $\{\hat{\sigma}_h\}_{\mathcal{H}}$ and $\{\hat{\mu}_h\}_{\mathcal{H}}$ in V^*. Consequently, there exist a subsequence $\mathcal{H}' \subset \mathcal{H}$ and a point $(y^*, \sigma^*, u^*, p^*, \mu^*) \in V \times V^* \times L^2(\Omega) \times V \times V^*$ such that for $h \in \mathcal{H}', h \to 0$ it holds

$$y_h \rightharpoonup y^* \quad \text{in } V, \quad p_h \rightharpoonup p^* \quad \text{in } V, \tag{4.2a}$$

$$u_h \rightharpoonup u^* \quad \text{in } L^2(\Omega), \tag{4.2b}$$

$$\hat{\sigma}_h \rightharpoonup^* \sigma^* \quad \text{in } V^*, \quad \hat{\mu}_h \rightharpoonup^* \mu^* \quad \text{in } V^*. \tag{4.2c}$$

Due to the Rellich–Kondrachov theorem V is compactly embedded in $L^2(\Omega)$ and hence, (4.2a) implies (4.1b),(4.1f).

For another subsequence, still denoted by $\mathcal{H}'$, we further deduce that for $h \in \mathcal{H}', h \to 0$ we have $y_h \to y^*$ and $p_h \to p^*$ pointwise almost everywhere. Hence, $y_h \leq 0, h \in \mathcal{H}'$, implies $y^* \leq 0$ almost everywhere (a.e.) in Ω.

Next, we show that the point $(y^*, \sigma^*, u^*, p^*, \mu^*)$ satisfies the state equation (2.13a), the adjoint state equation (2.13c), and (2.13d). Since $\{V_h\}_{\mathcal{H}}$ is limit dense in V, for any $v \in V$ we find a sequence $\{v_h\}_{\mathcal{H}}, v_h \in V_h, h \in \mathcal{H}$, such that $v_h \to v$ for $h \to 0$. Observing (4.2), for $h \in \mathcal{H}', h \to 0$, we deduce

$$a(y_h, v_h) \to a(y^*, v), \quad a(p_h, v_h) \to a(p^*, v),$$

$$(f + u_h, v_h)_{0,\Omega} \to (f + u^*, v)_{0,\Omega}, \quad (y^d - y_h, v_h)_{0,\Omega} \to (y^d - y^*, v)_{0,\Omega},$$

$$\langle\langle \sigma_h, v_h \rangle\rangle = \langle\langle \hat{\sigma}_h, v_h \rangle\rangle \to \langle \sigma^*, v \rangle, \quad \langle\langle \mu_h, v_h \rangle\rangle = \langle\langle \hat{\mu}_h, v_h \rangle\rangle \to \langle \mu^*, v \rangle.$$

Hence, passing to the limit in (3.17a) and (3.17c), we find that $(y^*, \sigma^*, u^*, p^*, \mu^*)$ satisfies (2.13a) and (2.13c).

The limit density of $\{V_h\}_{\mathcal{H}}$ in V further implies $u_h^d \to u^d, h \to 0$. Consequently, (3.17d) and (4.2) imply that (4.1d) holds true and that the pair (p^*, u^*) fulfills (2.13d).

Next, we verify $\sigma^* \in V_+^*$. Since $\{(V_h)_+\}_{\mathcal{H}}$ is limit dense in V_+, for any $v \in V_+$ there exists a sequence $\{v_h\}_{\mathcal{H}}, v_h \in (V_h)_+, h \in \mathcal{H}$, such that $v_h \to v$ as $h \to 0$. Observing $\sigma_h \in \mathcal{M}_+(\bar{\Omega})$ and (4.2c), we find

$$0 \leq \langle\langle \sigma_h, v_h \rangle\rangle = \langle \hat{\sigma}_h, v_h \rangle \to \langle \sigma^*, v \rangle,$$

whence $\langle \sigma^*, v \rangle$ for any $v \in V_+$.

In order to establish strong convergence of the states in V, due to (3.6b) we have

$$a(y_h, y_h) \leq a(y_h, v_h) + (f + u_h, y_h - v_h)_{0,\Omega}, \quad v_h \in V_h \cap V_-. \tag{4.3}$$

Since the sequence $\{V_h \cap V_-\}_{\mathcal{H}}$ is limit dense in V_-, there exists a sequence $\{v_h\}_{\mathcal{H}}, v_h \in V_h \cap V_-, h \in \mathcal{H}$, such that $v_h \to y^* \in V_-$ as $h \to 0$. Taking (2.4b) and (4.3) into account, it holds

$$\gamma \|y_h - y^*\|_{1,\Omega}^2 \leq a(y_h - y^*, y_h - y^*) = a(y_h, y_h) - a(y_h, y^*) - a(y^*, y_h - y^*)$$

$$\leq a(y_h, v_h) + (f + u_h, v_h)_{0,\Omega} - a(y_h, y^*) - a(y^*, y_h - y^*).$$

Due to the already proven assertions (4.1b),(4.1d) and in view of (4.2a) the right-hand side in the preceding inequality converges to zero which implies (4.1a). Moreover, observing (3.17b),(3.17f), and (4.1a), it follows that

$$0 = \langle \hat{\sigma}_h, y_h \rangle \to \langle \sigma^*, y^* \rangle, \quad 0 = \langle \hat{\mu}_h, y_h \rangle \to \langle \mu^*, y^* \rangle,$$

whence $\langle \sigma^*, y^* \rangle = \langle \mu^*, y^* \rangle = 0$.

For the proof of (4.1c), we note that the compact embedding of $L^2(\Omega)$ in V^* implies $u_h \to u^*$ in V^* as $\mathcal{H}' \ni h \to 0$. Since $A \in \mathcal{L}(V, V^*)$ is bounded, we obtain

$$\|\hat{\sigma}_h - \sigma^*\|_{V^*} \leq \|Ay_h - Ay^*\|_{V^*} + \|u_h - u^*\|_{V^*}$$
$$\leq \|A\|_{\mathcal{L}(V,V^*)} \|y_h - y^*\|_V + \|u_h - u^*\|_{V^*} \to 0 \; (h \to 0),$$

which implies (4.1c). Moreover, due to (3.17e),(4.1c), and (4.1e) we have

$$0 = \langle \hat{\sigma}_h, p_h \rangle \to \langle \sigma^*, p^* \rangle \; (\mathcal{H}' \ni h \to 0),$$

whence $\langle \sigma^*, p^* \rangle = 0$.

Next, we show $\langle \mu^*, p^* \rangle \geq 0$. To this end, setting $v_h = p_h$ in (3.17c) and observing $\langle\langle \mu_h, p_h \rangle\rangle \geq 0$, we find

$$0 \geq a(p_h, p_h) - (y^d - y_h, p_h)_{0,\Omega}. \tag{4.4}$$

Since the functional $v \in V \mapsto a(v, v)$ is lower semicontinuous and convex, it is weakly lower semicontinuous whence due to (4.2a)

$$a(p^*, p^*) \leq \liminf a(p_h, p_h).$$

On the other hand, the already proven assertions (4.1b),(4.1f) imply

$$(y^d - y_h, p_h)_{0,\Omega} \to (y^d - y^*, p^*)_{0,\Omega} \; (\mathcal{H}' \ni h \to 0).$$

Consequently, passing to the limit in (4.4) and taking into account that the triple (y^*, σ^*, u^*) satisfies (2.13c), we obtain

$$0 \geq a(p^*, p^*) - (y^d - y^*, p^*)_{0,\Omega} = -\langle \mu^*, p^* \rangle,$$

which proves $\langle \mu^*, p^* \rangle \geq 0$.

In order to verify that p^* satisfies (2.13e), we show

$$\langle \sigma^*, (p^*)^+ \rangle = \langle \sigma^*, (p^*)^- \rangle = 0, \tag{4.5}$$

which implies $p^* = 0$ in $\mathcal{C}^* = \text{int}(\text{supp}(\sigma^*))$ by Corollary 2.9. We note that (4.2a) gives rise to

$$(p_h)^+ \rightharpoonup (p^*)^+, \quad (p_h)^- \rightharpoonup (p^*)^- \text{ in } V \text{ as } \mathcal{H}' \ni h \to 0$$

(cf., e.g., [27]). Together with (3.17e), this leads to

$$0 = \langle\langle \sigma_h, (p_h)^+ \rangle\rangle \to \langle \sigma^*, (p^*)^+ \rangle, \quad 0 = \langle\langle \sigma_h, (p_h)^- \rangle\rangle \to \langle \sigma^*, (p^*)^- \rangle \; (\mathcal{H}' \ni h \to 0),$$

which proves (4.5).

It remains to show that (y^*, σ^*, u^*) is an almost C-stationary point and to prove (4.1h). In order to verify (4.1h), let $v \in C^1(\bar{\Omega})$. We have $y^*v \in V$ (cf., e.g., [13]). Since the sequence $\{S_h^{(1)}\}_{\mathcal{H}}$ is limit dense in $H^1(\Omega)$, there exists a sequence $\{v_h\}_{\mathcal{H}}, v_h \in S_h^{(1)}, h \in \mathcal{H}$, such that $v_h \to v \ (\mathcal{H} \ni h \to 0)$. Observing $v_h \in C(\bar{\Omega}), y_h \in C_0(\Omega)$, we have $v_h y_h \in C_0(\Omega), h \in \mathcal{H}$, which together with $(v_h y_h)|_T \in H^1(T), T \in \mathcal{T}_h(\Omega)$, implies $v_h y_h \in V, h \in \mathcal{H}$. Taking (4.1a) into account, we deduce $y_h v_h \to y^*v$ in V as $\mathcal{H}' \ni h \to 0$. Since $(y_h v_h)(a) = 0, a \in A_h$, it follows that

$$0 = \langle \hat{\mu}_h, y_h v_h \rangle \to \langle \mu^*, y^*v \rangle \ (\mathcal{H}' \ni h \to 0).$$

Hence, $\langle \mu^*, y^*v \rangle = 0$ which proves (4.1h), since $v \in C^1(\bar{\Omega})$ was chosen arbitrarily.

In order to prove (2.13i), we note that (3.17f) yields

$$\langle \hat{\mu}_h, v_h \rangle = 0, \quad v_h \in V_h \cap V_{\mathcal{I}_h \cup \mathcal{F}_h(y_h)}. \tag{4.6}$$

On the other hand, due to the pointwise a.e. convergence of $\{y_h\}_{\mathcal{H}'}$ to y^*, for sufficiently small $h_1 \in \mathcal{H}'$ we have

$$y_h < 0 \text{ a.e. in } \mathcal{I}^*, \quad \mathcal{H}' \ni h \leq h_1, \tag{4.7}$$

which shows $\mathcal{I}^* \subseteq \mathcal{I}_h$ for $h \leq h_1$. For $h \leq h_1$ we define

$$\tilde{\mathcal{I}}_h := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid \text{int}(T) \subseteq \mathcal{I}^*\},$$

such that $\tilde{\mathcal{I}}_h \subseteq \mathcal{I}^* \subseteq \mathcal{I}_h, \mathcal{H}' \ni h \leq h_1$. Since $\tilde{\mathcal{I}}_h$ may be empty, we choose $h_2 \in \mathcal{H}'$ sufficiently small so that $\tilde{\mathcal{I}}_h \neq \emptyset$ for $\mathcal{H}' \ni h \leq h_2$. Setting $h_3 := \min(h_1, h_2)$, we thus have

$$\emptyset \neq \tilde{\mathcal{I}}_h \subseteq \mathcal{I}^* \subseteq \mathcal{I}_h, \quad \mathcal{H}' \ni h \leq h_3. \tag{4.8}$$

Now, let $v \in C_{\mathcal{I}^*, 0} := \{v \in C_0(\Omega) \mid v|_{\mathcal{I}^*} \in C_0^\infty(\mathcal{I}^*), v|_{\Omega \setminus \mathcal{I}^*} = 0\}$ be chosen arbitrarily, but fixed. Since $\text{supp}(v) \subseteq \mathcal{I}^*$, there exists $h(v) \in \mathcal{H}', h(v) \leq h_3$, such that

$$\text{supp}(v) \subseteq \tilde{\mathcal{I}}_{h(v)} \subseteq \mathcal{I}^* \subseteq \mathcal{I}_h, \quad \mathcal{H}' \ni h \leq h(v).$$

Obviously, we have $v \in V_{\tilde{\mathcal{I}}_{h(v)}} \subseteq V_{\mathcal{I}^*, 0}$ and $\tilde{\mathcal{I}}_{h(v)} \subseteq \mathcal{I}_h, h \leq h(v)$, whence

$$V_h \cap V_{\tilde{\mathcal{I}}_{h(v)}} \subseteq V_h \cap V_{\mathcal{I}_h \cup \mathcal{F}_h(y_h)}, \quad h \leq h(v).$$

Observing (4.6), it follows that

$$\langle \hat{\mu}_h, v_h \rangle = 0, \quad v_h \in V_h \cap V_{\tilde{\mathcal{I}}_{h(v)}}, \ h \le h(v). \tag{4.9}$$

Since the sequence $\{V_h \cap V_{\tilde{\mathcal{I}}_{h(v)}}\}_{h \le h(v)} \subset V_{\tilde{\mathcal{I}}_{h(v)}}$ is limit dense in $V_{\tilde{\mathcal{I}}_{h(v)}}$, there exists a sequence $\{v_h\}_{h \le h(v)}, v_h \in V_h \cap V_{\tilde{\mathcal{I}}_{h(v)}}, h \le h(v)$, such that $v_h \to v$ as $h(v) \ge h \to 0$. In view of (4.2c) and (4.9), it follows that

$$0 = \langle \hat{\mu}_h, v_h \rangle \to \langle \mu^*, v \rangle \ (h(v) \ge h \to 0),$$

which gives $\langle \mu^*, v \rangle = 0, v \in C_{\mathcal{I}^*,0}$. The density of $C_{\mathcal{I}^*,0}$ in $V_{\mathcal{I}^*,0}$ implies (2.13i).

<div align="right">□</div>

5 A Posteriori Error Control

In this section, we want to derive a residual-type a posteriori error estimator for the discretization errors in the state, the adjoint state, and the control

$$e_{h,y} := y - y_h, \quad e_{h,p} := p - p_h, \quad e_{h,u} := u - u_h \tag{5.1}$$

that provides both an upper bound (reliability) and a lower bound (efficiency) up to consistency errors and data oscillations. The total discretization error $e_h := (e_{h,y}, e_{h,p}, e_{h,u})$ will be measured in the norm

$$|||e_h||| := \left(\|e_{h,y}\|_{1,\Omega}^2 + \|e_{h,p}\|_{1,\Omega}^2 + \|e_{h,u}\|_{0,\Omega}^2 \right)^{1/2}, \tag{5.2}$$

and we will show

$$\eta_h^2 - e_{h,eff}^c - \mathrm{osc}_{h,eff}^2 \lesssim |||e_h|||^2 \lesssim \eta_h^2 + e_{h,rel}^c + \mathrm{osc}_{h,rel}^2.$$

Here, η_h is the residual a posteriori error estimator, whereas $e_{h,rel}^c, e_{h,eff}^c$ and $\mathrm{osc}_{h,rel}, \mathrm{osc}_{h,eff}$ stand for the consistency errors and data oscillations associated with the reliability and efficiency estimates.

5.1 Components of the Reliability and Efficiency Estimates

In this subsection, we introduce the residual-type a posteriori error estimator consisting of element and edge residuals, discuss the consistency errors due to a mismatch in complementarity between the continuous and the discrete regime, and present the data oscillations.

5.1.1 Residual-Type a Posteriori Error Estimator

The residual-type a posteriori error estimator η_h is given by

$$\eta_h := \left((\eta_h^{(1)})^2 + (\eta_h^{(2)})^2 \right)^{1/2}, \tag{5.3}$$

where $\eta_h^{(1)}$ and $\eta_h^{(2)}$ consist of element residuals and edge residuals associated with the state equation (2.13a) and the adjoint state equation (2.13c)

$$\eta_h^{(1)} := \left(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} (\eta_T^{(1)})^2 + \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} (\eta_E^{(1)})^2 \right)^{1/2}, \tag{5.4a}$$

$$\eta_h^{(2)} := \left(\sum_{T \in \mathcal{T}_h(\mathcal{I}_h)} (\eta_T^{(2)})^2 + \sum_{E \in \mathcal{E}_{\mathcal{I}_h}} (\eta_E^{(2)})^2 \right)^{1/2}. \tag{5.4b}$$

In particular, the element residuals $\eta_T^{(\nu)}$ and the edge residuals $\eta_E^{(\nu)}$, $1 \leq \nu \leq 2$, are given by

$$\eta_T^{(1)} := h_T \, \|f + u_h\|_{0,T}, \quad , \eta_T^{(2)} := h_T \, \|y^d - y_h\|_{0,T}, \tag{5.5a}$$

$$\eta_E^{(1)} := h_E^{1/2} \, \|\nu_E \cdot [\nabla y_h]_E\|_{0,E}, \quad , \eta_E^{(2)} := h_E^{1/2} \, \|\nu_E \cdot [\nabla p_h]_E\|_{0,E}. \tag{5.5b}$$

5.1.2 Consistency Error (Mismatch in Complementarity)

We distinguish between reliability and efficiency related consistency errors.
Consistency Error for the Reliability Estimate.

$$e_{h,rel}^c := e_{h,\sigma}^{(1)} + e_{h,\sigma}^{(2)} + e_{h,\mu}^{(1)} + e_{h,\mu}^{(2)}, \tag{5.6}$$

where $e_{h,\sigma}^{(\nu)}, e_{h,\mu}^{(\nu)}, 1 \leq \nu \leq 2$, are given by

$$e_{h,\sigma}^{(1)} := \langle \tilde{\sigma}_h - \sigma, y - y_h \rangle, \quad e_{h,\sigma}^{(2)} := -\langle \tilde{\sigma}_h - \sigma, p - p_h \rangle, \tag{5.7a}$$

$$e_{h,\mu}^{(1)} := \langle \tilde{\mu}_h - \mu, y - y_h \rangle, \quad e_{h,\mu}^{(2)} := \langle \tilde{\mu}_h - \mu, p - p_h \rangle. \tag{5.7b}$$

Consistency Error for the Efficiency Estimate.

$$e_{h,eff}^c := \left(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} e_T^{(\sigma)} + \sum_{T \in \mathcal{T}_h(\mathcal{I}_h)} e_T^{(\mu)} + \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} e_{\omega_E}^{(\sigma)} + \sum_{E \in \mathcal{E}_{\mathcal{I}_h}} e_{\omega_E}^{(\mu)} \right), \tag{5.8}$$

where $e_T^{(\sigma)}, e_T^{(\mu)}$, and $e_{\omega_E}^{(\sigma)}, e_{\omega_E}^{(\mu)}$ are given by

$$e_T^{(\sigma)} := |(f_h + u_h)\, b_T|_{1,T}^{-1}\, \langle \sigma, (f_h + u_h)\, b_T \rangle, \qquad (5.9a)$$

$$e_T^{(\mu)} := |(y_h^d - y_h)\, b_T|_{1,T}^{-1}\, \langle \mu, (y_h^d - y_h)\, b_T \rangle, \qquad (5.9b)$$

$$e_{\omega_E}^{(\sigma)} := |\nu_E \cdot [\nabla y_h]_E\, b_E|_{1,\omega_E}^{-1}\, \langle \sigma, \nu_E \cdot [\nabla y_h]_E\, b_E \rangle, \qquad (5.9c)$$

$$e_{\omega_E}^{(\mu)} := -|\nu_E \cdot [\nabla p_h]_E\, b_E|_{1,T}^{-1}\, \langle \mu, \nu_E \cdot [\nabla p_h]_E\, b_E \rangle, \qquad (5.9d)$$

and b_T, b_E stand for the element and edge bubble functions.

5.1.3 Data Oscillations

As in case of the consistency errors, we distinguish between reliability and efficiency related data oscillations.

Data Oscillations for the Reliability Estimate.

$$\mathrm{osc}_{h,rel} := \Big(\sum_{T \in \mathcal{T}_h(\Omega)} \mathrm{osc}_T^2(u^d) \Big)^{1/2}, \qquad (5.10)$$

where $\mathrm{osc}_T(u^d)$ is given by

$$\mathrm{osc}_T(u^d) := \|u^d - u_h^d\|_{0,T}. \qquad (5.11)$$

Data Oscillations for the Efficiency Estimate.

$$\mathrm{osc}_{h,eff} := \Big(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} \mathrm{osc}_T^2(f) + \sum_{T \in \mathcal{T}_h(\mathcal{I}_h)} \mathrm{osc}_T^2(y^d) \Big)^{1/2}, \qquad (5.12)$$

where $\mathrm{osc}_T(f)$ and $\mathrm{osc}_T(y^d)$ are given by

$$\mathrm{osc}_T(f) := h_T\, \|f - f_h\|_{0,T}, \quad \mathrm{osc}_T(y^d) := h_T\, \|y^d - y_h^d\|_{0,T}. \qquad (5.13)$$

5.2 Reliability of the Error Estimator

Theorem 5.1. *Let (y, σ, u, p, μ) and $(y_h, \sigma_h, u_h, p_h, \mu_h)$ be solutions of* (2.13a)–(2.13g) *and* (3.17a)–(3.17f) *and let $\eta_h, e_{h,rel}^c, osc_{h,rel}$ be the residual-type error estimator, the consistency error, and the data oscillations as given by* (5.3),(5.6), *and* (5.10). *Then, it holds*

$$\||e_h|\|^2 \lesssim \eta_h^2 + e_{h,rel}^c + osc_{h,rel}^2. \tag{5.14}$$

The proof of Theorem 5.1 will be given by a series of lemmas.

We note that neither $e_{h,y}$ nor $e_{h,p}$ satisfy Galerkin orthogonality due to the presence of u, u_h in the right-hand sides of the continuous and discrete state equations (2.13a),(3.17a) and of y, y_h in the right-hand sides of the continuous and discrete adjoint state equations (2.13c),(3.17c). As in the case of the a posteriori error analysis of finite element approximations of control and/or state constrained distributed optimal control problems for second order elliptic PDEs, Galerkin orthogonality can be achieved with respect to an auxiliary state $y(u_h) \in V$ and an auxiliary adjoint state $p(y_h) \in V$ which are defined as the unique solutions of the variational equations

$$a(y(u_h), v) = (f + u_h, v)_{0,\Omega} - \langle \tilde{\sigma}_h, v \rangle, \quad v \in V, \tag{5.15a}$$

$$a(p(y_h), v) = (y^d - y_h, v)_{0,\Omega} - \langle \tilde{\mu}_h, v \rangle, \quad v \in V. \tag{5.15b}$$

In fact, it follows easily from (5.15a),(3.17a) and (5.15b),(3.17c) that

$$a(y(u_h) - y_h, v_h) = 0, \quad v_h \in V_h, \tag{5.16a}$$

$$a(p(y_h) - p_h, v_h) = 0, \quad v_h \in V_h. \tag{5.16b}$$

Lemma 5.2. *Under the assumptions of* Theorem 5.1 *let* $y(u_h), p(y_h)$ *be the auxiliary state and the auxiliary adjoint state as given by* (5.15a) *and* (5.15b) *and let* $\eta_h^{(1)}$ *and* $\eta_h^{(2)}$ *be the components of the residual a posteriori error estimator according to* (5.4a) *and* (5.4b)*. Then, it holds*

$$\|y(u_h) - y_h\|_{1,\Omega} \lesssim \eta_h^{(1)}, \tag{5.17a}$$

$$\|p(y_h) - p_h\|_{1,\Omega} \lesssim \eta_h^{(2)}. \tag{5.17b}$$

Proof. Denoting by P_h^C Clément's quasi-interpolation operator (cf., e.g., [44]), due to Proposition 3.9 and (5.16a) for $e := y(u_h) - y_h$ it holds

$$\|e\|_{1,\Omega}^2 \lesssim a(e, e) = r(e - P_h^C e), \tag{5.18}$$

where the residual $r(\cdot)$ is given by

$$r(v) := (f + u_h, v)_{0,\Omega} - \langle \tilde{\sigma}_h, v \rangle - a(y_h, v), \quad v \in V.$$

In view of the representation (3.26a) of the extension $\tilde{\sigma}_h$ of the discrete multiplier σ_h, by straightforward estimation we obtain

$$r(e - P_h^C e) \leq | \sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} (f + u_h, e - P_h^C e)_{0,T} | \tag{5.19}$$

$$+ | \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} (\nu_E \cdot [\nabla y_h]_E, e - P_h^C e)_{0,E} | + |F_h^{(\sigma)}(P_h^{SZ}(e - P_h^C e))|.$$

Taking advantage of the properties

$$\|e - P_h^C e\|_{0,T} \lesssim h_T \, |e|_{1,\omega_h^T}, \quad \|e - P_h^C e\|_{0,E} \lesssim h_T^{1/2} \, |e|_{1,\omega_h^E}$$

of Clément's quasi-interpolation operator, for the first two terms on the right-hand side of (5.19) it follows that

$$| \sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} (f + u_h, e - P_h^C e)_{0,T} | \; \leq \tag{5.20a}$$

$$\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} \|f + u_h\|_{0,T} \, \|e - P_H^C e\|_{0,T} \lesssim \sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} \eta_T^{(1)} \, |e|_{1,\omega_h^T},$$

$$| \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} (\nu_E \cdot [\nabla y_h]_E, e - P_h^C e)_{0,E} | \; \leq \tag{5.20b}$$

$$\sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} \|\nu_E \cdot [\nabla y_h]_E\|_{0,E} \, \|e - P_h^C e\|_{0,E} \lesssim \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} \eta_E^{(1)} \, |e|_{1,\omega_h^E}.$$

For the third term on the right-hand side in (5.19), in view of (3.26c) and the definition of the Scott–Zhang interpolation operator P_h^{SZ} we obtain

$$|F_h^{(\sigma)}(P_h^{SZ}(e - P_h^C e))| \; \leq \tag{5.21}$$

$$\sum_{T \in \mathcal{T}_h(\mathcal{F}_h(\sigma_h))} \left(\|f + u_h\|_{0,T} \sum_{a \in \mathcal{N}_h(T) \cap \mathcal{C}_h} \|(P_h^{SZ}(e - P_h^C e))(a)\varphi_h^{(a)}\|_{0,T} \right)$$

$$+ \sum_{E \in \mathcal{E}_{\mathcal{F}_h(\sigma_h)}} \|\nu_E \cdot [\nabla y_h]_E\|_{0,E} \, \|(P_h^{SZ}(e - P_h^C e))(a_E')\varphi_h^{(a_E')}\|_{0,E},$$

where a_E' stands for the single nodal point in $\mathcal{N}_h(E) \cap C_h, E \in \mathcal{E}_h(\mathcal{F}_h(\sigma_h))$. Using elementary properties of nodal basis functions

$$\|\varphi_h^{(a)}\|_{0,T} \lesssim h_T, \, a \in \mathcal{N}_h(T), \quad \|\varphi_h^{(a)}\|_{0,E} \lesssim h_E^{1/2}, \, a \in \mathcal{N}_h(E), \tag{5.22}$$

as well as the following property of P_h^{SZ} (see, e.g., [39])

$$|(P_h^{SZ}v)(a)| \lesssim h_T^{-1} \, \|v\|_{0,T}, \, a \in \mathcal{N}_h(T), \, v \in L^2(\Omega), \tag{5.23}$$

it follows that

$$
\sum_{a \in \mathcal{N}_h(T) \cap \mathcal{C}_h} |(P_h^{SZ}(e - P_h^C e))(a)| \, \|\varphi_h^{(a)}\|_{0,T} \lesssim \tag{5.24a}
$$

$$
h_T \sum_{a \in \mathcal{N}_h(T) \cap \mathcal{C}_h} |(P_h^{SZ}(e - P_h^C e))(a)| \lesssim
$$

$$
h_T \sum_{a \in \mathcal{N}_h(T) \cap \mathcal{C}_h} h_{T^{(a)}}^{-1} \|e - P_h^C e\|_{0,T_a} \lesssim h_T \sum_{a \in \mathcal{N}_h(T) \cap \mathcal{C}_h} |e|_{1,\omega_h^{T^{(a)}}},
$$

$$
\|(P_h^{SZ}(e - P_h^C e))(a'_E)\varphi_h^{(a'_E)}\|_{0,E} = \tag{5.24b}
$$

$$
|(P_h^{SZ}(e - P_h^C e))(a'_E)| \, \|\varphi_h^{(a'_E)}\|_{0,E} \lesssim
$$

$$
h_E^{1/2} h_{T^{(a'_E)}}^{-1} \|e - P_h^C e\|_{0,T^{(a'_E)}} \lesssim h_E^{1/2} |e|_{1,\omega_h^{T^{(a'_E)}}},
$$

where $T^{(a)}$ denotes the fixed element in ω_h^a which is used in the computation of the nodal coefficient $(P_h^{SZ}(e - P_h^C e))(a)$ (cf. (3.28)). Using (5.24a),(5.24b) in (5.21) yields

$$
|F_h^{(\sigma)}(P_h^{SZ}(e - P_h^C e))| \lesssim \tag{5.25}
$$

$$
\sum_{T \in \mathcal{T}_h(\mathcal{F}_h(\sigma_h))} \eta_T^{(1)} |e|_{1,\tilde{\omega}^T} + \sum_{E \in \mathcal{E}_{\mathcal{F}_h(\sigma_h)}} \eta_E^{(1)} |e|_{1,\omega_h^{T^{(a'_E)}}},
$$

where

$$
\tilde{\omega}^T := \bigcup_{a \in \mathcal{N}_h(T) \cap \mathcal{C}_h} \omega_h^{T^{(a)}}.
$$

Combining (5.20a),(5.20b), and (5.25), from (5.19) we deduce

$$
|r(e - P_h^C e)| \lesssim \sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} \eta_T^{(1)} |e|_{1,\hat{\omega}^T} + \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} \eta_E^{(1)} |e|_{1,\hat{\omega}^E}, \tag{5.26}
$$

where

$$
\hat{\omega}^T := \begin{cases} \tilde{\omega}^T \cup \omega_h^T \,, \, T \in \mathcal{T}_h(\mathcal{F}_h(\sigma_h)) \\ \omega_h^T \,, \text{ otherwise} \end{cases}, \quad \hat{\omega}^E := \begin{cases} \omega_h^{T^{(a'_E)}} \cup \omega_h^E \,, \, E \in \mathcal{E}_{\mathcal{F}_h(\sigma_h)} \\ \omega_h^E \,, \text{ otherwise} \end{cases}.
$$

Applying the Cauchy–Schwarz inequality in (5.26) and taking into account that ω_h^T and ω_h^E have a finite overlap, it follows that

$$|r(e - P_h^C e)| \lesssim \Big(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} (\eta_T^{(1)})^2 \Big)^{1/2} \Big(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} |e|_{1,\hat{\omega}_T}^2 \Big)^{1/2}$$

$$+ \Big(\sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} (\eta_E^{(1)})^2 \Big)^{1/2} \Big(\sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} |e|_{1,\tilde{\omega}_E}^2 \Big)^{1/2} \lesssim$$

$$\Big(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} (\eta_T^{(1)})^2 + \sum_{E \in \mathcal{E}_{\mathcal{Z}_h} \cup \mathcal{E}_{\mathcal{F}_h(\sigma_h)}} (\eta_E^{(1)})^2 \Big)^{1/2}.$$

$$\Big(\sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} |e|_{1,\hat{\omega}_T}^2 + \sum_{E \in \mathcal{E}_{\mathcal{Z}_h}} |e|_{1,\hat{\omega}_E}^2 \Big)^{1/2} \lesssim \eta_h^{(1)} |e|_{1,\Omega}.$$

Using the preceding inequality in (5.18) gives (5.17a).

For the proof of (5.17b) we set $e := p(y_h) - p_h$ and obtain

$$\|e\|_{1,\Omega}^2 \lesssim a(e,e) = r(e - P_h^C e), \tag{5.27}$$

where the residual $r(\cdot)$ is given by

$$r(v) := (y^d - y_h, v)_{0,\Omega} - \langle \tilde{\mu}_h, v \rangle - a(p_h, v), \quad v \in V.$$

The representation (3.26b) of the extension $\tilde{\mu}_h$ yields

$$r(e - P_h^C e) = \sum_{T \in \mathcal{T}_h(\Omega)} (y^d - y_h, e - P_h^C e)_{0,T} - \sum_{E \in \mathcal{E}_h(\Omega)} (\nu_E \cdot [\nabla p_h]_E, e - P_h^C e)_{0,E}$$

$$- \langle \tilde{\mu}_h, e - P_h^C e \rangle = \sum_{T \in \mathcal{T}_h(\mathcal{I}_h)} (y^d - y_h, e - P_h^C e)_{0,T}$$

$$- \sum_{E \in \mathcal{E}_{\mathcal{I}_h}} (\nu_E \cdot [\nabla p_h]_E, e - P_h^C e)_{0,E} - F_h^{(\mu)}(P_h^{SZ}(e - P_h^C e)).$$

The terms on the right-hand side can be estimated from above in much the same way as before resulting in

$$|r(e - P_h^C e)| \lesssim \eta_h^{(2)} |e|_{1,\Omega}, \tag{5.28}$$

which together with (5.27) allows to conclude. $\qquad\square$

Lemma 5.3. *Under the assumptions of* Theorem 5.1 *let* $y, y(u_h)$ *be the state and the auxiliary state and let* $p, p(y_h)$ *be the adjoint state and the auxiliary adjoint state. Further, let* $\eta_h^{(1)}$ *and* $\eta_h^{(2)}$ *be the components of the residual a posteriori error estimator according to* (5.4a) *and* (5.4b) *and let* $e_{h,\sigma}^{(1)}, e_{h,\mu}^{(2)}$ *be the consistency error terms given by* (5.7a),(5.7b)*. Then, it holds*

$$\|y - y(u_h)\|_{1,\Omega}^2 \lesssim \|e_{h,u}\|_{0,\Omega}^2 + (\eta_h^{(1)})^2 + e_{h,\sigma}^{(1)}, \tag{5.29a}$$

$$\|p - p(y_h)\|_{1,\Omega}^2 \lesssim \|e_{h,y}\|_{0,\Omega}^2 + (\eta_h^{(2)})^2 + e_{h,\mu}^{(2)}. \tag{5.29b}$$

Proof. Subtracting (5.15a) from (2.13a) yields

$$a(y - y(u_h), v) = (e_{h,u}, v)_{0,\Omega} + \langle \tilde{\sigma}_h - \sigma, v \rangle, \quad v \in V. \tag{5.30}$$

Choosing $v = y - y(u_h)$ and observing (2.4b), we get

$$\gamma \|y - y(u_h)\|_{1,\Omega}^2 \le a(y - y(u_h), y - y(u_h)) = \tag{5.31}$$

$$(e_{h,u}, y - y(u_h))_{0,\Omega} + \langle \tilde{\sigma}_h - \sigma, y_h - y(u_h) \rangle + e_{h,\sigma}^{(1)}.$$

The Cauchy–Schwarz inequality and Young's inequality give

$$|(e_{h,u}, y - y(u_h))_{0,\Omega}| \le \frac{\gamma}{4} \|y - y(u_h)\|_{0,\Omega}^2 + \frac{1}{\gamma} \|e_{h,u}\|_{0,\Omega}^2. \tag{5.32}$$

Moreover, if we choose $v = y_h - y(u_h)$ in (5.30), we obtain

$$\langle \tilde{\sigma}_h - \sigma, y_h - y(u_h) \rangle = (e_{h,u}, y(u_h) - y_h)_{0,\Omega} + a(y - y(u_h), y_h - y(u_h)).$$

Another application of the Cauchy–Schwarz inequality and Young's inequality yield

$$|\langle \tilde{\sigma}_h - \sigma, y_h - y(u_h) \rangle| \le \tag{5.33}$$

$$\frac{\gamma}{4} \|y - y(u_h)\|_{1,\Omega}^2 + \frac{2}{\gamma} \|y_h - y(u_h)\|_{1,\Omega}^2 + \frac{\gamma}{4} \|e_{h,u}\|_{0,\Omega}^2.$$

Using (5.32),(5.33) in (5.31) and setting

$$C_1 := \frac{\gamma^2 + 4}{2\gamma^2}, \quad C_2 := \frac{4}{\gamma^2}, \quad C_3 := \frac{2}{\gamma}, \tag{5.34}$$

it follows that

$$\|y - y(u_h)\|_{1,\Omega}^2 \le C_1 \|e_{h,u}\|_{0,\Omega}^2 + C_2 \|y_h - y(u_h)\|_{1,\Omega}^2 + C_3 e_{h,\sigma}^{(1)}. \tag{5.35}$$

The second term on the right-hand side in (5.35) can be estimated from above by (5.17a) which results in (5.29a).

The estimate (5.29b) can be established by using similar arguments. In fact, subtracting (5.15b) from (2.13c) yields

$$a(p - p(y_h), v) = -(e_{h,y}, v)_{0,\Omega} + \langle \tilde{\mu}_h - \mu, v \rangle, \quad v \in V. \tag{5.36}$$

Choosing $v = p - p(y_h)$ and $v = p_h - p(y_h)$, we obtain

$$\gamma \, \|p - p(y_h)\|^2_{1,\Omega} \leq a(p - p(y_h), p - p(y_h))$$
$$= (e_{h,y}, p(y_h) - p)_{0,\Omega} + \langle \tilde{\mu}_h - \mu, p_h - p(y_h) \rangle + e^{(2)}_{h,\mu},$$
$$\langle \tilde{\mu}_h - \mu, p_h - p(y_h) \rangle = (e_{h,y}, p_h - p(y_h))_{0,\Omega} + a(p - p(y_h), p_h - p(y_h)).$$

An application of the Cauchy–Schwarz inequality and Young's inequality gives

$$\|p - p(y_h)\|^2_{1,\Omega} \leq C_1 \, \|e_{h,y}\|^2_{0,\Omega} + C_2 \, \|p_h - p(y_h)\|^2_{1,\Omega} + C_3 \, e^{(2)}_{h,\mu}, \qquad (5.37)$$

from which (5.29b) can be deduced in view of (5.17b). $\qquad\qquad\square$

Lemma 5.4. *Under the assumptions of* Theorem 5.1 *let* $\eta_h, e^c_{h,rel}$, *and* $osc_{h,rel}$ *be the residual-type error estimator* (5.3), *the consistency error term* (5.6), *and the data oscillation* (5.10). *Then, it holds*

$$\|e_{h,u}\|^2_{0,\Omega} \lesssim \eta_h^2 + e^c_{h,rel} + osc^2_{h,rel}. \qquad (5.38)$$

Proof. Combining (2.13d) and (3.17d) we obtain

$$\|e_{h,u}\|^2_{0,\Omega} = (e_{h,u}, u - u_h)_{0,\Omega} \qquad\qquad (5.39)$$
$$= (e_{h,u}, u^d - u^d_h)_{0,\Omega} + (e_{h,u}, (u - u^d) - (u_h - u^d_h))_{0,\Omega}$$
$$= (e_{h,u}, u^d - u^d_h)_{0,\Omega} + \alpha^{-1} \, (e_{h,u}, p - p_h)_{0,\Omega}.$$

The first term on the right-hand side in (5.39) can be estimated from above by

$$|(e_{h,u}, u^d - u^d_h)_{0,\Omega}| \leq \frac{1}{4} \, \|e_{h,u}\|^2_{0,\Omega} + osc^2_h(u^d). \qquad (5.40)$$

The second term can be split according to

$$(e_{h,u}, p - p_h)_{0,\Omega} = (e_{h,u}, p - p(y_h))_{0,\Omega} + (e_{h,u}, p(y_h) - p_h)_{0,\Omega}. \qquad (5.41)$$

For the estimation of the first term on the right-hand side in (5.41) we choose $v = p - p(y_h)$ in (5.30) which gives

$$a(y - y(u_h), p - p(y_h)) = (e_{h,u}, p - p(y_h))_{0,\Omega} + \langle \tilde{\sigma}_h - \sigma, p - p(y_h) \rangle. \qquad (5.42)$$

On the other hand, choosing $v = y - y(u_h)$ in (5.36) yields

$$a(p - p(y_h), y - y(u_h)) = -(e_{h,y}, y - y(u_h))_{0,\Omega} + \langle \tilde{\mu}_h - \mu, y - y(u_h) \rangle. \qquad (5.43)$$

Combining (5.42) and (5.43) and using the symmetry of $(\cdot, \cdot)$, it follows that

$$(e_{h,u}, p - p(y_h))_{0,\Omega} = -(e_{h,y}, y - y(u_h))_{0,\Omega} + \qquad (5.44)$$

$$\langle \tilde{\sigma}_h - \sigma, p(y_h) - p_h \rangle + \langle \tilde{\mu}_h - \mu, y_h - y(u_h) \rangle + e_{h,\sigma}^{(2)} + e_{h,\mu}^{(1)}.$$

Now, choosing $v = p(y_h) - p_h$ in (5.30) and $v = y_h - y(u_h)$ in (5.36), for the second and third term on the right-hand side in (5.44) we find

$$\langle \tilde{\sigma}_h - \sigma, p(y_h) - p_h \rangle = -(e_{h,u}, p(y_h) - p_h)_{0,\Omega} + a(y - y(u_h), p(y_h) - p_h),$$

$$\langle \tilde{\mu}_h - \mu, y_h - y(u_h) \rangle = (e_{h,y}, y_h - y(u_h))_{0,\Omega} + a(p - p(y_h), y_h - y(u_h)),$$

and hence,

$$(e_{h,u}, p - p(y_h))_{0,\Omega} = -\|e_{h,y}\|_{0,\Omega}^2 - (e_{h,u}, p(y_h) - p_h)_{0,\Omega} + \qquad (5.45)$$

$$a(p - p(y_h), y_h - y(u_h)) + a(y - y(u_h), p(y_h) - p_h) + e_{h,\sigma}^{(2)} + e_{h,\mu}^{(1)}.$$

Using (5.45) in (5.41) results in

$$(e_{h,u}, p - p_h)_{0,\Omega} = a(p - p(y_h), y_h - y(u_h)) + \qquad (5.46)$$

$$a(y - y(u_h), p(y_h) - p_h) - \|e_{h,y}\|_{0,\Omega}^2 + e_{h,\sigma}^{(2)} + e_{h,\mu}^{(1)}.$$

For the first term on the right-hand side in (5.46), Young's inequality gives

$$|a(p - p(y_h), y_h - y(u_h))| \leq \frac{\varepsilon}{2} \|y_h - y(u_h)\|_{1,\Omega}^2 + \frac{1}{2\varepsilon} \|p - p(y_h)\|_{1,\Omega}^2.$$

Using (5.37) and choosing $\varepsilon = C_1/2$, we get

$$|a(p - p(y_h), y_h - y(u_h))| \leq \qquad (5.47)$$

$$\|e_{h,y}\|_{0,\Omega}^2 + \frac{C_2}{C_1} \|p_h - p(y_h)\|_{1,\Omega}^2 + \frac{C_1}{4} \|y_h - y(u_h)\|_{1,\Omega}^2 + \frac{C_3}{C_1} e_{h,\mu}^{(2)}.$$

The second term on the right-hand side in (5.46) can be estimated from above similarly:

$$|a(y - y(u_h), p(y_h) - p_h)| \leq \frac{\varepsilon}{2} \|p_h - p(y_h)\|_{1,\Omega}^2 + \frac{1}{2\varepsilon} \|y - y(u_h)\|_{1,\Omega}^2.$$

Observing (5.35), we choose $\varepsilon = 2C_1/\alpha$ and obtain

$$|a(y - y(u_h), p(y_h) - p_h)| \leq \frac{\alpha}{4} \|e_{h,u}\|_{0,\Omega}^2 + \frac{\alpha C_2}{4C_1} \|y_h - y(u_h)\|_{1,\Omega}^2 \qquad (5.48)$$

$$+ \frac{C_1}{\alpha} \|p_h - p(y_h)\|_{1,\Omega}^2 + \frac{\alpha C_3}{4C_1} e_{h,\sigma}^{(1)}.$$

Using (5.40) and (5.46)–(5.48) in (5.39), it follows that

$$\|e_{h,u}\|_{0,\Omega}^2 \lesssim \|p_h - p(y_h)\|_{1,\Omega}^2 + \|y_h - y(u_h)\|_{1,\Omega}^2 + e_{h,rel}^c + osc_{rel}^2. \qquad (5.49)$$

The assertion (5.38) follows from (5.49) by taking (5.17a),(5.17b) from Lemma 5.2 into account. □

Proof of Theorem 5.1. In view of

$$e_{h,y} = y - y(u_h) + y(u_h) - y_h,$$
$$e_{h,p} = p - p(y_h) + p(y_h) - p_h,$$

the estimate (5.14) follows from the preceding Lemmas 5.2, 5.3, and 5.4. □

5.3 Efficiency of the Error Estimator

Theorem 5.5. *Let (y, σ, u, p, μ) and $(y_h, \sigma_h, u_h, p_h, \mu_h)$ be solutions of (2.13a)–(2.13g) and (3.17a)–(3.17f) and let $\eta_h, e_{h,eff}^c, osc_{h,eff}$ be the residual-type error estimator, the consistency error, and the data oscillations as given by (5.3),(5.8), and (5.12). Then, it holds*

$$\eta_h^2 - e_{h,eff}^c - osc_{h,eff}^2 \lesssim \||e_h\||^2. \qquad (5.50)$$

The proof of Theorem 5.5 will be provided by the subsequent two lemmas taking into account the following well-known properties (cf., e.g., [44]) of the element bubble functions

$$\|q_h\|_{0,T}^2 \lesssim (q_h, q_h \, b_T)_{0,T}, \quad q_h \in P_1(T), \qquad (5.51a)$$

$$\|q_h \, b_T\|_{0,T} \lesssim \|q_h\|_{0,T}, \quad q_h \in P_1(T), \qquad (5.51b)$$

$$h_T^{-1} \|q_h\|_{0,T} \lesssim |q_h \, b_T|_{1,T} \lesssim h_T^{-1} \|q_h\|_{0,T}, \quad q_h \in P_1(T), \qquad (5.51c)$$

and of the edge bubble functions

$$\|q_h\|_{0,E}^2 \lesssim (q_h, q_h\, b_E)_{0,E}, \quad q_h \in P_1(E), \tag{5.52a}$$

$$\|q_h\, b_E\|_{0,E} \lesssim h_E^{1/2}\, \|q_h\|_{0,E}, \quad q_h \in P_1(E), \tag{5.52b}$$

$$h_E^{-1/2}\, \|q_h\|_{0,E} \lesssim |q_h\, b_E|_{1,\omega_E} \lesssim h_E^{-1/2}\, \|q_h\|_{0,E}, \quad q_h \in P_1(E). \tag{5.52c}$$

Lemma 5.6. *Under the assumptions of* Theorem 5.5 *let* $\eta_T^{(\nu)}, 1 \le \nu \le 2, e_T^\sigma, e_T^\mu,$ *and* $osc_T(f), osc_T(y^d)$ *be the element residuals* (5.5a)*, the consistency error terms* (5.9a),(5.9b)*, and the data oscillations* (5.13)*. Then, for all* $T \in \mathcal{T}_h(\mathcal{Z}_h)$ *it holds*

$$\eta_T^{(1)} \lesssim \|e_{h,y}\|_{1,T} + h_T\, \|e_{h,u}\|_{0,T} + e_T^\sigma + osc_T(f), \tag{5.53}$$

whereas for all $T \in \mathcal{T}_h(\mathcal{I}_h)$ *we have*

$$\eta_T^{(2)} \lesssim \|e_{h,p}\|_{1,T} + h_T\, \|e_{h,y}\|_{0,T} + e_T^\mu + osc_T(y^d). \tag{5.54}$$

Proof. Setting $\psi_T^\sigma := (f_h + u_h)\, b_T$, using (5.51a), $\Delta y_h|_T = 0$, Green's formula, and $\psi_T^\sigma|_{\partial T} = 0$, we obtain

$$h_T^2\, \|f_h + u_h\|_{0,T}^2 \lesssim h_T^2\, (f_h + u_h, \psi_T^\sigma)_{0,T} = \tag{5.55}$$

$$h_T^2\, (f_h + u_h + \Delta y_h, \psi_T^\sigma)_{0,T} = h_T^2\, (f_h + u_h, \psi_T^\sigma)_{0,T} - h_T^2\, a(y_h, \psi_T^\sigma).$$

On the other hand, since ψ_T^σ is an admissible test function in (2.13a), we have

$$a(y, \psi_T^\sigma) - (f + u, \psi_T^\sigma)_{0,T} + \langle \sigma, \psi_T^\sigma \rangle = 0. \tag{5.56}$$

Using (5.56) in (5.55), it follows that

$$h_T^2\, \|f_h + u_h\|_{0,T}^2 \lesssim h_T^2\, \left(a(y, \psi_T^\sigma) - (f + u, \psi_T^\sigma)_{0,T} + \langle \sigma, \psi_T^\sigma \rangle \right) - \tag{5.57}$$

$$h_T^2\, \left(a(y_h, \psi_T^\sigma) - (f_h + u_h, \psi_T^\sigma)_{0,T} \right) =$$

$$h_T^2\, \left(a(y - y_h, \psi_T^\sigma) - (f - f_h, \psi_T^\sigma)_{0,T} - (u - u_h, \psi_T^\sigma)_{0,T} + \langle \sigma, \psi_T^\sigma \rangle \right) \le$$

$$h_T^2\, \left(|e_{h,y}|_{1,T}\, |\psi_T^\sigma|_{1,T} + \|e_{h,u}\|_{0,T}\, \|\psi_T^\sigma\|_{0,T} + e_T^\sigma\, |\psi_T^\sigma|_{1,T} \right).$$

In view of (5.51b) and (5.51c), it holds

$$h_T^{-1}\, \|f_h + u_h\|_{0,T} \lesssim |\psi_T^\sigma|_{1,T} = |(f_h + u_h)\, b_T| \lesssim h_T^{-1}\, \|f_h + u_h\|_{0,T}, \tag{5.58}$$

$$\|\psi_T^\sigma\|_{0,T} \lesssim \|f_h + u_h\|_{0,T}.$$

Now, using (5.58) in (5.57), we get

$$h_T^2 \, \|f_h + u_h\|_{0,T}^2 \lesssim h_T \, \|f_h + u_h\|_{0,T} \left(\|e_{h,y}\|_{1,T} + h_T \, \|e_{h,u}\|_{0,T} + e_T^\sigma + \mathrm{osc}_T(f) \right).$$

Combining the preceding estimate with $\eta_T^{(1)} \le h_T \, \|f_h + u_h\|_{0,T} + \mathrm{osc}_T(f)$ yields (5.53). The assertion (5.54) can be shown by similar arguments. $\qquad\square$

Lemma 5.7. *Under the assumptions of* Lemma 5.6 *let* $\eta_E^{(v)}, 1 \le v \le 2$, *and* $e_{\omega_E}^\sigma, e_{\omega_E}^\mu$ *be the edge residuals and consistency error terms as given by* (5.5b) *and* (5.9c),(5.9d). *Further, for* $E = T_+ \cap T_-, T_\pm \in \mathcal{T}_h(\Omega)$ *let*

$$\eta_{\omega_E}^{(1)} := \eta_{T_+}^{(1)} + \eta_{T_-}^{(1)}, \quad \mathrm{osc}_{\omega_E}(f) := \mathrm{osc}_{T_+}(f) + \mathrm{osc}_{T_-}(f), \tag{5.59a}$$

$$\eta_{\omega_E}^{(2)} := \eta_{T_+}^{(2)} + \eta_{T_-}^{(2)}, \quad \mathrm{osc}_{\omega_E}(y^d) := \mathrm{osc}_{T_+}(y^d) + \mathrm{osc}_{T_-}(y^d). \tag{5.59b}$$

Then, for $E \in \mathcal{E}_{\mathcal{Z}_h}$ *we have*

$$\eta_E^{(1)} \lesssim \|e_{h,y}\|_{1,\omega_E} + h_E \, \|e_{h,u}\|_{0,\omega_E} + \eta_{\omega_E}^{(1)} + e_{\omega_E}^\sigma + \mathrm{osc}_{\omega_E}(f), \tag{5.60}$$

whereas for all $E \in \mathcal{E}_{\mathcal{T}_h}$ *it holds*

$$\eta_E^{(2)} \lesssim \|e_{h,p}\|_{1,\omega_E} + h_E \, \|e_{h,y}\|_{0,\omega_E} + \eta_{\omega_E}^{(2)} + e_{\omega_E}^\mu + \mathrm{osc}_{\omega_E}(y^d). \tag{5.61}$$

Proof. For $E \in \mathcal{E}_{\mathcal{Z}_h}$ we set $\psi_E^\sigma := \nu_E \cdot [\nabla y_h]_E \, b_E$. Then, (5.52a) implies

$$(\eta_E^{(1)})^2 = h_E \, \|\nu_E \cdot [\nabla y_h]_E\|_{0,E}^2 \lesssim h_E \, (\nu_E \cdot [\nabla y_h]_E, \psi_E^\sigma)_{0,E} \tag{5.62}$$

$$= h_E \, (\nu_{\partial T_+} \cdot \nabla y_h|_{\partial T_+}, \psi_E^\sigma)_{0,\partial T_+} + h_E \, (\nu_{\partial T_-} \cdot \nabla y_h|_{\partial T_-}, \psi_E^\sigma)_{0,\partial T_-},$$

where we have used that $\psi_E^\sigma|_{E'} = 0, E' \in \partial T_\pm \setminus \{E\}$. Further, Green's formula and $\Delta y_h|_{T_\pm} = 0$ yield

$$a_{T_\pm}(y_h, \psi_E^\sigma) = (\nabla y_h, \nabla \psi_E^\sigma)_{0,T_\pm} = (\nu_{\partial T_\pm} \cdot \nabla y_h|_{T_\pm}, \psi_E^\sigma)_{0,T_\pm}. \tag{5.63}$$

Using (5.63) in (5.62) gives

$$(\eta_E^{(1)})^2 \lesssim h_E \, a_{\omega_E}(y_h, \psi_E^\sigma). \tag{5.64}$$

Taking into account that ψ_E^σ is an admissible test function in (2.13a), we get

$$a_{\omega_E}(y, \psi_E^\sigma) - (f + u, \psi_E^\sigma)_{0,\omega_E} + \langle \sigma, \psi_E^\sigma \rangle = 0. \tag{5.65}$$

Combining (5.64) and (5.65), we obtain

$$(\eta_E^{(1)})^2 \lesssim h_E \, a(y_h - y, \psi_E^\sigma) + h_E \, (f_h + u_h, \psi_E^\sigma)_{0,\Omega_E} + \tag{5.66}$$
$$h_E \, (f - f_h, \psi_E^\sigma)_{0,\omega_E} + h_E \, (u - u_h, \psi_E^\sigma)_{0,\omega_E} - h_E \, \langle \sigma, \psi_E^\sigma \rangle$$
$$\leq h_E \, |y - y_h|_{1,\omega_E} \, |\psi_E^\sigma|_{1,\omega_E} + h_E \, \|\psi_E^\sigma\|_{0,\omega_E} \left(\|f_h + u_h\|_{0,\omega_E} + \right.$$
$$\left. \|u - u_h\|_{0,\omega_E} + \|f - f_h\|_{0,\omega_E} \right) + h_E \, e_{\omega_E}^\sigma \, |\psi_E^\sigma|_{1,\omega_E}.$$

Moreover, (5.52b) and (5.52c) imply

$$h_E^{-1/2} \, \|\nu_E \cdot [\nabla y_h]_E\|_{0,E} \lesssim |\psi_E^\sigma|_{1,\omega_E} = |\nu_E \cdot [\nabla y_h]_E \, b_E|_{1,\omega_E} \tag{5.67}$$
$$\lesssim h_E^{-1/2} \, \|\nu_E \cdot [\nabla y_h]_E\|_{0,E},$$
$$\|\psi_E^\sigma\|_{0,\omega_E} \lesssim h_E^{1/2} \, \|\nu_E \cdot [\nabla y_h]_E\|_{0,E}.$$

Using (5.67) in (5.66) yields

$$\eta_E^{(1)} \lesssim \|e_{h,y}\|_{1,\omega_E} + h_E \, \|e_{h,u}\|_{0,\omega_E} + h_E \, \|f_h + u_h\|_{0,\omega_E} + e_{\omega_E}^\sigma + \mathrm{osc}_{\omega_E}(f).$$

Due to the shape regularity of the triangulation, for $E \in \mathcal{E}_h(T)$ we have $h_E \lesssim h_T \lesssim h_E$ and hence,

$$h_E \, \|f_h + u_h\|_{0,\omega_E} \leq h_E \, \|f_h + u_h\|_{0,T_+} + h_E \, \|f_h + u_h\|_{0,T_-} \lesssim$$
$$h_{T_+} \, \|f_h + u_h\|_{0,T_+} + h_{T_-} \, \|f_h + u_h\|_{0,T_-} \lesssim \eta_{\omega_E}^{(1)}.$$

The preceding two estimates result in (5.60). The assertion (5.61) can be verified by similar arguments. $\qquad\square$

5.4   Estimation of the Consistency Error

In this subsection, we provide fully computable quantities for the approximation of the reliability and efficiency related consistency errors.

5.4.1   Approximation of Characteristic Functions

In this paragraph, following [15, 17, 28] in case of adaptive finite element approximations of control and/or state constrained optimally controlled second order elliptic boundary value problems, we provide approximations of the characteristic functions $\chi_{\mathcal{A}}$ and $\chi_{\mathcal{Z}}$ of the active set $\mathcal{A}$ and the zero set $\mathcal{Z}$ by means of the available finite

element solutions. Here and in the forthcoming paragraphs we will use realizations $\sigma_h', \mu_h' \in V_h$ of the discrete multipliers σ_h, μ_h with respect to the finite element spaces V_h according to

$$(\sigma_h', v_h)_{0,\Omega} = \langle\langle \sigma_h, v_h \rangle\rangle, \quad (\mu_h', v_h)_{0,\Omega} = \langle\langle \mu_h, v_h \rangle\rangle, \quad v_h \in V_h.$$

Moreover, we introduce a mesh function $\bar{h} \in S_h^{(1)}$ whose nodal values $\bar{h}(a)$ are given by averaging over local patches:

$$\bar{h}(a) := (\mathrm{card}(\omega_a))^{-1} \sum_{T \in \mathcal{T}_h(\omega_a)} h_T, \quad a \in \mathcal{N}_h(\bar{\Omega}).$$

The approximations of the characteristic functions are defined by means of

$$\chi_{h,\mathcal{A}}(a) := 1 - \frac{(\psi_h - y_h)(a)}{\gamma \bar{h}(a)^r + (\psi_h - y_h)(a)}, \quad a \in \mathcal{N}_h(\bar{\Omega}), \tag{5.68a}$$

$$\chi_{h,\mathcal{Z}}(a) := 1 - \frac{\sigma_h'(a)}{\gamma \bar{h}(a)^r + \sigma_h'(a)}, \quad a \in \mathcal{N}_h(\bar{\Omega}), \tag{5.68b}$$

where $0 < \gamma \leq 1$ and $r > 0$ are fixed. In case of uniform meshes with $\bar{h} \approx h = \max_{T \in \mathcal{T}_h(\bar{\Omega})} h_T$, the following result reflects the approximation properties of $\chi_{h,\mathcal{A}}$ and $\chi_{h,\mathcal{Z}}$.

Proposition 5.8. *For $0 \leq \varepsilon < 1$ and γ, r as in* (5.68a),(5.68b) *consider the partition*

$$\mathcal{I} \cap \mathcal{I}_h = \mathcal{I}_1 \cup \mathcal{I}_2,$$

where the sets $\mathcal{I}_\nu, 1 \leq \nu \leq 2$, are given by

$$\mathcal{I}_1 := \{x \in \mathcal{I} \mid 0 < \psi_h(x) - y_h(x) \leq \gamma h^{\varepsilon r}\}, \quad \mathcal{I}_2 := \{x \in \mathcal{I} \mid \psi_h(x) - y_h(x) > \gamma h^{\varepsilon r}\}.$$

Then, it holds

$$\|\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}}\|_{0,\omega} \begin{cases} = 0 & , \omega \subset \mathcal{A} \cap \mathcal{A}_h \\ < \min(|\omega|^{1/2}, \gamma^{-1}h^{-r}\|\psi_h - y_h\|_{0,\Omega} & , \omega \subset \mathcal{A} \cap \mathcal{I}_h \\ = |\omega|^{1/2} & , \omega \subset \mathcal{I} \cap \mathcal{A}_h \\ < |\omega|^{1/2} & , \omega \subset \mathcal{I}_1 \\ < |\omega|^{1/2}h^{r(1-\varepsilon)} & , \omega \subset \mathcal{I}_2 \end{cases}.$$

Proof. Without loss of generality we may assume $h \leq 1$. For the proof we distinguish several cases.

Case 1 ($\omega \subset \mathcal{A} \cap \mathcal{A}_h$): Obviously, $\chi_{\mathcal{A}}|_\omega = \chi_{h,\mathcal{A}}|_\omega = 1$.

Case 2 ($\omega \subset \mathcal{A} \cap \mathcal{I}_h$): We have $\chi_{\mathcal{A}}|_\omega = 1$ and hence,

$$(\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}})|_\omega = \frac{(\psi_h - y_h)|_\omega}{\gamma h^r + (\psi_h - y_h)|_\omega}.$$

Since $(\psi_h - y_h)|_\omega > 0$ and $\gamma h^r > 0$, it follows that

$$(\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}})|_\omega < \gamma^{-1} h^{-r}(\psi_h - y_h)|_\omega \quad \text{and} \quad (\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}})|_\omega < 1,$$

which allows to conclude.

Case 3 ($\omega \subset \mathcal{I} \cap \mathcal{A}_h$): The assertion follows readily from $\chi_{\mathcal{A}}|_\omega = 0$ and $\chi_{h,\mathcal{A}}|_\omega = 1$.

Case 4 ($\omega \subset \mathcal{I} \cap \mathcal{I}_h$): We have $\chi_{\mathcal{A}}|_\omega = 0$ and

$$(\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}})|_\omega = \frac{\gamma h^r}{\gamma h^r + (\psi_h - y_h)|_\omega}.$$

For $\omega \subset \mathcal{I}_1$ this implies $(\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}})|_\omega < 1$, and we conclude. On the other hand, for $\omega \subset \mathcal{I}_2$, taking $h \leq 1$ into account, we find

$$(\chi_{\mathcal{A}} - \chi_{h,\mathcal{A}})|_\omega < \min(1, h^{r(1-\varepsilon)}) = h^{r(1-\varepsilon)},$$

which proves the assertion.   □

5.4.2 Approximation of the Continuous Active/Inactive Sets

Based on the approximations $\chi_{h,\mathcal{A}}, \chi_{h,\mathcal{Z}}$ of the characteristic functions of the continuous sets $\mathcal{A}$ and $\mathcal{Z}$, we derive approximations of the continuous (strongly) active, biactive, inactive, and zero sets. To this end, for $0 < \kappa \leq 1$ and $0 < r' \leq r$ we first define nodal sets $\bar{A}_h, \bar{I}_h, \bar{C}_h, \bar{Z}_h$, and $\bar{B}_h$ as approximations of their continuous counterparts according to

$$\bar{A}_h := \{a \in \mathcal{N}_h(\bar{\Omega}) \mid \chi_{h,\mathcal{A}}(a) \geq 1 - \kappa \bar{h}(a)^{r'}\}, \quad \bar{I}_h := \mathcal{N}_h(\bar{\Omega}) \setminus \bar{A}_h,$$

$$\bar{C}_h := \left(\mathcal{N}_h(\Omega) \setminus \{a \in \mathcal{N}_h(\Omega) \mid \chi_{h,\mathcal{Z}}(a) \geq 1 - \kappa \bar{h}(a)^{r'}\}\right) \cap \bar{A}_h,$$

$$\bar{Z}_h := \mathcal{N}_h(\Omega) \setminus \bar{C}_h, \quad \bar{B}_h := \bar{A}_h \cap \bar{Z}_h.$$

These sets constitute a suitable basis for the specification of approximations $\bar{\mathcal{A}}_h$ of $\mathcal{A}, \bar{\mathcal{I}}_h$ of $\mathcal{I}, \bar{\mathcal{C}}_h$ of $\mathcal{C}$, and $\bar{\mathcal{Z}}_h$ of $\mathcal{Z}$ by means of

$$\bar{\mathcal{A}}_h := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{A}}_h^T\}, \; \bar{\mathcal{A}}_h^T := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subseteq \bar{A}_h\},$$

$$\tag{5.69a}$$

$$\bar{\mathcal{I}}_h := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{I}}_h^T \cup \bar{\mathcal{F}}_{y_h}^T\}, \tag{5.69b}$$

$$\bar{\mathcal{I}}_h^T := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subseteq \bar{I}_h\}, \ \bar{\mathcal{F}}_{y_h}^T := \mathcal{T}_h(\Omega) \setminus (\bar{\mathcal{A}}_h^T \cup \bar{\mathcal{I}}_h^T),$$

$$\bar{\mathcal{C}}_h := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{C}}_h^T\}, \ \bar{\mathcal{C}}_h^T := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subseteq \bar{C}_h\} \cup \tag{5.69c}$$

$$\{T \in \mathcal{T}_h(\Omega) \mid T \cap \Gamma \neq \emptyset \wedge \mathcal{N}_h(T) \cap \mathcal{N}_h(\Omega) \neq \emptyset \wedge T \subseteq \bar{\mathcal{A}}_h^T\},$$

$$\bar{\mathcal{Z}}_h := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{Z}}_h^T \cup \bar{\mathcal{F}}_{\sigma_h}^T\}, \tag{5.69d}$$

$$\bar{\mathcal{Z}}_h^T := \{T \in \mathcal{T}_h(\Omega) \mid \mathcal{N}_h(T) \subseteq \bar{Z}_h \cup \mathcal{N}_h(\Gamma)\}, \ \bar{\mathcal{F}}_{\sigma_h}^T := \mathcal{T}_h(\Omega) \setminus (\bar{\mathcal{C}}_h^T \cup \bar{\mathcal{Z}}_h^T).$$

The biactive set $\mathcal{B}$ and the free boundaries $\mathcal{F}(y)$ and $\mathcal{F}(\sigma)$ are approximated by

$$\bar{\mathcal{B}}_h := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{B}}_h^T\}, \ \bar{\mathcal{B}}_h^T := \bar{\mathcal{A}}_h^T \setminus \bar{\mathcal{C}}_h^T, \tag{5.69e}$$

$$\bar{\mathcal{F}}_{y_h} := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{F}}_{y_h}^T\}, \tag{5.69f}$$

$$\bar{\mathcal{F}}_{\sigma_h} := \bigcup \{T \in \mathcal{T}_h(\Omega) \mid T \in \bar{\mathcal{F}}_{\sigma_h}^T\}. \tag{5.69g}$$

In the documentation of the numerical results in the following Sect. 6, we will measure the quality of the approximation of the active set $\mathcal{A}$ and the strongly active set $\mathcal{C}$ by the a posteriori quantities

$$e_{\ell,\mathcal{A}}^{dva} := \|\chi_{\mathcal{A}_\ell} - \chi_{\bar{\mathcal{A}}_\ell}\|_{L^1(\Omega)}, \quad e_{\ell,\mathcal{C}}^{dva} := \|\chi_{\mathcal{C}_\ell} - \chi_{\bar{\mathcal{C}}_\ell}\|_{L^1(\Omega)}, \tag{5.70}$$

where the upper index 'dva' stands for 'discrete versus approximate', and compare them with the quantities

$$e_{\ell,\mathcal{A}}^{evd} := \|\chi_{\mathcal{A}} - \chi_{\mathcal{A}_\ell}\|_{L^1(\Omega)}, \quad e_{\ell,\mathcal{C}}^{evd} := \|\chi_{\mathcal{C}} - \chi_{\mathcal{C}_\ell}\|_{L^1(\Omega)}, \tag{5.71a}$$

$$e_{\ell,\mathcal{A}}^{eva} := \|\chi_{\mathcal{A}} - \chi_{\bar{\mathcal{A}}_\ell}\|_{L^1(\Omega)}, \quad e_{\ell,\mathcal{C}}^{eva} := \|\chi_{\mathcal{C}} - \chi_{\bar{\mathcal{C}}_\ell}\|_{L^1(\Omega)}. \tag{5.71b}$$

Here, the upper indices 'evd' and 'eva' mean 'exact versus discrete' and 'exact versus approximate'. Obviously, these latter quantities are only available, if the exact solution is known.

5.4.3 Approximation of the Continuous States and Multipliers

We derive approximations of the state y and the adjoint state p as well as various approximations of the multipliers σ and μ in terms of the approximations of the continuous active/biactive, strongly active, inactive, zero nodal points (sets) and free boundaries provided in the previous paragraph 5.4.2. Motivated by superconvergence results through local averaging (cf., e.g., [3]), we define approximations

$\bar{y}_h \in V_h$ of y and $\bar{p}_h \in V_h$ of p according to

$$\bar{y}_h(a) := \begin{cases} \operatorname{card}(\mathcal{N}_h(\omega_a))^{-1} \sum_{a' \in \mathcal{N}_h(\omega_a)} y_h(a') \, , \, a \in \bar{I}_h \\ \psi_h(a) \, , \, a \in \bar{A}_h \end{cases}, \tag{5.72a}$$

$$\bar{p}_h(a) := \begin{cases} \operatorname{card}(\mathcal{N}_h(\omega_a))^{-1} \sum_{a' \in \mathcal{N}_h(\omega_a)} p_h(a') \, , \, a \in \bar{Z}_h \\ 0 \, , \, a \in \bar{C}_h \end{cases}. \tag{5.72b}$$

Likewise, we define approximations σ_h'' and μ_h'' of σ and μ by means of

$$\sigma_h''(a) := \begin{cases} \operatorname{card}(\mathcal{N}_h(\omega_a))^{-1} \sum_{a' \in \mathcal{N}_h(\omega_a)} \sigma_h'(a') \, , \, a \in \bar{C}_h \\ 0 \, , \, a \in \bar{Z}_h \end{cases}, \tag{5.73a}$$

$$\mu_h''(a) := \begin{cases} \operatorname{card}(\mathcal{N}_h(\omega_a))^{-1} \sum_{a' \in \mathcal{N}_h(\omega_a)} \mu_h'(a') \, , \, a \in \bar{I}_h \\ 0 \, , \, a \in \bar{A}_h \end{cases}. \tag{5.73b}$$

Remark 5.9. The functions $\bar{y}_h, \bar{p}_h$ will replace y, p in the approximation of the consistency error $e^c_{h,rel}$, whereas σ_h'', μ_h'' will be used in the approximation of $e^c_{h,eff}$ and in a further form of the approximation of $e^c_{h,rel}$ (cf. paragraph 5.4.4).

For the approximation of the multipliers σ, μ in the consistency error $e^c_{h,rel}$ we will use an alternative approximation which relies on the structural properties of the multipliers. If the sets $\mathcal{C}$ and $\mathcal{A}$ are the union of a finite number of connected pairwise disjoint Lipschitz sets, for any $v \in V$ Proposition 2.1 guarantees the existence of sets $\tilde{\mathcal{C}}, \tilde{\mathcal{A}}$ and functions $v_{\mathcal{C}}^{ext}, v_{\mathcal{A}}^{ext} \in V$ such that $\mathcal{C} \subseteq \tilde{\mathcal{C}} \subseteq \Omega, \mathcal{A} \subseteq \tilde{\mathcal{A}} \subseteq \Omega$ and

$$\langle \sigma, v \rangle = \langle \sigma, v_{\mathcal{C}}^{ext} \rangle = (f + u, v_{\mathcal{C}}^{ext})_{0,\tilde{\mathcal{C}}} - (\nabla y, \nabla v_{\mathcal{C}}^{ext})_{0,\tilde{\mathcal{C}}},$$

$$\langle \mu, v \rangle = \langle \mu, v_{\mathcal{A}}^{ext} \rangle = (y^d - y, v_{\mathcal{A}}^{ext})_{0,\tilde{\mathcal{A}}} - (\nabla p, \nabla v_{\mathcal{A}}^{ext})_{0,\tilde{\mathcal{A}}}.$$

Employing the structural information provided in Proposition 2.16, we obtain

$$\langle \sigma, v \rangle = \tag{5.74a}$$

$$\left((f + u^d, v)_{0,\mathcal{C}} - (\nabla \psi, \nabla v)_{0,\mathcal{C}} \right) - \left((\Delta \psi, v_{\mathcal{C}}^{ext})_{0,(\tilde{\mathcal{C}} \setminus \mathcal{C}) \cap \mathcal{B}} + (\nabla \psi, \nabla v_{\mathcal{C}}^{ext})_{0,(\tilde{\mathcal{C}} \setminus \mathcal{C}) \cap \mathcal{B}} \right)$$

$$+ \left((f + u^d + \alpha^{-1} p, v_{\mathcal{C}}^{ext})_{0,(\tilde{\mathcal{C}} \setminus \mathcal{C}) \cap \mathcal{I}} - (\nabla y, \nabla v_{\mathcal{C}}^{ext})_{0,(\tilde{\mathcal{C}} \setminus \mathcal{C}) \cap \mathcal{I}} \right),$$

$$\langle \mu, v \rangle = (y^d - \psi, v)_{0,\mathcal{A}} \tag{5.74b}$$

$$+ \alpha \, (\nabla(\Delta \psi + f + u^d), \nabla v)_{0,\mathcal{B}} + \left((y^d - y, v_{\mathcal{A}}^{ext})_{0,\tilde{\mathcal{A}} \setminus \mathcal{A}} - (\nabla p, \nabla v_{\mathcal{A}}^{ext})_{0,\tilde{\mathcal{A}} \setminus \mathcal{A}} \right).$$

In order to provide a fully computable approximation, we replace the unknown sets $\mathcal{C}, \mathcal{B}, \mathcal{A}, \mathcal{I}$, and the unknown functions y, p by their previously defined approximations $\bar{\mathcal{C}}_h, \bar{\mathcal{B}}_h, \bar{\mathcal{A}}_h, \bar{\mathcal{I}}_h$, and $\bar{y}_h, \bar{p}_h$. Moreover, $\tilde{\mathcal{C}}, \tilde{\mathcal{A}}$ are chosen according to

$$\tilde{\mathcal{C}} := \bar{\mathcal{C}}_h \cup \bar{\mathcal{F}}_{\sigma_h}, \quad \tilde{\mathcal{A}} := \bar{\mathcal{A}}_h \cup \bar{\mathcal{F}}_{y_h}, \tag{5.75}$$

whereas $v_{\mathcal{C}}^{ext}, v_{\mathcal{A}}^{ext}$ are approximated by

$$v_{\bar{\mathcal{C}}_h}^{ext} := I_{\bar{\mathcal{C}}_h}(v_h), \quad v_{\bar{\mathcal{A}}_h}^{ext} := I_{\bar{\mathcal{A}}_h}(v_h), \quad v_h \in V_h. \tag{5.76}$$

Here, $I_{D_h}, D_h \subseteq \mathcal{N}_h(\Omega)$, is the operator from (3.19).

Using the previous approximations in (5.74) and assuming sufficient regularity of the data in $\bar{\mathcal{B}}_h$, we obtain the following approximations of the action of σ, μ on functions in V_h:

$$\langle \sigma, v_h \rangle \approx \langle \bar{\sigma}_h^{(1)}, v_h \rangle = \sum_{T \in \mathcal{T}_h(\bar{\mathcal{C}}_h)} \left((f + u^d, v_h)_{0,T} - (\nabla \psi, \nabla v_h)_{0,T} \right) \tag{5.77a}$$

$$- \sum_{T \in \mathcal{T}_h(\bar{\mathcal{F}}_{\sigma_h} \cap \bar{\mathcal{B}}_h)} \left((\Delta \psi, I_{\bar{\mathcal{C}}_h}(v_h))_{0,T} + (\nabla \psi, \nabla I_{\bar{\mathcal{C}}_h} v_h)_{0,T} \right)$$

$$+ \sum_{T \in \mathcal{T}_h(\bar{\mathcal{F}}_{\sigma_h} \cap \bar{\mathcal{I}}_h)} \left((f + u^d + \alpha^{-1} \bar{p}_h, I_{\bar{\mathcal{C}}_h}(v_h))_{0,T} - (\nabla \bar{y}_h, \nabla I_{\bar{\mathcal{C}}_h}(v_h))_{0,T} \right),$$

$$\langle \mu, v_h \rangle \approx \langle \bar{\mu}_h^{(1)}, v_h \rangle = \sum_{T \in \mathcal{T}_h(\bar{\mathcal{A}}_h)} (y^d - \psi, v_h)_{0,T} \tag{5.77b}$$

$$+ \alpha \sum_{T \in \mathcal{T}_h(\bar{\mathcal{B}}_h)} (\nabla(\Delta \psi + f + u^d), \nabla v_h)_{0,T}$$

$$+ \sum_{T \in \mathcal{T}_h(\bar{\mathcal{F}}_{y_h})} \left((y^d - \bar{y}_h, I_{\bar{\mathcal{A}}_h}(v_h))_{0,T} - (\nabla \bar{p}_h, \nabla I_{\bar{\mathcal{A}}_h}(v_h))_{0,T} \right).$$

As far as the regularity of the data is concerned, in the proof of Proposition 2.16 we have seen that $\Delta \psi \in L^2(\mathcal{B})$ and $\Delta \psi + f + u^d \in H^1(\mathcal{B})$. If $\bar{\mathcal{B}}_h \subset \mathcal{B}$ or else $\Delta \psi \in L^2(\bar{\mathcal{B}}_h)$ and $\Delta \psi + f + u^d \in L^2(\bar{\mathcal{B}}_h)$ hold true, (5.77a) and (5.77b) are well defined. Otherwise, employing the values of $\bar{y}_h$ and $\bar{p}_h$ in $\bar{\mathcal{B}}_h$, we can use the following simplification of the approximations of the action of σ, μ on functions in V_h:

$$\langle \sigma, v_h \rangle \approx \langle \bar{\sigma}_h^{(2)}, v_h \rangle = \sum_{T \in \mathcal{T}_h(\bar{\mathcal{C}}_h)} \Big((f + u^d, v_h)_{0,T} - (\nabla \psi, \nabla v_h)_{0,T} \Big) \tag{5.78a}$$

$$- \sum_{T \in \mathcal{T}_h(\bar{\mathcal{F}}_{\sigma_h} \cap \bar{\mathcal{B}}_h)} (\nabla \psi, \nabla I_{\bar{\mathcal{C}}_h}(v_h))_{0,T}$$

$$+ \sum_{T \in \mathcal{T}_h(\bar{\mathcal{F}}_{\sigma_h} \cap \bar{\mathcal{I}}_h)} \Big((f + u^d + \alpha^{-1} \bar{p}_h, I_{\bar{\mathcal{C}}_h}(v_h))_{0,T} - (\nabla \bar{y}_h, \nabla I_{\bar{\mathcal{C}}_h}(v_h))_{0,T} \Big),$$

$$\langle \mu, v_h \rangle \approx \langle \bar{\mu}_h^{(2)}, v_h \rangle = \sum_{T \in \mathcal{T}_h(\bar{\mathcal{A}}_h)} (y^d - \psi, v_h)_{0,T} - \sum_{T \in \mathcal{T}_h(\bar{\mathcal{B}}_h)} (\nabla \bar{p}_h, \nabla v_h)_{0,T}$$

$$\tag{5.78b}$$

$$+ \sum_{T \in \mathcal{T}_h(\bar{\mathcal{F}}_{y_h})} \Big((y^d - \bar{y}_h, I_{\bar{\mathcal{A}}_h}(v_h))_{0,T} - (\nabla \bar{p}_h, \nabla I_{\bar{\mathcal{A}}_h}(v_h))_{0,T} \Big).$$

5.4.4 Approximation of the Consistency Errors

For the consistency error $e_{h,rel}^c$ we will use three different types of approximations

$$e_{h,rel}^c \approx \bar{e}_{h,rel}^{c,(k)} := \bar{e}_{h,\sigma}^{1,(k)} + \bar{e}_{h,\sigma}^{2,(k)} + \bar{e}_{h,\mu}^{1,(k)} + \bar{e}_{h,\mu}^{2,(k)}, \quad 1 \leq k \leq 3. \tag{5.79}$$

For the first two approximations $\bar{e}_{h,rel}^{c,(k)}, 1 \leq k \leq 2$, we use the approximation of the multipliers by (5.77) and (5.78):

$$\bar{e}_{h,\sigma}^{1,(k)} := \langle \tilde{\sigma}_h - \bar{\sigma}_h^{(k)}, \bar{y}_h - y_h \rangle, \quad \bar{e}_{h,\sigma}^{2,(k)} := \langle \tilde{\sigma}_h - \bar{\sigma}_h^{(k)}, p_h - \bar{p}_h \rangle, \tag{5.80a}$$

$$\bar{e}_{h,\mu}^{1,(k)} := \langle \tilde{\mu}_h - \bar{\mu}_h^{(k)}, \bar{y}_h - y_h \rangle, \quad \bar{e}_{h,\mu}^{2,(k)} := \langle \tilde{\mu}_h - \bar{\mu}_h^{(k)}, \bar{p}_h - p_h \rangle. \tag{5.80b}$$

The third approximation $\bar{e}_{h,rel}^{c,(3)}$ is obtained by using the approximation of the multipliers by local averaging (cf. (5.73)):

$$\bar{e}_{h,\sigma}^{1,(3)} := \langle \tilde{\sigma}_h - \sigma_h'', \bar{y}_h - y_h \rangle, \quad \bar{e}_{h,\sigma}^{2,(3)} := \langle \tilde{\sigma}_h - \sigma_h'', p_h - \bar{p}_h \rangle, \tag{5.81a}$$

$$\bar{e}_{h,\mu}^{1,(3)} := \langle \tilde{\mu}_h - \mu_h'', \bar{y}_h - y_h \rangle, \quad \bar{e}_{h,\mu}^{2,(3)} := \langle \tilde{\mu}_h - \mu_h'', \bar{p}_h - p_h \rangle. \tag{5.81b}$$

Further, we compute upper bounds $\bar{e}_{h,rel}^{c,(k)}, 1 \leq k \leq 3$, according to

$$\bar{e}_{h,rel}^{c,(k)} \leq \bar{E}_{h,rel}^{c,(k)} := \bar{E}_{h,\sigma}^{1,(k)} + \bar{E}_{h,\sigma}^{2,(k)} + \bar{E}_{h,\mu}^{1,(k)} + \bar{E}_{h,\mu}^{2,(k)}, \quad 1 \leq k \leq 3, \tag{5.82}$$

where $\bar{E}_{h,\sigma}^{v,(k)}, \bar{E}_{h,\mu}^{v,(k)}, 1 \leq v \leq 2$, are given by summing up the absolute values of the elementwise contributions of $\bar{e}_{h,\sigma}^{v,(k)}, \bar{e}_{h,\mu}^{v,(k)}, 1 \leq v \leq 2$.

For the approximation of the consistency error $e^c_{h,eff}$ we use the approximation of the multipliers by local averaging as given by (5.73):

$$\bar{e}^c_{h,eff} \lesssim \bar{E}^c_{h,eff} := \sum_{T \in \mathcal{T}_h(\mathcal{Z}_h)} h_T^2 \, \|\sigma_h''\|_{0,T}^2 + \sum_{T \in \mathcal{T}_h(\mathcal{I}_h)} h_T^2 \, \|\mu_h''\|_{0,T}^2. \tag{5.83}$$

6 Numerical Results

In this section, we present numerical results for problems with and without strict complementarity illustrating the performance of the suggested finite element approximation. We note that for adaptively refined meshes it is appropriate to measure the decay in the error *err* in terms of the degrees of freedom (DOF) provided by the finite element mesh. In particular, if there exists a real number $\tau > 0$ such that $err = O(DOF^{-\tau})$, then τ is said to be the convergence rate of the error with respect to the degrees of freedom. In the numerical experiments, we are dealing with a hierarchy $\{\mathcal{T}_{h_n}(\Omega)\}_{n \in \mathbb{N}}$ of nested simplicial meshes with associated degrees of freedom $DOF(n)$. Denoting by $err(n), n \in \mathbb{N}$, the error with respect to the mesh $\mathcal{T}_{h_n}(\Omega)$, we refer to

$$\tau_n := \frac{\log(err(n-1)/err(n))}{\log(DOF(n)/DOF(n-1))}, \quad n \in \mathbb{N}, \tag{6.1}$$

as the experimental convergence rate in terms of the degrees of freedom. On a double logarithmic scale, the numbers τ_n correspond to the negative slopes of the lines connecting $\log(err(n-1))$ and $\log(err(n))$. In the subsequent numerical examples, we will compare these lines both for uniform refinement and adaptive refinement. In the regular case, we expect the slopes to be approximately the same, whereas for less regular solutions the slope for adaptive refinement is expected to be larger than in case of uniform refinement.

Example 6.1. We consider $A = -\Delta$ on the L-shaped domain $\Omega = (-2, 2)^2 \setminus ([0, 2] \times [-2, 0])$. In polar coordinates, given

$$y^*(r, \varphi) = -\gamma(r) \, r^{2/3} \, \sin(\tfrac{2}{3}\varphi),$$

$$\sigma^*(r) = \begin{cases} 1, r \geq \bar{r} := 0.5 \\ 0, \text{ otherwise} \end{cases}, \quad u^*(r, \varphi) = y^*(r, \varphi),$$

where

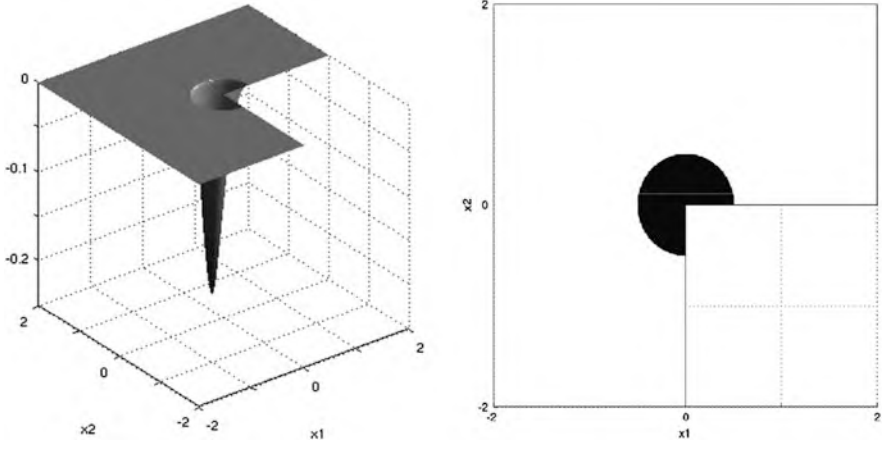$$\gamma(r) = \begin{cases} 0, r \geq \bar{r} \\ 16r^3 - 12r^2 + 1, \text{ otherwise} \end{cases},$$

Fig. 1 Example 6.1. Optimal state y^* (*left*) and inactive set $\mathcal{I}^*$, marked in *black* (*right*)
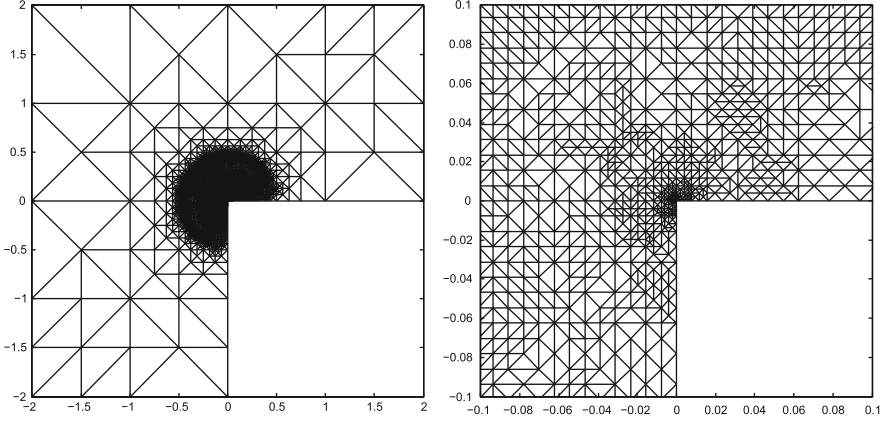

Fig. 2 Example 6.1. Final mesh (*left*) and zoom into the vicinity of the singularity at the origin (*right*)

it can be easily verified that the triple (y^*, σ^*, u^*) with the adjoint state $p^* = y^*$ and the multiplier $\mu^* = \sigma^*$ is an S-stationary point of (2.5) with respect to the data

$$y^d = \mu^* - \Delta p^* + y^*, \quad u^d = 0,$$
$$f = \sigma^* - \Delta y^* - p^*, \quad \alpha = 1, \quad \psi = 0.$$

Further, we have $\mathcal{I}^* = \{(r, \varphi) \mid r \in (0, \bar{r}), \varphi \in (0, 3\pi/2)\}$, $\mathcal{Z}^* = \mathcal{I}^*$, and hence, $\mathcal{B}^* = \emptyset$. The state y^* and the inactive set $\mathcal{I}^*$ are displayed in Fig. 1. The adaptively generated final mesh with 33468 DOFs and a zoom into the vicinity of the singularity of the state at the origin are shown in Fig. 2.
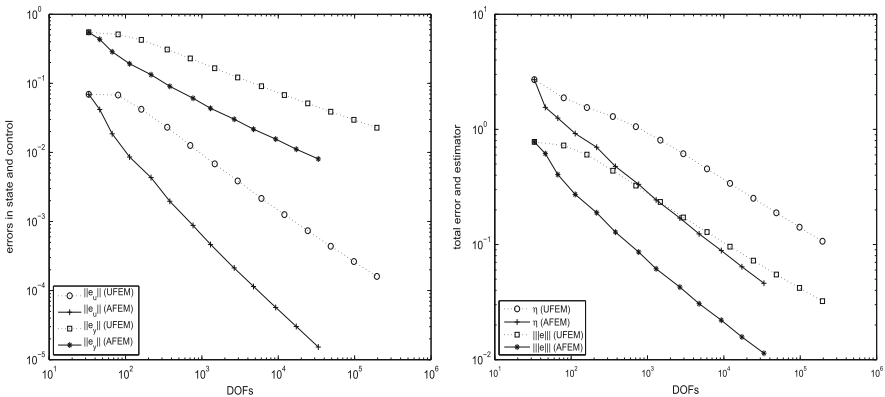
Fig. 3 Example 6.1. Convergence history: Decrease of the errors in the state $\|e_{h,y}\|_{1,\Omega}$ and in the control $\|e_{h,u}\|_{0,\Omega}$ as a function of the DOFs on a logarithmic scale (for uniform (UFEM) and adaptive (AFEM) refinement (*left*). Decrease of the estimator η_h and the total error $\||e_h\||$ as a function of the DOFs on a logarithmic scale (for uniform (UFEM) and adaptive (AFEM) refinement (*right*)

Table 1 Example 6.1: Experimental convergence rates (uniform and adaptive refinement)

n	$\|e_{h,y}\|_{1,\Omega}$		$\|e_{h,p}\|_{1,\Omega}$		$\|e_{h,u}\|_{0,\Omega}$		$\||e_h\||$	
	Unif.	Adapt.	Unif.	Adapt.	Unif.	Adapt.	Unif.	Adapt.
3	0.26	1.11	0.26	1.11	0.68	2.15	0.28	1.15
4	0.41	0.76	0.41	0.76	0.76	1.48	0.42	0.78
5	0.43	0.56	0.43	0.56	0.88	1.06	0.44	0.57
6	0.44	0.68	0.44	0.68	0.83	1.40	0.45	0.69
7	0.45	0.57	0.45	0.57	0.82	1.15	0.45	0.57
8	0.41	0.64	0.41	0.64	0.82	1.21	0.41	0.64
9	0.42	0.51	0.42	0.51	0.78	1.09	0.43	0.51
10	0.40	0.57	0.40	0.57	0.76	1.07	0.40	0.57
11	0.40	0.49	0.40	0.49	0.75	1.04	0.40	0.50
12	0.39	0.54	0.39	0.54	0.73	1.02	0.39	0.54
13	0.38	0.49	0.38	0.49	0.72	1.03	0.38	0.49

The convergence history is documented in Fig. 3 (left) which shows the decrease of the errors in the state $\|e_{h,y}\|_{1,\Omega}$ and in the control $\|e_{h,u}\|_{0,\Omega}$ as a function of the DOFs on a logarithmic scale both for uniform refinement (UFEM) and for adaptive refinement (AFEM). Likewise, Fig. 3 (right) shows the decrease of the total error $\||e_h\||$ and of the estimator η_h as a function of the DOFs on a logarithmic scale, again both for uniform refinement (UFEM) and for adaptive refinement (AFEM).

Table 1 contains the computed experimental convergence rates (cf. 6.1) for the approximation of the state, the adjoint state, the control, and the total error in case
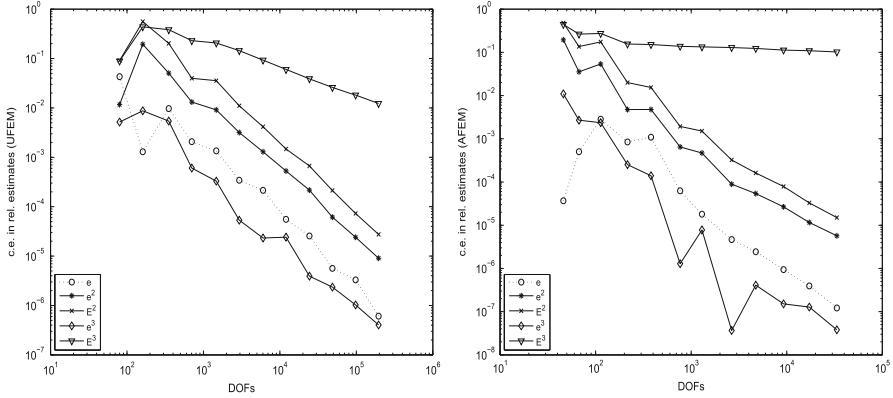
Fig. 4 Example 6.1. Decrease of the reliability related consistency error $e = |e^c_{h,rel}|$ (*dotted line*) and its estimates $e^k = |\bar{e}^{c,(k)}_{h,rel}|$, $E^k = \bar{E}^{c,(k)}_{h,rel}$, $2 \le k \le 3$, (*solid lines*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)

of uniform and adaptive refinement. We see that asymptotically the expected optimal convergence rates are achieved.

As far as the consistency errors and their estimates are concerned, we have to distinguish between the reliability related consistency errors $e^c_{h,rel}$ (cf. (5.6)) and the efficiency related consistency errors $e^c_{h,eff}$ (cf. (5.8)). Figure 4 displays the decay of $|e^c_{h,rel}|$ and its estimates $|\bar{e}^{c,(k)}_{h,rel}|$, $\bar{E}^{c,(k)}_{h,rel}$, $2 \le k \le 3$, as a function of the DOFs on a logarithmic scale for uniform refinement (left) and for adaptive refinement (right) (we note that $\bar{e}^{c,(1)}_{h,rel}$, $\bar{E}^{c,(1)}_{h,rel}$ and $\bar{e}^{c,(2)}_{h,rel}$, $\bar{E}^{c,(2)}_{h,rel}$ coincide for problems featuring strict complementarity which is the case in Example 1).

We observe that $|\bar{e}^{c,(2)}_{h,rel}|$ and $\bar{E}^{c,(2)}_{h,rel}$ provide upper bounds for $|e^c_{h,rel}|$ with approximately the same decay rates. On the other hand, $|\bar{e}^{c,(3)}_{h,rel}|$ slightly underestimates $|e^c_{h,rel}|$, whereas $\bar{E}^{c,(3)}_{h,rel}$ grossly overestimates $|e^c_{h,rel}|$ with an insufficient decay rate in particular for adaptive refinement.

Similarly, in Fig. 5 the decay of the efficiency related consistency errors $e^c_{h,eff}$ and their estimates $\bar{E}^{c,1}_{h,eff}$ are shown as functions of the DOFs on a logarithmic scale for uniform refinement (left) and adaptive refinement (right). After a pre-asymptotic phase, the estimates $\bar{E}^{c,1}_{h,eff}$ represent close upper bounds of $e^c_{h,eff}$ featuring essentially the same decay rates.
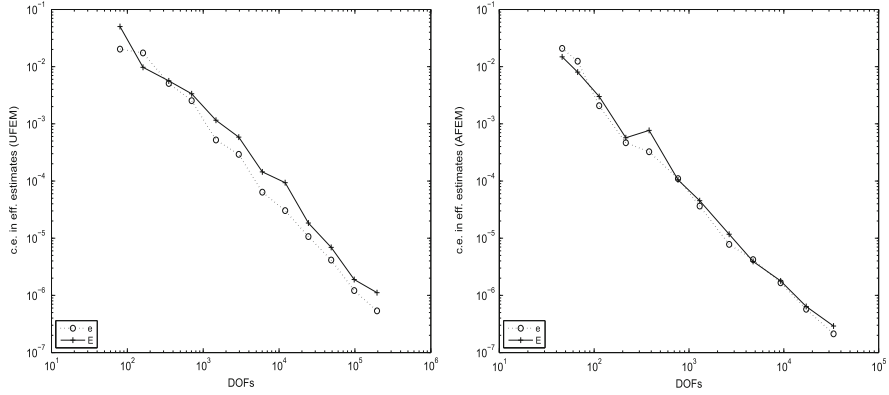
Fig. 5 Example 6.1. Decrease of the efficiency related consistency error $e = e^c_{h,eff}$ (*dotted line*) and its estimate $E = \bar{E}^{c,1}_{h,eff}$ (*solid line*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)

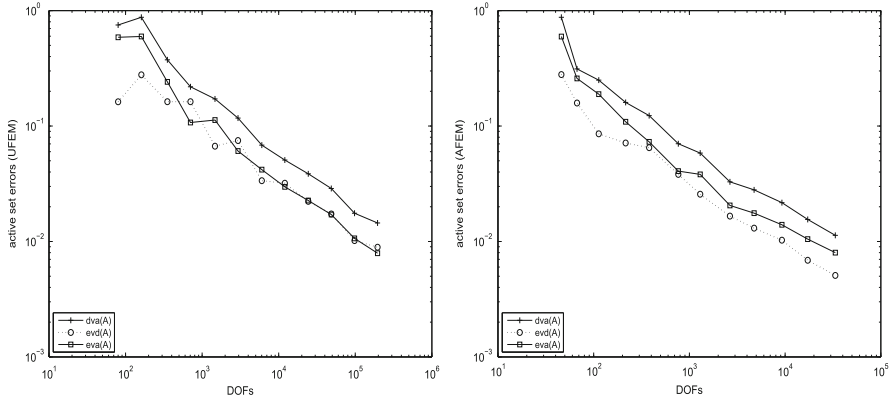**Fig. 6** Example 6.1. Approximation of the active set $\mathcal{A}$: quantities $e^{evd}_{h,\mathcal{A}}$ (*dotted line*) and $e^{dva}_{h,\mathcal{A}}$, $e^{eva}_{h,\mathcal{A}}$ (*solid lines*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)

Finally, Fig. 6 displays the decay of the errors with regard to the approximation of the active set $\mathcal{A}$ in terms of the quantities $e^{evd}_{h,\mathcal{A}}$, $e^{eva}_{h,\mathcal{A}}$, and $e^{dva}_{h,\mathcal{A}}$ (cf. (5.70),(5.71)). Recalling that the quantities $e^{evd}_{h,\mathcal{A}}$ and $e^{eva}_{h,\mathcal{A}}$ are the L^1-norms of the difference between the characteristic function of the continuous active set $\mathcal{A}$ on one hand and the characteristic function of the discrete active set $\mathcal{A}_h$ resp. the characteristic function of the approximate active set $\bar{\mathcal{A}}_h$ on the other hand, we see that the a posteriori quantity $e^{dva}_{h,\mathcal{A}}$ yields a close upper bound with approximately the same decay rates.

Example 6.2. The second example which has been considered in [18, 21] features a problem with lack of strict complementarity. We consider $A = -\Delta$ on $\Omega = (0,1)^2$. Given

$$y^*(x_1, x_2) = \begin{cases} -z_1(x_1)z_2(x_2) \,, \ (x_1, x_2) \in (0, 0.5) \times (0, 0.8) \\ 0 \,, \ \text{else} \end{cases},$$

$$\sigma^*(x_1, x_2) = 2 \max(0, -|x_1 - 0.8| - |(x_2 - 0.2)x_1 - 0.3| + 0.35),$$

$$u^*(x_1, x_2) = y^*(x_1, x_2),$$

where

$$z_1(x_1) := -4096\, x_1^6 + 6144\, x_1^5 - 3072\, x_1^4 + 512\, x_1^3,$$

$$z_2(x_2) := -244.140625\, x_2^6 + 585.9375\, x_2^5 - 468.75\, x_2^4 + 125\, x_2^3,$$

it can be easily verified that the triple (y^*, σ^*, u^*) with the adjoint state $p^* = y^*$ and the multiplier $\mu^* = \sigma^*$ is an S-stationary point of (2.5) with respect to the data

$$y^d = \mu^* - \Delta p^* + y^*, \quad u^d = 0,$$

$$f = \sigma^* - \Delta y^* - p^*, \quad \alpha = 1, \quad \psi = 0.$$

Further, we have $\mathcal{I}^* = \{(x_1, x_2) \mid (x_1, x_2) \in (0, 0.5) \times (0, 0.8)\}$, $\mathcal{C}^* = \{(x_1, x_2) \mid |x_1 - 0.8| + |(x_2 - 0.2)x_1 - 0.3| \le 0.35\}$, and hence, $\mathcal{B}^* = \Omega \setminus (\mathcal{I}^* \cup \mathcal{C}^*) \ne \varnothing$. The optimal state y^* and the optimal multiplier σ^* are shown in Fig. 7, whereas the inactive set $\mathcal{I}^*$ and the strongly active set $\mathcal{C}^*$ are displayed in Fig. 8. Figure 9 shows the adaptively generated mesh at level $n = 7$ with 2439 DOFs and the final mesh (level $n = 11$) with 34159 DOFs.
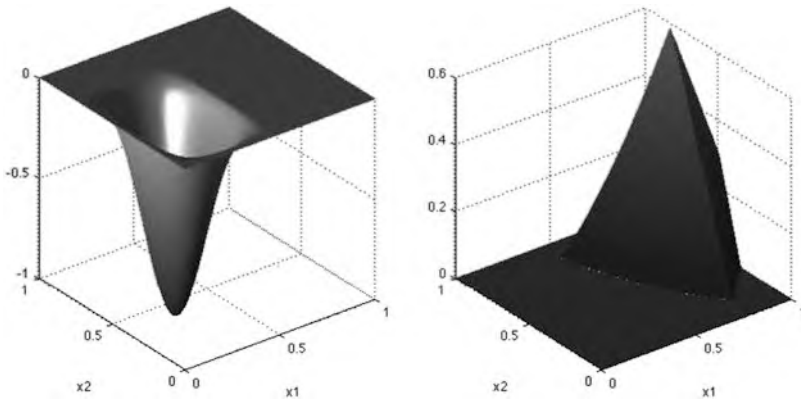

Fig. 7 Example 6.2. Optimal state y^* (*left*) and optimal multiplier σ^* (*right*)
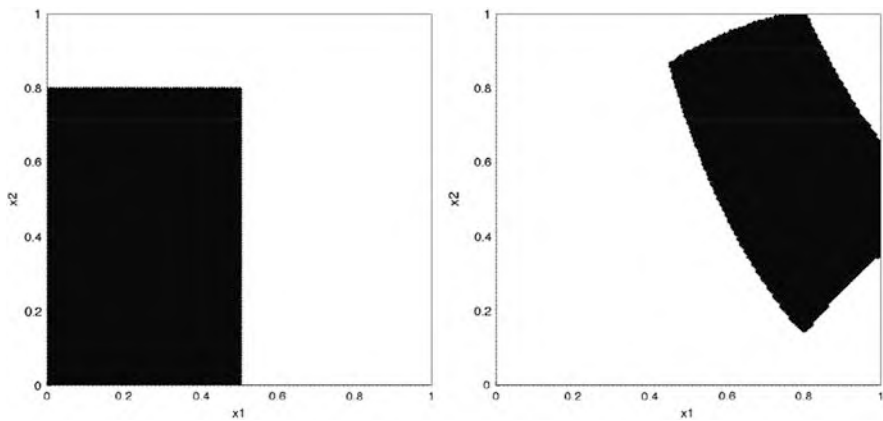
Fig. 8 Example 6.2. The inactive set $\mathcal{I}^*$ (*left*) and the strongly active set $\mathcal{C}^*$, both marked in *black* (*right*)
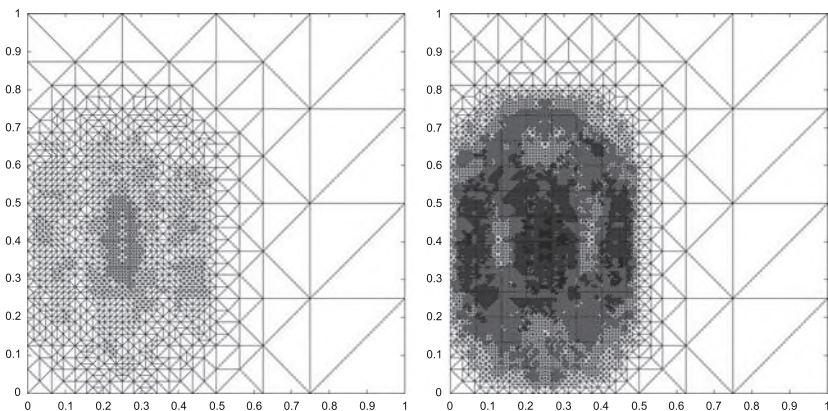

Fig. 9 Example 6.2. Mesh at refinement level $n = 7$ (*left*) and final mesh (*right*)

As in the first example, Fig. 10 displays the decrease of the errors in the state, in the control, in the total error, and in the estimator as functions of the DOFs on a logarithmic scale, whereas Table 2 contains the associated computed experimental convergence rates. Since the solution is smooth, uniform refinement is already optimal, i.e., in Table 2 we observe almost the same rates for uniform and adaptive refinement. However, as can be seen in Fig. 10, the error reductions are slightly less for adaptive refinement.
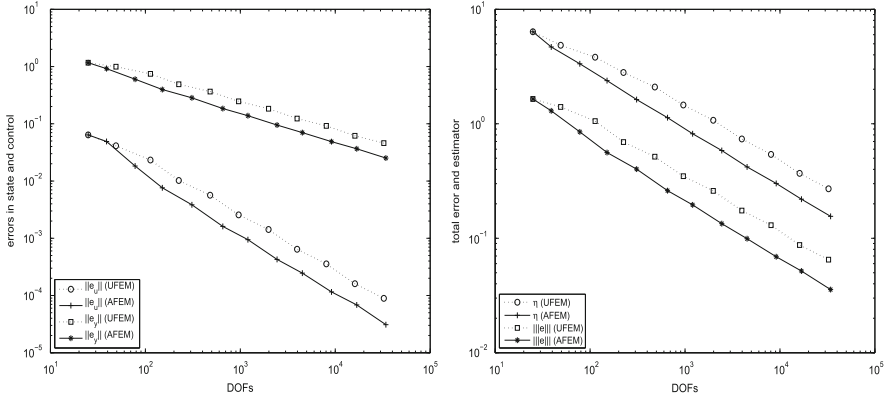
Fig. 10 Example 6.2. Convergence history: Decrease of the errors in the state $\|e_{h,y}\|_{1,\Omega}$ and in the control $\|e_{h,u}\|_{0,\Omega}$ as functions of the DOFs on a logarithmic scale (for uniform (UFEM) and adaptive (AFEM) refinement (*left*). Decrease of the estimator η_h and the total error $|||e_h|||$ as a function of the DOFs on a logarithmic scale (for uniform (UFEM) and adaptive (AFEM) refinement (*right*)

Table 2 Example 6.2: Experimental convergence rates (uniform and adaptive refinement)

n	$\|e_{h,y}\|_{1,\Omega}$		$\|e_{h,p}\|_{1,\Omega}$		$\|e_{h,u}\|_{0,\Omega}$		$|||e_h|||$	
	Unif.	Adapt.	Unif.	Adapt.	Unif.	Adapt.	Unif.	Adapt.
2	0.24	0.61	0.24	0.61	0.65	1.42	0.25	0.63
3	0.34	0.63	0.34	0.63	0.69	1.33	0.35	0.64
4	0.61	0.47	0.61	0.47	1.20	0.95	0.62	0.47
5	0.39	0.58	0.39	0.58	0.78	1.16	0.39	0.58
6	0.57	0.47	0.57	0.47	1.14	0.88	0.57	0.47
7	0.41	0.53	0.41	0.53	0.81	1.12	0.41	0.54
8	0.57	0.49	0.57	0.49	1.15	0.90	0.57	0.49
9	0.42	0.52	0.42	0.52	0.83	1.07	0.42	0.52
10	0.57	0.47	0.57	0.47	1.15	0.85	0.58	0.47
11	0.44	0.53	0.44	0.53	0.84	1.12	0.42	0.53

Figure 11 shows the decrease of the reliability related consistency error $e = |e_{h,rel}^c|$ (dotted line) and its estimates $e^k = |\bar{e}_{h,rel}^{c,(k)}|$, $E^k = \bar{E}_{h,rel}^{c,(k)}$, $1 \leq k \leq 3$, as functions of the DOFs on a logarithmic scale both for uniform refinement (left) and for adaptive refinement (right). We see a very similar behavior as in Example 1, i.e., for $1 \leq k \leq 2$, the quantities $e^k = |\bar{e}_{h,rel}^{c,(k)}|$ and $E^k = \bar{E}_{h,rel}^{c,(k)}$ provide close upper bounds, whereas $|\bar{e}_{h,rel}^{c,(3)}|$ underestimates and $\bar{E}_{h,rel}^{c,(3)}$ grossly overestimates the consistency error $|e_{h,rel}^c|$.
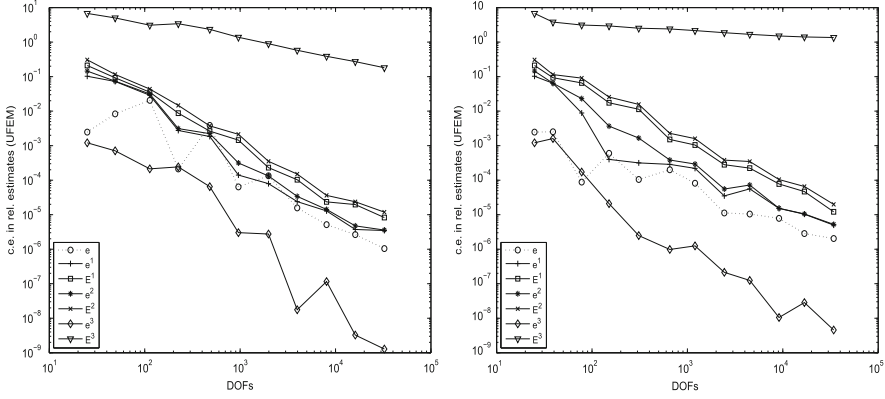
Fig. 11 Example 6.2. Decrease of the reliability related consistency error $e = |e^c_{h,rel}|$ (*dotted line*) and its estimates $e^k = |\bar{e}^{c,(k)}_{h,rel}|$, $E^k = \bar{E}^{c,(k)}_{h,rel}$, $1 \leq k \leq 3$, (*solid lines*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)
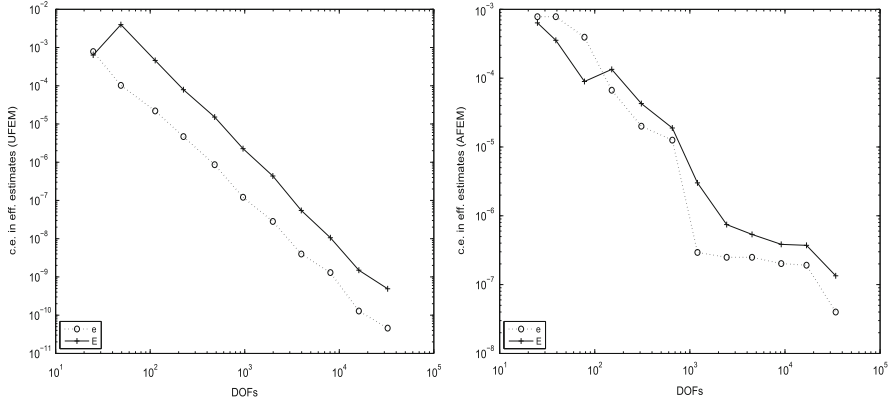

Fig. 12 Example 6.2. Decrease of the efficiency related consistency error $e = |e^c_{h,eff}|$ (*dotted line*) and its estimate $E = \bar{E}^{c,(1)}_{h,eff}$ (*solid line*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)

Figure 12 displays the decrease of the efficiency related consistency error $e^c_{h,eff}$ and its estimate $\bar{E}^{c,1}_{h,eff}$ as functions of the DOFs on a logarithmic scale for uniform refinement (left) and adaptive refinement (right). As in Example 1, after some pre-asymptotic phase in the adaptive regime, the estimates provide upper bounds of the consistency error.

Example 2 features the occurrence of a strongly active set $\mathcal{C}^*$ and hence, we are interested in how well the a posteriori quantities $e^{dva}_{h,\mathcal{A}}$ and $e^{dva}_{h,\mathcal{C}}$ coincide with $e^{eva}_{h,\mathcal{A}}$, $e^{evd}_{h,\mathcal{A}}$ and $e^{eva}_{h,\mathcal{C}}$, $e^{evd}_{h,\mathcal{C}}$, respectively. This is reflected in Figs. 13 and 14.
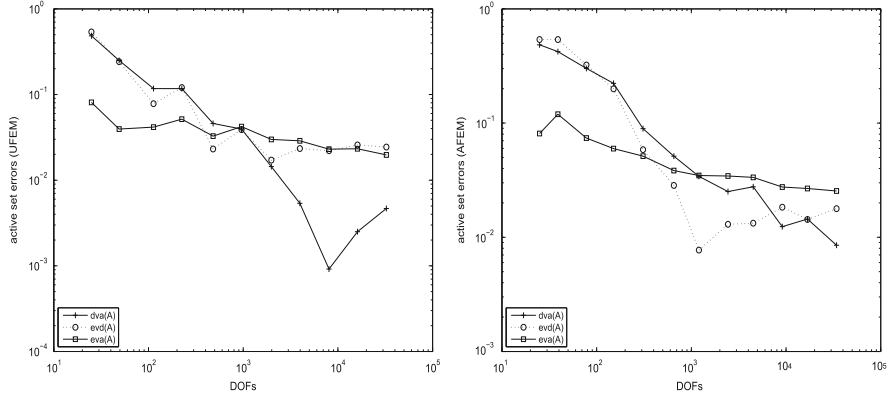
Fig. 13 Example 6.2. Approximation of the active set $\mathcal{A}$: quantities $e_{h,\mathcal{A}}^{evd}$ (*dotted line*) and $e_{h,\mathcal{A}}^{dva}, e_{h,\mathcal{A}}^{eva}$ (*solid lines*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)
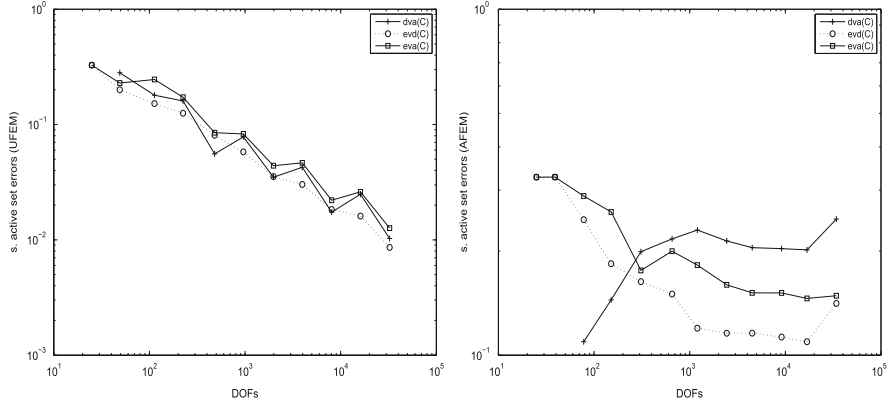

Fig. 14 Example 6.2. Approximation of the strongly active set $\mathcal{C}$: quantities $e_{h,\mathcal{C}}^{evd}$ (*dotted line*) and $e_{h,\mathcal{C}}^{dva}, e_{h,\mathcal{C}}^{eva}$ (*solid lines*) as functions of the DOFs on a logarithmic scale for uniform refinement (*left*) and adaptive refinement (*right*)

References

[1] R.A. Adams, J.J.F. Fournier, *Sobolev Spaces*, 2nd edn. (Academic, New York, 2003)

[2] M. Ainsworth, J.T. Oden, C.Y. Lee, Local a posteriori error estimators for variational inequalities. Numer. Math. Partial Differ. Equ. **9**, 23–33 (1993)

[3] I. Babuska, J. Whiteman, T. Strouboulis, *Finite Elements: An Introduction to the Method and Error Estimation*. (Oxford University Press, Oxford, 2011)

[4] V. Barbu, *Optimal Control of Variational Inequalities*. (Pitman, Boston/London/Melbourne, 1984)

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

[6] D. Braess, A posteriori error estimators for obstacle problems: another look. Numer. Math. **101**, 415–421 (2005)

[7] D. Braess, C. Carstensen, R.H.W. Hoppe, Convergence analysis of a conforming adaptive finite element method for an obstacle problem. Numer. Math. **107**, 455–471 (2007)

[8] D. Braess, C. Carstensen, R.H.W. Hoppe, Error reduction in adaptive finite element approximations of elliptic obstacle problems. J. Comp. Math. **27**, 148–169 (2009)

[9] S.C. Brenner, L.R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd edn. (Springer, New York, 2008)

[10] F. Facchinei, J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. (Springer, Berlin/Heidelberg/New York, 2003)

[11] A. Gaevskaya, Adaptive finite element methods for optimally controlled elliptic variational inequalities, Ph.D. thesis, Institute for Mathematics, University of Augsburg, 2013

[12] A. Gaevskaya, R.H.W. Hoppe, S. Repin, Functional approach to a posteriori error estimation for elliptic optimal control problems with distributed control. J. Math. Sci. **144**, 4535–4547 (2007)

[13] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*. (Pitman, Boston/London/Melbourne, 1985)

[14] A. Günther, M. Hinze, A posteriori error control of a state constrained elliptic control problem. J. Numer. Math. **16**, 307–322 (2008)

[15] M. Hintermüller, R.H.W. Hoppe, Goal-oriented adaptivity in control constrained optimal control of partial differential equations. SIAM J. Control Optim. **47**, 1721–1743 (2008)

[16] M. Hintermüller, R.H.W. Hoppe, Goal oriented mesh adaptivity for mixed control-state elliptic optimal control problems, in *Applied and Numerical Partial Differential Equations. Scientific Computing in Simulation, Optimization and Control in a Multidisciplinary Context*, ed. by W. Fitzgibbon, Y. Kuznetsov, P. Neittaanmäki, J. Périaux, O. Pironneau. Computational Methods in Applied Sciences, vol. 15. (Springer, Berlin/Heidelberg/New York, 2009)

[17] M. Hintermüller, R.H.W. Hoppe, Goal-oriented adaptivity in pointwise state constrained optimal control of partial differential equations. SIAM J. Control Optim. **48**, 5468–5487 (2010)

[18] M. Hintermüller, I. Kopacka, Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. SIAM J. Optim. **20**, 868–902 (2009)

[19] M. Hintermüller, R.H.W. Hoppe, Y. Iliash, M. Kieweg, An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. ESAIM Control Optim. Calc. Var. **14**, 540–560 (2008)

[20] M. Hintermüller, M. Hinze, R.H.W. Hoppe, Weak-duality based adaptive finite element methods for PDE-constrained optimization with pointwise gradient state-constraints. J. Comp. Math. **30**, 101–123 (2012)

[21] M. Hintermüller, R.H.W. Hoppe, C. Löbhard, Dual-weighted goal-oriented adaptive finite elements for optimal control of elliptic variational inequalities. ESAIM Control Optim. Calc. Var. **20**, 524–546 (2014)

[22] R.H.W. Hoppe, R. Kornhuber, Adaptive multilevel methods for obstacle problems. SIAM J. Numer. Anal. **31**, 301–323 (1994)

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

[24] C. Johnson, Adaptive finite element methods for the obstacle problem. Math. Models Methods Appl. Sci. **2**, 483–487 (1992)

[25] D. Kinderlehrer, G. Stampacchia, *An Introduction to Variational Inequalities and Its Applications*. (SIAM, Philadelphia, 2000)

[26] D. Klatte, B. Kummer, *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods, and Applications*. (Kluwer, Dordrecht, 2002)

[27] I. Kopacka, Mpecs/mpccs in functional space: first order optimality concepts, path-following and multilevel algorithms, Ph.D. thesis, Institute of Applied Mathematics, Karl-Franzens University at Graz, 2009

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

[29] Z.Q. Luo, J.S. Pang, D. Ralph, *Mathematical Programs with Equilibrium Constraints*. (Cambridge University Press, Cambridge, 1996)

[30] F. Mignot, Contrôle dans les inèquations variationelles elliptiques. J. Funct. Anal. **22**, 130–185 (1976)

[31] F. Mignot, J.P. Puel, Optimal control in some variational inequalities. SIAM J. Control Optim. **22**, 466–476 (1984)

[32] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I: Basic Theory*. (Springer, Berlin/Heidelberg/New York, 2006)

[33] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation II: Applications*. (Springer, Berlin/Heidelberg/New York, 2006)

[34] P. Neittaanmäki, J. Sprekels, D. Tiba, *Optimization of Elliptic Systems: Theory and Applications*. (Springer, Berlin/Heidelberg/New York, 2006)

[35] R. Nochetto, K.G. Siebert, A. Veeser, Pointwise a posteriori error control for elliptic obstacle problems. Numer. Math. **95**, 163–195 (2003)

[36] J. Outrata, M. Kocvara, J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. (Kluwer, Dordrecht, 1998)

[37] J. Outrata, J. Zowe, A numerical approach to optimization problems with variational inquality constraints. Math. Program. **68**, 105–130 (1995)

[38] H. Scheel, S. Scholtes, Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. Math. Oper. Res. **25**, 1–22 (2000)

[39] L.R. Scott, S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. **54**, 483–493 (1990)

[40] K.G. Siebert, A. Veeser, A unilaterally constrained quadratic minimization with adaptive finite elements. SIAM J. Optim. **18**, 260–289 (2007)

[41] R.S. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*. (World Scientific, River Edge, 2003)

[42] F.T. Suttmeier, On a direct approach to adaptive FE-discretizations for elliptic variational inequalities. J. Numer. Math. **13**, 73–80 (2005)

[43] A. Veeser, Efficient and reliable a posteriori error estimators for elliptic obstacle problems. SIAM J. Numer. Anal. **39**, 146–167 (2001)

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

[45] B. Vexler, W. Wollner, Adaptive finite elements for elliptic optimization problems with control constraints. SIAM J. Optim. **47**, 509–534 (2008)