A dual-weighted residual approach to goal-oriented adaptivity for optimal control of elliptic variational inequalities

M. Hintermüller  R.H.W. Hoppe
C. Loebhard

SFB-Report No. 2012-023  November 2012
SFB sponsors:

- Austrian Science Fund (FWF)
- University of Graz
- Graz University of Technology
- Medical University of Graz
- Government of Styria
- City of Graz
A DUAL-WEIGHTED RESIDUAL APPROACH TO
GOAL-ORIENTED ADAPTIVITY FOR OPTIMAL CONTROL OF
ELLIPTIC VARIATIONAL INEQUALITIES

M. HINTERMÜLLER, R.H.W. HOPPE, AND C. LÖBHARD

Abstract. A dual-weighted residual approach for goal-oriented adaptive finite elements for a class of optimal control problems for elliptic variational inequalities is studied. The development is based on the concept of C-stationarity. The overall error representation depends on primal residuals weighted by approximate dual quantities and vice versa as well as various complementarity mismatch errors. Also, a priori bounds for C-stationary points and associated multipliers are derived. Details on the numerical realization of the adaptive concept are provided and a report on numerical tests including the critical cases of biactivity are presented.

1. Introduction

In this paper we study goal-oriented adaptive finite element methods for a class of optimal control problems with variational inequality constraints. The latter problem class falls into the broader category of mathematical programs with equilibrium constraints, or MPECs for short, which, due to constraint degeneracy, cannot be treated by qualified Karush-Kuhn-Tucker-type theory for the derivation of stationarity systems; see [LPR96, OKZ98] in finite dimensions and [Bar83, Mor06a, Mor06b, NST06] in function space. The associated difficulties result from so-called biactivities in a regular setting, i.e. a situation where certain multipliers admit a pointwise interpretation, or – in the terminology of optimal control – the fact that the control-to-state mapping is in general only directionally differentiable. The latter fact renders the problem under investigation non-convex and non-differentiable. This structural property is responsible not only for the aforementioned theoretical challenges in describing stationarity, but it also causes problems in the design of reliable solution algorithms; see, e.g., the discussion in

2010 Mathematics Subject Classification. 49M25, 65K15, 90C33.

Key words and phrases. adaptive finite element method, C-stationarity, goal-oriented error estimation, mathematical programming with equilibrium constraints, optimal control of variational inequalities.

M.H. acknowledges support by the German Research Fund (DFG) through the Research Center MATHEON Project C28 and C31 and the SPP 1253 "Optimization with Partial Differential Equations", the Austrian Science Fund (FWF) through the START Project Y 305-N18 "Interfaces and Free Boundaries" and the SFB Project F32 04-N18 "Mathematical Optimization and Its Applications in Biomedical Sciences", and support through a J. Tinsley Oden Fellowship at the Institute for Computational Engineering and Sciences (ICES) at UT Austin, Texas, USA. R.H. has been supported by NSF-DMS 1115658, by the DFG-SPP 1253 and DFG-SPP 1506, by the BMBF projects "FROPT" and "MeFreSim", and by the ESF within the Research Networking Programme "OPTPDE". C.L. was supported by the DFG through SPP 1253 "Optimization with Partial Differential Equations".

1
in finite dimensions and [HIK02, HK09, HK11] in function space.

As our underlying problem class is originally posed in function space, the need for discretization for its numerical realization arises naturally. In this context one is interested in obtaining a specified accuracy in the approximate solution or a pre-defined target quantity by the least possible computational effort. Besides the development of efficient solvers, this concerns in particular the design of adaptive finite element methods for the discretization of the continuous problem or stationarity system. Based on a suitable error estimator or indicator, the latter technique aims at refining the discretization locally only in regions with large errors while keeping elements coarse wherever possible. This procedure ideally minimizes the degrees of freedom for computing an approximate solution to the original problem in order to achieve a user-specified accuracy.

Over the years different approaches to adaptive finite element methods (AFEMs) for the numerical treatment of partial differential equations have been studied. Here we only refer to the monographs [Ver96, BR03, Rep08, BWS11] and the many references therein. By now, a high level of sophistication has been reached. In contrast to this development, AFEM for elliptic variational inequalities, though well investigated, has not yet reached a comparable level; see [Joh92, Vee01, NSV03, Bra05, BCH07, BCH09, HHT11] and the references therein. This is due to the presence of the coincidence or active set with respect to the solution of the variational inequality, which, in general, causes a non-smooth dependence of the solution on the data and, thus, complicates the numerical treatment significantly. Recently, AFEM was successfully carried over to control and/or state constrained optimal control problems (not MPECs!); compare [LY00, VW08, HH08, HH10a, HH10b]. In particular, in [VW08, HH08, HH10a, HH10b] the dual-weighted residual (DWR) approach to goal-oriented adaptivity, which was pioneered by [BKR00] (see also the monograph [BR03]), was extended to the discretization of control and/or state constrained optimal control problems with partial differential equation constraints.

In the present paper, we develop a goal-oriented mesh refinement technique in the spirit of the DWR approach for the adaptive discretization of MPECs in function space. Based on the notion of a modified Lagrangian function associated with the MPEC or more precisely mathematical program with complementarity constraints (MPCC, for short), the so-called MPCC-Lagrangian, an error estimator (indicator) is derived which involves primal residuals weighted by dual variables and vice versa as well as error terms covering the mismatch in complementarity. The latter error indicators are relevant in the location of the coincidence sets which arise due to the variational inequality constraint. Finally, we point out that the literature on goal-oriented mesh adaptivity for MPCCs or MPECs in function space, and in particular for the problem class under investigation, appears essentially void.

The rest of this text is organized as follows. In Section 2 we outline the problem class under consideration in the original continuous form and also in its discrete version, and provide basic assertions as optimality systems and a priori estimates. Section 3 gives an error representation, which turns out to be a combination of weighted residuals and products reflecting the mismatch in complementarity. Although this representation is not an estimator, i.e., it depends on the (unknown) solutions of the continuous problem, it guides a way to an a posteriori error indicator. In Section 4 we derive fully a posteriori indicators for the error contribution on
every element of the discretization. The last Section provides the basic algorithms
used in our numerical solver and the numerical results in some examples.

Notation. Throughout this text, let $\Omega \subset \mathbb{R}^n$ be a bounded domain with $\partial \Omega$
denoting its boundary. We write $(z, w)_{\Omega}$ for the canonical scalar product in the
Lebesgue space $L^2(\Omega)$, and $\|z\|_{0, \Omega}$ for the induced norm. For $n \in \{1, 2, 3\}$, $H^1(\Omega)$ is
the Sobolev space of $L^2(\Omega)$-functions with gradients in $L^2(\Omega)^n$ and $H^1_0(\Omega) := \{y \in H^1(\Omega) | y|_{\partial \Omega} = 0\}$ endowed with the norm $\|z\|_{1, \Omega} := \|\nabla z\|_{0, \Omega}$.

For a function $z \in H^1(\Omega)$ and a set $\omega \subset \Omega$, we say that $z|_{\omega} \geq 0$ in the sense
of $H^1(\Omega)$, if there exists a sequence $(z_n)_{n \in \mathbb{N}} \subset W^{1, \infty}(\Omega)$ such that for all $x \in \omega$,
$z_n(x) \geq 0$ and $z_n \to z$ in $H^1(\Omega)$. With this definition, the relations “≤” and
“=” on subsets of $\Omega$ can be defined canonically on $H^1(\Omega)$ (cf. [KS80, p. 35]).
Furthermore, for a function $z \in H^1(\Omega)$ and a point $x \in \Omega$ we write $z(x) > 0$ in the
sense of $H^1(\Omega)$, if there exists a neighbourhood $U(x) \subset \Omega$ of $x$ and a function
$\varphi \in W^{1, \infty}(U(x))$ with $\varphi \geq 0$ and $\varphi(x) > 0$, such that $z|_{U(x)} - \varphi \geq 0$ in the sense
of $H^1(U(x))$. We abbreviate, for example, $\{z \geq 0\} := \{x \in \Omega | z(x) \geq 0\}$ and, for
functions $z \in L^2(\Omega)$, the above set is defined only up to a set of measure zero.

The dual space of $H^1_0(\Omega)$ is denoted by $H^{-1}(\Omega)$, and for $\lambda \in H^{-1}(\Omega)$ and
$z \in H^1_0(\Omega)$ the dual pairing is written as $\langle \lambda, z \rangle_{-1, 1, \Omega} := \lambda(z)$. For functionals
$z \in H^{-1}(\Omega)$ and a subset $\omega \subset \Omega$, we write $z|_{\omega} = 0$, if $\forall \phi \in H^1_0(\Omega)$ with $\phi|_{\Omega \setminus \omega} = 0$,
it holds that $\langle z, \phi \rangle_{-1, 1, \Omega} = 0$, and we write $z|_{\omega} \geq 0$, if $\forall \phi \in H^1_0(\Omega)$ with $\phi|_{\Omega \setminus \omega} = 0$
and $\phi|_{\Omega \setminus \omega} \geq 0$, it holds that $\langle z, \phi \rangle_{-1, 1, \Omega} \geq 0$.

2. Statement of the Problem and Stationarity

2.1. Problem definition. We consider the optimal control problem with a variational
inequality constraint given by

\begin{align}
(1a) & \quad \text{Minimize } J(y, u) := \frac{1}{2} \|y - y_d\|^2_{0, \Omega} + \frac{\nu}{2} \|u\|^2_{0, \Omega} \\
(1b) & \quad \text{over } (y, u) \in H^1_0(\Omega) \times L^2(\Omega), \\
(1c) & \quad \text{such that } y \geq 0, \\
(1d) & \quad \forall z \in H^1_0(\Omega), \quad z \geq 0, \quad a(y, z - y) \geq (u + f, z - y)_{0, \Omega},
\end{align}

i.e. we minimize an objective functional $J(y, u)$ depending on a state variable $y \in H^1_0(\Omega)$
with homogeneous Dirichlet boundary conditions and a control variable $u \in L^2(\Omega)$, both defined on a bounded domain $\Omega \subset \mathbb{R}^n$, $n \in \{1, 2, 3\}$. The given
function $y_d \in L^2(\Omega)$ is the desired state and $\nu > 0$ is the cost of the control action.

The constraint involves on the one hand an obstacle $\psi \equiv 0$ which bounds the
state $y$ from below, i.e., $y \geq 0$, and on the other hand a bounded, $H^1_0(\Omega)$-elliptic bilinear
bilinear form $a : H^1_0(\Omega) \times H^1_0(\Omega) \to \mathbb{R}$ which describes the influence of forces, such
as the fixed volume force $f \in L^2(\Omega)$ and the action of the control $u$, applied to the
state $y$.

The existence of a solution to problem \((1)\) can be established by arguments based
on infimizing sequences; see, e.g., [Bar81, MPS8].

Remark 2.1. We note that considering $\psi \equiv 0$ does not imply restrictions with
respect to the choice of the obstacle aside from $H^1_0(\Omega)$-regularity.

In the formulation of the first order optimality system and in the numerical solver
for the optimal control problem, a reformulation of the variational inequality \((1d)\)
by means of an additional slack variable $\xi$ is considered. For this purpose, let $y \in H^1_0(\Omega), y \geq 0$ and $u \in L^2(\Omega)$, and introduce the operator $A : H^1_0(\Omega) \to H^{-1}(\Omega)$ which satisfies $a(y, z) = \langle Ay, z \rangle_{-1,1,\Omega}$ for all $y, z \in H^1_0(\Omega)$ as well as a slack variable $\xi \in H^{-1}(\Omega)$ as follows,

$$\xi := Ay - u - f \in H^{-1}(\Omega).$$

Then, (1c)-(1d) is equivalent to

$$y \geq 0, \quad \xi \geq 0 \quad \text{and} \quad \langle \xi, y \rangle_{-1,1,\Omega} = 0.$$

Thus, for given data $f, y_d \in L^2(\Omega)$ on a domain $\Omega$ and $\nu > 0$, problem (1) can be rewritten as an optimal control problem with partial differential and complementarity constraints as follows: Find a solution $(y, u, \xi) \in H^1_0(\Omega) \times L^2(\Omega) \times H^{-1}(\Omega)$ to the problem

\begin{align}
\text{(2a)} & \quad \text{Minimize} \quad J(y, u) = \frac{1}{2} \|y - y_d\|_{0,\Omega}^2 + \frac{\nu}{2} \|u\|_{0,\Omega}^2 \\
\text{(2b)} & \quad \text{such that} \quad a(y, z) = (u + f, z)_{0,\Omega} + \langle \xi, z \rangle_{-1,1,\Omega}, \quad \forall z \in H^1_0(\Omega), \\
\text{(2c)} & \quad y \geq 0 \text{ in } H^1_0(\Omega), \quad \xi \geq 0 \text{ in } H^{-1}(\Omega) \text{ and } \langle \xi, y \rangle_{-1,1,\Omega} = 0.
\end{align}

Due to the complementarity structure of (2c) we call (2) a mathematical program with complementarity constraints, or an MPCC, for short.

**Remark 2.2** (Regularity of the data). In the situation of [Rod87, Thm. 5.2.2] we have $\|Ay^*\|_{0,\Omega} \leq C_R(\Omega)(\|f\|_{0,\Omega} + \|u^*\|_{0,\Omega})$ for the solution $y^*$ of the variational inequality in (1d) with control variable $u^* \in L^2(\Omega)$. Thus, if for instance $A = -\Delta$ on a convex domain $\Omega$ (or a domain with boundary $\delta\Omega$ of class $C^{1,1}$), then $y^*$ satisfies $y^* \in H^1_0(\Omega) \cap H^2(\Omega)$ (cf. [Rod87, Cor. 5.2.3]). Moreover, in the reformulation resulting in the linear complementarity problem (2b–2c) it holds that $\xi^* \in L^2(\Omega)$.

### 2.2. Stationarity conditions

In the error analysis in Section 3 we use stationarity conditions which are weaker than those stated in [MPS84], but yet much stronger than those obtained in [Bar84]. While one can guarantee strong stationarity for a solution of (2) according to [MPS84], here we only use the weaker notion of C(larke)-stationarity of the solution as only C-stationary points can be obtained with guarantee by current state-of-the-art solution algorithms. Obviously, the available stationarity condition is also essential for the development of a posteriori error estimators. For a more general definition of finite dimensional variants of stationarity concepts we refer to [SS00]. Function space versions of C-stationarity conditions in the regular case of Remark 2.2 which are more restrictive than the versions stated here can be found in [HK09].

**Definition 2.3** (Clarke or C-stationarity). A point $(y, u, \xi) \in H^1_0(\Omega) \times L^2(\Omega) \times H^{-1}(\Omega)$ is called C-stationary for problem (2), if there exist $p \in H^1_0(\Omega), \lambda \in H^{-1}(\Omega)$
and \( \mu \in H^1_0(\Omega) \) such that the following conditions hold,

\[
\begin{align*}
(3a) \quad \mathcal{A}y - u - \xi &= f \quad \text{in } H^{-1}(\Omega), \\
(3b) \quad \xi &\geq 0, \ y \geq 0, \langle \xi, y \rangle_{-1,1,\Omega} = 0, \\
(3c) \quad y - \lambda + \mathcal{A}^*p &= y_d \quad \text{in } H^{-1}(\Omega), \\
(3d) \quad \nu u - p &= 0 \quad \text{in } H^1_0(\Omega), \\
(3e) \quad -\mu - p &= 0 \quad \text{in } H^1_0(\Omega), \\
(3f) \quad \langle \lambda, y \rangle_{-1,1,\Omega} &= 0, \\
(3g) \quad \langle \xi, \mu \rangle_{-1,1,\Omega} &= 0, \\
(3h) \quad \langle \lambda, \mu \rangle_{-1,1,\Omega} &\geq 0.
\end{align*}
\]

**Definition 2.4** (Strong stationarity). A point \((y, u, \xi) \in H^1_0(\Omega) \times L^2(\Omega) \times H^{-1}(\Omega)\) is called *strongly stationary for problem (2)*, if it is C-stationary and if the multipliers \(\lambda \in H^{-1}(\Omega)\) and \(\mu \in H^1_0(\Omega)\) in Definition 2.3 satisfy the sign conditions

\[
\begin{align*}
(4a) \quad \forall \psi \in H^1_0(\Omega), \ \psi|_{y=0} &\geq 0 \quad \text{and} \quad \langle \xi, \psi \rangle_{-1,1,\Omega} = 0 : \ \langle \lambda, \psi \rangle_{-1,1,\Omega} \geq 0, \\
(4b) \quad \mu|_{y=0} &\geq 0.
\end{align*}
\]

We introduce \(X\) as the function space \(X := H^1_0(\Omega) \times L^2(\Omega) \times H^{-1}(\Omega) \times H^1_0(\Omega)\) with elements \(x = (y, u, \xi, p)\), and define the MPCC-Lagrangian \(\mathcal{L} : X \times H^{-1}(\Omega) \times H^1_0(\Omega) \to \mathbb{R}\) according to

\[
\begin{align*}
\mathcal{L}(y, u, \xi, p, \lambda, \mu) := J(y, u) + a(y, p) - (u + f, p)_{0,\Omega} - \langle \xi, p \rangle_{-1,1,\Omega} \\
- \langle \lambda, y \rangle_{-1,1,\Omega} - \langle \xi, \mu \rangle_{-1,1,\Omega}.
\end{align*}
\]

**Remark 2.5.** We note that if \((y^*, u^*, \xi^*)\) is C-stationary and \(p^*, \lambda^*, \mu^*\) are the associated dual variables from (3), then

\[
\mathcal{L}(x^*, \lambda^*, \mu^*) = J(y^*, u^*),
\]

and for all \(\delta x \in X\) it holds that

\[
\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*)(\delta x) = 0.
\]

**Remark 2.6.** The functional \(\mathcal{L}(\cdot, \lambda^*, \mu^*) : X \to \mathbb{R}\) is infinitely often Gâteaux differentiable at \(x \in X\). The second Gâteaux derivative \(\nabla_{xx} \mathcal{L}(x, \lambda^*, \mu^*)\) is independent of the argument \((x, \lambda^*, \mu^*)\) which for notational convenience will be omitted in the sequel.

### 2.3 A priori estimates for solutions of the optimality system.

This section is devoted to an a priori analysis of the dependence of a C-stationary point \((y^*, u^*, \xi^*)\) with dual variables \(p^*, \lambda^*\) and \(\mu^*\) on the problem input data \(\Omega, f, y_d,\) and \(\nu\). The following result states a priori bounds which are, for instance, relevant in the study of stability of solutions with respect to data perturbations.

For its formulation we invoke the following assumption concerning the data involved in (1). The coercive bilinear form \(a : H^1_0(\Omega) \times H^1_0(\Omega) \to \mathbb{R}\) admits the representation \(\mathcal{A} \in \mathcal{L}(H^1_0(\Omega), H^{-1}(\Omega))\) with coercivity constant \(\Sigma_\mathcal{A} > 0\) satisfying

\[
a(y, z) = \langle \mathcal{A}y, z \rangle_{-1,1,\Omega} \quad \text{and} \quad a(y, y) \geq \Sigma_\mathcal{A} \| y \|^2_{1,\Omega}
\]
for all \( y, z \in H^1_0(\Omega) \). Let \( C_A(\Omega) = \| A \|_{C(\Omega)} \) abbreviate the norm of the operator \( A \), and let \( A^* \) denote the adjoint of \( A \). Moreover, we denote the Friedrichs constant for the domain \( \Omega \) by \( C_F(\Omega) \) and the constant arising from the embedding \( L^2(\Omega) \hookrightarrow H^{-1}(\Omega) \) by \( C_{-1}(\Omega) \), i.e., for all \( y \in H^1_0(\Omega) \) and \( u \in L^2(\Omega) \), it holds that
\[
\| y \|_{0,\Omega} \leq C_F(\Omega) \| y \|_{1,\Omega} \quad \text{and} \quad \| u \|_{-1,\Omega} \leq C_{-1}(\Omega) \| u \|_{0,\Omega}.
\]

**Proposition 2.7.** A solution \((x^*, \lambda^*, \mu^*)\) of the stationarity system \((3)\) satisfies the following a priori bounds,
\[
(6) \quad \| y^* \|_{0,\Omega} \leq \| y_d \|_{0,\Omega} + \sqrt{\nu} \| f \|_{0,\Omega},
\]
\[
(7) \quad \| y^* \|_{1,\Omega} \leq \frac{C_F(\Omega)}{\Sigma_A} \left( \frac{1}{\sqrt{\nu}} \| y_d \|_{0,\Omega} + 2 \| f \|_{0,\Omega} \right),
\]
\[
(8) \quad \| u^* \|_{0,\Omega} \leq \frac{1}{\sqrt{\nu}} \| y_d \|_{0,\Omega} + \| f \|_{0,\Omega},
\]
\[
(9) \quad \| \xi^* \|_{-1,\Omega} \leq \left( \frac{C_F(\Omega)}{\Sigma_A} C_A(\Omega) + C_{-1}(\Omega) \right) \left( \frac{1}{\sqrt{\nu}} \| y_d \|_{0,\Omega} + 2 \| f \|_{0,\Omega} \right),
\]
\[
(10) \quad \| u^* \|_{1,\Omega} \leq \frac{C_F(\Omega)}{\Sigma_A} \left( \| y_d \|_{0,\Omega} + \sqrt{\nu} \| f \|_{0,\Omega} \right),
\]
\[
(11) \quad \| \lambda^* \|_{-1,\Omega} \leq \frac{1}{\sqrt{\nu}} \| y_d \|_{0,\Omega} + \sqrt{\nu} \| f \|_{0,\Omega}.
\]

Under the additional assumptions stated in Remark \((2.2)\) we obtain the estimate
\[
(12) \quad \| Ay^* \|_{0,\Omega} \leq C_R(\Omega) \left( \frac{1}{\sqrt{\nu}} \| y_d \|_{0,\Omega} + 2 \| f \|_{0,\Omega} \right).
\]

**Proof.** We start by multiplying the equations \((3a)\) and \((3c)\) by \( p^* \) and \( y^* \), respectively. Then, using the constraints \((3d)\)-\((3g)\), we obtain
\[
(13) \quad a(y^*, p^*) = (u^*, p^*)_{0,\Omega} + (f, p^*)_{0,\Omega} = \nu \| u^* \|_{0,\Omega}^2 + \nu (f, u^*)_{0,\Omega},
\]
\[
(14) \quad a(y^*, p^*) = (y_d, y^*)_{0,\Omega} - \| y^* \|_{0,\Omega}^2.
\]

Subtracting \((13)\) from \((14)\), rearranging terms and estimating the inner products one derives
\[
\| y^* \|_{0,\Omega}^2 + \nu \| u^* \|_{0,\Omega}^2 = (y_d, y^*)_{0,\Omega} - \nu (f, u^*)_{0,\Omega} \leq \frac{1}{2} \| y^* \|_{0,\Omega}^2 + \frac{1}{2} \| y_d \|_{0,\Omega}^2 + \nu \| u^* \|_{0,\Omega}^2 + \frac{\nu}{2} \| f \|_{0,\Omega}^2,
\]
and thus
\[
(15) \quad \| y^* \|_{0,\Omega}^2 + \nu \| u^* \|_{0,\Omega}^2 \leq \| y_d \|_{0,\Omega}^2 + \nu \| f \|_{0,\Omega}^2.
\]

This yields the \( L^2 \)-bounds for \( y^* \) and \( u^* \). Next, we multiply equation \((3a)\) by \( y^* \), recall that \((\xi^*, y^*)_{-1,1,\Omega} = 0 \) and deduce
\[
\| y^* \|_{1,\Omega}^2 \leq \frac{1}{\Sigma_A} \left( (u^*, y^*)_{0,\Omega} + (f, y^*)_{0,\Omega} \right) \leq \frac{C_F(\Omega)}{\Sigma_A} \left( \| u^* \|_{0,\Omega} + \| f \|_{0,\Omega} \right) \| y^* \|_{1,\Omega}.
\]

This, together with \((15)\), implies the bound on \( \| y^* \|_{1,\Omega} \), i.e.,
\[
\| y^* \|_{1,\Omega} \leq \frac{C_F(\Omega)}{\Sigma_A} \left( \| u^* \|_{0,\Omega} + \| f \|_{0,\Omega} \right) \leq \frac{C_F(\Omega)}{\Sigma_A} \left( \frac{1}{\sqrt{\nu}} \| y_d \|_{0,\Omega} + 2 \| f \|_{0,\Omega} \right).
\]
The bound on $\|\xi^*\|_{-1,\Omega}$ involves the constant $C_{-1}(\Omega)$ arising from the embedding $L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ and the norm $C_A(\Omega) = \|A\|_{L(H_0^1(\Omega),H^{-1}(\Omega))}$. In fact, from (3a), we obtain
\[
\|\xi^*\|_{-1,\Omega} = \|Ay^* - u^* - f\|_{-1,\Omega} \\
\leq C_A(\Omega) \|y^*\|_{1,\Omega} + C_{-1}(\Omega) \left(\|u^*\|_{0,\Omega} + \|f\|_{0,\Omega}\right) \\
\leq \left(\frac{C_F(\Omega)}{\Sigma_A} C_A(\Omega) + C_{-1}(\Omega)\right) \left(\frac{1}{\sqrt{\nu}} \|y_d\|_{0,\Omega} + 2 \|f\|_{0,\Omega}\right).
\]

The adjoint equation (3c) tested with $p^*$ and the sign condition (3b) yielding
\[
(\lambda^*,p^*)_{-1,\Omega} = -\langle \lambda^*,\mu^* \rangle_{-1,\Omega} \leq 0
\]
imply
\[
\|p^*\|^2_{1,\Omega} \leq \frac{1}{\Sigma_A} \left(\|y_d,\mu^*\|_{0,\Omega} - \langle y^*,p^* \rangle_{0,\Omega}\right),
\]
and in the same way as above,
\[
\nu \|u^*\|_{1,\Omega} = \|p^*\|_{1,\Omega} \leq \frac{C_F(\Omega)}{\Sigma_A} \left(2 \|y_d\|_{0,\Omega} + \sqrt{\nu} \|f\|_{0,\Omega}\right).
\]

The bound on $\|\lambda^*\|_{-1,\Omega}$ is derived analogously to the respective bound on $\xi^*$, i.e., one uses the adjoint equation to obtain
\[
\|\lambda^*\|_{-1,\Omega} = \|A^*p^* + y^* - y_d\|_{-1,\Omega} \\
\leq C_A(\Omega) \|p^*\|_{1,\Omega} + \|y^*\|_{-1,\Omega} + \|y_d\|_{-1,\Omega} \\
\leq C_A(\Omega) \|p^*\|_{1,\Omega} + C_{-1}(\Omega) \left(\|y^*\|_{0,\Omega} + \|y_d\|_{0,\Omega}\right) \\
\leq \left(\frac{C_A(\Omega)}{\Sigma_A} C_F(\Omega) + C_{-1}(\Omega)\right) \left(\sqrt{\nu} \|f\|_{0,\Omega} + 2 \|y_d\|_{0,\Omega}\right).
\]

This proves the first part of the proposition.

In the setting of Remark 2.2, we obtain
\[
\|Ay^*\|_{0,\Omega} \leq C_R(\Omega) \left(\|f\|_{0,\Omega} + \|u^*\|_{0,\Omega}\right) \leq C_R(\Omega) \left(2 \|f\|_{0,\Omega} + \frac{1}{\sqrt{\nu}} \|y_d\|_{0,\Omega}\right).
\]

This concludes the proof.

2.4. Discretization of the optimal control problem. We consider the discretization of the optimal control problem (1) by finite dimensional spaces $Y_h \subset H_0^1(\Omega)$ and $U_h \subset H_0^1(\Omega)$ for the state $y$ and the control $u$, respectively. Let $N_{\chi_h}$ and $N_{\psi_h}$ be sets with cardinality $|N_{\chi_h}| = \dim(Y_h)$ and $|N_{\psi_h}| = \dim(U_h)$. Further, we introduce $\{\varphi_z | z \in N_{\chi_h}\}$ and $\{\psi_z | z \in N_{\psi_h}\}$ as bases of $Y_h$ and $U_h$, respectively. Then, a function $y_h = \sum_{z \in N_{\chi_h}} y_{h,z} \varphi_z \in Y_h$ with $y_{h,z} \in \mathbb{R}$ can be identified by its component vector
\[
y_h = (y_{h,z})_{z \in N_{\chi_h}} \in \mathbb{R}^{|N_{\chi_h}|}.
\]
Similarly, we identify the function $u_h = \sum_{z \in N_{\psi_h}} u_{h,z} \psi_z \in U_h$ with
\[
u_h = (u_{h,z})_{z \in N_{\psi_h}}.
\]
For the ease of presentation, we assume that the basis of $Y_h$ yields $y_h \geq 0$ in $\Omega$ if and only if $y_{h,z} \geq 0$. A particular example of a discretization satisfying the above conditions is given in Section 4.1.
Next we define the stiffness matrix $A_h \in \mathbb{R}^{|N_{Y_h}| \times |N_{Y_h}|}$ and the mass matrix $M_h \in \mathbb{R}^{|N_{Y_h}| \times |N_{Y_h}|}$ representing the bilinear form $a$ on $Y_h$ and the $L^2$-scalar-product of functions in $Y_h$ with functions in $U_h$, respectively. Hence, the elements of these matrices are given by

$$(A_h)_{z \bar{z}} := a(\varphi_z, \varphi_{\bar{z}}) \quad \text{and} \quad (M_h)_{z \bar{v}} := (\varphi_z, \psi_{\bar{v}})_{0, \Omega}$$

for $z, \bar{z} \in N_{Y_h}$ and $v \in N_{U_h}$. Then, for all $y_h, w_h \in Y_h$ with vector representation $y_h, w_h$ and all $u_h \in U_h$ with vector representation $u_h$, it holds that

$$y_h^T A_h w_h = a(y_h, w_h) \quad \text{and} \quad y_h^T M_h u_h = (y_h, u_h)_{0, \Omega}.$$ 

Analogously, we define mass matrices $M^l_h$ and $M^f_h$ representing the inner products in $Y_h$ and $U_h$, respectively.

The data vectors $f_h = (f_h, z)_{z \in N_{Y_h}}$ and $y_{dh} = (y_{dh}, z)_{z \in N_{Y_h}}$ are defined by

$$f_h, z = \int_{\Omega} f \varphi_z \, dx, \quad y_{dh}, z = \int_{\Omega} y_d \varphi_z \, dx.$$ 

In the same way as in Section 2, we introduce a slack variable $\xi_h \in \mathbb{R}^{|N_{Y_h}|}$ as the vector

$$\xi_h := A_h y_h - M_h u_h - f_h.$$ 

The error representation in Section 3 requires an interpretation of the vector $\xi_h$ as an element $\xi_h$ in $H^{-1}(\Omega)$, such that for all $y_h \in Y_h \subset H^1_0(\Omega)$, we have

$$\int_{\Omega} \xi_h^T \xi_h = (\xi_h, y_h)_{-1,1,\Omega}.$$ 

Thus, for any $\xi_h \in \mathbb{R}^{|N_{Y_h}|}$ let $\xi_h := (M^\gamma_h)^{-1} \xi_h$ be the coefficients of the function $\xi_h$ in the basis $\{\varphi_z \mid z \in N_{Y_h}\}$, i.e.

$$\xi_h := \sum_{z \in N_{Y_h}} (M^\gamma_h)^{-1} \xi_h \varphi_z \in L^2(\Omega) \rightarrow H^{-1}(\Omega).$$ 

Then it holds that

$$y_h^T \xi_h = y_h^T M^\gamma_h (M^\gamma_h)^{-1} \xi_h = y_h^T M^\gamma_h \xi_h = (\xi_h, y_h)_{-1,1,\Omega} = (\xi_h, y_h)_{0,0,\Omega}.$$ 

In the following, we identify the vector $\xi_h$ (rather than $\hat{\xi}_h$) with the function $\xi_h$ and denote the associated discrete space by $\Xi_h$.

Consider an operator $\Pi$ on $L^2(\Omega)$ which satisfies

$$\int_{\Omega} \langle \Pi v, v_h \rangle_{0, \Omega} = \langle v, v_h \rangle_{0, \Omega} \quad \text{for all} \ v \in L^2(\Omega) \ \text{and} \ v_h \in Y_h.$$ 

The last step to the formulation of the discrete version of the optimal control problem considered in this text is the definition of the discrete objective functional $J_\Pi : H^1_0(\Omega) \times L^2(\Omega) \rightarrow \mathbb{R}$. Later, we will also use the discrete Lagrangian $L_\Pi : X \times H^{-1}(\Omega) \times H^1_0(\Omega) \rightarrow \mathbb{R}$. Thus, we define

$$J_\Pi(y, u) := \frac{1}{2} \|y - \Pi y_d\|^2_{0, \Omega, \Omega} + \frac{\nu}{2} \|u\|^2_{0, \Omega},$$ 

$$L_\Pi(x, \lambda, \mu) := \frac{1}{2} \|y - \Pi y_d\|^2_{0, \Omega} + \frac{\nu}{2} \|u\|^2_{0, \Omega} + a(y, p) - (u + \Pi f, p)_{0, \Omega} - \langle \xi, p \rangle_{-1,1,\Omega} - \langle \xi, \mu \rangle_{-1,1,\Omega} - \langle \lambda, y \rangle_{-1,1,\Omega}.$$ 

At this point, we leave it open to choose $\Pi$. For instance, one may consider $\Pi$ either to be equal to the identity on $L^2(\Omega)$ or the $L^2$-projection onto $Y_h$. 
The discrete optimal control problem finally reads: For given data \( f, y_d \) and \( \nu > 0 \), find a solution \((y_h, u_h, \xi_h) \in Y_h \times U_h \times \Xi_h \) to the problem

\[
\begin{align*}
(21a) \quad & \text{Minimize } J_\Pi(y_h, u_h) \\
(21b) \quad & \text{such that } a(y_h, \phi_h) - (u_h + \Pi f, \phi_h)_{0,\Omega} - \langle \xi_h, \phi_h \rangle_{-1,1,\Omega} = 0 \quad (\forall \phi_h \in Y_h), \\
(21c) \quad & y_h \geq 0, \quad \xi_h \geq 0 \quad \text{and} \quad \langle \xi_h, y_h \rangle_{-1,1,\Omega} = 0.
\end{align*}
\]

Utilizing standard techniques, one readily obtains the existence of a solution to (21). The vector representation of the latter satisfies the following discrete C-stationarity system (which does not depend on \( \Pi \)) with multipliers \( p_h, \lambda_h \) and \( \mu_h \in \mathbb{R}^{|N|_{\Xi_h}} \),

\[
\begin{align*}
(22a) \quad & A_h y_h - M_h u_h - \xi_h = f_h, \\
(22b) \quad & \xi \geq 0, \ y \geq 0, \ y^T \xi = 0, \\
(22c) \quad & M^T_h y_h - \lambda_h + A_h p_h = y_d, \\
(22d) \quad & \nu M^T_h u_h - M_h p_h = 0, \\
(22e) \quad & -\mu_h - p_h = 0, \\
(22f) \quad & \lambda_h = 0 \quad \text{if } y_{h,z} > 0, \quad \forall z \in N_{\Xi_h}, \\
(22g) \quad & \mu_h = 0 \quad \text{if } \xi_{h,z} > 0, \quad \forall z \in N_{\Xi_h}, \\
(22h) \quad & \lambda_h^T \mu_h \geq 0.
\end{align*}
\]

An optimal solution of (21) is strongly stationary, when it satisfies (22) and the following sign conditions on the biactive set,

\[
\begin{align*}
(23) \quad & \mu_{h,z} \geq 0 \quad \text{and} \quad \lambda_{h,z} \geq 0, \quad \text{if } y_{h,z} = 0 \quad \text{and} \quad \xi_{h,z} = 0, \quad \forall z \in N_{\Xi_h}.
\end{align*}
\]

The multipliers \( p_h \) and \( \mu_h \) may be interpreted as coefficient vectors pertinent to functions in \( Y_h \), whereas \( \lambda_h \) may be interpreted analogously to \( \xi_h \). Therefore, with these multipliers, we associate a tuple of functions \((p_h, \lambda_h, \mu_h) \in Y_h \times \Xi_h \times Y_h \). For the ease of notation, we abbreviate \( X_h = Y_h \times U_h \times \Xi_h \times Y_h \) and say that \((x^*_h, \lambda^*_h, \mu^*_h) \in X_h \times \Xi_h \times Y_h \) is C- or strongly stationary, if its vector representation is C- or strongly stationary.

**Remark 2.8.** Suppose that \((y^*_h, u^*_h, \xi^*_h) \) is C-stationary for (21) with associated multipliers \((p^*_h, \lambda^*_h, \mu^*_h) \in Y_h \times \Xi_h \times Y_h \). Then we have

\[
\begin{align*}
J(y^*_h, u^*_h) &= J_\Pi(y^*_h, u^*_h) + \frac{1}{2} (\|y_d\|_{0,\Omega}^2 - \|\Pi y_d\|_{0,\Omega}^2) \\
&= \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h) + \frac{1}{2}(\|y_d\|_{0,\Omega}^2 - \|\Pi y_d\|_{0,\Omega}^2) \\
&= \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h),
\end{align*}
\]

and for all \((x_h, \lambda_h, \mu_h) \in X_h \times \Xi_h \times Y_h \),

\[
\nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x_h) = \nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x_h) = 0.
\]

3. Primal-dual-weighted error representation

In our numerics, the adaptive mesh refinement will be guided by the error with respect to the objective functional, but other choices are conceivable. Hence, the latter represents the target quantity in our goal-oriented mesh adaption approach.
Proof. In view of (24), we have
\[ J(u^*_h, y^*_h) - J(u^*, y^*) = \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h) - \mathcal{L}(x^*, \lambda^*, \mu^*). \]
Taylor expansion of $\mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)$ at $x^*$ yields
\[ \mathcal{L}(x^*, \lambda^*, \mu^*) = \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h) + \nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x^*_h - x^*) \]
\[ + \frac{1}{2} \nabla_{xx} \mathcal{L}(x^*_h - x^*, x^*_h - x^*). \]
For the first term on the right-hand side in (28), we use the definition of $\mathcal{L}$ in (5) and the system (3) to obtain
\[ \mathcal{L}(x^*, \lambda^*, \mu^*) = \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h) - \mathcal{L}(x^*, \lambda^*, \mu^*). \]
Taking again advantage of (3), for the second term on the right-hand side in (28) we find
\[ \nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*)(x^*_h - x^*) = \langle \lambda^*, y^*_h \rangle_{-1,1,\Omega} + \langle \lambda^*, y^* \rangle_{-1,1,\Omega} + \langle \xi^*, \mu^* \rangle_{-1,1,\Omega} + \langle \xi^*, y^* \rangle_{-1,1,\Omega}. \]
Using (28), (30) in (27) gives the assertion.

The following lemma provides a representation of the second order Gâteaux derivative of the MPCC-Lagrangian.

Lemma 3.2. Under the same assumptions as in Theorem 3.1 for a tuple of discrete functions $\phi_h = (\delta x_h, \delta \lambda_h, \delta \mu_h)$ in $X_h \times \Xi_h \times Y_h$, with $\delta x_h = (\delta y_h, \delta u_h, \delta \xi_h, \delta \rho_h)$ it holds that
\[ \nabla_{xx} \mathcal{L}(x^*_h - x^*, x^*_h - x^*) = \nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x^*_h - x^* + \phi_h) \]
\[ - \langle \xi^*_h, \mu^* \rangle_{-1,1,\Omega} - \langle \xi^*, \mu^*_h \rangle_{-1,1,\Omega} - \langle \lambda^*, y^*_h \rangle_{-1,1,\Omega} - \langle \lambda^*, y^* \rangle_{-1,1,\Omega}. \]
Proof. Taylor expansion of $\nabla \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x^* - x^*_h)$ at the point $x^*$ and evaluation of $\nabla \mathcal{L}(x^*, \lambda^*_h, \mu^*_h)(x^* - x^*_h)$ leads to
\[ \nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x^* - x^*_h) = \nabla_x \mathcal{L}(x^*, \lambda^*_h, \mu^*_h)(x^* - x^*) + \nabla_{xx} \mathcal{L}(x^*_h - x^*, x^*_h - x^*) \]
\[ = \langle \xi^*_h, \mu^*_h \rangle_{-1,1,\Omega} + \langle \xi^*, \mu^*_h \rangle_{-1,1,\Omega} + \langle \lambda^*, y^*_h \rangle_{-1,1,\Omega} + \langle \lambda^*, y^* \rangle_{-1,1,\Omega} \]
\[ + \langle \xi^*, \mu^* \rangle_{-1,1,\Omega} + \nabla_{xx} \mathcal{L}(x^*_h - x^*, x^*_h - x^*), \]
which readily implies
\[ \nabla_{xx} \mathcal{L}(x^*_h - x^*, x^*_h - x^*) = \nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(x^* - x^*) - \langle \xi^*, \mu^* \rangle_{-1,1,\Omega} \]
\[ - \langle \xi^*, \mu^*_h \rangle_{-1,1,\Omega} - \langle \lambda^*, y^*_h \rangle_{-1,1,\Omega} - \langle \lambda^*, y^* \rangle_{-1,1,\Omega}. \]
According to (25), it holds that $\nabla_x \mathcal{L}(x^*_h, \lambda^*_h, \mu^*_h)(\phi_h) = 0$. This yields the assertion.
In combination with Theorem 3.1, the previous lemma gives rise to an error representation which will be the basis for the evaluation of the primal-dual weighted residuals.

**Theorem 3.3.** In addition to the assumptions in Theorem 3.1 let \( \delta x_h = (\delta y_h, \delta u_h, \delta \xi_h, \delta p_h) \in X_h \). Then, it holds that

\[
J(u^*, y^*) - J(u_h^*, y_h^*) = \frac{1}{2} a(y_h^*, p^* - \delta p_h) - \frac{1}{2} (u_h^* + f, p^* - \delta p_h)_{0,\Omega} - \frac{1}{2} \langle \xi_h^*, p^* - \delta p_h \rangle_{-1,1,\Omega} \\
+ \frac{1}{2} a(p_h^*, y^* - \delta y_h) + \frac{1}{2} (y_h - y_d, y^* - \delta y_h)_{0,\Omega} + \frac{1}{2} \langle \lambda_h^*, y^* - \delta y_h \rangle_{-1,1,\Omega} \\
+ \frac{1}{2} (\nu u_h^* - p_h^*, u^* - \delta u_h)_{0,\Omega} \\
+ \frac{1}{2} (\xi - \delta \xi_h, -\mu_h^* - p_h^*)_{-1,1,\Omega} \\
- \frac{1}{2} (\lambda^*, y_h^*)_{-1,1,\Omega} + \frac{1}{2} (\lambda_h^*, y^*)_{-1,1,\Omega} - \frac{1}{2} \langle \xi_h^*, \mu^* \rangle_{-1,1,\Omega} + \frac{1}{2} \langle \xi^*,\mu_h^* \rangle_{-1,1,\Omega}.
\]

**Proof.** The combination of (26) from Theorem 3.1 and (31) from Lemma 3.2 reveals

\[
J(u^*, y^*) - J(u_h^*, y_h^*) = \frac{1}{2} \nabla_x L(x_h^* - x^*, x_h^* - x^*) - \langle \lambda^*, y_h^* \rangle_{-1,1,\Omega} - \langle \xi_h^*, \mu^* \rangle_{-1,1,\Omega} \\
= \frac{1}{2} \nabla_x L(x_h^*, \lambda_h^*, \mu_h^*)\langle x^* - x_h^* - \phi_h \rangle - \frac{1}{2} \langle \xi_h^*, \mu^* \rangle_{-1,1,\Omega} + \frac{1}{2} \langle \xi^*, \mu_h^* \rangle_{-1,1,\Omega} \\
- \frac{1}{2} (\lambda^*, y_h^*)_{-1,1,\Omega} + \frac{1}{2} (\lambda_h^*, y^*)_{-1,1,\Omega}.
\]

Choosing \( \phi_h = \delta x_h - x_h^* = (\delta y_h - y_h^*, \delta u_h - u_h^*, \delta \xi_h - \xi_h^*, \delta p_h - p_h^*) \) and exploiting the structure of the MPCC-Lagrangian provides the assertion.

**Remark 3.4.** The terms (33a)-(33d) in Theorem 3.3 represent the dual and the primal error in the feasibility and optimality conditions (22a), (22c), (22d), (22e), whereas the remaining terms (33e) reflect the mismatch in complementarity as expressed by (22b) and (22f).

**Remark 3.5.** Note that in Theorem 3.1 and Theorem 3.3 one may alternatively consider the error \( J_H(y_h^*, u_h^*) - J(y^*, u^*) \). In this situation, the error representation includes additional data terms pertinent to replacing \( y_d \) by \( \Pi y_d \) in the objective, i.e.,

\[
J_H(y_h^*, u_h^*) - J(y^*, u^*) = J(y_h^*, u_h^*) - J(y^*, u^*) + \frac{1}{2} \| \Pi y_d \|_{0,\Omega}^2 - \frac{1}{2} \| y_d \|_{0,\Omega}^2.
\]

4. **Primal-Dual-Weighted A Posteriori Estimator**

Note that the representation of \( J(y^*, u^*) - J_H(y_h^*, u_h^*) \) is not fully a posteriori, which is, on the one hand, due to the weights \( x^* - \delta x_h \) and, on the other hand, due to the dependence of the terms (33e) on \( \lambda^*, y^*, \mu^* \) and \( \xi^* \). This fact prevents an immediate numerical realization of the representation in Theorem 3.3. In this section we deduce a fully a posteriori and local estimator from the terms in the error representation in Theorem 3.3.
4.1. Detailed discretization. Subsequently, we exemplarily consider specific choices for the bilinear form \( a \) and the discrete spaces. In fact, we set \( a(y, z) := \int_\Omega \nabla y \cdot \nabla z \, dx \) on \( H^1_0(\Omega) \times H^1_0(\Omega) \). Let \( \Omega \subset \mathbb{R}^2 \) be a polygonal bounded domain and let \( T \) be a triangulation of \( \Omega \) (into triangles) in the sense of [AH09, p. 394]. The set of all edges of triangles in \( T \) is denoted by \( E \), whereas \( \mathcal{N} \) is the set of nodes of triangles in \( T \). For a subset \( \omega \subset \Omega \), we write \( T(\omega) := \{ T \in T \mid T \subset \omega \} \), and, analogously, \( E(\omega) := \{ E \in E \mid E \subset \omega \} \) and \( \mathcal{N}(\omega) := \{ z \in \mathcal{N} \mid z \in \omega \} \).

By \( P_{1,0}(T) \) we denote the space of \( T \)-piecewise affine and globally continuous functions on \( \Omega \) with zero boundary conditions and the usual nodal basis

\[
\{ \varphi_z \in P_{1,0}(T) \mid z \in \mathcal{N}, \forall \bar{z} \in \mathcal{N}, \varphi_z(\bar{z}) = \delta_{z,\bar{z}} \}.
\]

We set \( Y_h = P_{1,0}(T) \) as well as \( U_h = P_{1,0}(T) \). Then, by [22d] and [22e], the terms (33e) and (33d) in the difference of the objectives yield zero. Indeed, \( M_h \) is an invertible matrix, and thus \( p_h^* = \nu u_h^* \) and \( \mu_h^* = -p_h^* \).

On every triangle \( T \in T \) we approximate \( y^*|_T \) (and \( p^*|_T \)) by the quadratic interpolant \( y_h^* \) based on the (at most) six nodes of \( \bigcup \{ S \in T \mid S \cap T \in T \cup E \} \) and \( p_h^* \) of \( p_h^* \), respectively (cf. [BR03]). The zero boundary condition guarantees the uniqueness of the interpolant also for triangles with edges on the boundary of \( \Omega \).

We define the edge jumps of a discrete function \( \phi_h \in Y_h \) with the \( T \)-piecewise constant gradient \( \nabla \phi_h \) as follows: For a triangle \( T \in T \) with an edge \( E \in \mathcal{E}(T) \) we denote by \( \nu_{T,E} \) the outer unit normal vector of \( T \), \( T_+ \), \( T_- \), and thus, the edge jump

\[
[\nabla \phi_h]_E := (\nabla \phi_h|_{T_+} - \nabla \phi_h|_{T_-}) \cdot \nu_{T_+,E} = (\nabla \phi_h|_{T_-} - \nabla \phi_h|_{T_+}) \cdot \nu_{T_-,E}
\]

is independent of the permutation of \( T_+ \) and \( T_- \). If \( E \subset \partial \Omega \), we define \( [\nabla \phi_h]_E := 0 \).

For \( z \in H^1_0(\Omega) \), \( T \)-piecewise integration by parts yields

\[
a(\phi_h, z) = \sum_{T \in T} \sum_{E \in (T \cap \partial \Omega)} \nabla \phi_h \cdot \nu_{T,E} \int_E z \, dS = \sum_{E \in \mathcal{E}} [\nabla \phi_h]_E \int_E z \, dS.
\]

Rearranging the edge terms according to triangles we obtain the identity

\[
a(\phi_h, z) = \sum_{T \in T} \sum_{E \in \mathcal{E}(T)} \frac{1}{2} [\nabla \phi_h]_E \int_E z \, dS = \sum_{T \in T} \frac{1}{2} [\nabla \phi_h]_{\partial T} \int_{\partial T} z \, dS.
\]

The total error estimator, which we derive below, has the structure

\[
\eta_h := \sum_{T \in T} \eta_T,
\]

where \( \eta_T = \eta_{PDE1,T} + \eta_{PDE2,T} + \eta_{CM,T} \). On every triangle \( T \), \( \eta_{PDE1,T} \) reflects the error contribution of the dual-weighted primal residual (33a) on \( T \), while \( \eta_{PDE2,T} \) contains the error contribution of the primal-weighted dual equation (33b) on \( T \). These estimators are derived in Section 4.2. The complementarity mismatch contribution to the error is contained in \( \eta_{CM,T} \), which is discussed in Section 4.3.
4.2. Primal-dual-weighted PDE residuals. For $\delta p_h \in P_{1,0}(T)$, we use (34) and the $L^2$-representation of $\xi_h$ due to (16) to write (33a) in Theorem 3.3 as

$$a(y_h^*, p^* - \delta p_h) - \int_T (u_h^* + f)(p^* - \delta p_h) \, dx - \langle \xi_h^*, p^* - \delta p_h \rangle_{-1,1,\Omega}$$

(36)

$$= \sum_{T \in \mathcal{T}} \left( \sum_{E \in \mathcal{E}(T)} \left( \frac{1}{2} [\nabla \phi_E]_E \int_E p^* - \delta p_h \, dS \right) - \int_T (u_h^* + f + \xi_h^*)(p^* - \delta p_h) \, dx \right).$$

We set $\delta p_h = \nu u_h^*$, estimate $p^*$ by $\nu \tilde{u}_h^*$ (cf. page 12) and recall (17) to obtain from (36)

$$\eta_{PDE1,T} := \frac{1}{2} \int_T \nu(u_h^* + f + \xi_h^*)(\tilde{u}_h^* - u_h^*) \, dx + \int_{\partial T} \frac{\nu}{2} [\nabla y_h^*]_{\partial T} (\tilde{u}_h^* - u_h^*) \, dS.$$

(37)

An analogue procedure applies to the weighted residual of the adjoint equation (33b), and we obtain

$$\eta_{PDE2,T} := \frac{1}{2} \int_T (y_h^* - y_d + \lambda_h^*)(\tilde{y}_h^* - y_h^*) \, dx + \int_{\partial T} \frac{\nu}{2} [\nabla u_h^*]_{\partial T} (\tilde{y}_h^* - y_h^*) \, dS.$$

(38)

As a consequence, the terms in

$$\eta_{PDE1} := \sum_{T \in \mathcal{T}} \eta_{PDE1,T}, \quad \eta_{PDE2} := \sum_{T \in \mathcal{T}} \eta_{PDE2,T}$$

are fully a posteriori. The integrals can be evaluated by Gauss quadrature rules which are exact for polynomials up to a certain degree.

4.3. Complementarity mismatch. The weighted complementarity residual is given by the sum

$$\frac{1}{2} \langle \xi^*, \mu_h^* \rangle_{-1,1,\Omega} - \langle \xi_h^*, \mu^* \rangle_{-1,1,\Omega} + \langle \lambda_h^*, y^* \rangle_{-1,1,\Omega} - \langle \lambda^*, y_h^* \rangle_{-1,1,\Omega},$$

and is next analyzed term by term. The resulting estimator is the sum of the a posteriori quantities defined in (39)-(42).

We begin our analysis with the terms that can be understood as $L^2$-products due to (16). In fact, considering $\langle \lambda_h^*, y^* \rangle_{-1,1,\Omega} = \langle \lambda_h^*, y_h^* \rangle_{0,\Omega}$, inserting $\lambda_h^*, y_h^* \rangle_{0,\Omega}$ and replacing $y^*$ by $\tilde{y}_h^*$ we obtain the a posteriori estimator

$$\eta_{CM1,T} := \frac{1}{2} \int_T (\tilde{y}_h^* - y_h^*) \lambda_h^* \, dx.$$

(39)

Further, subtracting $\langle \xi_h^*, \nu u_h^* \rangle_{0,\Omega}$ from $-\langle \xi_h^*, \mu^* \rangle_{-1,1,\Omega}$ and replacing $\mu^*$ by $-\nu \tilde{u}_h^*$ we find

$$\eta_{CM2,T} := \frac{1}{2} \int_T \nu(\tilde{u}_h^* - u_h^*) \xi_h^* \, dx.$$

(40)

Next, the term $\langle \xi^*, \mu_h^* \rangle_{-1,1,\Omega}$ which belongs to the mismatch of the continuous and discrete strongly active sets is analyzed. Due to $\langle \xi^*, \mu^* \rangle_{-1,1,\Omega} = 0$ following from (38), it holds that

$$\langle \xi^*, \mu_h^* \rangle_{-1,1,\Omega} = \langle \xi^*, \mu_h^* - \mu^* \rangle_{-1,1,\Omega}.$$

We exploit (3a) to conclude

$$\langle \xi^*, \mu_h^* - \mu^* \rangle_{-1,1,\Omega} = \langle Ay^* - u^* - f, \mu_h^* - \mu^* \rangle_{-1,1,\Omega}$$

$$= (-u^* - f, \mu_h^* - \mu^*)_{0,\Omega} + \langle \nabla y^*, \nabla(\mu_h^* - \mu^*) \rangle_{0,\Omega}.$$
Similarly to above, we use $\mu^* = -\nu u^*$ and $\mu^*_h = -\nu u^*_h$ whenever it appears in the weight and by $u^*_h$ in the residual, as well as we use $y^*_h$ instead of $y^*$. For $T \in T$ we obtain from (34)

$$\eta_{CM3,T} := \frac{\nu}{2} \left| -\int_T (u^*_h + f)(\bar{u}^*_h - u^*_h) \, dx + \int_{\partial T} \frac{1}{2} [\nabla y^*_h]_{\partial T} (\bar{u}^*_h - u^*_h) \, dS \right|.$$

Finally, for the term $\langle \lambda^*, y^*_h \rangle_{1,1,\Omega}$, by (3c) it holds that

$$-\langle \lambda^*, y^*_h \rangle_{1,1,\Omega} = \langle y^* + A^* p^* - y_d, y^*_h - y^*_h \rangle_{0,\Omega} + \langle \nabla p^*, \nabla (y^* - y^*_h) \rangle_{0,\Omega}.$$

We replace $y^*_h$ by $\bar{y}^*_h$ when it appears in the weight and by $y^*_h$ in the residual, and estimate

$$\eta_{CM4,T} := \frac{1}{2} \left| \int_T (y^*_h - y_d)(\bar{y}^*_h - y^*_h) \, dx + \int_{\partial T} \frac{\nu}{2} [\nabla u^*_h]_{\partial T} (\bar{y}^*_h - y^*_h) \, dS \right|,$$

where we have used integration by parts as in (34).

4.4. **Total Error.** Summarizing our above findings, we obtain the following estimator.

**Theorem 4.1.** Let $(x^*, \lambda^*, \mu^*) \in X \times H^{-1}(\Omega) \times H^1(\Omega)$ be a solution to problem (2) and let $(x^*_h, \lambda^*_h, \mu^*_h) \in X_h \times \Xi_h \times Y_h$ satisfy the conditions (22). With $\eta_{PDE1,T}$ from (37), $\eta_{PDE2,T}$ from (38), and $\eta_{CM,T} = \eta_{CM1,T} + \eta_{CM2,T} + \eta_{CM3,T} + \eta_{CM4,T}$ from (39)-(42), we estimate the local distribution of the difference $J(u^*, y^*) - J(u^*_h, y^*_h)$ by

$$\eta_h = \sum_{T \in T} \eta_{PDE1,T} + \eta_{PDE2,T} + \eta_{CM,T},$$

5. **Numerics**

We start by a brief overview of the adaptive finite element method (AFEM) and its subroutines in Section 5.1 before establishing the optimization algorithm yielding a C-stationary point of problem (21) on every adaptive refinement level in Section 5.2. Finally, Section 5.3 contains numerical results.

5.1. **Adaptive scheme.** AFEM typically follows the scheme displayed in Figure 1. In this context, the mesh-adaption process is guided iteratively by local error indicators relying on solutions of the considered problem on the current mesh. Elements with large local error indicators are marked for refinement, and a superset of the marked elements is actually refined to maintain a regular mesh. The overall algorithm repeats this cycle until a given complexity $N$ (e.g., the number of degrees of freedom in the SOLVE procedure) is reached.
Algorithm 1 AFEM for MPEC

Input: Triangulation $T_h$, data $f_h, y_{dh}$, initial state and control $y_h, u_h$, complexity $N$, bulk parameter $\theta$

1: loop
2: $(y_h, u_h, \xi_h, \lambda_h) = \text{solveMPEC}(T_h, y_h, u_h, \xi_h, f_h, y_{dh})$
3: $(\eta_T)_{T \in T_h} = \text{estimate}(T_h, y_h, u_h, \xi_h, \lambda_h, f_h, y_{dh})$
4: if $|\text{freenodes}(T_h)| > N$ then
5: return $(T_h, y_h, u_h, \xi_h, \lambda_h)$
6: end if
7: $M_h = \text{markbulk}(T_h, (\eta_T)_{T \in T_h}, \theta)$
8: $(T_h, y_h, u_h, \xi_h, f_h, y_{dh}) = \text{refineRGB}(T_h, M_h, y_h, u_h, \xi_h, f_h, y_{dh})$
9: end loop

We briefly discuss the respective steps of algorithm 1. The solution step 2 (subroutine solveMPEC) is described in detail in Section 5.2. The calculation of a local error indicator $\eta_h = \sum_{T \in T} \eta_T$ from the discrete solutions on the coarse mesh $T$ in step 3 (subroutine estimate) follows from Section 4. We use a bulk criterion (subroutine markbulk) to mark those elements in $T$ which are refined to obtain the next AFEM level. In step 7 this yields the set $M \subset T$ of minimal cardinality $|M|$ such that for a bulk parameter $\theta \in (0, 1)$ it holds that

$$\theta \eta_h \leq \sum_{T \in M} \eta_T.$$  

This criterion stops the marking procedure, relying on a greedy algorithm, as soon as the total error on $M$ has reached a fraction of $\theta$ of the total error on $T$.

In the refinement step 8 (subroutine refineRGB) one has to avoid hanging nodes and guarantee shape regularity of the triangulation. This is achieved by applying a closure algorithm, which marks additional triangles or edges; see also [Car04].

The AFEM level is indicated by $h$. In the adaptive scheme the maximum (local) mesh size is not required to converge to zero globally, as triangles may no longer get refined from some level $h$ on due to the error indicator.

5.2. Optimization algorithm. The discrete optimality system (22) is solved iteratively through a sequence of penalized problems (P). This strategy is inspired by [ATW05]. Recalling the mass and stiffness matrices of Section 2.4, on a fixed AFEM level $h$ and for a penalty parameter $\gamma > 0$ the penalized problem (P$_\gamma$) associated with (21) reads

$$\begin{align*}
\text{Minimize} & \quad \frac{1}{2} (y - M_h^{-1} y_{dh})^T M_h (y - M_h^{-1} y_{dh}) + \nu u^T M_h u + \gamma y^T \xi \\
\text{such that} & \quad A_h y = M_h u + f_h + \xi, \\
y & \geq 0 \quad \text{and} \quad \xi \geq 0.
\end{align*}$$

(P$_\gamma$) such that $A_h y = M_h u + f_h + \xi,$

$y \geq 0$ and $\xi \geq 0.$

Note that a feasible point $(y, u, \xi)$ of (P$_\gamma$) satisfies $y^T \xi \geq 0.$ Hence, (P$_\gamma$) penalizes violations of $y^T \xi = 0$ by $\gamma > 0.$ The necessary first order optimality system of (P$_\gamma$)
is given by

\begin{align}
\mathcal{R}_{\gamma,1} &:= M_h y + \gamma \xi + A_h p - \lambda - y_{d h} = 0, \\
\nu M_h u - M_h p & = 0, \\
\gamma y - p - \mu & = 0, \\
\mathcal{R}_{\gamma,2} &:= A_h y - \xi - M_h u - f_h = 0,
\end{align}

and the complementarity conditions

\begin{align}
\lambda &\geq 0, \quad y \geq 0, \quad \lambda \cdot y = 0, \\
\mu &\geq 0, \quad \xi \geq 0, \quad \mu \cdot \xi = 0.
\end{align}

We reduce the above system by eliminating the discrete adjoint state $p$ and the multiplier $\mu$. Using the max($0, \cdot$)-operator (componentwise) and an arbitrary, but fixed constant $c > 0$, we reformulate (43e)-(43f) as

\begin{align}
\mathcal{R}_{\gamma,3} &:= \lambda - \max\{0, \lambda - cy\} = 0, \\
\mathcal{R}_{\gamma,4} &:= \mu - \max\{0, \mu - c\xi\} = 0.
\end{align}

The resulting system (43a), (43d), (44) is solved by a semi-smooth Newton-method (algorithm 2); see [HIK02] for details on the latter. As an attempt towards globalization of the Newton solver, we employ a residual based strategy which adjusts the step size $\tau \in (0, 1)$ in algorithm 2 such that the total residual of system (43) decreases in each iteration. In this context, for $X = (y, u, \xi, \lambda)$, the residual $\text{res}_{\gamma}(X)$ in step 7 is the sum of four parts belonging to the four equations (43a), (43d), (44), that is,

\begin{align*}
\text{res}_{\gamma,1} &= |\mathcal{R}_{\gamma,1}^T A_h^{-1} \mathcal{R}_{\gamma,1}|^{\frac{1}{2}}, \\
\text{res}_{\gamma,2} &= |\mathcal{R}_{\gamma,2}^T A_h^{-1} \mathcal{R}_{\gamma,2}|^{\frac{1}{2}}, \\
\text{res}_{\gamma,3} &= |\mathcal{R}_{\gamma,3}|, \\
\text{res}_{\gamma,4} &= |\mathcal{R}_{\gamma,4}|, \\
\text{residual}_{\gamma}(X) &= \text{res}_{\gamma,1} + \text{res}_{\gamma,2} + \text{res}_{\gamma,3} + \text{res}_{\gamma,4}.
\end{align*}

In the associated step 7 of algorithm 2 we choose $\tau_0 = \Sigma_M$, the machine precision. Note that our globalization strategy is a simple Armijo-type backtracking which worked well in our numerical practice. For theoretical investigations and guaranteed convergence, however, a more complex strategy like a path search [FM99, Rai94] would be necessary.

Algorithm 3 corresponds to step 2 of algorithm 1. Concerning its convergence in finite dimensional spaces, we observe that MPEC-LICQ (see [ATW05, Def. 2]) is generically satisfied in the problem class under consideration. Thus, due to [ATW05], accumulation points of a series $(y_{\gamma}, u_{\gamma}, \xi_{\gamma})_{\gamma}$ of solutions to the respective penalized problems for nondecreasing $\gamma$ are either not feasible for problem 2 or satisfy C-stationarity. For completeness we also mention that under additional assumptions one may even get a strongly stationary point.
Algorithm 2 solvePenMPEC

Input: Data for system (43) including $\gamma$ and initial values $y_0, u_0, \xi_0$

1. Set $k = 0$, choose $c > 0, \varepsilon > 0, \zeta \in (0, 1), \kappa \in (0, 1)$

2. Set initial values $\lambda_0$ and $\mu_0$ according to equations (43a) and (43c)

3. loop

4. Calculate indices of active and inactive nodes
   
   $A(y_k) = \{ z \in N | \lambda_k(z) \geq cy_k(z) \}$, $I(y_k) = N \setminus A(y_k)$
   
   $A(\xi_k) = \{ z \in N | \mu_k(z) \geq c\xi_k(z) \}$, $I(\xi_k) = N \setminus A(\xi_k)$

5. Find the solution $X_{k+1}^o = (y_{k+1}^o, u_{k+1}^o, \xi_{k+1}^o, \lambda_{k+1}^o, \mu_{k+1}^o)$ to system (43a)-(43d) and set
   
   $y_{k+1}^o(A(y_k)) = 0, \lambda_{k+1}^o(I(y_k)) = 0$
   
   $\xi_{k+1}^o(A(\xi_k)) = 0, \mu_{k+1}^o(I(\xi_k)) = 0$

6. Initialize globalization strategy $\tau = 1, X_{k+1} = X_{k+1}^o$

7. while $\text{residual}_\gamma(X_{k+1}) > (1 - \kappa \tau)\text{residual}_\gamma(X_k)$ and $\tau > \tau_0 > 0$ do

8. $\tau = \zeta \tau$

9. $X_{k+1} = (1 - \tau)X_k + \tau X_{k+1}^o$

10. end while

11. if $\text{residual}_\gamma(X_{k+1}) \leq \varepsilon(\text{residual}_\gamma(X_0) + 1)$ then

12. return $y_{k+1}, u_{k+1}, \xi_{k+1}$

13. end if

14. Set $k = k + 1$

15. end loop

The residual in step 5 of algorithm 3 is computed as follows:

$$\mathcal{R}_1 := A_h y - M_h u - \xi - f_h, \quad \text{res}_1 = (\mathcal{R}_1^T A_h^{-1} \mathcal{R}_1)^{1/2},$$

$$\mathcal{R}_2 := M_h y - \lambda + \nu A_h u - y_{dh}, \quad \text{res}_2 = (\mathcal{R}_2^T A_h^{-1} \mathcal{R}_2)^{1/2},$$

$$\text{res}_3 = |u\xi|_1, \quad \text{res}_4 = |\xi - \max\{0, \xi - cy\}|_1, \quad \text{res}_5 = |y\lambda|_1,$$

(45) $\text{residual}(y, u, \xi, \lambda) = \text{res}_1 + \text{res}_2 + \text{res}_3 + \text{res}_4 + \text{res}_5.$
5.3. Numerical results. Finally, we present numerical results obtained by our AFEM approach. The bulk parameter in algorithm 1 is set to $\theta = 0.5$ in both examples.

Example 5.1. We tested our algorithm for solving [HK09, Example 6.1], where $\mathcal{A} = -\Delta$ on the square domain $\Omega = (0,1) \times (0,1)$ and

\[
\begin{align*}
  z_1(z_1) &= -4096x_1^6 + 6144x_1^5 - 3072x_1^4 + 512x_1^3, \\
  z_2(z_2) &= -244.140625x_2^3 + 585.9375x_2^2 - 468.75x_2 + 125x_2^2, \\
  y\Star(x_1, x_2) &= \begin{cases} z_1(z_1)z_2(z_2) & \text{in } (0,0.5) \times (0,0.8), \\
  0 & \text{else}, \end{cases} \\
  u\Star(x_1, x_2) &= y\Star(x_1, x_2), \\
  \xi\Star(x_1, x_2) &= 2 \max\{0, -|x_1 - 0.8| - |x_2 - 0.2|x_1 - 0.3| + 0.35\}. 
\end{align*}
\]

The data $f, y_d$ is set to

\[
f = -\Delta y\Star - u\Star - \xi\Star, \quad y_d = y\Star + \xi\Star - \nu \Delta u\Star.
\]

The parameter for the cost of the control is chosen as $\nu = 1$. The solution $(y\Star, u\Star, \xi\Star) = (y\Star, u\Star, \xi\Star)$ admits the regularity as stated in Remark 2.2.

On each refinement level in algorithm 1 the optimization routine in line 2 (algorithm 3) is employed with $\epsilon = 10^{-6}$, $\gamma = 10^{-3}$ and $\delta \gamma = 1.5$. The solver for the penalized problem, algorithm 2, is called with the parameter choices $c = 1$, $\epsilon = 10^{-6}$, $\gamma = 0.5$ and $\kappa = 10^{-4}$.

The solution calculated by our adaptive algorithm is shown in Figure 2. Owing to $u\Star = y\Star$, we plot only $y\Star_n$. For visualization purposes, the values of the multiplier vector $\xi_n\Star = -\lambda_n\Star$ in the nodes are plotted on a coarser mesh. The size of the circles indicates the respective (nonnegative) value of $\xi_n\Star$.

Figure 3 shows on the left-hand side the comparison of convergence of the error estimators as a function of the number of degrees of freedom in logarithmic scale for adaptive versus uniform refinement. The corresponding error estimators are denoted by superscript ‘A’ for ‘adaptive’ (solid lines) and ‘U’ for ‘uniform’ refinement (dashed lines). The right-hand side of Figure 3 shows the convergence of $J_h^A - J^\Star := J_h(y\Star_n, u\Star_n) - J(y\Star, u\Star)$ for discrete solutions $y\Star_n, u\Star_n$ on the adapted mesh compared to the convergence of $J_h^U - J^\Star := J_h(y\Star_n, u\Star_n) - J(y\Star, u\Star)$ for solutions $y\Star_n, u\Star_n$ on uniform meshes.
Figure 2. Solution graphs of the adaptive algorithm in Example 5.1 showing the state $y^*_h (= u^*_h$, left) and the slack variable $\xi^*_h (= -\lambda^*_h$, right). For the sake of clarity, $\xi^*_h$ is plotted on a coarser mesh.

Figure 3. Comparison of convergence of the estimators for adaptive (solid lines) and uniform refinement (dashed lines) for Example 5.1.

All parts of the estimators converge to zero, and in the adaptive variant of the algorithm, the estimated error is smaller compared to the estimated error on a uniform mesh with the same number of degrees of freedom (NrDOF). Although the convergence speed (i.e., the slope of the graphs in Figure 3) is not improved by adaptive mesh refinement owing to the high regularity of the solution, one can still observe a reduction of complexity when one aims at a certain accuracy in the objective value.

Table 1 is the convergence history of the adaptive refinement process including effectivity indices

$$\eta^A_h = \frac{\eta^A_h}{|J^A_h - J^*|}$$

on the respective AFEM levels $h$. Note that the value of $J^* := J(y^*, u^*)$ can be computed exactly from the solutions $y^*, u^*$ in this problem. The effectivity indices are smaller than 1 on coarse meshes, that is, we underestimate the real difference
Table 1. Convergence history of the adaptive method, AFEM levels 2, 6, 10, 14, 18 and 22 for Example 5.1 including effectivity indices.

<table>
<thead>
<tr>
<th>NrDOF</th>
<th>$\eta_h^A$</th>
<th>$J_h^A - J^*$</th>
<th>Eff.Ind.</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.03e+02</td>
<td>2.51e+02</td>
<td>0.411</td>
</tr>
<tr>
<td>27</td>
<td>6.38e+01</td>
<td>1.81e+02</td>
<td>0.352</td>
</tr>
<tr>
<td>225</td>
<td>1.33e+01</td>
<td>2.54e+01</td>
<td>0.524</td>
</tr>
<tr>
<td>2806</td>
<td>1.11e+00</td>
<td>1.99e+00</td>
<td>0.557</td>
</tr>
<tr>
<td>28311</td>
<td>1.11e-01</td>
<td>1.92e-01</td>
<td>0.581</td>
</tr>
<tr>
<td>235726</td>
<td>1.33e-02</td>
<td>1.23e-02</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Example 5.2. We consider again $A = -\Delta$ on the L-shaped domain $\Omega = (-1,0) \times (-1,1) \cup (0,1) \times (0,1)$ and define

$$y_d(x) = \begin{cases} -1 & \text{if } |x| > \frac{1}{10}, \\ 1 - 100x_1^2 - 50x_2^2 & \text{else}, \end{cases}$$

$$f(x) = \frac{1}{2} + \frac{1}{2}(x_1 - x_2),$$

$$\nu = 0.01.$$ 

On each refinement level in algorithm 1, the optimization routine in line 2 (algorithm 3) is employed with $\epsilon = 10^{-6}$, $\gamma = 10^{-2}$ and $\delta \gamma = 1.5$. The solver for the penalized problem, algorithm 2, is called with the parameter choices $c = 0.1$, $\epsilon = 10^{-6}$, $\zeta = 0.5$ and $\kappa = 10^{-4}$.

The solutions calculated by our adaptive algorithm are depicted in Figure 4. In Figure 5, we show again the convergence of the error estimator in adaptive versus uniform refinement (left), as well as the convergence comparison of $J_h(y^*_h, u^*_h) - J(y^*, u^*)$ (right). The lack of regularity of the solution due to the non-convexity of $\Omega$ seems to favor adaptive refinement when compared to uniform refinement. In fact, we observe a significantly better convergence rate for AFEM in this case.

Table 2 lists the data from the plots in Figure 5 together with effectivity indices for different AFEM levels. The effectivity indices show once again that the error estimators are not reliable, but converge to 1, similar to those in [BR03] for discretizing PDEs. The value of $J(y^*, u^*)$ used for the plot and the table is estimated by solutions on the adaptively generated mesh after additional uniform refinements. Finally, the adaptively generated mesh is depicted in Figure 6. We observe refinement in the non-convex corner as well as in the regions where $y^*_h$ is steep. Further refinement is found near the boundary between active and inactive sets, which is due to the complementarity mismatch error.
Figure 4. Solution graphs of the adaptive algorithm in Example 5.2: first row left state $y_h^\star$, right slack variable $\xi_h^\star$, second row left control $u_h^\star$ and right multiplier $\lambda_h^\star$.

Figure 5. Comparison of convergence of the estimators for adaptive (solid lines) and uniform refinement (dashed lines) for Example 5.2.

References

Table 2. Convergence history of the adaptive method, AFEM levels 2, 10, 17, 23 and 29 for Example 5.2 including effectivity indices.

Figure 6. Adaptively refined mesh for Example 5.2.
ADAPTIVITY IN OPTIMAL CONTROL OF VARIATIONAL INEQUALITIES


Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany
E-mail address: hint@math.hu-berlin.de

Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA
E-mail address: rohop@math.uh.edu

and: Institute of Mathematics, University of Augsburg, Universitätstraße 14, D-86152 Augsburg, Germany
E-mail address: hoppe@math.uni-augsburg.de

Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany
E-mail address: loebhard@math.hu-berlin.de