On Solving Elliptic Obstacle Problems by Compact Abs-Linearization

Olga Weiß¹ and Monika Weymuth²

¹Institut für Mathematik, Humboldt-Universität zu Berlin,
²Institut für Mathematik und Computergestützte Simulation, Universität der Bundeswehr München

Draft of January 24, 2021

Abstract

We consider optimal control problems governed by an elliptic variational inequality of the first kind, namely the obstacle problem. The variational inequality is treated by penalization which leads to optimization problems governed by a nonsmooth semi-linear elliptic PDE. The CALi algorithm is then applied for the efficient solution of these nonsmooth optimization problems. The special feature of the optimization algorithm CALi is the treatment of the nonsmooth Lipschitz-continuous operators abs, max and min, which allows to explicitly exploit the nonsmooth structure. Stationary points are located by appropriate decomposition of the optimization problem into so called smooth compact abs-linearized problems. Each of these compact abs-linearized problems can be solved by classical means. The comprehensive algorithmic concept is presented and its performance is discussed through examples.

Keywords: Variational Inequality, Obstacle Problem, Constrained Optimal Control, Finite Element Method, Nonsmooth Optimization, Abs-Linearization, A Priori Error Analysis

1 Introduction

In the present paper we consider an optimal control problem governed by an elliptic variational inequality of the first kind, more precisely the obstacle problem. There are various important applications which are modeled by means of variational inequalities as e.g., elastoplasticity or piezo electricity. Due to the variational inequality constraint these kind of problems are nonsmooth and non-convex which complicates their theoretical and numerical treatment.

The focus of this paper is put on a new algorithmic realization. Therefore we will not discuss the topic of necessary and sufficient optimality conditions which is already intensively investigated in the literature. We mention e.g., [22, 23, 16, 14, 25, 18, 28, 12, 2] and the references therein.

Various solution algorithms for optimal control of the obstacle problem already exist in the literature. A commonly used approach is to regularize or penalize the variational inequality to get a semi-linear partial differential equation (PDE), where the nonlinearity depends on the regularization parameter, see e.g., [3, 23, 16, 13, 15, 18, 26, 21] and the references therein. We will also follow this approach.

In [29] a new structure exploiting optimization method to solve optimal control problems subject to elliptic semi-linear and nonsmooth equations, the so called CALi algorithm, is proposed. In contrast to the usually applied smoothing and regularization techniques for nonsmooth optimization problems, this algorithm allows for an explicit exploitation of the structure caused by the nonsmoothness. For this purpose a special treatment of the absolute value operator, the so called compact abs-linearization, is applied, depending on the level, where the nonsmoothness occurs.

The purpose of this paper is to show that this algorithm is versatile and multifunctional, so that it can also be applied to optimal control problems governed by elliptic variational inequalities (VIs) of the first kind. The main goal of this paper is to elaborate and illustrate the adjustment of
this algorithm to exactly this class of nonsmooth optimization problems. Therefore the considered
optimization problems are reformulated into optimal control problems governed by nonsmooth
elliptic PDEs. The CALi algorithm is then applied for the efficient solution of these nonsmooth
optimization problems with the absolute value operator as the only source of nonsmoothness. Due
to reformulations based on the compact abs-linearization and well-known abs-linear reformulations
this covers also (but not exclusively) nonsmoothness given by the max and min operators.
The exploitation of the given data allows a targeted and optimal decomposition of the optimiza-
tion problem in order to compute stationary points. This approach is able to solve the considered
class of nonsmooth optimization problems in comparably less Newton steps and additionally main-
tains reasonable convergence properties. Numerical results for nonsmooth optimization problems
illustrate the proposed approach and its performance.
An appropriate decomposition of the optimization problem into so called smooth compact abs-
linearized problems allows to compute the solution of the corresponding optimization problem
constrained by VIs. Each of these compact abs-linearized problems can be solved by classical
means. The comprehensive algorithmic concept is presented and its performance is discussed
through examples.
In order to solve our problem numerically we discretize the semi-linear PDE arising by regular-
ization of the VI-constraint with the help of continuous, piecewise linear finite elements for the
state and piecewise constant functions for the control. Additional results of the present paper are
the convergence rates with respect to the regularization parameter for the error in the control and
the state. Our final error estimate contains information about the coupling of the regularization
parameter and the mesh size. Similar error estimates for another smoothing-scheme are established
in [26].

This paper has the following structure.
In Sec. 2, we introduce the considered problem class of obstacle problems, discuss its properties and
propose a suitable regularization which leads to an optimization problem with nonsmooth PDE
constraint. Furthermore, the solution operators corresponding to the original and the penalized
problem are also introduced and examined, as well as their relation. Sec. 3 presents a reformulation
of the nonsmooth optimization problem into a smooth one using the compact-abs-linearization to-
gether with a solution approach involving a penalty term and introduces the resulting optimization
algorithm. Moreover, the chosen discretization approach as well as the solution of the resulting
finite dimensional optimization problems are discussed. Sec. 4 deals with an investigation of error
estimates with respect to regularization and discretization. Numerical results for a collection of
test problems are presented and analyzed in Sec. 5. Finally, a conclusion and an outlook are given
in Sec. 6.

2 Preliminaries

2.1 Notation and Problem Statement
Throughout this work we use the standard notation $H^1_0(\Omega)$ and $W^{k,p}(\Omega)$, $k \in \mathbb{N}$, $1 \leq p \leq \infty$ for
the Sobolev spaces on a domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$. We refer to [1] for details of these spaces. As
usual the dual of $H^1_0(\Omega)$ w.r.t. the $L^2$-inner product is denoted by $H^{-1}(\Omega)$ and the symbol $\langle \cdot , \cdot \rangle$
denotes the dual pairing between $H^1_0(\Omega)$ and $H^{-1}(\Omega)$. The $L^2$-scalar product is denoted by $(\cdot , \cdot )$.
Moreover, we introduce the bilinear form $a : H^1_0(\Omega) \times H^1_0(\Omega) \to \mathbb{R}$ by

$$a(y,v) := \int_\Omega \nabla y \cdot \nabla v \, dx.$$  

(2.1)

The coercivity constant of $a$ will be denoted by $\beta$, i.e.,

$$a(v,v) \geq \beta \|v\|_{H^1_0(\Omega)}^2 \quad \forall v \in H^1_0(\Omega).$$  

We consider optimal control problems governed by an elliptic variational inequality of the first
kind. These kind of optimization problems are also known as the (elliptic) obstacle problem.

\[
\begin{align*}
\min_{(y,u) \in K \times L^2(\Omega)} J(y,u) := & \frac{1}{2} \| y - y_d \|^2_{L^2(\Omega)} + \frac{\alpha}{2} \| u \|^2_{L^2(\Omega)}, \\
\text{s.t.} \quad & (\nabla y, \nabla (v - y)) \geq (u + f, v - y) \quad \forall v \in K
\end{align*}
\]  

(2.2a) (2.2b)

where the \( \psi \) denotes a given obstacle and the cone \( K \) is defined by

\[
K = \{ v \in H^1_0(\Omega) : v \geq \psi \text{ a.e. in } \Omega \}.
\]

We impose the following assumptions on the data in Eq. (2.2):

i) \( \Omega \subset \mathbb{R}^d \) \((d = 1, 2, 3)\) is a bounded domain that is either convex and polygonal or has a \( C^{1,1} \)-boundary.

ii) The desired state satisfies \( y_d \in L^2(\Omega) \) and \( \alpha > 0 \) is a fixed real number.

iii) The obstacle \( \psi \) satisfies \( \psi \in W^{2,\infty}(\Omega), \psi \leq 0 \text{ a.e. on } \partial \Omega \).

iv) The given disturbance \( f \) is a function in \( L^2(\Omega) \).

Note that the condition \( \psi \leq 0 \text{ a.e. on } \partial \Omega \) is needed to ensure the existence of solutions for the obstacle problem.

### 2.2 Known Results and Penalization of the Obstacle Problem

In the following we summarize some known results about the variational inequality (2.2b) and the optimal control problem (2.2).

We start with an existence and uniqueness result.

**Lemma 2.1.** For every \( u \in H^{-1}(\Omega) \) the variational inequality (2.2b) has a unique solution \( y \in K \).

Moreover, the associated solution operator \( S : H^{-1}(\Omega) \to K \subset H^1_0(\Omega) \) mapping \( u \) to \( y \), is globally Lipschitz continuous with Lipschitz constant \( L = 1/\beta \) with \( \beta \) as in Eq. (2.1).

The proof is standard and can, for instance, be found in \([17]\].

It is important to note that the control-to-state operator \( S \) is in general not Gâteaux differentiable unless the biactive set

\[
\{ x \in \Omega : y(x) = \psi(x), -\Delta y(x) = u(x) \}
\]

has measure zero, see \([22]\].

The continuity of \( S \) and the weak lower semicontinuity of \( J \) imply the following result which can be found e.g., in \([3]\):

**Proposition 2.2.** There exists a globally optimal solution of (2.2) which is in general not unique due to the nonlinearity of \( S \).

In order to solve the variational inequality (2.2b) we use a common technique called penalization. The idea of this method is to approximate the variational inequality by a sequence of nonlinear equations. For details of penalization we refer to \([8, 17]\).

Using the max-function as penalty operator the variational inequality (2.2b) can be approximated by the penalized equation

\[
(\nabla y, \nabla v) - \frac{1}{\varepsilon}(\max(0, \psi - y), v) = \langle f + u, v \rangle \quad \forall v \in H^1_0(\Omega).
\]

(2.3)

For every \( u \in H^{-1}(\Omega) \), Eq. (2.3) has a unique solution \( y_\varepsilon(u) \) due to the monotonicity of the max-function (see e.g., \([8]\)). Therefore the associated solution operator \( S_\varepsilon : H^{-1}(\Omega) \to H^1_0(\Omega) \), mapping \( u \) to \( y_\varepsilon \), is well-defined.

**Lemma 2.3.** The operator \( S_\varepsilon \) is globally Lipschitz continuous with Lipschitz constant \( 1/\beta \) with \( \beta \) as in Eq. (2.1).
Proof. The proof is straightforward. We set $y^{(1)} = S_{\varepsilon}(u_1)$ and $y^{(2)} = S_{\varepsilon}(u_2)$ and insert $v = y^{(1)} - y^{(2)} \in H^1_0(\Omega)$ in (2.3) with $u = u_1$ and $u = u_2$. Subtracting the arising equalities and using the coercivity of $a(\cdot, \cdot)$ leads to
\[
\beta\|y^{(1)}_\varepsilon - y^{(2)}_\varepsilon\|_{H^1(\Omega)}^2 \leq a(y^{(1)}_\varepsilon - y^{(2)}_\varepsilon, y^{(1)}_\varepsilon - y^{(2)}_\varepsilon)
\]
\[
= (u_1 - u_2, y^{(1)}_\varepsilon - y^{(2)}_\varepsilon) + \frac{1}{\varepsilon}(\max(0, \psi - y^{(1)}_\varepsilon) - \max(0, \psi - y^{(2)}_\varepsilon), y^{(1)}_\varepsilon - y^{(2)}_\varepsilon)).
\]
The monotonicity of the max-function implies the claim.

The following result is well-known and can be found e.g., in [8].

Lemma 2.4. It holds that $S_{\varepsilon}(u) \rightharpoonup S(u)$ in $H^1_0(\Omega)$ as $\varepsilon \to 0$, where $S(u)$ denotes the solution of the variational inequality (2.2b) associated with $u$.

Theorem 2.5. Let $\{u_\varepsilon\}_{\varepsilon>0} \subset L^2(\Omega)$ be a sequence that converges weakly in $L^2(\Omega)$ to $u \in L^2(\Omega)$ as $\varepsilon \to 0$. Then we have the strong convergence
\[
S_{\varepsilon}(u_\varepsilon) \xrightarrow{\varepsilon \to 0} S(u) \quad \text{in} \quad H^1_0(\Omega).
\]

Proof. By the triangle inequality we have

\[
\|S_{\varepsilon}(u_\varepsilon) - S(u)\|_{H^1(\Omega)} \leq \|S_{\varepsilon}(u_\varepsilon) - S_{\varepsilon}(u)\|_{H^1(\Omega)} + \|S_{\varepsilon}(u) - S(u)\|_{H^1(\Omega)}.
\]

We observe that the second term tends to zero by Lem. 2.4. Moreover, due to Lem. 2.3 the first term can be estimated by
\[
\|S_{\varepsilon}(u_\varepsilon) - S_{\varepsilon}(u)\|_{H^1(\Omega)} \leq \frac{1}{\beta}\|u_\varepsilon - u\|_{H^{-1}(\Omega)}.
\]

By compact embeddings the right-hand side of (2.4) tends to zero for $\varepsilon \to 0$. 

Applying Eq. (2.3) the optimal control problem (2.2) can be approximated by
\[
\min_{(y, u) \in H^1_0(\Omega) \times L^2(\Omega)} J(y, u) := \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2,
\]

s.t. $(\nabla y, \nabla u) - \frac{1}{\varepsilon}(\max(0, \psi - y), v) = (f + u, v) \quad \forall v \in H^1_0(\Omega)$.

Due to the continuity of $S_{\varepsilon}$ and the weak lower semicontinuity of $J$ problem (2.5) has a globally optimal solution. Moreover, using Thm. 2.5 we can argue as in [18, proof of Thm. 3.14] to show that for each strictly locally optimal pair $(y^*, u^*)$ of Eq. (2.2) there is a family of local solutions $(y_\varepsilon, u_\varepsilon)$ of (2.5) that converges strongly to $(y^*, u^*)$ in $H^1_0(\Omega) \times L^2(\Omega)$.

3 The Compact Abs-Linearization

In this section we show how the nonsmooth optimization problem (2.5) can be reformulated into a closely related but smooth one. The reformulation is done by using the compact abs-linearization (CALi) which is introduced in [29] and based on the idea described in [9, 10]. Moreover, we explain how the algorithm CALi can be adapted to our problem class and finally we examine a finite element discretization of the reformulated optimization problem.

3.1 Reformulation of the Penalized Optimal Control Problem

Throughout this section the operator $\ell : L^2(\Omega) \to L^2(\Omega)$ is defined by
\[
\ell(y) := \max(0, y).
\]

The nonsmooth operator $\ell$ has the following properties:
Lemma 3.1. 1. The operator \( \ell(y)(x) = \ell(y(x)) \), \( \ell(\cdot) : L^2(\Omega) \to L^2(\Omega) \) denotes an (autonomous) Nemytzkii operator induced by some nonlinear and nonsmooth function \( \ell(t) : \mathbb{R} \to \mathbb{R} \) which satisfies the Carathéodory conditions, i.e., the mapping \( t \mapsto \ell(t) \) is continuous on \( \mathbb{R} \).

2. The function \( \ell : \mathbb{R} \to \mathbb{R} \) is monotonically increasing, and satisfies the boundedness condition
   \[
   |\ell(0)| \leq K \tag{3.1}
   \]
   for some constant \( 0 \leq K < \infty \). Furthermore \( \ell \) is globally Lipschitz continuous, i.e.,
   \[
   \exists L > 0 : |\ell(t_1) - \ell(t_2)| \leq L|t_1 - t_2| \quad \forall t_i \in \mathbb{R}. \tag{3.2}
   \]

3. The Nemytzkii operator \( \ell \) is directionally differentiable, i.e.,
   \[
   \left\| \frac{\ell(y + th) - \ell(y)}{t} - \ell'(y; h) \right\|_{L^2(\Omega)} \to 0 \text{ for } t \to 0_+, \quad \forall y, h \in L^2(\Omega), \tag{3.3}
   \]
   with \( \ell' \) being locally Lipschitz continuous and monotone.

4. The operator \( \ell \) can be expressed as a finite composition of the absolute value function and Fréchet differentiable operators.

Proof. The operator \( \ell(y) = \max(0,y) \) is a Nemytzkii operator induced by the pointwise non-linear and nonsmooth function \( \max(0, t) : \mathbb{R} \to \mathbb{R} \geq 0, t \mapsto \max(0, t) \), which satisfies the Carathéodory conditions. The function \( \max(0, t) \) is obviously monotonically increasing and satisfies the boundedness condition (3.1) with \( |\ell(0)| = |\max(0, 0)| = 0 =: K \). Furthermore, \( \max(0, t) \) is globally Lipschitz-continuous with constant \( L = 1 \).

Since the inducing function \( \ell(t) = \max(0, t) : \mathbb{R} \to \mathbb{R} \geq 0 \) is directionally differentiable with
   \[
   \max'((0, t); h) = \begin{cases} h, & \text{if } t > 0 \\ \max(0, h), & \text{if } t = 0 \\ 0, & \text{if } t < 0 \end{cases},
   \]
   it is also globally Lipschitz-continuous and monotone itself. Hence,
   \[
   |\max'((0, t); h)| \leq |h|
   \]
   for all directions \( h \in \mathbb{R} \). Then Lebesgue’s dominated convergence theorem implies also the directional differentiability of the induced Nemytzkii operator. This proves the assertions 1.-3. Assertion 4. follows from Prop. 3.2 below.

Note that by Lem. 3.1, points 1. and 2. it follows that also the associated Nemytzkii operator \( \ell : L^2(\Omega) \to L^2(\Omega) \) is Lipschitz-continuous and monotonically increasing.

For convenience of the reader we briefly explain the basic ideas of the structured evaluation and the compact abs-linearization described in [29, Def. 2.4 and 3.1]. For this purpose we introduce a new auxiliary function \( z \), called the switching function, for the argument of the absolute value function and \( \sigma \) for the sign of \( z \). The reformulation results in a representation, where all evaluations of the absolute value function can be clearly identified and exploited.

Before we introduce the structured evaluation, we want to recall the well-known reformulation
   \[
   \max(v, u) = (v + u + \text{abs}(v - u))/2. \tag{3.4}
   \]
Consider the nonsmooth Lipschitz-continuous operator \( \ell : H^1_0(\Omega) \to L^2(\Omega), \ell(y) = \max(0, y) \):
Proposition 3.2 (Structured evaluation for max(0, y)). An equivalent representation of $\ell(y) = \max(0, y)$ denoted by $\hat{\ell}$ can be obtained using the structured evaluation given by

$$
\begin{align*}
\hat{\ell}(y, \sigma z) &= \frac{1}{4}(y + \sigma z).
\end{align*}
$$

Note, that for the setting considered here, one has $z \in H^1(\Omega)$ and the function $\sigma$ is a Nemytskii operator defined by

$$
\sigma : H^1(\Omega) \rightarrow H^1(\Omega), \quad [\sigma(z)](x) = \text{sign}(z(x)) z(x) \quad \text{a.e. in } \Omega
$$

as functions of $z$. This choice ensures that $\sigma(z) = \text{abs}(z) \in H^1(\Omega)$ holds. Note that the Nemytskii operator $\sigma$ takes the values $-1, 0$ or $1$, i.e., $\sigma : \Omega \rightarrow \{-1, 0, 1\}$ and depends nonsmoothly on the switching function $z$. However, by applying the compact abs-linearization, defined below, one reformulates the operator equation of the optimization problem at hand into a smooth one.

Definition 3.3 (Compact Abs-Linearization). For a given nonsmooth operator equation involving the abs-operator or reformulated by the structured evaluation in the fashion of [29, Def. 2.4] and Prop. 3.2, the compact abs-linearization of the nonsmooth operator equation is obtained by fixing the involved $\sigma$ to a given $\hat{\sigma} \in L^2(\Omega)$, $\hat{\sigma} : \Omega \rightarrow \{-1, 1\}$.

Hence, using the compact abs-linearization, the resulting operator equation is smooth in both arguments $z$ and $\hat{\sigma}$, since the nonsmooth dependence of $\sigma$ on $z$ has been eliminated.

Note, that in the context of compact abs-linearization, the function $\hat{\sigma}$ takes only the values $1$ and $-1$, but no longer $0$. However, this does not influence the previous considerations, simply because if $z > 0$ and $\hat{\sigma} = +1$, or $z < 0$ and $\hat{\sigma} = -1$ respectively, then $\hat{\sigma} z = \text{abs}(z)$ is still valid. If $z = 0$, then even for $\hat{\sigma} \neq 0$ the relationship $\text{abs}(z) = \sigma z = 0$ is guaranteed. Hence, no longer considering zero as a value for $\hat{\sigma}$ does not pose any limitations. However fixing $\sigma$ to a certain function according to Def. 3.3 provides a linearization in the following sense. Since the dependency of $z$ and $\sigma$ has been removed, the term $\sigma z$ is now smooth and even linear in $z$.

Using the reformulation Eq. (3.4) for the max-function in Eq. (2.5b) as well as applying the compact-abs-linearization problem (2.5) can be reformulated into the smooth optimization problem

$$
\begin{align*}
\min_{y, z, u} & \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} , \quad \text{(3.5a)} \\
\text{s.t.} & \quad (\nabla y, \nabla v) - \frac{1}{2\nu}(\psi - y + \sigma z, v) = \langle u + f, v \rangle, \quad \forall v \in H^1_0(\Omega) \quad \text{(3.5b)} \\
& \quad (z - (\psi - y), v) = 0, \quad \forall v \in H^1_0(\Omega) \quad \text{(3.5c)} \\
& \quad \sigma z \geq 0 \quad \text{a.e. in } \Omega. \quad \text{(3.5d)}
\end{align*}
$$

Following the same procedure as in [29], we treat the inequality constraint (3.5d) with a penalty approach such that the objective function (3.5a) is modified to

$$
\begin{align*}
\min_{y, z, u} & \quad J(y, u) + \nu \int_{\Omega} \left( \max(-\sigma z, 0) \right)^4 \text{d}\Omega \quad \text{(3.6)}
\end{align*}
$$

with a penalty factor $\nu > 0$. In this framework, as well as in the remainder of this paper, $\nu$ describes a non-negative constant penalty parameter for the inequality condition on $\sigma z$. Here, the exponent 4 ensures that the target function is twice continuously differentiable despite the max function that is used for the formulation of the penalty function. This modified objective function is then coupled with the equality constraints (3.5b) and (3.5c) using Lagrange multipliers, resulting in the following Lagrange function

$$
\begin{align*}
\mathcal{L}(y, z, u, \lambda_{\text{PDE}}, \lambda_z) &= J(y, u) + \nu \int_{\Omega} \max(-\sigma z, 0)^4 \text{d}\Omega + \langle \nabla \lambda_{\text{PDE}}, \nabla y \rangle \quad \\
& \quad + \langle \lambda_{\text{PDE}}, -\frac{1}{2\nu}(\psi - y + \sigma z - (f + u)) + (\lambda_z, z - (\psi - y)) \rangle. \quad \text{(3.7)}
\end{align*}
$$
Here we assume that the optimization problem at hand fulfills some kind of constraint qualification in the sense of [4, 27] to ensure the existence of corresponding Lagrange multipliers and therefore also the Lagrange function. As already discussed in [29] this requirement is both justified and appropriate and is certainly met by the class of reformulated elliptic optimal control problems considered here, which satisfy the Slater condition, see e.g., [29, Rem. 2.17].

3.2 The Algorithm CALi

Before we introduce our algorithm for the solution of problem (3.5), we want to investigate a specific choice for the fixed \(\bar{\sigma}\).

Definition 3.4. For some \(\psi \in W^{2,\infty}(\Omega)\) and \(y_d \in L^2(\Omega)\) we denote by \(\bar{\sigma}_\psi\) which is defined by

\[
\bar{\sigma}_\psi = \text{sign}(\psi - y_d) = \begin{cases} +1, & \text{on } \Omega^+ := \{x \in \Omega : y_d(x) < \psi(x)\} \\ -1, & \text{on } \Omega^- := \{x \in \Omega : y_d(x) \geq \psi(x)\} \end{cases}, \tag{3.8}
\]

the fixed \(\bar{\sigma}\) with respect to the desired state \(y_d\) and the obstacle \(\psi\) according to compact abs-linearization in Def. 3.3 by virtue of the respective structured evaluation of \(\max(0, \psi - y_d)\).

As already discussed in [29], the choice of the specific \(\bar{\sigma}\) is crucial, since it determines the decomposition of the domain of the underlying optimization problem. In the following we want to emphasize the domain decomposition due to the compact abs-linearization and motivate the choice of \(\bar{\sigma}_\psi\). For this purpose we will first consider the domain decomposition given by \(-\text{sign}(y_d)\) with the desired state \(y_d\), i.e.,

\[
-\text{sign}(y_d(x)) = \begin{cases} -1, & \text{for } x \in \Omega^- \triangleq \{x \in \Omega : y_d(x) \geq 0\} \\ +1, & \text{for } x \in \Omega^+ \triangleq \{x \in \Omega : y_d(x) < 0\} \end{cases} \tag{3.9}
\]

Note that \(\bar{\sigma} = -\text{sign}(y_d)\) defined by Eq. (3.9) corresponds exactly to \(\bar{\sigma}_\psi\) for \(\psi \equiv 0\), i.e.,

\[
\bar{\sigma}_0 = \begin{cases} -1, & \text{for } x \in \Omega^- \\ +1, & \text{for } x \in \Omega^+ \end{cases}.
\]

Hence, every fixed \(\bar{\sigma}\) and especially \(\bar{\sigma}_\psi\) analogous to Eq. (3.9) correspondingly decomposes the domain \(\Omega\) into subdomains such that \(\Omega = \Omega^+ \cup \Omega^-\) with \(\bar{\sigma}(x) = +1\) on \(\Omega^+\) and \(\bar{\sigma}(x) = -1\) on \(\Omega^-\).

We consider the following example.

Example 3.5 (Domain Decomposition by \(\bar{\sigma}\)). In order to motivate the choice of a specific \(\bar{\sigma}\) and to illustrate the corresponding domain decomposition, we examine the following obstacle problem from [20], where we replaced the domain \(\Omega_1\) by \(\Omega\):

\[
\Omega = (0, 1)^2 \subseteq \mathbb{R}^2, \quad y_d(x_1, x_2) = -\sin(\pi x_1) \sin(\pi x_2), \quad f(x_1, x_2) = -2\pi^2 \sin(\pi x_1) \sin(\pi x_2)
\]

and \(\psi(x_1, x_2) = -0.25\)

Note that due to the chosen data the optimal state is given by

\[
y^* = \begin{cases} \psi, & \text{on } \Omega^+ := \{x \in \Omega : y_d(x) < \psi(x)\} \\ y_d, & \text{on } \Omega^- := \{x \in \Omega : y_d(x) \geq \psi(x)\} \end{cases}
\]

and due to choice of \(f\), \(y^*\) can be attained for all controls \(u \leq 0\) and hence especially for \(u^* \equiv 0\). Consequently, \(z^* = \psi - y^*\). However, this corresponds exactly to the function \(z_\psi\), which is given by

\[
z_\psi := \begin{cases} 0, & \text{on } \Omega^+ \\ \psi - y_d, & \text{on } \Omega^- \end{cases}.
\]
Therefore, we choose
\[ \bar{\sigma} = \bar{\sigma}_{\psi} = \begin{cases} +1, & \text{on } \Omega^+ \\ -1, & \text{on } \Omega^- \end{cases} \]
correspondingly. Based to these considerations we usually set \( \bar{\sigma} = \bar{\sigma}_{\psi} \) as defined in Def. 3.4. This choice is reasonable according to the specifications in [29] for the cases considered here whenever the function \( y^* = \max(\psi, y_d) \) can be reached as a feasible state.

Due to the decomposition of the domain \( \Omega \) into \( \Omega = \Omega^+ \cup \Omega^- \) provided by \( \bar{\sigma}_{\psi} \) Eq. (3.5b) can be reformulated as
\[
\int_{\Omega} \nabla y \cdot \nabla v \, dx - \frac{1}{2\varepsilon} \int_{\Omega^+} (\psi - y + \bar{\sigma}_{\psi} z) v \, dx - \frac{1}{2\varepsilon} \int_{\Omega^-} (\psi - y - z) v \, dx = \int_{\Omega} (u + f) v \, dx.
\]
Motivated by [29] and the previous examinations, we propose the method stated in Algo. 1 to solve the optimization problem (3.6) with constraints (3.5b)–(3.5c).

**Algorithm 1** CALi

**Input:** Initial values: \( \bar{\sigma}_{\psi}, y^0, z^0, u^0 \)
Parameter: \( \alpha, \nu, \varepsilon \geq 0 \)

Solve problem (3.6) subject to (3.5b)–(3.5c) for \( \bar{\sigma} = \bar{\sigma}_{\psi} \) to obtain \( y^*, z^*, u^*, \lambda_{PDE}^*, \lambda_z^* \)

**Output:** \( y^*, z^*, u^* \)

Since the proposed algorithm is essentially motivated by the special handling of the absolute value function, i.e., the compact abs-linearization, we call the resulting optimization algorithm CALi for Compact Abs-Lineaarization.

It should be noted that the solution of the smooth reformulated penalized problems can be accomplished with traditional methods of smooth optimization. For the numerical results shown in Sec. 5, we used a finite-element-approach based on FEniCS [19] to discretize the PDEs and to describe the other constraints in combination with a Newton method for the solution of the smooth modified compact abs-linearized problems.

For the initial state, control, switching function and signature function \( \bar{\sigma}_{\psi} \), the non-linear variational Lagrange problem is solved by Newton’s method using the derivatives calculated within FEniCS.

### 3.3 The Discrete System

We will discretize the considered systems with linear finite elements. For this purpose we introduce a family of meshes \( \{T^h\} \). The mesh \( T^h \) consists of open triangles \( T \) and the mesh width is defined by
\[
h := \max_{T \in T^h} h_T \quad \text{with} \quad h_T := \text{diam}(T).
\]
Moreover we assume that \( T^h \) is quasi-uniform in the sense of [5].

For the discretization of system (2.2) we introduce the space of piecewise linear functions
\[
V^h := \{ v^h \in H^1_0(\Omega) : v^h|_T \in P_1(T) \quad \forall T \in T^h \},
\]
where \( P_1 \) denotes the space of polynomials of degree \( \leq 1 \). The nodal basis of \( V^h \) is given by \( \{\xi_1, \ldots, \xi_n\} \). Moreover we define the space of piecewise constant functions by
\[
U^h := \text{span}\{e_T : T \in T^h\},
\]
where \( e_T : \Omega \rightarrow \mathbb{R} \) is the characteristic function for the simplex \( T \in \mathbb{T}^h \).

For a given function \( y^h \in V^h \) we denote by \( \mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n \) its vector of coefficients with respect to the basis \( \{\xi_1, \ldots, \xi_n\} \), i.e.,

\[
y^h(x) = \sum_{i=1}^{n} y_i \xi_i(x).
\]

Similarly, every discretized control function in the space \( U^h \) with \( u = (u_1, \ldots, u_m)^T \in \mathbb{R}^m \) can be written as

\[
u^h(x) = \sum_{i=1}^{m} u_i e_{T_i}(x),
\]

where \( m \) is the total number of elements \( T \) in the triangulation \( \mathbb{T}^h \).

One has to take into account that the operators max and \( \text{abs} \) are non-linear. The above representations yield the following discretization for some non-linear operator \( \ell \) (once again we can consider \( \ell(y) = \max(0, y) \) or \( \ell(y) = \text{abs}(y) \)):

\[
\ell(y^h) = \ell \left( \sum_{i=1}^{n} y_i \xi_i \right)
\]

and

\[
(\ell(y^h), v^h) = \int_{\Omega} \ell(y^h)v^h \, dx \approx \sum_{T \in \mathbb{T}^h} \int_{T} \ell(y^h)v^h \, dx.
\]

The integrals over the elements \( T \in \mathbb{T}^h \) are approximated by some quadrature formula

\[
\sum_{T \in \mathbb{T}^h} \int_{T} \ell(y^h)v^h \, dx \approx \sum_{T \in \mathbb{T}^h} \sum_{k=1}^{n_k} \omega_k \ell \left( \sum_{i=1}^{n} y_i \xi_i(x_k) \right) \sum_{j=1}^{n} \xi_j(x_k),
\]

with \( n_k \) quadrature points per element \( T \) and corresponding weights \( \omega_k \).

Hence, the naturally arising discretization for some nonsmooth operator equation like (2.5b) in the finite element context is per quadrature point, which increases the total number of absolute value evaluations.

The switching function \( z \) as well as the signature function \( \bar{\sigma} \) are discretized similarly to the state \( y \), such that \( z^h, \bar{\sigma}^h \in V^h \) with the coefficient vectors \( \bar{\sigma} = (\bar{\sigma}_1, \ldots, \bar{\sigma}_n)^T \in \mathbb{R}^n \) and \( z = (z_1, \ldots, z_n)^T \in \mathbb{R}^n \).

As seen in Eq. (3.11), we would like to point out that the number of nonsmooth functions \( \ell \) in the discretized problem is per quadrature point. As already discussed in [29], this is not in perfect alignment with a representation of some finite element function like the state \( y \). The consequent choice of this discretization leads to an increase of the polynomial degree due to the multiplication \( \bar{\sigma}z \) in the discretized representation of the operator \( \ell \) opposed to the operator \( \ell \). However, this type of discretization allows for a straightforward implementation with FEniCS.

Then, the discretization of the optimization problem (2.2) where the variational inequality is approximated by the penalized equation Eq. (2.5b) is given by:

\[
\min_{(y^h, u^h) \in V^h \times U^h} J(y^h, u^h),
\]

s.t. \( (\nabla y^h, \nabla v^h) - \frac{1}{\ell} \text{max}(\psi^h - y^h, 0^h), v^h) = (u^h, v^h) + (f^h, v^h) \quad \forall v^h \in V^h
\]

Here, we set \( \psi^h = I^h \psi \), where \( I^h : C(\bar{\Omega}) \rightarrow V^h \) denotes the standard Lagrange interpolation operator. Due to the continuous embedding \( W^{2, \infty}(\Omega) \hookrightarrow C(\bar{\Omega}) \), we have that \( I^h \psi \) is well-defined. Further \( f^h := P^h f \), where \( P^h : L^2(\bar{\Omega}) \rightarrow U^h \) is the \( L^2 \)-projection onto the space of piecewise constant functions.
Inserting Eq. (3.10) into Eq. (3.12b) and replacing $v$ by $\xi$ leads to:

$$\int_\Omega \sum_{k=1}^n \nabla \xi_j(x) \cdot \nabla \xi_k(x) y_k + \ell \left( \sum_{i=1}^m y_i \xi_i(x) \right) \xi_j(x) \, dx$$

$$= \int_\Omega \left( \sum_{s=1}^m u_s e_{T_s}(x) \right) \xi_j(x) \, dx + \int_\Omega \left( \sum_{s=1}^m f_{T_s}(x) \right) \xi_j(x) \, dx,$$

(3.13)

for $1 \leq j \leq n$. Here $\ell(y)$ is given by $\ell(y) = \max(\psi^h - y, 0)$ and hence the second term on the left hand side is particularly given by

$$\int_\Omega \ell \left( \sum_{i=1}^n y_i \xi_i(x) \right) \xi_j(x) \, dx = \int_\Omega \max \left( \sum_{i=1}^n (\psi_i - y_i) \xi_i(x), 0 \right) \xi_j(x) \, dx.$$

By defining

$$A_{jk} := \int_\Omega \nabla \xi_j(x) \cdot \nabla \xi_k(x) \, dx = (\nabla \xi_j, \nabla \xi_k),$$

$$b_k(y^h) := \int_\Omega \ell \left( \sum_{i=1}^n y_i \xi_i(x) \right) \xi_k(x) \, dx$$

and

$$g_j := \int_\Omega \left( \sum_{s=1}^m u_s e_{T_s}(x) \right) \xi_j(x) \, dx + \int_\Omega \left( \sum_{s=1}^m f_{T_s}(x) \right) \xi_j(x) \, dx,$$

Eq. (3.13) can be rewritten as

$$\sum_{k=1}^n A_{jk} y_k + b_k(y^h) = g_j.$$

Here $A_{jk}$ represent the entries of the stiffness matrix $A$. The discretization of the PDE results in a non-linear system of algebraic equations, which we abbreviate as

$$Ay + b(y) = (u^T + f^T)E,$$

(3.14)

with the control matrix $E_{ij} := (e_{T_i}, \xi_j)$ and $y = (y_1, \ldots, y_n)^T$ denoting the finite-element approximation belonging to the right-hand side given by the discrete control $u$. To this end, the function $y^h|_{T_k}$ on the linear element $T_k$ is realized in terms of its values at the vertices of $T_k$. Note that in the above algebraic system the vector $u$ and the matrices $A, E$ are constant since they are independent of the unknowns $y_1, \ldots, y_n$. However, as previously mentioned, this non-linear algebraic equation is assumed to be based on a reasonable approximation of the integral via quadrature. The resulting discretized objective functional reads as

$$\min_{(y,u) \in \mathbb{R}^n \times \mathbb{R}^m} J(y,u) = \frac{1}{2} (y - y_d)^T M (y - y_d) + \frac{\alpha}{2} u^T D u.$$

Herein $M \in \mathbb{R}^{n \times n}$ denotes the mass matrix $M_{ij} = (\xi_i, \xi_j)$ and $D$ the control mass matrix with the entries $D_{ij} = (e_{T_i}, e_{T_j})$, where $D$ is a diagonal matrix because the interior of the triangles are disjoint to each other. Analogous to the previously deduced discretization, the discrete version of the compact abs-linearized problem Eq. (3.5) with the added penalty approximation introduced in Eq. (3.6) is given by:

$$\min_{(y^h,z^h,u^h) \in \mathbb{V}^h \times \mathbb{V}^h \times \mathbb{U}^h} J(y^h,u^h) + \nu \int_\Omega \max \left( -\sigma^h z^h, 0 \right)^4 \, dx$$

s.t. $$(\nabla y^h, \nabla v^h) - \frac{1}{\nu} \left( (\psi^h - y^h + \sigma^h z^h, v^h) = (u^h, v^h) + (f^h, v^h), \forall v^h \in \mathbb{V}^h$$

$$\left( \psi^h - y^h - z^h, v^h \right) = 0 \quad \forall v^h \in \mathbb{V}^h.$$
Under consideration of Eq. (3.15), it becomes clear that the inequality constraint from Eq. (3.5d) is enforced per quadrature point via our penalty approach. We assume that the optimization problem Eq. (3.15) fulfills some kind of constraint qualification to ensure the existence of the Lagrange multipliers. The corresponding discrete Lagrange functional related to the penalized compact abs-linearized problem of system Eq. (3.15) is now given by

\[
\mathcal{L}(y^h, z^h, u^h, \lambda_{PDE}^h, \lambda_z^h) = J(y^h, u^h) + (\nabla \lambda_{PDE}^h \nabla y^h) + \nu \int_{\Omega} \left( \max(-\bar{\sigma}^h z^h, 0) \right)^4 dx + (\lambda_{PDE}^h, -\frac{1}{2} \left( \psi^h - y^h + \bar{\sigma}^h z^h \right) - u^h - f^h) + (\lambda_z^h, \psi - y^h - z^h).
\]

The KKT system corresponding to Eq. (3.16) with e.g., \(\bar{\sigma}^h = (\bar{\sigma}_\psi)^h\) is then solved with a non-linear variational Newton solver.

## 4 Error Estimates

In this section we will prove an error estimate for the \(L^2\)-error of the control, i.e., for \(\|u_{\varepsilon,h}^* - u^*\|_{L^2(\Omega)}\) under the assumption that a quadratic growth condition holds. Here \(u^*\) and \(u_{\varepsilon,h}^*\) \(\in L^2(\Omega)\), respectively, denote locally optimal solutions of (2.2) and (3.12), respectively. In order to derive an error bound we adapt the technique introduced in [21]. The proof is based on a quadratic growth condition and the \(L^2\)-error estimates for the state presented in [11] and [24], respectively.

In order to simplify the notation we introduce the reduced functionals

\[
g : L^2(\Omega) \to \mathbb{R}, \quad g(u) := J(S(u), u) \\
g_\varepsilon : L^2(\Omega) \to \mathbb{R}, \quad g_\varepsilon(u) := J(S_\varepsilon(u), u) \\
g_{\varepsilon,h} : L^2(\Omega) \to \mathbb{R}, \quad g_{\varepsilon,h}(u) := J(S_{\varepsilon,h}(u), u).
\]

Moreover, let \(u^*, u_{\varepsilon}^*\) and \(u_{\varepsilon,h}^*\) \(\in L^2(\Omega)\) be locally optimal solutions of (2.2), (3.5) and (3.12), respectively.

At this point we want to emphasize that the discrete optimization problem

\[
\begin{align*}
\min_{(y^h, z^h, u^h) \in V^h \times V^h \times U^h} & \quad J(y^h, u^h) \\
\text{s.t.} & \quad (\nabla y^h, \nabla v^h) - \frac{1}{\rho} \left( \psi^h - y^h + \bar{\sigma}^h z^h, v^h \right) = (u^h, v^h) + (f^h, v^h), \quad \forall v^h \in V^h \\
& \quad (\psi^h - y^h - z^h, v^h) = 0 \quad \forall v^h \in V^h \\
& \quad \bar{\sigma}^h z^h \geq 0 \quad \text{a.e. in } \Omega
\end{align*}
\]

is just a reformulation of problem (3.12) and consequently also the optimal state as well as the optimal control of the two problems are the same. Since problem (3.12) is more convenient for the error analysis, we always consider problem (3.12) instead of (4.1) in this section. However, the reader should be aware that the error estimates are also valid for (4.1).

We make the following assumptions for our error analysis:

**Assumption 4.1.**

i) \(u, f \in L^\infty(\Omega)\)

ii) There holds a quadratic growth condition, i.e., there are \(\rho, \delta > 0\) such that

\[
g(u^*) \leq g(u) - \delta \|u - u^*\|_{L^2(\Omega)}^2 \quad \forall u \in B_\rho(u^*),
\]

where \(B_\rho(u^*) := \{ u \in L^2(\Omega) : \|u - u^*\|_{L^2(\Omega)} \leq \rho \} \).

**Remark 4.2.** For the obstacle problem the quadratic growth condition (4.2) holds if \(u^*\) satisfies some second-order sufficient optimality conditions (cf. [18]).
We start with an $L^2$-error estimate of the regularization error for the state, which is proven in [11] for a convex and polygonal domain and in [24] for a domain with $C^{1,1}$-boundary.

**Theorem 4.3.** Let $y$ and $y_\varepsilon$ be the solutions of (2.2b) and (2.5b), respectively. Then there holds the estimate

$$\|y - y_\varepsilon\|_{L^\infty(\Omega)} \leq C\varepsilon$$

with a constant $C > 0$ independent of $\varepsilon$.

Next, we state an $L^2$-error estimate of the discretization error for the state, which is proven in [11].

**Theorem 4.4.** Let $\Omega$ be a bounded domain which is convex and polygonal. Moreover, let $y_\varepsilon$ and $u_{\varepsilon,h}$ be the solutions of (2.5b) respectively (3.12b). In addition assume that the obstacle satisfies $\psi < 0$ on the boundary. Then there exist constants $\varepsilon_d > 0$ and $h_d > 0$ such that for all $\varepsilon \leq \varepsilon_d$ and $h \leq h_d$ there holds

$$\|u_\varepsilon - u_{\varepsilon,h}\|_{L^2(\Omega)} \leq C h^2 \|L^\infty(\Omega) + \|u\|_{L^\infty(\Omega)} + \|\Delta\psi\|_{L^\infty(\Omega)}$$

with a constant $C > 0$ independent of $h$.

**Remark 4.5.** The same convergence rate is also shown in [24], where the boundary is assumed to be smooth, i.e., of class $C^{1,1}$. It is worth noting that the condition $\psi < 0$ on the boundary is not necessary if the boundary is smooth.

We continue with some preparatory lemmas which are needed for our error analysis. Throughout the remainder of this paper let $(\varepsilon_n, h_n)_{n \in \mathbb{N}} \subset \mathbb{R}^2_{\geq 0}$ denote a sequence converging to zero.

**Lemma 4.6.** Assume that $\psi < 0$ on the boundary, $\varepsilon_n \leq \varepsilon_d$ and $h_n \leq h_d$ with $\varepsilon_d$ and $h_d$ as in Thm. 4.4. Let $\{u_{\varepsilon_n,h_n}\}$ be a sequence that converges strongly in $H^{-1}(\Omega)$ to $u \in H^{-1}(\Omega)$ as $n \to \infty$. Denote the solution of the discretized equation $(3.12b)$ corresponding to $u_{\varepsilon_n,h_n}$ by $y_{\varepsilon_n,h_n}$ and the solution of (2.2b) corresponding to $u$ by $y$. Then $y_{\varepsilon_n,h_n} \to y$ in $H^1_0(\Omega)$.

**Proof.** We have

$$\|y - y_{\varepsilon_n,h_n}\|_{H^1(\Omega)} = \|S(u) - S_{\varepsilon_n,h_n}(u_{\varepsilon_n,h_n})\|_{H^1(\Omega)}$$

$$\leq \|S(u) - S(u_{\varepsilon_n,h_n})\|_{H^1(\Omega)} + \|S(u_{\varepsilon_n,h_n}) - S_{\varepsilon_n,h_n}(u_{\varepsilon_n,h_n})\|_{H^1(\Omega)}.$$ 

Theorems 4.3 and 4.4 imply

$$\|S(u_{\varepsilon_n,h_n}) - S_{\varepsilon_n,h_n}(u_{\varepsilon_n,h_n})\|_{H^1(\Omega)} \leq C(\varepsilon_n + h_n^2 \|\log h_n\|^2) \xrightarrow{n \to \infty} 0.$$ 

Moreover by Thm. 2.5 we have

$$\|S(u) - S(u_{\varepsilon_n,h_n})\|_{H^1(\Omega)} \xrightarrow{n \to \infty} 0,$$

which concludes the proof.

**Lemma 4.7.** Suppose that $u^*$ satisfies the quadratic growth condition (4.2) Then there is a sequence $\{u_{\varepsilon_n,h_n}^*\}$ of locally optimal solutions to (3.15) with $u_{\varepsilon_n,h_n}^* \to u^*$ in $L^2(\Omega)$ as $n \to \infty$.

**Proof.** Based on Lem. 4.6 the following proof is standard (see also [21, Lem. 5.5]), where an analogous result is proven. Nevertheless, for later purpose and for convenience of the reader, we sketch the arguments. Following the classical localization argument from [6], we define the following discrete problems:

$$\min_{u \in B_\rho(u^*)} y_{\varepsilon_n,h_n}(u),$$

where $B_\rho(u^*)$ denotes the closed $L^2$-ball from (4.2). By standard arguments the above problem admits a globally optimal solution for every $n < \infty$, denoted by $u_{\varepsilon_n,h_n}^*$. Due to the constraint this sequence is bounded in $L^2(\Omega)$ and thus admits a weakly convergent subsequence with limit $\tilde{u} \in L^2(\Omega)$, which, by compact embedding, converges strongly in $H^1(\Omega)$. By Lem. 4.6 the associated states $y_{\varepsilon_n,h_n} := S_{\varepsilon_n,h_n}(u_{\varepsilon_n,h_n}^*)$ converge strongly to $\tilde{y} := S(\tilde{u})$. The weak lower semicontinuity of the objective along with the isolated local optimality of $u^*$ implies $\tilde{u} = u^*$. Moreover, the Tikhonov term in the objective yields the norm convergence of $u_{\varepsilon_n,h_n}^*$ so that $u_{\varepsilon_n,h_n}^* \to u^*$ in $L^2(\Omega)$. This implies that $u_{\varepsilon_n,h_n}^*$ is in the interior of $B_\rho(u^*)$ for $n$ sufficiently large and therefore, $u_{\varepsilon_n,h_n}^*$ is a local solution of (3.15).
Lemma 4.8. Let \( \{u_{\varepsilon,h}^\ast\} \) be the sequence of Lem. 4.7. Then \( \{u_{\varepsilon,h}^\ast\} \) is uniformly bounded in \( H^1(\Omega) \).

Proof. Analogously to [7, Cor. 4.5] one can derive the following optimality system for (3.15): For every locally optimal solution \( u_{\varepsilon,h}^\ast \) of (3.15) with associated state \( y_{\varepsilon,h}^\ast \), there exist an adjoint state \( p_{\varepsilon,h}^\ast \in V^h \) and a multiplier \( \mu_{\varepsilon,h}^\ast \in L^\infty(\Omega) \) such that

\[
\begin{align*}
0 & = (p_{\varepsilon,h}^\ast, v^h) + \langle \mu_{\varepsilon,h}^\ast p_{\varepsilon,h}^\ast, v^h \rangle = (y_{\varepsilon,h}^\ast - y_d, v^h) \quad \forall v^h \in V^h \quad \text{(4.4)} \\
\mu_{\varepsilon,h}^\ast(x) & \in \partial_c \max(y_{\varepsilon,h}^\ast(x)) \quad \text{a.e. in } \Omega \quad \text{(4.5)} \\
p_{\varepsilon,h}^\ast(x) + \alpha u_{\varepsilon,h}^\ast(x) & = 0 \quad \text{a.e. in } \Omega,
\end{align*}
\]

where \( \partial_c \max : \mathbb{R} \to [0, \frac{1}{\varepsilon}] \) denotes the convex subdifferential of the function \( \xi(y) = -\frac{1}{\varepsilon} \max(\psi - y, 0) \). Testing equation (4.4) with \( p_{\varepsilon,h}^\ast \in V^h \), the coercivity of \( a \) and Hölder’s inequality leads to

\[
\beta \|p_{\varepsilon,h}^\ast\|^2_{H^1(\Omega)} \leq (\nabla p_{\varepsilon,h}^\ast, \nabla p_{\varepsilon,h}^\ast) = (y_{\varepsilon,h}^\ast - y_d, p_{\varepsilon,h}^\ast) - (\mu_{\varepsilon,h}^\ast p_{\varepsilon,h}^\ast, p_{\varepsilon,h}^\ast) \\
\leq \|y_{\varepsilon,h}^\ast - y_d\|_{L^2(\Omega)} \|p_{\varepsilon,h}^\ast\|_{H^1(\Omega)} + \|\mu_{\varepsilon,h}^\ast\|_{L^\infty(\Omega)} \|p_{\varepsilon,h}^\ast\|^2_{H^1(\Omega)}.
\]

Due to Eq. (4.5) and \( \|\mu_{\varepsilon,h}^\ast\|_{L^\infty(\Omega)} \leq \frac{1}{\varepsilon} \), we arrive at

\[
\|u_{\varepsilon,h}^\ast\|_{H^1(\Omega)} = \frac{1}{\alpha} \|p_{\varepsilon,h}^\ast\|_{H^1(\Omega)} \leq \frac{1}{\alpha|\beta - \frac{1}{\varepsilon}|} \|y_{\varepsilon,h}^\ast - y_d\|_{L^2(\Omega)} \|u_{\varepsilon,h}^\ast\|_{L^2(\Omega)} + \|y_d\|_{L^2(\Omega)}.
\]

The boundedness of \( y_{\varepsilon,h}^\ast \) in \( L^2(\Omega) \) implies the claim. \( \square \)

Theorem 4.9. Suppose that \( u^\ast \) satisfies the quadratic growth condition (4.2) and \( \psi < 0 \) on the boundary. Then there exist constants \( \varepsilon_d > 0 \) and \( h_d > 0 \) such that for all \( \varepsilon \leq \varepsilon_d \) and \( h \leq h_d \) one has

\[
\|u_{\varepsilon,h}^\ast - u^\ast\|_{L^2(\Omega)} \leq C \left( \varepsilon^{1/2} + h \log h \right)
\]

with constant \( C > 0 \) independent of \( \varepsilon \) and \( h \).

Proof. The proof follows the lines of [21, Thm. 5.8]. As seen in the proof of Lem. 4.7, \( u_{\varepsilon,h}^\ast \) is a global solution of (4.3) and therefore

\[
g_{\varepsilon,h}(u_{\varepsilon,h}^\ast) \leq g_{\varepsilon,h}(u^\ast) \quad \text{(4.6)}
\]

Moreover, for \( \varepsilon \) and \( h \) sufficiently small, we have \( u_{\varepsilon,h}^\ast \in B_{\mu}(u^\ast) \). Therefore the quadratic growth condition (4.2) and (4.6) imply

\[
\delta \|u_{\varepsilon,h}^\ast - u^\ast\|^2_{L^2(\Omega)} \leq g(u_{\varepsilon,h}^\ast) - g_{\varepsilon,h}(u_{\varepsilon,h}^\ast) + g_{\varepsilon,h}(u^\ast) - g(u^\ast) + g_{\varepsilon,h}(u_{\varepsilon,h}^\ast) - g_{\varepsilon,h}(u^\ast) \\
\leq |g(u_{\varepsilon,h}^\ast) - g_{\varepsilon,h}(u_{\varepsilon,h}^\ast)| + |g_{\varepsilon,h}(u^\ast) - g(u^\ast)| \quad \text{(4.7)}
\]

We split the first term of (4.7) into two terms

\[
|g(u_{\varepsilon,h}^\ast) - g_{\varepsilon,h}(u_{\varepsilon,h}^\ast)| \leq |g(u_{\varepsilon,h}^\ast) - g_{\varepsilon}(u_{\varepsilon,h}^\ast)| + |g_{\varepsilon}(u_{\varepsilon,h}^\ast) - g_{\varepsilon,h}(u_{\varepsilon,h}^\ast)| \quad \text{(4.8)}
\]

For the first term in (4.8) we get

\[
|g(u_{\varepsilon,h}^\ast) - g_{\varepsilon}(u_{\varepsilon,h}^\ast)| = \frac{1}{2} \left| \left| S(u_{\varepsilon,h}^\ast) - y_d \right|_{L^2(\Omega)} - \left| S_{\varepsilon}(u_{\varepsilon,h}^\ast) - S(u_{\varepsilon,h}^\ast) - y_d \right|_{L^2(\Omega)} \right| \\
\leq \frac{1}{2} \left| S_{\varepsilon}(u_{\varepsilon,h}^\ast) - S(u_{\varepsilon,h}^\ast) \right|_{L^2(\Omega)} + \left| S_{\varepsilon}(u_{\varepsilon,h}^\ast) - S(u_{\varepsilon,h}^\ast) \right|_{L^2(\Omega)} \left| S(u_{\varepsilon,h}^\ast) - y_d \right|_{L^2(\Omega)}.
\]

The boundedness of \( \{u_{\varepsilon,h}^\ast\} \) in \( H^1(\Omega) \) by Lem. 4.8 and the Lipschitz continuity of the operator \( S \) (cf. Lem. 2.1) imply that \( \|S(u_{\varepsilon,h}^\ast) - y_d\|_{L^2(\Omega)} \) is bounded. Hence, by Thm. 4.3 we obtain

\[
|g(u_{\varepsilon,h}^\ast) - g_{\varepsilon}(u_{\varepsilon,h}^\ast)| \leq C \varepsilon.
\]
Analogously we obtain for the second term in Eq. (4.8) the estimate
\[ |g_z(u^*_{e,h}) - g_z(u^*_{e,h})| \leq \frac{1}{2} \| S_z(u^*_{e,h}) - S_z(u^*_{e,h}) \|_{L^2(\Omega)}^2 + \| S_z(u^*_{e,h}) - S_z(u^*_{e,h}) \|_{L^2(\Omega)} \| S_z(u^*_{e,h}) - g_y \|_{L^2(\Omega)}. \]
Note that \( \| S_z(u^*_{e,h}) - g_y \|_{L^2(\Omega)} \) is bounded due to the boundedness of \( \{u^*_{e,h}\} \) in \( H^1(\Omega) \) and the Lipschitz continuity of the operator \( S_z \) (cf. Lem. 2.3). Thus Thm. 4.4 implies
\[ |g_z(u^*_{e,h}) - g_z(u^*_{e,h})| \leq C h^2 | \log h |^2. \]
Applying the same arguments to the second term of (4.7) completes the proof. \( \square \)

The previous theorem implies the following result:

**Corollary 4.10.** Let \( \varepsilon = C h^2 \leq \varepsilon_d \) for \( C > 0 \) arbitrary. Under the assumptions that \( u^* \) satisfies the quadratic growth condition (4.2) and \( \psi < 0 \) on the boundary, there exists \( h_d > 0 \) such that for all \( h \leq h_d \) it holds
\[ \| u^*_{e,h} - u^* \|_{L^2(\Omega)} \leq C h | \log h |. \]

## 5 Numerical Results

In this section we test the performance of the algorithm CALi for the numerical solution of optimization problems of the form Eq. (2.2). For this purpose we present three different test examples taken from [20], [15] and [2]. In all three examples the computational domain is chosen as the unit square \( \Omega = (0,1)^2 \) and for the fixed \( \bar{\sigma} \) we use \( \bar{\sigma} = \bar{\sigma}_e \) as introduced in Def. 3.3. For all tests we take \( y^0 = z^0 = u^0 \equiv 0 \) as initial guess for Newton’s method. We initialize our \( \varepsilon \)-homotopy with \( \varepsilon = 1.0 \) and decrease the value of the penalization parameter constantly until the linear system in Newton’s method is too ill-conditioned and Newton’s method does not converge in under 20 steps, where an absolute error of \( 10^{-12} \) is pursued within the respective Newton procedure. For each \( \varepsilon < 1.0 \) we take the solution of the compact abs-linearized and penalized problem, i.e., Eq. (3.6), at the preceding value of \( \varepsilon \) as starting value for the current Newton iteration.

Besides the number of Newton steps we also present the value \( \| \bar{\sigma} z - |z| \|_{L^2(\Omega)} \), which measures the violation of the condition \( \bar{\sigma} z = |z| \). Furthermore, following [2] the value \( \mu^- := \min_{k \in \mathcal{N}^-} (y_k - \psi_k) \), with \( \mathcal{N}^- := \{ k \in 1, \ldots, n : y_k - \psi_k < 0 \} \) is also documented, which denotes the violation of the obstacle constraint \( y \geq \psi \) a.e. in \( \Omega \). As the penalty parameter \( \varepsilon \) decreases \( \mu^- \) typically should tend to zero. We use the finite-element discretization introduced in Sec. 3. All the computations are done within the open source finite element environment FEniCS, version 2019.1.0, using the Python interface.

As a first example we consider once again Exam. 3.5.

**Example 5.1.** The obstacle problem is constructed with
\[ y_d(x_1,x_2) = -\sin(\pi x_1) \sin(\pi x_2), \quad f(x_1,x_2) = -2\pi^2 \sin(\pi x_1) \sin(\pi x_2), \quad \psi(x_1,x_2) = -0.25 \]
We have already discussed that the optimal state is given by
\[ y^* = \max(\psi, y_d) = \begin{cases} \psi, & \text{on } \Omega^+: = \{ x \in \Omega : y_d(x) < \psi(x) \} \\ y_d, & \text{on } \Omega^-: = \{ x \in \Omega : y_d(x) \geq \psi(x) \} \end{cases} \]
with \( y^*|_{\partial \Omega} = 0 \) and \( u^* \equiv 0 \).

The numerical results for this example considering different values for the parameters \( \alpha, \varepsilon \) and different mesh sizes are provided in Tab. 1. As expected we see that \( \mu^- \) tends to zero as \( \varepsilon \) decreases and the value \( \| \bar{\sigma} z - |z| \|_{L^2(\Omega)} \) is quite small. Moreover, we observe that except for \( \varepsilon = 1 \) only 2 Newton steps are needed to solve the problem.

Fig. 2 illustrates the desired state \( y_d \), the fixed \( \bar{\sigma}_e \), the solutions \( y, \lambda_{\partial \Omega} \) and \( u \) obtained with CALi using the parameters \( \alpha = 1.0, \varepsilon = 1e-06 \) and \( h = 7.728e-03 \) as well as the exact solution \( y^* = \max(y_d, \psi) \).
Figure 1: Convergence plot of the obstacle violation $|\mu^-|$ and the $L^2$-error for the state $y$ for Exam. 5.1 with $h = 5.05e-03$ and $\alpha = 1.0$ for decreasing $\varepsilon$ values.

Figure 2: Solution $y$ with adjoint $\lambda_{PDE}$ and control $u$ for Exam. 5.1 with $y_d$ and $\max(\psi, y_d)$ for $\sigma \equiv \sigma_{\psi}$. 
The data for the second example are chosen as:

\[ y_d = y^* + \xi^* - \alpha \Delta y^*, \quad f = -\Delta y^* - y^* - \xi, \quad \psi = 0 \]

with

\[ y^* = \begin{cases} 
160(x_1^3 - x_1^2 + 0.25x_1)(x_2^3 - x_2^2 + 0.25x_2), & \text{in } (0, 0.5)^2 \\
0, & \text{else}
\end{cases} \]

and

\[ \xi^* = \max \left(0, -2|x_1 - 0.8| - 2|x_1x_2 - 0.3| + 0.5\right), \]

according to [14, Exam. 5.1]. Note that by construction the optimal control is \( u^* = y^* \). The numerical results for this example considering different values for the parameters \( \alpha, \varepsilon \) and different

<table>
<thead>
<tr>
<th>( h )</th>
<th>( \alpha )</th>
<th>( \varepsilon )</th>
<th>( |y - y^*|_{L^2(\Omega)} )</th>
<th>( |\bar{\sigma}z - z|_{L^2(\Omega)} )</th>
<th>( \mu^- )</th>
<th># Newton</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.009e-02</td>
<td>1.0</td>
<td>1.0</td>
<td>2.17e-01</td>
<td>1.3e-05</td>
<td>-5.295e-01</td>
<td>6</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1.0</td>
<td>1e-02</td>
<td>7.09e-02</td>
<td>1.8e-05</td>
<td>-1.570e-02</td>
<td>2</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1.0</td>
<td>1e-04</td>
<td>3.87e-03</td>
<td>2.2e-05</td>
<td>-1.935e-03</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1.0</td>
<td>1.0</td>
<td>2.16e-01</td>
<td>1.6e-06</td>
<td>-7.188e-01</td>
<td>6</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1.0</td>
<td>1e-02</td>
<td>7.09e-02</td>
<td>1.2e-04</td>
<td>-5.281e-02</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1.0</td>
<td>1e-04</td>
<td>1.55e-03</td>
<td>2.1e-05</td>
<td>-1.970e-03</td>
<td>2</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-01</td>
<td>1.0</td>
<td>3.01e-01</td>
<td>1.3e-05</td>
<td>-7.181e-01</td>
<td>8</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-01</td>
<td>1e-02</td>
<td>7.09e-02</td>
<td>1.7e-04</td>
<td>-1.570e-02</td>
<td>2</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-01</td>
<td>1e-04</td>
<td>3.88e-03</td>
<td>2.3e-05</td>
<td>-1.969e-03</td>
<td>2</td>
</tr>
<tr>
<td>2.210e-02</td>
<td>1e-01</td>
<td>1.0</td>
<td>3.02e-01</td>
<td>2.7e-06</td>
<td>-7.193e-01</td>
<td>8</td>
</tr>
<tr>
<td>2.210e-02</td>
<td>1e-01</td>
<td>1e-02</td>
<td>7.10e-02</td>
<td>1.6e-04</td>
<td>-1.571e-02</td>
<td>2</td>
</tr>
<tr>
<td>2.210e-02</td>
<td>1e-01</td>
<td>1e-04</td>
<td>6.17e-03</td>
<td>2.0e-05</td>
<td>-1.971e-03</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-01</td>
<td>1.0</td>
<td>3.02e-01</td>
<td>1.5e-06</td>
<td>-7.194e-01</td>
<td>8</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-01</td>
<td>1e-02</td>
<td>7.10e-02</td>
<td>1.7e-03</td>
<td>-1.571e-02</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-01</td>
<td>1e-04</td>
<td>6.23e-03</td>
<td>2.1e-05</td>
<td>-1.971e-03</td>
<td>2</td>
</tr>
<tr>
<td>7.728e-03</td>
<td>1e-01</td>
<td>1.0</td>
<td>3.02e-01</td>
<td>6.3e-07</td>
<td>-7.194e-01</td>
<td>8</td>
</tr>
<tr>
<td>7.728e-03</td>
<td>1e-01</td>
<td>1e-02</td>
<td>7.10e-02</td>
<td>1.7e-03</td>
<td>-1.571e-02</td>
<td>2</td>
</tr>
<tr>
<td>7.728e-03</td>
<td>1e-01</td>
<td>1e-04</td>
<td>7.24e-03</td>
<td>2.4e-05</td>
<td>-1.970e-03</td>
<td>2</td>
</tr>
<tr>
<td>7.728e-03</td>
<td>1e-01</td>
<td>1e-06</td>
<td>1.44e-03</td>
<td>3.0e-08</td>
<td>-9.987e-04</td>
<td>2</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-03</td>
<td>1.0</td>
<td>3.11e-01</td>
<td>9.2e-07</td>
<td>-7.397e-01</td>
<td>11</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-03</td>
<td>1e-02</td>
<td>8.37e-02</td>
<td>9.5e-04</td>
<td>-1.938e-01</td>
<td>2</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-03</td>
<td>1e-04</td>
<td>6.13e-03</td>
<td>2.1e-05</td>
<td>-2.714e-03</td>
<td>2</td>
</tr>
<tr>
<td>3.009e-02</td>
<td>1e-03</td>
<td>1e-06</td>
<td>4.44e-03</td>
<td>3.2e-08</td>
<td>-1.006e-03</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-05</td>
<td>1.0</td>
<td>3.16e-01</td>
<td>9.3e-07</td>
<td>-7.500e-01</td>
<td>11</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-05</td>
<td>1e-02</td>
<td>2.82e-01</td>
<td>1.3e-06</td>
<td>-6.679e-01</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-05</td>
<td>1e-04</td>
<td>6.13e-03</td>
<td>2.0e-05</td>
<td>-2.714e-03</td>
<td>2</td>
</tr>
<tr>
<td>1.571e-02</td>
<td>1e-05</td>
<td>1e-06</td>
<td>4.07e-03</td>
<td>3.2e-08</td>
<td>-1.006e-03</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: Numerical results for Exam. 5.1.

Example 5.2. The data for the second example are chosen as:

\[ y_d = y^* + \xi^* - \alpha \Delta y^*, \quad f = -\Delta y^* - y^* - \xi, \quad \psi = 0 \]

with

\[ y^* = \begin{cases} 
160(x_1^3 - x_1^2 + 0.25x_1)(x_2^3 - x_2^2 + 0.25x_2), & \text{in } (0, 0.5)^2 \\
0, & \text{else}
\end{cases} \]
mesh sizes are provided in Tab. 2. Following [14] the results in Tab. 2 were computed with a Newton solver tolerance of $\frac{\varepsilon^2}{2}$. We observe that for all combinations of parameter values except for $\varepsilon = 1.0$ only one Newton iteration is required to obtain a residual norm below $10^{-12}$.

Figure 3: Solution $y$ (upper left), control $u$ (lower right) and adjoint $\lambda_{PDE}$ (lower left) for Exam. 5.2 with $\xi^*$ (upper right) for $\alpha = 1.0$ and $\varepsilon = 10^{-8}$.

In [14] this example has the title “lack of strict complementarity” due to the fact that the active set at the solution contains a subset where strict complementarity fails to hold, i.e., the biactive set has a positive measure. It is precisely this lack of strict complementarity that poses a challenge, since the active constraint gradients at the solution are linearly dependent.
Table 2: Numerical results for Exam. 5.2.

The numerical solutions for the parameters $\alpha = 1.0$ and $\epsilon = 10^{-8}$ are displayed in Fig. 3. We would like to point out that the solution of the state in Fig. 3 differs from Fig. 2 in [14] by a factor of $10^{-1}$ since the scaling factor is missing at the corresponding axis in [14, Fig. 2].

Fig. 4 shows the decay of the obstacle violation $|\mu^-|$ for Exam. 5.2 with $h = 7.728e-03$ and $\alpha = 1.0$ for decreasing $\epsilon$ values in a log-log-scale.
The theoretical results of Sec. 4 show that the overall error consists of two contributions, namely the regularization error, i.e. $\|y^*_\varepsilon - y^*\|_{L^2(\Omega)}$ resp. $\|u^*_\varepsilon - u^*\|_{L^2(\Omega)}$, and the discretization error, i.e. $\|y_{\varepsilon,h}^* - y^*\|_{L^2(\Omega)}$ resp. $\|u_{\varepsilon,h}^* - u^*\|_{L^2(\Omega)}$. In order to numerically ascertain the approximation properties of the presented solution method, these errors were computed for different penalization parameters and different mesh sizes. Fig. 5 shows the convergence plot for the $L^2$-regularization-errors of the state and the control with fixed mesh size $h$ and decreasing $\varepsilon$ values for Exam. 5.2 in a log-log-scale. Representational for the numerical test, Fig. 5 suggests an approximation order of $O(\varepsilon)$ for the $L^2$-error of the state and $O(\varepsilon^{1/2})$ for the $L^2$-error of the control which confirms our theoretical results derived in Sec. 4.

Moreover, we observed that for a penalty parameter $\varepsilon \leq h^2$, the approximation error remained...
almost constant, i.e., the approximation error is dominated by the discretization error provided that \( \varepsilon \) is sufficiently small, see e.g., Fig. 6. This observation is in agreement with the theory of Sec. 4.

\[ \| y_{\varepsilon} - y^* \| \]
\[ h = 3.009e-2 \]
\[ h = 1.571e-02 \]
\[ h = 7.728e-03 \]
\[ h = 4.562e-03 \]

Figure 6: Convergence plot of the \( L^2 \)-error of the state \( y \) for Exam. 5.2 with \( \alpha = 1.0 \) for decreasing \( \varepsilon \) for different mesh sizes \( h \).
Figure 7: Convergence plot of the $L^2$-error of the control $u$ for Exam. 5.2 with $\alpha = 1.0$ for decreasing $\varepsilon$ for different mesh sizes $h$.

Fig. 8 shows convergence plots for the $L^2$-errors $\|u_{\varepsilon,h}^* - y^*\|_{L^2(\Omega)}$ and $\|u_{\varepsilon,h}^* - u^*\|_{L^2(\Omega)}$ for Exam. 5.2 in a log-log-scale for decreasing mesh size $h$ and fixed parameter $\varepsilon = 10^{-9}$. The observations in Fig. 8 suggest an approximation order of $O(h)$ for the $L^2$-error of the control and $O(h^2)$ for the $L^2$-error of the state. Similar observations were made for other test cases, such that once again the theoretically determined results of Sec. 4 have been numerically verified.
Figure 8: Convergence plot of the discretization error for Exam. 5.2, i.e., $L^2$-error of the control $u$ (left) and the state $y$ (right) with $\alpha = 1.0$ and $\varepsilon = 10^{-9}$ for decreasing mesh size $h$.

The following example has a special feature in contrast to the previous ones, since the function $\max(\psi, y_d)$ does not provide the optimal signature function $\bar{\sigma}$ and thus $\bar{\sigma}_\psi$ does not yield the compact abs-linearized (CAL) problem formulation which provides the optimal solution $y^*$. We present this example as an outlook for further research on optimal strategies for generating $\bar{\sigma}$ or a sequence of $\bar{\sigma}$-signature functions together with an efficient switching strategy such that the corresponding final CAL problem provides the optimal solution.

Example 5.3. For this example of an obstacle problem we choose the following data:

\[
y_d(x_1, x_2) = -5x_1 - x_2 + 1, \quad f(x_1, x_2) = -x_1 + 0.5, \quad \psi(x_1, x_2) = 0.0.
\]

This example corresponds to Exam. 2 from [2]. The signature function $\bar{\sigma}$ was chosen as

\[
\bar{\sigma}(x_1, x_2) = \begin{cases} 
-1, & x_1 \leq 0.5 \\
+1, & \text{else}
\end{cases}
\text{ for } (x_1, x_2) \in \Omega.
\]

The numerical results for this example considering different values for the parameters $\alpha, \varepsilon$ and different mesh sizes are provided in Tab. 3. Once again, strict complementarity, however, is not satisfied, which makes this problem a further challenge.
Table 3: Numerical results for Exam. 5.3 with a Newton tolerance of $10^{-15}$.

Similar to [2] we observe a reduction in the absolute obstacle violation proportional to the reduction in the penalty parameter $\varepsilon$, see e.g., Fig. 9.
Although the procedure presented here and the corresponding algorithm works impeccably for a large class of optimization problems with obstacle conditions and allows an explicit structure exploitation, which records a reduction in the number of required Newton steps, there are also optimization problems like Exam. 5.3 where the choice and fixation of $\bar{\sigma}$ does not lead directly to the optimal CAL problem without further analysis and effort.

6 Conclusion

We investigated a regularization approach for obstacle optimization problems which results in optimal control problems constrained by a genuinely nonsmooth PDE. The presented and discussed solution method for this class of optimization problems is based on compact abs-linearization which enables the optimization without any substitute assumptions for the nonsmoothness. The key idea is to generate a suitable reformulation of the nonsmooth PDE constrained regularized problem, the so called compact abs-linearized problem which can be solved using conventional methods for smooth optimization problems.

The type of discretization employed here was also presented and critically examined. Moreover,
error estimates for the state and the control are derived, which contain information about the coupling of the regularization and the mesh size.

Finally, three different obstacle problems were considered and numerical results illustrating the performance of the presented algorithm were demonstrated and evaluated. The corresponding numerical results are very promising and also clearly confirm the theoretically derived error estimates.

The analysis and the numerical results have shown that it is useful as well as purposeful to rewrite the considered obstacle optimization problem by means of penalization and compact abs-linearization into a smooth but strongly related problem. By profoundly choosing the signature function $\bar{\sigma}$ one obtains a subproblem of the penalized nonsmooth problem, which is itself smooth and can be solved efficiently and effectively by means of standard optimization methods. In this paper we have shown that in certain cases the choice of $\bar{\sigma}_\psi = \text{sign}(\psi - y_d)$ provides the optimal signature function. However, the choice of the signature function $\bar{\sigma}$ seems to be a delicate issue in general.

We therefore suggest and consider further research on a strategy for a generally advantageous choice of $\bar{\sigma}$ as well as a strategy for switching from one $\bar{\sigma}^i$, i.e., a concrete subproblem to the next $\bar{\sigma}^{i+1}$ by cleverly switching certain signs on certain areas in the underlying domain. By means of such a procedure the compact abs-linearization can be successively applied. The resulting problems are smooth again and can be solved as before with the usual methods of smooth optimization.

Another aspect of further research comprises the consideration of suitable regularization methods for optimization problems constrained by variational inequalities of the second kind into similar nonsmooth PDE constrained optimization problems and applying the algorithm CALi for their efficient solution.

References


