Deutsche Forschungsgemeinschaft
Priority Programme 1962

Solving Non-Smooth Semi-Linear Optimal Control Problems with Abs-Linearization

Olga Weiß, Stephan Schmidt, Andrea Walther

SPP 1962 | Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-093r

received on April 14, 2020

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

Solving Non-Smooth Semi-Linear Elliptic Optimal Control Problems with Abs-Linearization

Olga Weiß[1], Stephan Schmidt[1], Andrea Walther[1]

[1]Institut für Mathematik, Humboldt-Universität zu Berlin

March 19, 2020

Abstract

We investigate optimization problems with a non-smooth partial differential equation as constraint, where non-smoothness is assumed to be caused by the functions abs, min and max. For the efficient as well as robust solution of such problems, we propose a new optimization method based on abs-linearization, i.e., a special handling of the non-smoothness without regularization. The key idea of this approach is the determination of stationary points by an appropriate decomposition of the original non-smooth problem into several smooth so-called branch problems. Each of these branch problems can be solved by classical means. The exploitation of corresponding optimality conditions for the smooth case identifies the next branch and thus yields a successive reduction of the objective value. This approach is able to solve the considered class of non-smooth optimization problems without any regularization of the non-smoothness and additionally maintains reasonable convergence properties. Numerical results for non-smooth optimization problems illustrate the proposed approach and its performance.

Keywords: Non-Smooth Optimization, Abs-Linearization, PDE Constrained Optimization, Non-Smooth PDE, Elliptic Optimal Control Problem

1 Introduction

Non-smooth PDE-constrained optimization problems are known to be difficult to handle, theoretically as well as algorithmically. The difficulty usually lies in the fact that no adjoint equation in the classical sense can be derived, which has a direct consequence on the development of algorithms, since no reduced gradient is available for first-order methods. In this paper we assume that the non-smoothness in the semi-linear elliptic state equation is caused by a non-smooth superposition operator which can be decomposed into a finite number of smooth functions and non-smooth Lipschitz continuous operators abs, min and max. The presented algorithm takes advantage of this structural assumption and specifically exploits the non-smooth structure in the interest of solving the underlying optimization problem.

Non-smooth optimization problems with a partial differential equation (PDE) as constraint that involves the mentioned non-smooth non-linear functions arise in many modern applications. For example, a corresponding semi-linear elliptic partial differential equation describes the deflection of a stretched thin membrane partially covered by water, see [17]. Furthermore, a similar non-smooth partial differential equation arises in free boundary problems for a confined plasma, see, e.g. [17, 20]. Even nowadays, the optimization of such problems is challenging. Therefore, often either the non-smoothness is regularized, i.e., the non-differentiable term is replaced by a suitable smooth approximation to avoid dealing with the non-smoothness (see e.g. [3] and [8]) to apply an algorithm suitable for smooth optimization or the semi-smooth Newton method is used. For example, in [6] a method from a semi-smooth Newton method is proposed to solve a specific non-smooth optimization problem including the max operator.

Despite the fact that the model problems considered in this article are of a less demanding nature as for instance the ones considered in [20], their treatment is an essential step in understanding the problem class and developing an appropriate structure-exploiting algorithm. For this purpose, we propose an alternative algorithm that is not based on the semi-smooth Newton method and that explicitly exploits the non-smoothness.
In the finite dimensional setting the unconstrained minimization of piecewise smooth functions by successive abs-linearization without any regularization for the non-smoothness was studied by Griewank, Walther and co-authors in [10, 12, 13] and related work. There, it is always assumed that the non-smoothness of the considered optimization problem stems from evaluations of the absolute value function only. Using well-known reformulations, this covers the maximum and the minimum functions as well as complementarity problems. In [24] we already extended and adapted the algorithmic idea of the approach in finite dimensions to the infinite dimensional case, i.e., to PDE-constrained optimization problems with non-smooth objective functionals. Although the algorithm SALMIN presented in [24] can also handle non-smooth optimization problems in function spaces by explicitly exploiting the non-smooth structure, it is not applicable to the optimization problems considered in this paper. The main difficulty involves the already mentioned challenge that the non-smoothness appears in the state equation and thus no adjoint equation in the classical sense as well as no classical reduced problem formulation can be derived. It is also important to note that the local model generated in [24] for the non-smooth case does not support the classical chain rule. Hence, one cannot directly handle the reduced unconstrained formulation. Therefore, we propose here a penalty-based approach to treat the PDE constraint explicitly. Nevertheless, we follow the idea for the finite dimensional case in that the key point of the optimization method under consideration is the location of stationary points by an appropriate decomposition of the original problem into smooth so-called branch problems. Each of these branch problems can be solved by classical methods for smooth PDE-constrained optimization. Then, the exploitation of standard optimality conditions for the smooth case determines the next branch problem and ensures the reduction of the target function value. In deriving necessary optimality conditions, the difficulty lies in the fact that, while the solution domain of the PDE is compact, the number and location of the solutions is unknown. For this reason, a direct approach, i.e., first-discretize-then-optimize, is presented for the numerical solution of the optimization problems.
The paper is organized as follows. In Sec. 2, we introduce the considered problem class, discuss its properties and propose a reformulation of the first order necessary optimality conditions. The resulting smooth branch problems will be presented in Sec. 3. This includes a solution approach involving a penalty term and an analysis of the corresponding optimality conditions. Sec. 4 summarizes the resulting optimization algorithm. Furthermore, the chosen discretization approach as well as the corresponding solution of the subproblems is discussed. Numerical results for a collection of test problems are presented and analysed in Sec. 5. Finally, a conclusion and an outlook are given in Sec. 6.

2 The Problem Class, its Properties and a Reformulation

In order to illustrate ideas, we consider the following prototypical example in the further course of this paper, where we focus on real valued functions defined on a Lipschitz domain $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$. As a model problem we consider PDE-constrained optimization problems of the form

$$\begin{aligned} \min_{(y,u)\in H_0^1(\Omega)\times L^2} \quad & j(y) + \frac{\alpha}{2}\|u\|_{L^2}^2 \\ \text{s.t.} \quad & -\Delta y + \ell(y) - u = 0 \quad \text{in } \Omega \end{aligned} \tag{1}$$

with a convex and twice continuously Fréchet differentiable functional $j : H_0^1(\Omega) \to \mathbb{R}$ and a semi-linear elliptic PDE constraint.
The special and at the same time challenging feature of Eq. (1) is the non-smoothness in the state equation which is caused by the non-smooth operator $\ell : H_0^1(\Omega) \to L^2(\Omega)$. For the exact assumptions on the model problem as well as the definition of the operator ℓ we refer to Assumption 2.1 given next. Throughout the paper, we assume that the model problem Eq. (1) has the following properties:

Assumptions 2.1.

i. The domain $\Omega \subset \mathbb{R}^n, n \in \mathbb{N}$, *is a Lipschitz domain.*

ii. The functional $j : H_0^1(\Omega) \to \mathbb{R}$ *is convex, twice continuously Fréchet differentiable and bounded from below.*

iii. The operator $\ell : H_0^1(\Omega) \to L^2(\Omega)$ *denotes the Nemytskij operator induced by an operator, which is bounded and measurable in* $x \in \Omega$ *for every fixed* y*, monotone in* y *for almost every* $x \in \Omega$ *and locally Lipschitz continuous.*

iv. The operator ℓ *can be expressed as finite composition of the absolute value function and Fréchet differentiable operators.*

v. The constant $\alpha > 0$ *is a given Tikhonov parameter.*

We will denote the Nemitzkii operator as well as the inducing operator having other domains and regions with the same symbol, ℓ. Note that point iv. in Ass. 2.1 refers to the fact that the Lipschitz-continuous operator ℓ can be given by the structured evaluation presented in Def. 2.4.

The Solution Operator

The solution or control-to-state operator $S(u) = y$ corresponding to the non-smooth state equation given in Eq. (1) plays an impotent role in the analysis of the overall optimization problem. Therefore we will state here some main properties and results:

Lemma 2.2 (The solution operator)**.** *Let* $\ell : H_0^1(\Omega) \to L^2(\Omega)$ *be a non-smooth operator satisfying Ass. 2.1 and* $S(u) = y$ *the solution operator associated with the PDE in Eq. (1). Then* S *has the following properties.*

i. The control-to-state operator $S(u) = y$ *associated with Eq. (1) is a non-smooth operator.*

ii. S *is well-defined, bijective and globally Lipschitz continuous as a function from* $L^2(\Omega)$ *to the image space* $\{v \in H_0^1(\Omega) | \Delta v \in L^2(\Omega)\}$.

iii. S *is directionally but not Gâteaux differentiable*

iv. $S : L^2 \to \{v \in H_0^1(\Omega) | \Delta v \in L^2(\Omega)\}$ *is Hadamar directional differentiable for all points and in all directions in* $L^2(\Omega)$.

Proof. For the proof of these assertions, one can use similar arguments as in [6, Prop. 2.1]. For a detailed proof we refer the reader to [25]. □

On the Model Problem

A frequently used functional j as part of the cost functional is the tracking-type functional

$$j(y) = \frac{1}{2}\|y - y_d\|_{L^2}^2$$

with some given function y_d usually denoting a given desired state. In the further course we will therefore concentrate on the following optimization problem

$$\begin{aligned} \min_{(y,u)\in H_0^1(\Omega)\times L^2(\Omega)} \ & \mathcal{J}(y,u), \quad \text{with } \mathcal{J}(y,u) = \frac{1}{2}\|y-y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \\ \text{s.t.} \quad & -\Delta y + \ell(y) - u = 0 \ \text{ in } \Omega\,. \end{aligned} \tag{2}$$

However, this does not impose any restrictions. It should be noted that the algorithm proposed in Sec. 4 of this paper is not limited to this specific class of semi-linear PDEs or this kind of objective functionals. Instead, the arguments can easily be adapted to more general cases with, for example, a general linear elliptic differential operator of second order instead of the Laplacian operator, as well as to the more general objective functional considered in Eq. (1). Therefore, Sec. 5 presents also numerical results for other differential operators. Nevertheless, to illustrate the idea of the algorithm we restrict ourselves here to this class of semi-linear elliptic PDEs.
In addition to the assumptions on the non-smooth state equation given in Ass. 2.1, it can easily be observed that the tracking type objective functional $\mathcal{J} : H_0^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ in Eq. (2) is weakly lower semi-continuous and twice continuously Fréchet-differentiable.
One particular example of this class of model problems of non-smooth semi-linear elliptic optimal control problems, where $\ell(y) = \max(0, y)$, can be found in [6]. There the authors also show that the resulting non-smooth control-to-state operator is directionally differentiable. They also precisely characterize its Bouligand subdifferentials, derive first-order optimality conditions using the Bouligand subdifferentials and use the directional derivative of the control-to-state mapping to establish strong stationarity conditions.

Lemma 2.3. *For all $u \in H^{-1}(\Omega)$ the PDE of the optimization problem (2) is non-linear, well posed and has a unique solution $y \in H_0^1(\Omega)$. Furthermore, the optimal control problem (2) admits at least one solution.*

Proof. The proof applies standard arguments for monotone operators. For a detailed proof we refer the reader to [25]. □

Although the objective functional itself is convex, the optimization problem (2) is not convex, which is why the existence of several locally optimal controls has to be taken into account. Also, due to the non-convexity of the above optimal control problem, necessary first-order optimality conditions are no longer sufficient and the consideration of sufficient second-order optimality conditions becomes necessary if one wants to compute an actual minimal point. However, in this paper we limit our considerations to stationary points and an alternative way of ensuring a minimum, hence second-order conditions, will not be the subject of this paper.
For the optimization we have to take into account, that it is usually not possible to realize arbitrary large controls $u \in L^2(\Omega)$. Therefore control constraints yielding a bounded and convex set of admissible controls can be introduced into the model problem. In addition to the methods presented here, the handling of such control constraints may include standard optimal control methods for control constraints [21] or the application of an additional penalty term similar to Eq. (18). However, this is not directly dealt with in this paper.

Reformulating the PDE Constraint

Now, we introduce an essential reformulation of the PDE constraint based on the idea described in [12, 11]. For this purpose, we consider the Nemytskij operator ℓ which is defined by the non-linear part of the PDE. Inspired by the finite dimensional approach of Griewank and Walther, we assume that the non-smooth operator ℓ can be described as a composition of elemental functions that are either continuously Fréchet differentiable or the absolute value operator. Subsequently, consecutive continuously Fréchet differentiable elemental functions can be conceptually combined to obtain a representation, where all evaluations of the absolute value function can be clearly identified and exploited, see Def. 2.4.

Definition 2.4 (Structured evaluation)**.** *Let $\ell : H_0^1(\Omega) \to \mathbb{R}$ be some non-smooth Lipschitz continuous operator satisfying Ass. 2.1. Then, an equivalent representation of ℓ denoted by $\hat{\ell}$ can be obtained using the structured evaluation given by*

$$\begin{aligned} \left.\begin{aligned} z_i &= \psi_i(y, (\sigma_j z_j)_{j<i}) \\ \sigma_i &= \mathrm{sign}(z_i) \end{aligned}\right\} & \, i = 1, \dots, s \\ \hat{\ell}(y, \sigma z) = \psi_{s+1}(y, (\sigma_i z_i)_{1 \leqslant i \leqslant s}) \quad & \textit{with} \quad \sigma z = (\sigma_1 z_1, \dots, \sigma_s z_s). \end{aligned}$$

It should be noted that the notation $(\sigma_j z_j)_{j<i}$ indicates that ψ_i might depend also explicitly on the previously defined switching functions z_j with $j < i$. Hence, the switching function z_1 is defined as the argument of the first absolute value evaluation, i.e., as $\psi_1(y)$.

In the finite dimensional case, one has $z_i \in \mathbb{R}$ and therefore $\sigma_i \in \{-1, 0, 1\}$. For the infinite dimensional setting considered here, one has $z_i \in H^1(\Omega)$ and the functions σ_i are also Nemytskij operators defined by

$$\sigma_i : H^1(\Omega) \to L^2(\Omega), \qquad [\sigma_i(z_i)](x) = \mathrm{sign}(z_i(x)) z_i(x) \quad \text{a.e. in } \Omega$$

as functions of z_i. This choice ensures that $\sigma_i(z_i) = \mathrm{abs}(z_i) \in L^2(\Omega)$ holds. From now on, we will use the notation $\hat{\ell}(y, \sigma z) = \ell(y)$ for $\sigma z = (\sigma_1 z_1, \dots, \sigma_s z_s)$ to refer explicitly to this particular representation of the non-smooth part $\ell(y)$ based on the auxiliary functions z_i and σ_i, $1 \leqslant i \leqslant s$. It follows from the representation in Tab. 2.4 that ℓ is locally Lipschitz continuous. Hence, ℓ and therefore also the equivalent $\hat{\ell}(y, \sigma z)$ are also continuous due to the assumed smoothness of ψ_i, $i = 1, \dots, s$, [16, Theo. 3.15] and [26, Cha. 1]. Note, the operator $\hat{\ell}(.,.)$ is not smooth in z since σ depends non-Fréchet differentiably on z. However, and this is important to note, the new function $\hat{\ell}(.,.)$ is smooth i.e., Fréchet differentiable, in its two arguments y and σz, due to the chosen formulation. This fact will be exploited later to define the smooth branch problems.

Using the well-known reformulations

$$\begin{aligned} \min(v, u) &= (v + u - \mathrm{abs}(v - u))/2 \;\; \text{and} \\ \max(v, u) &= (v + u + \mathrm{abs}(v - u))/2 \;, \end{aligned} \tag{3}$$

a large class of non-smooth functions is covered by this function model.

Example 2.5. *Consider the non-smooth operator $\ell(y) = \max(5y,\ y|y|)$. Exploiting the identities (3), we can reformulate ℓ as a function in terms of the absolute value function and smooth elemental functions in the following way:*

$$\ell(y) = \max(5y,\ y|y|) = \tfrac{1}{2}\Big(5y + y|y| + \big|5y - y|y|\big|\Big) \,.$$

The corresponding structured evaluation for $\ell(y) = \max(5y,\ y|y|)$ is given by

$$\begin{aligned} z_1 &= \psi_1(y) &&= y \\ \sigma_1 &= \mathrm{sign}(z_1) && \\ z_2 &= \psi_2(y, \sigma_1 z_1) &&= 5y - y\sigma_1 z_1 \\ \sigma_2 &= \mathrm{sign}(z_2) && \\ \hat{\ell}(y, \sigma z) &= \psi_3(y, \sigma z) &&= \tfrac{1}{2}\Big(5y + y\sigma_1 z_1 + \sigma_2 z_2\Big) \end{aligned}$$

Inserting the formulation $\hat{\ell}(y, \sigma z)$ with the auxiliary functions σ_i and z_i of ℓ into the original optimal control problem (2), one obtains for the functions $y \in H_0^1(\Omega), z \in [H^1(\Omega)]^s$ and $u \in L^2(\Omega)$ the smooth optimization problem with state constraints

$$\begin{aligned} \min_{y,z,u,\sigma} \quad & \tfrac{1}{2}\|y - y_d\|_{L^2}^2 + \tfrac{\alpha}{2}\|u\|_{L^2}^2 \\ \text{s.t.} \quad & (\nabla v, \nabla y)_{L^2} + (\hat{\ell}(y, \sigma z) - u, v)_{L^2} = 0 \;\; \forall v \in H_0^1(\Omega) \\ & \left.\begin{aligned} (\psi_i(y, (\sigma_j z_j)_{j<i}) - z_i, v)_{L^2} &= 0 \;\; \forall v \in H_0^1(\Omega) \\ \sigma_i z_i &\geqslant 0 \;\; \text{a.e. in } \Omega \\ \sigma_i : \Omega &\to \{-1, 0, 1\} \end{aligned}\right\} \quad \forall\ i = 1, \dots, s \,. \end{aligned} \tag{4}$$

Here, $[H^1(\Omega)]^s$ denotes the product $H^1(\Omega) \times \cdots \times H^1(\Omega)$ of the Hilbert spaces the switching function $z = (z_1, \dots, z_s)$ lives in.
In addition to Ass. 2.1 we also assume that the given non-smooth operator ℓ fulfills the following property.

Assumptions 2.6. *Let $\ell : H_0^1(\Omega) \to L^2(\Omega)$ fulfill Ass. 2.1 and $\hat{\ell}(y, \sigma z)$ be given by Exam. 2.5. We assume that ℓ and $\hat{\ell}$ are such that the corresponding constraints of optimization problem (4) fulfill some kind of constraint qualification (e.g., in the sense of [4, 5, 21]) to ensure that the Lagrange function and the corresponding Lagrange-Multipliers are well-defined.*

The following lemma characterizes the essential relation between the solutions of the original optimization problem (2) and the, according to Tab. 2.4 reformulated, optimization problem (4) with additional equality and inequality constraints for the auxiliary functions z and σ.

Lemma 2.7. *A pair $(y^*, u^*) \in L^2(\Omega) \times H_0^1(\Omega)$ with $y^* := y^*(u^*)$ is a local solution to the original optimization problem (2) if and only if $(y^*, z^*, u^*, \sigma^*)$ with $\sigma_i^* = \mathsf{sign}(z_i^*)$ and $z_i^* = \psi_i(y, (\sigma_j^* z_j^*)_{j<i})$ for $1 \leqslant i \leqslant s$ is a local solution of the optimization problem (4).*

Proof. Assume that u^* and the corresponding $y^* := y^*(u^*)$ are solutions of the original optimization problem (2). Considering the equivalent reformulation of the operator ℓ into $\hat{\ell}$ by Tab. 2.4 and defining the auxiliary functions z_i^* and σ_i^* by

$$z_i^* = \psi_i(y, (\sigma_j^* z_j^*)_{j<i}), \qquad \sigma_i^* = \mathsf{sign}(z_i^*) \qquad \forall\ i = 1, \dots, s\,, \tag{5}$$

it follows that (y^*, z^*, u^*, σ) is a local solution of the optimization problem (4) if $\sigma_i = \sigma_i^*$ holds. Here, the additional equality and inequality constraints for the definitions of the additional functions z_i^* and σ_i^*, $1 \leqslant i \leqslant s$, ensure that $\sigma_i^*(z_i^*) = \mathsf{abs}(z_i^*) \in L^2(\Omega)$ is valid for $1 \leqslant i \leqslant s$.
On the other hand, assume that $(y^*, z^*, u^*, \sigma^*)$, with σ_i^* defined by Eq. (5), is a local solution for optimization problem (4). Then $\sigma_i^*(z_i^*) = \mathsf{abs}(z_i^*) \in L^2(\Omega)$ is valid for $1 \leqslant i \leqslant s$ and one can replace in Eq. (4) σ_i accordingly, as well as z_i by $\psi_i(y, (\sigma_j z_j)_{j<i})$ for $1 \leqslant i \leqslant s$, taking the second equality condition in Eq. (4) into account. This then yields the optimal control problem (2) with solution (y^*, u^*). ☐

This observation motivates the optimization algorithm (Algo. 1) proposed in this paper, i.e., the solution of a sequence of smooth subproblems of the form Eq. (4) to solve the original non-smooth optimization problem (2). Note that the derivation of meaningful optimality conditions for Eq. (4) does not succeed with classical methods because of the non-smooth dependence of σ on z. However, if σ^* is known and $\sigma \equiv \sigma^*$ is fixed accordingly, the optimality conditions can be derived using classical Karush–Kuhn–Tucker (KKT) theory, since the non-smooth dependence of σ and z is removed.
To determine the sequence of branch problems to be solved, we examine the necessary optimality conditions for Eq. (4) with fixed functions $\sigma_i \equiv \sigma_i^*$ according to Eq. (5). Using standard KKT theory for smooth PDE-constrained optimization problems [14], i.e., introducing corresponding Lagrange multipliers λ_{PDE}, $\lambda = (\lambda_1, \dots, \lambda_s)$, and $\mu = (\mu_1, \dots, \mu_s)$, one obtains for the Lagrangian

$$\begin{aligned}\mathcal{L}(y, z, u, \lambda_{PDE}, \lambda, \mu) =& \mathcal{J}(y, u) + \left(\nabla \lambda_{PDE}, \nabla y\right)_{L^2} + \left(\lambda_{PDE}, \hat{\ell}(y, \sigma z) - u\right)_{L^2} \\ &+ \sum_{i=1}^{s} \left(\lambda_i, \psi_i(y, (\sigma_j z_j)_{j<i}) - z_i\right)_{L^2} - \sum_{i=1}^{s} \left(\mu_i, \sigma_i z_i\right)_{L^2}\end{aligned}$$

at the optimal point for $\sigma_i = \sigma_i^*$ the first order necessary conditions

$$
\begin{aligned}
0 = D_y\mathcal{L}(\delta_y) \quad &= \tfrac{\partial \mathcal{J}}{\partial y}\delta_y + \left(\nabla\lambda_{PDE}, \nabla\delta_y\right)_{L^2} + \left(\lambda_{PDE}, \tfrac{\partial\hat{\ell}}{\partial y}\delta_y\right)_{L^2} \\
&\quad + \sum_{i=1}^{s}\left(\lambda_i, \tfrac{\partial\psi_i(y,(\sigma_j z_j)_{j<i})}{\partial y}\delta_y\right)_{L^2} \qquad \forall \delta_y \in H^{-1} && (6)\\
0 = D_u\mathcal{L}(\delta_u) \quad &= \tfrac{\partial \mathcal{J}}{\partial u}\delta_u - \left(\lambda_{PDE}, \delta_u\right)_{L^2} \qquad \forall \delta_u \in (L^2)^* && (7)\\
0 = D_{\lambda_{PDE}}\mathcal{L}(\delta_{\lambda_{PDE}}) &= \left(\nabla\delta_{\lambda_{PDE}}, \nabla y\right)_{L^2} + \left(\delta_{\lambda_{PDE}}, \hat{\ell} - u\right)_{L^2} \qquad \forall \delta_{\lambda_{PDE}} \in H^{-1} && (8)\\
0 = D_{\lambda_i}\mathcal{L}(\delta_{\lambda_i}) \quad &= \left(\delta_{\lambda_i}, \psi_i(y,(\sigma_j z_j)_{j<i}) - z_i\right)_{L^2} \qquad \forall \delta_{\lambda_i}, i = 1,\dots,s, \delta_{\lambda_i} \in C(\bar{\Omega})^* && (9)\\
0 = D_{z_k}\mathcal{L}(\delta_{z_k}) \quad &= \left(\lambda_{PDE}, \sigma_k \tfrac{\partial\hat{\ell}(y,\sigma z)}{\partial z_k}\delta_{z_k}\right)_{L^2} - \left(\lambda_k, \delta_{z_k}\right)_{L^2} \\
&\quad + \sum_{i=k+1}^{s}\left(\lambda_i, \sigma_k \tfrac{\partial\psi_i(y,(\sigma_j z_j)_{j<i})}{\partial z_k}\delta_{z_k}\right)_{L^2} \\
&\quad - (\mu_k, \sigma_k\delta_{z_k})_{L^2} \qquad \forall \delta_{z_k}, k = 1,\dots,s, \delta_{z_k} \in L^2 && (10)
\end{aligned}
$$

$$
\begin{aligned}
0 &= (\mu_i, \sigma_i z_i)_{L^2} \qquad i = 1,\dots,s \\
0 &\leqslant \mu_i \qquad i = 1,\dots,s\,, && (11)
\end{aligned}
$$

where the arguments of $\mathcal{L}$ are omitted for brevity. Note that in these equations one obtains extra factors σ_k due to the chain rule. Rearranging the terms in the integrals, the condition (10) yields for $k = 1,\dots,s$

$$0 = \sigma_k \frac{\partial\hat{\ell}(y,\sigma z)}{\partial z_k}\lambda_{PDE} - \lambda_k + \sum_{i=k+1}^{s}\sigma_k \frac{\partial\psi_i(y,(\sigma_j z_j)_{j<i})}{\partial z_k}\lambda_i - \sigma_k\mu_k\,.$$

In this case the right-hand side represents the zero function in the corresponding Hilbert space. Applying σ_k and exploiting the non-negativity of μ_k according to Eq. (11), one obtains

$$0 \leqslant \mu_k|\sigma_k| = r(\sigma_k, y, z, \lambda) \qquad \text{a.e. in } \Omega\,, \tag{12}$$

with

$$r(\sigma_k, y, z, \lambda) := |\sigma_k|\frac{\partial\hat{\ell}(y,\sigma z)}{\partial z_k}\lambda_{PDE} - \sigma_k\lambda_k + \sum_{i=k+1}^{s}|\sigma_k|\frac{\partial\psi_i(y,(\sigma_j z_j)_{j<i})}{\partial z_k}\lambda_i\,. \tag{13}$$

We will use inequality (12) later to define the sequence of subproblems to be solved.

3 Defining and Solving the Branch Problems

Definition 3.1 (Abs-Linearization)**.** *For a given structured evaluation and the resulting operator $\hat{\ell}$ described in Def. 2.4 the Abs-Linearization is obtained by fixing all σ_i for $1 \leqslant i \leqslant s$ to given $\bar{\sigma}_i \in L^2(\Omega)$, $\bar{\sigma}_i : \Omega \to \{-1, 1\}$ for $1 \leqslant i \leqslant s$.*

Using the abs-linearization, the resulting operator $\hat{\ell}(.,\bar{\sigma}.)$ is smooth in both arguments. Now, everything is prepared to introduce the main idea of the new optimization algorithm. For fixed functions $\bar{\sigma}_i \in L^2(\Omega)$, $\bar{\sigma}_i : \Omega \to \{-1, 1\}$ for $1 \leqslant i \leqslant s$, we define for $(y, z, u) \in H_0^1(\Omega) \times [L^2(\Omega)]^s \times L^2(\Omega)$ the branch problem

$$\min_{y,z,u} \quad \mathcal{J}(y, u) \tag{14}$$

$$
\begin{aligned}
\text{s.t.} \quad (\nabla v, \nabla y)_{L^2} + (\hat{\ell}(y, \bar{\sigma}z) - u, v)_{L^2} &= 0 \ \forall v \in H_0^1(\Omega) && (15)\\
(\psi_i(y, (\bar{\sigma}_j z_j)_{j<i}) - z_i, v_i)_{L^2} &= 0 \ \forall v_i \in H_0^1(\Omega) \quad \forall\, i = 1,\dots,s && (16)\\
\bar{\sigma}_i z_i &\geqslant 0 \text{ a.e. in } \Omega \quad \forall\, i = 1,\dots,s\,. && (17)
\end{aligned}
$$

All functions occurring in this branch problem are smooth in the variables y, u and z because the function $\hat{\ell}(.,.)$ is smooth in its arguments as mentioned already in the last section. Therefore, standard smooth optimization methods can be used to solve the branch problem (14)–(17). Naturally, the question arises how to chose the functions $\bar{\sigma}_i$, $1 \leqslant i \leqslant s$, such that the solutions of the branch problems approach the solution of the original non-smooth problem (2). A corresponding strategy will be derived in this section.

The Lagrangian with Bi-quadratic Penalty

As mentioned already above, so far the solution of the non-smooth optimization problem using a reduced formulation for simulation based approaches, which results essentially from applying the implicit functions theorem, is not possible due to the lack of the classical chain rule as well as the non-smoothness of the control-to-state operator associated with the non-smooth state equation. For this reason, we propose here a penalty-based approach to solve the optimization problem (14)–(17), where the constraints (15) and (16) are handled explicitly. Methods based on a reduced formulation will be subject of future research.
From a formal point of view, we treat the inequality constraints (17) with a penalty approach such that the target function (14) is modified to

$$\min_{y,z,u} \quad \mathcal{J}(y,u) + \nu \int_\Omega \sum_{i=1}^{s} \Big(\max(-\bar{\sigma}_i z_i, 0) \Big)^4 \, d\Omega \tag{18}$$

with a penalty factor $\nu > 0$. In this context, as well as in the further course, ν describes a non-negative constant penalty parameter for the inequality conditions on $\sigma_i z_i$. Note that the resulting penalty term does not serve for regularizing the non-smoothness into a smoother term but to treat the inequality constraint. Here, we chose the exponent 4 to ensure that the target function is twice continuously differentiable despite the max function that is used for the formulation of the penalty function. This modified target function is then coupled with the equality constraints by means of Lagrange multipliers yielding the Lagrangian

$$\begin{aligned} \mathcal{L}^p(y,z,u,\lambda_{PDE},\lambda_1,\dots,\lambda_s) &= \mathcal{J}(y,u) + \big(\nabla\lambda_{PDE}, \nabla y\big)_{L^2} + \big(\lambda_{PDE}, \hat{\ell}(y,\bar{\sigma} z) - u\big)_{L^2} \\ &\quad + \sum_{i=1}^{s} \big(\lambda_i, \psi_i(y,(\bar{\sigma}_j z_j)_{j<i}) - z_i\big)_{L^2} + \nu \int_\Omega \sum_{i=1}^{s} \Big(\max(-\bar{\sigma}_i z_i, 0) \Big)^4 \, d\Omega \,. \end{aligned} \tag{19}$$

A similar penalty approach was studied in [23], where the logarithm was used as barrier function. Here, we employ the max-function since we have to evaluate the penalty function also at 0.

Example 3.2. *We consider again* $\ell(y) = \max(5y,\ y|y|) = \frac{1}{2}\Big(5y + y|y| + \big|5y - y|y|\big|\Big)$. *For the reformulated optimization problem given by*

$$\begin{aligned} \min_{(y,z,u)\in H_0^1\times H_0^1\times L^2} \quad & \tfrac{1}{2}\|y - y_d\|_{L^2}^2 + \tfrac{\alpha}{2}\|u\|_{L^2}^2 \\ \text{s.t.} \quad & (\nabla v, \nabla y)_{L^2} + \Big(\tfrac{1}{2}\big(5y + y\bar{\sigma}_1 z_1 + \bar{\sigma}_2 z_2\big) - u, v\Big)_{L^2} = 0 \ \forall v \in H_0^1(\Omega) \\ & (y - z_1, v)_{L^2} = 0 \ \forall v \in H_0^1(\Omega) \\ & (5y - y\bar{\sigma}_1 z_1 - z_2, v)_{L^2} = 0 \quad \forall v \in H_0^1(\Omega) \\ & \bar{\sigma}_1 z_1 \geqslant 0 \ \ a.e \ in \ \Omega \\ & \bar{\sigma}_2 z_2 \geqslant 0 \ \ a.e \ in \ \Omega \,, \end{aligned}$$

one obtains the Lagrangian

$$
\begin{aligned}
\mathcal{L}^p(y, z, u, &\lambda_{PDE}, \lambda_1, \lambda_2) \\
&= \mathcal{J}(y, u) + \left(\nabla\lambda_{PDE}, \nabla y\right)_{L^2} + \left(\lambda_{PDE}, \tfrac{1}{2}\left(5y + y\bar{\sigma}_1 z_1 + \bar{\sigma}_2 z_2\right) - u\right)_{L^2} \\
&\quad + \left(\lambda_1, y - z_1\right)_{L^2} + \left(\lambda_2, 5y - y\bar{\sigma}_1 z_1 - z_2\right)_{L^2} \\
&\quad + \nu \int_\Omega \sum_{i=1}^{2} \left(\max(-\bar{\sigma}_i z_i, 0)\right)^4 \, d\Omega\ .
\end{aligned}
$$

Deriving Necessary Optimality Conditions

A simple comparison of the equivalently reformulated problem and the penalty branch problem suggests that a solution to the penalty branch problem is feasible for the original problem, if it satisfies the condition

$$
\sigma_i z_i = \mathrm{abs}(z_i) \text{ a.e in } \Omega\ . \tag{20}
$$

However, this provides no statement about the optimality of the solution. In addition, if the condition Eq. (20) is not met, only a statement about the selected penalty parameter can be made, but no efficient strategy for switching the fixed σ_i can be derived. The strategy presented below, on the other hand, guarantees a descent in the objective value due to Farkas' lemma, which is stated and proved in Thm. 4.1.
For a branch problem, the first-order necessary optimality conditions can now be derived from the Lagrangian (19) by using once more standard KKT theory for smooth PDE-constrained optimization problems given that some regularity conditions are satisfied at the local minimum. This yields as necessary first order conditions the equations

$$
\begin{aligned}
0 = D_y\mathcal{L}^p(\delta_y) &= \tfrac{\partial\mathcal{J}}{\partial y}\delta_y + \left(\nabla\lambda_{PDE}, \nabla\delta_y\right)_{L^2} + \left(\lambda_{PDE}, \tfrac{\partial\hat{\ell}}{\partial y}\delta_y\right)_{L^2} \\
&\quad + \sum_{i=1}^{s} \left(\lambda_i, \tfrac{\partial\psi_i(y,(\bar{\sigma}_j z_j)_{j<i})}{\partial y}\delta_y\right)_{L^2} && \forall\delta_y && (21)\\
0 = D_u\mathcal{L}^p(\delta_u) &= \tfrac{\partial\mathcal{J}}{\partial u}\delta_u - \left(\lambda_{PDE}, \delta_u\right)_{L^2} && \forall\delta_u && (22)\\
0 = D_{\lambda_{PDE}}\mathcal{L}^p(\delta_{\lambda_{PDE}}) &= \left(\nabla\delta_{\lambda_{PDE}}, \nabla y\right)_{L^2} + \left(\delta_{\lambda_{PDE}}, \hat{\ell}\,\right)_{L^2} - \left(\delta_{\lambda_{PDE}}, u\right)_{L^2} && \forall\delta_{\lambda_{PDE}} && (23)\\
0 = D_{\lambda_i}\mathcal{L}^p(\delta_{\lambda_i}) &= \left(\delta_{\lambda_i}, \psi_i(y, (\bar{\sigma}_j z_j)_{j<i}) - z_i\right)_{L^2} && \forall\delta_{\lambda_i}, 1\leqslant i\leqslant s && (24)\\
0 = D_{z_k}\mathcal{L}^p(\delta_{z_k}) &= \left(\lambda_{PDE}, \bar{\sigma}_k \tfrac{\partial\hat{\ell}(y,\bar{\sigma}z)}{\partial z_k}\delta_{z_k}\right)_{L^2} - \left(\lambda_k, \delta_{z_k}\right)_{L^2} \\
&\quad + \sum_{i=k+1}^{s} \left(\lambda_i, \bar{\sigma}_k \tfrac{\partial\psi_i(y,(\bar{\sigma}_j z_j)_{j<i})}{\partial z_k}\delta_{z_k}\right)_{L^2} && && (25)\\
&\quad + \nu\int_\Omega -4\bar{\sigma}_k \max(-\bar{\sigma}_k z_k, 0)^3\ \delta_{z_k}\ d\Omega && \forall\delta_{z_k}, 1\leqslant k\leqslant s
\end{aligned}
$$

As one can easily see, the optimality conditions (6)–(9) coincide with the optimality conditions (21)–(24). The following relation between the KKT point of the branch problem Eqs. (14)–(17) and the original optimization problem Eq. (2) can be derived.

Lemma 3.3. *Let $\bar{y} \in H^1_0(\Omega)$, $\bar{z} = (\bar{z}_1, \dots, \bar{z}_s) \in [H^1_0(\Omega)]^s$, $\bar{u} \in L^2(\Omega)$, $\bar{\lambda}_{PDE} \in H^{-1}$ and $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_s) \in [(L^2)^*]^s$. Assume that the conditions Eq. (21)–(25) hold for $(\bar{y}, \bar{z}, \bar{u}, \bar{\lambda}_{PDE}, \bar{\lambda})$ together with Eq. (12) for all $1 \leqslant k \leqslant s$. Then the pair $(\bar{y}, \bar{z}, \bar{u}, \bar{\lambda}_{PDE}, \bar{\lambda})$ is a KKT point of the optimization problem Eq. (4). Furthermore, $(\bar{y}, \bar{u})$ is a stationary point of the original problem Eq. (2).*

Proof. We again point out that the optimality conditions (6)–(9) for the branch problem coincide with the optimality conditions (21)–(24) for the reformulated problem Eq. (4). Hence, if one computes a solution of the slightly modified branch problem with the target function

(18) and the constraints (15)–(16), the necessary first order conditions (6)–(9) of the original optimization problem are already satisfied. Consequently, the only condition to verify is Eq. (12). Since the expressions on the right-hand side are completely independent of the Lagrange multiplier μ of the original optimization problem, one can compute this quantity also for the solution of the modified branch problem. If it is non-negative, the computed solution $(\bar{y}, \bar{z}, \bar{u})$ of the modified branch problem fulfills the necessary first order conditions of the original optimization problem for the chosen functions $\bar{\sigma}_i \in L^2(\Omega)$. Hence, $(\bar{y}, \bar{z}, \bar{u})$ is a stationary point of Eq. (4) and by Lem. 2.7 it is also a stationary point of Eq. (2). □

Note, that if Eq. (12) is does not hold, i.e, the expression on the right-hand side is negative, it is a very natural strategy to choose the index k for which the right-hand side of the condition (12) is minimal, to modify the corresponding $\bar{\sigma}_k$ appropriately and to solve the then newly defined branch problem by the same strategy. Due to the structure of Eq. (12), the Lagrange multiplier λ_k identifies the regions where the sign of $\bar{\sigma}_k$ has to be changed to obtain a reduction in the function value. Obviously, other strategies to choose the index k as alternatives to the greedy approach described here might be applied as well.

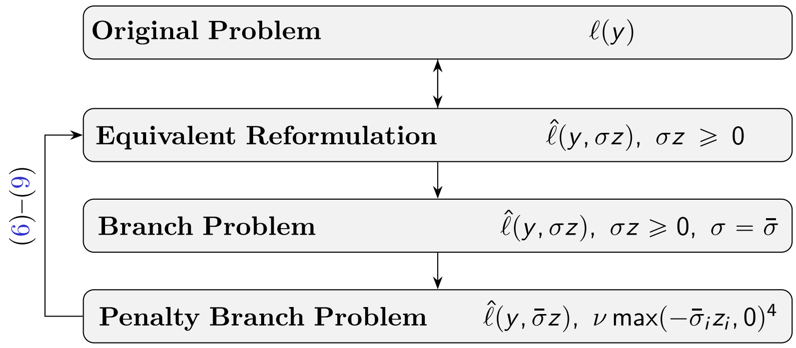

Figure 1: The different optimization problems and how they relate to each other
.

Fig. 1 illustrates the nature of the relationships between the different problem formulations that are derived and discussed in this paper.

4 The Resulting Optimization Algorithm

Motivated by the observations of the last section, we propose the following method stated in Algo. 1 to solve optimal control problems with non-smooth PDEs of the class considered here as constraints. Since the proposed algorithm is essentially motivated by the special handling of the absolute value function, i.e., the abs-linearization, we call the resulting optimization algorithm SALi for Successive Abs-Linearization. Note that the formulation of the algorithm is done in the function space. Therefore, up to this point one can use the own method of choice to solve the smooth modified branch problems. Following standard practice for PDE-constrained optimization, we develop the algorithm in a function space setting. This has the advantage that the associated algorithms are often able to provide mesh-independent convergence for a variety of conform discretizations, see for example [1, 2, 15, 22]. As can be seen from the numerical results in Sec. 5, mesh independence is also an important feature of the algorithm presented here.
For the numerical results shown in the next section, 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 branch problems.
For the initial state, control and parameters $\bar{\sigma}_i$, the non-linear variational Lagrange problem is solved by Newton's method using the derivatives calculated within FEniCS.
The computed solution is examined according to the switching rule and the branch problem

Algorithm 1

Input: Initial values: $\bar{\sigma}^0 = (\bar{\sigma}_1^0, \ldots, \bar{\sigma}_s^0), y^0, z^0 = (z_1^0, \ldots, z_s^0), u^0$
Parameter: $\alpha > 0, \nu > 0, i = 0$
for $i = 0, 1, \ldots$ **do**
Solve branch problem (18) with constraints (15)–(16) to obtain $y^i, z^i, u^i, \lambda_{PDE}^i, \lambda^i$
if Eq. (12) holds for $k = 1, \ldots, s$ **then**
y^i, z^i, u^i stationary for original optimal control problem, stop
else
$\kappa = \operatorname{argmax}\{k \in \{1, \ldots, s\} : -r(\bar{\sigma}_k^i, y, z, \lambda)\}$, where $r(.)$ is given by Eq. (13)
Switch branch using λ_κ to define $\bar{\sigma}_\kappa^{i+1}$
Set $\bar{\sigma}_k^{i+1} = \bar{\sigma}_k^i$ for $k = 1, \ldots, s,\ k \neq \kappa$
end if
$i = i + 1$
end for

is modified by updating the corresponding σ_i. Here again the update strategy is based on Eq. (12).
Hereinafter we derive a heuristic for the switching strategy, which is presented in Algo. 2, and provides the desired effect, as we will see in the numerical results. By exploiting the essence of Eq. (12), a beneficial and comparatively easy way to implement an update strategy for the parameters $\bar{\sigma}_i$ can be created. Reformulation of Eq. (12) and application of σ_k provides

$$0 \leqslant \left(\lambda_{PDE}, |\sigma_k| \tfrac{\partial \hat{\ell}(y,\sigma z)}{\partial z_k} \delta_{z_k}\right)_{L^2} - \left(\lambda_k \sigma_k, \delta_{z_k}\right)_{L^2} + \sum_{i=k+1}^{s} \left(\lambda_i, |\sigma_k| \tfrac{\partial \psi_i(y,(\sigma_j z_j)_{j<i})}{\partial z_k} \delta_{z_k}\right)_{L^2} \ \forall \delta_{z_k}, \forall k = 1, \ldots, s.$$

This condition is violated if and only if there exists an index $k \in \{1, \ldots, s\}$ such that $\sigma_k = \operatorname{sign}(\lambda_k)$ with

$$0 > \left(\lambda_{PDE}, |\sigma_k| \tfrac{\partial \hat{\ell}(y,\sigma z)}{\partial z_k} \delta_{z_k}\right)_{L^2} - \left(|\lambda_k|, \delta_{z_k}\right)_{L^2} + \sum_{i=k+1}^{s} \left(\lambda_i, |\sigma_k| \tfrac{\partial \psi_i(y,(\sigma_j z_j)_{j<i})}{\partial z_k} \delta_{z_k}\right)_{L^2}.$$

Since $|\sigma_k| \equiv 1$, this is equivalent to

$$0 > \left(\lambda_{PDE}, \tfrac{\partial \hat{\ell}(y,\sigma z)}{\partial z_k} \delta_{z_k}\right)_{L^2} - \left(|\lambda_k|, \delta_{z_k}\right)_{L^2} + \sum_{i=k+1}^{s} \left(\lambda_i, \tfrac{\partial \psi_i(y,(\sigma_j z_j)_{j<i})}{\partial z_k} \delta_{z_k}\right)_{L^2}. \tag{26}$$

Again, as already discussed in Sec. 3 it is a natural strategy to choose the index k for which the right-hand side in Eq. (26) is minimal. Since this is significantly influenced by the Lagrange multiplier λ_k, we use this as an indicator to switch from the current branch problem defined by $\bar{\sigma}_k^i$ to the next one defined by $\bar{\sigma}_k^{i+1}$ by switching the signs of $\bar{\sigma}_k^i$ in the regions where the corresponding $|\lambda_k^i|$ is largest. For this purpose, the Lagrange multipliers λ_i corresponding to the solution of the current branch problem are projected to the adequate finite dimensional function space and their L^∞-norm is computed in order to determine the Lagrange multiplier λ_k with maximum influence on Eq. (26). If this maximum value (almost) vanishes, a stationary point is already reached and the algorithm stops. Otherwise, the sign of the corresponding discretized $\bar{\sigma}_k$ is switched at those mesh points where $|\lambda_k|$ is large and exceeds a certain threshold. Certainly, there are also other update strategies, however this is the one we have chosen and which has show convincing results, as we will see in the further course of this report. We would like to emphasize that the vanishing Lagrange multiplier λ_k corresponds to the equality constraint Eq. (16) for the definition of the switching function z_k. The termination condition due to this vanishing Lagrange multiplier is based on the requirement that the associated equality constraint Eq. (16) is satisfied naturally at the solution. Nevertheless, the

Algorithm 2

Input: Solved branch problem (18) for $\bar{\sigma}^k$ with solution $y^k, z^k, u^k, \lambda^k_{PDE}, \lambda^k$; $\mathcal{L}(y^k, z^k, u^k, \lambda^k_{PDE}, \lambda^k)$, $\mathcal{J}(y^k, u^k)$, $\mathcal{L}(y^{k-1}, z^{k-1}, u^{k-1}, \lambda^{k-1}_{PDE}, \lambda^{k-1})$, $\mathcal{J}(y^{k-1}, u^{k-1})$
Parameter: $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$
if $\|\lambda_k\|_\infty < \varepsilon_1$ holds for $k = 1, \dots, s$ **then**
  y^i, z^i, u^i stationary for original optimal control problem, Stop
end if
if $|\mathcal{L}(y^k, z^k, u^k, \lambda^k_{PDE}, \lambda^k) - \mathcal{J}(y^k, u^k)| < \varepsilon_2$ and $|\mathcal{J}(y^{k-1}, u^{k-1}) - \mathcal{J}(y^k, u^k)| < \varepsilon_3$ holds **then**
  Stop
else
  $\kappa = \operatorname{argmax}\{k \in \{1, \dots, s\} : \|\lambda_k\|_\infty\}$
  Switch branches, e.g., switch sign $\bar{\sigma}^{i+1}_\kappa$ where $|\lambda_\kappa|$ is large
  Set $\bar{\sigma}^{i+1}_k = \bar{\sigma}^i_k$ for $k = 1, \dots, s$, $k \neq \kappa$
end if

Lagrange multiplier do not necessarily have to vanish at the stationary point. For this reason, it is additionally checked whether the admissibility is maintained by calculating the difference between the Lagrange function of the current branch problem at the calculated solution and the original target function (without penalty term). Admissibility is achieved if this difference is close to zero. The algorithm stops if admissibility is reached and the objective function value does not improve significantly compared to the previous values, even if $\|\lambda_k\|_\infty$ is not close to zero for all $1 \leqslant k \leqslant s$.
Note that if no switching occurs, then the branch problem with an optimal solution was reached, and the algorithm stops. Otherwise, the branch problem is updated accordingly and a new solution is computed by once again solving the non-linear variational Lagrange problem by applying Newton's method. This way a successive reduction in the objective function value is observed. This can also be seen in Fig. 2(a) for one example.
Despite the fact that this heuristic works well in practice, we will continue to develop our existing approach further and adapt it for the calculation of Eq. (12) and a related systematic switching strategy of the branch problems.
In what follows we will deal with the reduction of the target function value after each branch problem switch and examine it in more detail.

Theorem 4.1 (Decent Direction)**.** *Consider the optimization problem Eq. (2) with a non-smooth operator $\ell : H^1_0(\Omega) \to L^2(\Omega)$ satisfying Ass. 2.6. Then a solution $(\bar{y}, \bar{z}, \bar{u})$ for the associated penalty optimization problem corresponding to $\bar{\sigma}$ with objective functional Eq. (18) and the constraints Eqs. (15)–(16) is already an optimal solution for Eq. (2) or there exists a σ° such that the solution $(y^\circ, z^\circ, u^\circ)$ to the corresponding penalty branch problem is feasible for Eq. (2) and satisfies $J(y^\circ, u^\circ) \leqslant J(\bar{y}, \bar{u})$.*

Proof. According to Ass. 2.6 the considered optimization problem satisfies some kind of constraint qualification. The constraint qualification allows the use of the Farkas alternative (see e.g. Appendix A as well as [7, 9, 18]) to provide necessary conditions of the KKT type. The Farkas alternative then especially yields that, if the solution to the penalty branch problem fails to be an optimal solution to the original problem formulation, i.e., Eq. (12) does not hold, then there exists some descent direction yielding descent in the objective value.
According to Lem. 3.3 the solution $\bar{w} := (\bar{y}, \bar{z}, \bar{u})$ for the associated penalty optimization problem is already a solution for the original optimization problem (2) if Eq. (10) —and hence condition (12)— holds. Since the optimality conditions Eqs. (6)–(9) coincide with the optimality conditions Eqs. (21)–(24), they hold for $\bar{w}$. Hence, the only potential non vanishing components in the derivative of $\mathcal{L}^p(\bar{w}, \lambda_{PDE}, \lambda)$ are the derivative with respect to the switching

functions $z_i, i = 1, \dots, s$. Let K be the cone that is spanned by the equality and active inequality constraints Eqs. (15)–(17), $A : L^2(\Omega) \to L^2(\Omega)^*$ the operator defined by the derivatives of the constraints and $b := DJ(\bar{w})$ the derivative of the objective functional $J(\bar{w}) := J(\bar{y}, \bar{u})$ at the considered point $\bar{w}$. Note, that for the active inequality constraints we have $z_i = 0$. Hence, the extended Farkas Lemma, Lem. A.1, yields that the solution to the current penalty branch problem either fulfills Eq. (12) and therefore all optimality conditions for the original optimal control problem are met and the current iterate is already a solution to the original problem, or there exists a direction v^* such that

$$\langle DJ(\bar{w}), v^* \rangle < 0 \,. \tag{27}$$

The latter implies a descent direction for the objective functional and hence the existence of a $\sigma^\circ \neq \bar{\sigma}$ such that the solution to the corresponding penalty branch problem satisfies $J(y^\circ, u^\circ) \leqslant J(\bar{y}, \bar{u})$. Once again the Farkas alternative implies that this solution is either an optimal solution to the original optimization problem or there exists a decent direction in the sense of Eq. (27). This proofs the assertion. □

Thus, it should be noted that because of this choice, the corresponding penalty branch problem either already provides an optimal solution of the original problem or, by using Eq. (26), a descent direction. It is precisely this circumstance that is exploited in the switching strategy and the heuristic, explained previously, and thus leads iteratively to a reduction in the objective functional and ultimately to a stationary point.

Finite Dimensional Formulation

If one considers the discretized and hence finite dimensional problem, the convergence of the algorithm follows immediately. This is due to the fact that the bounded original problem was decomposed into a large but finite number of discretized branch problems and the objective function value decreases with each iteration step. Therefore, a stationary point is reached after finitely many steps.
One might suspect that the update strategy for σ is somehow reminiscent of active-set strategies. However, let us highlight one more time that the σ update strategy depends only on the Lagrange multiplier associated with the specific equality constraint corresponding to the definition of the specific switching function z.
In the last paragraph, the algorithm was presented and explained in the continuous function space setting. Now, the natural question is how to put this into practice and, especially, how to solve the individual branch problems. For the numerical treatment of the optimal control problem (18) with the constraints given by Eqs. (15) and (16), the Lagrange equation (19) will be discretized. For this purpose we apply a standard finite element method with piecewise linear and continuous ansatz functions for the discretization of the functions y and $z_i\sigma_i,\ i = 1, \dots, s$, and piecewise constant ansatz functions for the control u. The resulting problem is solved by the Galerkin method within the open source finite element environment FEniCS.
Below we will discuss the spatial discretization of the constrained optimization problem (2), which will result in a large-scale non-linear optimization problem. We focus on finite element approaches with a quasi uniform triangulation $\mathbb{T}^h = \{T_1, \dots, T_m\}$, the vector space of test functions $V^h := \{v^h \in C^0(\bar{\Omega}) : v^h|_{T_j} \in \mathcal{P}_1(T)\ \forall T \in \mathbb{T}^h, v^h|_{\partial\Omega} = 0\} = \mathrm{span}\{\xi_1, \dots, \xi_n\}$ and the discrete control space $U^h := \mathrm{span}\{e_T : T \in \mathbb{T}^h\}$ where $e_T : \Omega \to \mathbb{R}$ denotes the characteristic function for the simplex $T \in \mathbb{T}^h$. The superscript h denotes the mesh size of the triangulation and is given by

$$h := \max_{T \in \mathbb{T}^h} \mathrm{diam}(T) \,.$$

Then the discretization of Eq. (2) can be stated as

$$\min_{(y^h, u^h) \in V^h \times U^h} \quad J(y^h, u^h) \tag{28}$$

$$\text{s.t.} \quad (\nabla y^h, \nabla v^h)_\Omega + (\ell(y^h), v^h)_\Omega - (u^h, v^h)_\Omega = 0 \quad \forall v^h \in V^h \,. \tag{29}$$

For a given function $y^h \in V^h$ we denote by $\mathbf{y} = (y_1, \dots, y_n)^T \in \mathbb{R}^n$ its vector of coefficients with respect to the basis $\{\xi_1, \dots, \xi_n\}$, 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 $\mathbf{u} = (u_1, \dots, u_m)^T \in \mathbb{R}^m$ 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 $\mathbb{T}^h$. Taking into account that the operator ℓ is non-linear, the above representations yield the following discretizations:

$$\ell(y^h) = \ell(\sum_{i=1}^{n} y_i \xi_i)$$

and

$$(\ell(y^h), v^h)_{L^2} = \int_\Omega \ell(y^h) v^h dx \approx \sum_{T \in \mathbb{T}^h} \int_T \ell(y^h) v^h dx \ . \tag{30}$$

The integrals over the elements $T \in \mathbb{T}$ are approximated by some quadrature formula

$$\sum_{T \in \mathbb{T}^h} \int_T \ell(y^h) v^h dx \approx \sum_{T \in \mathbb{T}^h} \sum_{k=1}^{n_k} \omega_k \ \ell\left(\sum_{i=1}^{n_i} y_i \xi_i(x_k)\right) \sum_{j=1}^{n_j} \xi_j(x_k) \ , \tag{31}$$

with n_k quadrature points per element T and corresponding weights ω_k.
Hence, the naturally arising discretization for the non-smooth operator from Definition 2.4 in the finite element context is per quadrature point.This increases the number of absolute value evaluations, but not the way they are nested.
As seen in Eq. (31), we would like to point out that the number of non-smooth functions ℓ in the discretized problem is per quadrature point. Compared with our substitution strategy Tab. 2.4, this is not in perfect alignment with a representation by a finite element function like the state y. Consequentially, the choice of this discretization and the execution of the equivalent reformulation according to Tab. 2.4 leads to an increase of the polynomial degree due to the multiplication $\bar{\boldsymbol{\sigma}}_\mathbf{i} \cdot \mathbf{z_i}$ in the discretized representation of the operator $\hat{\ell}$ in contrast to the operator ℓ. However, this specific discretization allows for a straight forward implementation with FEniCS.
Inserting Eq. (30) into Eq. (29) and replacing v by ξ leads to:

$$\int_\Omega \sum_{k=1}^{n} \nabla \xi_j(x) \cdot \nabla \xi_k(x) y_k + \ell\big(\sum_{i=1}^{n} y_i \xi_i(x)\big) \xi_j(x) dx = \int_\Omega \left(\sum_{s=1}^{m} u_s e_{T_s}(x)\right) \xi_j(x) dx \ , \tag{32}$$

for $i \leqslant j \leqslant n$. By defining

$$A_{jk} := \int_\Omega \nabla \xi_j(x) \cdot \nabla \xi_k(x) dx = (\nabla \xi_j, \nabla \xi_k)_\Omega,$$
$$b_k(y^h) := \int_\Omega \ell\big(\sum_{i=1}^{n} y_i \xi_i(x)\big) \xi_k(x) dx$$

and

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

Eq. (32) 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

$$A\mathbf{y} + \mathbf{b}(\mathbf{y}) = \mathbf{u}^T E \ , \tag{33}$$

with the control matrix $E_{ij} := (e_{T_i}, \xi_j)$ and $\mathbf{y} = (y_1, \dots, y_n)^T$ 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 point values at preselected sets of nodes scattered along the boundary of T_k. Note that in the above algebraic system the vector $\mathbf{u}$ and the matrices A, E are constant since they are independent of the unknown $y_1, \dots, y_n$. However, as previously mentioned, this non-linear algebraic equation is assumed to be based on a reasonable approximation of the integral via quadrature.
Hence, the resulting discretized objective functional reads as

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

Herein $M \in \mathbb{R}^{n\times n}$ denotes the mass matrix $M_{ij} = (\xi_i, \xi_j)_\Omega$ and D the control mass matrix with the entries $D_{ij} = (e_{T_i}, e_{T_j})_\Omega$, where D is a diagonal matrix because the interior of the triangles are disjunct to each other.
Similarly, to the previously derived discretization, the discrete counterpart to branch problem Eq. (14)–(17) is given by:

$$\begin{aligned}
&\min_{(y^h,z^h,u^h)\in V^h\times[V^h]^s\times U^h} \quad J(y^h, u^h) + \nu \int_\Omega \max\left(-\bar{\sigma}_i^h z_i^h, 0\right)^4 dx \\
&\text{s.t.} \quad (\nabla y^h, \nabla v^h)_\Omega + \left(\hat{\ell}(y^h, \bar{\sigma}^h z^h), v^h\right)_\Omega = (u^h, v^h)_\Omega, \qquad \forall v^h \in V^h \\
&\qquad\qquad (z_i^h - \psi_i(y^h, (\bar{\sigma}_j^h z_j^h)_{j<i}), v^h)_\Omega = 0 \qquad \forall\ 1 \leqslant i \leqslant s,\ \forall v^h \in V^h \ .
\end{aligned} \tag{34}$$

Note that $z^h = (z_1^h, \dots, z_s^h) \in [V^h]^s$. Hence, it becomes clear that the inequality constraint from Eq. (17) is enforced per quadrature point via our penalty approach.

The assumptions for the non-smooth operator ℓ are carried over from the continuous setting into the discrete and hence once again we assume that the optimization problem Eq. (34) fulfills some kind of constraint qualification to ensure that the Lagrange function and the Lagrange multipliers are well-defined, i.e., the existence of the Lagrange multipliers is ensured. The corresponding discrete Lagrange functional related to the penalty branch problem of system Eq. (34) is now given by

$$\begin{aligned}
\mathcal{L}^p(y^h, z^h, u^h, \lambda^h_{PDE}, \lambda^h) = \mathcal{J}(y^h, u^h) + \left(\nabla\lambda^h_{PDE}, \nabla y^h\right)_\Omega + \nu \int_\Omega \sum_{i=1}^{s} \left(\max(-\bar{\sigma}_i^h z_i^h, 0)\right)^4 dx \\
+ \left(\lambda^h_{PDE}, \hat{\ell}(y^h, \bar{\sigma} z^h) - u^h\right)_\Omega + \sum_{i=1}^{s} \left(\lambda_i^h, \psi_i(y^h, (\bar{\sigma}_j^h z_j^h)_{j<i}) - z_i^h\right)_\Omega \ .
\end{aligned} \tag{35}$$

The KKT system corresponding to Eq. (35)is then solved with a non-linear variational Newton solver. To determine the sequence of branch problems to be solved, we apply the already explained switching method in its discrete version corresponding to the discretization described above.

5 Numerical Results

For the numerical tests we considered two-dimensional examples defined below in Case 1 to Case 4. In each example Ω was chosen to be the unit square, and we take as an initial guess $y \equiv 0, u \equiv 0, z_1 \equiv 0, z_2 \equiv 0$. Furthermore, $\bar{\sigma}_1$ and $\bar{\sigma}_2$ are chosen such that they fit the ones defined by the desired state y_d. We terminate the iteration if either the L^2-Norm of the Lagrange multipliers λ_i becomes less than 10^{-9} and therefore no further switching between branch problems is done, or if the difference between the Lagrange function value, which includes is the same as the bi-quadratic penalty terms and the original objective functional, becomes less than 10^{-12}. The latter implicitly ensures that the sign condition $\bar{\sigma}_i z_i \geqslant 0$ is correctly adhered to. All calculations were performed with FEniCS, version 2019.1.0, using the Python interface.

Case 1

$$
\begin{aligned}
&\min_{(y,u)} \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \\
&\text{s.t.} \quad -\Delta y + \max(0, y) - u = f \quad \text{in } \Omega = (0,1)^2 , \\
&\text{with } y_d(x_1, x_2) = \begin{cases} ((x_1 - \frac{1}{2})^4 + \frac{1}{2}(x_1 - \frac{1}{2})^3)\sin(\pi x_2), & \text{if } x_1 \leqslant \frac{1}{2} \\ 0, & \text{otherwise} , \end{cases}
\end{aligned}
$$

where $f \in L^2(\Omega)$ is chosen on the right hand side such that $-\Delta y_d + \max(0, y_d) = f$ is fulfilled.

Case 2

$$
\begin{aligned}
&\min_{(y,u)} \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \\
&\text{s.t.} \quad -\Delta y + \max(5y, y|y|) - u = 0 \quad \text{in } \Omega , \\
&\text{with } y_d(x_1, x_2) = \frac{\sin\left(10\pi((x_1-\frac{1}{2})^2+(x_2-\frac{1}{2})^2)\right)}{\sqrt{\frac{1}{100}+(x_1-\frac{1}{2})^2+(x_2-\frac{1}{2})^2}} - 1 .
\end{aligned}
$$

Case 3

$$
\begin{aligned}
&\min_{(y,u)} \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \\
&\text{s.t.} \quad -\varepsilon\Delta y + \max(5y, y|y|) - u = 0 \quad \text{in } \Omega , \\
&\text{with } y_d(x_1, x_2) = \min\left(\max\left(\left|x_1 - \tfrac{1}{2}\right|, \left|x_2 - \tfrac{1}{2}\right|\right) - \tfrac{1}{4},\ 0\right) \text{ and } \varepsilon \geqslant 0, \text{ const.}
\end{aligned}
$$

Case 4

$$
\begin{aligned}
&\min_{(y,u)} \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \\
&\text{s.t.} \quad -\Delta y + \min(y, y|y|) - u = 0 \quad \text{in } \Omega , \\
&\text{with } \ y_d(x_1, x_2) = (x_1 - \tfrac{1}{2})^3 \cos(\pi x_2) .
\end{aligned}
$$

The numerical results for these cases, considering different values of the mesh size denoted by h, the penalty parameter α for the control in the objective functional, and the penalty parameter ν in the bi-quadratic penalty term, are presented in Tab. 1–4. Herein the abbreviation #SC denotes the total number of of switches between branch problems for the given parameter setting.

It can be observed that in almost all cases only a few Newton iterations are needed to solve the problem and to compute the stationary point.

A commonly used method for solving such non-smooth problems are semi-smooth Newton-like methods. Therefore, we also provide a comparison with results obtained with a semi-smooth Newton approach.
Case 1 represents an example taken from [6]. The parameters were adopted accordingly and the mesh size was reconstructed to match the one used in [6] in the best possible way. Tab. 1 shows a comparison between the non-regularized approach presented here and the proposed semi-smooth Newton's method in [6]. It can be observed that in the more involved example, according to [6], the approach presented here requires only one single Newton step and no switches between branch problems to compute the optimal solution. The semi-smooth Newton method on the other hand requires an average of three to five steps for the considered problem. Tab. 2.4 shows also the quality of the resulting approximation which is given the relative error $\|y_h - y\|_{L^2} \,/\, \|y\|_{L^2}$.
The fact that SALi does not require any switches between branch problems is mainly due to the fact that the reformulation described in Tab. 2.4 makes it possible to exploit as much information as possible given by the optimization problem and in particular by the given desired state y_d. The initial choice of the σ_i motivated by the desired state already provides the perfect guess for the σ_i. Since the desired state is reachable by the given state equation no switches between branch problems are required and the optimal solution can be computed by solving the initial branch problem, which is already the final one. Therefore, the convergence in just one Newton step in Tab. 1 is not surprising either. Due to the desired state y_d being reachable, we know the respective branch of the corresponding absolute value beforehand. Hence the state equation could also be written as

$$-\Delta y + \tfrac{1}{2}(y + \mathsf{sign}(y_d)y) - u = f \ .$$

Note that this might be a different PDE than the original one, but in the tracking type optimization context both optimization problems will attain the same solution. Modifying the state equation this way, we receive a quadratic objective with a linear constraint in y and hence convergence in just one Newton step. For similar reason we can observe convergence in two Newton steps in Tab. 5.
It is important to highlight that having knowledge about the optimal branch given by σ^* corresponding to $\mathsf{sign}(y_d)$ beforehand and using this to define the initial values accordingly, the operator $r(\sigma^*, y, z, \lambda)$ given by Eq. (13) is already non-negative. Thus, the described switching method based on $r(\sigma_k, y, z, \lambda)$ leads to the desired result, i.e., no switching, since the initial $\bar{\sigma} = \sigma^*$ has already the correct sign.

			SALi			[6]
h	α	ν	$\frac{\|y_d - y_h\|_{L^2}}{\|y_d\|_{L^2}}$	#SC	# Newton	# Newton
3.009e-02	1e-4	50	5.764e-04	0	1	4
1.537e-02	1e-4	50	1.514e-04	0	1	5
7.728e-03	1e-4	50	3.790e-05	0	1	3
3.885e-03	1e-4	50	9.663e-06	0	1	3
3.009e-02	1e-4	100	5.764e-04	0	1	4
1.537e-02	1e-4	100	1.514e-04	0	1	5
7.728e-03	1e-4	100	3.790e-05	0	1	3
3.885e-03	1e-4	100	9.663e-06	0	1	3
7.728e-03	1e-4	500	3.790e-05	0	1	3
7.728e-03	1e-2	100	8.106e-05	0	1	2
7.728e-03	1e-3	100	6.609e-05	0	1	2
7.728e-03	1e-5	100	1.237e-05	0	1	5
7.728e-03	1e-6	100	3.056e-06	0	1	no conv.

Table 1: Numerical results in Case 1.

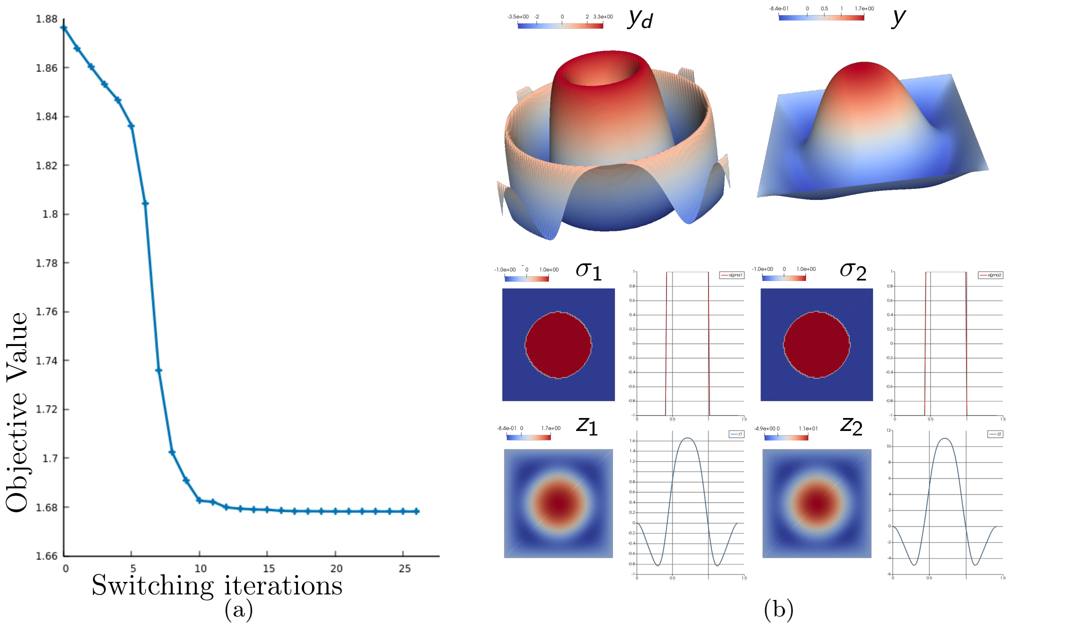

Figure 2: (a) History of the objective function value with respect to the branch problem switches corresponding to the parameters given in the first row in Tab. 2. (b) Final iteration step with final branch problem and resulting solution for y, z_1 and z_2.

The numerical results for Case 2 are given in Tab. 2. In this demanding case, where a genuine non-linear and non-smooth operator in the PDE and an unreachable target function y_d occur, comparatively more switches between branch problems and also more Newton iterations are needed to compute the minimal solution.
It should be noted that in each test case, it is verified that the condition $\sigma z = \mathsf{abs}(z)$ in the integral sense holds for the resulting z. This is shown here only in an exemplary fashion for the second case in Fig. 2 and has also been computed for Case 4 in Tab. 5, where the maximal value of $\|\sigma_i z_i - |z_i|\|_{L^2}$ for $i = 1, 2$ is always fairly close to zero. Fig. 2(b) illustrates the last iteration step in σ_1, σ_2 as well as the resulting states y, z with an over line plot for the z and σ components showing how the prescribed signs are observed. The target function y_d is shown in the top left corner. In Fig. 2(a) one can see how the successive exploitation of the corresponding dual variables leads to the next branch problem which results in a successive reduction in the objective function value for all considered cases.
The method presented here also allows the treatment of optimization problems of the considered problem class with non-smooth target functions y_d as given in Case 3. Such target functions are not achievable due to the PDE constraint with the Laplace operator as differential operator. Nevertheless, in the example considered in Case 3, with non damped Laplacian, i.e., $\varepsilon \equiv 1$, no switches and only two Newton steps are required to calculate the minimum solution. The numerical results are given in Tab. 3.
However, if the Laplace operator is attenuated by a positive factor $\varepsilon < 1$, also less regular solutions for y are achievable. Tab. 4 shows the numerical results for different values for ε in Case 3.

h	α	ν	Objective	$\frac{\|\|y-y_d\|\|}{\|\|y_d\|\|}$	#SC	#Newton
1.537e-02	1e-4	100	1.678	8.027e-01	26	132
1.159e-02	1e-4	100	1.677	8.025e-01	28	141
7.071e-03	1e-4	100	1.677	8.023e-01	35	165
1.537e-02	1e-4	500	1.678	8.027e-01	33	118
1.159e-02	1e-4	500	1.677	8.025e-01	36	185
7.071e-03	1e-4	500	1.677	8.023e-01	46	225
1.159e-02	1e-6	100	0.379	2.963e-01	22	118
7.071e-03	1e-6	100	0.377	2.937e-01	34	178
1.159e-02	1e-6	500	0.381	2.976e-01	15	92
7.071e-03	1e-6	500	0.378	2.962e-01	31	175
7.071e-03	1e-7	500	0.127	1.800e-01	49	317

Table 2: Numerical results for smooth but non reachable y_d (Case 2).

h	α	Objective	$\|\|y-y_d\|\|_{L^2}$	$\max_{i=1,2}\{\|\|\sigma_i z_i - \|z_i\|\|\|_{L^2}\}$	#SC	# Newt.
7.071e-03	1e-2	1.158e-03	4.568e-02	8.9e-30	0	2
7.071e-03	1e-3	7.679e-04	3.431e-02	5.1e-28	0	2
1.537e-02	1e-4	3.889e-04	2.133e-02	2.0e-08	0	2
1.159e-02	1e-4	3.886e-04	2.132e-02	2.0e-08	0	2
7.071e-03	1e-4	3.885e-04	2.131e-02	1.9e-08	0	2
1.159e-02	1e-4	3.886e-04	2.132e-02	2.0e-08	0	2
7.071e-03	1e-4	3.885e-04	2.131e-02	1.9e-08	0	2
1.159e-02	1e-6	2.283e-05	3.690e-03	1.4e-06	0	3
7.071e-03	1e-6	2.277e-05	3.677e-03	1.4e-06	0	3
1.159e-02	1e-7	4.431e-06	1.574e-03	7.5e-06	0	3
7.071e-03	1e-7	4.397e-06	1.559e-03	7.3e-06	0	3

Table 3: Numerical results in Case 3 for $\varepsilon \equiv 1$.

h	ε	α	Objective	$\frac{\|\|y-y_d\|\|}{\|\|y_d\|\|}$	#SC	# Newton
1.537e-02	1e-1	1e-4	3.886e-04	4.181e-01	0	2
1.159e-02	1e-1	1e-4	3.886e-04	4.179e-01	0	2
7.071e-03	1e-1	1e-4	3.884e-04	4.177e-01	0	2
1.537e-02	1e-2	1e-4	3.888e-04	4.181e-01	0	2
1.159e-02	1e-2	1e-4	3.886e-04	4.179e-01	0	2
7.071e-03	1e-2	1e-4	3.884e-04	4.177e-01	0	2
1.537e-02	1e-4	1e-4	3.888e-04	4.181e-01	0	2
1.159e-02	1e-4	1e-4	3.886e-04	4.179e-01	0	2
7.071e-03	1e-4	1e-4	3.884e-04	4.177e-01	0	2
7.071e-03	1e-6	1e-4	3.886e-04	4.179e-01	0	2

Table 4: Numerical results in Case 3 with damped Laplacian.

h	α	Objective	$\frac{\|y-y_h\|_{L^2}}{\|y\|_{L^2}}$	$\max_{i=1,2}\{\|\sigma_i z_i - \lvert z_i\rvert\|_{L^2}\}$	#SC	#Newt.
2.8e-02	1e-02	5.572e-04	9.97e-01	1.1e-11	0	2
2.8e-02	1e-03	5.434e-04	9.73e-01	3.8e-10	0	2
2.8e-02	1e-04	4.750e-04	8.76e-01	5.0e-09	0	2
2.8e-02	1e-06	2.183e-04	5.54e-01	1.5e-08	0	2
1.4e-02	1e-02	5.565e-04	9.97e-01	9.7e-12	0	2
1.4e-02	1e-03	5.426e-04	9.73e-01	7.7e-11	0	2
1.4e-02	1e-04	4.737e-04	8.75e-01	6.2e-10	0	2
1.4e-02	1e-06	2.143e-04	5.47e-01	1.8e-09	0	2
1.4e-02	1e-08	7.463e-05	3.22e-01	2.6e-09	0	2
7.7e-03	1e-02	5.564e-04	9.97e-01	1.2e-11	0	2
7.7e-03	1e-03	5.425e-04	9.73e-01	2.9e-11	0	2
7.7e-03	1e-04	4.735e-04	8.75e-01	4.8e-12	0	2
7.7e-03	1e-06	2.133e-04	5.45e-01	1.2e-11	0	2
7.7e-03	1e-08	7.183e-05	3.12e-01	1.7e-11	0	2

Table 5: Numerical results for Case 4.

One may wonder how the algorithm reacts if one cannot extract the correct branch problem or the correct signs of the σ_i from the underlying a priori information as in Cases 1 and 3. We also tested Case 1 with some random initial $\bar{\sigma}_i$. Tab. 6 shows the numerical results for Case 1 with initial $\bar{\sigma}_1$ given by

$$\bar{\sigma}_1 = \begin{cases} +1, & \text{if } x_1 + x_2 > 1 \\ -1, & \text{otherwise .} \end{cases}$$

This shows in particular that the presented algorithm and more precisely the associated switching method performs as intended, even if we do not start with the proper branch problem. In any case Algo. 1 converges properly. However, provided the optimization problem provides already the correct sign for σ as for instance with tracking type objective functionals, it is always advisable to exploit this for the initial $\bar{\sigma}$.

h	α	ν	$\frac{\|y_d-y_h\|_{L^2}}{\|y_d\|_{L^2}}$	#SC	# Newton
3.009e-02	1e-4	50	5.711e-04	1	3
1.537e-02	1e-4	50	4.430e-04	1	3
7.728e-03	1e-4	50	4.911e-04	1	3
3.885e-03	1e-4	50	5.103e-04	1	3
3.009e-02	1e-4	100	5.750e-04	1	3
1.537e-02	1e-4	100	4.577e-04	1	3
7.728e-03	1e-4	100	5.108e-04	1	3
3.885e-03	1e-4	100	5.335e-04	1	3
7.728e-03	1e-4	500	6.775e-04	1	3
7.728e-03	1e-2	100	1.519e-03	1	3
7.728e-03	1e-3	100	1.027e-03	1	3
7.728e-03	1e-5	100	2.220e-04	1	3
7.728e-03	1e-6	100	1.403e-04	1	3

Table 6: Numerical results in Case 1 with initial σ not corresponding to $sgn(y_d)$.

As additional observation, Tab. 1, Tab. 2 and Tab. 5 suggest a further special property of the SAli algorithm, namely mesh independence. Regardless of the mesh size, the behavior for the relative error $\|y_d - y\|_{L^2} / \|y_d\|_{L^2}$ with respect to different parameters α remains the same. The mesh independence can also be observed in Tab. 1 for Case 1 since independent of the mesh size only one Newton step is required. The same applies to Case 4 presented in Tab. 5. There, however, independent of the mesh size always two Newton steps are required, since the nesting of the absolute value occurring in case 4 leads to a quadratic term appearing in the PDE. Moreover, with Tab. 5 it is clearly evident that in each parameter setting the desired condition $\bar{\sigma}_i z_i = \mathsf{abs}(z_i)$ for $i \in \{1, \dots, s\}$ is met in the integral sense.

6 Conclusion and Outlook

We presented a new approach based on successive abs-linearization for the solution of optimization problems constrained by non-smooth PDEs. For the considered class of genuinely non-smooth problems, this approach enables the optimization without any substitute assumptions and regularizations for the non-smoothness. The key idea is to appropriately decompose the non-smooth problem into smooth branch problems, which can be solved by classical smooth optimization problems. Optimality conditions for the considered formulations were derived and discussed. Solving the current branch problem, exploiting standard optimality conditions for the smooth case as well as using an indicator strategy to determine the next branch problem, which results in a successive reduction in the objective function value for all tested cases and leads to the minimal solution. By treating the inequality condition with a bi-quadratic penalty approach the sign condition could easily be incorporated into the algorithmic framework. The type of discretization employed here was also presented and critically examined. Finally, several non-smooth PDE-constrained problems that fit into the considered setting were discussed. The corresponding numerical results clearly show also the resulting mesh independence of the presented method. However, a comprehensive convergence analysis for the continuous case as well as the more detailed analytical investigation of the optimality conditions and their classification, such as the comparison with the optimality conditions of MPECS, remain the subject of current research.

Acknowledgments

This work was partially funded by the DFG priority program SPP 1962 within the project "Shape Optimization for Maxwell's Equations Including Hysteresis Effects in the Material Laws (HyLa)".

A Farkas Alternative in Hilbert Spaces

We recall an extended version of the Farkas' Lemma, which gives necessary and sufficient conditions on the solvability of a linear system. The following is derived from [7].

Lemma A.1 (Farkas' Lemma)**.** *Let H and V be some Hilbert spaces over a domain $\Omega \subset \mathbb{R}^n$, and denote by H^* and V^* the topological dual of H and V respectively. Furthermore, let K be a convex cone in H and $A : H \to V$ a bounded linear operator. If $A(K)$ is weakly closed, then the following are equivalent:*

(a) The system $Ax = b$ has a solution $x \in K^*$.

(b) $\langle b, v^* \rangle \geqslant 0$ for all v^* with $A^* v^* \in K^*$.

For a more general version of the Farkas lemma see for instance [9] and [18].
Note, that the equivalence in Lem. A.1 also indicates, that only one of the two following properties can hold:

(a) The system $Ax = b$ has a solution $x \in K^*$.

(b') $\exists v^*$ with $A^* v^* \in K^*$ such that $\langle b, v^* \rangle < 0$.

This is known as the Farkas alternative.

References

[1] E.L. Allgower, K. Bohmer, F.A. Potra, and W.C. Rheinboldt. A mesh-independence principle for operator equations and their discretizations. *SIAM Journal on Numerical Analysis*, 23(1):160–169, 1986.

[2] I. Antal and J. Karátson. A mesh independent superlinear algorithm for some nonlinear nonsymmetric elliptic systems. *Computers & Mathematics with Applications*, 55(10):2185–2196, 2008.

[3] V. Barbu. *Optimal Control of Variational Inequalities.* Pitman, Boston, 1984.

[4] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer Series in Operations Research and Financial Engineering. Springer New York, 2013.

[5] E. Börgens, C. Kanzow, P. Mehlitz, and G. Wachsmuth. New constraint qualifications for optimization problems in banach spaces based on cone continuity properties. Technical report, SPP1962 Preprint SPP1962-128, 2019.

[6] C. Christof, C. Clason, C. Meyer, and S. Walther. Optimal control of a non-smooth semilinear elliptic equation. *Mathematical Control and Related Fields*, 8(1):247–276, 2018.

[7] B. Craven and J. Koliha. Generalizations of Farkas' theorem. *SIAM Journal on Mathematical Analysis*, 8:983–997, 1977.

[8] J.C. De Los Reyes. On the optimal control of some nonsmooth distributed parameter systems arising in mechanics. *GAMM-Mitteilungen*, 40(4):268–286, 2018.

[9] N. Dinh and V. Jeyakumar. Farkas' lemma: three decades of generalizations for mathematical optimization. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, 22(1), 2014.

[10] S. Fiege, A. Walther, and A. Griewank. An algorithm for nonsmooth optimization by successive piecewise linearization. *Mathematical Programming, Series A*, 2018. DOI: 10.1007/s10107-018-1273-5.

[11] A. Griewank. On stable piecewise linearization and generalized algorithmic differentiation. *Optimization Methods and Software*, 28(6):1139–1178, 2013.

[12] A. Griewank and A. Walther. First and second order optimality conditions for piecewise smooth objective functions. *Optimization Methods and Software, 31(5):904-930*, 2016.

[13] A. Griewank and A. Walther. Relaxing kink qualifications and proving convergence rates in piecewise smooth optimization. *SIAM Journal on Optimization*, 29(1):262–289, 2019.

[14] M. Gunzburger. *Perspectives in Flow Control and Optimization.* SIAM, 2003.

[15] M. Hintermüller and M. Ulbrich. A mesh-independence result for semismooth Newton methods. *Mathematical Programming*, 101(1):151–184, 2004.

[16] J. Jahn. *Introduction to the Theory of Nonlinear Optimization.* Springer, 2007.

[17] F. Kikuchi, K. Nakazato, and T. Ushijima. Finite element approximation of a nonlinear eigenvalue problem related to MHD equilibria. *Japan Journal of Applied Mathematics*, 1(2):369–403, 1984.

[18] J.-B. Lasserre. A Farkas lemma without a standard closure condition. *SIAM Journal on Control and Optimization*, 35, 1997.

[19] A. Logg, K.-A. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Element Method.* Springer, 2012.

[20] R. Temam. A non-linear eigenvalue problem: the shape at equilibrium of a confined plasma. *Archive for Rational Mechanics and Analysis*, 60:51–73, 1975.

[21] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*. Graduate Studies in Mathematics. American Mathematical Society, 2010.

[22] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. Society for Industrial and Applied Mathematics, 2011.

[23] M. Ulbrich and S. Ulbrich. Primal-dual interior-point methods for PDE-constrained optimization. *Mathematical Programming*, 117(1):435–485, 2009.

[24] A. Walther, O. Weiß, A. Griewank, and S. Schmidt. Nonsmooth optimization by successive abs-linearisation in function spaces. *Applicable Analysis*, 2020. https://doi.org/10.1080/00036811.2020.1738397.

[25] O. Weiß. *Nonsmooth Optimization by Abs-Linearization in Reflexive Function Spaces*. PhD thesis, Humboldt-Universität zu Berlin, 2020. In preparation.

[26] E. Zeidler. *Applied Functional Analysis. Applications to Mathematical Physics*. Springer, 1995.