Deutsche Forschungsgemeinschaft
Priority Programme 1962

A Topological Derivative-Based Algorithm to Solve Optimal Control Problems with $L^0(\Omega)$ Control Cost

Daniel Wachsmuth

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

Preprint Number SPP1962-201

received on November 22, 2022

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/

A topological derivative-based algorithm to solve optimal control problems with $L^0(\Omega)$ control cost

Daniel Wachsmuth*

November 22, 2022

Abstract. In this paper, we consider optimization problems with L^0-cost of the controls. Here, we take the support of the control as independent optimization variable. Topological derivatives of the corresponding value function with respect to variations of the support are derived. These topological derivatives are used in a novel algorithm. In the algorithm, topology changes happen at large values of the topological derivative. Convergence results are given.

Keywords. Topological derivative, sparse optimal control, L^0 optimization.

MSC classification. 49M05, 49K40, 65K10

1 Introduction

In this paper we are interested in the following optimal control problem: Minimize

$$\min \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \beta\|u\|_0$$

over all (y, u) satisfying

$$-\Delta y = u \quad \text{a.e. on } \Omega$$

and

$$u_a \le u \le u_b.$$

Here, $\|u\|_0$ is the measure of the support of u. This optimal control problem can be interpreted in the context of optimal actuator placement: Find a (possibly small) set $A \subseteq \Omega$ such that controls supported on A can still minimize a certain objective functional.

In this work, we will take the support of the control u as own optimization variable $A \subseteq \Omega$. In addition, we will allow for a more general control problem as above. The abstract problem we are interested in is: Minimize with respect to $u \in L^2(\Omega)$ and $A \subseteq \Omega$ the functional

$$J(u, A) := \frac{1}{2}\|S(\chi_A u) - y_d\|^2_H + \int_\Omega g(u(x)) + \chi_A(x)\beta(x)\,\mathrm{d}x, \tag{1.1}$$

*Institut für Mathematik, Universität Würzburg, 97074 Würzburg, Germany, daniel.wachsmuth@mathematik.uni-wuerzburg.de. This research was partially supported by the German Research Foundation DFG under project grant Wa 3626/3-2.

where $S : L^2(\Omega) \to H$ is a solution operator of a linear partial differential equation, H is a Hilbert space, $y_d \in H$ is given, $g : \mathbb{R} \to \bar{\mathbb{R}}$ is a strongly convex function, and $\beta \in L^1(\Omega)$ is non-negative.

Given $A \subseteq \Omega$, the functional $u \mapsto J(u, A)$ admits minimizers, and we can study the value function

$$J(A) := \min_{u \in L^2(\Omega)} J(u, A). \tag{1.2}$$

We will investigate topological derivatives of the value function. In additions, we are interested in the shape optimization problem

$$\min_{A \subseteq \Omega} J(A). \tag{1.3}$$

The topological derivative is the main result of Theorem 4.2. It can be extended to non-strongly convex g, see Theorem 5.5. These results generalize available results in the literature [2, 5, 4, 10], as we allow for non-smooth g and incorporate control constraints. In comparison to earlier work, we will use less smoothness assumptions, in particular no continuity of controls and adjoints is required.

The concept of topological derivatives goes back to the seminal work [13]. It was applied to an optimal control problem in [14], which is different than ours: there, the observation term in the cost functional was taken on the set A, i.e., the cost functional contained $\frac{1}{2}\|y - y_d\|^2_{L^2(A)}$. In these works, asymptotic analysis with respect to radius of small inclusions/exclusions was performed. Minimax-differentiability to compute topological derivatives was applied to optimal control problems in [4, 5, 16]. These results cannot be applied to problems with control constraints. In addition, the abstract theory only allows to compute the topological derivative at one fixed point, which necessitates continuity assumptions on that point. In our proof, we get the topological derivative at almost all $x \in \Omega$ at once using the Lebesgue differentiation theorem. Moreover, we can allow for control constraints and non-smooth functions g. Let us also mention [10], where the goal was to obtain controls that are robust with respect to perturbations of the initial state of the state. One of the motivations of this work was the question, how control constraints can be incorporated in the setting of [10]. It would be interesting to see whether our approach also works in this robust control framework.

In addition to the development of the topological derivative, we also investigate a novel algorithm to solve the problem at hand. In the algorithm, variations of a given set $A_k \subseteq \Omega$ are performed at points, where the topological derivative has the wrong sign and has large absolute value. We incorporate a line-search technique involving a step-size $t > 0$, which is required to satisfy an Armijo-like descent condition. We emphasize that small values of the parameter may still lead to topological changes of the set far away from the current boundary. This is different to the (simplified) level-set method as considered, e.g., in [2, 3, 8, 11]. There, small values of t cannot result in topology changes but only lead to boundary variations. This observation was the second motivation of this work: to develop an algorithm, where the line-search corresponds to the derivation of the topological derivative in the following way: Let $A_{k,t}$ be the candidate for the next iterate produced using the step-size t. Then the difference quotient $\frac{1}{t}(J(A_{k,t}) - J(A))$ should converge for $t \searrow 0$ to an expression resembling the topological derivative. Such a result is not true for level set methods. Our method is described in Section 6, see Algorithm 6.4. The mentioned estimate of $\frac{1}{t}(J(A_{k,t}) - J(A))$ is Lemma 6.2.

We also give a convergence result in Theorem 6.7. Here, we get the following interesting result: if the sequence of characteristic functions (χ_{A_k}) of the iterates $A_k \subseteq \Omega$ does *not* converge strongly in $L^1(\Omega)$, then the sequence (A_k) is a minimizing sequence for (1.3), see Corollary 6.8. To the best of our knowledge there are no such convergence results in the literature. The sole exception being [2], where the convergence analysis is done for a lower-semicontinuous envelope of the level-set functional. In fact, the question of convergence of topological derivative-based methods is mentioned as an open problem in [11, Section 5].

Notation

We will denote the Lebesgue measure of a measurable set $A \subseteq \mathbb{R}^d$ by $|A|$. For $r > 0$ and $x \in \mathbb{R}^d$, let $B_r(x)$ be the open ball with radius r centered at x. Its Lebesgue measure will be denoted by $|B_r|$. We set $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$. For a function $g : \mathbb{R} \to \bar{\mathbb{R}}$, we set $\operatorname{dom} g := \{u \in \mathbb{R} : \; g(u) < +\infty\}$. The subdifferential of a convex function g at u will be denoted by $\partial g(u)$. We will write $x^+ := \max(x, 0)$ and $x^- := \min(x, 0)$ for $x \in \mathbb{R}$.

2 Assumptions and preliminary results

Throughout this paper, we will work with the following assumptions concerning the problem (1.1)

(A1) $\Omega \subseteq \mathbb{R}^d$ is Lebesgue measurable with $|\Omega| < \infty$.

(A2) H is a real Hilbert space, $S \in \mathcal{L}(L^2(\Omega), H)$, $y_d \in H$.

(A3) $g : \mathbb{R} \to \bar{\mathbb{R}}$ is proper, convex, lower semi-continuous. In addition, $g(u) \geq 0$ for all $u \in \mathbb{R}$, and $g(u) = 0$ if and only if $u = 0$.

(A4) There is $\mu > 0$ such that

$$\mu\lambda(1-\lambda)|u-v|^2 + g(\lambda u + (1-\lambda)v) \leq \lambda g(u) + (1-\lambda) g(v) \quad \forall u, v \in \operatorname{dom} g.$$

(A5) There is $q > 6$ such that $S^*S \in \mathcal{L}(L^2(\Omega), L^q(\Omega))$, where $S^* \in \mathcal{L}(H, L^2(\Omega))$ denotes the Hilbert space-adjoint of S,

(A6) $\beta \in L^1(\Omega)$.

Let us comment on these assumptions. As we plan to use the Lebesgue differentiation theorem, we assume that the underlying measure space is induced by the Lebesgue measure of $\mathbb{R}^d$ in **(A1)**. Conditions **(A2)**, **(A3)**, **(A4)** imply the well-posedness of the problem $\min_{u \in L^2(\Omega)} J(u, A)$ for fixed A. Assumption **(A4)** is strong convexity of the function g. The results of the paper are still valid in the non-strong convex case ($\mu = 0$) under slightly strengthened assumptions on g and S, we will comment on this in Section 5. Condition **(A5)** implies that certain remainder terms in the expansion of topological derivatives are of higher order, see Theorem 4.2.

We will explicitly mention in upcoming, important results (theorems and propositions), which of these assumptions are used. If the strong convexity assumption is not mentioned then μ can be taken equal to zero.

2.1 Existence of minimizers of J for fixed A

Let $A \subseteq \Omega$ measurable be given. Here, we consider the problem

$$\min_{u \in L^2(\Omega)} J(u, A). \tag{$\mathbf{P}_A$}$$

where J is given by (1.1). Note that due to the construction of J and **(A3)**, we have

$$J(\chi_A u, A) \le J(u, A) \tag{2.1}$$

for all $u \in L^2(\Omega)$.

Due to strong convexity of g and $g(0) = 0$ by **(A3)**, **(A4)**, we have

$$g(u) \ge \mu |u|^2 \quad \forall u \in \operatorname{dom} g. \tag{2.2}$$

Proposition 2.1. *Assume **(A1)**, **(A2)**, **(A3)**, **(A4)**. Let $A \subseteq \Omega$ measurable be given. Then there is a uniquely determined minimizer u_A of ($\mathbf{P}_A$).*

Moreover, $u_A = 0$ almost everywhere on $\Omega \setminus A$.

Proof. Due to (2.2), minimizing sequences of $J(\cdot, A)$ are bounded in $L^2(\Omega)$. In addition, $u \mapsto J(u, A)$ is weakly lower semi-continuous from $L^2(\Omega)$ to $\mathbb{R}$ because of **(A2)** and **(A3)**. The existence of solutions follows now by standard arguments. Uniqueness of solutions is a consequence of strong convexity of g **(A4)**. The last claim follows from (2.1). □

Note that the last claim implies

$$\chi_A u_A = u_A. \tag{2.3}$$

In all what follows, we will not make use of the unique solvability of ($\mathbf{P}_A$). We will just use that u_A is any solution of ($\mathbf{P}_A$).

2.2 Optimality conditions for ($\mathbf{P}_A$)

Let $A \subseteq \Omega$ measurable be given, and let u_A be a solution of ($\mathbf{P}_A$). Let us denote the associated state by

$$y_A := S(\chi_A u) \tag{2.4}$$

and adjoint state by

$$p_A := S^*(y_A - y_d) = S^*(S(\chi_A u) - y_d). \tag{2.5}$$

Let $u \in L^2(\Omega)$ and $B \subseteq \Omega$ be given. Let $y := S(\chi_B u)$. Then by elementary calculations, we find

$$\begin{aligned} \frac{1}{2}\|y - y_d\|_H^2 - \frac{1}{2}\|y_A - y_d\|_H^2 &= (y_A - y_d, y - y_A)_H + \frac{1}{2}\|y - y_A\|_H^2 \\ &= (p_A, \chi_B u - \chi_A u_A) + \frac{1}{2}\|y - y_A\|_H^2. \end{aligned} \tag{2.6}$$

For $B = A$, we get

$$\frac{1}{2}\|y - y_d\|_H^2 - \frac{1}{2}\|y_A - y_d\|_H^2 = (p_A, \chi_A(u - u_A)) + \frac{1}{2}\|y - y_A\|_H^2.$$

Hence, $\chi_A p_A \in L^2(\Omega)$ is the Frechet derivative of $u \mapsto \frac{1}{2}\|S(\chi_A u) - y_d\|_H^2$ at u_A.

Proposition 2.2. *Assume (A1), (A2), (A3). Let $A \subseteq \Omega$ measurable and let u_A by the unique solution of $(\mathbf{P}_A)$. Let p_A be given by (2.5). Then it holds*

$$-\chi_A(x)p_A(x) \in \partial g(u(x)) \quad \text{for almost all } x \in \Omega \tag{2.7}$$

and

$$u_A(x) = \operatorname*{arg\,min}_{u \in \mathbb{R}} \chi_A(x)p_A(x) \cdot u + g(u) \quad \text{for almost all } x \in \Omega. \tag{2.8}$$

Proof. Let us denote $G(u) := \int_\Omega g(u(x))\,\mathrm{d}x$. As argued above, $\chi_A p_A \in L^2(\Omega)$ is the Frechet derivative of $u \mapsto \frac{1}{2}\|S(\chi_A u) - y_d\|_H^2$ at u_A. Then by well-known results, see, e.g., [7, Proposition II.2.2], we get $-p_A \in \partial G(u_A)$. Using [12, Theorem 3A], this is equivalent to the pointwise a.e. inclusion (2.7), which in turn is equivalent to (2.8). ☐

The condition (2.8) can be interpreted as Pontryagin's maximum principle for $(\mathbf{P}_A)$.

2.3 Boundedness results for solutions of $(\mathbf{P}_A)$

In this section, we will derive bounds on (u_A, y_A, p_A) that are uniform with respect to $A \subseteq \Omega$.

Lemma 2.3. *There is $M > 0$ such that*

$$\|y_A - y_d\|_H + \|u_A\|_{L^2(\Omega)} \leq M$$

for all $A \subseteq \Omega$.

Proof. This follows directly from $J(A, u_A) \leq J(A, 0)$ and (2.2). ☐

Corollary 2.4. *There is $P > 0$ such that*

$$\|u_A\|_{L^q(\Omega)} \leq P, \quad \|p_A\|_{L^q(\Omega)} \leq P$$

for all $A \subseteq \Omega$, where q is from (A5).

Proof. First, we have

$$\|p_A\|_{L^q(\Omega)} \leq \|S^*S\|_{\mathcal{L}(L^2(\Omega),L^q(\Omega))}\|u_A\|_{L^2(\Omega)} \leq \|S^*S\|_{\mathcal{L}(L^2(\Omega),L^q(\Omega))}M =: P,$$

with M as in Lemma 2.3. Using (2.8) with $u = 0$, **(A3)**, and (2.2), we have for almost all $x \in \Omega$

$$\mu|u_A(x)|^2 \leq g(u_A(x)) \leq -p_A(x)u_A(x)$$

which implies $\mu|u_A(x)| \leq |p_A(x)|$ and $\|u_A\|_{L^q(\Omega)} \leq P$. ☐

3 Analysis of the value function

In this section, we will investigate stability properties of $A \mapsto (u_A, y_A, p_A)$, where y_A and p_A solve (2.4) and (2.5). The goal is to derive formulas for the topological derivative of $A \mapsto J(A)$, where $J(A)$ is the value function defined in (1.2) by

$$J(A) = \min_{u \in L^2(\Omega)} J(u, A).$$

For brevity, we refer to tuples (u_A, y_A, p_A), where u_A solves $(\mathbf{P}_A)$ and y_A, p_A are given by (2.4) and (2.5) as solutions of $(\mathbf{P}_A)$.

3.1 Sensitivity analysis of ($\mathbf{P}_A$) with respect to A

Let us start with the following preliminary expansion.

Lemma 3.1. *Let $A, B \subseteq \Omega$, and let (u_A, y_A, p_A) and (u_B, y_B, p_B) be solutions of ($\mathbf{P}_A$) and ($\mathbf{P}_B$). Then it holds*

$$J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|_H^2 \\ = \int_\Omega g(u_A) - g(u_B) + \chi_A p_A(u_A - u_B) + (\chi_A - \chi_B)(\beta + p_A u_B)\,\mathrm{d}x.$$

Proof. Doing the expansion of $y \mapsto \frac{1}{2}\|y - y_d\|_H^2$ similarly as in (2.6), we have

$$J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|_H^2 \\ = \int_\Omega g(u_A) - g(u_B) + p_A(\chi_A u_A - \chi_B u_B) + (\chi_A - \chi_B)\beta\,\mathrm{d}x. \quad (3.1)$$

In addition, we have due to (2.3)

$$\begin{aligned}\int_\Omega p_A(\chi_A u_A - \chi_B u_B)\,\mathrm{d}x &= \int_\Omega p_A(u_A - u_B)\,\mathrm{d}x \\ &= \int_\Omega \chi_A p_A(u_A - u_B) + (1 - \chi_A)p_A(u_A - u_B)\,\mathrm{d}x \\ &= \int_\Omega \chi_A p_A(u_A - u_B) - (\chi_B - \chi_A)p_A u_B\,\mathrm{d}x,\end{aligned}$$

which is the claim. □

Lemma 3.2. *Let $A, B \subseteq \Omega$, and let (u_A, y_A, p_A) and (u_B, y_B, p_B) be solutions of ($\mathbf{P}_A$) and ($\mathbf{P}_B$). Then it holds for almost all $x \in \Omega$*

$$\mu|u_B(x) - u_A(x)| \le |p_B(x) - p_A(x)|.$$

Proof. Due to strong convexity of g by **(A4)** and (2.7), we have for almost all $x \in \Omega$

$$\mu|u_B(x) - u_A(x)|^2 \le -(\chi_A(x)p_A(x) - \chi_B(x)p_B(x))(u_A(x) - u_B(x)),$$

which proves the claim. □

Lemma 3.3. *Let $A, B \subseteq \Omega$, and let (u_A, y_A, p_A) and (u_B, y_B, p_B) be solutions of ($\mathbf{P}_A$) and ($\mathbf{P}_B$). Then it holds*

$$\mu\|u_B - u_A\|_{L^2(\Omega)}^2 + \|y_B - y_A\|_H^2 \le \int_\Omega (\chi_A - \chi_B)(p_A u_B - p_B u_A)\,\mathrm{d}x$$

*with μ from **(A4)**.*

Proof. Due to Lemma 3.1, we have

$$J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|_H^2 \\ = \int_\Omega g(u_A) - g(u_B) + \chi_A p_A(u_A - u_B) + (\chi_A - \chi_B)(\beta + p_A u_B)\,\mathrm{d}x$$

as well as

$$J(B,u_B) - J(A,u_A) + \frac{1}{2}\|y_A - y_B\|_H^2$$
$$= \int_\Omega g(u_B) - g(u_A) + \chi_B p_B(u_B - u_A) + (\chi_B - \chi_A)(\beta + p_B u_A)\,\mathrm{d}x.$$

Adding both equations gives

$$\|y_A - y_B\|_H^2 = \int_\Omega (\chi_A p_A - \chi_B p_B)(u_A - u_B) + (\chi_A - \chi_B)(p_A u_B - p_B u_A)\,\mathrm{d}x.$$

Due to strong convexity of g by **(A4)** and (2.7), we have

$$\int_\Omega (\chi_A p_A - \chi_B p_B)(u_A - u_B) \le -\mu\|u_A - u_B\|_{L^2(\Omega)}^2, \tag{3.2}$$

and the claim is proven. □

Note that the previous result remains true with $\mu = 0$ in the non-convex case. Now we can prove the main result of this section, which is a stability estimate of solutions of $(\mathbf{P}_A)$ with respect to variations of A (or χ_A). In the proof, we will use the fact that for characteristic functions

$$\|\chi_A - \chi_B\|_{L^s(\Omega)} = \|\chi_A - \chi_B\|_{L^1(\Omega)}^{\frac{1}{s}} \quad \forall s \in (1,\infty).$$

Theorem 3.4. *Assume **(A1)**, **(A2)**, **(A3)**, **(A4)**, **(A5)**. There is a constant $K > 0$ such that for all $A, B \subseteq \Omega$*

$$\|p_A - p_B\|_{L^q(\Omega)} + \|u_A - u_B\|_{L^2(\Omega)} + \|y_B - y_A\|_{L^2(\Omega)} \le K\|\chi_A - \chi_B\|_{L^1(\Omega)}^{\frac{1}{2}-\frac{1}{q}},$$

*where (u_A, y_A, p_A) and (u_B, y_B, p_B) are solutions of $(\mathbf{P}_A)$ and $(\mathbf{P}_B)$, and q is from **(A5)**.*

Proof. From **(A5)**, we find

$$\|p_A - p_B\|_{L^q(\Omega)} \le \|S^*S\|_{\mathcal{L}(L^2(\Omega),L^q(\Omega))}\|u_A - u_B\|_{L^2(\Omega)}.$$

Define $\mu' := \mu/\|S^*S\|_{\mathcal{L}(L^2(\Omega),L^q(\Omega))}^2$. Let s be such that $\frac{1}{s} + \frac{1}{q} + \frac{1}{2} = 1$. From the inequality of Lemma 3.3, we obtain with Hölder's inequality

$$\begin{aligned}
&\frac{\mu'}{2}\|p_A - p_B\|_{L^q(\Omega)}^2 + \frac{\mu}{2}\|u_B - u_A\|_{L^2(\Omega)}^2 + \|y_B - y_A\|_H^2 \\
&\qquad \le \mu\|u_B - u_A\|_{L^2(\Omega)}^2 + \|y_B - y_A\|_H^2 \\
&\qquad \le \int_\Omega (\chi_A - \chi_B)(p_A u_B - p_B u_A)\,\mathrm{d}x \\
&\qquad \le \|\chi_A - \chi_B\|_{L^s(\Omega)}(\|p_A\|_{L^q(\Omega)}\|u_B - u_A\|_{L^2(\Omega)} \\
&\qquad\qquad + \|p_A - p_B\|_{L^q(\Omega)}\|u_A\|_{L^2(\Omega)}) \\
&\qquad \le (P + M)\|\chi_A - \chi_B\|_{L^1(\Omega)}^{\frac{1}{s}}(\|u_B - u_A\|_{L^2(\Omega)} + \|p_A - p_B\|_{L^q(\Omega)}),
\end{aligned}$$

where P and M are from Corollary 2.4 and Lemma 2.3, and the claim is proven. □

3.2 Expansions of the value function

Let us define $H : \mathbb{R} \times \mathbb{R} \to \bar{\mathbb{R}}$ by

$$H(u,p) := p \cdot u + g(u).$$

This function reminds of the Hamiltonian of optimal control problems. In addition, we need its infimum with respect to u,

$$\min_{u\in\mathbb{R}} H(u,p) = \min_{u\in\mathbb{R}} (p\cdot u + g(u)) = -\sup_{u\in\mathbb{R}} (-p\cdot u - g(u)) = -g^*(-p),$$

where g^* is the convex conjugate to g. The existence of this minimum follows from the coercivity of g, see (2.2). Let us denote this function by $\bar{H}$, i.e.,

$$\bar{H}(p) := \min_{u\in\mathbb{R}} H(u,p) = -g^*(-p).$$

We will need some Lipschitz estimate of $\bar{H}$.

Lemma 3.5. *Let $A, B \subseteq \Omega$, and let (u_A, y_A, p_A) and (u_B, y_B, p_B) be solutions of ($\mathbf{P}_A$) and ($\mathbf{P}_B$). Then we have*

$$\|\bar{H}(p_A) - \bar{H}(p_B)\|_{L^{q/2}(\Omega)} \le P\|p_A - p_B\|_{L^q(\Omega)} \le PK\|\chi_A - \chi_B\|_{L^1(\Omega)}^{\frac{1}{2}-\frac{1}{q}},$$

where P and K are from Corollary 2.4 and Theorem 3.4, respectively.

Proof. Let $p_1, p_2 \in \mathbb{R}$ be given. Let $u_i = \arg\min_{v\in[u_a,u_b]} H(p_i, v)$ for $i = 1,2$. Then we get by the properties of $\bar{H}$

$$\bar{H}(p_1) \le H(p_1,u_2) = (p_1 - p_2)u_2 + H(p_2,u_2) = (p_1-p_2)u_2 + \bar{H}(p_2).$$

This implies

$$\bar{H}(p_2) \le -(p_1 - p_2)u_1 + \bar{H}(p_1)$$

by exchanging (p_1, u_1) and (p_2, u_2) in the above estimate. Summarizing, we obtain

$$|\bar{H}(p_1) - \bar{H}(p_2)| \le |p_1 - p_2| \max(|u_1|, |u_2|).$$

Using Corollary 2.4 yields the claim. □

We will proceed with the following expansion of the value function. Note that in the non-strongly convex case, i.e., without assuming **(A4)**, the claim is valid with $\mu = 0$.

Lemma 3.6. *Let $A, B \subseteq \Omega$, and let (u_A, y_A, p_A) and (u_B, y_B, p_B) be solutions of ($\mathbf{P}_A$) and ($\mathbf{P}_B$). Then it holds*

$$\begin{aligned} J(A,u_A) - J(B,u_B) &+ \frac{\mu}{2}\|u_B - u_A\|_{L^2(A\cap B)}^2 + \frac{1}{2}\|y_B - y_A\|_H^2 \\ &\le \int_{A\setminus B} \beta + \bar{H}(p_A)\,\mathrm{d}x + \int_{B\setminus A} -\beta - \bar{H}(p_B) - (p_A - p_B)u_B\,\mathrm{d}x. \end{aligned}$$

Proof. From Lemma 3.1 we get

$$\begin{aligned} J(A, u_A) - J(B, u_B) &+ \frac{1}{2}\|y_B - y_A\|_H^2 \\ &= \int_\Omega g(u_A) - g(u_B) + \chi_A p_A (u_A - u_B) + (\chi_A - \chi_B)(\beta + p_A u_B)\,\mathrm{d}x. \end{aligned} \tag{3.3}$$

We will now split the integral on the right-hand side into integrals on $A \cap B$, $A \setminus B$, and $B \setminus A$. This is sufficient as the integrand vanishes outside of $A \cup B$. For the integral on $A \cap B$, we can use the optimality condition (2.7) as well as the strong convexity of g to obtain

$$\int_{A\cap B} g(u_A) - g(u_B) + \chi_A p_A (u_A - u_B)\,\mathrm{d}x \le -\frac{\mu}{2}\|u_B - u_A\|^2_{L^2(A\cap B)}. \tag{3.4}$$

Moreover, u_B vanishes on $A \setminus B$, while u_A vanishes on $B \setminus A$. This allows to simplify

$$\begin{aligned} \int_\Omega g(u_A) - g(u_B) + \chi_A p_A (u_A - u_B)\,\mathrm{d}x &\le -\frac{\mu}{2}\|u_B - u_A\|^2_{L^2(A\cap B)} \\ &\quad + \int_{A\setminus B} g(u_A) + p_A u_A\,\mathrm{d}x - \int_{B\setminus A} g(u_B)\,\mathrm{d}x \end{aligned} \tag{3.5}$$

In addition, we have

$$\begin{aligned} \int_\Omega (\chi_A - \chi_B)(\beta + p_A u_B)\,\mathrm{d}x &= \int_{A\setminus B} \beta + p_A u_B\,\mathrm{d}x - \int_{B\setminus A} \beta + p_A u_B\,\mathrm{d}x \\ &= \int_{A\setminus B} \beta\,\mathrm{d}x - \int_{B\setminus A} \beta + p_A u_B\,\mathrm{d}x. \end{aligned} \tag{3.6}$$

Applying (3.4), (3.5), and (3.6), in (3.3), results in the upper bound

$$\begin{aligned} J(A, u_A) - J(B, u_B) &+ \frac{\mu}{2}\|u_B - u_A\|^2_{L^2(A\cap B)} + \frac{1}{2}\|y_B - y_A\|_H^2 \\ &\le \int_{A\setminus B} \beta + p_A u_A + g(u_A)\,\mathrm{d}x + \int_{B\setminus A} -\beta - p_A u_B - g(u_B)\,\mathrm{d}x. \end{aligned}$$

Using $\bar{H}$, this can be written as

$$\begin{aligned} J(A, u_A) - J(B, u_B) &+ \frac{\alpha}{2}\|u_B - u_A\|^2_{L^2(A\cap B)} + \frac{1}{2}\|y_B - y_A\|_H^2 \\ &\le \int_{A\setminus B} \beta + \bar{H}(p_A)\,\mathrm{d}x + \int_{B\setminus A} -\beta - \bar{H}(p_B) - (p_A - p_B)u_B\,\mathrm{d}x, \end{aligned}$$

which is the claim. □

The next result is the main result of this section. It gives an expansion of the value function $J(A)$ together with an remainder term that is of higher order in $\|\chi_A - \chi_B\|_{L^1(\Omega)}$.

Theorem 3.7. *Assume* ***(A1)***, ***(A2)***, ***(A3)***, ***(A4)***, ***(A5)***, ***(A6)***. *Let* $A, B \subseteq \Omega$, *and let* (u_A, y_A, p_A) *and* (u_B, y_B, p_B) *be solutions of* $(\mathbf{P}_A)$ *and* $(\mathbf{P}_B)$. *Then it holds*

$$\left| J(A, u_A) - J(B, u_B) - \int_\Omega (\chi_A - \chi_B)(\beta + \bar{H}(p_B)) \,\mathrm{d}x \right| + \frac{\mu}{2} \|u_B - u_A\|^2_{L^2(A\cap B)} + \frac{1}{2}\|y_B - y_A\|^2_H \leq 2PK \|\chi_A - \chi_B\|^{\frac{3}{2}-\frac{3}{q}}_{L^1(\Omega)},$$

where P, K, q *are from Corollary 2.4, Theorem 3.4, and* ***(A5)***, *respectively.*

Proof. Using the result of Lemma 3.6, we get

$$\begin{aligned} J(A, u_A) - J(B, u_B) &+ \frac{\mu}{2} \|u_B - u_A\|^2_{L^2(A\cap B)} + \frac{1}{2}\|y_B - y_A\|^2_H \\ &\leq \int_{A\setminus B} \beta + \bar{H}(p_A) \,\mathrm{d}x + \int_{B\setminus A} -\beta - \bar{H}(p_B) - (p_A - p_B)u_B \,\mathrm{d}x \\ &= \int_\Omega (\chi_A - \chi_B)(\beta + \bar{H}(p_B)) \,\mathrm{d}x \\ &\quad + \int_{A\setminus B} \bar{H}(p_A) - \bar{H}(p_B) \,\mathrm{d}x + \int_{B\setminus A} (p_A - p_B) u_B \,\mathrm{d}x. \end{aligned}$$

The latter two integrals can be estimated using Lemma 3.5, Corollary 2.4, the property $|A \setminus B| + |B \setminus A| = \|\chi_A - \chi_B\|_{L^1(\Omega)}$, and Theorem 3.4 as follows

$$\begin{aligned} \int_{A\setminus B} \bar{H}(p_A) - \bar{H}(p_B) \,\mathrm{d}x + \int_{B\setminus A} (p_A - p_B) u_B \,\mathrm{d}x & \\ \leq P \|\chi_A - \chi_B\|^{1-\frac{2}{q}}_{L^1(\Omega)} \|p_A - p_B\|_{L^q(\Omega)} & \\ \leq PK \|\chi_A - \chi_B\|^{\frac{3}{2}-\frac{3}{q}}_{L^1(\Omega)}. & \end{aligned} \tag{3.7}$$

This results in the upper bound

$$\begin{aligned} J(A, u_A) - J(B, u_B) &+ \frac{\mu}{2} \|u_B - u_A\|^2_{L^2(A\cap B)} + \frac{1}{2}\|y_B - y_A\|^2_H \\ &\leq \int_\Omega (\chi_A - \chi_B)(\beta + \bar{H}(p_B)) \,\mathrm{d}x + PK \|\chi_A - \chi_B\|^{\frac{3}{2}-\frac{3}{q}}_{L^1(\Omega)}. \end{aligned} \tag{3.8}$$

To obtain a lower bound, we use the result of Lemma 3.6 but with the roles of A and B reversed (and multiplying the resulting inequality by -1), which yields

$$\begin{aligned} J(A, u_A) - J(B, u_B) &- \frac{\mu}{2} \|u_B - u_A\|^2_{L^2(A\cap B)} - \frac{1}{2}\|y_B - y_A\|^2_H \\ &\geq \int_{A\setminus B} \beta + \bar{H}(p_A) - (p_A - p_B)u_A \,\mathrm{d}x + \int_{B\setminus A} -\beta - \bar{H}(p_B) \,\mathrm{d}x. \end{aligned}$$

With the help of Lemma 3.5, Corollary 2.4, and Theorem 3.4, we can estimate

$$\int_{A\setminus B} \bar{H}(p_A) - \bar{H}(p_B) - (p_A - p_B)u_A \,\mathrm{d}x \leq 2PK \|\chi_A - \chi_B\|^{\frac{3}{2}-\frac{3}{q}}_{L^1(\Omega)}, \tag{3.9}$$

which gives the lower bound

$$J(A,u_A) - J(B,u_B) - \frac{\mu}{2}\|u_B - u_A\|^2_{L^2(A\cap B)} - \frac{1}{2}\|y_B - y_A\|^2_H \\ \geq \int_\Omega (\chi_A - \chi_B)(\beta + \bar{H}(p_B))\,\mathrm{d}x - 2PK\|\chi_A - \chi_B\|^{\frac{3}{2}-\frac{3}{q}}_{L^1(\Omega)}.$$

Both inequalities together prove the claim. □

As a by-product of the previous proof, we get the improved stability estimate

$$\|u_B - u_A\|^2_{L^2(A\cap B)} + \|y_B - y_A\|^2_H \leq K'\|\chi_A - \chi_B\|^{\frac{3}{2}(\frac{1}{2}-\frac{1}{q})}_{L^1(\Omega)},$$

which improves the exponent from Theorem 3.4 by a factor $\frac{3}{2}$.

Remark 3.8. *If $S^* \in \mathcal{L}(H, L^q(\Omega))$ then the estimate can improved to*

$$\|u_B - u_A\|^2_{L^2(A\cap B)} + \|y_B - y_A\|^2_H \leq K'\|\chi_A - \chi_B\|^{2(\frac{1}{2}-\frac{1}{q})}_{L^1(\Omega)}$$

by estimating $\|p_A - p_B\|_{L^q(\Omega)}$ against $\|y_A - y_B\|_H$ in the estimates (3.7) and (3.9).

4 Topological derivatives

Definition 4.1. *Let $B \subseteq \Omega$. Then the topological derivative of J at B at the point x is defined by*

$$DJ(B)(x) = \begin{cases} \lim_{r\searrow 0} \dfrac{J(B\cup B_r(x)) - J(B)}{|B_r|} & \text{if } x \notin B \\ \\ \lim_{r\searrow 0} \dfrac{J(B\setminus B_r(x)) - J(B)}{|B_r|} & \text{if } x \in B. \end{cases}$$

The existence of the topological derivative is now a consequence of the expansion in Theorem 3.7 and the Lebesgue differentiation theorem.

Theorem 4.2. *Assume **(A1)**, **(A2)**, **(A3)**, **(A4)**, **(A5)**, **(A6)**. Let $B \subseteq \Omega$, and let (u_B, y_B, p_B) be a solution of $(\mathbf{P}_B)$.*

Then for almost all $x \in \Omega$ the topological derivative $DJ(B)(x)$ exists, and is given by

$$DJ(B)(x) = \sigma(B,x)(\beta(x) + \bar{H}(p_B(x)))$$

with

$$\sigma(B,x) := \begin{cases} +1 & \text{if } x \notin B \\ -1 & \text{if } x \in B. \end{cases}$$

Proof. Let $x_0 \in B$. Let $r > 0$. Define $A(x_0, r) := B \setminus B_r(x_0)$. Then it follows $\chi_{A(x_0,r)} - \chi_B = -\chi_{B\cap B_r(x_0)}$, which implies $\|\chi_{A(x_0,r)} - \chi_B\|_{L^1(\Omega)} \leq |B_r|$. Using this in the result of Theorem 3.7, we find

$$\left| J(A(x_0,r)) - J(B) + \int_{B\cap B_r(x_0)} \beta + \bar{H}(p_B)\,\mathrm{d}x \right| \leq 2PK|B_r|^{\frac{3}{2}-\frac{3}{q}}. \tag{4.1}$$

Let us now define

$$v(x_0, r) := \frac{1}{|B_r|} \int_{B_r(x_0)} \chi_B \cdot (\beta + \bar{H}(p_B))\, \mathrm{d}x.$$

By the Lebesgue differentiation theorem, we have

$$\lim_{r \searrow 0} v(x, r) = \chi_B(x) \cdot (\beta(x) + \bar{H}(p_B(x)))$$

for almost all $x \in \Omega$. This implies together with (4.1)

$$\lim_{r \searrow 0} \frac{J(A(x, r)) - J(B)}{|B_r|} = -(\beta(x) + \bar{H}(p_B(x)))$$

for almost all $x \in B$. Here we used that $\frac{3}{2} - \frac{3}{q} > 1$ by **(A5)**. This proves the claim for $x \in B$.

The claim for $x \notin B$ can be proven completely analogously: this time we set $A(x_0, r) := B \cup B_r(x_0)$ for $x_0 \notin B$, which implies $\chi_{A(x_0,r)} - \chi_B = \chi_{B_r(x_0) \setminus B}$, resulting in the different sign of the topological derivative. □

Note that in contrast to other works, we do not need to impose continuity of u_B near x_0 as in [10, Corollary 4.1], nor do we need to argue by Hölder continuity of the adjoint as in [1, Corollary 3.2].

We can now formulate a necessary optimality condition for (1.3) using the topological derivative.

Theorem 4.3. *Assume **(A1)**, **(A2)**, **(A3)**, **(A4)**, **(A5)**, **(A6)**. Let B be a solution of (1.3). Then*

$$\beta + \bar{H}(p_B) \leq 0 \qquad \text{a.e. on } B$$

and

$$\beta + \bar{H}(p_B) \geq 0 \qquad \text{a.e. on } \Omega \setminus B.$$

Proof. The result follows immediately from Theorem 4.2. □

Remark 4.4. *Using the celebrated Ekeland's variational principle [6], a following result can be proven for ϵ-solutions: There is an ϵ-solution, such that optimality conditions are satisfied up to ϵ. We briefly sketch the proof.*

Let V be the metric space of characteristic functions χ_B, $B \subseteq \Omega$ measurable, supplied with the $L^1(\Omega)$-metric, which makes it a complete space. Applying [6, Theorem 1.1] with $\epsilon > 0$ and $\lambda = 1$ there is $B_\epsilon \subseteq \Omega$ such that

$$J(B_\epsilon) \leq \inf_{B \subseteq \Omega} J(B) + \epsilon \tag{4.2}$$

and

$$J(A) \geq J(B_\epsilon) - \epsilon \|\chi_A - \chi_{B_\epsilon}\|_{L^1(\Omega)} \tag{4.3}$$

for all $A \subseteq \Omega$. Owing to (4.2) the set B_ϵ is then an ϵ-solution of (1.3). Due to inequality (4.3), we can consider variations of $J(B_\epsilon)$ to obtain estimates of the topological derivative:

For $x_0 \in \Omega$ and $r > 0$, define $A(x_0, r)$ as in the proof of Theorem 4.2. Then $\frac{1}{|B_r|}(J(A(x_0, r)) - J(B_\epsilon)) \geq -\epsilon$ by (4.3), which results in $DJ(B_\epsilon)(x_0) \geq -\epsilon$ for *almost all x_0. Using the expression of the topological derivative of Theorem 4.2 implies*

$$\beta + \bar{H}(p_{B_\epsilon}) \leq \epsilon \qquad a.e. \text{ on } B_\epsilon$$

and

$$\beta + \bar{H}(p_{B_\epsilon}) \geq -\epsilon \qquad a.e. \text{ on } \Omega \setminus B_\epsilon.$$

In addition, the defect in the optimality condition of Theorem 4.3 can be used to get an error estimate as follows.

Lemma 4.5. *Assume **(A1)**, **(A2)**, **(A3)**, **(A6)**. Let $A \subseteq \Omega$, and let (u_A, y_A, p_A) be a solution of $(\mathbf{P}_A)$. Let the defect δ_A be defined by*

$$\delta_A := \int_A \left(\beta + \bar{H}(p_A)\right)^+ \,\mathrm{d}x - \int_{\Omega\setminus A} \left(\beta + \bar{H}(p_A)\right)^- \,\mathrm{d}x.$$

Then we have

$$J(A) - \inf_{B \subseteq B} J(B) \leq \delta_A.$$

If B is a solution of (1.3) then we have the error estimate

$$J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|_H^2 + \frac{\mu}{2}\|u_B - u_A\|_{L^2(A\cap B)}^2 \leq \delta_A.$$

Proof. Let $B \subseteq \Omega$ and (u_B, y_B, p_B) be a solution of $(\mathbf{P}_B)$. By Lemma 3.1, we have

$$\begin{aligned} &J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|_H^2 \\ &\qquad = \int_\Omega g(u_A) - g(u_B) + \chi_A p_A (u_A - u_B) + (\chi_A - \chi_B)(\beta + p_A u_B) \,\mathrm{d}x. \end{aligned}$$

Using (3.4), we obtain

$$\begin{aligned} J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|_H^2 \leq -\frac{\mu}{2}\|u_B - u_A\|_{L^2(A\cap B)}^2 \\ + \int_{A\setminus B} g(u_A) + p_A u_A + \beta \,\mathrm{d}x - \int_{B\setminus A} g(u_B) + \beta + p_A u_B \,\mathrm{d}x. \end{aligned}$$

Employing the definition of $\bar{H}$ yields

$$\begin{aligned} J(A, u_A) - J(B, u_B) &+ \frac{1}{2}\|y_B - y_A\|_H^2 + \frac{\mu}{2}\|u_B - u_A\|_{L^2(A\cap B)}^2 \\ &\leq \int_{A\setminus B} \beta + \bar{H}(p_A) \,\mathrm{d}x - \int_{B\setminus A} \beta + H(u_B, p_A) \,\mathrm{d}x \\ &\leq \int_{A\setminus B} \beta + \bar{H}(p_A) \,\mathrm{d}x - \int_{B\setminus A} \beta + \bar{H}(p_A) \,\mathrm{d}x \\ &\leq \int_A \left(\beta + \bar{H}(p_A)\right)^+ \,\mathrm{d}x - \int_{\Omega\setminus A} \left(\beta + \bar{H}(p_A)\right)^- \,\mathrm{d}x. \end{aligned}$$

If B is a solution of (1.3) then the claim follows. Otherwise, we take the supremum of $-J(B, u_B)$ on the left-hand side. □

5 The non-strongly convex case

Let us briefly comment on the non-strongly convex case. That is, we no longer assume the strong convexity of g as in **(A4)**. We will replace **(A4)** and **(A5)** by the following two assumptions.

(A4') $\operatorname{dom} g$ is a bounded subset of $\mathbb{R}$,

(A5') There is $q > 3$ such that $S^* \in \mathcal{L}(H, L^q(\Omega))$, where $S^* \in \mathcal{L}(H, L^2(\Omega))$ denotes the Hilbert space-adjoint of S.

(A4') implies the solvability of $(\mathbf{P}_A)$. In addition, solutions u_A of $(\mathbf{P}_A)$ will be $L^\infty(\Omega)$. Due to the missing strong convexity, we have to replace the assumption on S^*S in **(A5)** by an assumption on S^*. The $L^\infty(\Omega)$-regularity of optimal controls will allows us to work with a smaller exponent q in **(A5')** when compared to **(A5)**.

Note that we do not add assumptions that imply unique solvability of $(\mathbf{P}_A)$.

Proposition 5.1. *Let $A \subseteq \Omega$ measurable be given. Then there is a minimizer u_A of $(\mathbf{P}_A)$. Moreover, $\chi_A u_A$ is also a minimizer of $(\mathbf{P}_A)$.*

Proof. Due to **(A4')** minimizing sequences of $u \mapsto J(A, u)$ are bounded in $L^\infty(\Omega)$. Then the proof of existence follows as in Proposition 2.1. The last claim is a consequence of (2.1). □

In the sequel, we will assume that a solution u_A of $(\mathbf{P}_A)$ satisfies $\chi_A u_A = u_A$. Due to the previous result, this is not restriction at all, as for every minimizer u_A also $\chi_A u_A$ is a minimizer. Let us start with a replacement of Lemma 2.3 and Corollary 2.4.

Lemma 5.2. *There is $M > 0$ and $P' > 0$ such that*

$$\|y_A - y_d\|_H \leq M$$

and

$$\|u_A\|_{L^\infty(\Omega)} \leq P', \quad \|p_A\|_{L^q(\Omega)} \leq P'$$

*for all $A \subseteq \Omega$ and all solutions (u_A, y_A, p_A) of $(\mathbf{P}_A)$. Here, q is as in **(A5')**.*

Proof. The bound of y_A can be obtained as in Lemma 2.3, the bounds of u_A and p_A are consequences of **(A4')** and **(A5')**. □

Due to the missing strong convexity of g, we cannot expect stability of controls as in Theorem 3.4. Here, we have the following replacement.

Theorem 5.3. *Assume **(A1)**, **(A2)**, **(A3)**, **(A4')**, **(A5')**. Then there is a constant $K' > 0$ such that for all $A, B \subseteq \Omega$*

$$\|p_A - p_B\|_{L^q(\Omega)} + \|y_B - y_A\|_{L^2(\Omega)} \leq K' \|\chi_A - \chi_B\|_{L^1(\Omega)}^{\frac{1}{2} - \frac{1}{2q}},$$

*where (u_A, y_A, p_A) and (u_B, y_B, p_B) are solutions of $(\mathbf{P}_A)$ and $(\mathbf{P}_B)$, and q is from **(A5')**.*

Proof. From **(A5')**, we get $\|p_A - p_B\|_{L^q(\Omega)} \le \|S^*\|_{\mathcal{L}(H,L^q(\Omega))}\|y_A - y_B\|_{L^2(\Omega)}$. Define $\mu' := 1/\|S^*\|^2_{\mathcal{L}(H,L^q(\Omega))}$. Let q' be such that $\frac{1}{q'} + \frac{1}{q} = 1$. From the inequality of Lemma 3.3, we obtain with Hölder's inequality

$$\begin{aligned}\frac{\mu'}{2}\|p_A - p_B\|^2_{L^q(\Omega)} + \frac{1}{2}\|y_B - y_A\|^2_H &\le \int_\Omega (\chi_A - \chi_B)(p_A u_B - p_B u_A)\,\mathrm{d}x \\ &\le 2(P')^2\|\chi_A - \chi_B\|^{1-\frac{1}{q}}_{L^1(\Omega)},\end{aligned}$$

where P' is from Lemma 5.2, and the claim is proven. □

This stability result has to replace Theorem 3.4 in the proof of Theorem 3.7. The result corresponding to the latter theorem now reads as follows.

Theorem 5.4. *Assume* ***(A1)***, ***(A2)***, ***(A3)***, ***(A4')***, ***(A5')***, ***(A6)***. *Let $A, B \subseteq \Omega$, and let (u_A, y_A, p_A) and (u_B, y_B, p_B) be solutions of $(\mathbf{P}_A)$ and $(\mathbf{P}_B)$. Then it holds*

$$\begin{aligned}\left| J(A,u_A) - J(B,u_B) - \int_\Omega (\chi_A - \chi_B)(\beta + \bar{H}(p_B))\,\mathrm{d}x \right| + \frac{1}{2}\|y_B - y_A\|^2_H \\ \le 2P'K'\|\chi_A - \chi_B\|^{\frac{3}{2}(1-\frac{1}{q})}_{L^1(\Omega)},\end{aligned}$$

where P', K', q are from Lemma 5.2 and ***(A5')***, *respectively.*

Proof. We can proceed exactly as in the proof of Theorem 3.7 with $\mu = 0$. Only the estimates (3.7) and (3.9) have to be modified. The estimate of Lemma 3.5 has to be changed to

$$\|\bar{H}(p_A) - \bar{H}(p_B)\|_{L^q(\Omega)} \le P'\|p_A - p_B\|_{L^q(\Omega)} \tag{5.1}$$

using the $L^\infty(\Omega)$-bound of Lemma 5.2, as well as the estimate of $\bar{H}$ from the proof of Lemma 3.5. Note that due to **(A4')**, $\bar{H}(p)$ is well-defined and finite for all $p \in \mathbb{R}$.

Then the error term of (3.7) can be estimated using Lemma 3.5 and Theorem 5.3 as

$$\begin{aligned}\int_{A\setminus B} \bar{H}(p_A) - \bar{H}(p_B)\,\mathrm{d}x + \int_{B\setminus A} (p_A - p_B)u_B\,\mathrm{d}x \\ \le P'\|\chi_A - \chi_B\|^{1-\frac{1}{q}}_{L^1(\Omega)}\|p_A - p_B\|_{L^q(\Omega)} \\ \le P'K'\|\chi_A - \chi_B\|^{\frac{3}{2}(1-\frac{1}{q})}_{L^1(\Omega)}.\end{aligned} \tag{5.2}$$

The error contribution from (3.9) can be estimated similarly as

$$\int_{A\setminus B} \bar{H}(p_A) - \bar{H}(p_B) - (p_A - p_B)u_A\,\mathrm{d}x \le 2P'K'\|\chi_A - \chi_B\|^{\frac{3}{2}(1-\frac{1}{q})}_{L^1(\Omega)}.$$

The claimed estimate can now be obtained with the same arguments as in the proof of Theorem 5.4. □

Theorem 5.5. *Assume* ***(A1)***, ***(A2)***, ***(A3)***, ***(A4')***, ***(A5')***, ***(A6)***. *Let $B \subseteq \Omega$.*

Then for almost all $x \in \Omega$ the topological derivative $DJ(B)(x)$ exists, and it is given by the expression in Theorem 4.2.

6 Optimization method based on the topological derivative

In this section, we introduce a new optimization algorithm that is motivated by the work on the topological derivative. Here, we work under the set of assumptions of Theorem 4.2 or Theorem 5.5. We assume that we can choose $q = +\infty$ in **(A5)** or **(A5')**.

Let $A_k \subseteq \Omega$ be a given iterate together with solutions of $(\mathbf{P}_{A_k})$ denoted by (y_k, u_k, p_k). Let us define the residual in the optimality condition of Theorem 4.3 as

$$\rho_k := \chi_{A_k}(\beta + \bar{H}(p_k))^+ + \chi_{\Omega\setminus A_k}(\beta + \bar{H}(p_k))^-. \tag{6.1}$$

Note that

$$\delta_{A_k} = \|\rho_k\|_{L^1(\Omega)},$$

with δ as in Lemma 4.5.

New iterates A_{k+1} will now be defined by adding/removing points to/from A_k, where the absolute value of ρ_k is large. We will achieve this in the following way. Given $t \in (0, 1)$, define

$$\begin{aligned}\tilde{A}_{k,t} := (A_k \setminus \{x \in \Omega :\ \rho_k(x) \geq (1-t)\|\rho_k\|_{L^\infty(\Omega)}\})\\ \cup \{x \in \Omega :\ \rho_k(x) \leq -(1-t)\|\rho_k\|_{L^\infty(\Omega)}\}\end{aligned} \tag{6.2}$$

Using the expression of the topological derivative $DJ(A_k)$ from Theorem 4.2, we have

$$\rho_k = (DJ(A_k))^-$$

and

$$\begin{aligned}\tilde{A}_{k,t} = (A_k \setminus \{x \in \Omega :\ DJ(A_k)(x) \leq -(1-t)\|(DJ(A_k))^-\|_{L^\infty(\Omega)}\})\\ \cup \{x \in A_k \setminus \Omega :\ \ DJ(A_k)(x) \leq -(1-t)\|(DJ(A_k))^-\|_{L^\infty(\Omega)}\}.\end{aligned}$$

Hence, a point x belongs to the symmetric difference of A_k and $\tilde{A}_{k,t}$ if the topological derivative $DJ(A_k)(x)$ is negative and has relatively large absolute value. In difference to other methods, even small step-sizes t may result in a change of the topology. That is, domain variations are not reduced to boundary variations for small t.

Now, we determine the step-size $t \in (0, 1)$ and $A_{k,t} \subseteq \tilde{A}_{k,t}$ such that the conditions

$$\|\chi_{A_{k,t}} - \chi_{A_k}\|_{L^1(\Omega)} \leq c_1 t \tag{6.3}$$

and

$$J(A_{k,t}) \leq J(A_k) + \sigma \int_\Omega (\chi_{A_{k,t}} - \chi_{A_k})\rho_k \,\mathrm{d}x \tag{6.4}$$

are satisfied, where $c_1 > 0$ and $\sigma \in (0, 1)$ is given. The first condition is to ensure that remainder terms in the expansions of Theorem 3.7 or Theorem 5.4 are of higher order. The second condition is inspired by the Armijo descent condition from nonlinear optimization. It can be replaced by a non-monotone linesearch with obvious modifications.

We start with the following observation, which shows that (6.4) ensures descent of J.

Lemma 6.1. *Let $A_{k,t}$ be defined as above. Then it holds for all $t \in (0,1)$*

$$\int_\Omega (\chi_{A_{k,t}} - \chi_{A_k})\rho_k \,\mathrm{d}x \le -(1-t)\|\rho_k\|_{L^\infty(\Omega)}\|\chi_{A_{k,t}} - \chi_{A_k}\|_{L^1(\Omega)}.$$

Proof. The claim follows directly from the definition of $A_{k,t}$ and (6.2)

$$\begin{aligned}\int_\Omega (\chi_{A_{k,t}} - \chi_{A_k})\rho_k \,\mathrm{d}x &= \int_{A_{k,t}\setminus A_k} \rho_k \,\mathrm{d}x - \int_{A_k\setminus A_{k,t}} \rho_k \,\mathrm{d}x \\ &\le -(1-t)\|\rho_k\|_{L^\infty(\Omega)}\|\chi_{A_{k,t}} - \chi_{A_k}\|_{L^1(\Omega)}.\end{aligned}$$

□

In addition, we have the following estimate of $J(A_{k,t}) - J(A_k)$.

Lemma 6.2. *There is $R > 0$ and $\tau > 0$ such that*

$$J(A_{k,t}) - J(A_k) \le \int_\Omega (\chi_{A_{k,t}} - \chi_{A_k})\rho_k \,\mathrm{d}x + R\|\chi_{A_{k,t}} - \chi_{A_k}\|^{1+\tau}_{L^1(\Omega)}.$$

Proof. This is a direct consequence of Theorems 3.7 and 5.4, and (6.2). □

Similar results are proven for the level set method in [1, Section 5]. There it was shown that the decrease of the cost functional for small step-sizes is determined by the values of the topological derivative at the boundary of the current set. This is in contrast to our result: the decrease of the functional is determined by large values of the topological derivative, and domain variations may happen far away from the boundary of A_k even for small step-sizes.

Now, we are in the position to prove that a step-size satisfying (6.3) and (6.4) exists. Again, we use an Armijo-type approach.

Lemma 6.3. *Define*

$$t_k := \max\{\tau^l : \ l = 0, 1, 2, \dots \ \textit{such that } (\tau^l, A_{k,\tau^l}) \textit{ satisfies } (6.3) \textit{ and } (6.4)\}.$$

Then t_k is well-defined.

Proof. Given $t > 0$ we can choose $A_{k,t} \subseteq \tilde{A}_{k,t}$ satisfying (6.3). For such $A_{k,t}$ we have

$$\begin{aligned}&J(A_{k,t}) - J(A_k) - \sigma \int_\Omega (\chi_{A_{k,t}} - \chi_{A_k})\rho_k \,\mathrm{d}x \\ &\le (1-\sigma)\int_\Omega (\chi_{A_{k,t}} - \chi_{A_k})\rho_k \,\mathrm{d}x + R\|\chi_{A_{k,t}} - \chi_{A_k}\|^{1+r}_{L^1(\Omega)} \\ &\le \Big(R\|\chi_{A_{k,t}} - \chi_{A_k}\|^r_{L^1(\Omega)} - (1-\sigma)(1-t)\Big)\|\rho_k\|_{L^\infty(\Omega)}\|\chi_{A_{k,t}} - \chi_{A_k}\|_{L^1(\Omega)}.\end{aligned}$$

Due to (6.3), the right-hand side will be negative for sufficiently small t. □

These ideas lead to the following algorithm.

Algorithm 6.4. 1. *Choose $\tau \in (0,1)$, $\sigma \in (0,1)$, $A_0 \subseteq \Omega$. Set $k := 0$.*

2. *Compute a solution (u_k, y_k, p_k) of $(\mathbf{P}_{A_k})$.*

3. Compute ρ_k as in (6.1).

4. Determine (t_k, A_{k+1}) satisfying (6.3) and (6.4) using the line-search strategy of Lemma 6.3.

5. Set $k := k + 1$, go to step 2.

Due to Lemma 6.3, the algorithm is well-defined. The iteration can be terminated if $\|\rho_k\|_{L^1(\Omega)}$ is small enough. This is motivated by Lemma 4.5: if $\|\rho_k\|_{L^1(\Omega)}$ is less than some tolerance $\epsilon > 0$ then A_k is an ϵ-solution of (1.3). If not terminated, the algorithm will produce an infinite sequence of sets (A_k), such that $(J(A_k))$ is monotonically decreasing. We have the following basic convergence result.

Lemma 6.5. *Let (A_k) be the iterates of Algorithm 6.4 with step sizes (t_k) and ρ_k from (6.1). Then it holds*

$$\sum_{k=0}^{\infty} \int_\Omega |(\chi_{A_{k+1}} - \chi_{A_k})\rho_k| \, \mathrm{d}x < +\infty$$

and

$$\sum_{k=0}^{\infty} (1 - t_k)\|\rho_k\|_{L^\infty(\Omega)} \|\chi_{A_{k+1}} - \chi_{A_k}\|_{L^1(\Omega)} < +\infty.$$

Proof. Due to the descent condition (6.4) and the result of Lemma 6.1, we have the chain of inequalities

$$\begin{aligned} J(A_{k+1}) - J(A_k) &\le \sigma \int_\Omega (\chi_{A_{k+1}} - \chi_{A_k})\rho_k \, \mathrm{d}x \\ &\le -\sigma(1 - t_k)\|\rho_k\|_{L^\infty(\Omega)} \|\chi_{A_{k+1}} - \chi_{A_k}\|_{L^1(\Omega)} \le 0. \end{aligned}$$

Since J is bounded from below, we can sum the above inequalities for $k = 0, \ldots,$ which proves the claim. □

Let us prove an estimate of ρ_k in terms of the expressions in the claim of the previous lemma.

Lemma 6.6. *Let (A_k) be the iterates of Algorithm 6.4 with step sizes (t_k) and ρ_k from (6.1). Then it holds*

$$\int_\Omega |\rho_k| \le \int_\Omega |(\chi_{A_{k+1}} - \chi_{A_k})\rho_k| \, \mathrm{d}x + (1 - t_k)\|\rho_k\|_{L^\infty(\Omega)} \|1 - |\chi_{A_{k+1}} - \chi_{A_k}|\|_{L^1(\Omega)}.$$

Proof. By construction, we have $|\rho_k| \le (1 - t_k)\|\rho_k\|_{L^\infty(\Omega)}$ on $A_k \cap A_{k+1}$ and on $(\Omega \setminus A_k) \cap (\Omega \setminus A_{k+1})$. This proves the claim. □

The first term in the estimate of the previous lemma converges to zero by Lemma 6.5. Consequently, the term $(1 - t_k)\|\rho_k\|_{L^\infty(\Omega)}$ is the crucial quantity when studying the convergence.

Theorem 6.7. *Let (A_k) be the iterates of Algorithm 6.4 with step sizes (t_k) and ρ_k from (6.1). Then it holds:*

1. *If* $\liminf_{k\to\infty}(1-t_k)\|\rho_k\|_{L^\infty(\Omega)} > 0$ *then* (χ_{A_k}) *converges in* $L^1(\Omega)$ *to* χ_A *for some* $A \subseteq \Omega$*. In addition,* $\|\rho_k\|_{L^1(\Omega)} \to \delta_A > 0$ *for* $k \to \infty$*, where* δ_A *is as in Lemma 4.5. The set* A *is a* δ_A*-solution of* (1.3)*.*

2. *If* $\liminf_{k\to\infty}(1-t_k)\|\rho_k\|_{L^\infty(\Omega)} = 0$ *then* $\liminf_{k\to\infty}\|\rho_k\|_{L^1(\Omega)} = 0$*, and* (A_k) *is a minimizing sequence of* (1.3)*.*

Proof. Suppose $\liminf_{k\to\infty}(1-t_k)\|\rho_k\|_{L^\infty(\Omega)} > 0$. Then due to Lemma 6.5 we get $\sum_{k=0}^\infty \|\chi_{A_{k+1}} - \chi_{A_k}\|_{L^1(\Omega)} < +\infty$. This implies that (χ_{A_k}) is a Cauchy sequence in $L^1(\Omega)$, and consequently converges in $L^1(\Omega)$ to a characteristic function χ_A.

Let (u_A, y_A, p_A) be a solution of ($\mathbf{P}_A$). By Theorems 3.4 and 5.3, the adjoint states converge in $L^q(\Omega)$, i.e., $p_k \to p_A$ in $L^q(\Omega)$ with q from **(A5)** or **(A5')**. Due to Lemma 3.5 or (5.1), we get the convergence of $\bar H(p_k) \to \bar H(p_A)$ in $L^1(\Omega)$, which implies the convergence $\|\rho_k\|_{L^1(\Omega)} \to \delta_A$. The assumptions imply $\liminf_{k\to\infty}\|\rho_k\|_{L^\infty(\Omega)} > 0$, hence $\delta_A > 0$. By Lemma 4.5, A is a δ_A-solution of (1.3).

Now, let us assume $\liminf_{k\to\infty}(1-t_k)\|\rho_k\|_{L^\infty(\Omega)} = 0$. Then Lemma 6.6 implies $\liminf_{k\to\infty}\|\rho_k\|_{L^1(\Omega)} = 0$. In addition, Lemma 4.5 shows

$$\liminf_{k\to\infty}(J(A_k) - \inf_{B\subseteq B} J(B)) \le \liminf_{k\to\infty} \delta_{A_k} = \liminf_{k\to\infty}\|\rho_k\|_{L^1(\Omega)} = 0.$$

Consequently, (A_k) contains a minimizing sequence of (1.3). Since $(J(A_k))$ is monotonically decreasing, (A_k) is a minimizing sequence. □

Let us note that the convergence theorem is valid under the assumptions of Lemma 4.5, as long as the algorithm produces a sequence of iterates. The full set of assumptions was only needed to guarantee that the line-search is well-defined, cf., see Lemma 6.3, which uses Lemma 6.2.

Let us add the following remarkable consequence of the theorem: if the sequence of iterates (χ_k) does not converge strongly, the sequence (A_k) is a minimizing sequence.

Corollary 6.8. *Let* (A_k) *be the iterates of Algorithm 6.4 with step sizes* (t_k) *and* ρ_k *from* (6.1)*. Suppose* (χ_k) *does not converge strongly in* $L^1(\Omega)$*. Then* (A_k) *is a minimizing sequence for* (1.3)*.*

Proof. Due to the first case of Theorem 6.7, the assumption implies $\liminf_{k\to\infty}(1-t_k)\|\rho_k\|_{L^\infty(\Omega)} = 0$, and the claim follows by the second case of Theorem 6.7. □

7 Numerical experiments

7.1 Optimal control problem with L^0-control cost

Let us report about numerical results of the application of Algorithm 6.4 to the following problem: Minimize

$$\min \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \beta\|u\|_0$$

over all $(y, u) \in H^1_0(\Omega) \times L^2(\Omega)$ satisfying

$$-\Delta y = u \qquad \text{a.e. on } \Omega$$

and

$$u_a \le u \le u_b \quad \text{a.e. on } \Omega.$$

This corresponds to the abstract setting with the choices $S := (-\Delta)^{-1} : L^2(\Omega) \to H_0^1(\Omega) \hookrightarrow L^2(\Omega)$, $H := L^2(\Omega)$, $g(u) := \frac{\alpha}{2}u^2 + I_{[u_a,u_b]}(u)$, $\beta(x) := \beta$. Here, I_C denotes the indicator function of the convex set C, defined by $I_C(x) = 0$ for $x \in C$ and $I_C(x) = +\infty$ for $x \notin C$. The assumptions are all satisfied. In particular g is strongly convex with modulus $\mu := \alpha$. Assumption Item **(A5)** and Item **(A5')** are satisfied with $q = \infty$ due to Stampacchia's result [15].

We choose $\Omega = (0,1)$. We used a standard finite-element discretization on a shape-regular mesh on Ω. State and adjoint variables (i.e., y, p) were discretized using continuous piecewise linear functions, while the control variable was discretized using piecewise constant functions. Let us remark that for the finest discretization, the control functions have $2,000,000$ degrees of freedom. The subproblems ($\mathbf{P}_A$) were solved by a semismooth Newton implementation. The parameters in the line-search of Algorithm 6.4 were chosen to be $c_1 = 1$, $\tau = 0.5$, and $\sigma = 0.1$. The algorithm was stopped if one of the following condition was fulfilled: $\|\rho_k\|_{L^\infty(\Omega)} \le 10^{-12}$, the support of ρ_k contained ≤ 3 elements, or the line-search failed to find a valid step-size. Termination due to the latter condition can happen if the relevant quantities in (6.4), are very small so that errors in the inexact solve of the sub-problem ($\mathbf{P}_A$) are of the same order.

In addition, we used the following data

$$y_d(x_1, x_2) = 10x_1 \sin(5x_1)\cos(7x_2),\ \alpha = 0.01,\ \beta = 0.01,\ u_a = -4,\ u_b = +4,$$

which was also used in [9, 17]. The computed optimal control, which is obtained by the last iterate of Algorithm 6.4 on the finest mesh, can be seen in Figure 1. Due to the presence of the L^0-term in the objective, the control is zero on a relatively large part of Ω.

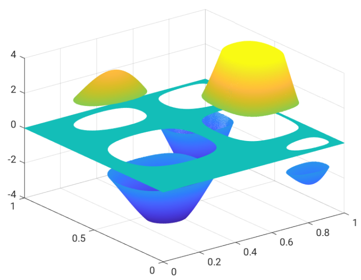

Figure 1: Solution for $h = 2.24 \cdot 10^{-3}$, Section 7.1

The results of the computations for different meshes can be seen in Table 1. There, h denotes the mesh-size of the triangulation, J denotes the value of the functional J at the final iterate, similarly $\|\chi\|_{L^1(\Omega)}$ is the size of the support of the optimal control, and $\|\rho\|_{L^1(\Omega)}$ is the error estimate from the topological derivative at the final iteration. The values corresponding to the mesh-size $h = 2.83 \cdot 10^{-3}$ are in agreement with those from [17]. For this example, all computations stopped due to the support of ρ_k containing less than 3 elements. In addition, for this example, the step-size $t = 1$ was always accepted. Algorithm Algorithm 6.4 was started with the initial choice $A_0 = \Omega$. As can be seen from Table 1, the optimal values of J and $\|\chi\|_{L^1(\Omega)}$ converge for $h \searrow 0$, and $\|\rho\|_{L^1(\Omega)} \to 0$ for $h \searrow 0$. According to Theorem 6.7, this strongly suggests that the iterates are a minimizing sequence of (1.3).

h	J	$\|\chi\|_{L^1(\Omega)}$	$\|\rho\|_{L^1(\Omega)}$
$4.42 \cdot 10^{-2}$	4.712	0.43896	$4.33 \cdot 10^{-3}$
$2.21 \cdot 10^{-2}$	5.054	0.44299	$2.12 \cdot 10^{-8}$
$1.13 \cdot 10^{-2}$	5.216	0.44352	$2.09 \cdot 10^{-8}$
$5.66 \cdot 10^{-3}$	5.299	0.44432	$2.04 \cdot 10^{-8}$
$2.83 \cdot 10^{-3}$	5.340	0.44455	$2.11 \cdot 10^{-11}$
$1.41 \cdot 10^{-3}$	5.360	0.44460	$4.05 \cdot 10^{-11}$

Table 1: Results of optimization, Section 7.1

Let us report about the influence of the choice of the initial guess $A_0 \subseteq \Omega$. Here we chose the following set of parameters: y_d was as above, and

$$\alpha = 0.001, \ \beta = 0.1, \ u_a = -40, \ u_b = +40.$$

For this example, the method returned the same solution independent of the initial guess. We depicted the iteration history for different choices of A_0 in Figure 2. In general, the method was faster when starting from $A_0 = \Omega$ than from $A_0 = \emptyset$. As one can see from Figure 2, the convergence of $\|\rho\|_{L^1(\Omega)}$ is stable with respect to mesh refinement.

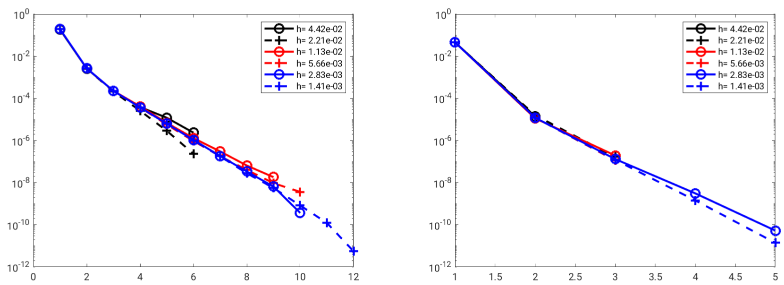

Figure 2: Comparison of iteration history of $\|\rho_k\|_{L^1(\Omega)}$ for different choice of A_0: $A_0 = \emptyset$ (left), $A_0 = \Omega$ (right), Section 7.1

7.2 Binary control problems

Following the idea of [1, 2], we will apply our algorithm to a binary control problem, where controls only can take values in $\{0, +1\}$. The problem considered in [1, 2] reads: Minimize

$$\min \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \nu\|u\|_{L^1(\Omega)}$$

over all $(y, u) \in H^1_0(\Omega) \times L^2(\Omega)$ satisfying

$$-\Delta y = u \quad \text{a.e. on } \Omega$$

and

$$u(x) \in \{0, 1\} \text{ f.a.a. } x \in \Omega.$$

Hence u itself is a characteristic function of type χ_A. And the above problem can be written in our setting as: Minimize

$$J(A, u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \nu \int_A \mathrm{d}x$$

over all $(y, u) \in H^1_0(\Omega) \times L^2(\Omega)$ satisfying

$$-\Delta y = \chi_A u \quad \text{a.e. on } \Omega$$

and the (trivial) constraint

$$u = 1 \text{ a.e. on } \Omega.$$

This setting fits into our framework with $\beta = \nu$, $g(u) := I_{\{1\}}(u)$. However, the assumption $g(0) = 0$ is not valid, and the crucial relation (2.3) does not hold. Still we can compute the topological derivative as follows. The solution of $u \mapsto J(A, u)$ is given by $u_A \equiv 1$, which greatly simplifies the computations of Section 3. And we have the following result concering the topological derivative of the value function.

Theorem 7.1. *The topological derivative $DJ(B)(x)$ of the value function of the binary control problem exists for almost all $x \in \Omega$, and is given by*

$$DJ(B)(x) = \sigma(B, x)(\beta(x) + p_B(x))$$

with $\sigma(B, x)$ as in Theorem 4.2.

Proof. The result of Lemma 3.1 in this situation has to be modified to

$$J(A, u_A) - J(B, u_B) + \frac{1}{2}\|y_B - y_A\|^2_H = \int_\Omega (\chi_A - \chi_B)(\beta + p_A)\,\mathrm{d}x,$$

where we have used $\chi_A u_A - \chi_B u_B = \chi_A - \chi_B$ in (3.1). Since $p_A - p_B = S^*S(\chi_A - \chi_B)$, we have the estimate $\|p_A - p_B\|_{L^\infty(\Omega)} \leq c\|\chi_A - \chi_B\|_{L^2(\Omega)} = c\|\chi_A - \chi_B\|^{1/2}_{L^1(\Omega)}$, which replaces the result of Theorem 3.4. Now the claim can be proven as in the proof of Theorem 4.2. $\square$

The topological derivative coincides with the result [1, Corollary 3.2]. The computation of the topological derivative does not involve the solution of any optimization problem: given A, only y_A and p_A have to be computed.

Let us report about the results for the following choice of parameters, corresponding to Case 3 in [1, Section 9]:

$$y_d = 0.05, \quad \nu = 0.002.$$

The computed control on the finest discretization can be seen in Figure 3, which agrees with [1, Figure 4]. The results of the optimization runs for different discretizations can be seen in Table 2. In all cases, the algorithm stopped due to a failed line-search. Nevertheless, the error quantity $\|\rho\|_{L^1(\Omega)}$ is very small, and is decreasing with decreasing mesh-size. According to Theorem 6.7 this indicates that the algorithm produces a minimizing sequence.

h	J	$\|\chi\|_{L^1(\Omega)}$	$\|\rho\|_{L^1(\Omega)}$
$6.99 \cdot 10^{-2}$	$1.799 \cdot 10^{-3}$	1.63770	$3.33 \cdot 10^{-7}$
$3.49 \cdot 10^{-2}$	$1.872 \cdot 10^{-3}$	1.63818	$2.21 \cdot 10^{-8}$
$1.79 \cdot 10^{-2}$	$1.909 \cdot 10^{-3}$	1.63802	$1.51 \cdot 10^{-9}$
$8.94 \cdot 10^{-3}$	$1.928 \cdot 10^{-3}$	1.63805	$1.28 \cdot 10^{-11}$
$4.47 \cdot 10^{-3}$	$1.938 \cdot 10^{-3}$	1.63802	$8.15 \cdot 10^{-12}$
$2.24 \cdot 10^{-3}$	$1.943 \cdot 10^{-3}$	1.63802	$2.38 \cdot 10^{-13}$

Table 2: Results of optimization, Section 7.2

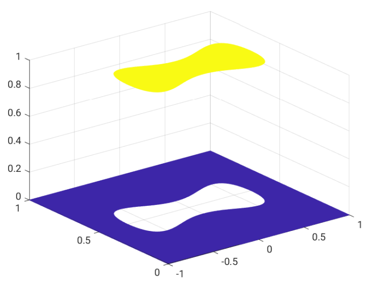

Figure 3: Solution for $h = 2.24 \cdot 10^{-3}$, Section 7.2

References

[1] S. Amstutz. "A semismooth Newton method for topology optimization". In: *Nonlinear Anal.* 73.6 (2010), pp. 1585–1595. DOI: 10.1016/j.na.2010.04.065.

[2] S. Amstutz. "Analysis of a level set method for topology optimization". In: *Optim. Methods Softw.* 26.4-5 (2011), pp. 555–573. DOI: 10.1080/10556788.2010.521557.

[3] S. Amstutz and H. Andrä. "A new algorithm for topology optimization using a level-set method". In: *J. Comput. Phys.* 216.2 (2006), pp. 573–588. DOI: 10.1016/j.jcp.2005.12.015.

[4] M. C. Delfour. "Control, shape, and topological derivatives via minimax differentiability of Lagrangians". In: *Numerical methods for optimal control problems*. Vol. 29. Springer INdAM Ser. Springer, Cham, 2018, pp. 137–164.

[5] M. C. Delfour and K. Sturm. "Parametric semidifferentiability of minimax of Lagrangians: averaged adjoint approach". In: *J. Convex Anal.* 24.4 (2017), pp. 1117–1142.

[6] I. Ekeland. "On the variational principle". In: *J. Math. Anal. Appl.* 47 (1974), pp. 324–353. DOI: 10.1016/0022-247X(74)90025-0.

[7] I. Ekeland and R. Témam. *Convex analysis and variational problems*. Vol. 28. Classics in Applied Mathematics. SIAM, Philadelphia, PA, 1999. DOI: 10.1137/1.9781611971088.

[8] P. Fulmański, A. Laurain, J.-F. Scheid, and J. Sokołowski. "A level set method in shape and topology optimization for variational inequalities". In: *Int. J. Appl. Math. Comput. Sci.* 17.3 (2007), pp. 413–430. DOI: 10.2478/v10006-007-0034-z.

[9] K. Ito and K. Kunisch. "Optimal control with $L^p(\Omega)$, $p \in [0,1)$, control cost". In: *SIAM J. Control Optim.* 52.2 (2014), pp. 1251–1275. DOI: 10.1137/120896529.

[10] D. Kalise, K. Kunisch, and K. Sturm. "Optimal actuator design based on shape calculus". In: *Math. Models Methods Appl. Sci.* 28.13 (2018), pp. 2667–2717. DOI: 10.1142/S0218202518500586.

[11] A. A. Novotny, J. Sokołowski, and A. Żochowski. "Topological derivatives of shape functionals. Part II: First-order method and applications". In: *J. Optim. Theory Appl.* 180.3 (2019), pp. 683–710. DOI: 10.1007/s10957-018-1419-x.

[12] R. T. Rockafellar. "Integral functionals, normal integrands and measurable selections". In: *Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975)*. Lecture Notes in Math., Vol. 543. Springer, Berlin, 1976, pp. 157–207.

[13] J. Sokołowski and A. Żochowski. "On the topological derivative in shape optimization". In: *SIAM J. Control Optim.* 37.4 (1999), pp. 1251–1272. DOI: 10.1137/S0363012997323230.

[14] J. Sokołowski and A. Żochowski. "Topological derivative for optimal control problems". In: *Control Cybernet.* 28.3 (1999). Recent advances in control of PDEs, pp. 611–625.

[15] G. Stampacchia. "Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus". In: *Ann. Inst. Fourier (Grenoble)* 15.fasc. 1 (1965), pp. 189–258.

[16] K. Sturm. "First-order differentiability properties of a class of equality constrained optimal value functions with applications". In: *J. Nonsmooth Anal. Optim.* 1 (2020), Paper No. 6034, 25. DOI: 10.15407/dopovidi2020.01.024.

[17] D. Wachsmuth. “Iterative hard-thresholding applied to optimal control problems with $L^0(\Omega)$ control cost”. In: *SIAM J. Control Optim.* 57.2 (2019), pp. 854–879. DOI: 10.1137/18M1194602.