SIAM J. OPTIM.
Vol. 30, No. 3, pp. 1922–1953

EFFICIENT TECHNIQUES FOR SHAPE OPTIMIZATION WITH VARIATIONAL INEQUALITIES USING ADJOINTS*

DANIEL LUFT†, VOLKER H. SCHULZ†, AND KATHRIN WELKER‡

Abstract. In general, standard necessary optimality conditions cannot be formulated in a straightforward manner for semismooth shape optimization problems. In this paper, we consider shape optimization problems constrained by variational inequalities of the first kind, so-called obstacle-type problems. Under appropriate assumptions, we prove existence of adjoints for regularized problems and convergence to adjoints of the unregularized problem. Moreover, we derive shape derivatives for the regularized problem and prove convergence to a limit object. Based on this analysis, an efficient optimization algorithm is devised and tested numerically.

Key words. semismooth optimization, variational inequality, obstacle problem, shape optimization, numerical methods, adjoint methods

AMS subject classifications. 65K15, 49Q10, 49M29, 35Q93, 35J86, 49J40

DOI. 10.1137/19M1257226

1. Introduction.

We consider shape optimization problems constrained by variational inequalities (VI) of the first kind, so-called obstacle-type problems. Applications are manifold and arise whenever a shape is to be constructed in a way not to violate constraints for the state solutions of partial differential equations depending on a geometry to be optimized. Just think of a heat equation depending on a shape, where the temperature is not allowed to surpass a certain threshold. This example is basically the model problem that we are formulating in section 2. Applications of general VIs include contact problems in solid state mechanics, viscoplasticity and network equilibrium problems, and thus a wide range of industrial problems (cf. [38, 1, 33, 14]).

Shape optimization problem constraints in the form of VIs are challenging, since classical constraint qualifications for deriving Lagrange multipliers generically fail. Therefore, not only the development of stable numerical solution schemes but also the development of suitable first order optimality conditions is an issue.

By usage of tools of modern analysis, such as monotone operators in Banach spaces, significant results on properties of solution operators of VIs have been achieved since the 1960s (cf. [6, 7, 28]). However, there are very few approaches in the literature to the problem class of VI constrained shape optimization problems so far. In [27], shape optimization of two-dimensional (2D) elasto-plastic bodies is studied, where the shape is simplified to a graph such that one dimension can be written as a function of the other. The nontrivial existence of solutions of VI constrained shape optimization

*Received by the editors April 18, 2019; accepted for publication (in revised form) April 27, 2020; published electronically July 16, 2020.
https://doi.org/10.1137/19M1257226
Funding: This work has been partly supported by the German Research Foundation (DFG) within the priority program "Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization," SPP 1962/1 and SPP 1962/2, under contract Schu 804/15-1 and WE 6629/1-1, by the DFG research training group 2126 Algorithmic Optimization, and by the BMBF (Bundesministerium für Bildung und Forschung) within the collaborative project GIVEN (FKZ: 05M18UTA).
†Department of Mathematics, University of Trier, Trier, 54286 Germany (luft@uni-trier.de, volker.schulz@uni-trier.de).
‡Helmut Schmidt University/University of the Federal Armed Forces, Faculty of Mechanical Engineering, 22043 Hamburg, Germany (welker@hsu-hh.de).

problems is discussed in [10, 44]. In [44, Chapter 4], shape derivatives of elliptic VI constrained problems are presented in the form of solutions to VIs themselves. In [35], shape optimization for 2D graph-like domains is investigated. Also [29, 30] present existence results for shape optimization problems which can be reformulated as optimal control problems, whereas [12, 16] show existence of solutions in a more general set-up. In [36, 37], level-set methods are proposed and applied to graph-like 2D problems. Moreover, [20] presents a regularization approach to the computation of shape and topological derivatives in the context of elliptic VIs, thus circumventing the numerical problems in [44, Chapter 4]. Recently, in [17], a sensitivity analysis was performed for a class of semilinear VIs and a strong convergence property is shown for the material derivative. Furthermore, state-shape derivatives are established under regularity assumptions.

In this paper, we aim at optimality conditions for VI constrained shape optimization along the lines of optimality conditions for VI constrained optimal control problems as in [18, 19, 21]. In general, standard necessary optimality conditions cannot be formulated in a straightforward manner for semismooth shape optimization problems. Under appropriate assumptions, we prove existence of adjoints and convergence of adjoints resulting from regularized VIs. These analytical results are also verified numerically. Moreover, convergence of shape derivatives related to the smoothed problem is shown and the limit object is identified. Furthermore, we build on the resulting optimality conditions and devise an optimization algorithm giving specific numerical results. This algorithm no longer depends on smoothing strategies as in [15]. In [15], a shape optimization method based on a regularized variant of the VI has been devised and it is observed that the performance of this algorithm strongly depends on the tightness of the obstacle. This problem no longer arises with the strategy developed in the present paper. On the contrary, the algorithm gets even faster the more degrees of freedom are constrained by the obstacle.

This paper is structured as follows. In section 2, we formulate the VI constrained shape optimization model with general elliptic coefficients on which we focus in this paper. The necessary optimality conditions, including the existence of adjoint variables under certain regularity assumptions to the model problem, are formulated in section 3. In section 4, we formulate an algorithm to solve the model problem, based on these analytical results and compare numerically this approach with several regularized strategies.

2. Problem class. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain equipped with a sufficiently smooth boundary $\partial\Omega$, where $n \in \mathbb{N}$ is the dimension. For typical applications $n = 2$ or $n = 3$. This domain is assumed to be partitioned into a subdomain $\Omega_{\text{out}} \subset \Omega$ and an interior domain $\Omega_{\text{int}} \subset \Omega$ with boundary $\Gamma_{\text{int}} := \partial\Omega_{\text{int}}$ such that $\Omega_{\text{out}} \sqcup \Omega_{\text{int}} \sqcup \Gamma_{\text{int}} = \Omega$, where $\sqcup$ denotes the disjoint union. The closure of Ω is denoted by $\bar{\Omega}$. We consider Ω depending on Γ_{int}, i.e., $\Omega = \Omega(\Gamma_{\text{int}})$. Figure 1 illustrates this situation. In the following, the boundary Γ_{int} of the interior domain is called the *interface* and an element of an appropriate shape space $\mathcal{X}$ (cf. Remark 1). In contrast to the outer boundary $\partial\Omega$, which is assumed to be fixed, the inner boundary Γ_{int} is variable. If Γ_{int} changes, then the subdomains $\Omega_{\text{int}}, \Omega_{\text{out}} \subset \Omega$ change in a natural manner.

Let $\nu > 0$ be an arbitrary constant. For the objective function

$$J(y, \Omega) := \mathcal{J}(y, \Omega) + \mathcal{J}_{\text{reg}}(\Omega) := \frac{1}{2} \int_\Omega |y - \bar{y}|^2 \ dx + \nu \int_{\Gamma_{\text{int}}} 1 \ ds \tag{1}$$

we consider the shape optimization problem

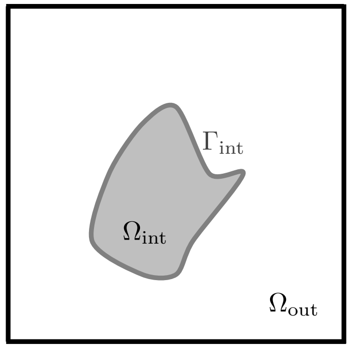

FIG. 1. *Example of a domain* $\Omega = \Omega_{out} \sqcup \Gamma_{int} \sqcup \Omega_{int}$.

$$\min_{\Gamma_{\text{int}} \in \mathcal{X}} \; J(y, \Omega) \tag{2}$$

constrained by the obstacle-type VI

$$a(y, v - y) \geq \langle f, v - y \rangle \quad \forall v \in K := \{\theta \in H_0^1(\Omega) \colon \theta(x) \leq \varphi(x) \text{ in } \Omega\}, \tag{3}$$

where $y \in K$ is the solution of the VI, $f \in L^2(\Omega)$ is explicitly dependent on the shape, $\langle \cdot, \cdot \rangle$ denotes the duality pairing, and $a(\cdot, \cdot)$ is a general strongly elliptic, i.e., coercive, symmetric bilinear form

$$\begin{aligned} &a \colon H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}, \\ &\qquad (y, v) \mapsto \int_\Omega \sum_{i,j} a_{i,j} \partial_i y \partial_j v + \sum_i d_i (\partial_i y v + y \partial_i v) + byv \, dx \end{aligned} \tag{4}$$

defined by coefficient functions $a_{i,j}, d_j, b \in L^\infty(\Omega)$, fulfilling the weak maximum principle. However, the results of this paper still remain correct if symmetry of $a(.,.)$ is dropped as an assumption by simple modifications of proofs.

With the tracking-type objective $\mathcal{J}$ the model is fitted to data measurements $\bar{y} \in H^1(\Omega)$. The second term $\mathcal{J}_{\text{reg}}$ in the objective function J is a perimeter regularization. A perimeter regularization is frequently used to overcome ill-posedness of inverse problems; e.g., [3] investigates the regularization and numerical solution of geometric inverse problems related to linear elasticity. In (3), φ denotes an obstacle which needs to be an element of $L^1_{\text{loc}}(\Omega)$ such that the set of admissible functions K is nonempty (cf. [44]). If additionally $\partial\Omega$ is Lipschitzian and $\varphi \in H^1(\Omega)$ with $\varphi_{|\partial\Omega} \geq 0$, then there is a unique solution to (3) satisfying $y \in H_0^1(\Omega)$, given that the assumptions from above hold (cf. [22, 9, 46]). Further, (3) can be equivalently expressed as

$$a(y, v) + (\lambda, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega), \tag{5}$$

$$\begin{aligned} \lambda &\geq 0 \quad \text{in } \Omega, \\ y &\leq \varphi \quad \text{in } \Omega, \\ \lambda(y - \varphi) &= 0 \quad \text{in } \Omega \end{aligned} \tag{6}$$

with $(\cdot, \cdot)_{L^2(\Omega)}$ denoting the L^2-scalar product and $\lambda \in L^2(\Omega)$.

It is well known, e.g., from [9], that under these assumptions there exists a unique solution y to the obstacle-type VI (3) and an associated Lagrange multiplier λ. The existence of solutions of any shape optimization problem is a nontrivial question. Shape optimization problems constrained by VIs are especially challenging because, in general, it is not guaranteed that an adjoint state can be introduced (cf. [44, Example in Chapter 1, Chapter 4]). An essential theoretical tool for the study of the existence of solutions is the derivation of optimality conditions, i.e., in particular, the formulation of an adjoint equation. Therefore, section 3 investigates the model problem analytically, also in view of formulating a numerically applicable algorithm in section 4.

Remark 1. The interface Γ_{int} is an element of an appropriate shape space. Please note that there exists no common shape space suitable for all applications. It should be mentioned that the existence of shape derivatives and their form is not dependent on the explicit choice of a shape space, hence only requirements noted in the relevant theorems are necessary. From a computational point of view one has to deal with polygonal shape representations arising in the setting of constrained shape optimization. This is owed to the fact that finite element methods usually discretize the models. In this paper, we use a Steklov–Poincaré metric as introduced in [42]. Such metrics can be considered, e.g., on the space B_e (cf. [32]) or more generally on the space of $H^{1/2}$-shapes (cf. [48]). In [43], it is outlined that this is an essential step toward applying efficient FE solvers. Of course, it is possible to choose other shape space models, but this is beyond the scope of this paper.

3. Convergence results for adjoints and shape derivatives. We assume the situation mentioned in section 2, which is also found in [23], giving us $\lambda \in L^2(\Omega)$. It can be easily verified that this in turn gives the possibility to summarize the conditions (6) equivalently into a single condition of the form

$$\lambda = \max\big(0, \lambda + c(y - \varphi)\big) \quad \text{for any } c > 0. \tag{7}$$

The direct handling of general obstacle-type VIs formulated as in (5)–(6), with (6) being equivalently substitutes by (7), poses several challenges. One challenge of the solution of (5) is the occurrence of distributional numerical iterates for λ in $H^{-1}(\Omega)$ when an augmented Lagrangian approach is applied to (5) constrained by (7), despite the analytical solution λ having $L^2(\Omega)$-regularity. For a more detailed discussion of this, see [23, p. 2]. In order to circumvent the occurrence of distributions in the solution scheme, the authors of [23] introduce a relaxation for relation (7) with a given regularization parameter $\alpha \in (0, 1)$,

$$\lambda = \alpha \cdot \max\big(0, \lambda + \varrho(y - \varphi)\big) \quad \text{for any } \varrho > 0, \tag{8}$$

which in turn is equivalent to

$$\lambda = \max\big(0, \bar{\lambda} + c(y - \varphi)\big) \quad c \in (0, \infty) \tag{9}$$

if $\bar{\lambda} = 0$ and $c = \frac{\varrho\alpha}{1-\alpha}$, where $\bar{\lambda} \in L^2(\Omega)$ can be motivated by updates of the augmented Lagrangian. This results in the equation

$$a(y_c, v) + \big(\max\big(0, \bar{\lambda} + c(y_c - \varphi)\big), v\big)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \qquad \forall v \in H_0^1(\Omega), \tag{10}$$

which in the following is called the *regularized state equation* or *relaxed obstacle problem.* Explicit dependence on λ is avoided, making the resulting semilinear elliptic equation tractable, for example, by semismooth Newton methods; see, e.g.,

[23]. Moreover, the authors of [23] prove L^2-convergence of the regularized multiplier $\max\big(0, \bar\lambda + c\cdot(y_c-\varphi)\big)$ to the original λ for their proposed semismooth Newton method.

With problem (10) we are still left to solve a nonlinear, semismooth problem, giving rise to problems concerning existence of adjoints for the shape optimization problem. Hence, standard smoothing strategies can be applied to render this problem smooth enough to show existence of adjoints and to apply techniques such as Newton iterations.

In light of [41] and [8], we pose the following assumptions on the smoothed max-function, which from now on is called $\max_\gamma\colon \mathbb{R} \to [0,\infty)$, with $\gamma > 0$ being the smoothing parameter.

Assumption 1 (on smoothed max-function).
(i) $\max_\gamma \in C^1(\Omega)$ for all $\gamma > 0$.
(ii) There exists a function $g\colon (0,\infty) \to [0,\infty)$ with $g(\gamma) \to 0$ as $\gamma \to \infty$ s.t. $|\max_\gamma(x) - \max(0,x)| \le g(\gamma)$ for all $x \in \mathbb{R}$ and for all $\gamma > 0$.
(iii) $\max'_\gamma(x) \in [0,1]$ and monotonically nondecreasing for all $x \in \mathbb{R}$ and all $\gamma > 0$.
(iv) $\max'_\gamma$ converges uniformly to 0 on $(-\infty,-\delta)$ and 1 on (δ,∞) for all $\delta > 0$ for $\gamma \to \infty$.

In the following, let sign_γ denote the derivative of $\max_\gamma$. An example satisfying these assumptions is given in (41). Applying $\max_\gamma$ instead of max in (10) gives the following equation, which we call the *fully regularized state equation* in the subsequent sections:

$$a(y_{\gamma,c}, v) + \big(\max_\gamma\big(\bar\lambda + c(y_{\gamma,c}-\varphi)\big), v\big)_{L^2(\Omega)} = (f,v)_{L^2(\Omega)} \quad \forall v \in H^1_0(\Omega). \tag{11}$$

So linearizing the corresponding Lagrangian with respect to $y_{\gamma,c}$ results in the typical adjoint equation

$$\begin{aligned} &a(p_{\gamma,c}, v) + c\cdot\big(\mathrm{sign}_\gamma(\bar\lambda + c(y_{\gamma,c}-\varphi))\cdot p_{\gamma,c}, v\big)_{L^2(\Omega)} \\ &\quad = -(y_{\gamma,c} - \bar y, v)_{L^2(\Omega)} \quad \forall v \in H^1_0(\Omega) \end{aligned} \tag{12}$$

(see, e.g., [18] or [41] in the context of optimal control).

Remark 2. As in [41], smoothness of the state equation (11) in $y_{\gamma,c}$ guarantees existence of solutions to the linearized equation (12) for a given $L^2(\Omega)$ right-hand side and, thus, existence of adjoints in the case of the considered tracking-type objective functional (1).

3.1. State and adjoint equation. We first show that solutions of (11) converge strongly in H^1 to solutions of (5)–(6) for $\gamma, c \to \infty$. This is proven in [41] for stronger assumptions on the smoothed function $\max_\gamma$ and under $\gamma = c$. Since we rely on the general case $\gamma \ne c$ for the proofs in ongoing discussions, we state an appropriate result. The first part of the following theorem is in analogy to [8, Lemma 4.2]. However, the difference is that we consider general elliptic bilinear forms and—more importantly—a modified argument in the maximum function resulting in different regularized state equations. These generalizations are necessary for our further analytical investigations leading to an adjoint equation.

Proposition 3.1 (H^1-convergence of the state). *Let $y_{\gamma,c}$, y_c, and y be solutions to* (11), (10), *and* (5), *respectively. Here, $a(\cdot,\cdot)$ is chosen by an elliptic bilinear form as in* (4) *on a bounded, open domain $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary, $f \in L^2(\Omega)$,* *and* $\gamma, c > 0$*. Moreover, assume* $\varphi \in H^1(\Omega)$*,* $\bar{\lambda} \in L^2(\Omega)$ *and let* $\max_\gamma\colon \mathbb{R} \to \mathbb{R}$ *satisfy Assumption* 1*.*

Then (11) *and* (10) *possess unique solutions and*

$$y_{\gamma,c} \to y_c \text{ in } H^1(\Omega) \quad \text{as } \gamma \to \infty; \tag{13}$$

$$y_c \to y \text{ in } H^1(\Omega) \quad \text{as } c \to \infty. \tag{14}$$

Proof. We prove statement (13) of the theorem. For a proof of statement (14), we refer to [23, Theorem 3.1].

We start by ensuring the existence of solutions to (11) and (10). For this, we show that the Nemetskii operator defined by

$$\Phi_\gamma\colon H^1(\Omega) \to L^2(\Omega), y \mapsto \max_\gamma\big(\bar{\lambda} + c \cdot (y - \varphi)\big) \tag{15}$$

is a monotone operator for all $\gamma, c > 0$. Due to Assumption 1, it is clear that $\max_\gamma\colon \mathbb{R} \to \mathbb{R}$ is a pointwise monotone function, implying that $\max_\gamma\colon H^1(\Omega) \to H^1(\Omega), y \mapsto \max_\gamma(y)$ is a monotone operator. Since

$$\Psi_c\colon H^1(\Omega) \to H^1(\Omega),\ y \mapsto \bar{\lambda} + c \cdot (y - \varphi)$$

is an affine linear operator, and thus monotone, the composition $\max_\gamma \circ \Psi_c = \Phi_\gamma$ is also monotone. The same argument holds for the nonsmoothed operator

$$\Phi\colon H^1(\Omega) \to L^2(\Omega), y \mapsto \max\big(0, \bar{\lambda} + c \cdot (y - \varphi)\big).$$

Therefore, applying the Minty–Browder theorem for monotone operators yields the existence of unique solutions to (11) and (10) in $H^1(\Omega)$ for all $f \in L^2(\Omega)$ if Ω is bounded and we operate in Hilbert spaces.

Now, we prove the second convergence (13). For fixed $c > 0$, let $y_{\gamma,c}$ and y_c be solutions to (11) and (10), respectively. Assumption 1(ii) together with the monotonicity of Φ, the coercivity of $a(\cdot,\cdot)$ with constant $K > 0$, and $y_{\gamma,c} - y_c \in H^1(\Omega)$ acting as a test function yields

$$\begin{aligned}
0 &\le K \cdot ||y_{\gamma,c} - y_c||^2_{H^1(\Omega)} \\
&\le a(y_{\gamma,c} - y_c, y_{\gamma,c} - y_c) \\
&\le a(y_{\gamma,c} - y_c, y_{\gamma,c} - y_c) \\
&\quad + \Big(\max\big(0, \bar{\lambda} + c(y_{\gamma,c} - \varphi)\big) - \max\big(0, \bar{\lambda} + c(y_c - \varphi)\big), y_{\gamma,c} - y_c\Big)_{L^2(\Omega)} \\
&= \Big(\max_\gamma\big(\bar{\lambda} + c(y_{\gamma,c} - \varphi)\big) - \max\big(0, \bar{\lambda} + c(y_{\gamma,c} - \varphi)\big), y_{\gamma,c} - y_c\Big)_{L^2(\Omega)} \\
&\le \int_\Omega |\max_\gamma\big(\bar{\lambda} + c(y_{\gamma,c} - \varphi)\big) - \max\big(0, \bar{\lambda} + c(y_{\gamma,c} - \varphi)\big)| \cdot |y_{\gamma,c} - y_c|\, dx \\
&\le g(\gamma) \cdot \mathrm{vol}(\Omega)^{\frac{1}{2}} \cdot ||y_{\gamma,c} - y_c||_{H^1(\Omega)},
\end{aligned}$$

which gives the desired convergence (13). ☐

The following definition is needed to state the first main result of this paper, the convergence of adjoints.

Definition 3.2. *Let* $\Omega \subset \mathbb{R}^n$ *be a bounded, open domain with Lipschitz boundary. A set* $A \subseteq \Omega$ *is called regularly decomposable if there exists an* $N \in \mathbb{N}$ *and path-connected, bounded, open* $A_i \subset \Omega$ *with Lipschitz boundaries* ∂A_i *such that* $A = \sqcup_{i=1}^N \bar{A}_i$ *is a disjoint union.*

With this definition it is possible to formulate the first main theorem concerning the convergence of adjoints corresponding to the fully regularized problems and characterization of the limit object.

THEOREM 3.3 (convergence of the adjoints). *Let $\Omega \subset \mathbb{R}^n$ for $n \leq 4$ be a bounded, open domain with Lipschitz boundary. Moreover, let the following assumptions be satisfied:*

(i) *$\varphi \in H^1(\Omega)$, $f \in L^2(\Omega)$, $\bar{y} \in H^1(\Omega)$ and coefficient functions $a_{i,j}, d_j, b \in L^\infty(\Omega)$ in (5)–(6);*

(ii) *the active set $A = \{x \in \Omega \mid y - \varphi \geq 0\}$ corresponding to (5)–(6) is regularly decomposable;*

(iii) *$A_c := \{x \in \Omega \mid \bar{\lambda} + c \cdot (y_c - \varphi) \geq 0\}$ is regularly decomposable and*

$$A_c \subseteq A \quad \forall c > 0, \tag{16}$$

where y_c solves the regularized state equation (10);

(iv) *the following convergence holds:*

$$\|\mathrm{sign}_\gamma(\bar{\lambda} + c \cdot (y_{\gamma,c} - \varphi)) - \mathrm{sign}(\bar{\lambda} + c \cdot (y_c - \varphi))\|_{L^1(\Omega)} \to 0 \quad \text{for } \gamma \to \infty. \tag{17}$$

Then the adjoints $p_{\gamma,c} \to p_c$ in $H_0^1(\Omega)$ for $\gamma \to \infty$ for all $c > 0$, where p_c is the solution to

$$a(p_c, v) + c \cdot \int_\Omega \mathbb{1}_{A_c} \cdot p_c \cdot v \, dx = -\int_\Omega (y_c - \bar{y}) \cdot v \, dx \quad \forall v \in H_0^1(\Omega). \tag{18}$$

Moreover, there exists $p \in H^{-1}(\Omega)$ to (5)–(6) and p is representable as an H_0^1-function given by the extension of $\tilde{p} \in H_0^1(\Omega \setminus A)$ to $\bar{\Omega}$, i.e.,

$$p = \begin{cases} \tilde{p} & \text{in } \Omega \setminus A, \\ 0 & \text{in } A, \end{cases} \tag{19}$$

where $\tilde{p} \in H_0^1(\Omega \setminus A)$ is the solution of the elliptic problem

$$a_{\Omega\setminus A}(\tilde{p}, v) = -\int_{\Omega\setminus A} (y - \bar{y}) v \, dx \qquad \forall v \in H_0^1(\Omega \setminus A) \tag{20}$$

with

$$\begin{aligned} &a_{\Omega\setminus A} : H_0^1(\Omega \setminus A) \times H_0^1(\Omega \setminus A) \to \mathbb{R}, \\ &(\tilde{p}, v) \mapsto \int_{\Omega\setminus A} \sum_{i,j} a_{i,j} \partial_i \tilde{p} \partial_j v + \sum_i d_i (\partial_i \tilde{p} v + \tilde{p} \partial_i v) + b \tilde{p} v \, dx \end{aligned} \tag{21}$$

being the restriction of bilinear form $a(\cdot,\cdot)$ to $\Omega \setminus A$.

Further, the solutions p_c of (18) converge strongly in $H_0^1(\Omega)$ to the H_0^1-representation of p.

Proof. Let us consider the regularized problem (11) for $\gamma, c > 0$. Existence and uniqueness of solutions $y_{\gamma,c}, p_{\gamma,c}$ of regularized state and adjoint are guaranteed by application of the Minty–Browder theorem in analogy to Theorem 3.3 for $y_{\gamma,c}$ and the Lax–Milgram theorem for $p_{\gamma,c}$, respectively.

This proof consists of two main parts:

1. showing the H^1-convergence of the smoothed to the nonsmoothed regularized adjoint $p_{\gamma,c} \to p_c$ for $\gamma \to \infty$,
2. analyzing the limit PDE (18) for $c \to \infty$ and proving that $p_c \to p$ in $H^1(\Omega)$ for $c \to \infty$, where p is defined as in (19).

To 1. We start to show the H^1-convergence of the smoothed to the nonsmoothed regularized adjoint $p_{\gamma,c} \to p_c$ for $\gamma \to \infty$.

The assumption (17) of L^1-convergence of $\operatorname{sign}_\gamma(\bar\lambda + c \cdot (y_{\gamma,c} - \varphi))$ is equivalent to L^p-convergence for all $p \in [1,\infty)$ in our setting, since

$$\begin{aligned}&\|\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) - \operatorname{sign}(\bar\lambda + c\cdot(y_c-\varphi))\|_{L^p(\Omega)}\\ &\quad\le \|\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) - \operatorname{sign}(\bar\lambda + c\cdot(y_c-\varphi))\|_{L^1(\Omega)}^{1/p} \to 0 \quad \text{for } \gamma\to\infty\end{aligned}$$

by monotony of the integral and Assumption 1(ii)–(iv). Denote by $S_{\gamma,c}\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ the linear operator corresponding to the left-hand side of the smoothed adjoint equation (12) and $S_c\colon H^1(\Omega) \to H^{-1}(\Omega)$ the one to (18). We establish convergence of $S_{\gamma,c}$ to S_c in the operator norm. In the following, we apply Hölder's inequality, and moreover we use L^p-convergence of $\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi))$ for all $p\in[1,\infty)$ as well as boundedness of $\operatorname{sign}_\gamma$ and sign.

Further, since we are in the situation $\Omega \subset \mathbb{R}^n$ for $n \le 4$, we have the following embedding with embedding constant $C > 0$ (cf. [40, Theorem 4.12, Part I, Case C]):

$$H_0^1(\Omega) \hookrightarrow L^4(\Omega) \quad \text{for } n \le 4. \tag{22}$$

Combining all this yields

$$\begin{aligned}&\|S_{\gamma,c} - S_c\|_{\mathrm{op}}\\ &\quad= \sup_{\substack{g\in H_0^1(\Omega)\\ \|g\|=1}} \sup_{\substack{h\in H_0^1(\Omega)\\ \|h\|=1}} c\cdot |((\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) - \operatorname{sign}(\bar\lambda + c\cdot(y_c-\varphi)))\cdot g, h)_{L^2(\Omega)}|\\ &\quad\le \sup_{\substack{g\in H_0^1(\Omega)\\ \|g\|=1}} \sup_{\substack{h\in H_0^1(\Omega)\\ \|h\|=1}} c\cdot \big\|\big(\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi))\\ &\qquad\qquad - \operatorname{sign}(\bar\lambda + c\cdot(y_c-\varphi))\big)\big\|_{L^2(\Omega)}\cdot\|g\|_{L^4(\Omega)}\cdot\|h\|_{L^4(\Omega)}\\ &\quad\le C^2\cdot c\cdot\|\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) - \operatorname{sign}(\bar\lambda + c\cdot(y_c-\varphi))\|_{L^2(\Omega)} \to 0 \quad\text{for } \gamma\to\infty,\end{aligned}$$

which gives the desired convergence in the operator norm. Using analyticity of the inversion $\mathcal{I}\colon S \mapsto S^{-1}$ in the domain of invertible, bounded, linear operators given in our setting, convergence of the solution operators $S_{\gamma,c}^{-1} \to S_c^{-1}$ in operator norm is implied immediately; see, e.g., [49, p. 237]. Combining this with the convergence of $y_{\gamma,c} \to y_c$ in $H_0^1(\Omega)$ established by Theorem 3.3 yields

$$\begin{aligned}&\|p_{\gamma,c} - p_c\|_{H_0^1(\Omega)}\\ &\quad= \|-S_{\gamma,c}^{-1}(y_{\gamma,c}-\bar y) + S_c^{-1}(y_c-\bar y)\|_{H_0^1(\Omega)}\\ &\quad\le \|S_{\gamma,c}^{-1}(y_{\gamma,c}-\bar y) - S_{\gamma,c}^{-1}(y_c-\bar y)\|_{H_0^1(\Omega)} + \|S_{\gamma,c}^{-1}(y_c-\bar y) - S_c^{-1}(y_c-\bar y)\|_{H_0^1(\Omega)}\\ &\quad\le \|S_{\gamma,c}^{-1}\|_{\mathrm{op}}\|y_{\gamma,c}-y_c\|_{H_0^1(\Omega)} + \|S_{\gamma,c}^{-1} - S_c^{-1}\|_{\mathrm{op}}\|y_c-\bar y\|_{H_0^1(\Omega)} \to 0 \quad\text{for } \gamma\to\infty\end{aligned}$$

since $\|S_{\gamma,c}^{-1}\|_{\mathrm{op}}$ can be bounded due to convergence.

To 2. Next, we analyze the limit PDE (18) for $c \to \infty$. We show that $p_c \to p$ in $H^1(\Omega)$ for $c\to\infty$, where p is defined as in (19). For this, we first notice that our assumption concerning regular decomposability of $A = \{x\in\Omega \mid y-\varphi \ge 0\}$ ensures that $\partial A = \{x \in \Omega \mid y - \varphi = 0\}$ forms a $C^{0,1}$-manifold embedded in Ω. This in turn leads to well definedness of the restricted bilinear form $a_{\Omega \setminus A}(.,.)$ and the well-posedness of the variational problem (20) and, thus, of $p \in H_0^1(\Omega)$. Our next step is to show

$$p_c \to p \quad \text{in } H^1(\Omega) \text{ for } c \to \infty. \tag{23}$$

To show this, we artificially constrain problem (18) to $A \subseteq \Omega$. So denote by $a_A(\cdot,\cdot)$ the restriction of the bilinear form $a(\cdot,\cdot)$ to $A \subseteq \Omega$, defined in analogy to (21). The corresponding restricted problem becomes

$$a_A(p_{c_{|A}}, v) + c \cdot \int_A \mathbb{1}_{A_c} \cdot p_{c_{|A}} \cdot v \, dx = -\int_A (y_c - \bar{y}) \cdot v \, dx \quad \forall v \in H_0^1(A), \tag{24}$$

where the Dirichlet condition $p_{c_{|A}} = p_c$ on ∂A is incorporated in the usual way. Dividing by $c > 0$ gives an equivalent equation in the sense that a solution $p_{c_{|A}} \in H^1(A)$ to (24) also solves the equivalent equation

$$\frac{1}{c} \cdot a_A(p_{c_{|A}}, v) + \int_A \mathbb{1}_{A_c} \cdot p_{c_{|A}} \cdot v \, dx = -\frac{1}{c} \cdot \int_A (y_c - \bar{y}) \cdot v \, dx \quad \forall v \in H_0^1(A). \tag{25}$$

The differential operator corresponding to the left-hand side of the equivalent equation (25) is given by

$$S_{A,c} \colon H^1(A) \to H^{-1}(A),\ p \mapsto \frac{1}{c} \cdot a_A(p,\cdot) + (\mathbb{1}_{A_c} \cdot p, \cdot)_{L^2(A)}. \tag{26}$$

Next, we show that the differential operators $S_{A,c}$ converge in the linear operator norm $\|\cdot\|_{\mathrm{op}}$ with the limit operator

$$S_A \colon H^1(A) \to H^{-1}(A),\ p \mapsto (p,\cdot)_{L^2(A)}. \tag{27}$$

Hence

$$\begin{aligned}
&\|S_{A,c} - S_A\|_{\mathrm{op}} \\
&\quad = \sup_{\substack{g \in H_0^1(A) \\ \|g\|=1}} \sup_{\substack{h \in H_0^1(A) \\ \|h\|=1}} \left| \frac{1}{c} \cdot a_A(g,h) - \int_{A \setminus A_c} g \cdot h \, dx \right| \\
&\quad \le \sup_{\substack{g \in H_0^1(A) \\ \|g\|=1}} \sup_{\substack{h \in H_0^1(A) \\ \|h\|=1}} \Bigg(\frac{1}{c} \Big(\sum_{i,j} \|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j \|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)} \Big) \\
&\qquad\qquad \cdot \|g\|_{H_0^1(A)} \cdot \|h\|_{H_0^1(A)} \\
&\qquad\qquad + \mathrm{vol}(A \setminus A_c)^{\frac{1}{2}} \cdot \|g\|_{L^4(A)} \cdot \|h\|_{L^4(A)} \Bigg) \\
&\quad = \frac{1}{c} \left(\sum_{i,j} \|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j \|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)} \right) + C^2 \cdot \mathrm{vol}(A \setminus A_c)^{\frac{1}{2}} \\
&\quad \to 0 \quad \text{for } c \to \infty,
\end{aligned}$$

due to (22) and since $\mathrm{vol}(A \setminus A_c) \to 0$ for $c \to \infty$, which would otherwise contradict $y_c \to y$ in $H_0^1(\Omega)$. We can now apply a similar argument as in step 1, namely the analyticity of the inversion operator $\mathcal{I}\colon S \mapsto S^{-1}$, giving us convergence of the solution operators $S_{A,c}^{-1} \to S_A^{-1}$ in $\|\cdot\|_{\mathrm{op}}$. Also notice that we can obtain the sequence of solutions $p_{c_{|A}}$ by solving (25) with the corresponding right-hand sides $-\frac{1}{c}(y_c - \bar{y})$ instead of the original equation (24) and that the right-hand sides converge to 0 in $H^1(\Omega)$ as $c \to \infty$, as y_c is convergent by Proposition 3.1. We conclude

$$\begin{aligned}
0 &\leq \|p_c\|_{H^1(A)} = \|S_{A,c}^{-1}\left(-\frac{1}{c}(y_c - \bar{y})\right)\|_{H^1(A)} \\
&\leq \frac{1}{c}\|(S_{A,c}^{-1} - S_A^{-1})(y_c - \bar{y})\|_{H^1(A)} + \frac{1}{c}\|S_A^{-1}(y_c - \bar{y})\|_{H^1(A)} \\
&\leq \frac{1}{c}\big(\|S_{A,c}^{-1} - S_A^{-1}\|_{\mathrm{op}} + \|S_A^{-1}\|_{\mathrm{op}}\big) \cdot \big(\|y_c - y\|_{H^1(A)} + \|\bar{y}\|_{H^1(A)}\big) \\
&\to 0 \quad \text{for } c \to \infty.
\end{aligned}$$

For the proof of convergence it remains to address the convergence of p_c on $\Omega \setminus A$. We can artificially restrict (18) to $\Omega \setminus A$ by imposing the Dirichlet boundary $p_{c|A}$ on ∂A, since ∂A forms a $C^{0,1}$-submanifold of Ω as we assumed regular decomposability (cf. Definition 3.2) of the active set A. To distinguish the corresponding bilinear forms, we denote the restricted bilinear form by $a_{\Omega\setminus A}$. Since the unrestricted bilinear form $a(\cdot,\cdot)$ is strongly elliptic, coercivity for some constant $K > 0$ also holds for $a_{\Omega\setminus A}$. This together with Hölder's inequality, assumption $A_c \subseteq A$ for all $c > 0$, and the fact that $p_c - \tilde{p} \in H_0^1(\Omega \setminus A)$ can act as a test function gives

$$\begin{aligned}
0 &\leq K\|p_c - \tilde{p}\|_{H^1(\Omega\setminus A)}^2 \leq a_{\Omega\setminus A}(p_c - \tilde{p}, p_c - \tilde{p}) = a_{\Omega\setminus A}(p_c, p_c - \tilde{p}) - a_{\Omega\setminus A}(\tilde{p}, p_c - \tilde{p}) \\
&= -c\int_{\Omega\setminus A} \mathbb{1}_{A_c} p_c (p_c - \tilde{p})\, dx - \int_{\Omega\setminus A} (y_c - \bar{y})(p_c - \tilde{p})\, dx + \int_{\Omega\setminus A} (y - \bar{y})(p_c - \tilde{p})\, dx \\
&= \int_{\Omega\setminus A} (y - y_c)(p_c - \tilde{p})\, dx \leq \|y_c - y\|_{H^1(\Omega)}\|p_c - \tilde{p}\|_{H^1(\Omega\setminus A)},
\end{aligned}$$

where $\tilde{p} \in H^1(\Omega \setminus A)$ is defined as in (20). This results in

$$p_c \to p \quad \text{in } H_0^1(\Omega \setminus A) \quad \text{for } c \to \infty \tag{28}$$

due to our assumptions and $y_c \to y$ in $H^1(\Omega)$ as by Proposition 3.1. Together with (23) this gives the desired convergence $p_c \to p$ in $H_0^1(\Omega)$. ∎

There are two nontrivial assumptions in Theorem 3.3: assumptions (iii) and (iv). In the following, we formulate two remarks in which we address these assumptions (cf. Remark 3 for (iii) and Remark 4 for (iv)).

Remark 3. It is possible to fulfill assumption (16) on inclusion of the active sets $A_c \subset A$ by choosing a sufficient $\bar{\lambda} \in L^2(\Omega)$. To be more precise, if we assume $\varphi \in H^2(\Omega)$, we can choose $\bar{\lambda} := \max\{0, f - S\varphi\}$ with S being the differential operator corresponding to the elliptic bilinear form $a(\cdot,\cdot)$ in (5), guaranteeing feasibility $y_{c_1} \leq y_{c_2} \leq y \leq \varphi$ for all $0 < c_1 \leq c_2$. For the proof of this, we refer to [23, section 3.2].

Remark 4. Assumption 17 ensures that convergence of sign_γ is compatible with convergence of $y_{\gamma,c}$ for $\gamma \to \infty$. To give a working example, we verify this assumption in the numerical section under (45) for several demonstrative cases.

Remark 5. The limit object $p \in H_0^1(\Omega)$ of the adjoints $p_{\gamma,c}$ as defined in (19) is the solution of an elliptic problem (20) on a domain $\Omega \setminus A$ with topological dimension greater than 0. This can be exploited in numerical computations, for instance, by a fat boundary method for finite elements on domains with holes as proposed by the authors of [34].

Remark 6. We remind the reader that p is not necessarily an adjoint to the original problem (1) constrained by (5) and (6), but merely a solution of a part of the limit of the optimality conditions for the regularized problem. For a discussion of a similar phenomenon in the context of optimal control, we refer the interested reader to [8, section 4.2].

3.2. Shape derivatives. In this section, we apply our convergence results for the regularized state and adjoint equations to derive similar convergence results for the shape derivatives of the shape optimization problem constrained by the fully regularized state equation (11). In general, shape derivatives of the unregularized VI constrained shape optimization problems do not exist (cf., e.g., [44, Chapter 1.1]). Nevertheless, we show existence of shape derivatives for the shape optimization problem constrained by the fully regularized VI (11). Then, a limiting object corresponding to the unregularized equation (5)–(6) is derived.

In the following, we split the main results into two theorems, the first one being the shape derivative for the fully regularized equation and the second one being convergence of the former for $\gamma, c \to \infty$.

The shape derivative of a general shape functional H at Ω in the direction of a sufficiently smooth vector field V is denoted by $DH(\Omega)[V]$. For the definition of shape derivatives or a detailed introduction into shape calculus, we refer to the monographs [11, 44]. In general, we have to deal with so-called material and shape derivatives of generic functions $h\colon \Omega \to \mathbb{R}$ in order to derive shape derivatives of objective shape functions. For their definitions and more details we refer to the literature, e.g., [39]. In the following, we denote the material derivative of h by $\dot{h}$ or $D_m(h)$, and the shape derivative of h in the direction of a vector field V is denoted by h'.

Remark 7. In this section, we only consider the shape functional J defined in (1) without regularization term $\mathcal{J}_{reg}$, i.e., we focus only on $\mathcal{J}$. The shape derivative of J is given by the sum of the shape derivative of $\mathcal{J}$ and $\mathcal{J}_{reg}$, where $D\mathcal{J}_{reg}(\Omega)[V] = \nu \int_{\Gamma_{int}} \kappa \langle V, n\rangle \, ds$ with $\kappa := div_{\Gamma_{int}}(n)$ denoting the mean curvature of Γ_{int}. Please note that the objective functional and the shape derivative in correlation with the regularized VI (11) depend on the parameters γ and c. In order to denote this dependency, we use the notation $\mathcal{J}_{\gamma,c}$ and $D\mathcal{J}_{\gamma,c}(\Omega)[V]$ for the objective functional and its shape derivative, respectively.

We state the first theorem, which presents the shape derivative of the objective functional $\mathcal{J}$ defined in (1) constrained by the fully regularized VI (11).

THEOREM 3.4. *Assume the setting of the shape optimization problem formulated in section* 2. *Let the assumptions of Theorem* 3.3 *hold. Moreover, let* $M := (a_{i,j})_{i,j=1,2,\dots,n}$ *be the matrix of coefficient functions to the leading order terms in* (4). *Assume* $y_{\gamma,c}, p_{\gamma,c} \in W^{1,4}(\Omega)$, $a_{ij}, b_i, d \in L^\infty(\Omega) \cap W^{1,4}(\Omega)$, *and* $f \in H^1(\Omega)$. *Furthermore, let* $D_m(y_{\gamma,c}), D_m(p_{\gamma,c}) \in H_0^1(\Omega)$ *for all* $\gamma, c > 0$. *Then the shape derivatives of* $\mathcal{J}$ *defined in* (1) *constrained by a fully regularized VI* (11) *in the direction of a vector field* $V \in H_0^1(\Omega, \mathbb{R}^n)$ *exist and are given by*

$$
\begin{aligned}
(29) \qquad & D\mathcal{J}_{\gamma,c}(\Omega)[V] \\
&= \int_\Omega -(y_{\gamma,c} - \bar{y})\nabla \bar{y}^T V - \nabla y_{\gamma,c}^T (\nabla V^T M - \nabla M \cdot V + M^T \nabla V)\nabla p_{\gamma,c} \\
&\quad + (\nabla b^T V) y_{\gamma,c} p_{\gamma,c} + y_{\gamma,c} \cdot ((\nabla d^T V)^T \nabla p_{\gamma,c} - d^T(\nabla V \nabla p_{\gamma,c}))
\end{aligned}
$$

$$
\begin{aligned}
&+ p_{\gamma,c} \cdot ((\nabla d^T V)^T \nabla y_{\gamma,c} - d^T(\nabla V \nabla y_{\gamma,c}))\\
&- c \cdot \operatorname{sign}_\gamma(\bar\lambda + c \cdot (y_{\gamma,c} - \varphi)) \cdot \nabla\varphi^T V \cdot p_{\gamma,c} - \nabla f^T V p_{\gamma,c}\\
&+ \operatorname{div}(V)\Big(\frac{1}{2}(y_{\gamma,c} - \bar y)^2 + b y_{\gamma,c} p_{\gamma,c} + \sum_{i,j} a_{i,j}\partial_i y_{\gamma,c}\partial_j p_{\gamma,c}\\
&\qquad + \sum_i d_i(\partial_i y_{\gamma,c} p_{\gamma,c} + y_{\gamma,c}\partial_i p_{\gamma,c})\\
&\qquad + \max\nolimits_\gamma\big(\bar\lambda + c\cdot(y_{\gamma,c} - \varphi)\big)p_{\gamma,c} - f p_{\gamma,c}\Big)\, dx\,.
\end{aligned}
$$

Proof. Let us consider the shape optimization problem with fully regularized state equations with parameters $\gamma, c > 0$ as in (11) and fixed shape Γ_{int} to derive a corresponding shape derivative. The first part of the proof, consisting of the existence of shape derivatives $D\mathcal{J}_{\gamma,c}$ for all $\gamma, c > 0$, is found in Appendix A.

As the second part of the proof, we derive the shape derivative expression. Note that (29) and following integrals are well defined due to our assumptions on integrability combined with (22) and dimension $n \le 4$. By applying standard shape calculus techniques (cf. [4, 47]) to the target functional part of the Lagrangian we get

$$
\begin{aligned}
&D\Big(\frac{1}{2}\int_\Omega (y_{\gamma,c} - \bar y)^2 dx\Big)[V]\\
&= \int_\Omega (y_{\gamma,c} - \bar y)(D_m(y_{\gamma,c}) - D_m(\bar y)) + \frac{1}{2}\operatorname{div}(V)(y_{\gamma,c} - \bar y)^2 dx\\
&= \int_\Omega (y_{\gamma,c} - \bar y) D_m(y_{\gamma,c}) dx + \int_\Omega -(y_{\gamma,c} - \bar y)\nabla\bar y^T V + \frac{1}{2}\operatorname{div}(V)(y_{\gamma,c} - \bar y)^2 dx
\end{aligned}
\tag{30}
$$

since the target $\bar y \in L^2(\Omega)$ does not depend on the shape. Next, as similarly found in, e.g., [47], we calculate the shape derivative of the bilinear form $a(\cdot,\cdot)$. For avoiding confusion with the active sets A and A_c, we call the coefficient matrix $(a_{i,j})_{i,j}$ of the leading order parts of the bilinear form M. As before we have

$$
\begin{aligned}
D(a(y_{\gamma,c}, p_{\gamma,c}))[V] = \int_\Omega D_m(a(y_{\gamma,c}, p_{\gamma,c})) + \operatorname{div}(V)\Bigg(&\sum_{i,j} a_{i,j}\partial_i y_{\gamma,c}\partial_j p_{\gamma,c}\\
&+ \sum_i d_i(\partial_i y_{\gamma,c} p_{\gamma,c} + y_{\gamma,c}\partial_i p_{\gamma,c}) + b y_{\gamma,c} p_{\gamma,c}\Bigg) dx.
\end{aligned}
\tag{31}
$$

We use linearity, chain rules, product rules, and gradient identities for the material derivative $D_m(\cdot)$, as found in [4], to reformulate $D_m\big(a(y_{\gamma,c}, p_{\gamma,c})\big)$. For readability, we analyze each term individually. We start with the leading order terms:

$$
\begin{aligned}
&D_m\left(\sum_{i,j} a_{i,j}\partial_i y_{\gamma,c}\partial_j p_{\gamma,c}\right)\\
&= \sum_{i,j} D_m(a_{i,j})\partial_i y_{\gamma,c}\partial_j p_{\gamma,c} + a_{i,j} D_m(\partial_i y_{\gamma,c})\partial_j p_{\gamma,c} + a_{i,j}\partial_i y_{\gamma,c} D_m(\partial_j p_{\gamma,c})\\
&= \sum_{i,j} D_m(a_{i,j})\partial_i y_{\gamma,c}\partial_j p_{\gamma,c} + a_{i,j}\Big((\partial_i D_m(y_{\gamma,c}) - \sum_k \partial_k y_{\gamma,c}\partial_i V_k)\partial_j p_{\gamma,c}\\
&\qquad + \partial_i y_{\gamma,c}(\partial_j D_m(p_{\gamma,c}) - \sum_k \partial_k p_{\gamma,c}\partial_j V_k)\Big)
\end{aligned}
$$

$$
\begin{aligned}
=&\sum_{i,j}\Big(\nabla a_{i,j}^T V\partial_i y_{\gamma,c}\partial_j p_{\gamma,c}+a_{i,j}\partial_i D_m(y_{\gamma,c})\partial_j p_{\gamma,c}+a_{i,j}\partial_i y_{\gamma,c}\partial_j D_m(p_{\gamma,c})\\
&\quad -a_{i,j}(\partial_i V^T\nabla y_{\gamma,c})\partial_j p_{\gamma,c}-a_{i,j}\partial_i y_{\gamma,c}(\partial_j V^T\nabla p_{\gamma,c})\Big)\\
=&\nabla y_{\gamma,c}^T(\nabla M^T V)\nabla p_{\gamma,c}+\sum_{i,j}\Big(a_{i,j}\partial_i D_m(y_{\gamma,c})\partial_j p_{\gamma,c}+a_{i,j}\partial_i y_{\gamma,c}\partial_j D_m(p_{\gamma,c})\Big)\\
&-\nabla y_{\gamma,c}^T(\nabla V^T M)\nabla p_{\gamma,c}-\nabla y_{\gamma,c}^T(M\nabla V)\nabla p_{\gamma,c}.
\end{aligned}
$$

For the first order terms of $a(\cdot,\cdot)$ we only compute $y_{\gamma,c}d^T\nabla p_{\gamma,c}$, since calculations are analogous for the second term by switching the roles of $y_{\gamma,c}$ and $p_{\gamma,c}$. We get

$$
\begin{aligned}
&D_m(y_{\gamma,c}d^T\nabla p_{\gamma,c})\\
&\quad = D_m(y_{\gamma,c})d^T\nabla p_{\gamma,c}+y_{\gamma,c}\cdot\sum_i(D_m(d_i)\partial_i p_{\gamma,c}+d_i D_m(\partial_i p_{\gamma,c}))\\
&\quad = D_m(y_{\gamma,c})d^T\nabla p_{\gamma,c}\\
&\qquad +\sum_i\Big(y_{\gamma,c}(\nabla d_i^T V)\partial_i p_{\gamma,c}+y_{\gamma,c}d_i\partial_i D_m(p_{\gamma,c})-\sum_k(y_{\gamma,c}d_i\partial_k p_{\gamma,c}\partial_i V_k)\Big)\\
&\quad = D_m(y_{\gamma,c})d^T\nabla p_{\gamma,c}+y_{\gamma,c}(\nabla d^T V)^T\nabla p_{\gamma,c}+y_{\gamma,c}d^T\nabla D_m(p_{\gamma,c})-y_{\gamma,c}d^T(\nabla V\nabla p),
\end{aligned}
$$

where we again use shape independence of the coefficient functions of $a(\cdot,\cdot)$. For the term of order zero we apply the product rule for material derivatives and shape independence of coefficient functions:

$$
D_m(by_{\gamma,c}p_{\gamma,c})=(\nabla b^T V)y_{\gamma,c}p_{\gamma,c}+bD_m(y_{\gamma,c})p_{\gamma,c}+by_{\gamma,c}D_m(p_{\gamma,c}).
$$

Combining these formulas, plugging them into (30), and collecting all material derivatives of $y_{\gamma,c}$ and $p_{\gamma,c}$ result in the shape derivative of the bilinear form $a(\cdot,\cdot)$:

$$
\begin{aligned}
&D\Big(a(y_{\gamma,c},p_{\gamma,c})\Big)[V]\\
&\quad = a(D_m(y_{\gamma,c}),p_{\gamma,c})+a(y_{\gamma,c},D_m(p_{\gamma,c}))\\
&\qquad +\int_\Omega \nabla y_{\gamma,c}^T(\nabla M^T V-\nabla V^T N-M\nabla V)\nabla p_{\gamma,c}\\
&\qquad + y_{\gamma,c}\cdot((\nabla d^T V)^T\nabla p_{\gamma,c}-d^T(\nabla V\nabla p_{\gamma,c}))\\
&\qquad + p_{\gamma,c}\cdot((\nabla d^T V)^T\nabla y_{\gamma,c}-d^T(\nabla V\nabla y_{\gamma,c}))+(\nabla b^T V)y_{\gamma,c}p_{\gamma,c}\\
&\qquad +\operatorname{div}(V)\left(\sum_{i,j}a_{i,j}\partial_i y_{\gamma,c}\partial_j p_{\gamma,c}+\sum_i d_i(\partial_i y_{\gamma,c}p_{\gamma,c}+y_{\gamma,c}\partial_i p_{\gamma,c})+by_{\gamma,c}p_{\gamma,c}\right)dx.
\end{aligned}
$$

The shape derivative of the term including $\max_\gamma$ is calculated by chain rule, which is applicable since we assume sufficient smoothness of $\max_\gamma$:

$$
\begin{aligned}
&D\Big((\max\nolimits_\gamma(\bar\lambda+c\cdot(y_{\gamma,c}-\varphi)),p_{\gamma,c})_{L^2(\Omega)}\Big)[V]\\
&\quad =\int_\Omega D_m\Big((\max\nolimits_\gamma(\bar\lambda+c\cdot(y_{\gamma,c}-\varphi)),p_{\gamma,c})_{L^2(\Omega)}\Big)\\
&\qquad +\operatorname{div}(V)\max\nolimits_\gamma(\bar\lambda+c\cdot(y_{\gamma,c}-\varphi))p_{\gamma,c}dx,
\end{aligned}
$$

$$
\begin{aligned}
D_m&\Big((\max_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)), p_{\gamma,c})_{L^2(\Omega)}\Big)\\
&= \big(\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) D_m(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)), p_{\gamma,c}\big)_{L^2(\Omega)}\\
&\quad + \big(\max_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)), D_m(p_{\gamma,c})\big)_{L^2(\Omega)}\\
&= -c\cdot\big(\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi))\nabla\varphi^T V, p_{\gamma,c}\big)_{L^2(\Omega)}\\
&\quad + \big(c\cdot \operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi))p_{\gamma,c}, D_m(y_{\gamma,c})\big)_{L^2(\Omega)}\\
&\quad + \big(\max_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)), D_m(p_{\gamma,c})\big)_{L^2(\Omega)},
\end{aligned}
$$

due to $D_m(\bar\lambda) = 0$ and $D_m(\varphi) = \nabla\varphi^T V$, as φ is invariant under perturbation of the domain by problem definition. The shape derivative of the last term in the Lagrangian (50) is given by a simple product rule

$$
D\big((f, p_{\gamma,c})_{L^2(\Omega)}\big)[V] = (D_m(f), p_{\gamma,c})_{L^2(\Omega)} + \big(f, D_m(p_{\gamma,c})\big)_{L^2(\Omega)}.
$$

We now use the assumptions $D_m(y_{\gamma,c}), D_m(p_{\gamma,c}) \in H_0^1(\Omega)$. If we rearrange the terms with $D_m(y_{\gamma,c})$ and $D_m(p_{\gamma,c})$ acting as test functions and applying the saddle point conditions, which means that the state equation (11) and adjoint equation (12) are fulfilled, the terms consisting of $D_m(y_{\gamma,c})$ and $D_m(p_{\gamma,c})$ cancel. By adding all terms of (50), the shape derivative $D\mathcal{J}_{\gamma,c}(\Omega)[V]$ as in (29) is established. ∎

Remark 8. One can fulfill the assumptions of the averaged adjoint theorem and, thus, guarantee existence of the shape derivative without a computation of the material derivatives $D_m(y_{\gamma,c}), D_m(p_{\gamma,c})$. The assumption $D_m(y_{\gamma,c}), D_m(p_{\gamma,c}) \in H_0^1(\Omega)$ for all $\gamma, c > 0$ in Theorem 3.4 is only needed in order to calculate the shape derivative expression (29).

Remark 9. The assumption $D_m(y_{\gamma,c}) \in H_0^1(\Omega)$ in Theorem 3.4 is needed to apply the saddle point conditions and get the closed form of the shape derivative. For this, it is sufficient that $y_{\gamma,c} \in H_0^2(\Omega)$. For example, this regularity can be ensured by additionally assuming $a_{i,j} \in C^1(\bar\Omega)$, $d \equiv 0$, $b \equiv 0$ for the coefficients of the strongly elliptic bilinear form $a(\cdot,\cdot)$ together with $\varphi > 0$ and choosing $\bar\lambda \in L^\infty(\Omega)$. The latter two assumptions imply that the maximal monotone Nemetskii operator (15) is equal to 0 for $y_{\gamma,c} = 0$ and sufficiently large $\gamma, c > 0$. In combination with the former assumptions, [5, Theorem A.1] can be applied to get $y_{\gamma,c} \in H_0^2(\Omega)$ for all sufficiently large $\gamma, c > 0$.

Remark 10. The assumption $D_m(p_{\gamma,c}) \in H_0^1(\Omega)$ in Theorem 3.4 can be fullfilled, e.g., by assuming additional regularity $a_{i,j} \in C^1(\bar\Omega)$ of the leading coefficients of the bilinear form $a(.,.)$ and C^2-regularity of the boundary $\partial\Omega$. This together with the fact that $c\cdot\operatorname{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) \in L^\infty(\Omega)$ acts as part of the coefficient function of the zero order terms permits application of a regularity theorem for linear elliptic problems (cf. [13, Theorem 4, p. 317]) giving $p_{\gamma,c} \in H^2(\Omega)$. This in turn guarantees $D_m(p_{\gamma,c}) \in H_0^1(\Omega)$.

Next, we formulate the second main theorem of this section, which states the convergence of the shape derivatives of the fully regularized problem.

THEOREM 3.5. *Assume the setting of the shape optimization problem formulated in* (2). *Let the assumptions of Theorem* 3.3 *hold and* $\varphi \in H^2(\Omega)$. *Moreover, let* $M := (a_{i,j})_{i,j=1,2}$ *be the matrix of coefficient functions to the leading order terms in* (4). *Then, for all* $V \in H_0^1(\Omega, \mathbb{R}^n)$, *the shape derivatives* $D\mathcal{J}_{\gamma,c}(\Omega)[V]$ *in* (29) *converge to* $D\mathcal{J}(\Omega)[V]$ *for* $\gamma, c \to \infty$, *where*

$$
\begin{aligned}
&D\mathcal{J}(\Omega)[V] \\
&:= \int_\Omega - (y-\bar{y})\nabla\bar{y}^T V - \nabla y^T(\nabla V^T M - \nabla M\cdot V + M^T\nabla V)\nabla p \\
&\qquad + y\cdot((\nabla d^T V)^T\nabla p - d^T(\nabla V\nabla p)) + p\cdot((\nabla d^T V)^T\nabla y - d^T(\nabla V\nabla y)) \\
&\qquad + (\nabla b^T V)yp - \nabla f^T V p \\
&\qquad + \operatorname{div}(V)\Big(\frac{1}{2}(y_{\gamma,c}-\bar{y})^2 + \sum_{i,j} a_{i,j}\partial_i y\partial_j p \\
&\qquad + \sum_i d_i(\partial_i yp + y\partial_i p) + byp - fp\Big)\,dx + \int_A (\varphi-\bar{y})\nabla\varphi^T V\,dx.
\end{aligned} \tag{32}
$$

Proof. We see that (29) already resembles (32) except for the two terms

$$T_1(V) := -c\cdot\int_\Omega \operatorname{sign}_\gamma(\bar{\lambda} + c\cdot(y_{\gamma,c}-\varphi))\cdot\nabla\varphi^T V\cdot p_{\gamma,c}\,dx, \tag{33}$$

$$T_2(V) := \int_\Omega \operatorname{div}(V)\cdot\max\nolimits_\gamma\big(\bar{\lambda} + c\cdot(y_{\gamma,c}-\varphi)\big)\cdot p_{\gamma,c}\,dx. \tag{34}$$

We proceed in two steps. First, we show convergence for T_1 and T_2 as restricted operators on $C_0^\infty(\Omega,\mathbb{R}^n)$. Second, we show that the limiting operators can be continuously extended to $H_0^1(\Omega,\mathbb{R}^n)$.

Let $V\in C_0^\infty(\Omega,\mathbb{R}^n)$. By this, we have $\operatorname{div}(V)\cdot p_{\gamma,c}, \nabla\varphi^T V \in H_0^1(\Omega)$ for all $\gamma, c>0$, which enables us to use these functions as test functions for the state and adjoint equations. This leads to

$$
\begin{aligned}
T_1(V) &= -c\cdot\int_\Omega \operatorname{sign}_\gamma(\bar{\lambda} + c\cdot(y_{\gamma,c}-\varphi))\cdot\nabla\varphi^T V\cdot p_{\gamma,c}\,dx \\
&= a(p_{\gamma,c},\nabla\varphi^T V) + (y_{\gamma,c}-\bar{y},\nabla\varphi^T V)_{L^2(\Omega)} \\
&\to a(p,\nabla\varphi^T V) + (y-\bar{y},\nabla\varphi^T V)_{L^2(\Omega)} =: \tilde{T}_1(V) \quad \text{for } \gamma,c\to\infty
\end{aligned}
$$

and

$$
\begin{aligned}
T_2(V) &= \int_\Omega \operatorname{div}(V)\cdot\max\nolimits_\gamma\big(\bar{\lambda} + c\cdot(y_{\gamma,c}-\varphi)\big)\cdot p_{\gamma,c}\,dx \\
&= -a(y_{\gamma,c},p_{\gamma,c}\cdot\operatorname{div}(V)) + (f,p_{\gamma,c}\cdot\operatorname{div}(V))_{L^2(\Omega)} \\
&\to -a(y,p\cdot\operatorname{div}(V)) + (f,p\cdot\operatorname{div}(V))_{L^2(\Omega)} =: \tilde{T}_2(V) \quad \text{for } \gamma,c\to\infty
\end{aligned}
$$

due to Theorem 3.3, Proposition 3.1, and our assumption $\varphi\in H^2(\Omega)$.

Next, we lift the convergence from $V\in C_0^\infty(\Omega,\mathbb{R}^n)$ to $H_0^1(\Omega,\mathbb{R}^n)$ by continuous extension. Since $C_0^\infty(\Omega,\mathbb{R}^n)$ is a dense subspace of $H_0^1(\Omega,\mathbb{R}^n)$ and the latter being the completion of the former by the $\|\cdot\|_{H_0^1(\Omega,\mathbb{R}^n)}$ norm, it is sufficient to show that the limits of $T_1(V_n)$ and $T_2(V_n)$ form a Cauchy sequence for a given Cauchy sequence $(V_n)_{n\in\mathbb{N}}\subset C_0^\infty(\Omega,\mathbb{R}^n)$ under the $\|\cdot\|_{H_0^1(\Omega,\mathbb{R}^n)}$-norm. So let $(V_n)_{n\in\mathbb{N}}\subset C_0^\infty(\Omega,\mathbb{R}^n)$ with $\|V_n-V_m\|_{H_0^1(\Omega,\mathbb{R}^n)}\to 0$ for $m,n\to\infty$. For the limit of T_1 we have

$$
\begin{aligned}
&|\tilde{T}_1(V_n)-\tilde{T}_1(V_m)| \\
&\quad = |a(p,\nabla\varphi^T(V_n-V_m)) + (y-\bar{y},\nabla\varphi^T(V_n-V_m))_{L^2(\Omega)}| \\
&\quad \le \left(\sum_{i,j}\|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j\|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \int_\Omega \sum_{i,j} \partial_i p \partial_j (\nabla\varphi^T (V_n - V_m)) \\
&+ \sum_i (\partial_i p \nabla\varphi^T (V_n - V_m) + p\partial_i \nabla\varphi^T (V_n - V_m)) \\
&+ p\nabla\varphi^T (V_n - V_m) dx + \int_\Omega (y - \bar{y}) \nabla\varphi^T (V_n - V_m) dx \\
&= \left(\sum_{i,j} \|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j \|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)} \right) \\
&\quad \cdot \int_\Omega \sum_{i,j} \left(\partial_i p \big((\partial_j \nabla\varphi)^T (V_n - V_m) + \nabla\varphi^T \partial_j (V_n - V_m) \big) \right) \\
&\quad + p\nabla\varphi^T (V_n - V_m) dx + \int_\Omega (y - \bar{y}) \nabla\varphi^T (V_n - V_m) dx \\
&\leq \left(\sum_{i,j} \|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j \|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)} \right) \cdot \|p\|_{H_0^1(\Omega)} \\
&\quad \cdot C \cdot \left(\sum_{i,j} \big(\|(\partial_j \nabla\varphi)^T (V_n - V_m)\|_{L^1(\Omega)} + \|\nabla\varphi^T \partial_j (V_n - V_m)\|_{L^1(\Omega)} \big) \right. \\
&\qquad \left. + \|\nabla\varphi^T (V_n - V_m)\|_{L^1(\Omega)} \right) \\
&\quad + C \cdot \|\varphi\|_{H^1(\Omega)} \cdot \|y - \bar{y}\|_{L^2(\Omega)} \cdot \|V_n - V_m\|_{H_0^1(\Omega,\mathbb{R}^n)} \\
&\leq C \cdot \left(\sum_{i,j} \|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j \|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)} \right) \cdot \|p\|_{H_0^1(\Omega)} \\
&\qquad \cdot \|\varphi\|_{H^1(\Omega)} \cdot 9 \cdot \|V_n - V_m\|_{H_0^1(\Omega,\mathbb{R}^n)} \\
&\quad + C \cdot \|\varphi\|_{H^1(\Omega)} \cdot \|y - \bar{y}\|_{L^2(\Omega)} \cdot \|V_n - V_m\|_{H_0^1(\Omega,\mathbb{R}^n)} \to 0 \quad \text{for } m, n \to \infty.
\end{aligned}
$$

Here, we use integration by parts, Gauss's theorem, $p_{|\partial\Omega} = 0$, $(V_n - V_m)_{|\partial\Omega} = 0$, and $L^2(\Omega) \hookrightarrow L^1(\Omega)$ with constant $C > 0$ as in (55). Thus, $(\tilde{T}_1(V_n))_{n\in\mathbb{N}}$ forms a Cauchy sequence and, therefore, gives a value for the continuous extension of $\tilde{T}_1$ for the limit of V_n in $H_0^1(\Omega, \mathbb{R}^n)$. For T_2 we use the same techniques, leading to

$$
\begin{aligned}
&|\tilde{T}_2(V_n) - \tilde{T}_2(V_m)| \\
&\quad = |-a(y, p \cdot \operatorname{div}(V_n - V_m)) + (f, p \cdot \operatorname{div}(V_n - V_m))_{L^2(\Omega)}| \\
&\quad \leq \left(\sum_{i,j} \|a_{i,j}\|_{L^\infty(\Omega)} + \sum_j \|d_j\|_{L^\infty(\Omega)} + \|b\|_{L^\infty(\Omega)} \right) \\
&\qquad \cdot 13 \cdot C \cdot \|y\|_{H_0^1(\Omega)} \cdot \|p\|_{H_0^1(\Omega)} \cdot \|V_n - V_m\|_{H_0^1(\Omega,\mathbb{R}^n)} \\
&\qquad + C \cdot \|f\|_{L^2(\Omega)} \cdot \|p\|_{H_0^1(\Omega)} \cdot \|V_n - V_m\|_{H_0^1(\Omega,\mathbb{R}^n)} \to 0 \quad \text{for } m, n \to \infty.
\end{aligned}
$$

With these convergences T_1, T_2 converge to the continuously extended limit objects, which we from now on denote by T_1, T_2, for all $V \in H_0^1(\Omega, \mathbb{R}^n)$. Next, we simplify the sum of these two limiting objects. Let $V \in C_0^\infty(\Omega, \mathbb{R}^n)$. Then

$$
\begin{aligned}
&T_1(V) + T_2(V)\\
&\quad = a(p, \nabla\varphi^T V) + (y - \bar{y}, \nabla\varphi^T V)_{L^2(\Omega)}\\
&\qquad - a(y, p \cdot \operatorname{div}(V)) + (f, p \cdot \operatorname{div}(V))_{L^2(\Omega)}\\
&\quad = a_{\Omega\setminus A}(p, \nabla\varphi^T V) + (y - \bar{y}, \nabla\varphi^T V)_{L^2(\Omega\setminus A)}\\
&\qquad + a_A(p, \nabla\varphi^T V) + (y - \bar{y}, \nabla\varphi^T V)_{L^2(A)} + (\lambda, p \cdot \operatorname{div}(V))_{L^2(\Omega)}\\
&\quad = (\varphi - \bar{y}, \nabla\varphi^T V)_{L^2(A)},
\end{aligned}
\tag{35}
$$

where we use the definition of p, complementary slackness of $\lambda \in L^2(\Omega)$, test function properties of $\nabla\varphi^T V$ and $p \cdot \operatorname{div}(V)$, and the state and adjoint equations. We apply again a continuity argument to gain this identity for all $V \in H_0^1(\Omega, \mathbb{R}^n)$. We see that the limit object in (35) is exactly the missing term in the limit of the shape derivatives $D\mathcal{J}_{\gamma,c}(\Omega)[V]$ (cf. (32)). ∎

Remark 11. Theorems 3.4 and 3.5 are also valid when $f \in H^1(\Omega)$ or $\varphi \in H^2(\Omega)$ depend explicitly on the shape Ω with shape derivatives $f', \varphi' \in H_0^1(\Omega)$. Then the shape derivatives need to be modified accordingly by replacing terms including $\nabla f^T V$ and $\nabla\varphi^T V$ by $\nabla f^T V + f'$ and $\nabla\varphi^T V + \varphi'$. Further, Theorems 3.4 and 3.5 remain valid for piecewise constant $f \in L^\infty(\Omega)$ depending on the shape Ω by adjusting the proofs applying integral splitting techniques as found in [47, Remark 4.21, Theorem 4.23].

Remark 12. It is common knowledge that by pushing the obstacle φ to infinity, i.e., $\varphi(x) \uparrow \infty$ for all $x \in \Omega$, the state equation representing the VI (5) becomes a regular elliptic PDE in weak formulation

$$a(y, v) = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega)$$

due to (6). This means that we encounter shape optimization problems with elliptic PDE constraints. Formula (32) remains valid by applying $A = \varnothing$, giving a shape derivative for a general elliptic problem.

Remark 13. The limiting object (32) is in general not the shape derivative of the unregularized problem. It can be regarded as part of the limit system arising during convergence of the optimality systems of the fully regularized problem. Finding a framework in shape optimization to describe the type of this limiting object will be part of further research and is beyond the scope of this article. For readers interested in a treatise on limiting systems of optimality systems in nonsmooth optimal control we recommend [8].

Remark 14. The limiting objects of the convergence results for adjoint variables (cf. Theorem 3.3) and shape derivatives (cf. Theorem 3.4) can be put into relation by conditions resembling C-stationarity, e.g., as found in [19, Definition 4.1.].

Using our terminology, it is necessary for C-stationarity conditions to hold that a $\xi \in H^{-1}(\Omega)$ exists such that the adjoint equation can be formulated in the form

$$a(p, v) + \langle \xi, v \rangle = -((y - \bar{y}), v)_{L^2(\Omega)}. \tag{36}$$

We can define such a $\xi \in H^{-1}(\Omega)$ by emulating the definition of p in (19), including enforcement of the Dirichlet condition $p = 0$ on ∂A with Nitsche's method using boundary terms (cf. [24]). The state equation, corresponding complementarity conditions, and the design equation, which in our setting can be viewed as the shape derivative identity (29), hold in analogy to the cited definition of C-stationarity. The remaining conditions

$$\langle \xi, p\rangle \geq 0 \quad \text{and} \quad p = 0 \ \ a.e. \ \text{in } \{\xi > 0\}, \tag{37}$$

by the definitions of ξ and p, are satisfied as well. It is worth mentioning that—to the knowledge of the authors—no types of C-stationarity-like conditions for optimality of VI constrained shape optimization problems have been investigated or defined before. By defining C-stationarity in this context, as outlined above, we can sum up the theorems by stating that the solutions of the regularized equations converge to a C-stationary system.

4. Algorithmic aspects and numerical investigations. In this section, we put the theoretical treatise highlighted in the previous section into numerical practice on domains in $\mathbb{R}^2$. We employ a steepest descent algorithm with backtracking linesearch in order to perform the optimization procedures with various regularized as well as unregularized versions of the specialized VI (see (39)). Also, we propose a way to incorporate the unregularized approach in an algorithm and compare it to the different regularizations.

For convenience, we specialize the more general constraint (5) to a Laplacian version:

$$\min_{\Gamma_{\text{int}}} \frac{1}{2}\int_\Omega |y-\bar{y}|^2 \ dx + \nu \int_{\Gamma_{\text{int}}} 1 \ ds \tag{38}$$

$$\begin{aligned} \text{s.t.} \quad \int_\Omega \nabla y^T \nabla v \ dx + \langle \lambda, v\rangle &= \int_\Omega fv \ dx \quad \forall v \in H_0^1(\Omega),\\ \lambda &\geq 0 \quad \text{in } \Omega,\\ y &\leq \varphi \quad \text{in } \Omega,\\ \lambda(y-\varphi) &= 0 \quad \text{in } \Omega. \end{aligned} \tag{39}$$

We use $\nu = 10^{-5}$ for all computations in this section. As the right-hand side of the state equation in (38) we choose the following piecewise constant function $f \in L^\infty(\Omega)$ defined by

$$f(x) = \begin{cases} -10 & \text{for } x \in \Omega_{\text{out}},\\ 100 & \text{for } x \in \Omega_{\text{in}}. \end{cases} \tag{40}$$

For calculations of the smoothed state and adjoint we have to specify $\max_\gamma$ satisfying Assumption 1. For demonstration purposes, we choose a similar smoothing procedure as in [23, section 2]:

$$\max\nolimits_\gamma(x) = \begin{cases} \max(0,x) & \text{for } x \in \mathbb{R}\setminus[-\frac{1}{\gamma},\frac{1}{\gamma}],\\ \frac{\gamma}{4}x^2 + \frac{1}{2}x + \frac{1}{4\gamma} & \text{else.} \end{cases} \tag{41}$$

A different, more regular smoothing is, e.g., given in [41, (1.10)]. Both smoothing techniques mentioned satisfy Assumption 1. For the sake of completeness, we also give the first derivative formula

$$\text{sign}_\gamma(x) = \begin{cases} 0 & \text{for } x \in (-\infty, -\frac{1}{\gamma}),\\ \frac{\gamma}{2}x + \frac{1}{2} & \text{for } x \in [-\frac{1}{\gamma},\frac{1}{\gamma}],\\ 1 & \text{for } x \in (\frac{1}{\gamma},\infty). \end{cases} \tag{42}$$

In this setting, the shape derivative (32) simplifies to

$$
\begin{aligned}
&D\mathcal{J}(\Omega)[V] \\
&= \int_\Omega -(y-\bar{y})\nabla\bar{y}^T V - \nabla y^T(\nabla V^T + \nabla V)\nabla p \\
&\qquad + \operatorname{div}(V)\left(\frac{1}{2}(y-\bar{y})^2 + \nabla y^T\nabla p - fp\right)\,dx + \int_A (\varphi-\bar{y})\nabla\varphi^T V\,dx
\end{aligned} \tag{43}
$$

and analogously the shape derivative for the fully regularized equation in (29). Notice that the shape derivative of the perimeter regularization is also included in our computations (cf. Remark 7).

In the following numerical experiments, we consider two different obstacles:

$$\varphi_1(x) = 0.5 \quad \text{and} \quad \varphi_2(x) = 5e^{-x_1-1}. \tag{44}$$

The calculations are performed with Python using the finite element package FEniCS. For detailed informations on FEniCS, we refer to [2] and [31]. As an initial shape we choose a centered circle with radius 0.15, illustrated in Figure 7. The computational grid of the initial shape, which is embedded in the hold-all-domain $(0,1)^2 \subset \mathbb{R}^2$, consists of 2184 vertices with 4206 cells, having a maximum cell diameter of 0.0359 and a minimum cell diameter of 0.018. The algorithm employed for the shape optimization is summarized in Algorithm 2. In the following, we describe the algorithm and the chosen parameters in detail.

The target data $\bar{y} \in L^2(\Omega)$ is computed by using the mesh of the target interface to calculate a corresponding state solution of (39) by the semismooth Newton method proposed in [23]. These are visualized in Figure 2 for both obstacles φ_1 and φ_2. We apply the same method for calculating state variables y in the unregularized optimization approach.

For the regularized and smoothed states $y_{\gamma,c}$ and y_c we use a Newton and a semismooth Newton method provided by the FEniCS package in order to solve the linear systems assembled by using first order polynomials on the computational grids.

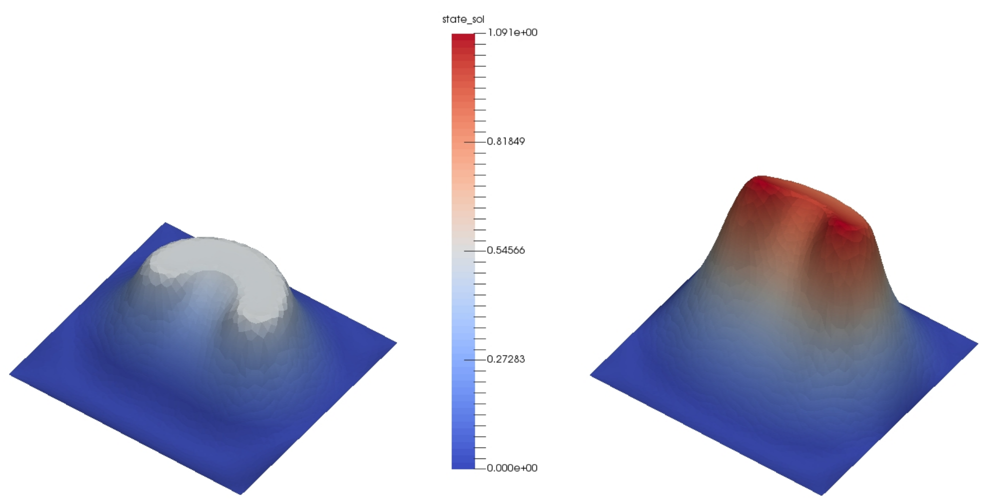

Fig. 2. *Solutions $\bar{y}$ to the VI in the target shape. On the left: $\varphi_1 = 0.5$. On the right: $\varphi_2 = 5e^{-x_1-1}$.*

All state calculations in our routines are performed with a stopping criterion of $\varepsilon_{\text{state}} = 3.e-4$ for the error norms. In light of Remark 3 we choose $\bar{\lambda} = \max\{0, f + \Delta\varphi\}$, which is possible due to sufficient regularity of φ_1 and φ_2.

To ensure assumptions of Theorems 3.3, 3.4, and 3.5, it is necessary to fulfill

$$\|\text{sign}_\gamma(\bar{\lambda} + c \cdot (y_{\gamma,c} - \varphi)) - \text{sign}(\bar{\lambda} + c \cdot (y_c - \varphi))\|_{L^1(\Omega)} \to 0 \quad \text{for } \gamma \to \infty. \tag{45}$$

We calculate the corresponding norm using various $c > 0$ and both φ_1 and φ_2 on refined meshes having 212, 642 vertices, 423, 682 cells, and maximum and minimum cell diameter of 0.0038 and 0.0015, respectively. An example convergence plot can be found in Figure 3. We want to point out that as $\gamma \to \infty$, the norm in (45) converges to an $\varepsilon > 0$ which is close to 0. This is due to numerical errors resulting from the state equation, since their solution determines the active set, which is needed to calculate the values of sign and sign_γ. The functions, whose L^1-norms are of interest, are illustrated in Figure 4 on a refined mesh. Furthermore, we observe that these functions and hence the errors go to 0 for ever finer grid widths and more precisely

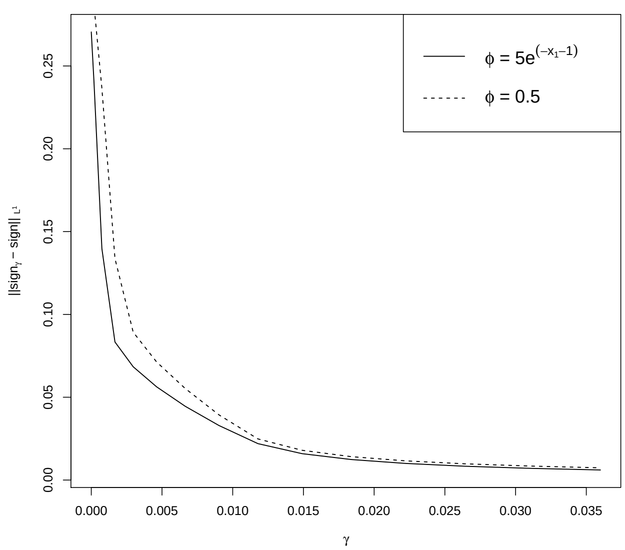

FIG. 3. *Convergence plots for* $\|\text{sign}_\gamma(\bar{\lambda}+c\cdot(y_{\gamma,c}-\varphi))-\text{sign}(\bar{\lambda}+c\cdot(y_c-\varphi))\|_{L^1(\Omega)}$ *as a function of* γ.

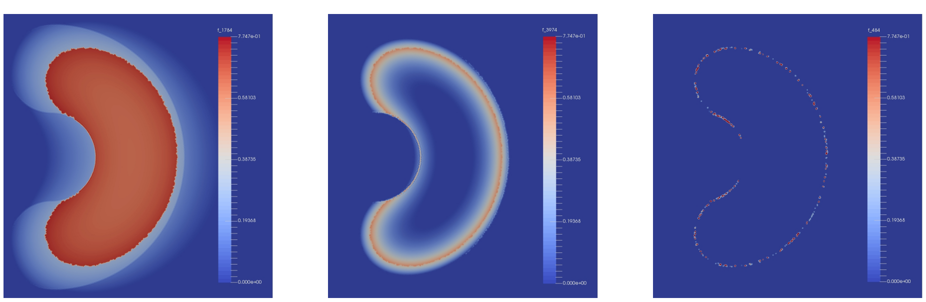

FIG. 4. *Graphs of* $\text{sign}_\gamma\big(\bar{\lambda}+c\cdot(y_{\gamma,c}-\varphi)\big)-\text{sign}\big(\bar{\lambda}+c\cdot(y_c-\varphi)\big)$ *as functions of* $x \in \Omega$ *calculated on the refined target mesh with* $c = 1000$. *From left to right:* $\gamma = 0.00075$, $\gamma = 0.009$, *and* $\gamma = 10$.

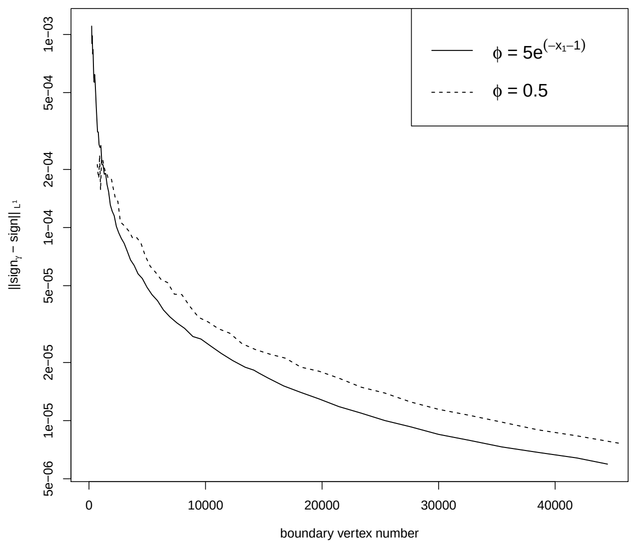

FIG. 5. *Graphs of* $\|\mathrm{sign}_\gamma(\bar\lambda + c\cdot(y_{\gamma,c}-\varphi)) - \mathrm{sign}(\bar\lambda + c\cdot(y_c-\varphi))\|_{L^1(\Omega)}$ *as functions of the number of mesh vertices at the boundary of* A *calculated on refined meshes with* $c = 10^5$, $\gamma = 10^8$.

calculated states $y_{\gamma,c}, y$. This is supported by a study successively evaluating the mentioned L^1-norms in the circular start shape, as depicted in Figure 7, on meshes generated by adaptive refinement at the boundary of the active set for large $\gamma, c > 0$. The results are seen in Figure 5.

The adjoints $p_{\gamma,c}$ and p_c are calculated by solving (12) and (18) with first order elements by using the FEniCS standard linear algebra back end solver PETSc.

Calculating the limit p of the adjoints $p_{\gamma,c}$ as in (19) and (20) is performed in several steps. First, a linear system corresponding to

$$\begin{aligned} -\Delta p &= -(y-\bar y) && \text{in } \Omega, \\ p &= 0 && \text{on } \partial\Omega \end{aligned} \tag{46}$$

is assembled without incorporation of information from the active set A. Afterward, the vertex indices corresponding to the points in the active set $A = \{x \in \Omega \mid y - \varphi \geq 0\}$ are collected by checking the condition

$$y(x) - \varphi(x) \geq -\varepsilon_{\mathrm{adj}} \tag{47}$$

for some error bound $\varepsilon_{\mathrm{adj}} > 0$. The error bound $\varepsilon_{\mathrm{adj}}$ is incorporated since y is feasibly approximated by y_i with the semismooth Newton method from [23], i.e., $y_i \leq \varphi$ for all $i \in \mathbb{N}$. After this, the collected vertex indices are used to incorporate the Dirichlet boundary conditions $p = 0$ in A into the linear system corresponding to (46). To solve the resulting system, we use the same procedures as used to solve for $p_{\gamma,c}, p_c$, i.e., the standard PETSc back end conjugate gradient solver. An exemplary solution p of the unregularized adjoint equation is illustrated in Figure 6. We want to point out that the active set and consequently the zero level set resulting from the Dirichlet conditions can be observed in Figure 6.

To calculate gradients $U \in H_0^1(\Omega, \mathbb{R}^2)$ used in the steepest descent method for solving (38), we use a Steklov–Poincaré metric induced by the linear elasticity equation,

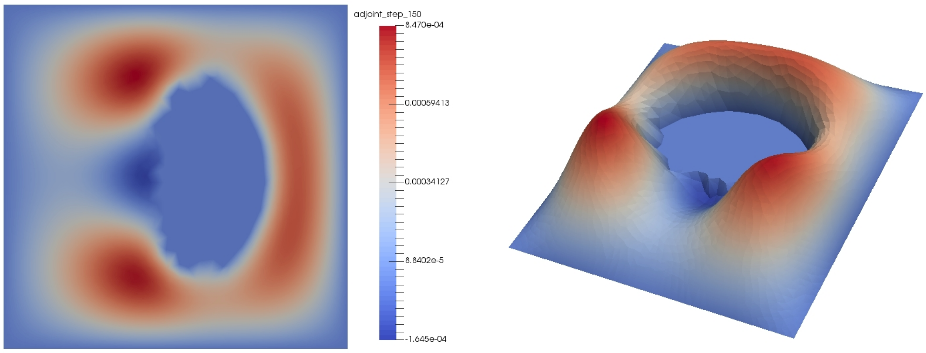

FIG. 6. *Solution of the adjoint p in step 150 for the unregularized equation (19) and $\varepsilon_{adj} = 10^{-9}$.*

Algorithm 1: Backtracking linesearch.

$\tilde{U} \leftarrow U_k$
while $\mathcal{J}\big(y_{\tilde{U}}, \mathcal{T}_{\tilde{U}}(\Omega_k)\big) > 0.995 \cdot \mathcal{J}\big(y_k, \Omega_k\big)$ **do**
 $\tilde{U} \leftarrow 0.5 \cdot \tilde{U}$
end
$\Omega_{k+1} \leftarrow \mathcal{T}_{\tilde{U}}(\Omega_k)$

as proposed in [42]. In particular, we assemble the shape derivatives given in Theorems 3.4 and 3.5 as the right-hand side of the linear elasticity equation

$$
\begin{aligned}
\int_\Omega \sigma(U) : \epsilon(V)\, dx = DJ(\Omega)[V] \qquad \forall V \in H_0^1(\Omega, \mathbb{R}^2),\\
\sigma(U) :&= \lambda_{\text{elas}} \operatorname{Tr}(U) I + 2\mu_{\text{elas}} \epsilon(U),\\
\epsilon(U) :&= \frac{1}{2}(\nabla U^T + \nabla U),\\
\epsilon(V) :&= \frac{1}{2}(\nabla V^T + \nabla V)
\end{aligned}
\tag{48}
$$

with the so-called Lamé parameters λ_{elas} and μ_{elas}. Here, we choose $\lambda_{\text{elas}} = 0$ and μ_{elas} as the solution of the Poisson problem

$$
\begin{aligned}
-\Delta \mu_{\text{elas}} &= 0 && \text{in } \Omega,\\
\mu_{\text{elas}} &= \mu_{\max} && \text{on } \Gamma,\\
\mu_{\text{elas}} &= \mu_{\min} && \text{on } \partial\Omega
\end{aligned}
\tag{49}
$$

for $\mu_{\max}, \mu_{\min} > 0$. As a physical interpretation, this enables us to control stiffness of the grid by choosing $\mu_{\max}$ and $\mu_{\min}$ in order to influence μ_{elas} acting as a coefficient function in the linear elasticity equation (48). Thus, larger values of $\mu_{\max}$ lead to more stiffness at the interface Γ and larger values of $\mu_{\min}$ to more stiffness at the boundary $\partial\Omega$ of the hold-all domain Ω. For our calculations, we choose $\mu_{\min} = 0$ and $\mu_{\max} = 25$ for φ_1 and $\mu_{\max} = 55$ for φ_2. It is important to notice that we set all right-hand-side values of (48) which do not have a neighboring vertex on the interface to 0. For a more detailed discussion of this we refer to [42].

To complete the description of our optimization we briefly explain the linesearch we will employ in our numerical calculations. We use a simple backtracking linesearch

Algorithm 2: Shape optimization for unregularized VI constraints with safeguard strategy.

Set $\Omega_0, \varphi, f, \bar{\lambda}, \bar{y}, \gamma, c$
while $\|D\mathcal{J}(\Omega_k)\| > \varepsilon_{\text{shape}}$ ***or*** $\|D\mathcal{J}_{\gamma,c}(\Omega_k)\| > \varepsilon_{\text{shape}}$ **do**
  Calculate state y_k with tolerance $\varepsilon_{\text{state}}$
  Calculate 'adjoint' p_k
    Assemble adjoint system (46) neglecting active set
    Collect vertex indices of active set by (47)
    Implement Dirichlet conditions of active set
    Solve modified adjoint linear system
  Calculate $\|D\mathcal{J}(\Omega_k)\|$ and shape gradient U_k
    Assemble gradient system (48)
    Set $D\mathcal{J}(\Omega_k)[V] = 0$ on all vertices without support at interface Γ_{int}
    Solve for gradient U_k
  Perform backtracking linesearch (algorithm 1) to get $\tilde{U}_k$
  if linesearch fails to give descent direction $\tilde{U}_k$ **or** $\|D\mathcal{J}(\Omega_k)\| \leq \varepsilon_{shape}$:
    Calculate fully regularized state $y_{\gamma,c}$
    Calculate fully regularized adjoint $p_{\gamma,c}$
    Calculate $\|D\mathcal{J}_{\gamma,c}(\Omega_k)\|$
    if $\|D\mathcal{J}_{\gamma,c}(\Omega_k)\| > \varepsilon_{shape}$
      Calculate fully regularized gradient $U_{\gamma,c}$ by $D\mathcal{J}_{\gamma,c}(\Omega_k)$
      Perform backtracking linesearch (algorithm 1) to get $\tilde{U}_k$
  $\Omega_{k+1} \leftarrow \mathcal{T}_{\tilde{U}_k}(\Omega_k)$
end

with sufficient descent criterion, where U_k denotes the shape derivative calculated at the corresponding interface in Ω_k in step number k, $\mathcal{T}_{\tilde{U}}(\Omega_k) := \{y \in \mathbb{R}^2 : y = x + \tilde{U}(x) \text{ for some } x \in \Omega_k\}$ the linearized vector transport by $\tilde{U}$, and $y_{\tilde{U}}$ the state solution in $\mathcal{T}_{\tilde{U}}(\Omega_k)$.

We summarize our approach in Algorithm 2 for the unregularized procedures, resulting exemplary shape iterates of this routine are shown in Figure 7. The regularized and smoothed procedures work analogously by modifying the state, adjoint, and shape derivative equations. The calculations of $p_{\gamma,c}, p_c$ are straightforward and do not need the additional steps outlined before and in Algorithm 2 for the unregularized p. In the design of Algorithm 2 a safeguarding technique is employed. This stems from the fact that the limit of shape derivatives $D\mathcal{J}$ from (32) is in general not the true shape derivative of the initial problem; see Remark 14. Hence an additional test of the convergence criterion for the fully regularized shape derivative $\mathcal{D}\mathcal{J}_{\gamma,c}$ is performed after the convergence by $D\mathcal{J}$. If no convergence is detected by $D\mathcal{J}_{\gamma,c}$, $D\mathcal{J}_{\gamma,c}$ will be used to calculate a further descent direction, as the latter is a true shape derivative by Theorem 3.4. Further, the safeguarding acts as a safety when the adjoint limit object p_k is flawed due to erroneous allocation of the active set A_k as discussed with (47) in the beginning of this section. The smoothed model is not prone to this effect and hence acts as a substitute model for further gradient calculations.

In our calculations the safeguard was never activated by not finding a descent direction during the linesearch procedure, indicating that the shape derivative limiting object $D\mathcal{J}$ is acting appropriately for a shape derivative substitute, making the safeguard for this purpose obsolete. Still, for coarse grids or imprecise calculations of

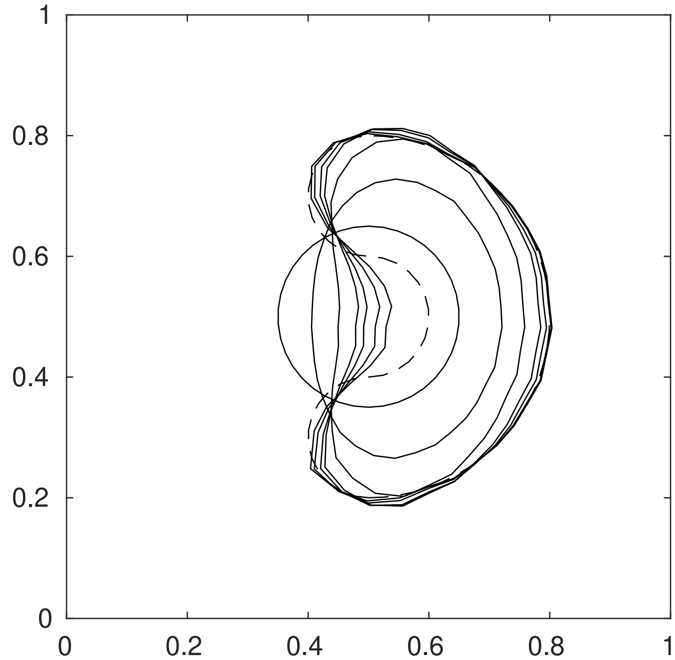

FIG. 7. *Interfaces of steps* $0, 50, 150, 320, 450, 750, 1200$ *of the unregularized optimization procedure using* $\varepsilon_{adj} = 10^{-9}$ *and* $\varphi_2 = 5e^{-x_1-1}$. *The target interface is represented with dotted lines; the start interface is the centered circle.*

the state y_k the safeguarding is activated at convergence, indicating a nonnegligible difference in $\|D\mathcal{J}_{\gamma,c}\|$ and $\|D\mathcal{J}\|$. This is only due to false allocation of the active set A_k, resulting in inaccurate p_k and $\|D\mathcal{J}(\Omega_k)\|$. For grids with maximum cell diameter 0.01 or less and error tolerance $\varepsilon_{\text{state}} < 1.e-7$ for the state calculation, the errors in active set allocation are sufficiently small for $\|D\mathcal{J}(\Omega_k)\| \approx \|D\mathcal{J}_{\gamma,c}(\Omega_k)\|$ and the safeguarding to not be activated at convergence in all our examples.

Our findings concerning convergence of the various shape optimization approaches, using the unregularized approach for various φ_{adj}, as well as regularized approaches with different parameters $\gamma, c > 0$, are displayed in Figure 8 for $\varphi_1 = 0.5$ and in Figure 9 for $\varphi_2 = 5e^{-x_1-1}$. Morphed shapes arising during the optimization procedure are plotted in Figure 7 for the unregularized approach using $\varepsilon_{\text{adj}} = 10^{-9}$. It can be seen in the plots that there are vanishing differences between approaches using fully regularized calculation with sufficiently large γ and c, regularized ones with large c, and the unregularized one. For smaller regularization parameters γ and c, the solved state and adjoint equations begin to differ from the original problem and, thus, slow down convergence, and for very small γ and c there is no convergence at all.

The convergence behavior of the unregularized method strongly depends on the selection of the active set. When the state solution y is not calculated with sufficient precision the numerical errors lead to misclassification of vertex indices. Hence wrong Dirichlet conditions are incorporated in the adjoint system, creating errors in the adjoint. This makes the gradient sensitive to error for smaller ε_{adj}, as can be seen by the slight roughness of the target graphs in Figures 8 and 9 for $\varepsilon_{\text{adj}} = 10^{-9}$ and $\varepsilon_{\text{adj}} = 0.01$. In order to compensate this, the condition for checking active set indices (47) can be relaxed by increasing ε_{adj}. This increases the likelihood of

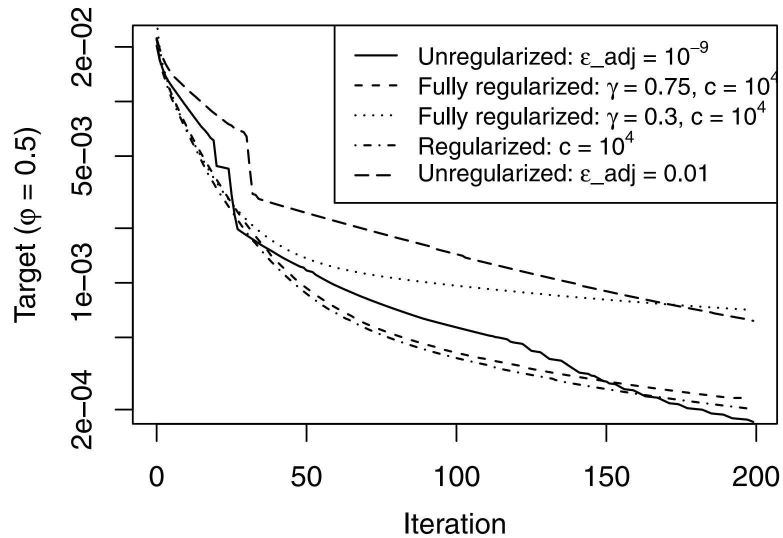

FIG. 8. *Convergence plot of target functional $\mathcal{J}$ values for different regularization and unregularized approaches for obstacle $\varphi_1 = 0.5$ using steepest descent.*

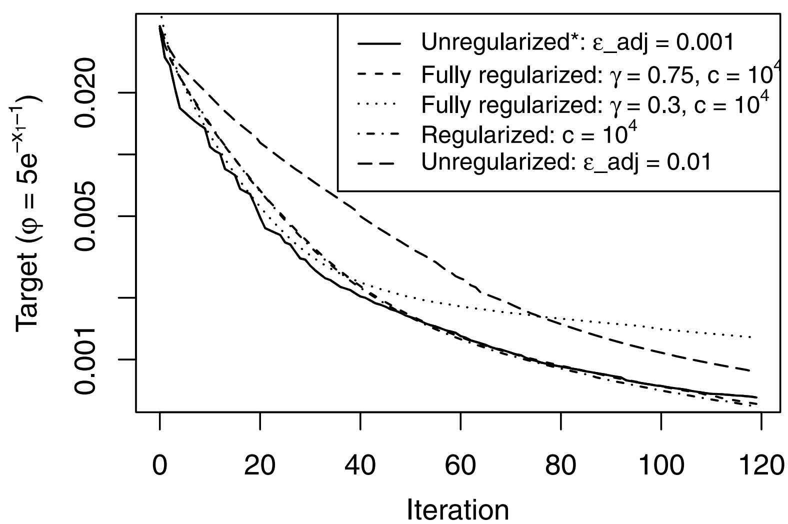

FIG. 9. *Convergence plot of target functional $\mathcal{J}$ values for different regularization and unregularized approaches for obstacle $\varphi_2 = 5e^{-x_1-1}$. Unregularized* uses a lower tolerance $\varepsilon_{state} = 0.00001$ for the state calculation. Notice that regularized and fully regularized approaches for $\gamma = 0.75, c = 10^4$ are almost indistinguishable.*

correctly classifying the true active indices, while also increasing the likelihood of misclassification of inactive indicies. Such a relaxation can lead to errors in the adjoint increasing with ε_{adj}, thus trading convergence speed for robustness, also visible in Figures 8 and 9. Of course, this gets less feasible for highly oscillatory obstacle φ and state y, as well as state solves with high tolerance $\varepsilon_{\text{state}}$.

In order to circumvent this, it is obviously sufficient to decrease error tolerance $\varepsilon_{\text{state}}$ of the state calculation. An exemplary result of this can be seen in Figure 9 under unregularized*, where we decreased the error tolerance to $\varepsilon_{\text{state}} = 4.e-5$. Nevertheless, an additional decrease of $\varepsilon_{\text{state}}$ comes with more computational cost, whereas with an increase of ε_{adj} the robustness is paid by loss of convergence speed.

It is worth mentioning that implementing the unregularized state and adjoint becomes especially numerically exploitable with higher resolution meshes and more strongly binding obstacles φ, i.e., larger active sets A. This is possible by sparse solvers due to the incorporation of Dirichlet conditions on the active set, as we have proposed, or by a fat boundary method as in [34]. This is especially advantageous for large systems resulting from fine resolution grids, as exploitability of sparsity and accuracy of our method increase at the same time.

So in contrast to the method proposed in [15], where performance slows down for more active obstacle φ, we do not notice unusual slowdown in performance with the methods proposed in this article and even offer the possibility to actually benefit numerically from more binding obstacle φ.

5. Conclusion. Shape optimization for VIs is more challenging than both elliptic shape optimization and optimal control for VIs. In this paper, we derive optimality conditions for shape optimization in the context of VIs in the flavor of optimal control problems. Regularized variants are studied and limiting conditions derived. This gives rise to highly efficient optimization algorithms. In the future general investigations of necessary optimality criteria for VI constrained shape optimization like C-stationarity are conceivable. Also large-scale multidimensional computational comparisons of the presented method in comparison to other state-of-the-art methods is of particular interest.

Appendix A. Proof of existence of shape derivatives in Theorem 3.4. What follows is a proof for the existence of shape derivatives $D\mathcal{J}_{\gamma,c}$ for all $\gamma, c > 0$ under the assumptions of Theorem 3.4.

Proof. For the proof of existence $D\mathcal{J}_{\gamma,c}$ for all $\gamma, c > 0$, we will employ the so-called averaged adjoint approach, as found in [26, Chapter 7, Theorem 7.1]. We roughly follow a proof found in [26, Chapter 7, Theorem 7.2], but have to give a proof ourselves, since the situation in this paper differs from the one mentioned previously, e.g., as we are not having a bounded Nemetskii operator as the nonlinearity in the semilinear state equation.

Let $\gamma, c > 0$. By definition, the Lagrangian function corresponding to our problem is given by

$$\begin{aligned}\mathcal{L}_{\gamma,c}(\Omega, y_{\gamma,c}, p_{\gamma,c}) = \frac{1}{2}\int_\Omega (y_{\gamma,c} - \bar{y})^2 dx + a(y_{\gamma,c}, p_{\gamma,c}) \\ + \left(\max\nolimits_\gamma\left(\bar{\lambda} + c \cdot (y_{\gamma,c} - \varphi)\right), p_{\gamma,c}\right)_{L^2(\Omega)} - (f, p_{\gamma,c})_{L^2(\Omega)}.\end{aligned} \tag{50}$$

We have to prove the assumptions (H0)–(H3) needed to apply the averaged adjoint theorem as found in [26, Chapter 7, Theorem 7.1] in order to guarantee shape differentiability. For convenience, we do not state all the lengthy assumptions here, and we refer the interested reader to [26, Chapter 7, Theorem 7.1]. Let us fix a deformation vector field V and denote the domains deformed by Ω_t for deformation parameters $t \in [0, \tau]$ and some $\tau > 0$ small enough, such that the corresponding deformation $\mathcal{T}_t$ is bijective. For the rest of the existence proof, we will drop γ, c as the subscripts of $y_{\gamma,c}$ and $p_{\gamma,c}$ for readability purposes, still knowing we are in the fully regularized situation. Define

$$G : [0, \tau] \times H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}, (t, y, p) \mapsto \mathcal{L}(\Omega_t, y \circ \mathcal{T}_t^{-1}, p \circ \mathcal{T}_t^{-1}) \tag{51}$$

for the deformed domain Ω_t resulting from application of the deformation $\mathcal{T}_t$ in direction V parametrized by $t \in [0, \tau]$.

The first assumption (H0) concerns well behavedness of (51). First notice that the function in (51) is both differentiable in the state y and adjoint p, which can be seen after applying the transformation theorem to (51). Thus for the set

$$X(t) := \left\{ \hat{y} \in H_0^1(\Omega) \mid \inf_{y \in H_0^1(\Omega)} \sup_{p \in H_0^1(\Omega)} G(t, y, p) \; = \sup_{p \in H_0^1(\Omega)} G(t, \hat{y}, p) \right\} \tag{52}$$

we have

$$X(t) = \{y^t\} \subset H_0^1(\Omega) \quad \forall t \in [0, \tau], \tag{53}$$

with $y^t = y_t \circ \mathcal{T}_t$ being the retraction of the unique solution $y_t \subset H_0^1(\Omega_t)$ of the fully regularized state equation (11) on the deformed domain Ω_t as by the use of the Minty–Browder theorem as portrayed in the proof of Proposition 3.1.

Further, the Lagrangian in direction V as a function of averaged states inserted

$$[0, 1] \to \mathbb{R}, \; s \mapsto G(t, sy^t + (1 - s)y^0, p) \tag{54}$$

is absolutely continuous in $s \in [0, 1]$. For this, we make use of the fact that

$$L^q(\Omega) \hookrightarrow L^p(\Omega) \quad \text{for } 1 \leq p < q \leq \infty \tag{55}$$

for bounded $\Omega \subset \mathbb{R}^n$ with a constant depending on p and q. So absolute continuity is satisfied, as we have existing derivatives of the integrand due to Assumption 1(i) and integrability of the integrand due to

$$\begin{aligned}
&\|\mathrm{max}_\gamma(\bar{\lambda} + c \cdot (z - \varphi))\|_{L^2(\Omega)} \\
&\quad \leq \|\mathrm{max}_\gamma(\bar{\lambda} + c \cdot (z - \varphi)) - \max(0, \bar{\lambda} + c \cdot (z - \varphi))\|_{L^2(\Omega)} \\
&\quad + \| \max(0, \bar{\lambda} + c \cdot (z - \varphi))\|_{L^2(\Omega)} \\
&\quad \leq \mathrm{vol}(\Omega)^{\frac{1}{2}} g(\gamma) + \|\bar{\lambda} + c \cdot (z - \varphi)\|_{L^2(\Omega)} \\
&\quad \leq \mathrm{vol}(\Omega)^{\frac{1}{2}} g(\gamma) + \|\bar{\lambda} + c \cdot z\|_{L^2(\Omega)} + C \cdot c \cdot \|\varphi\|_{H_0^1(\Omega)} \\
&\quad \leq \mathrm{vol}(\Omega)^{\frac{1}{2}} g(\gamma) + \|\bar{\lambda}\|_{L^2(\Omega)} + C \cdot c \cdot (\|\varphi\|_{H_0^1(\Omega)} + \|z\|_{H_0^1(\Omega)}) < \infty
\end{aligned} \tag{56}$$

for all $z \in L^2(\Omega)$ and all $z \in H_0^1(\Omega)$ as by Assumption 1(ii), Ω being bounded and $H_0^1(\Omega) \hookrightarrow L^4(\Omega) \hookrightarrow L^2(\Omega)$ with constant C by (22) and (55). Further, the directional derivative mapping

$$[0, 1] \to \mathbb{R}, \; s \mapsto \frac{\partial}{\partial y} G(t, sy^t + (1 - s)y^0, p; \tilde{p}) \tag{57}$$

is integrable for all $\tilde{p} \in H_0^1(\Omega)$, since

$$\begin{aligned}
&\left\| \frac{\partial}{\partial y} G(t, sy^t + (1 - s)y^0, p; \tilde{p}) \right\|_{L^1(0,1)} \\
&\quad = \int_0^1 |a_t(p, \tilde{p}) + c \cdot \Big(\mathrm{sign}_\gamma \big(\bar{\lambda} + c \cdot (sy^t + (1 - s)y^0 - \varphi)\big) p, \tilde{p} \Big)_{L^2(\Omega)} \\
&\qquad + (sy^t + (1 - s)y^0 - \bar{y}, \tilde{p})_{L^2(\Omega)}| \, ds
\end{aligned}$$

$$
\begin{aligned}
&\leq \int_0^1 B_t \cdot \|p\|_{H_0^1(\Omega)} \|\tilde{p}\|_{H_0^1(\Omega)} + c \cdot \|p\|_{H_0^1(\Omega)} \|\tilde{p}\|_{H_0^1(\Omega)} \\
&\quad + (s\|y^t\|_{H_0^1(\Omega)} + (1-s)\|y^0\|_{H_0^1(\Omega)} + \|\bar{y}\|_{L^2(\Omega)}) \|\tilde{p}\|_{H_0^1(\Omega)} \, ds \\
&= (B_t + c)\|p\|_{H_0^1(\Omega)} \|\tilde{p}\|_{H_0^1(\Omega)} + \Bigg(\frac{1}{2}\|y^t\|_{H_0^1(\Omega)} + \frac{1}{2}\|y^0\|_{H_0^1(\Omega)} \\
&\qquad\qquad + \|\bar{y}\|_{L^2(\Omega)} \Bigg) \|\tilde{p}\|_{H_0^1(\Omega)} < \infty
\end{aligned}
$$

by Hölder's inequality, Assumption 1(iii) and $a_t(.,.)$ being the bilinear form defined by retraction of $a(.,.)$ from Ω_t to Ω bound with constants $B_t > 0$. This, and the easy-to-verify fact that G is an affine linear function in p, gives us (H0). We remind the careful reader that the Jacobians created by retraction of Ω_t to Ω are to be implicitly included in scalar products and norms above for the calculations to be valid. We don't explicitly state these for readability.

Next we introduce the set of so-called averaged adjoints

(58)
$$
Y(t, y^t, y^0) := \left\{ q \in H_0^1(\Omega) \mid \int_0^1 \frac{\partial}{\partial y} G(t, sy^t + (1-s)y^0, q; \tilde{p}) \, ds = 0 \;\; \forall \tilde{p} \in H_0^1(\Omega) \right\}.
$$

We manipulate the averaged adjoint equation found in (58) by interchanging integrals

$$
\begin{aligned}
0 &= \int_0^1 \frac{\partial}{\partial y} G(t, sy^t + (1-s)y^0, q; \tilde{p}) \, ds \\
&= \int_0^1 a_t(q, \tilde{p}) + c \cdot \left(\operatorname{sign}_\gamma (\bar{\lambda} + c \cdot (sy^t + (1-s)y^0 - \varphi)) q, \tilde{p} \right)_{L^2(\Omega)} \\
&\qquad + \left(sy^t + (1-s)y^0 - \bar{y}, \tilde{p} \right)_{L^2(\Omega)} \, ds \\
&= a_t(q, \tilde{p}) + c \cdot \left(\left(\left(\int_0^1 \operatorname{sign}_\gamma (\bar{\lambda} + c \cdot (sy^t + (1-s)y^0 - \varphi) \right) ds) q, \tilde{p} \right)_{L^2(\Omega)} \\
&\qquad + \left(\frac{1}{2} y^t + \frac{1}{2} y^0 - \bar{y}, \tilde{p} \right)_{L^2(\Omega)},
\end{aligned} \tag{59}
$$

which is an elliptic PDE with an additional positive $L^\infty(\Omega)$ coefficient function term for the zeroth order terms, where we again omitted explicit statement of Jacobians. By Assumption 1(iii), the additional coefficient term for the zeroth order terms in the averaged adjoint equation (59) satisfies

$$
0 \leq \int_0^1 \operatorname{sign}_\gamma (\bar{\lambda}_t + c \cdot (sy^t + (1-s)y^0 - \varphi_t)) ds \leq 1 \qquad \forall t \in [0, \tau], \tag{60}
$$

which results in coercivity and boundedness of the corresponding bilinear form of the averaged adjoint equation. This lets us apply the Lax–Milgram lemma, resulting in existence of a unique solution for the averaged adjoint equation we will denote by $q^t \in H_0^1(\Omega)$ for all $t \in [0, \tau]$. Thus we have the identity $Y(t, y^t, y^0) = \{q^t\} \subset H_0^1(\Omega)$, which together with (53) ensures condition (H2).

We also notice that the derivatives of $\frac{\partial}{\partial t} G$ exist and can be explicitly calculated after application of the transformation theorem, giving us (H1).

To apply the averaged adjoint theorem from [25], it remains to address (H3), which is satisfied in our case by application of [45, Lemma 4.1], if for the unique solutions of the state and adjoint equation $y^0 \in X(0)$ and $q^0 \in Y(0, y^0, y^0)$ and a given sequence $(t_n)_{n\in\mathbb{N}} \subset [0,\tau]$ converging to zero, we can find a subsequence $(t_{n_k})_{k\in\mathbb{N}} \subseteq (t_n)_{n\in\mathbb{N}}$ with $q^{t_{n_k}} \in Y(t_{n_k}, y^t_{n_k}, y^0)$ such that

$$\lim_{\substack{k\to\infty \\ t\searrow 0}} \frac{\partial}{\partial t} G(t, y^0, q^{t_{n_k}}) = \frac{\partial}{\partial t} G(0, y^0, q^0). \tag{61}$$

We will mimic parts of the argumentation found in the proof of [45, Theorem 5.1] customized to our situation, which is slightly different from the one found in [45, Theorem 5.1] or [26, Theorem 7.2].

Consider the solutions $y^0 \in X(0)$ and $q^0 \in Y(0, y^0, y^0)$ and a sequence $(t_n)_{n\in\mathbb{N}} \subset [0,\tau]$ converging to zero.

First notice that by monotony of the Nemetskii operator (15) of the concerning semilinear state equation we have

$$(\max\nolimits_\gamma(\bar\lambda + c\cdot(z-\varphi)), z)_{L^2(\Omega)} \geq (\max\nolimits_\gamma(\bar\lambda - c\cdot\varphi), z)_{L^2(\Omega)} \qquad \forall z \in L^2(\Omega). \tag{62}$$

This in turn, together with the coercivity of the retracted bilinearform $a_t(.,.)$ with constant $K_t > 0$, (56), and choosing $y^t \in X(t) \subset H^1_0(\Omega)$ as a test function, gives us

$$\begin{aligned} 0 &\leq \|y^t\|^2_{H^1_0(\Omega)} \leq K_t \cdot a_t(y^t, y^t) \\ &= K_t \int_\Omega f\cdot y^t - \max\nolimits_\gamma(\bar\lambda + c\cdot(y^t-\varphi))\cdot y^t \, dx \\ &\leq K_t \int_\Omega f\cdot y^t - \max\nolimits_\gamma(\bar\lambda - c\cdot\varphi))\cdot y^t \, dx \\ &\leq K_t \cdot (\|f\|_{L^2(\Omega)} + \|\max\nolimits_\gamma(\bar\lambda - c\cdot\varphi))\|_{L^2(\Omega)}) \cdot \|y^t\|_{H^1_0(\Omega)} < \infty, \end{aligned} \tag{63}$$

again omitting Jacobians. Dividing by $\|y^t\|_{H^1_0(\Omega)}$, using the convergence $\mathcal{T}_t$ to the identity for $t \downarrow 0$, and taking a supremum we achieve

$$\|y^t\|_{H^1_0(\Omega)} \leq \sup_{t\in\{t_n\}} \Big(K_t \cdot (\|f\|_{L^2(\Omega)} + \|\max\nolimits_\gamma(\bar\lambda - c\cdot\varphi))\|_{L^2(\Omega)}) \Big) =: M < \infty, \tag{64}$$

bounding the norms by a constant $0 < M < \infty$ independent of $t \in [0,\tau]$. Recognize that the norms still implicitly depend on t, since Jacobians are to be included.

In the same line of argumentation we can confirm the boundedness of $\|q^t\|_{H^1_0(\Omega)}$. For this, we apply the first inequality of (60) to get

$$\begin{aligned} 0 &\leq \|q^t\|^2_{H^1_0(\Omega)} \leq K_t \cdot a_t(q^t, q^t) \\ &= -K_t \Bigg(c\cdot \left(\left(\int_0^1 \operatorname{sign}_\gamma(\bar\lambda + c\cdot(sy^t + (1-s)y^0 - \varphi))ds \right) \cdot q^t, q^t \right)_{L^2(\Omega)} \\ &\qquad + \left(\frac{1}{2} y^t + \frac{1}{2} y^0 - \bar y, q^t \right)_{L^2(\Omega)} \Bigg) \\ &\leq K_t (\|y^t\|_{H^1_0(\Omega)} + \|y^0\|_{H^1_0(\Omega)} + \|\bar y\|_{L^2(\Omega)}) \|q^t\|_{H^1_0(\Omega)}. \end{aligned}$$

By finally using (64) we arrive at

$$\|q^t\|_{H^1_0(\Omega)} \leq \sup_{t\in\{t_n\}} \Big(K_t \cdot (M + \|y^0\|_{H^1_0(\Omega)} + \|\bar y\|_{L^2(\Omega)}) \Big) < \infty. \tag{65}$$

As we have established bound (64), we can choose a subsequence $(t_{n_k})_{k\in\mathbb{N}} \subseteq (t_n)_{n\in\mathbb{N}}$ such that $y^{t_{n_k}} \rightharpoonup z$ weakly in $H_0^1(\Omega)$ for $k \to \infty$ and some $z \in H_0^1(\Omega)$.

Further, using the convergence of the retracted functions $\bar{\lambda}_t$ and φ_t in $L^2(\Omega)$ for $t \downarrow 0$, we can uniformly and independently of t_{n_k} bound

$$\begin{aligned}&\|\max_\gamma(\bar{\lambda} + c \cdot (y^{t_{n_k}} - \varphi))\|_{L^2(\Omega)}\\ &\quad \le \mathrm{vol}(\Omega)^{\frac{1}{2}} g(\gamma) + c \cdot M + \sup_{t\in\{t_{n_k}\}} \|\bar{\lambda} - c\cdot\varphi\|_{L^2(\Omega)} < \infty\end{aligned}$$

by using (56) and (64). By having boundedness of the coercive bilinear forms $a_t(.,.)$, (64), and smoothness in $\max_\gamma$ by Assumption 1(i) we are able to apply Lebesgue's dominated convergence theorem to the retracted state equations, giving us $y_{t_{n_k}} \rightharpoonup y^0$ weakly in $H_0^1(\Omega)$ due to the unique solution guaranteed by the Minty–Browder theorem.

Applying the same routine due to (65), we can choose a subsequence of $\{t_{n_k}\}$, which we will again call $\{t_{n_k}\}$ by abuse of notation, such that $q^{t_{n_k}} \rightharpoonup u$ weakly in $H_0^1(\Omega)$ for some $u \in H_0^1(\Omega)$. Then uniform boundedness (60), the previously established weak convergence $y_{t_{n_k}} \rightharpoonup y^0$, and (65) yield applicability of Lebesgue's theorem for inserted t_{n_k} in (59). For $k \to \infty$, the limit equation of (59) is the fully regularized adjoint equation (12), which has a unique solution by the Lax–Milgram lemma. Thus we have that $q^{t_{n_k}} \rightharpoonup q^0 = p_{\gamma,c} \in Y(0, y^0, y^0)$ weakly in $H_0^1(\Omega)$ with the previously established weak convergence of $q^{t_{n_k}}$ and continuity of sign_γ by Assumption 1(i).

Now we have found a subsequence $\{t_{n_k}\} \subseteq \{t_n\}$ such that $q^{t_{n_k}} \rightharpoonup q^0$ weakly in $H_0^1(\Omega)$. Using the transformation theorem, $G(t, y^0, q^{t_{n_k}})$ from (51) can be stated as an integral in Ω with integrands being differentiable in $t \in [0,\tau]$. The derivative $\frac{\partial}{\partial t} G(t, y^0, q^{t_{n_k}})$ is weakly continuous in its first and last arguments, hence the weak convergence $q^{t_{n_k}} \rightharpoonup q^0$ implies (61), which is condition (H3) from [26, Theorem 7.1]. All assumptions (H0)–(H3) for the averaged adjoint theorem [26, Theorem 7.1] are satisfied, finally guaranteeing existence of shape derivatives $D\mathcal{J}_{\gamma,c}$ for all $\gamma, c > 0$. □

Acknowledgments. The authors are indebted to Leonhard Frerick (Trier University) for many helpful comments and discussions about functional analytical aspects of convergence. Furthermore, the paper profited form remarks of Gerd Wachsmuth on an earlier version.

REFERENCES

[1] A. Capatina, *Variational Inequalities and Frictional Contact Problems*, Springer, New York, 2014.

[2] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells, *The FEniCS project version* 1.5, Arch. Numer. Softw., 3 (2015), https://doi.org/10.11588/ans.2015.100.20553.

[3] H. B. Ameur, M. Burger, and B. Hackl, *Level set methods for geometric inverse problems in linear elasticity*, Inverse Problems, 20 (2004), pp. 673–696.

[4] M. Berggren, *A unified discrete-continuous sensitivity analysis method for shape optimization*, in Applied and Numerical Partial Differential Equations, Comput. Methods Appl. Sci., Springer, New York, 2010, pp. 25–39.

[5] J. Bonnans and D. Tiba, *Pontryagin's principle in the control of semilinear elliptic variational inequalities*, Appl. Math. Optim., 23 (1991), pp. 299–312.

[6] H. Brézis, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, in Contributions to Nonlinear Functional Analysis, Academic Press, New York, 1971, pp. 101–156.

[7] H. Brézis and G. Stampacchia, *Sur la regularité de la solution d'inéquations elliptiques*, Bull. Soc. Math. France, 96 (1968), pp. 153–180.

[8] C. Christof, C. Clason, C. Meyer, and S. Walther, *Optimal control of a non-smooth semilinear elliptic equation*, Math. Control Relat. Fields, 8 (2018), pp. 247–276.

[9] G. S. D. Kinderlehrer, *An Introduction to Variational Inequalities and Their Applications*, Classics in Appl. Math. 31, SIAM, Philadelphia, 1980.

[10] M. Delfour and J.-P. Zolésio, *Shape sensitivity analysis via min max differentiability*, SIAM J. Control Optim., 26 (1988), pp. 834–862.

[11] M. Delfour and J.-P. Zolésio, *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*, 2nd ed., Adv. Design Control 22, SIAM, Philadelphia, 2001.

[12] Z. Denkowski and S. Migorski, *Optimal shape design for hemivariational inequalities*, Univ. Iagel. Acta Math., 36 (1998), pp. 81–88.

[13] L. Evans, *Partial Differential Equations*, AMS, Providence, RI, 1993.

[14] F. Giannessi and A. Maugeri, *Variational Inequalities and Network Equilibrium Problems*, Springer, New York, 1995.

[15] B. Führ, V. Schulz, and K. Welker, *Shape optimization for interface identification with obstacle problems*, Vietnam J. Math., 46 (2018), pp. 96–985, doi:10.1007/s10013-018-0312-0.

[16] L. Gasiński, *Mapping method in optimal shape design problems governed by hemivariational inequalities*, in Shape Optimization and Optimal Design, J. Cagnol, M. Polis, and J.-P. Zolésio, eds., Lect. Notes Pure Appl. Math. 216, Marcel Dekker, New York, 2001, pp. 277–288.

[17] C. Heinemann and K. Sturm, *Shape optimization for a class of semilinear variational inequalities with applications to damage models*, SIAM J. Math. Anal., 48 (2016), pp. 3579–3617.

[18] M. Hintermüller, *An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities*, Inverse Problems, 24 (2008), 034017.

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

[20] M. Hintermüller and L. Laurain, *Optimal shape design subject to elliptic variational inequalities*, SIAM J. Control Optim., 49 (2011), pp. 1015–1047.

[21] M. Hintermüller and T. Surowiec, *First-order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis*, SIAM J. Optim., 21 (2011), pp. 1561–1593.

[22] K. Ito and K. Kunisch, *Optimal control of elliptic variational inequalities*, Appl. Math. Optim., 41 (2000), pp. 343–364.

[23] K. Ito and K. Kunisch, *Semi-smooth Newton methods for variational inequalities of the first kind*, ESIAM Math. Model. Numer. Anal., 37 (2003), pp. 41–62.

[24] M. Juntunen and R. Stenberg, *Nitsche's method for general boundary conditions*, Math. Comp., 78 (2009), pp. 1353–1374.

[25] K. Sturm, *On Shape Optimization with Non-linear Partial Differential Equations*, Ph.D. thesis, 2015.

[26] K. Sturm, *Shape differentiability under non-linear PDE constraints*, in New Trends in Shape Optimization, A. Pratelli and G. Leugering, eds., Internat. Ser. Numer. Math. 166, Springer, New York, 2015, pp. 271–300.

[27] M. Kocvara and J. Outrata, *Shape optimization of elasto-plastic bodies governed by variational inequalities*, in Boundary Control and Variation, J.-P. Zolésio, ed., Lect. Notes Pure Appl. Math. 163, Marcel Dekker, New York, 1994, pp. 261–271.

[28] J. L. Lions and G. Stampacchia, *Variational inequalities*, Commun. Pure Appl. Math., 20 (1967), pp. 493–519.

[29] W. Liu and J. Rubio, *Optimal shape design for systems governed by variational inequalities, part 1: Existence theory for the elliptic case*, J. Optim. Theory Appl., 69 (1991), pp. 351–371.

[30] W. Liu and J. Rubio, *Optimal shape design for systems governed by variational inequalities, part 2: Existence theory for the evolution case*, J. Optim. Theory Appl., 69 (1991), pp. 373–396.

[31] A. Logg, K.-A. Mardal, and G. N. Wells, *Automated Solution of Differential Equations by the Finite Element Method*, Lect. Notes Comput. Sci. Eng. 84, Springer, New York, 2012, https://doi.org/10.1007/978-3-642-23099-8.

[32] M. Bauer, M. Bruveris and P. W. Michor, *Overview of the geometries of shape spaces and diffeomorphism groups*, J. Math. Imaging Vision, 50 (2014), pp. 60–97.

[33] M. Cocou, *Existence of solutions of a dynamic Signorini's problem with nonlocal friction in viscoelasticity*, Z. Angew. Math. Phys., 53 (2002), pp. 1099–1109.

[34] B. Maury, *A fat boundary method for the Poisson problem in a domain with holes*, J. Sci. Comput., 16 (2001), pp. 319–339.

[35] A. Myśliński, *Domain optimization for unilateral problems by an embedding domain method*, in Shape Optimization and Optimal Design, J. Cagnol, M. Polis, and J.-P. Zolésio, eds., Lect. Notes Pure Appl. Math. 216, Marcel Dekker, New York, 2001, pp. 355–370.

[36] A. Myśliński, *Level set approach for shape optimization of contact problems*, in Proceedings of the European Congress on Computational Methods in Applied Sciences and Engineering, P. Neittaanmäki, T. Rossi, K. Majava, O. Pironneau, and I. Lasiecka, eds., 2004.

[37] A. Myśliński, *Level set method for shape and topology optimization of contact problems*, in IFIP Conference on System Modeling and Optimization, Springer, New York, 2007, pp. 397–410.

[38] N. Kikuchi and J. T. Oden, *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*, Stud. Appl. Numer. Math. 8, SIAM, Philadelphia, 1988.

[39] A. Novotny and J. Sokolowski, *Topological Derivatives in Shape Optimization*, Springer, New York, 2013.

[40] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Pure Appl. Math. 140, Elsevier, New York, 2003.

[41] A. Schiela and D. Wachsmuth, *Convergence analysis of smoothing methods for optimal control of stationary variational inequalities*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 771–787.

[42] V. Schulz, M. Siebenborn, and K. Welker, *Efficient PDE constrained shape optimization based on Steklov–Poincaré type metrics*, SIAM J. Optim., 26 (2016), pp. 2800–2819.

[43] M. Siebenborn and K. Welker, *Algorithmic aspects of multigrid methods for optimization in shape spaces*, SIAM J. Sci. Comput., 39 (2017), pp. B1156–B1177.

[44] J. Sokolowski and J. Zolésio, *Introduction to Shape Optimization*, Comput. Math. 16, Springer, New York, 1992.

[45] K. Sturm, *On Shape Optimization with Non-Linear Partial Differential Equations*, Ph.D. thesis, Technische Universität Berlin, 2015.

[46] G. Troianiello, *Elliptic Differential Equations and Obstacle Problems*, Springer, New York, 2013.

[47] K. Welker, *Efficient PDE Constrained Shape Optimization in Shape Spaces*, Ph.D. thesis, Universität Trier, 2016.

[48] K. Welker, *Suitable Spaces for Shape Optimization*, https://arxiv.org/abs/1702.07579, 2017.

[49] E. Zeidler, *Applied Functional Analysis: Main Principles and Their Applications*, Appl. Math. Sci. 109, Springer, New York, 2012.