ruhr.paD
UA Ruhr Zentrum für
partielle Differentialgleichungen

Strong Stationarity Conditions for a Class of Optimization Problems Governed by Variational Inequalities of the 2nd Kind

J. C. de los Reyes and Ch. Meyer

Preprint 2015-07

STRONG STATIONARITY CONDITIONS FOR A CLASS OF OPTIMIZATION PROBLEMS GOVERNED BY VARIATIONAL INEQUALITIES OF THE 2ND KIND

J. C. DE LOS REYES[‡] AND C. MEYER[§]

Abstract. We investigate optimality conditions for optimization problems constrained by a class of variational inequalities of the second kind. Based on a nonsmooth primal-dual reformulation of the governing inequality, the differentiability of the solution map is studied. Directional differentiability is proved both for finite-dimensional and function space problems, under suitable assumptions on the active set. A characterization of B- and strong stationary optimal solutions is obtained thereafter. Finally, based on the obtained first-order information, a trust-region algorithm is proposed for the solution of the optimization problems.

Key words. Variational inequalities, optimality conditions, mathematical programs with equilibrium constraints.

1. Introduction. Optimization problems with variational inequality constraints have been intensively investigated in the last years with many important applications in focus. Problems in contact mechanics, phase separation or elastoplasticity are some of the most relevant application examples. Special analytical and numerical techniques have been developed for characterizing and finding optima of such problems, mainly in the finite-dimensional case (see [18] and references therein).

In the function space framework much of the work has been devoted to optimization problems constrained by variational inequalities of the first kind:

$$\min\ j(y,u) \tag{1.1a}$$

$$\text{subject to: } (Ay, v-y) \geq (u, v-y),\ \text{for all } v \in K, \tag{1.1b}$$

where $A : V \mapsto V^*$ is an elliptic operator and $K \subset V$ is closed convex set. Such obstacle type structure has allowed to develop an analytical machinery for such kind of problems. In addition, different type of stationarity concepts have been investigated in that framework (C-, B-, M- and strong stationary points). The utilized proof techniques include regularization approaches as well as differentiability properties (directional, conic) of the solution map or elements of set valued analysis (see e.g. [1, 2, 10–12, 16, 17, 19, 20, 22, 23]).

For problems involving variational inequalities of the second kind:

$$\min\ j(y,u) \tag{1.2a}$$

$$\text{subject to: } (Ay, v-y) + \varphi(v) - \varphi(y) \geq (u, v-y),\ \text{for all } v \in V, \tag{1.2b}$$

with φ continuous and convex, only weak results have been obtained in the past, due to the very general structure (see e.g. [1–3, 21]). In [4] a special class of problems were investigated, where a richer structure of the nondifferentiability was exploited. Nonsmooth terms of the type $\varphi(y) = \int_S |By|\ ds$ were considered there and, by using a tailored regularization approach, a more detailed optimality system was obtained.

[‡]Research Center on Mathematical Modelling (MODEMAT), Escuela Politécnica Nacional, Quito-Ecuador

[§]Faculty of Mathematics, Technische Universität Dortmund, Dortmund-Germany.

The results were then extended to problems in fluid mechanics [5], image processing [7] and elastoplasticity [6]. Thanks to the availability of primal and dual formulations in elastoplasticity, the kind of optimality systems obtained in [4] were proved to be equivalent to C-stationary optimality systems in optimization problems constrained by variational inequalities of the first kind, see [6].

In this paper we aim to characterize further stationary points by investigating differentiability properties of the solution map. In that spirit B- and strong stationarity conditions are in focus. To avoid problems related to the regularity of the variables, we start by considering the finite-dimensional case. A reformulation of the variational inequality as a nonsmooth system of primal dual equations enables us to take difference quotients and prove directional differentiability of the finite-dimensional solution operator.

The technique is then extended to the function space setting. Since in this context the regularity of the functions as well as the structure of the active set play a crucial role, special functional analysis and measure theoretical methods have to be considered. As a preparatory step, the Lipschitz continuity of the solution operator from $L^p(\Omega) \to L^\infty(\Omega)$ is proved by using Stampacchia's technique. The directional differentiability of the solution map is then proved by assuming that the active set has a special structure, namely that it consists of the union of a regular subdomain of positive measure and a set of zero capacity (see Assumption 3.16 below). With the directional differentiability at hand, the characterization of B-stationarity points is carried out thereafter. The theoretical part of the paper ends with the derivation of strong stationarity conditions by an adaptation of the method of proof introduced by [20] for optimal control of the obstacle problem.

In the last part of the paper the first order information related to the directional derivative is utilized within a trust-region algorithm for the solution of the VI-constrained optimization problem. The computed derivative information is treated as an inexact descent direction, which is inserted into the trust-region framework to get robust iterates. The performance of the resulting algorithm is tested on a representative test problem, showing the suitability of the approach.

2. Differentiability for a finite dimensional VI of second kind. We start by considering the following prototypical VI in $\mathbb{R}^n$:

$$\langle Ay,\, v - y\rangle + |v|_1 - |y|_1 \geq \langle u,\, v - y\rangle \quad \forall\, v \in \mathbb{R}^n. \tag{2.1}$$

Throughout this section $\langle .,\, .\rangle = \langle .,\, .\rangle_{\mathbb{R}^n}$ denotes the Euclidean scalar product. Moreover, $A \in \mathbb{R}^{n\times n}$ is positive definite and $|v|_1 = \sum_{i=1}^n |v_i|$. Existence and uniqueness for (2.1) for arbitrary right hand sides $u \in \mathbb{R}^n$ follows by classical arguments due to the maximal monotonicity of $A + \partial|\,.\,|_1$.

Definition 2.1. *We denote the solution mapping associated to* (2.1) *by* $S : \mathbb{R}^n \ni u \mapsto y \in \mathbb{R}^n$.

By Introducing a slack variable $q := u - Ay \in \mathbb{R}^n$, we see that (2.1) is equivalent to

$$q \in \partial|\cdot|_1(y), \tag{2.2}$$

where $\partial|\cdot|_1$ denotes the convex subdifferential of $\mathbb{R}^n \ni v \mapsto |v|_1 \in \mathbb{R}$. Evaluating the

subdifferential in (2.2) leads to the following system of nonsmooth equations

$$\begin{cases} Ay + q = u & \\ q_i y_i = |y_i|, & i = 1, 2, ..., n \\ \max\{|q_i|, 1\} = 1, & i = 1, 2, ..., n. \end{cases} \tag{2.3}$$

In order to derive a directional derivative for S, consider a perturbed version of (2.1), given by

$$\begin{aligned} Ay^t + q^t &= u + t\,h \\ q_i^t y_i^t &= |y_i^t|, \quad i = 1, 2, ..., n \\ \max\{|q_i^t|, 1\} &= 1, \quad i = 1, 2, ..., n, \end{aligned} \tag{2.4}$$

which leads to the following nonsmooth system for the difference quotient:

$$\begin{aligned} A\,\frac{y^t - y}{t} + \frac{q^t - q}{t} &= h \\ \frac{q_i^t y_i^t - q_i y_i - (|y_i^t| - |y_i|)}{t} &= 0, \quad i = 1, 2, ..., n \\ \frac{\max\{|q_i^t|, 1\} - \max\{|q_i|, 1\}}{t} &= 0, \quad i = 1, 2, ..., n. \end{aligned} \tag{2.5}$$

In the sequel, we will pass to the limit in (2.5) to obtain the relations determining the directional derivative of S. For this purpose we test the VI associated with (2.4), given by

$$\langle Ay^t,\, v - y^t\rangle + |v|_1 - |y^t|_1 \geq \langle u + th,\, v - y^t\rangle \quad \forall\, v \in \mathbb{R}^n, \tag{2.6}$$

with $v = y$. If we test (2.1) with $v = y^t$ and add both inequalities, we arrive at

$$\lambda_{\min}(A)\left|\frac{y^t - y}{t}\right|^2 \leq \left\langle \frac{y^t - y}{t},\, A\,\frac{y^t - y}{t}\right\rangle \leq \left\langle h,\, \frac{y^t - y}{t}\right\rangle,$$

where $|\,.\,| = |\,.\,|_{\mathbb{R}^n}$ denotes the euclidian norm and $\lambda_{\min}(A) > 0$ is the smallest eigenvalue of A. Thus

$$\left|\frac{y^t - y}{t}\right| \leq \frac{1}{\lambda_{\min}(A)}\,|h| < \infty,$$

and so there exists a converging subsequence, w.l.o.g. $\left\{\frac{y^t - y}{t}\right\}_{t>0}$ itself, such that

$$\frac{y^t - y}{t} \xrightarrow{t\searrow 0} \eta. \tag{2.7}$$

In Theorem 2.6 below we will see that the limit η is unique so that the whole sequence $\{(y^t - y)/t\}$ converges. This justifies to assume the convergence of the whole sequence right from the beginning. By definition of q we have

$$\frac{q^t - q}{t} = h - A\,\frac{y^t - y}{t} \xrightarrow{t\searrow 0} h - A\eta =: \lambda, \tag{2.8}$$

which in particular implies $q^t \to q$.

LEMMA 2.2. *For all $i = 1, 2, ..., n$ there holds*

$$\frac{q_i^t y_i^t - q_i y_i - (|y_i^t| - |y_i|)}{t} \xrightarrow{t \searrow 0} \lambda_i y_i + q_i \eta_i - \mathrm{abs}'(y_i; \eta_i) \tag{2.9}$$

$$\frac{\max\{|q_i^t|, 1\} - \max\{|q_i|, 1\}}{t} \xrightarrow{t \searrow 0} \max'(|q_i|; \mathrm{abs}'(q_i; \lambda_i)). \tag{2.10}$$

Herein abs' *and* $\max'$ *denote the directional derivatives of* $\mathbb{R} \ni x \mapsto |x| \in \mathbb{R}$ *and* $\mathbb{R} \ni x \mapsto \max\{x, 1\} \in \mathbb{R}$, *i.e.,*

$$\mathrm{abs}'(a; b) = \begin{cases} \mathrm{sign}(a) b, & a \neq 0, \\ |b|, & a = 0, \end{cases} \qquad \textit{and} \qquad \max'(a; b) = \begin{cases} 0, & a < 1, \\ b, & a > 1, \\ \max\{b, 0\}, & a = 1. \end{cases}$$

Proof. The mapping $g : \mathbb{R}^2 \to \mathbb{R}$, $g(a, b) := a\, b - |a|$ is directionally differentiable and Lipschitz continuous, and thus Hadamard-differentiable, cf. [**?**]. Moreover, one has

$$\frac{q_i^t y_i^t - q_i y_i - (|y_i^t| - |y_i|)}{t} = \frac{g\big((y_i, q_i) + t(\eta_i, \lambda_i) + r(t)\big) - g(y_i, q_i)}{t}$$

with

$$r(t) := (y_i^t, q_i^t) - (y_i, q_i) - t(\eta_i, \lambda_i) = o(t)$$

according to (2.7) and (2.8). Therefore the Hadamard-differentiability yields (2.9).

To prove (2.10) observe that, by the same argument as before, $x \mapsto \max\{x, 1\}$ is also Hadamard-differentiable. Thanks to the chain rule for Hadamard-derivatives the mapping $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \max\{|x|, 1\}$ is thus Hadamard-differentiable, too. The same reasoning as above then yields (2.10). □

REMARK 2.3. *Note that*

$$\max'(|a|; \mathrm{abs}'(a; b)) = \begin{cases} 0, & |a| < 1 \\ \mathrm{sign}(a) b, & |a| > 1 \\ \max\{0, ab\}, & |a| = 1. \end{cases}$$

In view of (2.7), (2.8), and Lemmas 2.2, we can pass to the limit as $t \searrow 0$ in (2.5) and obtain in this way:

$$A\eta + \lambda = h \tag{2.11a}$$

$$\lambda_i y_i + q_i \eta_i = \mathrm{abs}'(y_i; \eta_i), \quad i = 1, 2, ..., n \tag{2.11b}$$

$$\max\{0, q_i \lambda_i\} = 0 \quad \text{for all } i \in \{1, ..., n\} \text{ with } |q_i| = 1. \tag{2.11c}$$

(Note that the case $|q_i| > 1$ is obsolete.) The system (2.11) will lead to a VI satisfied by the limit η. To see this, we have to reformulate (2.11) in the following way:

LEMMA 2.4. *The system* (2.11) *is equivalent to*

$$A\eta + \lambda = h \tag{2.12a}$$

$$\lambda_i = 0 \quad \textit{for all } i \in \{1, ..., n\} \textit{ with } y_i \neq 0 \tag{2.12b}$$

$$\eta_i = 0 \quad \textit{for all } i \in \{1, ..., n\} \textit{ with } |q_i| < 1 \tag{2.12c}$$

$$\eta_i q_i \geq 0 \quad \textit{for all } i \in \{1, ..., n\} \textit{ with } y_i = 0,\ |q_i| = 1 \tag{2.12d}$$

$$\lambda_i q_i \leq 0 \quad \textit{for all } i \in \{1, ..., n\} \textit{ with } y_i = 0,\ |q_i| = 1. \tag{2.12e}$$

Proof. (2.11) ⇒ (2.12):
It is evident that

$$\max\{0, q_i\lambda_i\} = 0 \text{ if } |q_i| = 1 \quad \Longleftrightarrow \quad q_i\lambda_i \leq 0 \text{ if } |q_i| = 1, \tag{2.13}$$

which implies (2.12e). Next, let $i \in \{1, ..., n\}$ such that $y_i \neq 0$. Then

$$q_i = \frac{y_i}{|y_i|} = \operatorname{sign}(y_i),$$

and hence (2.11b) yields $\lambda_i y_i = 0$, which in turn gives (2.12b) due to $y_i \neq 0$. Now take $i \in \{1, ..., n\}$ with $|q_i| < 1$ arbitrary. Then we have $y_i = 0$, and hence (2.11b) implies $q_i\eta_i = |\eta_i|$. Because of $|q_i| < 1$ this results in (2.12c). To show (2.12d), let $i \in \{1, ..., n\}$ with $y_i = 0$ and $|q_i| = 1$ be arbitrary. Then (2.11b) gives $q_i\eta_i = |\eta_i| \geq 0$.

(2.12) ⇒ (2.11):
Due to (2.12b) and (2.12e) we have $\lambda_i q_i \leq 0$ whenever $|q_i| = 1$, which, in view of (2.13), implies (2.11c). Because of (2.12b), we have

$$\lambda_i y_i + \eta_i q_i = \eta_i q_i \quad \forall\, i = 1, ..., n. \tag{2.14}$$

Now, if $y_i \neq 0$, then $q_i = \operatorname{sign}(y_i)$ and thus $\eta_i q_i = \operatorname{sign}(y_i)\eta_i$. If $y_i = 0$ and $|q_i| < 1$, then, by (2.12c), we obtain $\eta_i q_i = 0 = |\eta_i|$. If finally $y_i = 0$ and $|q_i| = 1$, then (2.12d) implies $\eta_i q_i = |\eta_i||q_i| = |\eta_i|$. In summary (2.11b) is verified, which yields the assertion. □

System (2.12) is not yet complete, since there is still one relation missing to derive the VI fulfilled by η. The missing part is stated in the following lemma.

LEMMA 2.5. *There holds*

$$\eta_i\lambda_i = 0 \quad \textit{for all } i \in \{1, ..., n\} \textit{ with } y_i = 0,\ |q_i| = 1.$$

Proof. Let $i \in \{1, ..., n\}$ with $y_i = 0$ and $|q_i| = 1$ be arbitrary. W.l.o.g. we assume that $q_i = 1$. The case $q_i = -1$ can be discussed analogously. If $\eta_i = 0$, the assertion is trivially fulfilled. So let $\eta_i \neq 0$. By (2.12d) and $q_i = 1$ we then have $\eta_i > 0$. Due to (2.7) this implies

$$\frac{y_i^t - y_i}{t} > 0 \quad \text{for } t > 0 \text{ sufficiently small}$$

and thus, due to $y_i = 0$,

$$y_i^t > 0 \quad \text{for } t > 0 \text{ sufficiently small.}$$

Consequently, $q_i^t = \operatorname{sign}(y_i^t) = 1$ for $t > 0$ sufficiently small and hence, since $q_i = 1$ by assumption,

$$\lambda_i = \lim_{t \searrow 0} \frac{q_i^t - q_i}{t} = 0,$$

which gives the assertion. □

Now we have everything at hand to prove the main result of this section, i.e., the directional differentiability of $S : u \mapsto y$.

THEOREM 2.6. *The solution mapping S of (2.1) is directionally differentiable at every point $u \in \mathbb{R}^n$ and the directional derivative $\eta = S'(u; h)$ in direction $h \in \mathbb{R}^n$ solves the following* VI of first kind*:*

$$\eta \in K(y), \quad \langle A\eta,\, v - \eta\rangle \geq \langle h,\, v - \eta\rangle \quad \forall\, v \in K(y) \tag{2.15}$$

where $K(y)$ is the convex cone defined by

$$K(y) := \{v \in \mathbb{R}^n : v_i = 0 \text{ if } |q_i| < 1,\ v_i q_i \geq 0 \text{ if } y_i = 0,\ |q_i| = 1\}. \tag{2.16}$$

Proof. Define the biactive set by

$$\mathcal{B} := \{i \in \{1, ..., n\} : y_i = 0,\ |q_i| = 1\}.$$

First we show that the limit η solves (2.15). We already know that η satisfies (2.12) and in addition $\eta_i \lambda_i = 0$ if $y_i = 0$ and $|q_i| = 1$. Thus (2.12c) and (2.12d) imply $\eta \in K(y)$, i.e., feasibility of η. Now let $v \in K(y)$ be arbitrary. Then (2.12b), $v \in K(y)$, and (2.12e) yield

$$\langle \lambda,\, v\rangle = \sum_{i \in \mathcal{B}} \lambda_i v_i = \sum_{i \in \mathcal{B}} \lambda_i \underbrace{q_i q_i}_{=1} v_i \leq 0. \tag{2.17}$$

Similarly, we infer from (2.12b), $\eta \in K(y)$, and Lemma 2.5 that

$$\langle \lambda,\, \eta\rangle = \sum_{i \in \mathcal{B}} \lambda_i v_i = 0.$$

Therefore, if we multiply (2.12a) with $v - \eta$, then we arrive at

$$\langle h,\, v - \eta\rangle = \langle A\eta,\, v - \eta\rangle + \langle \lambda,\, v\rangle - \langle \lambda,\, \eta\rangle \leq \langle A\eta,\, v - \eta\rangle,$$

so that the limit η indeed solves (2.15).

Since A is positive definite and $K(y)$ is convex and closed, the operator $A + \partial I_{K(y)}(.) : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is maximal monotone, where $I_{K(y)}$ denotes the indicator function of the set $K(y)$. Thus there is a unique solution of (2.15). Since every accumulation point η of the difference quotient $(y^t - y)/t$ solves (2.15), the limit is thus unique and consequently a well-known argument gives the convergence of the whole sequence. ∎

COROLLARY 2.7. *Let the biactive set have zero cardinality, i.e. $y_i = 0$ implies $|q_i| < 1$. Then S is Gâteaux-differentiable, i.e. $S'(u; h)$ is linear and continuous w.r.t. h, and $\eta = S'(u)h$ is given by the unique solution of the following linear system:*

$$\eta_i = 0 \quad \text{for all } i \in \{1, ..., n\} \text{ with } y_i = 0 \tag{2.18}$$

$$\sum_{j : y_j \neq 0} A_{ij} \eta_j = h_i \quad \text{for all } i \in \{1, ..., n\} \text{ with } y_i \neq 0. \tag{2.19}$$

Proof. If the biactive set has zero cardinality, then (2.12c) implies (2.18). Moreover, (2.12b) immediately yields (2.19). Since A is positive definite, the same holds for $A_{\mathcal{I}} := (A_{ij})_{i,j \in \mathcal{I}}$ with $\mathcal{I} := \{i \in \{1, ..., n\} : y_i \neq 0\}$. Thus $A_{\mathcal{I}}$ is invertible and $\eta_{\mathcal{A}} = A_{\mathcal{I}}^{-1} h_{\mathcal{I}}$. Together with (2.18), i.e. $\eta_{\{1,...,n\} \setminus \mathcal{I}} = 0$, this implies that η is uniquely determined by (2.18) and (2.19). Moreover, due to the invertibility of $A_{\mathcal{I}}$, η depends continuously on h as claimed. ∎

3. Weak differentiability for a VI of second kind in function space. Next we extend the result of the preceding section to a VI of second kind in function space. For this purpose, let $\Omega \subset \mathbb{R}^d$, $d \geq 1$, be a bounded domain with regular boundary satisfying the cone condition. We consider the following prototypical VI of second kind:

$$\langle Ay,\, v - y\rangle + \int_\Omega |v|\,dx - \int_\Omega |y|\,dx \geq \langle u,\, v - y\rangle \quad \forall\, v \in V, \tag{VI2}$$

where we abbreviated $V := H_0^1(\Omega)$. From now on $\langle .,\, .\rangle$ denotes the dual pairing in V. Furthermore $A : V \to V^*$ stands for the following linear second-order elliptic differential operator:

$$Ay = \sum_{i=1}^{d} \Big(\sum_{j=1}^{d} \frac{\partial}{\partial x_i} a_{ij} \frac{\partial y}{\partial x_j} + b_i \frac{\partial y}{\partial x_i} \Big) + \gamma\, y, \tag{3.1}$$

where $a_{ij}, b_i, \gamma \in L^\infty(\Omega)$, $i, j = 1, .., d$, are such that A is coercive, i.e.

$$\langle Ay,\, y\rangle \geq \alpha \,\|y\|_V^2, \tag{3.2}$$

with a constant $\alpha > 0$. In addition, we require

$$\gamma \geq 0. \tag{3.3}$$

Moreover, $u \in V^*$ is given a inhomogeneity.

The plan of this section is as follows. First we state some well known results for (VI2) concerning existence, uniqueness, and an equivalent reformulation by means of a complementarity-like system. Then we introduce a perturbed problem, similar to (2.4), and derive several auxiliary results for the associated difference quotients and their (weak) limits. In order to show an infinite dimensional analogon to (2.12b), we unfortunately need to assume some properties of the active set, see Assumption 3.16 below. Based on this assumption, we can derive a weak directional differentiability result, similar to Theorem 2.6 (see Theorem 3.19 below).

LEMMA 3.1. *For every $u \in V^*$ there exists a unique solution $y \in V$ of (VI2), which we denote by $y = S(u)$. The associated solution operator $S : V^* \to V$ is globally Lipschitz continuous, i.e., there exists a constant $L > 0$ such that*

$$\|S(u_1) - S(u_2)\|_V \leq L\, \|u_1 - u_2\|_{V^*} \quad \forall\, u_1, u_2 \in V^*. \tag{3.4}$$

Proof. Existence and uniqueness for (VI2) follows by standard arguments from the maximal monotonicity of $A + \partial\|.\|_{L^1(\Omega)}$, see for instance [1]. To prove the Lipschitz continuity we test the VI for $y_1 = S(u_1)$ with $y_2 = S(u_2)$ and vice versa and add the arising inequalities to obtain

$$\langle A(y_1 - y_2),\, y_1 - y_2\rangle \leq \langle u_1 - u_2,\, y_1 - y_2\rangle.$$

The coercivity of A then yields the result. ∎

REMARK 3.2. *Sometimes we will use S with different domains and ranges, which may be inferred from the context.*

By standard arguments based on Fenchel duality or the Hahn-Banach theorem, the VI in (VI2) can be rewritten in terms of a complementarity-like system, see e.g. [4]. In this way the following result is obtained:

LEMMA 3.3. *For every $u \in V^*$ there exists a unique function $q \in L^2(\Omega)$ such that the unique solution $y \in V$ of (VI2) fulfills the following complementarity-like system:*

$$\langle Ay,\, v\rangle + \int_\Omega q\, v\, dx = \langle u,\, v\rangle \quad \forall\, v \in V \tag{3.5a}$$

$$q(x)y(x) = |y(x)|, \quad |q(x)| \le 1 \quad \text{a.e. in } \Omega. \tag{3.5b}$$

The function q is called slack function in all what follows, and we will refer to (3.5b) as slackness condition in the sequel.

Next let $h \in V^*$ be arbitrary and $\{t_n\} \subset \mathbb{R}^+$ be an arbitrary sequence of positive numbers tending to 0. We denote the solutions to the VI associated to $u + t_n h$ by y_n, i.e.,

$$\langle Ay_n,\, v - y_n\rangle + \int_\Omega |v|\, dx - \int_\Omega |y_n|\, dx \ge \langle u + t_n h,\, v - y_n\rangle \quad \forall\, v \in V. \tag{3.6}$$

The associated slack function is analogously denoted by $q_n \in L^2(\Omega)$, i.e.

$$\begin{aligned} \langle Ay_n,\, v\rangle + \int_\Omega q_n\, v\, dx &= \langle u + t_n h,\, v\rangle \quad \forall\, v \in V, \\ q_n(x)y(x) = |y_n(x)|, &\quad |q_n(x)| \le 1 \quad \text{a.e. in } \Omega. \end{aligned} \tag{3.7}$$

By Lemma 3.1 it holds

$$\left\| \frac{y_n - y}{t_n} \right\|_V \le \|h\|_{V^*}$$

and thus there is a weakly convergent subsequence, denoted the same, and a limit point $\eta \in V$ such that

$$\frac{y_n - y}{t_n} \rightharpoonup \eta \text{ in } V. \tag{3.8}$$

This simplification of notation will be justified by the uniqueness of the weak limit η, which implies the weak convergence of the whole sequence by a well-known argument (see Theorem 3.19 below). For the slack functions we obtain

$$\int_\Omega \frac{q_n - q}{t_n}\, v\, dx = \langle h,\, v\rangle - \left\langle A\, \frac{y_n - y}{t_n},\, v \right\rangle \to \langle h - A\eta,\, v\rangle \quad \forall\, v \in V,$$

i.e.,

$$\frac{q_n - q}{t_n} \rightharpoonup \lambda \text{ in } V^*,$$

with $\lambda = h - A\eta$. Note that it is in general not possible to show the boundedness of $(q_n - q)/t_n$ in any Lebesgue space so that one cannot expect λ to be more regular.

Next consider the first equation in the slackness condition (3.5b) for y and y_n. By multiplying these equations with $1/t_n$ and an arbitrary $\varphi \in C_0^\infty(\Omega)$, integrating over Ω, and taking the difference, we arrive at

$$\int_\Omega \frac{q_n - q}{t_n}\, y_n\, \varphi\, dx + \int_\Omega \frac{y_n - y}{t_n}\, q\, \varphi\, dx = \int_\Omega \frac{|y_n| - |y|}{t_n}\, \varphi\, dx, \quad \forall\, \varphi \in C_0^\infty(\Omega). \tag{3.9}$$

In order to pass to the limit in this relation, we have to define the following sets:

DEFINITION 3.4. *We define –up to sets of zero measure–*

$$\begin{aligned} \mathcal{A} &:= \{x \in \Omega : y(x) = 0\}, & \mathcal{A}_s &:= \{x \in \Omega : |q(x)| < 1\} \\ \mathcal{I} &:= \{x \in \Omega : y(x) \neq 0\}, & \mathcal{B} &:= \{x \in \Omega : |q(x)| = 1,\ y(x) = 0\} \\ \mathcal{I}^+ &:= \{x \in \Omega : y(x) > 0\}, & \mathcal{I}^- &:= \{x \in \Omega : y(x) < 0\} \\ \mathcal{B}^+ &:= \{x \in \Omega : q(x) = 1,\ y(x) = 0\}, & \mathcal{B}^- &:= \{x \in \Omega : q(x) = -1,\ y(x) = 0\}. \end{aligned} \tag{3.10}$$

The set $\mathcal{A}$ is called active set*, while $\mathcal{A}_s$ is the* strongly active set*. Moreover, we call $\mathcal{I}$ and $\mathcal{B}$* inactive *and* biactive set*, respectively.*

Note that

$$\Omega = \mathcal{A} \cup \mathcal{I} \quad \text{and} \quad \mathcal{A} = \mathcal{A}_s \cup \mathcal{B},$$

due to (3.5b). The next lemma covers the directional differentiability of the L^1-norm. Its proof is straightforward and therefore postponed to Appendix A.

LEMMA 3.5. *For every $\varphi \in L^\infty(\Omega)$ it holds*

$$\int_\Omega \frac{|y_n(x)| - |y(x)|}{t_n}\, \varphi(x)\, dx \to \int_\Omega \operatorname{abs}'\big(y(x); \eta(x)\big)\, \varphi(x)\, dx,$$

where abs′ *again denotes the directional derivative of the absolute value.*

Together with Lemma 3.5 the weak convergence of $(q_n - q)/t_n$ in V^* and $(y_n - y)/t_n$ in V and the strong convergence of y_n to y in V allow to pass to the limit in (3.9), which results in

$$\langle \lambda,\, y\, \varphi \rangle + \int_\Omega \eta\, q\, \varphi\, dx = \int_\Omega \operatorname{abs}'(y; \eta)\, \varphi\, dx \quad \forall\, \varphi \in C_0^\infty(\Omega). \tag{3.11}$$

Using this relation, we can prove the following result, which is just the infinite dimensional counterpart to (2.12c) and (2.12d):

LEMMA 3.6. *There holds*

$$\eta(x) = 0 \quad \text{a.e., where } |q(x)| < 1 \tag{3.12}$$

$$\eta(x)\, q(x) \geq 0 \quad \text{a.e., where } |q(x)| = 1 \text{ and } y(x) = 0. \tag{3.13}$$

Proof. Let $\varphi \in C_0^\infty(\Omega)$ with $\varphi \geq 0$ a.e. in Ω be arbitrary. The slackness condition (3.5b) implies for all $n \in \mathbb{N}$ that

$$\frac{q_n(x) - q(x)}{t_n}\, y(x) \leq 0 \quad \text{a.e. in } \Omega. \tag{3.14}$$

Indeed, if $y(x) = 0$, then the assertion is trivial. If $y(x) > 0$, then (3.5b) implies $q(x) = 1$ and thus $q_n(x) - q(x) \leq 0$, since $|q_n(x)| \leq 1$. Analogously, if $y(x) < 0$, then $q(x) = -1$ and hence $q_n(x) - q(x) \geq 0$ is obtained. All in all we have thus proven (3.14). Therefore we have

$$\langle \lambda,\, y\, \varphi \rangle = \lim_{n \to \infty} \int_\Omega \frac{q_n - q}{t_n}\, y\, \varphi\, dx \leq 0,$$

and thus (3.11) yields

$$\int_\Omega \eta\, q\, \varphi\, dx \geq \int_\Omega \mathrm{abs}'(y;\eta)\, \varphi\, dx \quad \forall\, \varphi \in C_0^\infty(\Omega) \text{ with } \varphi \geq 0.$$

The fundamental lemma of the calculus of variations thus yields

$$\eta(x)\, q(x) \geq \mathrm{abs}'(y;\eta)(x) \quad \text{a.e. in } \Omega,$$

which by definition of $\mathrm{abs}'(y;\eta)$ in turn gives

$$\eta(x)\, q(x) \geq |\eta(x)| \quad \text{a.e. in } \mathcal{A}.$$

Since $|q(x)| \leq 1$ a.e. in Ω, this results in

$$\eta(x)\, q(x) = |\eta(x)| \quad \text{a.e. in } \mathcal{A}. \tag{3.15}$$

As the slackness conditions in (3.5b) implies $\{x \in \Omega : |q(x)| < 1\} \subset \{x \in \Omega : y(x) = 0\}$, the result follows immediately from (3.15). ∎

LEMMA 3.7. *There holds* $\langle \lambda, \eta \rangle \geq 0$.

Proof. By inserting the definition of the slack variable q into (VI2) one obtains

$$\int_\Omega q(v - y)\, dx \leq \int_\Omega |v|\, dx - \int_\Omega |y|\, dx \quad \forall\, v \in V \tag{3.16}$$

and an analogous inequality for q_n and y_n. Inserting $y_n \in V$ in this inequality and y in the corresponding one for q_n and y_n, adding both inequalities and dividing by t_n^2 yields

$$\int_\Omega \frac{q_n - q}{t_n} \frac{y_n - y}{t_n}\, dx \geq 0.$$

Since A is elliptic and bounded, the mapping $V \ni w \mapsto \langle Aw, w \rangle \in \mathbb{R}$ is convex and continuous and thus weakly lower semicontinuous. The equations for q and q_n and the weak convergence of $(y_n - y)/t_n$ in V therefore imply

$$\begin{aligned}
0 &\leq \liminf_{n\to\infty} \int_\Omega \frac{q_n - q}{t_n} \frac{y_n - y}{t_n}\, dx \\
&\leq \limsup_{n\to\infty} \int_\Omega \frac{q_n - q}{t_n} \frac{y_n - y}{t_n}\, dx \\
&= \limsup_{n\to\infty} \left(\left\langle h, \frac{y_n - y}{t_n} \right\rangle - \left\langle A\Big(\frac{y_n - y}{t_n}\Big), \frac{y_n - y}{t_n} \right\rangle \right) \\
&\leq \lim_{n\to\infty} \left\langle h, \frac{y_n - y}{t_n} \right\rangle - \liminf_{n\to\infty} \left\langle A\Big(\frac{y_n - y}{t_n}\Big), \frac{y_n - y}{t_n} \right\rangle \\
&\leq \langle h, \eta \rangle - \langle A\eta, \eta \rangle = \langle \lambda, \eta \rangle.
\end{aligned}$$

∎

The most delicate issue, when transferring the finite dimensional findings of Section 2 to the function space setting, is to verify the conditions (2.12a) and (2.12e) on λ. To do so, we first prove that S is Lipschitz continuous in $L^\infty(\Omega)$, provided that the right hand sides in (VI2) are more regular. We employ the well-known technique of

Stampacchia based on the following lemma, whose proof is presented in Appendix B for convenience of the reader.

LEMMA 3.8 (Stampacchia). *For every function $w \in V$ and every $k \geq 0$, the function w_k defined by*

$$w_k(x) := \begin{cases} w(x) - k, & w(x) \geq k \\ 0, & |w(x)| < k \\ w(x) + k, & w(x) \leq -k \end{cases} \tag{3.17}$$

is an element of V. Furthermore, if there is a constant $\alpha > 0$ such that

$$\alpha \|w_k\|^2_{H^1(\Omega)} \leq \int_\Omega f\, w_k \, dx \quad \forall\, k \geq 0 \tag{3.18}$$

with a function $f \in L^p(\Omega)$, $p > \max\{d/2, 1\}$, then w is essentially bounded and there exists a constant $c > 0$ so that

$$\|w\|_{L^\infty(\Omega)} \leq c\, \|f\|_{L^p(\Omega)}. \tag{3.19}$$

LEMMA 3.9. *There exists a constant $K > 0$ such that*

$$\|S(u_1) - S(u_2)\|_{L^\infty(\Omega)} \leq K\, \|u_1 - u_2\|_{L^p(\Omega)}$$

for all $u_1, u_2 \in L^p(\Omega)$ with $p > \max\{d/2, 1\}$. Here we identified $u \in L^p(\Omega)$ with an element of V^*.

Proof. We apply Lemma 3.8 to $w := y_1 - y_2$ with $y_i = S(u_i)$, $i = 1, 2$. To this end we shall verify (3.18) with $f = u_1 - u_2$. For this purpose let $v \in V$ be arbitrary and test the VI for y_1 with $y_1 - v$ and the one for y_2 with $y_2 + v$ and add the arising inequalities to obtain:

$$\langle A(y_1 - y_2),\, v\rangle + \int_\Omega \big(|y_1| + |y_2| - |y_1 - v| - |y_2 + v|\big) dx \leq \int_\Omega (u_1 - u_2) v \, dx \quad \forall\, v \in V. \tag{3.20}$$

Next let $k \geq 0$ be arbitrary and define $w_k = (y_1 - y_2)_k$ as in (3.17). In the following we will prove that

$$I(x) := |y_1(x)| + |y_2(x)| - |y_1(x) - w_k(x)| - |y_2(x) + w_k(x)| \geq 0 \quad \text{a.e. in } \Omega, \tag{3.21}$$

by a simple distinction of cases.

1st case: $|y_1(x) - y_2(x)| < k$:
In this case we have $w_k(x) = 0$ and thus (3.21) is trivially fulfilled with equality.

2nd case: $y_1(x) - y_2(x) \geq k$:
Now we obtain $w_k(x) = y_1(x) - y_2(x) - k$ and consequently

$$I(x) = |y_1(x)| + |y_2(x)| - |y_2(x) + k| - |y_1(x) - k|.$$

If $y_1(x) \geq k$ and $y_2(x) \leq -k$, then

$$I(x) = |y_1(x)| + |y_2(x)| + y_2(x) + k - y_1(x) + k \geq 2k \geq 0.$$

If $y_1(x) \le k$ and $y_2(x) \ge -k$, then

$$I(x) = |y_1(x)| + |y_2(x)| - y_2(x) - k + y_1(x) - k \ge 2\big(y_1(x) - y_2(x) - k\big) \ge 0,$$

where we used $y_1(x) - y_2(x) \ge k$ for the last estimate.
If $y_1(x) \ge k$ and $y_2(x) \ge -k$, then

$$I(x) = |y_1(x)| + |y_2(x)| - y_2(x) - y_1(x) \ge 0.$$

If finally $y_1(x) \le k$ and $y_2(x) \le -k$, then

$$I(x) = |y_1(x)| + |y_2(x)| + y_2(x) + y_1(x) \ge 0,$$

which gives the assertion of (3.21) for this case.

3rd case: $y_1(x) - y_2(x) \le -k$*:*
In this case we get that $y_2(x) - y_1(x) \ge k$ and thus $I(x) = |y_1(x)| + |y_2(x)| - |y_2(x) - k| - |y_1(x) + k|$. Interchanging the roles of $y_1(x)$ and $y_2(x)$ and repeating the arguments for the second case immediately yields (3.21) in the third case.

Let us now define $\mathcal{A}_k := \{x \in \Omega : |w(x)| \ge k\}$. From the first part of Lemma 3.8 we get that $w_k \in V$ and so we are allowed to insert w_k as test function in (3.20). Owing to the coercivity of A, the definition of w_k in (3.17), (3.3), and (3.21), we then obtain

$$\begin{aligned}
\alpha \, \|w_k\|^2_{H^1(\Omega)} &\le \langle A w_k, \, w_k \rangle \\
&= \int_{\mathcal{A}_k} \Big[\sum_{i=1}^d \Big(\sum_{j=1}^d a_{ij} \frac{\partial w_k}{\partial x_j} \, \frac{\partial w_k}{\partial x_j} dx + b_i \, \frac{\partial w_k}{\partial x_i} \, w_k + \gamma \, w_k^2 \Big] \mathrm{d}x \\
&\le \int_{\Omega} \Big[\sum_{i=1}^d \Big(\sum_{j=1}^d a_{ij} \frac{\partial w}{\partial x_j} \, \frac{\partial w_k}{\partial x_j} dx + b_i \, \frac{\partial w}{\partial x_i} \, w_k + \gamma \, w \, w_k \Big] \mathrm{d}x \\
&= \langle A w, \, w_k \rangle = \langle A(y_1 - y_2), \, w_k \rangle \le \int_{\Omega} (u_1 - u_2) w_k \, dx,
\end{aligned}$$

which is (3.18) with $f = u_1 - u_2$. Since $k \ge 0$ was arbitrary, all conditions of Lemma 3.8 are satisfied so that it can be applied and gives the desired result. □

REMARK 3.10. *Since* $S(0) = 0$*, it immediately follows from Lemma 3.9 that*

$$\|S(u)\|_{L^\infty(\Omega)} \le c \|u\|_{L^p(\Omega)}.$$

COROLLARY 3.11. *If* $u, h \in L^p(\Omega)$ *with* $p > \max\{d/2, 1\}$*, then*

$$\frac{y_n - y}{t_n} \rightharpoonup^* \eta \ \text{ in } L^\infty(\Omega),$$

which implies $\eta \in L^\infty(\Omega)$*.*

Proof. By Lemma 3.9 $(y_n - y)/t_n$ is bounded in $L^\infty(\Omega)$. Thus, there is a subsequence converging weakly-$*$ to a element $\tilde{\eta} \in L^\infty(\Omega)$. This subsequence therefore converges weakly in $L^2(\Omega)$ and in view of (3.8) we find

$$\int_\Omega \eta \, v \, dx = \int_\Omega \tilde{\eta} \, v \, dx, \quad \forall \, v \in L^2(\Omega).$$

The fundamental lemma of the calculus of variations implies $\tilde{\eta} = \eta$ a.e. in Ω. Since the subsequential weak limit is therefore independent of the subsequence, a standard argument implies weak-$*$ convergence of the whole sequence as claimed. ∎

Based on the Lipschitz continuity of S in Lemma 3.9, we can prove a first result towards an infinite dimensional counterpart to (2.12a).

LEMMA 3.12. *Assume that $u, h \in L^p(\Omega)$ with $p > \max\{d/2, 1\}$. Let moreover $\rho > 0$ be arbitrary and define –up to sets of measure zero–*

$$\mathcal{A}_\rho := \{x \in \Omega : y(x) \in [-\rho, \rho]\}.$$

Then for all $v \in V$ with $v(x) = 0$ a.e. in $\mathcal{A}_\rho$ there holds

$$\langle \lambda, v \rangle = 0.$$

Proof. Let $\rho > 0$ and $v \in V$ with $v(x) = 0$ a.e. in $\mathcal{A}_\rho$ be arbitrary. Thanks to Lemma 3.9 we have

$$\|y_n - y\|_{L^\infty(\Omega)} \le K\, t_n\, \|h\|_{L^p(\Omega)} \to 0. \tag{3.22}$$

Therefore, for almost all $x \in \Omega$ with $y(x) > \rho$, it follows that

$$y_n(x) \ge y(x) - |y(x) - y_n(x)| \ge \rho - \|y - y_n\|_{L^\infty(\Omega)} \ge \frac{\rho}{2} > 0, \quad \forall\, n \ge N_1,$$

with $N_1 \in \mathbb{N}$ depending on ρ but not on x. Therefore, thanks to (3.5b), we have for all $n \ge N_1$ that

$$q_n(x) = \frac{y_n(x)}{|y_n(x)|} = 1 \quad \Rightarrow \quad \frac{q_n(x) - q(x)}{t_n} = 0 \quad \text{f.a.a. } x \in \Omega \text{ with } y(x) > \rho, \tag{3.23}$$

where we used that $q(x) = 1$ due to $y(x) > \rho > 0$. Completely analogously one can show the existence of $N_2 \in \mathbb{N}$, only depending on ρ, such that

$$\frac{q_n(x) - q(x)}{t_n} = 0 \quad \text{f.a.a. } x \in \Omega \text{ with } y(x) < -\rho$$

for all $n \ge N_2$. Therefore, since $v(x) = 0$ a.e., where $y(x) \in [-\rho, \rho]$, we obtain

$$\int_\Omega \frac{q_n - q}{t_n}\, v\, dx = 0 \quad \forall\, n \ge \max\{N_1, N_2\}.$$

The convergence $(q_n - q)/t_n \rightharpoonup \lambda$ in V^* thus implies the assertion. ∎

The aim is now to drive ρ in Lemma 3.12 to zero. This however requires several additional assumptions. The first one covers the regularity of y and q.

ASSUMPTION 3.13.

1. *We assume that the solution $y = S(u)$ is continuous.*
2. *The slack function is continuous, i.e. $q \in C(\bar{\Omega})$.*

REMARK 3.14. *Let us point out that Assumption 3.13(1) is not restrictive at all. Indeed, Lemma 3.3 implies that y solves $Ay = u - q$ and, if $u \in L^2(\Omega)$, then y thus solves a second-order elliptic equation with right hand side in $L^2(\Omega)$. For problems of this*

type, standard regularity theory yields continuity of the solution under mild assumptions on the data, see for instance [8]. In contrast to this, Assumption 3.13(2) cannot be guaranteed in general. Nevertheless, multiple numerical observations indicate that q is often continuous.

If Assumption 3.13 is satisfied, i.e. if y and q have continuous representatives, then we can define the sets in Definition 3.4 in a pointwise manner, i.e., not only up to sets of zero measure. The sets arising in this way are denoted by the same symbols, and we always mean these sets in all what follows when writing $\mathcal{A}$, $\mathcal{I}$, $\mathcal{B}$ etc.

LEMMA 3.15. *Under Assumption 3.13 the sets $\mathcal{I}^+$ and $\mathcal{I}^-$ are strictly separated, i.e., there exists $\delta > 0$ such that*

$$\operatorname{dist}(\mathcal{I}^+, \mathcal{I}^-) := \min\left\{ |x - z|_{\mathbb{R}^d} : x \in \overline{\mathcal{I}^+}, z \in \overline{\mathcal{I}^-} \right\} > \delta.$$

Proof. Since $\bar{\Omega}$ is compact, Assumption 3.13(2) implies that q is uniformly continuous. From the slackness condition (3.5b) we infer $q = 1$ in $\mathcal{I}^+$ so that the uniform continuity of q yields the existence of $\delta > 0$ with

$$q(x) \geq 1/2 \quad \text{for all } x \in \mathcal{I}^+ + B(0, \delta). \tag{3.24}$$

Hence, due to $q = -1$ on $\mathcal{I}^-$ by (3.5b), this gives the assertion. ∎

In addition to Assumption 3.13, we need the following rather restrictive assumption on the active set.

ASSUMPTION 3.16. *The active set $\mathcal{A} = \{x \in \Omega : y(x) = 0\}$ satisfies the following conditions:*

1. *$\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_0$, where $\mathcal{A}_1$ has positive measure and $\mathcal{A}_0$ has zero capacity.*
2. *$\mathcal{A}_1$ is closed and possesses non-empty interior. Moreover, it holds $\mathcal{A}_1 = \overline{\operatorname{int}(\mathcal{A}_1)}$.*
3. *For the set $\mathcal{J} := \Omega \setminus \mathcal{A}_1$ it holds*

$$\partial\mathcal{J} \setminus (\partial\mathcal{J} \cap \partial\Omega) = \partial\mathcal{A}_1 \setminus (\partial\mathcal{A}_1 \cap \partial\Omega), \tag{3.25}$$

and both $\mathcal{A}_1$ and $\mathcal{J}$ are supposed to have regular boundaries. That is the connected components of $\mathcal{J}$ and $\mathcal{A}_1$ have positive distance from each other and the boundaries of each of them satisfies the cone condition.

Figures 3.1 and 3.2 illustrate Assumption 3.16 in the two-dimensional case.

With the help of Assumption 3.13 and 3.16 we can now prove the following infinite dimensional counterpart to (2.12a):

LEMMA 3.17. *Let $u, h \in L^p(\Omega)$, $p > \max\{d/2, 1\}$, be given. Assume that u is such that Assumptions 3.13 and 3.16 are fulfilled. Then*

$$\langle \lambda, v \rangle = 0 \quad \textit{for all } v \in V \textit{ with } v(x) = 0 \textit{ a.e. in } \mathcal{A}$$

holds true.

Proof. Let $v \in V$ with $v(x) = 0$ a.e. in $\mathcal{A}$ be arbitrary. By Assumption 3.16(3) there are linear and continuous trace operators $\tau_j : H^1(\Omega) \to L^2(\partial\mathcal{J})$ and $\tau_a : H^1(\Omega) \to L^2(\partial\mathcal{A}_1)$. Due to $v = 0$ a.e. in $\mathcal{A}_1$, we have $\tau_a v = 0$ and, by (3.25) and $v \in V$, thus

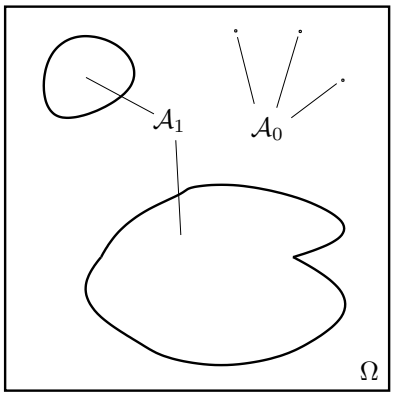

FIG. 3.1. *Active set satisfying Assumption 3.16*

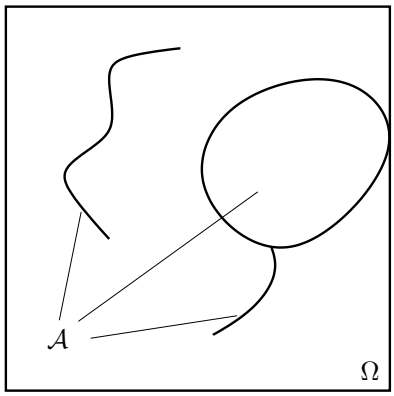

FIG. 3.2. *Active set not feasible for Assumption 3.16*

$\tau_j v = 0$. Since $\partial\mathcal{J}$ is regular, there exists a sequence $\{\varphi_k\}_{k\in\mathbb{N}} \subset C_0^\infty(\mathcal{J})$ with $\varphi_k \to v$ in $H^1(\mathcal{J})$, see e.g. [9, Lemma 1.33]. In particular it holds

$$\omega_k := \operatorname{supp}(\varphi_k) \subset\subset \mathcal{J}.$$

We extend φ_k by zero outside $\mathcal{J}$ to obtain a function in $C_0^\infty(\Omega)$, which we denote by the same symbol for simplicity. Because of $v = 0$ a.e. in $\mathcal{A}_1$ it follows that

$$\varphi_k \overset{k\to\infty}{\longrightarrow} v \quad \text{in } V. \tag{3.26}$$

By construction we have $\mathcal{J} \subset \mathcal{I} \cup \mathcal{A}_0$. Since $\mathcal{A}_0$ has zero capacity, there is a sequence $\{w_m\}_{m\in\mathbb{N}} \subset V$ and a sequence of open neighborhoods of $\mathcal{A}_0$, denoted by $\{\mathcal{U}_m\}_{m\in\mathbb{N}} \subset \Omega$, such that

$$w_m \geq 0 \text{ a.e. in } \Omega, \quad w_m = 1 \text{ a.e. in } \mathcal{U}_m, \quad w_m \overset{m\to\infty}{\longrightarrow} 0 \text{ in } H^1(\Omega).$$

Now let $k, m \in \mathbb{N}$ be fixed but arbitrary and define

$$\mathcal{I}_m^+ := (\omega_k \setminus \mathcal{U}_m) \cap \mathcal{I}^+, \quad \mathcal{I}_m^- := (\omega_k \setminus \mathcal{U}_m) \cap \mathcal{I}^-.$$

Since $\mathcal{U}_m$ is open, $\omega_k \setminus \mathcal{U}_m$ is closed. Moreover, in view of $\mathcal{J} = \mathcal{I} \cup \mathcal{A}_0$, it holds $\omega_k \setminus \mathcal{U}_m \subset \mathcal{I}$. Thus, Lemma 3.15 and the boundedness of Ω yield that $\mathcal{I}_m^+$ and $\mathcal{I}_m^-$ are compact. The continuity of y therefore implies that there is $\xi \in \mathcal{I}_m^+$ such that

$$y(\xi) = \min_{x\in\mathcal{I}_m^+} y(x)$$

and, due to $\xi \in \mathcal{I}^+$, one obtains $\rho_m^+ := y(\xi) > 0$. Analogously one derives $\rho_m^- := \max_{x\in\mathcal{I}_m^-} y(x) < 0$. As in the proof of Lemma 3.12 one proves the existence of $N_m^+ \in \mathbb{N}$ such that for all $n \geq N_m^+$ there holds

$$\frac{q_n(x) - q(x)}{t_n} = 0 \quad \text{f.a.a. } x \in \Omega \text{ with } y(x) \geq \rho_m^+,$$

see (3.23). Clearly, there is $N_m^- \in \mathbb{N}$ so that the same equation holds for every $n \geq N_m^-$ and almost all $x \in \Omega$ with $y(x) \leq \rho_m^-$. Consequently, we obtain

$$\frac{q_n - q}{t_n} = 0 \quad \text{a.e. in } \omega_k \setminus \mathcal{U}_m = \mathcal{I}_m^+ \cup \mathcal{I}_m^-, \tag{3.27}$$

provided that $n \geq N_m := \max\{N_m^+, N_m^-\}$.

Thanks to (3.27) and $w_m = 1$ a.e. in $\mathcal{U}_m$, it follows

$$\begin{aligned}\int_\Omega \frac{q_n - q}{t_n}\,\varphi_k\,w_m\,dx &= \int_{\omega_k\setminus\mathcal{U}_m} \frac{q_n - q}{t_n}\,\varphi_k\,w_m\,dx + \int_{\mathcal{U}_m} \frac{q_n - q}{t_n}\,\varphi_k\,w_m dx \\ &= \int_{\mathcal{U}_m} \frac{q_n - q}{t_n}\,\varphi_k\,dx \qquad \forall\, n \geq N_m.\end{aligned} \tag{3.28}$$

On the other hand $\varphi_k w_m \in V$ is a feasible test function for (3.5a) and (3.7). If we insert this test function and subtract the arising equation, then (3.28) together with Hölder's inequality and Sobolev embeddings yield

$$\begin{aligned}&\int_{\mathcal{U}_m} \frac{q_n - q}{t_n}\,\varphi_k\,dx = \int_\Omega \frac{q_n - q}{t_n}\,\varphi_k\,w_m\,dx \\ &= -\int_\Omega \nabla\Big(\frac{y_n - y}{t_n}\Big)\cdot\nabla(\varphi_k w_m)\,dx + \int_\Omega h\,\varphi_k\,w_m\,dx \\ &\leq 2\Big\|\frac{y_n - y}{t_n}\Big\|_{H^1(\Omega)}\|w_m\|_{H^1(\Omega)}\|\varphi_k\|_{W^{1,\infty}(\Omega)} + c\,\|h\|_{L^2(\Omega)}\|w_m\|_{H^1(\Omega)}\|\varphi_k\|_{H^1(\Omega)}\end{aligned}$$

for all $n \geq N_m$. Therefore, in view of (3.27), the weak convergence (and thus boundedness) of $(y_n - y)/t_n$ gives

$$\int_\Omega \frac{q_n - q}{t_n}\,\varphi_k\,dx = \int_{\mathcal{U}_m} \frac{q_n - q}{t_n}\,\varphi_k\,dx \leq c\,\|w_m\|_{H^1(\Omega)}\|\varphi_k\|_{W^{1,\infty}(\Omega)}$$

for all $n \geq N_m$ and thus

$$\langle\lambda,\,\varphi_k\rangle = \lim_{n\to\infty}\int_\Omega \frac{q_n - q}{t_n}\,\varphi_k\,dx \leq c\,\|w_m\|_{H^1(\Omega)}\|\varphi_k\|_{W^{1,\infty}(\Omega)}.$$

Due to $w_m \to 0$ in $H^1(\Omega)$, passing to the limit $m \to \infty$ yields $\langle\lambda,\,\varphi_k\rangle \leq 0$. The above arguments also apply to $-\varphi_k$ so that $\langle\lambda,\,\varphi_k\rangle = 0$. Since $k \in \mathbb{N}$ was arbitary, this equation holds for every $k \in \mathbb{N}$ and thus we can pass to the limit $k \to \infty$. The convergence in (3.26) then gives the assertion. □

Similarly to (2.16), we define

$$\begin{aligned}\mathcal{K}(y) &:= \{v \in V :\ v(x) = 0 \text{ a.e. in } \mathcal{A}_s,\ v(x)q(x) \geq 0 \text{ a.e. in } \mathcal{B}\} \\ &= \{v \in V :\ v(x) = 0 \text{ a.e., where } |q(x)| < 1, \\ &\qquad\qquad v(x)q(x) \geq 0 \text{ a.e., where } |q(x)| = 1 \text{ and } y(x) = 0\}\end{aligned} \tag{3.29}$$

This set will be the feasible set of the VI belonging to the directional derivative of S (see Theorem 3.19 below). As seen in the proof of Theorem 2.6, in the finite dimensional setting, there holds $\lambda^\top v \leq 0$ for all $v \in K(y)$, see (2.17). The infinite dimensional analogon is also true, provided that Assumptions 3.13 and 3.16 hold, as the following lemma shows.

LEMMA 3.18. *Let $u, h \in L^p(\Omega)$ with $p > \max\{d/2, 1\}$ be given, and assume that u is such that Assumptions 3.13 and 3.16 are fulfilled. Then there holds*

$$\langle\lambda,\,v\rangle \leq 0 \quad \textit{for all } v \in \mathcal{K}(y).$$

Proof. Let $v \in \mathcal{K}(y)$ be fixed but arbitrary. Due to $\mathcal{A}_s \cup \mathcal{B} \cup \mathcal{I} = \Omega$ and $v(x) = 0$ a.e. in $\mathcal{A}_s$, we obtain

$$\int_\Omega \frac{q_n - q}{t_n}\, v\, dx = \int_{\mathcal{B}^+} \frac{q_n - 1}{t_n}\, v\, dx + \int_{\mathcal{B}^-} \frac{q_n + 1}{t_n}\, v\, dx + \int_{\mathcal{I}} \frac{q_n - q}{t_n}\, v\, dx, \tag{3.30}$$

Since $q_n \in [-1, 1]$ a.e. in Ω and $q\, v \geq 0$ a.e. in $\mathcal{B}$, which implies $v \geq 0$ a.e. in $\mathcal{B}^+$ and $v \leq 0$ a.e. in $\mathcal{B}^-$, we can further estimate

$$\int_\Omega \frac{q_n - q}{t_n}\, v\, dx \leq \int_{\mathcal{I}} \frac{q_n - q}{t_n}\, v\, dx = \int_{\mathcal{J}} \frac{q_n - q}{t_n}\, v\, dx,$$

where $\mathcal{J}$ is the set from Assumption 3.16(3). For the last equality we used that $\mathcal{J} = \mathcal{I} \cup \mathcal{A}_0$ and $\mathcal{A}_0$ has zero capacity, thus zero Lebesgue-measure.

We now prove that $\mathcal{J} = \mathcal{J}^+ \cup \mathcal{J}^-$, where $\mathcal{J}^+$ and $\mathcal{J}^-$ possess regular boundaries and coincide with $\mathcal{I}^+$ and $\mathcal{I}^-$ up to sets of zero capacity. For this purpose, we show $\overline{\mathcal{J}} = \overline{\mathcal{I}}$. Due to $\mathcal{I} \subset \mathcal{J}$, we clearly have $\overline{\mathcal{I}} \subseteq \overline{\mathcal{J}}$. Let $\xi \in \overline{\mathcal{J}}$ be arbitrary. Then there is a sequence $\{x_k\}_{k\in\mathbb{N}} \subset \mathcal{J}$ so that $x_k \to \xi$. If $\{x_k\}$ contains a subsequence in $\mathcal{I}$, we immediately obtain $\xi \in \overline{\mathcal{I}}$. So assume the contrary, i.e., in view of $\mathcal{J} = \mathcal{I} \cup \mathcal{A}_0$, $x_k \in \mathcal{A}_0$ for all $k \in \mathbb{N}$ sufficiently large. W.l.o.g. we assume $\{x_k\} \subset \mathcal{A}_0$ for the whole sequence. Since $\mathcal{A}_0$ has zero capacity, thus zero measure, there is, for each x_k, a sequence $\{x_k^{(m)}\}_{m\in\mathbb{N}} \subset \Omega \setminus \mathcal{A}_0$ with $x_k^{(m)} \to x_k$ for $m \to \infty$. Since $x_k^{(m)} \notin \mathcal{A}_0$, we have either $x_k^{(m)} \in \mathcal{A}_1$ or $x_k^{(m)} \in \mathcal{I}$. If $\{x_k^{(m)}\}$ would contain a subsequence in $\mathcal{A}_1$, then the closedness of $\mathcal{A}_1$ would imply $x_k \in \mathcal{A}_1$ in contradiction to $x_k \in \mathcal{A}_0$. Thus we may w.l.o.g. assume that $\{x_k^{(m)}\} \subset \mathcal{I}$. Therefore, there is a diagonal sequence $\{x_k^{(m(k))}\} \subset \mathcal{I}$ converging to ξ, which gives $\xi \in \overline{\mathcal{I}}$. Hence we have shown

$$\overline{\mathcal{J}} = \overline{\mathcal{I}} = \overline{\mathcal{I}^+} \cup \overline{\mathcal{I}^-}$$

with $\mathcal{I}^+$ and $\mathcal{I}^-$ as defined in (3.10). Since $\overline{\mathcal{I}^+}$ and $\overline{\mathcal{I}^-}$ have positive distance from each other by Lemma 3.15, there exist sets $\mathcal{J}^+, \mathcal{J}^-$ such that $\mathcal{J}^+ \cup \mathcal{J}^- = \mathcal{J}$ and $\operatorname{dist}(\mathcal{J}^+, \mathcal{J}^-) > \delta$. Moreover, thanks to Lemma 3.15 and $\mathcal{J} = \mathcal{I} \cup \mathcal{A}_0$ with $\operatorname{cap}(\mathcal{A}_0) = 0$, $\mathcal{J}^+$ differs from $\mathcal{I}^+$ only by a set of zero capacity and the same holds for $\mathcal{J}^-$ and $\mathcal{I}^-$. Finally, because of $\operatorname{dist}(\mathcal{J}^+, \mathcal{J}^-) > \delta$, Assumption 3.16(3) yields that $\mathcal{J}^+$, $\mathcal{J}^-$, $\Omega \setminus \mathcal{J}^+$, and $\Omega \setminus \mathcal{J}^-$ possess regular boundaries. (This actually implies that $\mathcal{J}^\pm = \operatorname{int}(\overline{\mathcal{I}^\pm})$.)

Since $\mathcal{J}^+$ differs from $\mathcal{I}^+$ only on a set of zero measure, the definition of $\mathcal{I}^+$ and the slackness condition (3.5b) imply $q = 1$ a.e. in $\mathcal{J}^+$, and analogously $q = -1$ a.e. in $\mathcal{J}^-$. Thus (3.30) can be further estimated by

$$\begin{aligned}
\int_\Omega \frac{q_n - q}{t_n}\, v\, dx &\leq \int_{\mathcal{J}^+} \underbrace{\frac{q_n - 1}{t_n}}_{\leq 0} \underbrace{\max\{0, v\}}_{\geq 0}\, dx + \int_{\mathcal{J}^+} \frac{q_n - q}{t_n} \min\{0, v\}\, dx \\
&\quad + \int_{\mathcal{J}^-} \underbrace{\frac{q_n + 1}{t_n}}_{\geq 0} \underbrace{\min\{0, v\}}_{\geq 0}\, dx + \int_{\mathcal{J}^-} \frac{q_n - q}{t_n} \max\{0, v\}\, dx \\
&\leq \int_{\mathcal{J}^+} \frac{q_n - q}{t_n} \min\{0, v\}\, dx + \int_{\mathcal{J}^-} \frac{q_n - q}{t_n} \max\{0, v\}\, dx.
\end{aligned} \tag{3.31}$$

Next we show that $\min\{0, v\} \in H_0^1(\mathcal{J}^+)$ and $\max\{0, v\} \in H_0^1(\mathcal{J}^-)$. The proof of Lemma 3.15 shows

$$\big(\mathcal{I}^+ + B(0,\varepsilon)\big) \setminus \mathcal{I}^+ \subset \{x \in \Omega : q(x) \geq 1/2,\ y(x) = 0\} \subset \mathcal{A}_s \cup \mathcal{B}^+, \tag{3.32}$$

see (3.24). Because of $v \in \mathcal{K}(y)$ we have $q\, v \geq 0$ a.e. in $\mathcal{A}_s \cup \mathcal{B}^+$ and thus (3.32) gives $v \geq 0$ a.e. in $\big(\mathcal{I}^+ + B(0,\varepsilon)\big) \setminus \mathcal{I}^+$. Since $\mathcal{I}^+$ and $\mathcal{J}^+$ differ only up to a set of zero measure, we thus get

$$\min\{0, v\} = 0 \quad \text{a.e. in } (\mathcal{J}^+ + B(0,\varepsilon)) \setminus \mathcal{J}^+.$$

The regularity of $\partial\mathcal{J}^+$ and $\partial(\Omega \setminus \mathcal{J}^+)$ therefore gives

$$\min\{0, v(x)\} = 0 \quad \text{a.e. on } \partial\mathcal{J}^+,$$

and thus $\min\{0, v\} \in H_0^1(\mathcal{J}^+)$. An analogous argument shows that $\max\{0, v\} \in H_0^1(\mathcal{J}^-)$. Due to the zero trace and the regularity of $\partial\mathcal{J}^+$ by Assumption 3.16(3), we can extend $\min\{0, v\}$ by zero outside $\mathcal{J}^+$ to obtain a function in V, i.e., $\chi_{\mathcal{J}^+} \min\{0, v\} \in V$, where $\chi_{\mathcal{J}^+}$ denotes the characteristic function of $\mathcal{J}^+$. Thus the weak convergence $(q_n - q)/t_n \rightharpoonup \lambda$ in V^* gives

$$\int_{\mathcal{J}^+} \frac{q_n - q}{t_n} \min\{0, v\}\, dx = \int_\Omega \frac{q_n - q}{t_n} \chi_{\mathcal{J}^+} \min\{0, v\}\, dx \to \langle \lambda,\ \chi_{\mathcal{J}^+} \min\{0, v\}\rangle.$$

Since $\chi_{\mathcal{J}^+} \min\{0, v\} = 0$ a.e. in $\mathcal{A} \subset \Omega \setminus \mathcal{J}^+$, Lemma 3.17 yields $\langle \lambda,\ \chi_{\mathcal{J}^+} \min\{0, v\}\rangle = 0$. Analogously

$$\int_{\mathcal{J}^-} \frac{q_n - q}{t_n} \max\{0, v\}\, dx \to \langle \lambda,\ \chi_{\mathcal{J}^-} \max\{0, v\}\rangle = 0$$

is obtained. Therefore, in view of (3.31), we finally arrive at $\langle \lambda,\ v\rangle \leq 0$ and, since $v \in \mathcal{K}(y)$ was arbitrary, this proves the assertion. □

Now we are finally in the position to prove the main result of this section covering the "weak directional differentiability" of the solution operator associated with the VI in (VI2).

THEOREM 3.19. *Let $u, h \in L^p(\Omega)$ with $p > \max\{d/2, 1\}$ be given. Suppose further that Assumptions 3.13 and 3.16 are fulfilled by $y = S(u)$ and the associated slack variable q. Then there holds*

$$\frac{S(u + t\,h) - S(u)}{t} \rightharpoonup \eta \quad \text{in } V, \quad \text{as } t \searrow 0, \tag{3.33}$$

where $\eta \in V$ solves the following VI of first kind:

$$\eta \in \mathcal{K}(y), \quad \langle A\eta,\ v - \eta\rangle \geq \langle h,\ v - \eta\rangle \quad \forall\, v \in \mathcal{K}(y) \tag{3.34}$$

with $\mathcal{K}(y)$ as defined in (3.29).

Proof. Lemma 3.6 yields $\eta \in \mathcal{K}(y)$. Furthermore, since $A\eta + \lambda = h$, Lemmas 3.7 and 3.18 give

$$\langle A\eta,\ v - \eta\rangle - \langle h,\ v - \eta\rangle = \langle \lambda,\ \eta\rangle - \langle \lambda,\ v\rangle \geq 0$$

for all $v \in \mathcal{K}(y)$, which is just the VI in (3.34).

Since $\mathcal{K}(y)$ is nonempty, convex, and closed and A is bounded and coercive, standard arguments yields existence and uniqueness for this VI of first kind. Thus the weak limit η is unique, which implies the weak convergence of the whole sequence. ∎

DEFINITION 3.20. *With a little abuse of notation we call the weak limit η in* (3.33) weak directional derivative *and denote it by $\eta = S'_w(u;h)$.*

REMARK 3.21. *If $\mathcal{B}$ has zero measure, then $\mathcal{K}(y)$ turns into*

$$\mathcal{K}(y) = \{v \in V : \ v(x) = 0 \ \text{a.e. in } \mathcal{A}_s\},$$

i.e., a linear and closed subspace of V. Thus, in this case, (3.34) *becomes an equation. If $\mathcal{A}_s$ possesses a regular boundary, then this equation is equivalent to*

$$A\eta = h \quad \text{in } \mathcal{I} \quad \text{and} \quad \eta = 0 \quad \text{a.e. in } \mathcal{A} = \mathcal{A}_s.$$

REMARK 3.22. *It is very likely that Theorem 3.19 could be proven without the restrictive Assumption 3.16, if the weak limit η would satisfy the conditions in* (3.12) *and* (3.13) *not only almost everywhere, but* quasi-everywhere *in Ω. In this case, the feasible set of* (3.34) *would read*

$$\begin{aligned}\mathcal{K} := \{v \in V : \ & v(x) = 0 \ \text{q.e., where } |q(x)| < 1,\\ & v(x)q(x) \geq 0 \ \text{q.e., where } |q(x)| = 1 \ \text{and } y(x) = 0\}.\end{aligned}$$

However, unfortunately, so far we have neither been able to show that (3.15) *holds quasi everywhere, nor to establish a counterexample which demonstrates that this is wrong in general. This question gives rise to future research.*

4. Bouligand stationarity. With the differentiability result of Theorem 3.19 at hand, it is now straightforward to establish first-order optimality conditions in purely primal form for optimization problems governed by (VI2). To be more precise, we consider an optimization problem of the form

$$\left.\begin{aligned} &\min \quad J(y,u)\\ &\text{s.t.} \quad \langle Ay,\, v-y\rangle + \int_\Omega |v|\,dx - \int_\Omega |y|\,dx \geq \langle u,\, v-y\rangle \quad \forall\, v \in V\\ &\text{and} \quad u \in U_{\text{ad}}, \end{aligned}\right\} \tag{4.1}$$

where $U_{\text{ad}} \subset L^p(\Omega)$, $p > \max\{d/2, 1\}$, is nonempty, closed, and convex.

As shown in [10, Lemma 3.9], weak convergence of the difference quotient associated with the control-to-state mapping $S : u \mapsto y$ is sufficient to prove that the reduced objective, defined by

$$j : L^p(\Omega) \to \mathbb{R}, \quad j(u) := J(S(u), u),$$

is directionally differentiable. This allows us to formulate the following purely primal optimality conditions, which, in case of optimal control of VIs of first kind, are known as Bouligand stationarity conditions.

THEOREM 4.1. *Let $p > \max\{d/2, 1\}$ and assume that J is Fréchet-differentiable from $V \times L^p(\Omega)$ to $\mathbb{R}$. Suppose moreover that $\bar{u} \in U_{\text{ad}}$ is a local optimal solution of* (4.1),

such that $\bar{y} = S(\bar{u})$ and the associated slack variable $\bar{q}$ satisfy Assumptions 3.13 and 3.16. Then the following primal stationarity conditions are fulfilled:

$$\partial_y J(\bar{y}, \bar{u})\eta + \partial_u J(\bar{y}, \bar{u})(u - \bar{u}) \geq 0 \quad \forall\, u \in U_{\mathrm{ad}}, \tag{4.2}$$

where $\eta \in V$ solves (3.34) with $\mathcal{K}(y) = \mathcal{K}(\bar{y})$ and $h = u - \bar{u}$.

Proof. As mentioned above, [10, Lemma 3.9] and Theorem 3.19 imply that $u \mapsto j(u)$ is directionally differentiable in every direction $h \in L^p(\Omega)$ with directional derivative $j'(\bar{u}; h) = \partial_y J(\bar{y}, \bar{u}) S'_w(\bar{u}; h) + \partial_u J(\bar{y}, \bar{u})h$. Local optimality of $\bar{u}$ yields $j'(\bar{u}; u - \bar{u}) \geq 0$, which is the assertion. ∎

Next we derive a variant of the above optimality condition based on the cone tangent to the admissible set of (4.1). As a result, we obtain an optimality condition which can be interpreted as the counterpart of the implicit programming approach in the discussion of finite dimensional MPECs, see [18, Section 3.3]. Note that such similarities have already been observed in [10].

LEMMA 4.2. *Assume that $\bar{u} \in L^p(\Omega)$, $p > \max\{d/2, 1\}$, is such that Assumptions 3.13 and 3.16 are fulfilled. Suppose moreover that the sequences $\{u_n\} \subset L^p(\Omega)$ and $\{t_n\} \subset \mathbb{R}^+$ satisfy*

$$t_n \searrow 0, \quad \frac{u_n - \bar{u}}{t_n} \rightharpoonup h \quad \text{in } L^p(\Omega).$$

Then

$$\frac{S(u_n) - S(\bar{u})}{t_n} \rightharpoonup S'_w(\bar{u}; h) \quad \text{in } V.$$

Proof. By adding a zero we obtain

$$\frac{S(u_n) - S(\bar{u})}{t_n} = \frac{S(u_n) - S(\bar{u} + t_n\, h)}{t_n} + \frac{S(\bar{u} + t_n\, h) - S(\bar{u})}{t_n}.$$

While the latter addend converges weakly to $S'_w(\bar{u}; h)$ by Theorem 3.19, the Lipschitz continuity of S by Lemma 3.1 yields for the first addend that

$$\left\| \frac{S(u_n) - S(\bar{u} + t_n\, h)}{t_n} \right\|_V \leq L \left\| \frac{u_n - \bar{u}}{t_n} - h \right\|_{V^*} \to 0,$$

where we used the compactness of the embedding $L^p(\Omega) \hookrightarrow V^*$. ∎

We define the tangent cone to the admissible set of (4.1) as follows:

DEFINITION 4.3 (Tangent cone). *For given $u \in U_{\mathrm{ad}}$ we define the tangent cone at u by*

$$\mathcal{T}(u) := \Big\{ (\eta, h) \subset V \times L^p(\Omega) : \exists\, \{u_n\}_{n\in\mathbb{N}} \subset U_{\mathrm{ad}},\, \{t_n\} \subset \mathbb{R}^+ \text{ such that } \frac{u_n - u}{t_n} \rightharpoonup h \text{ in } L^p(\Omega) \quad \text{and} \quad \frac{S(u_n) - S(u)}{t_n} \rightharpoonup \eta \text{ in } V \Big\}.$$

Since the VI in (4.1) is uniquely solvable such that y is determined by u, this cone coincides with the standard tangent cone in finite dimensions, except that we replace

strong by weak convergence. Next consider the VI in (3.34) associated with the directional derivative of S at $\bar{u}$. Due to the coercivity of A, this VI does clearly not only admit a unique solution for right hand sides in $L^p(\Omega)$, but also for inhomogeneities in V^*. We denote the associated solution operator by $G : V^* \to V$, i.e.

$$\eta = G(h) \quad :\Longleftrightarrow \quad \eta \in \mathcal{K}(\bar{y}), \quad \langle A\eta,\, v - \eta\rangle \geq \langle h,\, v - \eta\rangle \quad \forall\, v \in \mathcal{K}(\bar{y}). \tag{4.3}$$

Furthermore, owing again to the coercivity of A this operator is Lipschitz continuous, i.e.

$$\|G(h_1) - G(h_2)\|_V \leq \frac{1}{\alpha}\,\|h_1 - h_2\|_{V^*} \quad \forall\, h_1, h_2 \in V^*, \tag{4.4}$$

where α is the coercivity constant of A. This enables us to show the following

THEOREM 4.4. *Suppose that the assumptions of Theorem 4.1 are fulfilled with a local optimum $\bar{u} \in U_{\mathrm{ad}}$ of (4.1). Then there holds*

$$\partial_y J(\bar{y}, \bar{u})\eta + \partial_u J(\bar{y}, \bar{u})h \geq 0 \quad \forall\, (\eta, h) \in \mathcal{T}(\bar{u}). \tag{4.5}$$

Proof. If $h_n \rightharpoonup h$ in $L^p(\Omega)$ and consequently $h_n \to h$ in V^*, then (4.4) gives $G(h_n) \to G(h)$ in V. Since $G(h) = S'_w(\bar{u}; h)$ for $h \in L^p(\Omega)$, this implies that $L^p(\Omega) \ni h \mapsto S'_w(\bar{u}; h) \in V$ is completely continuous. Now let $(\eta, h) \in \mathcal{T}(\bar{u})$ be arbitrary. Hence there is $\{u_n\} \in U_{\mathrm{ad}}$ so that $(u_n - \bar{u})/t_n \rightharpoonup h$ in $L^p(\Omega)$. As seen above, $S'_w(\bar{u}; .)$ is the solution operator of a VI of first kind with the cone $\mathcal{K}(\bar{y})$ as feasible set. Hence, $S'_w(\bar{u}; .)$ is positively homogeneous such that Theorem 4.1 yields

$$\partial_y J(\bar{y}, \bar{u}) S'_w\Big(\bar{u}; \frac{u_n - \bar{u}}{t_n}\Big) + \partial_u J(\bar{y}, \bar{u})\Big(\frac{u_n - \bar{u}}{t_n}\Big) \geq 0. \tag{4.6}$$

The complete continuity of $S'_w(\bar{u}; .)$ together with Lemma 4.2 implies

$$S'_w\Big(\bar{u}; \frac{u_n - \bar{u}}{t_n}\Big) \to S'_w(\bar{u}; h) = \eta \quad \text{in } V.$$

Due to the weak continuity of $\partial_u J(\bar{y}, \bar{u})$ the second addend in (4.6) converges to $\partial_u J(\bar{y}, \bar{u})h$, which completes the proof. ∎

5. Strong stationarity. In this section we aim at deriving optimality conditions which, in contrast to the ones presented in Section 4, also involve dual variables. Given the differentiability result and the Bouligand stationarity conditions in Theorem 4.1, we can follow the lines of [20]. For this purpose we have to require the following assumptions concerning the quantities in the optimal control problem 4.1:

ASSUMPTION 5.1. *We suppose that $U_{\mathrm{ad}} = L^2(\Omega)$. Moreover, J is continously Fréchet-differentiable from $V \times L^2(\Omega)$ to $\mathbb{R}$.*

In order to be able to utilize our differentiability result we furthermore assume the following:

ASSUMPTION 5.2. *Assume that $\bar{u}$ is a local optimum such that the associated state $\bar{y}$ and the associated slack variable $\bar{q}$ satisfy Assumptions 3.13 and 3.16.*

LEMMA 5.3. *Under Assumptions 5.1 and 5.2 there exists a $\bar{p} \in V$ such that*

$$\partial_y J(\bar{y}, \bar{u})G(h) - \langle \bar{p},\, h\rangle \geq 0 \quad \forall\, h \in V^*$$

with G as defined in (4.3).

Proof. By Theorem 4.1 and $S'_w(\bar{u}; h) = G(h)$ for $h \in L^2(\Omega)$, there holds

$$\partial_y J(\bar{y}, \bar{u}) G(h) + \partial_u J(\bar{y}, \bar{u}) h \geq 0 \quad \forall\, h \in L^2(\Omega). \tag{5.1}$$

which, together with (4.4), gives in turn

$$\partial_u J(\bar{y}, \bar{u}) h \leq \|\partial_y J(\bar{y}, \bar{u})\|_{V^*} \frac{1}{\alpha} \|h\|_{V^*} \quad \forall\, h \in L^2(\Omega).$$

Therefore, by the Hahn-Banach theorem, the linear functional $\partial_u J(\bar{y}, \bar{u}) : L^2(\Omega) \to \mathbb{R}$ can be extended to a linear and bounded functional on V^*, which we identify with a function $\bar{p} \in V$, i.e.

$$\langle \bar{p},\, h \rangle = -\partial_u J(\bar{y}, \bar{u}) h \quad \forall\, h \in L^2(\Omega).$$

The density of $L^2(\Omega) \hookrightarrow V^*$ in combination with (5.1) then gives the assertion. ∎

Next define $q \in V$ as solution of

$$\langle A^* q,\, v \rangle = \langle \partial_y J(\bar{y}, \bar{u}),\, v \rangle \quad \forall\, v \in V,$$

which is well defined because of the coercivity of A. Furthermore, we introduce the operator $\Pi : V \to \mathcal{K}(\bar{y})$ by

$$\Pi := G \circ A.$$

Note that Π can be interpreted as A-projection on $\mathcal{K}(\bar{y})$. It is straightforward to see the following properties of Π:

$$\begin{aligned} &\Pi \text{ as well as } I - \Pi \text{ are idempotent,} \\ &\Pi \circ (I - \Pi) = (I - \Pi) \circ \Pi = 0, \end{aligned} \tag{5.2}$$

and, as $\mathcal{K}(\bar{y})$ is a convex cone,

$$\langle A(I - \Pi)\xi,\, \Pi(\xi) \rangle = 0 \quad \forall\, \xi \in V. \tag{5.3}$$

Moreover by construction, we find $G = \Pi \circ A^{-1}$. Thus Lemma 5.3 implies for every $h \in V^*$ that

$$\begin{aligned} 0 &\leq \partial_y J(\bar{y}, \bar{u}) G(h) - \langle \bar{p},\, h \rangle \\ &= \langle G(h),\, A^* q \rangle - \langle A A^{-1} h,\, \bar{p} \rangle \\ &= \langle A\Pi(A^{-1}h),\, q - \bar{p} \rangle - \langle A(I - \Pi)(A^{-1}h),\, \bar{p} \rangle \\ &= \langle \Pi(A^{-1}h),\, A^*(q - \bar{p}) \rangle \\ &\quad - \langle A(I - \Pi)(A^{-1}h),\, \Pi(\bar{p}) \rangle - \langle A(I - \Pi)(A^{-1}h),\, (I - \Pi)(\bar{p}) \rangle. \end{aligned} \tag{5.4}$$

If we insert $h = A(I - \Pi)\bar{p} \in V^*$, then (5.2) and (5.3) yield

$$\langle A(I - \Pi)\bar{p},\, (I - \Pi)\bar{p} \rangle \leq 0.$$

The coercivity of A then implies $\bar{p} = \Pi(\bar{p})$ and thus $\bar{p} \in \mathcal{K}(\bar{y})$, i.e.

$$\begin{aligned} \bar{p}(x) &= 0 \quad \text{a.e., where } |\bar{q}(x)| < 1, \\ \bar{p}(x)\bar{q}(x) &\geq 0 \quad \text{a.e., where } |\bar{q}(x)| = 1 \text{ and } \bar{y}(x) = 0. \end{aligned}$$

Next we define $z \in V$ by

$$\langle Az,\, v\rangle = \langle v,\, A^*(\bar p - q)\rangle \quad \forall\, v \in V \tag{5.5}$$

and insert $h = A\Pi(z) \in V$ in (5.4). Together with (5.2), (5.5), and (5.3), we obtain in this way

$$0 \le \langle \Pi(z),\, A^*(q - \bar p)\rangle = -\langle Az,\, \Pi(z)\rangle = -\langle A\Pi(z),\, \Pi(z)\rangle$$

so that $\Pi(z) = G(Az) = 0$ by the coercivity of A. Consequently the definition of G in (4.3) leads to

$$\langle Az,\, v\rangle \le 0 \quad \forall\, v \in \mathcal{K}(\bar y) \quad \Longrightarrow \quad \langle A^*\bar p,\, v\rangle \le \langle A^* q,\, v\rangle = \langle \partial_y J(\bar y, \bar u),\, v\rangle \quad \forall\, v \in \mathcal{K}(\bar y).$$

By defining $\bar\mu := g'(\bar y) - A^*\bar p \in V^*$ we therefore arrive at

$$\begin{aligned} A^*\bar p &= \partial_y J(\bar y, \bar u) - \mu \quad \text{in } V^* \\ \langle \bar\mu,\, v\rangle &\ge 0 \quad \forall\, v \in \mathcal{K}(\bar y). \end{aligned}$$

All in all we have thus proven the following:

THEOREM 5.4. *Assume that Assumption 5.1 holds. Suppose moreover that $\bar u$ is a local optimum which satisfies Assumption 5.2. Then there exists an adjoint state $\bar p \in V$ and a multiplier $\mu \in V^*$ such that the following* strong stationarity system *is fulfilled:*

$$A\bar y + \bar q = \bar u \quad \text{in } V^* \tag{5.6a}$$

$$\bar q(x)\,\bar y(x) = |\bar y(x)|, \quad |\bar q(x)| \le 1 \quad \text{a.e. in } \Omega \tag{5.6b}$$

$$A^*\bar p = \partial_y J(\bar y, \bar u) - \mu \quad \text{in } V^* \tag{5.6c}$$

$$\bar p \in \mathcal{K}(\bar y), \quad \langle \bar\mu,\, v\rangle \ge 0 \quad \forall\, v \in \mathcal{K}(\bar y) \tag{5.6d}$$

$$\bar p + \partial_u J(\bar y, \bar u) = 0 \tag{5.6e}$$

with $\mathcal{K}(\bar y)$ as defined in (3.29).

REMARK 5.5. *A comparable result for optimal control problems governed by VIs of the first kind is known as strong stationarity conditions, see [11]. This is why we have chosen the same terminology here.*

REMARK 5.6. *We compare the optimality system (5.6) with results from [4] obtained via regularization and subsequent limit analysis. The optimality system obtained in [4] coincides with (5.6) except that (5.6d) is replaced by*

$$\langle \bar\mu,\, \bar p\rangle \ge 0, \quad \langle \bar\mu,\, \bar y\rangle = 0. \tag{5.7}$$

However, thanks to the definition of $\mathcal{K}(\bar y)$ and $\pm\bar y \in \mathcal{K}(\bar y)$, these relations are an immediate consequence of (5.6d). The optimality system in (5.6) is therefore sharper compared to the one obtained via regularization. We point out however that the analysis in [4] does not require the restrictive Assumptions 3.16 and 3.13 and in addition applies to more general VIs of the second kind.

6. An inexact trust-region algorithm. In this section we propose an inexact trust-region algorithm for the solution of the finite-dimensional optimization problem:

$$\min J(y,u) \tag{6.1}$$

$$\text{subject to: } \langle Ay, v-y\rangle + g|v|_1 - g|y|_1 \geq \langle u, v-y\rangle, \text{ for all } v \in \mathbb{R}^n, \tag{6.2}$$

with $g > 0$. The main difficulty of the method consists in computing a descent direction along which the algorithm has to perform the next step. In the case of an empty biactive set, the derivative information is given by (2.18)-(2.19). From the latter, existence of an adjoint state can be proved and an adjoint calculus may be performed.

Since the information so obtained does not necessarily correspond to an element of the subdifferential, in case of a non-empty biactive set, we apply a trust-region scheme to provide robust iterates. In this context the adjoint related gradient is considered as an inexact version of a descent direction. Since in the applications we focus on, the biactive set is either empty or very small, such an approach is justified from the numerical point of view.

Indeed, by assuming that the biactive set

$$B = \{i : y_i = 0, |q_i| = 1\},$$

is empty, the solution operator is Gâteaux differentiable and the directional derivative $\eta = S'(u)h$ corresponds to the solution of the following system of equations:

$$\eta_i = 0 \text{ for } i : y_i = 0,$$

$$\sum_{j:y_j \neq 0} A_{i,j}\eta_j = h_i \text{ for } i : y_i \neq 0.$$

To simplify the description of the algorithm, we confine ourselves to a quadratic cost functional of the form $J(y,u) = 1/2\,\|y-z\|^2 + \alpha/2\,\|u\|^2$, where $\|\,.\,\|$ denotes the Euclidian norm and $z \in \mathbb{R}^n$ is a given desired state. Considering the reduced cost functional

$$j(u) = \frac{1}{2}\|S(u) - z\|^2 + \frac{\alpha}{2}\|u\|^2,$$

the directional derivative is given by

$$j'(u)h = (S(u) - z, S'(u)h) + \alpha(u,h) = \sum_i (y_i - z_i)\eta_i + \alpha \sum_i u_i h_i.$$

Let us recall that the inactive set is given by $\mathcal{I} := \{i \in \{1,\dots,n\} : y_i \neq 0\}$. By reordering the indices such that the active and inactive ones occur in consecutive order, and defining the adjoint state $p \in \mathbb{R}^n$ as the solution to the system:

$$\begin{pmatrix} I & 0 \\ 0 & A_{\mathcal{I}}^T \end{pmatrix} p = y - z,$$

where $A_{\mathcal{I}}$ corresponds to the block of A with indexes i, j such that $y_i \neq 0, y_j \neq 0$, we obtain that

$$j'(u)h = \sum_{i\in\mathcal{I}} p_i h_i + \alpha \sum_i u_i h_i$$

or, equivalently, $j'(u) = \begin{cases} \alpha u_i & \text{if } i \notin \mathcal{I} \\ p_i + \alpha u_i & \text{if } i \in \mathcal{I}. \end{cases}$

Before stating the trust-region algorithm, let us introduce some notation to be used. The quadratic model of the reduced cost function is given by

$$q_k(s) = j(u_k) + g_k^T s + \frac{1}{2} s^T H_k s,$$

where $g_k = j'(u_k)$ and H_k is a matrix with second order information, obtained with the BFGS method. The trust region radius is denoted by Δ_k and the actual and predicted reductions are given by

$$ared_k(s^k) := j(u_k) - j(u_k + s^k) \text{ and } pred_k(s^k) = j(u_k) - q_k(s^k), \text{ respectively.}$$

The quality indicator is computed by

$$\rho_k(s^k) = \frac{ared_k(s^k)}{pred_k(s^k)}.$$

The resulting trust region algorithm (of dogleg type) is given through the following steps:

Trust region algorithm.

1. Choose the parameter values $0 < \eta_1 < \eta_2 < 1$, $0 < \gamma_0 < \gamma_1 < 1 < \gamma_2$, $\Delta_{min} \geq 0$.
2. Choose the initial iterate $x_0 \in \mathbb{R}^n$ and the trust region radius $\Delta_0 > 0$, $\Delta_0 \geq \Delta_{min} \geq 0$.
3. Compute the Cauchy step $s_c^k = -t^* g_k$, where

$$t^* = \begin{cases} \dfrac{\Delta_k}{||g_k||}, & \text{if } \; g_k^\top H_k g_k \leq 0 \\ \min\left(\dfrac{||g_k||^2}{g_k^\top H_k g_k}, \dfrac{\Delta_k}{||g_k||} \right), & \text{if } \; g_k^\top H_k g_k > 0 \end{cases}$$

   and the Newton step $s_n^k = -H_k^{-1} g_k$. If s_n^k satisfies the fraction of Cauchy decrease:

   $$\exists \delta \in (0,1] \text{ and } \beta \geq 1 \text{ such that } \|s^k\| \leq \beta \Delta_k \text{ and } pred_k(s^k) \geq \delta \; pred_k(s_c^k).$$

   then $s^k = s_n^k$, else $s^k = s_c^k$.
4. If $\varrho_k(s^k) > \eta_2$, then

   $$u_{k+1} = u_k + s_k, \quad \Delta_{k+1} \in [\Delta_k, \gamma_2 \Delta_k]$$

   Else if $\varrho_k(s^k) \in (\eta_1, \eta_2)$, then

   $$u_{k+1} = u_k + s_k, \quad \Delta_{k+1} \in [\max(\Delta_{min}, \gamma_1 \Delta_k), \Delta_k]$$

   Else if $\varrho_k(s^k) \leq \eta_1$, then

   $$u_{k+1} = u_k, \quad \Delta_{k+1} \in [\gamma_0 \Delta_k, \gamma_1 \Delta_k]$$

   Repeat until stopping criteria.

6.1. Example. We consider as test example the following finite-dimensional optimization problem:

$$\min\ J(y,u) = \frac{1}{2}\|y-z\|^2 + \frac{\alpha}{2}\|u\|^2 \tag{6.3}$$

$$\text{subject to: } \langle Ay, v-y\rangle + g|v|_1 - g|y|_1 \geq \langle u, v-y\rangle, \text{ for all } v\in\mathbb{R}^n, \tag{6.4}$$

where A corresponds to the finite differences discretization matrix of the negative Laplace operator in the two dimensional domain $\Omega =]0,1[^2$, $z = 10\sin(5x_1)\cos(4x_2)$ stands for the desired state and α and g are positive constants. It is expected that as g becomes larger the solution becomes sparser.

For solving (6.4) within the trust region algorithm a semismooth Newton method is used. The method is built upon a Huber-type regularization of the l_1-norm, cf. [14, eq. (7.14)], and the use of dual information. Specifically, we consider the solution of the regularized inequality:

$$Ay + q = u \tag{6.5}$$

$$q - h_\gamma(y) = 0, \tag{6.6}$$

where $(h_\gamma(y))_i = g\frac{\gamma y_i}{\max(g,\gamma|y_i|)}$. Considering a generalized derivative of the max function, the following system has to be solved in each semismooth Newton iteration:

$$A\delta_y + \delta_q = u - Ay - q \tag{6.7}$$

$$\delta_q - \frac{\gamma\delta_y}{\max(g,\gamma|y|)} + \operatorname{diag}(\chi_{\mathcal{I}_\gamma})\frac{\gamma^2 y^T\delta_y}{\max(g,\gamma|y|)^2}\frac{y}{|y|} = -q + h_\gamma(y), \tag{6.8}$$

where $\left(\chi_{\mathcal{I}_\gamma}\right)_i := \begin{cases} 1 & \text{if } \gamma|y_i| \geq g, \\ 0 & \text{if not.} \end{cases}$, $\max(g,\gamma|y|) := (\max(g,\gamma|y_1|),\ldots,\max(g,\gamma|y_n|))^T$ and the division is to be understood componentwise. By using dual information in the iteration matrix (as in [13], [5]) the following modified version of (6.8) is obtained:

$$\delta_q - \frac{\gamma\delta_y}{\max(g,\gamma|y|)} + \operatorname{diag}(\chi_{\mathcal{I}_\gamma})\frac{\gamma^2 y^T\delta_y}{\max(g,\gamma|y|)^2}\frac{q}{\max(g,|q|)} = -q + h_\gamma(y). \tag{6.9}$$

This leads to a globally convergent iterative algorithm, which converges locally with superlinear rate.

The used trust region parameter values are $\eta_1 = 0.25$, $\eta_2 = 0.75$, $\gamma_1 = 0.5$, $\gamma_2 = 1.5$ and $\beta = 1$. For the parameter values $\alpha = 0.0001$ and $g = 15$, and the mesh size step $h = 1/80$, the algorithm requires a total number of 35 iterations to converge, for a stopping criteria given by $\|u_{k+1} - u_k\| \leq 1e-4$. The optimized state is shown in Figure 6.1, where a large zone where the state takes value zero can be observed.

The algorithm was also tested for other values of the parameters α and g, yielding the convergence behaviour registered in Table 6.1. Although the considered derivative information was inexact, the trust-region approach yields convergence in a relatively small number of iterations.

Further descent type directions to be used in the context of the trust-region methodology, as well as the convergence theory of the combined approach, will be investigated in future work.

Appendix A. Directional derivative of the L^1-norm.

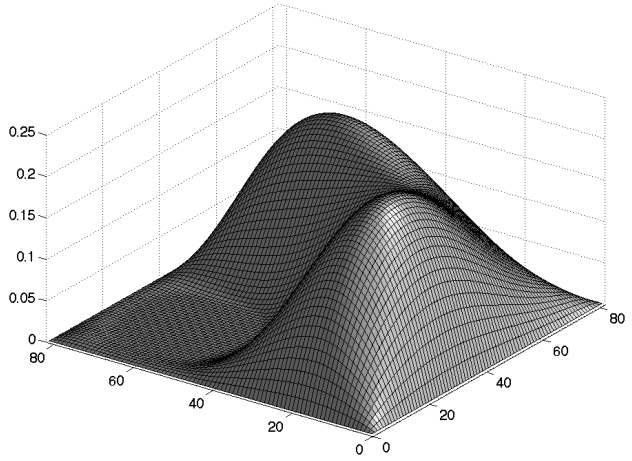

FIG. 6.1. *Optimized state: on the left corner the sparse structure of the solution can be observed.*

α \ g	1	5	10	15
0,1	20	28	53	-
0,01	23	24	27	32
0,001	33	48	54	31
0,0001	69	70	62	34

TABLE 6.1
Number of trust-region iterations for different α and g values. Mesh size step $h = 1/40$.

Proof of Lemma 3.5. We consider the mapping

$$g : L^1(\Omega) \ni y \mapsto \int_\Omega |y(x)|\,\varphi(x)\,dx \in \mathbb{R}.$$

It is easily seen that g is Lipschitz continuous. Moreover, for arbitrary $y, \eta \in L^1(\Omega)$, the directional differentiability of $\mathbb{R} \ni r \mapsto |r| \in \mathbb{R}$ yields

$$\frac{|y(x) + t_n\eta(x)| - |y(x)|}{t_n} \to \mathrm{abs}'\big(y(x); \eta(x)\big),$$

and, since almost all points in Ω are common Lebesgue points of y and η, this pointwise convergence holds almost everywhere in Ω. Due to

$$-2|\eta(x)| \le \frac{|y(x) + t_n\eta(x)| - |y(x)|}{t_n} - \mathrm{abs}'\big(y(x); \eta(x)\big) \le 2|\eta(x)| \quad \text{a.e. in } \Omega,$$

Lebesgue dominated convergence theorem thus gives

$$\frac{|y + t_n\eta| - |y|}{t_n} \to \mathrm{abs}'(y; \eta) \text{ in } L^1(\Omega),$$

which in turn implies the directional differentiability of g with

$$g'(y; \eta) = \int_\Omega \mathrm{abs}'(y(x); \eta(x))\,\varphi(x)\,dx.$$

Consequently g is Hadamard-differentiable and hence

$$\int_\Omega \Big(\frac{|y_n| - |y|}{t_n} - \mathrm{abs}'(y;\eta) \Big) \varphi \, dx = \frac{g(y + t_n\eta + r(t_n)) - g(y)}{t_n} - g'(y;\eta) \to 0,$$

since

$$r(t_n) := y_n - y - t_n\eta$$

so that $\|r(t_n)\|_{L^1(\Omega)} = o(t_n)$ thanks to (3.8) and the compact embedding $V \hookrightarrow L^1(\Omega)$. ∎

Appendix B. Boundedness for functions in $H^1(\Omega)$.

For convenience of the reader, we prove Lemma 3.8. The arguments are classical and go back to [15].

Proof of Lemma 3.8. The truncated function defined in (3.17) is equivalent to

$$w_k(x) = w(x) - \min\big((\max(w(x), -k), k\big)$$

and therefore [15, Theorem A.1] implies $w_k \in V$.

It remains to verify the L^∞-bound in (3.19). If $d = 1$, then the assertion follows directly from (3.18) and the Sobolev embedding $H^1(\Omega) \hookrightarrow L^\infty(\Omega)$.

So assume that $d \geq 2$. Then let $k \geq 0$ be given and set $A(k) := \{x \in \Omega \mid |w(x)| \geq k\}$. Note that $w_k(x) = 0$ a.e. in $\Omega \setminus A(k)$. Next let $h \geq k$ be arbitrary so that $w(x) \geq h \geq k$ a.e. in $A(h)$. Then Sobolev embeddings give that

$$\begin{aligned} \|w_k\|^2_{H^1(\Omega)} \geq c\,\|w_k\|^2_{L^m(\Omega)} &= c\Big(\int_{A(k)} \big||w| - k\big|^m dx \Big)^{2/m} \\ &\geq c \int_{A(h)} (h-k)^m dx^{2/m} = c\,(h-k)^2 |A(h)|^{2/m}, \end{aligned} \tag{B.1}$$

where $m = 2d/(d-2)$, see e.g. ... On the other hand, (3.18) implies

$$\alpha\,\|w_k\|^2 \leq \int_{A(k)} f\,w_k\,dx \leq \|f\|_{L^{m'}(A(k))}\,\|w_k\|_{L^m(A(k))} \leq c\,\|f\|_{L^{m'}(A(k))}\,\|w_k\|_{H^1(\Omega)},$$

where m' is the conjugate exponent to m, i.e. $1/m + 1/m' = 1$. Note that

$$m' = \frac{m}{m-1} = \frac{d}{d/2+1} \leq \frac{d}{2} < p, \quad \text{if } d \geq 2,$$

and thus $f \in L^{m'}(\Omega)$ by the assumption on f in Lemma 3.8. Together with Young's inequality, then Hölder's inequality yields

$$\|w_k\|^2 \leq c\Big(\int_{A(k)} |f|^{m'}\,dx \Big)^{2/m'} \leq c\,\|f\|^2_{L^p(\Omega)}\,|A(k)|^{2r/m'} \tag{B.2}$$

with $r = p/(p-m') \geq 1$ so that $r' = r/(r-1) = p/m'$. By setting

$$s = \frac{m}{m'}\,r = \frac{p}{(m'-1)(p-m')} \tag{B.3}$$

we infer from (B.1) and (B.2) that

$$|A(h)|^{2/m} \le c\,\|f\|^2_{L^p(\Omega)}\,\frac{1}{(h-k)^2}\left(|A(h)|^{2/m}\right)^s \quad \text{for all } h > k \ge 0. \tag{B.4}$$

Since $m > 2$, we have $m' < 2$ and therefore $(m'-1)(p-m') < p - m' < p$ such that (B.3) gives in turn $s > 1$. In this case, according to [15, Lemma B.1], it follows from (B.4) that the nonnegative and non-increasing function $\mathbb{R} \ni h \mapsto |A(h)|^{2/m} \in \mathbb{R}$ admits a zero at

$$h^* = 2^{s/(s-1)}\sqrt{c\,|\Omega|^{2(s-1)/m}}\,\|f\|_{L^p(\Omega)}.$$

By definition, $|A(h^*)| = 0$ is equivalent to $|w(x)| \le h^*$ a.e. in Ω, which yields the assertion. □

Acknowledgement. The authors would like to thank Gerd Wachsmuth (TU Chemnitz) for his hint concerning strong stationarity.

This work was supported by a DFG grant within the Collaborative Research Center SFB 708 (*3D-Surface Engineering of Tools for Sheet Metal Forming – Manufacturing, Modeling, Machining*), which is gratefully acknowledged.

REFERENCES

[1] V. Barbu. *Analysis and Control of nonlinear infinite dimensional systems*. Academic Press, New York, 1993.

[2] Maïtine Bergounioux. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM Journal on Control and Optimization*, 36(1):273–289, 1998.

[3] Joseph Frédéric Bonnans and Eduardo Casas. An extension of Pontryagin's principle for state-constrained optimal control of semilinear elliptic equations and variational inequalities. *SIAM Journal on Control and Optimization*, 33(1):274–298, 1995.

[4] Juan Carlos De los Reyes. Optimal control of a class of variational inequalities of the second kind. *SIAM Journal on Control and Optimization*, 49:1629–1658, 2011.

[5] Juan Carlos De los Reyes. Optimization of mixed variational inequalities arising in flows of viscoplastic materials. *Computational Optimization and Applications*, 52:757–784, 2012.

[6] Juan Carlos De los Reyes, Roland Herzog, and Christian Meyer. Optimal control of static elastoplasticity in primal formulation. Ergebnisberichte des Instituts für Angewandte Mathematik 474, TU Dortmund, 2013.

[7] Juan Carlos De los Reyes and Carola-Bibiane Schönlieb. Image denoising: Learning the noise model via nonsmooth PDE-constrained optimization. *Inverse Problems & Imaging*, 7(4), 2013.

[8] Lawrence C. Evans. *Partial differential equations: Graduate studies in Mathematics*. American Mathematical Society, 1998.

[9] H Gajewski, K Gröger, and K Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Akademie-Verlag, Berlin, 1978.

[10] Roland Herzog, Christian Meyer, and Gerd Wachsmuth. B-and strong stationarity for optimal control of static plasticity with hardening. *SIAM Journal on Optimization*, 23(1):321–352, 2013.

[11] M. Hintermüller and I. Kopacka. Mathematical programs with complementarity constraints in function space: C-and strong stationarity and a path-following algorithm. *SIAM Journal on Optimization*, 20(2):868–902, 2009.

[12] M. Hintermüller, B. Mordukhovich, and T. Surowiec. Several approaches for the derivation of stationarity conditions for elliptic mpecs with upper-level control constraints. *Math. Prog. A*, to appear.

[13] M. Hintermüller and G. Stadler. An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM J. Sci. Comput.*, 28(1):1–23 (electronic), 2006.

[14] Peter J. Huber. *Robust statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., Hoboken, NJ, 1981.

[15] David Kinderlehrer and Guido Stampacchia. *An introduction to variational inequalities and their applications*, volume 31. SIAM, 2000.

[16] K. Kunisch and D. Wachsmuth. Path-following for optimal control of stationary variational inequalities. *Computational Optimization and Applications*, 51:1345–1373, 2012.

[17] K. Kunisch and D. Wachsmuth. Sufficient optimality conditions and semi-smooth newton methods for optimal control of stationary variational inequalities. *ESAIM: Control, Optimisation and Calculus of Variations*, 180:520–547, 2012.

[18] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical programs with equilibrium constraints*. Cambridge University Press, Cambridge, 1996.

[19] F. Mignot. Controle dans les inéquations variationelles elliptiques. *Journal of Functional Analysis*, 22:130–185, 1976.

[20] F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM J. Control Optim.*, 22(3):466–476, 1984.

[21] J. V. Outrata. A generalized mathematical program with equilibrium constraints. *SIAM J. Control Optim.*, 38(5):1623–1638 (electronic), 2000.

[22] Jiří Outrata, Jiří Jarušek, and Jana Stará. On optimality conditions in control of elliptic variational inequalities. *Set-Valued and Variational Analysis*, 19(1):23–42, 2011.

[23] A. Schiela and D. Wachsmuth. Convergence analysis of smoothing methods for optimal control of stationary variational inequalities. *ESAIM Math. Model. Numer. Anal.*, 47(3):771–787, 2013.