# A Bundle-Free Implicit Programming Approach for a Class of MPECs in Function Space

M. Hintermüller* and T. Surowiec†

**Abstract**

Using a standard first-order optimality condition for nonsmooth optimization problems, a general framework for a descent method is developed. This setting is applied to a class of mathematical programs with equilibrium constraints in function space from which a new algorithm is derived. Global convergence of the algorithm is demonstrated in function space and the results are then illustrated by numerical experiments.

**Key words.** elliptic variational inequality, elliptic MPEC, implicit programming, optimal control of variational inequalities, nonsmooth optimization

**AMS subject classifications.** 49J52, 49K20, 49K21, 65Kxx, 90C33, 49M05

## 1 Introduction

In recent times, there has been a significant interest in the optimal control of variational inequalities. Following the terminology from mathematical programming, such models have become known as mathematical programs with equilibrium constraints (MPECs) in function space.

Much is known about the theoretical aspects of these models, see, e.g., the monograph [4] as well as [5, 6, 32]. However, efforts to derive efficient numerical methods have been more recent, see the discussion below.

We sketch a standard technique to motivate our approach. Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $n \in \{1, 2, 3\}$, with sufficiently regular boundary $\partial\Omega$. Let $\alpha > 0$, and fix $y_d, f \in L^2(\Omega)$. We then consider:

$$\begin{aligned} &\min \tfrac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \tfrac{\alpha}{2}||u||^2_{L^2(\Omega)} \text{ over } (u, y) \in L^2(\Omega) \times H^1_0(\Omega) \\ &\text{subject to (s.t.) } y \in \operatorname{argmin}_z\{\tfrac{1}{2}\textstyle\int_\Omega |\nabla z|^2 dx - \int_\Omega (u + f) z dx \,|\, z \geq 0,\ a.e.\, \Omega\}. \end{aligned} \tag{1}$$

Despite convexity and differentiability of the data, the bilevel structure of (1) generally guarantees a lack of linearity, convexity, and differentiability of the control-to-state mapping. Moreover, any reformulation as an NLP would involve a degenerate constraint system.

*Dept. of Mathematics, Humboldt University of Berlin and Dept. of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, E-Mail:hint@math.hu-berlin.de

†Dept. of Mathematics, Humboldt University of Berlin, E-Mail:surowiec@math.hu-berlin.de

By penalizing the "lower-level" constraint $y \geq 0,\ a.e.\,\Omega$ with a smooth parameter-dependent Nemytski-type operator induced by $\rho_\varepsilon : \mathbb{R} \to \mathbb{R}$, one obtains a sequence of penalized regularized nonlinear problems depending on the penalization parameter $\gamma > 0$ and regularization parameter $\varepsilon > 0$ defined by

$$\begin{array}{l} \min \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)} \text{ over } (u, y) \in L^2(\Omega) \times H^1_0(\Omega) \\ \text{s.t. } -\Delta y + \gamma \rho_\varepsilon(-y) = u + f. \end{array} \tag{2}$$

This results in the optimal control of a semilinear partial differential equation; a topic which has been more deeply studied and developed, see e.g., [42] and the references therein.

There are several drawbacks to this approach:

- In the limit, sequences of stationary points for (2) do not in general satisfy the first-order optimality conditions for (1) as derived by Mignot and Puel in [32], so called strongly stationary points, (and rederived using alternative methods in [34]); rather to a weaker set of conditions, i.e. typically a form of C-stationarity (see Section 3 below).
- In non-redundant cases, i.e., when $y_{\gamma,\varepsilon}$ is not feasible for finite $\gamma$ and $\varepsilon > 0$, one never actually obtains a true feasible point of (1).
- Even with path-following one must solve a sequence of (large-scale, upon discretization) nonlinear systems or nonlinear nonconvex optimization problems (one for each parameter pair $(\gamma, \varepsilon)$, as $\gamma \to +\infty$, $\varepsilon \downarrow 0$); see, e.g., [20, 21, 39].

Interestingly, it is not always the case that the solution mapping for the lower-level problem in (1) is nonsmooth at *every* point. Indeed, if the control-to-state mapping is Lipschitz, then the operator is Gâteaux differentiable up to small "$\sigma$-porous" sets, see [2, 35]. In fact, the MPEC may be locally "smooth" near a solution. In such cases, this "adapted penalty" approach regularizes a smooth problem.

The adapted penalty approach has a long history going back to early work of Yvon [45, 46] and J.L. Lions [27, 28]. It was used extensively by Barbu in [4] to derive optimality conditions as well as in [5, 6, 24]. The method is still in use, as can be seen, e.g., in [18, 11] (in different, but somewhat related contexts). In [24], a numerical method was provided. However, as noted in Remark 6.1(iii), their algorithm appears to converge only when biactivity is absent. This implicitely states that the algorithm is likely to converge only if there is no nonsmoothness present, as reflected in their numerical experiments. In [21], a semismooth-Newton-based multigrid method was developed guaranteeing "$\mathcal{E}$-almost-C-stationarity" in the limit when approaching the limit point by a sequence of stationary points only. Afterwards in [39], under the assumption that one works with sequences of stationary points that *converge to a (local) minimizer of* (1), it was shown that the limit point of the approximation method will satisfy "C-stationary" conditions. Actual convergence rates for the approximation error were also obtained, assuming each of the stationary points is a global solution of the relaxed problem. Nevertheless, the currently available numerical approaches for the optimal control of semilinear PDEs generally cannot guarantee these additional assumptions.

A related reformulation of (1) involves the introduction of a slack variable $\lambda \in H^{-1}(\Omega)$ such that one considers the complementarity problem

$$-\Delta y - \lambda = u + f, (\text{ in } H^{-1}(\Omega)),\ y \geq 0,\ a.e.\,\Omega,\ \lambda \geq 0, (\text{ in } H^{-1}(\Omega)),\ \langle \lambda, y \rangle = 0.$$

instead of the variational inequality. Relaxation techniques in this context, see e.g. [37, 40] (n-dim) and [20] ($\infty$-dim), typically involve replacing $\langle \lambda, y \rangle = 0$ by $\langle \lambda, y \rangle \leq \gamma^{-1}$ and proceeding as above. Another possibility would be to use a so-called NCP-function, however, this is only a viable option in certain regular cases and might still be problematic from an algorithmic point-of-view in function space settings.

In contrast, we develop here an iterative framework in infinite dimensional Hilbert spaces applicable to problems such as (1) *without regularizing* the MPEC. While in the worst case the new method yields a form of a C-stationary point, it may provide more desirable results as seen in the numerical experiments, such as strong- or B-stationary points. The approach is inspired by results in finite-dimensions, cf. the monographs [33] and [29] and the references therein, as well as the papers [12, 13, 40]. Our results combine concepts from generalized differentiation and classical approximation results for regularizing variational inequalities as found throughout [14].

The paper is structured as follows. In Section 2, a class of non-smooth optimization problems that encompasses a large class of MPECs is introduced. After recalling a primal first-order optimality condition and noting its equivalence to the zero solving a particular auxiliary problem, a regularized auxiliary problem is introduced whose solutions amount to descent directions for the original problem. An associated line search is introduced, which, in combination with the regularized auxiliary problem, is used to define a conceptual algorithm. In Section 3, we reduce our focus to the study of problems similar to (1). After developing some results for the approximation of the directional derivative of the solution mapping of a variational inequality, we derive an implementable algorithm and discuss its convergence properties. In order to ensure global convergence in all possible settings, the algorithm is augmented by an additional interior loop. In Section 4, we illustrate the results from Section 3 via numerical experiments.

# 2 A General Framework and a Conceptual Algorithm

We begin with the following abstract optimization problem

$$\begin{aligned} &\min J(u, y) \text{ over } (u, y) \in H \times V \\ &s.t. \quad y = S(Bu), \end{aligned} \tag{3}$$

where $V$ and $H$ are Hilbert spaces and $V \hookrightarrow H \equiv H^* \hookrightarrow V^*$ represents a Gelfand triple, $J : H \times V \to \mathbb{R}$ is continuously Fréchet differentiable,bounded from below, and weakly lower semi-continuous; $B : H \to H$ is a bounded linear operator, i.e., $B \in \mathcal{L}(H)$, and is compact from $H$ to $V^*$, $S : V^* \to V$ is Lipschitz continuous and Hadamard directionally differentiable. Moreover, we assume that $J(\cdot, S(B\cdot)) : H \to \mathbb{R}$ is coercive.

Recall that a mapping $\Phi$ between Banach spaces $X$ and $Y$ is said to be Hadamard directionally differentiable, provided the following limits exist for all $x \in X$:

$$\lim_{\substack{t \to 0^+ \\ h' \to_X h}} t^{-1}(\Phi(x + th') - \Phi(x)), \ \forall h \in X. \tag{4}$$

Additionally, if $\Phi$ is Lipschitz continuous and Hadamard directionally differentiable, then the Hadamard directional derivative coincides with the standard (Gâteaux) directional derivative

$\Phi'(x,\cdot)$, which is also Lipschitz continuous (cf. Prop. 2.49 in [7]). In particular, there exists a small-$o$ function, dependent on $x$, such that $o(t)/t \to 0$ as $t \to 0^+$ with

$$\Phi(x+th) = \Phi(x) + t\Phi'(x;h) + o(t).$$

We refer the reader to Chapter 2 in [7] as well as [41] for more on this and related concepts.

It is often convenient to consider (3) in reduced form:

$$\min \mathcal{J}(u) := J(u, S(Bu)) \text{ over } u \in H \tag{5}$$

We always use $\mathcal{J}$ to represent the reduced objective functional, regardless of the setting. Under these data assumptions, if $(\bar{u}, \bar{y})$ is a solution to (3), then (cf. [22, Theorem 2.1])

$$J'(\bar{u};h) = J_y(\bar{u},\bar{y})S'(B\bar{u};Bh) + J_u(\bar{u},\bar{y})h \geq 0, \forall h \in H. \tag{6}$$

When $S$ is the solution mapping of a variational inequality, then (3) represents an MPEC. In such a case, feasible points $(\bar{u}, \bar{y})$ to (3) that satisfy (6) are referred to as B-stationary points. By definition B-stationary points are good candidates for being at least locally optimal solutions to (3). This property is lacking by the majority of the available stationarity concepts for MPECs. Regardless of the true structure of $S$, we will henceforth refer to feasible points $(\bar{u}, \bar{y})$ to (3) that satisfy (6) as being B-stationary.

In order to compute B-stationary points, we need to reformulate (6). Let $y = S(u)$. Then (6) leads to the auxiliary optimization problem (AP):

$$\min J_y(u,y)S'(Bu;Bh) + J_u(u,y)h \text{ over } h \in H, \tag{7}$$

which has the solution $0 \in H$ if $(u, y)$ is a locally optimal solution of (3). Nevertheless, solving (7) can be difficult. In particular, if $(u, y)$ does not solve (3), then (7) need not have a solution. Therefore, we add the quadratic form $q : H \times H \to \mathbb{R}$ to the objective functional in (7) with $q(h, h) := (Qh, h)_H$ for all $h \in H$ and $Q \in \mathcal{L}(H)$ uniformly coercive and bounded and refer to this problem as the regularized auxiliary problem (RAP):

$$\min F(h) := \tfrac{1}{2}q(h,h) + J_y(u,y)S'(Bu;Bh) + J_u(u,y)h \text{ over } h \in H. \tag{8}$$

In order to show that (8) has a solution, we need the following technical lemma.

**Lemma 2.1.** *Let $X$ be a real Banach space and $K \subset X$ be a proper closed cone. Define the positively homogeneous functional $P : X \to \mathbb{R}$ such that $K \subset \operatorname{dom} P$. Finally, for the bilinear form $b : X \times X \to \mathbb{R}$, we assume there exist constants $c_1$ and $c_2$ with $c_2 \geq c_1 > 0$ such that*

$$c_1||h||_X^2 \leq b(h,h), \ \forall h \in X \quad \text{and} \quad b(h,h') \leq c_2||h||_X \cdot ||h'||_X, \ \forall h, h' \in X.$$

*Then zero minimizes $P + b$ over $K$ if and only if $0 \in X$ minimizes $P$ over $K$.*

*Proof.* If zero minimizes $P$ over $K$, then, by the assumptions on $b$, zero must also minimize $P+b$ over $K$. Conversely, suppose zero minimizes $P+b$ and that $\exists h \neq 0$ in $K$ with $0 > P(h)$. By assumption, $0 \leq P(h) + tb(h,h)$, $\forall t > 0$. Then by taking $t = \alpha\frac{-P(h)}{b(h,h)}$, $\alpha > 0$, it holds for all $\alpha < 1$ that $P(h) + tb(h,h) = (1-\alpha)P(h) < 0$; a contradiction. The assertion follows. $\square$

In light of Lemma 2.1, we have the following corollary.

**Corollary 2.2** (**Consistency and Equivalence of B-Stationarity using the RAP**). *If $(\bar{u}, \bar{y})$ is a locally optimal solution to (3), then 0 solves (8). Moreover, $0 \in H$ solves (7) if and only if $0 \in H$ solves (8).*

**Proposition 2.3** (**Properties of the RAP**). *Let $(u, y)$ be a feasible point for (3). Then the following properties hold.*

1. *The corresponding RAP (8) has an optimal solution $\bar{h} \in H$.*
2. *If $\bar{h} \neq 0$, then it is a proper descent direction for the objective function $\mathcal{J}(u)$ in (5).*
3. *If the directional derivative $S'(Bu; B\cdot)$ is bounded and linear, then $\bar{h}$ is unique.*

*Proof.* Let $(u, y)$ be feasible for (3) and consider (8). Then $F$ is weakly lower semicontinuous, since $q$ is convex and continuous from $H \times H \to \mathbb{R}$, $B : H \to V^*$ is completely continuous, $S'(Bu; \cdot) : V^* \to V$ is Lipschitz continuous, and $J_u(u, y)\cdot$ is continuous and linear from $H \to \mathbb{R}$. It remains to show that $F$ is coercive with respect to the $H$-norm in order to prove the existence of a solution $\bar{h}$ to (8) (cf. Theorem 3.2.2 in [3]). Using the coercivity of $q$, there exists a constant $c_1 > 0$ such that

$$||h||_H^{-1}(\frac{1}{2}q(h,h) + J_y(u,y)S'(Bu;Bh) + J_u(u,y)h) \geq$$
$$c_1||h||_H + ||h||_H^{-1}J_y(u,y)S'(Bu;Bh) + ||h||_H^{-1}J_u(u,y)h.$$

Since the terms $J_y(u, y)S'(Bu; B\frac{h}{||h||_H}) + J_u(u, y)\frac{h}{||h||_H}$ are bounded for all sequences $\{h\} \subset H$, they cannot diverge towards $-\infty$. It follows that (8) has a solution, which is unique when $S'(Bu; B\cdot)$ is bounded and linear, due to the strong convexity of the resulting problem.

Finally, if $\bar{h}$ solves (8), but $\bar{h} \neq 0$. Then $F(\bar{h}) \leq F(0) = 0$. Given the assumptions on $q$, we then deduce

$$J_y(u, y)S'(Bu; B\bar{h}) + J_u(u, y)\bar{h} \leq -c_1||\bar{h}||_H^2 < 0. \tag{9}$$

This completes the proof. □

Given the properties of the RAP (8) proved in Proposition 2.3, we can formulate a conceptual algorithm. We use the notation $(u, y) \in \operatorname{gph} S$ to represent $y \in S(u)$.

**Algorithm 1** Conceptual Algorithm

**Input:** $u_0 \in H$; $\epsilon \geq 0$; $k := 0$

1: Set $y_0 = S(Bu_0)$.
2: Solve (8) with $(u, y) = (u_0, y_0)$ to obtain $h_1$.
3: **while** $||h_k||_H > \epsilon$ **do**
4: Compute $(u_{k+1}, y_{k+1})$ via a line search in direction $h_{k+1}$.
5: Solve (8) with $(u, y) = (u_{k+1}, y_{k+1})$ to obtain $h_{k+1}$.
6: Set $k := k + 1$.
7: **end while**

The stopping criterion for Algorithm 1 relates to the B-stationarity of the point $(u_k, y_k)$. This contrasts with standard stopping criteria in nonsmooth optimization, see e.g., [30]. Nevertheless, it should be kept in mind that the conceptual algorithm serves as a tool for the development of an implementable algorithm, as finding $h_k$ is nontractable in truly nonsmooth settings. Proposition 2.3 2. allows us to use a standard line search; we let $\varphi(\tau) := \mathcal{J}(u+\tau h)$. Note that $\varphi'(0) = J_y(u,y)S'(Bu;Bh) + J_u(u,y)h$.

**Algorithm 2** Line Search (Armijo)

**Input:** $\tau_0 > 0$; $c \in (0,1)$; $\nu \in (0,1)$; $l = 0$;

1: **while** $\varphi(\tau_l) > \varphi(0) + \nu\tau_l\varphi'(0)$ **do**
2: Set $\tau_{l+1} = c\tau_l$; $l := l + 1$.
3: **end while**

We first show that the line search stops in a finite number of steps.

**Lemma 2.4.** *For any $u \in H$ and $h \in H$ such that $(u,y) \in \operatorname{gph} S$ and $h$ solves the corresponding RAP (8), the line search stops in a finite number of steps.*

*Proof.* Assume $h \neq 0$, otherwise the statement holds trivially. Suppose $\nexists \tau > 0$ such that $\varphi(\tau) \leq \varphi(0) + \nu\tau\varphi'(0)$. Then $\forall \tau > 0$

$$\mathcal{J}(u) + \nu\tau(J_y(u,y)S'(Bu;Bh) + J_u(u,y)h) < \mathcal{J}(u+\tau h).$$

Expanding $\mathcal{J}$ at $u$ in direction $h$, we obtain

$$\mathcal{J}(u) + \nu\tau(J_y(u,y)S'(Bu;Bh) + J_u(u,y)h) < \mathcal{J}(u+\tau h) = \mathcal{J}(u) + \tau(J_y(u,y)S'(Bu;Bh) + J_u(u,y)h) + o(\tau).$$

Since $h$ solves (8), it follows from (9), with $\bar{h} = h$, that there exists $\varepsilon > 0$ such that

$$\varepsilon \leq (\nu - 1)(J_y(u,y)S'(Bu;Bh) + J_u(u,y)h) < \frac{o(\tau)}{\tau}, \ \forall \tau > 0,$$

a contradiction. □

**Theorem 2.5** (**Convergence of the Conceptual Algorithm**). *For each $k \in \mathbb{N}$, let $\tau_k$ represent the largest $\tau_l$ generated by the line search such that $\varphi(\tau_l) \leq \varphi(0) + \nu\tau_l\varphi'(0)$ and let $\widehat{\tau}_k$ be the smallest step size $\tau_l > 0$ such that $\varphi(\tau) > \varphi(0) + \nu\tau\varphi'(0)$ with $u := u_k$ and $h := h_k$. The conceptual algorithm, using the above line search, converges for any given starting point $u_0 \in H$ with $\epsilon = 0$ provided one of the following conditions holds:*

1. *There exists some $\underline{\tau} > 0$ such that for all $k$ sufficiently large $\tau_k \geq \underline{\tau}$.*
2. *The step sizes $\tau_k \downarrow 0$ and*

$$\limsup_{k \to +\infty} \frac{J(u_k + \widehat{\tau}_k h_k) - J(u_k) - \widehat{\tau}_k J'(u_k; h_k)}{\widehat{\tau}_k} \leq 0. \tag{10}$$

*Proof.* We first prove convergence assuming 1. Suppose that $\{u_k\} \subset H$, $\{\tau_k\} \subset \mathbb{R}_+$ and $\{h_k\} \subset H$ are sequences generated by the conceptual algorithm and line search. It follows from (9) and $u_{k+1} = u_k + \tau_k h_k$, that $\{\mathcal{J}(u_k)\}_k$ is monotonically decreasing. Since $\{\mathcal{J}(u_k)\}_k$ is bounded from below, $\{\mathcal{J}(u_{k+1}) - \mathcal{J}(u_k)\}_k$ is a null sequence. Then using (9) and $\varphi(\tau_k) \leq \varphi(0) + \nu\tau_k\varphi'(0)$, it follows that $||h_k||_H \to 0$.

The proof of 2. is also fairly standard. Since $\tau_k \downarrow 0$ it must also hold that $\widehat{\tau}_k \downarrow 0$. Moreover, by (9), we have $J'(u_k; h_k) \leq -c_1||h_k||_H^2$, for some $c_1 > 0$. Then by definition of $\widehat{\tau}_k$, we have

$$\widehat{\tau}_k^{-1}(\mathcal{J}(u_k + \widehat{\tau}_k h_k) - \mathcal{J}(u_k) - \widehat{\tau}_k\mathcal{J}'(u_k; h_k)) > (\nu - 1)\mathcal{J}'(u_k; h_k) \geq (1 - \nu)c_1||h_k||_H^2.$$

But then, by assumption, we have

$$0 \geq \limsup_{k\to+\infty} \widehat{\tau}_k^{-1}(\mathcal{J}(u_k + \widehat{\tau}_k h_k) - \mathcal{J}(u_k) - \widehat{\tau}_k\mathcal{J}'(u_k; h_k)) \geq \limsup_{k\to+\infty}(1 - \nu)c_1||h_k||_H^2 \geq \liminf_{k\to+\infty}(1 - \nu)c_1||h_k||_H^2 \geq 0.$$

Hence, $||h_k||_H \to 0$. □

The stepsize requirement 1. is admittedly strong. The asymtotic condition 2. is related to notions of semismoothness. However, it would be difficult to check numerically. Rather than imposing (10), we will add a robustification step to the algorithm, if it appears that $\tau_k \downarrow 0$ in order to ensure global convergence.

Perhaps the central difficulty in our approach is finding a robust descent direction $h_k$. We propose to handle this as follows :

- When $S$ is smooth, i.e. when strict complementarity holds, then we can directly solve the RAP, which reduces to a quadratic program.
- When strict complementarity fails, we approximate the directional derivative and show that a proper descent direction can be readily calculated.
- If it appears that $\tau_k \to 0$ but $||h_k|| \not\to 0$ or $||h_k|| \to 0$ "too" slowly, then we resort to a smoothing step to obtain the new iterate.

# 3 Optimal Control of a Unilateral Obstacle Problem

This section is split into four subsections. We define a model problem that fits into the general framework in Subsection 3.1. In order to obtain proper descent directions when $S$ is nonsmooth, we derive approximation results for the directional derivative in Subsection 3.2. In Subsection 3.3, we present an algorithm, which converges under the step size assumption from Theorem 2.5. Finally, in Subsection 3.4, we augment the new algorithm by adding a robustification step in order to obtain convergence without the step size assumption.

## 3.1 Definition and Properties of the Model

Define $\Omega \subset \mathbb{R}^n$, $n \in \{1,2,3\}$, to be an open bounded domain, which is either convex polyhedral or has a $C^{1,1}$-boundary. The spaces from the general setting are assumed to be defined by

$$V := H_0^1(\Omega), \quad H := L^2(\Omega), \quad V^* := H^{-1}(\Omega).$$

See [1] for the definition of the Lebesgue spaces $L^p(\Omega)$, as well as the Sobolev space $H_0^1(\Omega)$, its dual $H^{-1}(\Omega)$, and $H^2(\Omega)$. Under these assumptions the embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$ is continuous, dense, injective, and compact, and thus, they consistute a Gelfand triple (cf. Chapter 17.1 in [44]). We denote the Lebesgue measure of a set $\mathcal{B} \subset \mathbb{R}^d$ by $m(\mathcal{B})$.

We henceforth assume that $A$ is a symmetric second-order linear elliptic operator associated with the bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\langle Av, w\rangle := a(v,w) = \sum_{i,j=1}^{l} \int_\Omega a_{ij} \frac{\partial v}{\partial x_j} \frac{\partial w}{\partial x_i} dx + \int_\Omega cvw dx, \ \forall v, w \in H_0^1(\Omega),$$

where $c \geq 0$ is in $L^\infty(\Omega)$, $a_{ij} \in C^{0,1}(\bar{\Omega})$, i.e., Lipschitz continuous on the closure of $\Omega$, with $\sum_{i,j} w_i a_{ij} w_j \geq \xi |w|^2_{\mathbb{R}^n}$ for all $w \in \mathbb{R}^n$ and some real number and $\xi \geq 0$. Thus, if $y \in H^2(\Omega)$, then $Ay \in L^2(\Omega)$. The condition $Ay \in L^2(\Omega)$ is needed to ensure the increased regularity of the multiplier $\lambda \in L^2(\Omega)$ in the associated complementarity problem.

Finally, the closed convex subset $M \subset H_0^1(\Omega)$ will always be defined by

$$M := \left\{ y \in H_0^1(\Omega) \,|\, y \geq 0, \ a.e.\, \Omega \right\}.$$

For any $y \in M$, we define the subsets

$$\mathcal{A}(y) := \{ x \in \Omega \,|\, y(x) = 0 \} \text{ and } \mathcal{I}(y) := \Omega \setminus \mathcal{A}(y)$$

i.e., the active and inactive sets for the lower-level problem. Since $H^2(\Omega) \hookrightarrow C(\overline{\Omega})$ is continuous, $\mathcal{A}(y)$ will always be compact in $\text{int}(\Omega)$.

Given these basic data assumptions, we consider the following class of optimization problems:

$$\begin{aligned} &\min J(u,y) := \tfrac{1}{2}||y - y_d||^2_{L^2} + \tfrac{\alpha}{2}||u||^2_{L^2} \text{ over } (u,y) \in L^2(\Omega) \times H_0^1(\Omega), \\ &s.t. \quad Ay + N_M(y) \ni Bu + f. \end{aligned} \tag{11}$$

Here, $\alpha > 0$, $f \in L^2(\Omega)$, $B \in \mathcal{L}(L^2(\Omega))$ is completely continuous from $L^2(\Omega) \to H^{-1}(\Omega)$, and $N_M(y)$ represents the convex normal cone defined by

$$N_M(y) := \begin{cases} \left\{ \lambda \in H^{-1}(\Omega) \,\middle|\, \langle \lambda, y' - y\rangle_{H^{-1}, H_0^1} \leq 0, \ \forall y' \in M \right\}, & \text{if } y \in M \\ \emptyset, & \text{otherwise.} \end{cases}$$

Though we use here the case $y \geq 0$, the current setting includes the case $y \geq \psi$, provided $\psi \in H^2(\Omega)$ with $\psi|_{\partial\Omega} \leq 0$. The boundary requirement of $\psi$ is needed to ensure that $M \neq \emptyset$. Note that we can still obtain $Ay \in L^2(\Omega)$ under the weaker assumptions: $A\psi$ is a regular Borel measure and $\sup\{A\psi, 0\} \in L^2(\Omega)$ (in the sense of measures). In both cases $Ay \in L^2(\Omega)$.

In such a setting, the obstacle can be "absorbed" by the righthand side of the variational inequality (see Theorem II.1 and Corollary II.3 in [8]).

The mapping $S : H^{-1}(\Omega) \to H_0^1(\Omega)$ will always denote the solution mapping of the variational inequality in (11), i.e.,

$$S(w) := \left\{ y \in H_0^1(\Omega) \,|\, Ay + N_M(y) \ni w + f \right\}, \ \forall w \in H^{-1}(\Omega).$$

It is well known that $S$ is Lipschitz continuous from $H^{-1}(\Omega)$ into $H_0^1(\Omega)$ (see e.g., Chapter 3 in [25]).

Clearly, the objective function $J(u, y)$ in the MPEC (11) satisfies all requirements of the objective functional in the general framework. Finally, we refer to Proposition 4.5 in [23], see also the earlier works [31, 16], in which it was argued that $S$ is Hadamard directionally differentiable. Therefore, the MPEC (11) fits into the general framework of Section 2.

We recall that $S'(w; \cdot)$ is directly characterized by its graph:

$$\operatorname{gph} S'(w; \cdot) = \left\{ (h, d) \in H^{-1}(\Omega) \times H_0^1(\Omega) \,\middle|\, Ad + N_{\mathcal{K}(y,v)}(d) \ni h \right\}, \tag{12}$$

where $v := w + f - Ay \in N_M(y)$ and $\mathcal{K}(y, v)$ represents the so-called critical cone defined by

$$\mathcal{K}(y, v) := T_M(y) \cap \{v\}^{\perp}.$$

Here, $T_M(y)$ is the tangent cone to $M$ at $y$ given by (see e.g., Theorem 6.57 in [7] )

$$\begin{aligned} T_M(y) &= \left\{ d \in H_0^1(\Omega) \,\middle|\, \exists t_k \downarrow 0, \, \exists d_k \to_{H_0^1} d : y + t_k d_k \geq 0, \ a.e.\, \Omega \right\} \\ &= \left\{ d \in H_0^1(\Omega) \,|\, d \geq 0, \ \text{q.e.} \ \{x \in \Omega \,|y(x) = 0\} \right\}, \end{aligned} \tag{13}$$

where $q.e.$ stands for quasi-everywhere, see Appendix A.

Given a feasible point $(u, y)$ of (11), the B-stationarity conditions are (cf. [22, Theorem 2.1]):

$$\alpha(u, h)_{L^2} + (y - y_d, S'(Bu; Bh))_{L^2} \geq 0, \forall h \in L^2(\Omega). \tag{14}$$

Accordingly, by letting $Q \in \mathcal{L}(L^2(\Omega))$ represent some uniformly coercive bounded linear operator, we can define the bilinear form $q : L^2(\Omega) \times L^2(\Omega) \to \mathbb{R}$ by $q(h, h) := (Qh, h)_{L^2}$. Then it follows from (14) that the RAP associated with (11) is given by:

$$\min \tfrac{1}{2} q(h, h) + \alpha(u, h)_{L^2} + (y - y_d, S'(Bu; Bh))_{L^2} \text{ over } h \in L^2(\Omega). \tag{15}$$

## 3.2 Obtaining Descent Directions

We seek descent directions for the (reduced) objective functional $\mathcal{J}(u)$ defined by

$$\mathcal{J}(u) := \frac{1}{2} ||S(Bu) - y_d||_{L^2}^2 + \frac{\alpha}{2} ||u||_{L^2}^2.$$

If $S'(Bu; Bh)$ is linear is in $h$, then we solve a quadratic program. Otherwise we replace $S'(Bu; Bh)$ by an approximation, which is associated with a smooth version of the RAP. We note that our method does not require solving the smoothed RAP and that it does not rely on decision processes with combinatorial complexity. Throughout the following discussion, $(u, y)$ will be a feasible point for (11).

Under the standing assumptions, $y = S(Bu) \in H^2(\Omega) \cap H_0^1(\Omega)$ (cf. the standard regularity results in [25, 36]). Moreover, $A : H^2(\Omega) \cap H_0^1(\Omega) \to L^2(\Omega)$, $f \in L^2(\Omega)$, and

$B \in \mathcal{L}(L^2(\Omega))$. Then there exists a "multiplier" $\lambda \in N_M(y)$ such that $\lambda \in L^2(\Omega)$ and $(y, \lambda)$ solves the mixed complementarity problem

$$Ay - \lambda = Bu + f, \quad \lambda - \max(0, \lambda - y) = 0, \tag{16}$$

for a fixed $u \in L^2(\Omega)$. Since $\lambda$ can be interpreted pointwise, it follows from the second equation in (16) that

$$\lambda \geq 0, \ a.e. \, \mathcal{A}(y) \text{ and } \lambda = 0, \ a.e. \, \mathcal{I}(y). \tag{17}$$

Moreover, using $\mathcal{A}(y)$, the strongly active and biactive sets can be explicitly defined by

$$\mathcal{A}^+(y, \lambda) := \{x \in \mathcal{A}(y) \, | \lambda(x) > 0 \} \text{ and } \mathcal{A}^0(y, \lambda) := \{x \in \mathcal{A}(y) \, | \lambda(x) = 0 \},$$

respectively. We leave off the arguments if it is clear in context. We state the following known result for the sake of completeness.

**Proposition 3.1.** *Let $u \in L^2(\Omega)$ and $(y, \lambda)$ solve (16), the critical cone can be written*

$$\mathcal{K}(y, \lambda) := \left\{ d \in H_0^1(\Omega) \, \middle| d \geq 0, \ q.e. \, \mathcal{A}, \ d = 0, \ a.e. \, \mathcal{A}^+ \right\}.$$

*Proof.* This follows from [7, Theorems 6.44, 6.57]. □

We henceforth invoke the following assumptions:

$$\mathcal{A} = \overline{\text{int}(\mathcal{A})}, \tag{18a}$$

$$\text{If } m(\mathcal{A}^0) = 0, \text{ then } \text{cap}(\mathcal{A}^0) = 0 \text{ or } S'(Bu; Bh) = d = 0 \ q.e. \, \mathcal{A}. \tag{18b}$$

We will refer to the case in (18b) as "strict complementarity." Note that $S'(Bu; Bh) = d \in H_0^1(\Omega)$ solves

$$-\Delta v = Bh, \text{ in } H^{-1}(\mathcal{I}), \ v = 0, \ q.e. \, \partial\mathcal{I}, \tag{19}$$

see e.g., [31, Théorème 6.6] or [7, Corollary 6.60]. Otherwise, i.e., when $m(\mathcal{A}^0) > 0$, we say that strict complementarity fails.

Given the linearity of $S'$ under strict complementarity, we can obtain a descent direction by solving a quadratic program. This leads to the following result.

**Theorem 3.2 (Descent Directions under Strict Complementarity).** *Let $u \in L^2(\Omega)$ and $(y, \lambda)$ solve (16). If strict complementarity holds, then*

$$h = Q^{-1}(p - \alpha u),$$

*is a proper descent direction. Here, $p \in H_0^1(\Omega)$ such that $p = 0, \ q.e. \, \mathcal{A}.$ and*

$$-\Delta p = y_d - y, \ \text{in } H^{-1}(\mathcal{I}), \ p = 0, \ q.e. \, \partial\mathcal{I}, \tag{20}$$

*Proof.* This follows from (18a)-(18b) and the optimality conditions for (15). □

In order to handle a lack of strict complementarity, we propose a method in which

- the objective $F$ in (15) is replaced by a smoother mapping $F_\gamma$, and

- the existence of a sufficiently large penalty parameter $\gamma > 0$ is shown such that $-\nabla F_\gamma(0)$ is a proper descent direction for $\mathcal{J}(u)$.

In what follows let $u \in L^2(\Omega)$ and $(y, \lambda)$ solve (16). We use $\chi_{\mathcal{A}^+}$ and $\chi_{\mathcal{A}}$ to represent the characteristic functions of the associated strongly active and active sets, respectively. In addition, we let $\max_\varepsilon(0, \cdot)$ be the following (global) $C^2$-smoothing of the maximum operator:

$$\max_\varepsilon(0, \varphi) := \begin{cases} \varphi - \frac{\varepsilon}{2} & \text{if } \varphi \geq \varepsilon, \\ \frac{\varphi^3}{\varepsilon^2} - \frac{\varphi^4}{2\varepsilon^3} & \text{if } \varphi \in (0, \varepsilon), \\ 0 & \text{if } \varphi \leq 0. \end{cases} \qquad \varphi \in \mathbb{R},\ \varepsilon > 0.$$

Referring to [21, Lemma 2.1], we see that $\max_\varepsilon(0, \cdot)$ is non-negative, convex, monotone and Lipschitz continuous. We then approximate $d = S'(Bu; Bh)$ by a class of penalized problems:

$$\min \frac{1}{2}\langle Ad, d\rangle_{H^{-1}, H_0^1} - (Bh, d)_{L^2} + \frac{\gamma}{2}||\chi_{\mathcal{A}^+} d||^2_{L^2} + \gamma \int_\Omega \chi_{\mathcal{A}} \max_\varepsilon(0, -d)dx, \ \text{ over } d \in H_0^1(\Omega). \tag{21}$$

Due to convexity and differentiability, $d_{\gamma,\varepsilon}$ solves (21) if and only if it solves:

$$Ad + \gamma\chi_{\mathcal{A}^+} d - \gamma\chi_{\mathcal{A}} \max_\varepsilon(0, -d)' = Bh. \tag{22}$$

We will see that $d_{\gamma,\varepsilon}$ converges strongly in $H_0^1(\Omega)$ to some $d^* \in H_0^1(\Omega)$ along some subsequence of $\gamma \to +\infty$. For the sake of argument, let $\{d_k\}$ denote this sequence. Consider now that for any $\varphi \in H_0^1(\Omega)$ such that $\varphi \geq 0,\ q.e.\, \mathcal{A}$ and $\varphi = 0,\ a.e.\, \mathcal{A}^+$, we have:

$$\langle Bh, \varphi\rangle = \langle Ad_k, \varphi\rangle + \gamma_k\langle \chi_{\mathcal{A}^+} d_k, \varphi\rangle - \gamma_k\langle \chi_{\mathcal{A}} \max_\varepsilon(0, -d_k)', \varphi\rangle = \\ \langle Ad_k, \varphi\rangle - \gamma_k\langle \chi_{\mathcal{A}} \max_\varepsilon(0, -d_k)', \varphi\rangle \leq \langle Ad_k, \varphi\rangle$$

Passing to the limit, it holds that $\langle Bh, \varphi\rangle \leq \langle Ad^*, \varphi\rangle$.

In addition, we show that $d^* = 0,\ a.e.\, \mathcal{A}^+$ and $d^* \geq 0,\ a.e.\, \mathcal{A}$. So $d^*$ solves the same variational inequality as $S'(Bu; Bh)$. Applying Lemma A.2 to $\text{int}(\mathcal{A})$, it follows that $d^* \geq 0$, $q.e.$ on $\text{int}(\mathcal{A})$. In order to extend this to $\partial\mathcal{A} \cap \Omega$, we henceforth assume the following:

$$\text{If } m(\mathcal{A}^0) > 0, \text{ then } \exists \nu > 0, \exists \gamma' > 0 : \forall \gamma \geq \gamma', d_\gamma \geq 0,\ a.e. \mathcal{A}_\nu := \{x \in \Omega \,|\text{dist}(x, \mathcal{A}) < \nu\}. \tag{23}$$

This is reminiscent of a condition on the adjoint state used in [26, Assumption 1′(iii′)] to derive second-order optimality conditions for a related control problem. Under (18a)-(23), we now investigate the consistency of this approximation as $\gamma \to +\infty$ and $\varepsilon \to 0^+$.

**Proposition 3.3** (**Consistency of the Approximation**). *Let $u \in L^2(\Omega)$ and $(y, \lambda)$ solve (16) , and $\varepsilon > 0$. Then for all sequences $\{\gamma_n\}$ with $\gamma_n \to +\infty$ and $\{d_n\}$ with $d_n := d_{\gamma_n,\varepsilon}$, there exists a subsequence $\{d_k\}$ with $d_k := d_{\gamma_{n_k},\varepsilon}$ such that*

$$d_k \overset{H_0^1}{\to} d = S'(Bu; Bh).$$

*Proof.* For simplicity, let $\beta_\varepsilon(\cdot) := \max_\varepsilon(0, -\cdot)$. Fix some $n \in \mathbb{N}$ and let $\gamma_n$ and $d_n$ be as in the hypotheses. By definition, it follows that

$$\frac{1}{2}\langle Ad_n, d_n\rangle_{H^{-1},H_0^1} - (Bh, d_n)_{L^2} + \frac{\gamma_n}{2}||\chi_{\mathcal{A}^+}d_n||_{L^2}^2 + \gamma_n \int_\Omega \chi_{\mathcal{A}}\beta_\varepsilon(d_n)dx \leq$$
$$\frac{1}{2}\langle Ad, d\rangle_{H^{-1},H_0^1} - (Bh, d)_{L^2} + \frac{\gamma_n}{2}||\chi_{\mathcal{A}^+}d||_{L^2}^2 + \gamma_n \int_\Omega \chi_{\mathcal{A}}\beta_\varepsilon(d)dx,\ \forall d \in H_0^1(\Omega). \quad (24)$$

Setting $d = 0$ in (24) and rearranging terms, it follows from the coercivity of $A$, $\beta_\varepsilon(\cdot) \geq 0$, and the continuity of $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ that there exists a constant $\zeta > 0$ such that $||d_n||_{H_0^1} \leq \zeta||Bh||_{L^2}$. Hence, $\{d_n\}$ is bounded in $H_0^1(\Omega)$ and there exists a subsequence $\{d_{n_k}\}$ and a $d^* \in H_0^1(\Omega)$ such that $d_{n_k} \rightharpoonup d^*$ in $H_0^1(\Omega)$.

Next by using a fixed arbitrary $d \in \mathcal{K}(y, \lambda)$ in (24), we obtain a constant $c > 0$ such that

$$\frac{1}{2}\langle Ad_{n_k}, d_{n_k}\rangle_{H^{-1},H_0^1} - (Bh, d_{n_k})_{L^2} + \frac{\gamma_{n_k}}{2}||\chi_{\mathcal{A}^+}d_{n_k}||_{L^2}^2 + \gamma_{n_k} \int_\Omega \chi_{\mathcal{A}}\beta_\varepsilon(d_{n_k})dx \leq c.$$

Dividing through by $\gamma_{n_k}$ and passing to the limit inferior, it follows from the weak lower semicontinuity of $\langle A\cdot, \cdot\rangle_{H^{-1},H_0^1}$ that

$$\frac{1}{2}||\chi_{\mathcal{A}^+}d_{n_k}||_{L^2}^2 + \int_\Omega \chi_{\mathcal{A}}\beta_\varepsilon(d_{n_k})dx \to 0 \text{ as } n_k \to +\infty, \quad (25)$$

In light of (23), $d^* \in \mathcal{K}(y, \lambda)$. Substituting $d^*$ into the right side of (24), we obtain

$$\frac{1}{2}\langle Ad_{n_k}, d_{n_k}\rangle_{H^{-1},H_0^1} - (Bh, d_{n_k})_{L^2} + \frac{\gamma_{n_k}}{2}||\chi_{\mathcal{A}^+}d_{n_k}||_{L^2}^2 + \gamma_{n_k} \int_\Omega \chi_{\mathcal{A}}\beta_\varepsilon(d_{n_k})dx \leq$$
$$\frac{1}{2}\langle Ad^*, d^*\rangle_{H^{-1},H_0^1} - (Bh, d^*)_{L^2}.$$

For the symmetric bilinear form $a(\cdot, \cdot) := \langle A\cdot, \cdot\rangle$, we have

$$a(d_{n_k}, d_{n_k}) - a(d^*, d^*) = a(d_{n_k} - d^*, d_{n_k} - d^*) + 2a(d^*, d_{n_k} - d^*). \quad (26)$$

Then using $\beta_\varepsilon(\cdot) \geq 0$, the compactness of $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$, and rearranging terms, we have

$$\frac{\xi}{2}||d_{n_k} - d^*||_{H_0^1}^2 \leq (Bh, d_{n_k} - d^*)_{L^2} - 2a(d^*, d_{n_k} - d^*),$$

from which we conclude that $d_{n_k} \to d^*$ in $H_0^1(\Omega)$.

Finally, given $d_{n_k} \to d^*$ in $H_0^1(\Omega)$, we fix an arbitrary $d \in \mathcal{K}(y, \lambda)$ and substitute it into the right side of (24). After passing to the limit, we have

$$\frac{1}{2}\langle Ad^*, d^*\rangle_{H^{-1},H_0^1} - (Bh^*, d^*)_{L^2} \leq \frac{1}{2}\langle Ad, d\rangle_{H^{-1},H_0^1} - \langle Bh^*, d\rangle_{H^{-1},H_0^1}.$$

Hence, $d^* = S'(Bu; Bh)$. □

**Remark 3.4.** *Based on Proposition 3.3, it is not necessary for $\varepsilon \downarrow 0$ to ensure the convergence of the approximating derivatives $d_{\gamma,\varepsilon}$*

The next result will be needed in order to verify that there exist penalty parameters $\gamma$ such that the proposed direction is a true descent direction for $\mathcal{J}$.

**Proposition 3.5** (**Continuity Properties**). *For any $u \in L^2(\Omega)$, $(y, \lambda)$ solving (16) , and fixed $\gamma$ and $\varepsilon > 0$, the mapping $h \mapsto d_{\gamma,\varepsilon}$ is completely Lipschitz continuous from $L^2(\Omega)$ into $H_0^1(\Omega)$.*

*Proof.* Let $\beta_\varepsilon(\cdot) := \max_\varepsilon(0, -\cdot)$. Since $\int_\Omega \chi_{\mathcal{A}}\beta_\varepsilon(\cdot)dx$ is a proper, convex, differentiable functional its derivative $\chi_{\mathcal{A}}\beta_\varepsilon(d)'$ at some point $d \in H_0^1(\Omega)$ is monotone. Therefore, $\langle \chi_{\mathcal{A}}(\beta_\varepsilon'(d_1) - \beta_\varepsilon'(d_2)), d_1 - d_2\rangle_{H^{-1},H_0^1} \geq 0$, for all $d_1, d_2 \in H_0^1(\Omega)$. Now let $w_1, w_2 \in H^{-1}(\Omega)$ define $d_1$ and $d_2$ to be the unique solutions satisfying the following equations

$$Ad_1 + \gamma\chi_{\mathcal{A}^+}d_1 + \gamma\chi_{\mathcal{A}}\beta_\varepsilon'(d_1) = w_1 \text{ and } Ad_2 + \gamma\chi_{\mathcal{A}^+}d_2 + \gamma\chi_{\mathcal{A}}\beta_\varepsilon'(d_2) = w_2,$$

respectively. Subtracting the second equation from the first and then multiplying both sides by $d_1 - d_2$, we obtain

$$\begin{aligned}\langle A(d_1 - d_2), d_1 - d_2\rangle_{H^{-1},H_0^1} + \gamma||\chi_{\mathcal{A}^+}(d_1 - d_2)||_{L^2}^2 + \langle \chi_{\mathcal{A}}(\beta_\varepsilon'(d_1) - \beta_\varepsilon'(d_2)), d_1 - d_2\rangle_{H^{-1},H_0^1} \\ = \langle w_1 - w_2, d_1 - d_2\rangle_{H^{-1},H_0^1}.\end{aligned}$$

Using the non-negativity of the terms on the lefthand side of the equation and the coercivity of $A$, we deduce the existence of some $\xi > 0$ such that$||d_1 - d_2||_{H_0^1} \leq \xi||w_1 - w_2||_{H^{-1}}$. The assertion follows from the compactness of $B$ from $L^2(\Omega)$ to $H^{-1}(\Omega)$. □

Finally, we will need the following result on the smoothness of the mapping $h \mapsto d_{\gamma,\varepsilon}$ in Proposition 3.5.

**Proposition 3.6** (**Differentiability Properties**). *For any $u \in L^2(\Omega)$, $(y, \lambda)$ solving (16), and fixed $\gamma$ and $\varepsilon > 0$, the mapping $h \mapsto d_{\gamma,\varepsilon}$ is Gâteaux differentiable from $L^2(\Omega)$ into $H_0^1(\Omega)$.*

*Proof.* We use the continuous differentiability of the penalty term and the Implicit Function Theorem (IFT). For some $w \in H^{-1}(\Omega)$, let $d_{\gamma,\varepsilon}(w)$ solve $Ad+\gamma\chi_{\mathcal{A}^+}d-\gamma\chi_{\mathcal{A}}\max_\varepsilon(0, -d)' = w$. Next, define $\mathcal{F}_{\gamma,\varepsilon}(w, d) := Ad + \gamma\chi_{\mathcal{A}^+}d - \gamma\chi_{\mathcal{A}}\max_\varepsilon(0, -d)' - w$. Based on the assumptions on $\max_\varepsilon$, $\mathcal{F}_{\gamma,\varepsilon}$ is continuously Fréchet differentiable with respect to its first and second arguments denoted, respectively, by

$$D_1\mathcal{F}_{\gamma,\varepsilon}(w, d) = -I, \quad D_2\mathcal{F}_{\gamma,\varepsilon}(w, d) = A + \gamma\chi_{\mathcal{A}^+} + \gamma\chi_{\mathcal{A}}\max_\varepsilon(0, -d)''.$$

Here, $I$ represents the identity on $H^{-1}(\Omega)$. Since $D_2\mathcal{F}_{\gamma,\varepsilon}(w, d)$ is (for each $d$) a uniformly coercive bounded linear operator from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$ and hence, invertible, it follows from the IFT that the derivative of $d$ with respect to $w$ (in direction $\eta \in H^{-1}(\Omega)$) is the unique solution $\delta(\eta) \in H_0^1(\Omega)$ of the following linear equation

$$A\delta + \gamma\chi_{\mathcal{A}^+}\delta + \gamma\chi_{\mathcal{A}}\max_\varepsilon(0, -d)''\delta = \eta.$$

□

Let $d_{\gamma,\varepsilon}(h)$ be the smooth approximation of $S'(Bu;Bh)$ considered in Propositions 3.3, 3.5, 3.6. Following (15), we define the fully regularized auxiliary problem for some fixed $\gamma$ and $\varepsilon > 0$, ($\text{FRAP}_{\gamma,\varepsilon}$) by

$$\min F_{\gamma,\varepsilon}(h) := \tfrac{1}{2}q(h,h) + \alpha(u,h)_{L^2} + (y - y_d, d_{\gamma,\varepsilon}(h))_{L^2} \text{ over } h \in L^2(\Omega). \qquad (\text{FRAP}_{\gamma,\varepsilon})$$

We continue by proving some basic properties of $\text{FRAP}_{\gamma,\varepsilon}$.

**Proposition 3.7** (**Existence and B-stationarity**). *Let $u \in L^2(\Omega)$ and $(y,\lambda)$ solve (16). For fixed arbitrary $\gamma$ and $\varepsilon > 0$, $FRAP_{\gamma,\varepsilon}$ has a solution $h$. Moreover, if for some $\gamma > 0$, $h = 0$ solves $FRAP_{\gamma,\varepsilon}$, then $(u,y)$ is B-stationary.*

*Proof.* It can be easily shown that $F_{\gamma,\varepsilon}$ is coercive and weakly lower-semicontinuous by using the results of Proposition 3.5 and the fact that $d_{\gamma,\varepsilon}(0) = 0$, from which it follows that $\text{FRAP}_{\gamma,\varepsilon}$ has a solution $h_{\gamma,\varepsilon}$. If $h = 0$ solves $\text{FRAP}_{\gamma,\varepsilon}$, then

$$0 \le \frac{1}{2}q(h,h) + \alpha(u,h)_{L^2} + (y - y_d, d_{\gamma,\varepsilon}(h))_{L^2},\ \forall h \in L^2(\Omega).$$

Passing to the limit, along an appropriate subsequence, as $\gamma \to +\infty$, it follows from Proposition 3.3 that

$$0 \le \frac{1}{2}q(h,h) + \alpha(u,h)_{L^2} + (y - y_d, S'(Bu;Bh))_{L^2},\ \forall h \in L^2(\Omega).$$

By Lemma 2.1, $(u,y)$ is B-stationary. □

As mentioned earlier, we are not interested in *solving* $\text{FRAP}_{\gamma,\varepsilon}$, rather we wish to exploit certain properties in order to obtain a descent direction.

**Proposition 3.8** (**Using $F_{\gamma,\varepsilon}$ to Characterize Descent Directions**). *Let $u \in L^2(\Omega)$, $(y,\lambda)$ solve (16), and fix $\gamma$ and $\varepsilon > 0$. Suppose $0 \in L^2(\Omega)$ is not a solution to $FRAP_{\gamma,\varepsilon}$ and $h \in L^2(\Omega)$. If there exists a $\beta > 0$ such that $F_{\gamma,\varepsilon}(\beta h) < 0$, then there exists $\hat{\gamma} > 0$ such that $h$ is a proper descent direction for $\mathcal{J}$ at $u$.*

*Proof.* Suppose that $0 \in L^2(\Omega)$ is not a solution to the $\text{FRAP}_{\gamma,\varepsilon}$. Then since $F_{\gamma,\varepsilon}(0) = 0$, there must exist a function $\tilde{h} \in L^2(\Omega)$ such that $\frac{1}{2}q(\tilde{h},\tilde{h}) + \alpha(u,\tilde{h})_{L^2} + (y - y_d, d_{\gamma,\varepsilon}(\tilde{h}))_{L^2} < 0$.

Fixing $\tau > 0$, it follows from the directionally differentiability of $\mathcal{J}$, the previous inequality, and the coercivity of $q(\cdot,\cdot)$ that

$$\begin{aligned}
\mathcal{J}(u + \tau\tilde{h}) - \mathcal{J}(u) &= \mathcal{J}(u) + \tau\mathcal{J}'(u;\tilde{h}) + o(\tau) - \mathcal{J}(u) \\
&= \tau(\mathcal{J}'(u;\tilde{h}) + \frac{o(\tau)}{\tau}) \\
&= \tau\left(\alpha(u,\tilde{h})_{L^2} + (y - y_d, S'(Bu;B\tilde{h}))_{L^2} + \frac{o(\tau)}{\tau}\right) \\
&\le \tau\left(-\frac{1}{2}q(\tilde{h},\tilde{h}) + (y - y_d, S'(Bu;B\tilde{h}) - d_{\gamma,\varepsilon}(\tilde{h}))_{L^2} + \frac{o(\tau)}{\tau}\right) \\
&\le \tau\left(-\frac{1}{2}c_1||\tilde{h}||^2_{L^2} + (y - y_d, S'(Bu;B\tilde{h}) - d_{\gamma,\varepsilon}(\tilde{h}))_{L^2} + \frac{o(\tau)}{\tau}\right).
\end{aligned}$$

Proposition 3.3 implies that $d_{\gamma,\varepsilon}(\tilde{h}) \to S'(Bu; B\tilde{h})$ strongly in $H_0^1(\Omega)$ along a subsequence of $\gamma \to +\infty$, denoted still by $\gamma$ and therefore, strongly in $L^2(\Omega)$. Hence, there exist $\tilde{\gamma} > 0$ such that

$$||d_{\tilde{\gamma},\varepsilon}(\tilde{h}) - S'(Bu; B\tilde{h})||_{L^2} < \frac{1}{||y - y_d||_{L^2}} \cdot \frac{c_1}{4}||\tilde{h}||_{L^2}^2, \tag{27}$$

provided that $\|y - y_d\|_{L^2} > 0$; otherwise the descent property follows readily. Thus, for $\gamma = \tilde{\gamma}$, it holds that

$$\frac{\mathcal{J}(u + \tau\tilde{h}) - \mathcal{J}(u)}{\tau} \leq -\frac{c_1}{4}||\tilde{h}||_{L^2}^2 + \frac{o(\tau)}{\tau} \overset{\tau \downarrow 0}{\Rightarrow} \mathcal{J}'(u; \tilde{h}) \leq -\frac{c_1}{4}||\tilde{h}||_{L^2}^2.$$

□

Proposition 3.8 raises the need to find an $h$ such that $F_{\gamma,\varepsilon}(h) < 0$. We provide a suggestion in the next result.

**Theorem 3.9** (**Descent Directions without Strict Complementarity**). *Let $u \in L^2(\Omega)$,$(y, \lambda)$ solve* (16)*, and suppose that $p \in H_0^1(\Omega)$ solves the adjoint equation given by*

$$Ap + \gamma\chi_{\mathcal{A}^+}p = y_d - y. \tag{28}$$

*Then there exists a finite $\gamma > 0$ such that*

$$h = (B^*p - \alpha u)$$

*is a proper descent direction for $\mathcal{J}$ at $u$ provided $p \neq \alpha u$.*

*Proof.* If $\nabla F_{\gamma,\varepsilon}(0) \neq 0$, then 0 does not solve $\text{FRAP}_{\gamma,\varepsilon}$. Moreover, it follows from the differentiability of $d_{\gamma,\varepsilon}$ that

$$\nabla F_{\gamma,\varepsilon}(0) = \alpha u - B^*p,$$

where $p$ solves (28). Setting $h = -\nabla F_{\gamma,\varepsilon}(0) = B^*p - \alpha u$ we have $(\nabla F_{\gamma,\varepsilon}(0), h) < 0$. Hence $h$ is a proper descent direction for $F_{\gamma,\varepsilon}$ at zero.

Finally, if $F_{\gamma,\varepsilon}(h) < 0$, then $\tau = 1$. Otherwise, suppose $F_{\gamma,\varepsilon}(h) = 0$ and define $\psi(t) := F_{\gamma,\varepsilon}(0 + th)$. Since $\psi'(0) = F'_{\gamma,\varepsilon}(0)h < 0$ and $\psi$ is continuous, there must exist some $\tau > 0$ such that $\psi(\tau) = F_{\gamma,\varepsilon}(\tau h) < 0$. If $F_{\gamma,\varepsilon}(h) > 0$, then by continuity and $\psi'(0) < 0$, there exists some $\tau^* > 0$ such that $\psi(\tau^*) = F_{\gamma,\varepsilon}(\tau^* h) = 0$. Reverting to the previous argument with $\tau^* h$ in place of $h$ completes the proof. □

## 3.3 An Iterative Procedure for Calculating Stationary Points

The theoretical results of Subsection 3.2 motivate Algorithm 3. We momentarily postpone a discussion of the proper stopping criterion. All model data, i.e., $\Omega$, $y_d$, $\alpha$, $B$, $A$, $f$, $Q$, will be treated globally. We occasionally use $\mathcal{A}_k^0 := \mathcal{A}^0(y_k, \lambda_k)$, $\mathcal{A}_k^+ := \mathcal{A}^+(y_k, \lambda_k)$, $\mathcal{A}_k := \mathcal{A}(y_k, \lambda_k)$, $\mathcal{I}_k := \mathcal{I}(y_k)$, for readability. In addition, we refer to the case when $m(\mathcal{A}_k^0) = 0$ as the "smooth case;" otherwise we refer to the "nonsmooth case."

**Algorithm 3** Descent Method for MPECs I

**Input:** $u_0 \in L^2(\Omega)$; $\gamma_0 > 0$; $\varepsilon > 0$; $k := 0$; $\rho_2 > 1$;

```
(y_0, λ_0) solves (16) with u = u_0.
while stopping criterion not fulfilled do
    if m(A_k^0) = 0 then
        Set h_k = Q^{-1}(B*p_k − αu_k), where p_k solves (20) with y = y_k.
        Compute (u_{k+1}, y_{k+1}, λ_{k+1}) via Algorithm 2 in direction h_k.
    else
        Set h_k = B*p_k − αu_k, where p_k solves (28) with y = y_k, γ = γ_k.
        while (27) fails do
            Choose γ̃_k > γ_k.
            Set h_k = B*p_k − αu_k, where p_k solves (28) with y = y_k, γ = γ̃_k.
            Set γ_k = γ̃_k.
        end while
        Compute (u_{k+1}, y_{k+1}, λ_{k+1}) via Algorithm 2 in direction h_k
        Choose γ_{k+1} > ρ_2 γ_k.
    end if
    Set k := k + 1.
end while
```

**Remark 3.10.** *When $m(\mathcal{A}_k^0) > 0$, we always use $d_{\gamma_k,\varepsilon}(h_k)$ in place of $d(h_k)$ in the line search at iteration $k$.*

**Remark 3.11** (**Practical aspects of the line search**). *If $Q$ is the identity and strict complementarity holds, then $\mathcal{J}'(u_k; h_k) = \alpha(u_k, h_k) - (B^*p_k, h_k) = -\|B^*p_k - \alpha u_k\|^2$. Thus, in smooth settings there is no need to calculate $d_k = S(Bu_k; Bh_k)$ whenever we choose $Q = Id$.*

**Remark 3.12** (**Evaluating** (27)). *In practice, one could avoid excess computational effort by assuming $d_{\gamma,\varepsilon}(h) \to S'(Bu; Bh)$ like $o(\gamma^{-1/2})$ and first choosing $\hat{\gamma}_k$ such that*

$$(\hat{\gamma}_k)^{-\rho_1} \leq \frac{1}{||y - y_d||_{L^2}} \cdot \frac{c_1}{4} ||\tilde{h}||_{L^2}^2$$

*for some $\rho_1 > 1/2$ and then setting $\tilde{\gamma}_k := \max(\hat{\gamma}_k, \rho_2\gamma_k)$ for some fixed $\rho_2 > 1$.*

Note that Algorithm 3 assumes $B^*p_k \neq \alpha u_k$ for all $k$. If for some finite $k$, $B^*p_k = \alpha u_k$, then the algorithm would potentially stop. We thus need to study the stationarity of $(u_k, y_k)$.

**Definition 3.13** (**Stationarity Conditions**). *A feasible point $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ for the MPEC (11) is called a C-stationary point, if there exist multipliers $\bar{\lambda} \in L^2(\Omega)$, $\bar{p} \in H_0^1(\Omega)$, and $\bar{\mu} \in H^{-1}(\Omega)$ such that*

$$\alpha\bar{u} = B^*\bar{p}, \tag{29}$$

$$A\bar{p} = y_d - \bar{y} + \bar{\mu}, \tag{30}$$

$$A\bar{y} - \bar{\lambda} = B\bar{u} + f, \tag{31}$$

*where the multipliers satisfy the following conditions:*

$$0 \leq \bar{\lambda} \perp \bar{y} \geq 0, \ a.e.\, \Omega, \tag{32}$$
$$0 = \bar{p}, \ a.e.\, \mathcal{A}^+(\bar{y}, \bar{\lambda}), \tag{33}$$
$$0 = \langle \bar{\mu}, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi = 0 \ a.e.\, \mathcal{A}(\bar{y}), \tag{34}$$
$$0 \geq \langle \bar{\mu}, \bar{p} \rangle_{H^{-1}, H_0^1}. \tag{35}$$

*If in addition,*

$$0 \geq \langle \bar{\mu}, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi \geq 0 \ a.e.\, \mathcal{A}^0(\bar{y}), \tag{36}$$
$$0 \leq \bar{p}, \ a.e.\, \mathcal{A}^0(\bar{y}, \bar{\lambda}), \tag{37}$$

*then $(\bar{u}, \bar{y})$ is said to be strongly stationary.*

The notion of S-stationarity differs from the system obtained in [31, 32]: The *a.e.* requirement on $p$ in (33) and (37) is less restrictive than a *q.e.* requirement, and, conversely, the conditions on $\mu$ in (34) and (36) are more restrictive than when requiring *q.e.* conditions on $\varphi$. Nevertheless, the conditions stated here are typically guaranteed in numerics and limiting arguments, see e.g., [46, 45, 4, 24, 20, 21, 22, 38]. Note also that in the presence of strict complementarity, (36)-(37) hold vacuously.

**Theorem 3.14** (**Stationarity at Finite Termination**). *In the context of Algorithm 3, suppose there exists $k_0 \in \mathbb{N}$ such that $B^* p_{k_0} = u_{k_0}$.*

1. *If strict complementarity holds, then by defining $\mu_{k_0} \in H^{-1}(\Omega)$ such that*

$$\langle \mu_{k_0}, \varphi \rangle_{H^{-1}, H_0^1} = 0, \forall \varphi \in H_0^1(\Omega) : \varphi = 0, \ a.e.\, \mathcal{A}(y_{k_0}),$$

*the point $(u_{k_0}, y_{k_0}, \lambda_{k_0}, p_{k_0}, \mu_{k_0})$ satisfies (29)-(37). Hence, is $(u_{k_0}, y_{k_0})$ is strongly stationary.*

2. *If strict complementarity fails, then using $\mu_{k_0} = -\gamma_{k_0} \chi_{\mathcal{A}_{k_0}^+} p_{k_0}$ one has*

$$\alpha u_{k_0} = B^* p_{k_0}, \tag{38}$$
$$A p_{k_0} = y_d - y_{k_0} + \mu_{k_0}, \tag{39}$$
$$A y_{k_0} - \lambda_{k_0} = B u_{k_0} + f, \tag{40}$$

*along with the conditions:*

$$0 \leq \lambda_{k_0} \perp y_{k_0} \geq 0, \ a.e.\, \Omega, \tag{41}$$
$$0 = \mu_{k_0}, \ a.e.\, \mathcal{A}^0(y_{k_0}, \lambda_{k_0}) \cup \mathcal{I}(y_{k_0}), \tag{42}$$
$$0 \geq (\mu_{k_0}, p_{k_0})_{L^2}. \tag{43}$$

**Remark 3.15.** *It can be shown that in the current setting, strong stationarity and B-stationarity are equivalent concepts.*

*Proof.* Both statements follow immediately by definition of the algorithm and $\mu_{k_0}$. □

When strict complementarity fails, conditions (38)-(43) are not enough to argue that $(u_{k_0}, y_{k_0})$ is C-stationary since (33) is missing. In such a situation, a new adjoint state $p'_{k_0}$ and penalty parameter $\gamma'_{k_0}$ should be obtained by first choosing $\gamma'_{k_0} > \gamma_{k_0}$, keeping $\mathcal{A}^+_{k_0}$, $u_{k_0}$, and $y_{k_0}$ as before, and then resolving the adjoint equation. If, once again, $h_{k_0} := B^* p'_{k_0} - u_{k_0} = 0$ then we repeat this process until $h_{k_0} \neq 0$ and resume the iteration. Of course if $\ker B^* = \{0\}$, then $B^* p'_{k_0} = B^* p_{k_0}$ implies $p_{k_0} = p'_{k_0}$. This is only possible if either $m(\mathcal{A}^+_{k_0}) = 0$ or $p_{k_0} = 0$, *a.e.* $\mathcal{A}^+_k$, in which case $(u_{k_0}, y_{k_0})$ is C-stationary. Next, we show that if $\ker B^*$ is nontrivial, then an adjoint state $p$ and muliplier $\mu$ can be found such that $(u_{k_0}, y_{k_0})$ is a C-stationary point.

**Theorem 3.16** (**A Special Case of C-stationarity**). *Let $\gamma_k \to +\infty$. Suppose that the sequence*

$$\{(u_k, y_k, \lambda_k, p_k, h_k)\} \subset L^2(\Omega) \times H^1_0(\Omega) \times L^2(\Omega) \times H^1_0(\Omega) \times L^2(\Omega)$$

*is defined by repeating 1.-5. as $k \to +\infty$ in the following recursion rule:*

*0. $u_0 \in L^2(\Omega)$ and $k = 0$.*

*1. $(y_k, \lambda_k)$ solves (16) with $u = u_k$.*

*2. $p_k$ solves: $[A + \gamma_k \chi_{\mathcal{A}^+_k}]p = y_d - y_k$, in $p$.*

*3. $h_k = B^* p_k - \alpha u_k$.*

*4. $(u_{k+1}, y_{k+1}, \lambda_{k+1})$ obtained via a line search in direction $h_k$.*

*5. $k = k + 1$.*

*If $\ker B^*$ is nontrivial and there exists a $k_0 \in \mathbb{N}$ such that $h_k = 0$ for all $k \geq k_0$, then $u_k = u_{k_0}$, $y_k = y_{k_0}$, $\lambda_k = \lambda_{k_0}$, for all $k \geq k_0$ and there exists a subsequence of $\{k_l\}$ along with $p \in H^1_0(\Omega)$ and $\mu \in H^{-1}(\Omega)$ with $p_{k_l} \rightharpoonup_{H^1_0} p$ and $\gamma_{k_l} \chi_{\mathcal{A}^+_{k_l}} p_{k_l} \rightharpoonup_{H^{-1}} \mu$ such that $(u_{k_0}, y_{k_0}, \lambda_{k_0}, p, \mu)$ satisfies conditions (29) - (35). In other words, $(u_{k_0}, y_{k_0})$ is C-stationary.*

*Proof.* The triple $(u_k, y_k, \lambda_k) = (u_{k_0}, y_{k_0}, \lambda_{k_0})$ for all $k \geq k_0$ follows from the fact that the line search yields $u_{k+1} = u_k$ for $k \geq k_0$. In addition, this implies that $\mathcal{A}^+_k = \mathcal{A}^+_{k_0}$ and $\mathcal{A}^0_k = \mathcal{A}^0_{k_0}$ for all $k \geq k_0$. If follows that (29), (31), (32) hold. In what follows, we leave off the subscripts on these fixed quantities.

For an arbitrarily fixed $k \geq k_0$, multiplying the adjoint equation in 2. by $p_k$ yields

$$\langle A p_k, p_k \rangle_{H^{-1}, H^1_0} + \gamma_k (\chi_{\mathcal{A}^+} p_k, p_k)_{L^2} = (y_d - y_k, p_k)_{L^2}. \tag{44}$$

We leave off the subscripts from the $H^1_0$-duality pairings and $L^2$-inner products for readability.

Since $A$ is coercive, there exists a $\xi_1 > 0$ such that $\xi_1 ||p_k||^2_{H^1_0} \leq (y_d - y, p_k)$. Since $k \geq k_0$ and the embedding $H^1_0(\Omega) \hookrightarrow L^2(\Omega)$ is continuous, the sequence $\{p_k\} \subset H^1_0(\Omega)$ is bounded. Hence, there exists a $p \in H^1_0(\Omega)$ and a subsequence of $\{p_k\}$, denoted still by $k$, such that $p_k \rightharpoonup_{H^1_0} p$. It follows from (44) that there exists a $\xi_2 > 0$ such that

$$||\chi_{\mathcal{A}^+} p_k||_{L^2} \leq \frac{\xi_2}{\gamma_k} ||y - y_d||_{L^2}.$$

Therefore, as neither $\mathcal{A}^+$ nor $y$ change as $k \to +\infty$, $p = 0$, *a.e.* $\mathcal{A}^+(y, \lambda)$. Moreover, the compactness of the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ implies that $p_k \to_{L^2} p$. This yields (33), whereas (29) follows from the fact that $\alpha u = B^* p_k$ for all $k \geq k_0$.

Since $p_k \rightharpoonup_{H_0^1} p$, it holds that $\langle Ap_k, \varphi \rangle + (y - y_d, \varphi) \to \langle Ap, \varphi \rangle + (y - y_d, \varphi)$ for all $\varphi \in H_0^1(\Omega)$. Then since $\langle Ap_k, \varphi \rangle + (y - y_d, \varphi) = -\gamma_k \langle \chi_{\mathcal{A}^+} p_k, \varphi \rangle$, $\forall \varphi \in H_0^1(\Omega)$, there must exist a $\mu \in H^{-1}(\Omega)$ such that $-\gamma \chi_{\mathcal{A}^+} p_k \rightharpoonup_{H^{-1}} \mu$. Whence we have (30).

Now let $\varphi \in H_0^1(\Omega)$ such that $\varphi = 0$, *a.e.* $\mathcal{A}(y)$. Then since $-\gamma \langle \chi_{\mathcal{A}^+} p_k, \varphi \rangle = 0$, $\forall k$, it holds that $\langle \mu, \varphi \rangle = 0$, $\forall \varphi \in H_0^1(\Omega) : \varphi = 0$, *a.e.* $\mathcal{A}(y)$. This yields (34).

Finally, it follows from (44), in conjunction with $p_k \rightharpoonup_{H_0^1} p$, that

$$\langle \mu, p \rangle = \langle Ap, p \rangle + (y - y_d, p) \leq \liminf_{k \to \infty} \langle Ap_k, p_k \rangle + (y - y_d, p_k) = \liminf_{k \to \infty} -\gamma_k \langle \chi_{\mathcal{A}^+} p_k, p_k \rangle \leq 0,$$

from which (35) follows. This completes the proof. □

Using Theorem 3.16, we can only guarantee C-stationarity of the iterates in the nonsmooth loop for a pathological case. Nevertheless, we will see that the algorithm guarantees a slightly weakened version of C-stationarity introduced by the first author and I. Kopacka in [20] asymtotically.

**Definition 3.17** (**$\mathcal{E}$-Almost-C-Stationarity**). *A feasible point $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ to the MPEC (11) is called an $\mathcal{E}$-almost C-stationary point, if there exist multipliers $\bar{\lambda} \in L^2(\Omega)$, $\bar{p} \in H_0^1(\Omega)$, and $\bar{\mu} \in H^{-1}(\Omega)$ such that (29)-(31), (32)-(33) and (35) hold, and in place of (34) one has $\forall \sigma > 0$, there exists a subset $\mathcal{O}_\sigma \subset \mathcal{I}(\bar{y})$ with $m(\mathcal{O}_\sigma) \leq \sigma$ such that*

$$\langle \mu^*, \phi \rangle = 0, \ \forall \phi \in H_0^1(\Omega) : \phi = 0, \ a.e. \, \mathcal{A}(\bar{y}) \cup \mathcal{O}_\sigma. \tag{45}$$

Note that these conditions were first derived using $\mathcal{E}$gorov's theorem, whereas our argument is slightly different. We will see in the following result that another form of stationarity is possible, which makes direct use of the generalized gradient of Clarke. We restate the definition here for convenience.

**Definition 3.18.** *Let $X$ be a real Banach space and $f : X \to \mathbb{R}$ locally Lipschitz near $x$. Then the generalized directional derivative of $f$ at $x$ in direction $v \in X$, denoted $f^\circ(x; v)$ is defined by*

$$f^\circ(x; v) := \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y + tv) - f(y)}{t}.$$

*The generalized gradient of $f$ at $x$, denoted $\partial^\circ f(x)$, is the subset of $X^*$, given by*

$$\partial^\circ f(x) := \{ x^* \in X^* \, | \langle x^*, v \rangle \leq f^\circ(x; v), \ \forall v \in X \}.$$

For more on Clarke's theory of nonsmooth analysis, we refer the reader to [10].

**Theorem 3.19** (**Convergence of Algorithm 3**). *Let $\{u_k\}$ be a sequence of controls generated by Algorithm 3 and assume that Algorithm 3 does not finitely terminate. Then $\{u_k\}$ is uniformly bounded in $L^2(\Omega)$. Assume in addition that there exists a scalar $\tau^* > 0$ such that the stepsizes $\{\tau_k\}$ generated by Algorithm 2 satisfy $\tau_k > \tau^*$ for all sufficiently large $k$.*

1. *If along a subsequence* $\{k_l\}$, $m(\mathcal{A}^0_{k_l}) = 0$, *then there exists a further subsequence* $\{k_{l_m}\}$ *along with a* $u^* \in L^2(\Omega)$ *such that* $u_{k_{l_m}} \xrightarrow{L^2} u^*$ *with*

$$0 \in \partial^\circ \mathcal{J}(u^*), \text{ or equivalently } \mathcal{J}^\circ(u^*; v) \geq 0, \ \forall v \in L^2(\Omega). \tag{46}$$

*In addition there exist* $y^* \in H^2(\Omega) \cap H^1_0(\Omega)$, $p^* \in H^1_0(\Omega)$, $\mu^* \in H^{-1}(\Omega)$, $\lambda^* \in L^2(\Omega)$ *such that* $(u^*, y^*, \lambda^*, p^*, \mu^*)$ *fulfills* (29)-(35) *with* (34) *replaced by:*

$$\langle \mu^*, y^* \rangle_{H^{-1}, H^1_0} = 0. \tag{47}$$

2. *If along a subsequence* $\{k_l\}$, $m(\mathcal{A}^0_{k_l}) > 0$, *then there exists a further subsequence* $\{k_{l_m}\}$ *along with a* $u^* \in L^2(\Omega)$ *such that* $u_{k_{l_m}} \xrightarrow{L^2} u^*$. *In addition, there exist* $y^* \in H^2(\Omega) \cap H^1_0(\Omega)$, $p^* \in H^1_0(\Omega)$, $\mu^* \in H^{-1}(\Omega)$, $\lambda^* \in L^2(\Omega)$ *such that* $(u^*, y^*, \lambda^*, p^*, \mu^*)$ *is* $\mathcal{E}$*-almost-C stationary.*

**Remark 3.20.** *Note that if* $\mathcal{J}$ *is regular in the sense of Clarke [10, Definition 2.3.4] at* $u^*$, *e.g.,* $\mathcal{J}$ *is convex on a neighborhood of* $u^*$, *then* (46) *implies*

$$\mathcal{J}'(u^*; v) \geq 0, \ \forall v \in L^2(\Omega).$$

*In other words, we have found a B-stationary point.*

*Proof.* Let $\{u_k\}$ and $\{h_k\}$ be generated by Algorithm 3 for which $\{\tau_k\}$ is the sequence of associated stepsizes. It follows from Propositions 3.2 and 3.9 that there exists a $c_1 > 0$, independent of the iterations, such that

$$\mathcal{J}(u_{k+1}) - \mathcal{J}(u_k) \leq -\nu \tau_k c_1 ||h_k||^2_{L^2}, \ \forall k \geq 0.$$

Given $\{\mathcal{J}(u_k)\}$ is strictly monotonically decreasing and bounded from below by zero, there exists $\mathcal{J}^* \geq 0$ such that $\mathcal{J}(u_k) \to \mathcal{J}^*$. By definition of $\mathcal{J}$:

$$\frac{\alpha}{2} ||u_k||^2_{L^2} \leq \mathcal{J}(u_k) < \mathcal{J}(u_{k-1}) < \cdots < \mathcal{J}(u_0).$$

Hence, $\{u_k\}$ is bounded (regardless of the behavior of $\tau_k$).

In light of the additional assumptions on $\tau_k$, we extend the inequality derived from the line search to obtain

$$\mathcal{J}(u_{k+1}) - \mathcal{J}(u_k) \leq -\nu \tau_k c_1 ||h_k||^2_{L^2} \leq -\nu \tau^* c_1 ||h_k||^2_{L^2}, \ \forall k \geq 0,$$

Clearly, $\{\mathcal{J}(u_{k+1}) - \mathcal{J}(u_k)\}_{k \geq 0}$ is a null sequence in $\mathbb{R}$. Thus, $h_k \xrightarrow{L^2} 0$.

Suppose now that there exists a subsequence $\{k_l\}$ such that $m(\mathcal{A}^0_{k_l}) = 0$. Since $L^2(\Omega)$ is reflexive, there exists a subsequence $\left\{u_{k_{l_m}}\right\}$ and an element $u^* \in L^2(\Omega)$ such that $u_{k_{l_m}} \xrightarrow{L^2} u^*$. By definition, $y_{k_{l_m}} = S(Bu_{k_{l_m}})$. As $S(B\cdot)$ is completely continuous, $y_{k_{l_m}} \to y^* = S(Bu^*)$ in $H^1_0(\Omega)$. The increased regularity follows from the smoothness of $\partial\Omega$. Continuing, we have $p_{k_{l_m}} = 0, \ a.e. \, \mathcal{A}(y_{k_{l_m}})$ and, from (20), $\langle A p_{k_{l_m}}, p_{k_{l_m}} \rangle = (y_d - y_{k_{l_m}}, p_{k_{l_m}})$. Due to the coercivity

of $A$, $\left\{p_{k_{l_m}}\right\}_m$ is bounded in $H_0^1(\Omega)$. Taking $p_n := p_{k_{l_{m_n}}}$ and $p^* \in H_0^1(\Omega)$ such that $p_n \overset{H_0^1}{\rightharpoonup} p^*$, it follows from the compactness of $B$ that $u_n := \alpha^{-1}B^*p_n$ converges strongly in $L^2(\Omega)$ to $u^*$, i.e.,

$$\alpha u^* = B^*p^*.$$

By noting that $\mathcal{J}'(u_k; v) = \langle \nabla \mathcal{J}(u_k), v\rangle$ for all $v \in L^2(\Omega)$. Hence, $\langle \nabla \mathcal{J}(u_k), v\rangle \leq \mathcal{J}^\circ(u_k; v)$ for all $v \in L^2(\Omega)$. Given $\nabla \mathcal{J}(u_n) = \alpha u_n - B^*p_n \overset{L^2}{\to} 0$, it follows from the strong convergence of $u_n$ and the upper-semicontinuity of the Clarke directional derivative that $0 \leq \mathcal{J}^\circ(u^*; v)$ for all $v \in L^2(\Omega)$. Whence we have (46).

Building on the results in 2., we continue the argument by noting that $\lambda_{k_{l_m}} = Ay_{k_{l_m}} - Bu_{k_{l_m}} - f$ converges strongly in $H^{-1}(\Omega)$ to $Ay^* - Bu^* - f =: \lambda^*$. Moreover, by the regularity theory for variational inequalities and the assumptions on $A$, we also have that $Ay_{k_{l_m}}$ is uniformly bounded in $L^2(\Omega)$, see [8]. Therefore, $\lambda_{k_{l_m}} := Ay_{k_{l_m}} - Bu_{k_{l_m}} - f$ converges weakly in $L^2(\Omega)$, along a subsequence, to $\lambda^* := Ay^* - Bu^* - f$. The state equation (31) and complementarity relation (32) follow. In the sequel, we use the subscript $n$ which refers to the subsequence $\left\{k_{l_{m_n}}\right\}$. Letting $\mu_n := Ap_n + y_n - y_d$, it follows that $\mu_n \overset{H^{-1}}{\rightharpoonup} \mu^*$ as $m \to +\infty$. In which case

$$Ap^* = y_d - y^* + \mu^*,$$

i.e. (30) holds. To prove (33), let $(\cdot)_+ := \max(0, \cdot)$, understood pointwise. Clearly, $(\cdot)_+ : L^2(\Omega) \to L^2(\Omega)$ is Lipschitz continuous with modulus 1. Since $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, $p_n \to p^*$ in $L^2(\Omega)$. Then $(p_n)_+ \to_{L^2} (p^*)_+$ and $(-p_n)_+ \to_{L^2} (-p^*)_+$. Given $\lambda_n \rightharpoonup \lambda^*$ in $L^2(\Omega)$, $(\lambda_n, (p_n)_+) \to (\lambda^*, (p^*)_+)$ and $(\lambda_n, (-p_n)_+) \to (\lambda^*, (-p^*)_+)$. Hence

$$0 = (\lambda_n, (p_n)_+)_{L^2} = \int_\Omega \lambda_n (p_n)_+ dx \to \int_\Omega \lambda^* (p^*)_+ dx$$

Hence, $p^* \leq 0$, $a.e.\, \mathcal{A}^+(y^*, \lambda^*)$. Conducting the same argument with $(-p_n)_+$, we see that $p^* \geq 0$, $a.e.\, \mathcal{A}^+(y^*, \lambda^*)$, (33) follows. Lastly, we see that $\langle \mu_k, y_k\rangle = 0$ for all $k$. Hence, $\langle \mu^*, y^*\rangle = 0$.

By letting $\mu_k := -\gamma_k \chi_{\mathcal{A}_k^+} p_k$ in the nonsmooth case, i.e., $m(\mathcal{A}_k^0) > 0$, the arguments above may be easily adapted. In addition, we can obtain $\mathcal{E}$-almost-C stationarity using the following argument for $\mu_k$. Without loss of generality we use the subscript $k$ (based on the arguments above, the actual subscript should be $k_{l_{m_n}}$)

We first define the following sequence of subsets of $\Omega$:

$$M_1^k := \mathcal{A}(y_k, \lambda_k) \cap \mathcal{I}(y^*) = \{x \in \Omega \,|\, y^*(x) > 0,\, y_k(x) = 0,\, \lambda_k(x) > 0\}.$$

Since $y_k \to y^*$ in $H_0^1(\Omega)$ it follows from $||y_k - y^*||_{L^2}^2 = ||y_k - y^*||_{L^2(M_1^k)}^2 + ||y_k - y^*||_{L^2(\Omega \setminus M_1^k)}^2$, that $||y^*||_{L^2(M_1^k)}^2 = ||y_k - y^*||_{L^2(M_1^k)}^2 \to 0$ and hence, $m(M_1^k) \to 0$.

Continuing, we define the sequence $\{\alpha_k\}$ of non-negative real numbers by setting $\alpha_k := m(M_1^k)$. As argued above, $\{\alpha_k\}$ is a null sequence. Without loss of generality, assume that $\alpha_k$ is strictly monotonically decreasing, otherwise we take a subsequence and continue denoting the sequence members by $\alpha_k$. Now let $\sigma > 0$ and define the subsequence $\{\alpha_{N_l}\}_{l \geq 0}$, by letting $N_l \in \mathbb{N}$ such that $|\alpha_{N_l}| < \frac{\sigma}{2}\left(\frac{1}{2}\right)^l$, $l = 0, 1, 2, \ldots$. For any fixed $L \in \mathbb{N}$ it holds

that $\sum_{l=0}^{L}|\alpha_{N_l}| < \frac{\sigma}{2}\sum_{l=0}^{L}\left(\frac{1}{2}\right)^l = \frac{\sigma}{2}\left(\frac{1-1/2^L}{1-1/2}\right)$. Thus, $\sum_{l=0}^{\infty}|\alpha_{N_l}| = \lim_{L\to+\infty}\sum_{l=0}^{L}|\alpha_{N_l}| \leq \lim_{L\to+\infty}\frac{\sigma}{2}\left(\frac{1-1/2^L}{1-1/2}\right) = \sigma$. Next, define the subset $\mathcal{O}_\sigma$ of $\mathcal{I}(y^*)$ by $\mathcal{O}_\sigma := \bigcup_{l=0}^{\infty}\mathcal{A}(y_{N_l},\lambda_{N_l})\cap\mathcal{I}(y^*)$. Due to the countable subadditivity of the Lebesgue measure, it follows that $m(\mathcal{O}_\sigma) \leq \sum_{l=0}^{\infty}|\alpha_{N_l}| \leq \sigma$. Now let $\varphi \in H_0^1(\Omega)$ such that $\varphi = 0,\ a.e.\,\mathcal{A}(y^*)\cup\mathcal{O}_\sigma$. Then since

$$\gamma_{N_l}\int_\Omega \chi_{\mathcal{A}_{N_l}} p_{N_l}\varphi dx = \gamma_{N_l}\left[\int_{\mathcal{A}(y_{N_l},\lambda_{N_l})\cap\mathcal{I}(y^*)} p_{N_l}\varphi dx + \int_{\mathcal{A}(y_{N_l},\lambda_{N_l})\cap\mathcal{A}(y^*)} p_{N_l}\varphi dx\right] = 0$$

and $\gamma_k\int_\Omega \chi_{\mathcal{A}_k}p_k\varphi dx \to \langle\mu^*,\varphi\rangle$, it holds for all $\sigma > 0$ that there exists a subset $\mathcal{O}_\sigma \subset \mathcal{I}(y^*)$ with $m(\mathcal{O}_\sigma)\leq\sigma$ such that $\langle\mu^*,\varphi\rangle = 0,\ \forall\varphi\in H_0^1(\Omega):\varphi = 0,\ a.e.\,\mathcal{A}(y^*)\cup\mathcal{O}_\sigma$.

□

**Remark 3.21** (**Condition** (46) **and C-stationarity**). *Though at a first glance the characterization* (46) *appears to have very little in common with C-stationarity conditions, it has been shown (in a finite dimensional context) that they are equivalent under suitable assumptions: Existence of a Lipschitz localization of the solution map S of the VI and a generalized LICQ condition for MPCC, [43, Theorem 2.12, Theorem 2.13].*

**Remark 3.22** (**Choosing a Proper Stopping Criterion**). *Based on the theoretical results in Theorems 3.16 and 3.19 and Remark 3.21, it appears that the ideal stopping criterion should be based on the residual of a C-stationarity system,* res_c. *One could also choose the stopping criterion using S-stationarity, however, this should only be done if* $m(\mathcal{A}_k^0) = 0$. *In practice, one works with a discretization of the MPEC. As such, there is no difference between C-stationary and* $\mathcal{E}$*-almost-C-stationarity. Therefore, using C-stationarity conditions is justified. From now on, we check* res_c $\leq$ tol *with* tol$\geq 0$. *In the convergence proofs,* tol$= 0$, *as is usually done.*

Without the stepsize assumption, it can only be ensured that $\tau_k c_1\nu||h_k||_{L^2}^2 \to 0$. Though this does not rule out the possibility that $||h_k||_{L^2}\to 0$. Referring to the proof of Theorem 3.19, we see that this is crucial not only for the termination of the algorithm, but also, it ensures that $\alpha u^* = B^*p^*$.

In order to anticipate this problem, we add an additional step to the algorithm which checks the convergence rate of $\{\tau_k\}$ and $\{||h_k||_{L^2}^2\}$. Given a null sequence $\{q_k\}$ with $q_k\downarrow 0$. We monitor the critical behavior by defining $\sigma_0 := 0$ and for $k\in\mathbb{N}$:

$$\sigma_k := \begin{cases} 1+\sigma_{k-1}, & \tau_k < q_k||h_k||_{L^2}^2 > 0, \\ \sigma_{k-1}, & \text{else.}\end{cases} \tag{48}$$

If at the $k^{th}$ iteration, $\sigma_k$ is "large", then we revert to a "robustification step," calculate a new control, reset the counter $\sigma_k$ and "restart" Algorithm 3.

## 3.4 The Robustification Step

If $\tau_k\downarrow 0$ faster than $||h_k||\to 0$, then we use $\zeta\max_\eta(0,-\cdot)$ with $\zeta,\eta>0$ to smooth the obstacle problem. This transforms the MPEC into the following control problem

$$\min \mathcal{J}_{\zeta,\eta}(u) := \frac{1}{2}||y_{\zeta,\eta}(u)-y_d||_{L^2}^2 + \frac{\alpha}{2}||u||_{L^2}^2 \text{ over } u\in L^2(\Omega). \tag{49}$$

Here, $y_{\zeta,\eta}(u)$ is the unique solution to the problem

$$Ay + \zeta \max_\eta(0, -y) = Bu + f$$

and as before $\max_\eta(0, -\cdot)$ is the global $C^2$-smoothing of the $\max(0, \cdot)$ function. Using standard optimality theory for the optimal control of semilinear equations, see e.g. [42], one arrives at the following first-order necessary optimality conditions for (49): If $u_{\zeta,\eta}$ is a local minimizer of $\mathcal{J}_{\zeta,\eta}$ then there exists an adjoint state $p_{\zeta,\eta} \in H_0^1(\Omega)$ and $y_{\zeta,\eta} \in H_0^1(\Omega)$ such that

$$\alpha u_{\zeta,\eta} = B^* p_{\zeta,\eta}, \tag{50}$$

$$Ap_{\zeta,\eta} - \zeta \max_\eta(0, -y_{\zeta,\eta})' p_{\zeta,\eta} = y_d - y_{\zeta,\eta}, \tag{51}$$

$$Ay_{\zeta,\eta} + \zeta \max_\eta(0, -y_{\zeta,\eta}) = Bu_{\zeta,\eta} + f. \tag{52}$$

The system (50)-(52) can be solved with a standard Newton iteration, thus guaranteeing local superlinear convergence under the usual assumptions.

**Algorithm 4** Descent Method for MPECs II

**Input:** $u_0 \in L^2(\Omega)$; $\gamma_0 > 0$; $\varepsilon > 0$; $\zeta_0 > 0$; $\eta > 0$; $k := 0$; $\rho_2 > 0$; $\rho_3 > 0$;$\{q_k\}$, $q_k \downarrow 0$; $\overline{\sigma} > 0$.

1: $(y_0, \lambda_0)$ solves (16) with $u = u_0$.
2: **while** stopping criterion not fulfilled **do**
3:  **if** $m(\mathcal{A}_k^0) = 0$ **then**
4:  Compute $h_k = Q^{-1}(B^* p_k - \alpha u_k)$, where $p_k$ solves (20) with $y = y_k$.
5:  Compute $(u_{k+1}, y_{k+1}, \lambda_{k+1})$ via Algorithm 2 in direction $h_k$.
6:  **else**
7:  Set $h_k = B^* p_k - \alpha u_k$, where $p_k$ solves (28) with $y = y_k$, $\gamma = \gamma_k$.
8:  **while** (27) fails **do**
9:  Choose $\tilde{\gamma}_k > \gamma_k$.
10:  Set $h_k = B^* p_k - \alpha u_k$, where $p_k$ solves (28) with $y = y_k$, $\gamma = \tilde{\gamma}_k$.
11:  Set $\gamma_k = \tilde{\gamma}_k$.
12:  **end while**
13:  Choose $\gamma_{k+1} > \rho_2 \gamma_k$.
14:  Compute $(u_{k+1}, y_{k+1}, \lambda_{k+1})$ via Algorithm 2 in direction $h_k$
15:  **end if**
16:  Compute $\sigma_{k+1}$ as in (48).
17:  **if** $\sigma_{k+1} > \overline{\sigma}$ **then**
18:  Compute $\tilde{u}_k$ by solving (50)-(52) with $\zeta = \zeta_k$.
19:  Set $\zeta_{k+1} > \rho_3 \zeta_k$, $u_{k+1} = \tilde{u}_k$, $\sigma_{k+1} = 0$.
20:  Let $(y_{k+1}, \lambda_{k+1})$ solve (16) with $u = u_{k+1}$.
21:  **end if**
22:  Set $k := k + 1$.
23: **end while**

**Theorem 3.23** (**Convergence of Algorithm 4**). *Suppose Algorithm 4 generates an infinite sequence of controls $\{u_k\}$. In addition, assume that $\{u_k\}$ is bounded and the robustification step appears infinitely often. Then there exists a subsequence $\{u_{k_l}\}$ and a control*

*$u^* \in L^2(\Omega)$ such that $u_{k_l} \overset{L^2}{\to} u^*$, where $(u^*, y^*)$, with $y^* = S(Bu^*)$, is an $\mathcal{E}$-almost-C-stationarity. If the robustification step appears only finitely often, then the results of Theorem 3.19 hold.*

*Proof.* Given the boundedness of $\{u_k\}$, the proof is identical to that of [21, Theorem 3.4]. □

**Remark 3.24.** *Under the additional assumptions made in [39], e.g., $u^*$ is an isolated locally optimal solution of (11), one can show that $(u^*, y^*)$ satisfies a stronger type of C-stationarity (without the $\mathcal{E}$-almost part).*

# 4 Numerical Experiments

Unlike in the finite dimensional literature, the test set for MPECs in function space is rather small. This makes a comparison to the few existing solution methods rather difficult. For this reason, we select some examples from the literature (see [20, 21]) for our numerical experiments. We present two examples which are nonsmooth at the solution (Examples 4.1,4.2) and one in which only the smooth case occurs (Example 4.3). In both cases, the algorithm performs very well, in particular, it yields an S-stationary point (upon inspection of the numerical result after successful termination of the algorithm). We compare the performance of our algorithm to those in [20, 21] for these examples in the Conclusions.The robustification step was never needed. All examples were solved using MATLAB running on a Dell Optiplex 780, Intel(R) Core(TM)2 Quad CPU Q9550 @ 2.83 GHz.

In all our examples, $\Omega = (0,1) \times (0,1)$. We set $A = -\Delta$, $B = Id_{L^2(\Omega)}$, and discretize $A$ using finite differences with a standard 5-point-stencil. Furthermore, we implement a nested grid strategy, i.e., we solve on a given mesh and pass from this (coarse) mesh to the next (finer) mesh by using a 9-point prolongation stencil, cf. [15]. Whenever $||h_k||_{L^2} \leq$ 1E-6, the residual of the entire C-stationarity system was calculated and the iteration was terminated if the latter was also below 1E-6.

The variational inequality, i.e. the mixed-complementarity problem, was solved using a primal-dual active set strategy, which is known to be equivalent to a nonsmooth Newton step that converges globally and locally at a superlinear rate on each mesh, cf. [19]. We based the stopping criterion on the residual of the entire complementarity system, using the discrete $L^2$-norm, stopping whenever a tolerance of 1E-8 was reached. In order to calculate $d_{\gamma,\varepsilon}$ we employed a standard Newton step, using a tolerance of 1E-8 and starting with $d \equiv 0$. We initialize $\gamma =$ 1E4 and $\varepsilon =$ 1E-3 and increase $\gamma$ by 10 every time a nonsmooth step is needed. In every instance where an approximation of the directional derivative was required, the condition (27), which guarantees descent in nonsmooth cases, held. The performance of the algorithm can be seen in the tables and figures below. See Table 1 for a key to the tables.

**Example 4.1.** This example is taken from Example 5.1 in [21]. Note that strict complementarity fails at the solution. Letting

$$y^\dagger(\mathbf{x}_1, \mathbf{x}_2) = \begin{cases} 1600(\mathbf{x}_1^3 - \mathbf{x}_1^2 + 0.25\mathbf{x}_1)(\mathbf{x}_2^3 - \mathbf{x}_2^2 + 0.25\mathbf{x}_2) & \text{in } (0, 0.5) \times (0, 0.5), \\ 0 & \text{else,} \end{cases}$$

$$\xi^\dagger(\mathbf{x}_1, \mathbf{x}_2) = \max(0, -2|\mathbf{x}_1 - 0.8| - 2|\mathbf{x}_1\mathbf{x}_2 - 0.3| + 0.5),$$

| | |
|---|---|
| DoF | Degrees of freedom |
| Line Searches | Number of times Algorithm 2 was employed |
| Residual of C-stat. | Maximum residual of C-stationarity system |
| S-stat. | "Is the current point S-stationary (Y/N)." |
| Lin. Solves | Total number of linear systems solved |
| ns | Number of nonsmooth steps |
| s | Number of smooth steps |

Table 1: Key for Tables 2, 5, and 8

we set $f = -\Delta y^\dagger - y^\dagger - \xi^\dagger$, and $y_d = y^\dagger + \xi^\dagger - \alpha\Delta y^\dagger$. In addition, $\alpha = 1$, $u_0 = 0$, and $q(h,h) = ||h||^2_{L^2}$. The run data is presented in Tables 2, 3, 4.

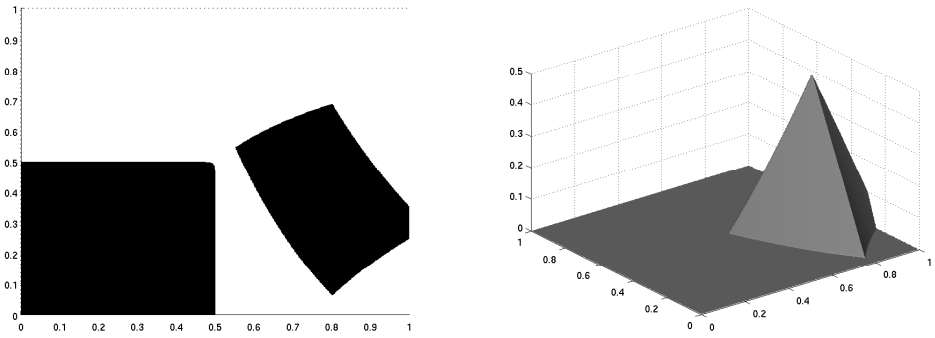

Figure 1: (l.) Biactive Set (white) and (r.) Slack $\lambda^*$ for Example 4.1

**Example 4.2.** In this example, compare with Example 6.1 in [20], we consider yet another case in which strict complementarity fails. Here, we set $u_0 = 0$, $\alpha = 1$, and define $f$ and $y_d$ as follows.

$$\begin{aligned}
z_1(\mathbf{x}_1,\mathbf{x}_2) &= -4096\mathbf{x}_1^6 + 6144\mathbf{x}_2^5 - 3072\mathbf{x}_1^4 + 512\mathbf{x}_2^3,\\
z_2(\mathbf{x}_1,\mathbf{x}_2) &= -244.140625\mathbf{x}_1^6 + 585.9375\mathbf{x}_2^5 - 468.75\mathbf{x}_2^4 + 125\mathbf{x}_2^3,\\
y^*(\mathbf{x}_1,\mathbf{x}_2) &= \begin{cases} z_1(\mathbf{x}_1,\mathbf{x}_2)z_2(\mathbf{x}_1,\mathbf{x}_2) & (\mathbf{x}_1,\mathbf{x}_2) \in ]0,0.5[\times]0,0.8[,\\ 0 & \text{otherwise,}\end{cases}\\
u^* &= y^*,\\
\xi^*(\mathbf{x}_1,\mathbf{x}_2) &= 2\max(0, -|\mathbf{x}_1 - 0.8| - |\mathbf{x}_1\mathbf{x}_2 - 0.2| - 0.3 + 0.35),\\
f &= -\Delta y^* - u^* - \xi^*,\\
y_d &= y^* + \xi^* - \alpha\Delta u^*.
\end{aligned}$$

The run data is presented in Tables 5, 6, 7.

**Example 4.3.** In this example, taken from Example 6.2 in [20], $m(\mathcal{A}_k^0) = 0$ for all $k$. As such, the algorithm reduces to a type of descent method and no smoothing steps are needed.

| DoF | Line Searches | Residual of C-stat. | S-stat. | Lin. Solves | ns | s |
|---|---|---|---|---|---|---|
| 9 | 2 | 1.1959e-08 | Y | 25 | 2 | 0 |
| 49 | 2 | 4.6183e-09 | Y | 29 | 2 | 0 |
| 225 | 2 | 3.7916e-09 | Y | 33 | 2 | 0 |
| 961 | 1 | 5.7524e-07 | Y | 24 | 1 | 0 |
| 3969 | 1 | 4.6164e-07 | Y | 21 | 1 | 0 |
| 16129 | 1 | 4.113e-07 | Y | 18 | 1 | 0 |
| 65025 | 1 | 3.7407e-07 | Y | 15 | 1 | 0 |
| 261121 | 1 | 3.4004e-07 | Y | 15 | 1 | 0 |

Table 2: Performance of Algorithm 4 in Example 4.1.

| ( iter , mesh ) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 3 | 4 | 5 | 5 | 5 | 4 | 4 |
| 1 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 |
| 2 | 2 | 2 | 2 | | | | | |

Table 3: Behavior of Nonsmooth Newton Iteration/Primal-Dual-Active-Set Strategy for Solving the Variational Inequality in Example 4.1. Table entry $(i, j)$ represents number of steps to convergence during $i^{th}$ call to solver on $j^{th}$ mesh with DoF $= (2^{j+1} - 1)^2$.

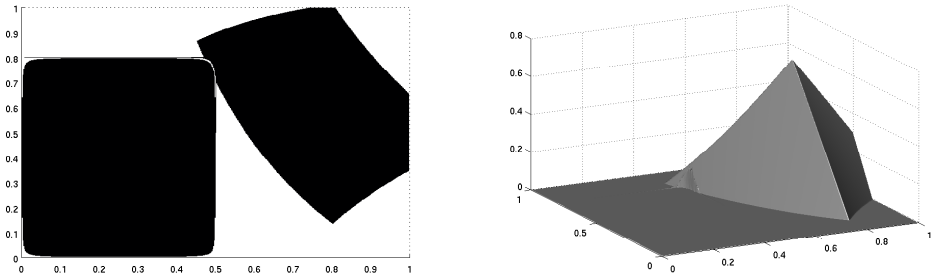

Figure 2: (l.) Biactive Set (white) and (r.) Slack $\lambda^*$ for Example 4.2

The behavior, typical of a descent method, is illustrated in Tables 8 and 9. Though the number of calls to the solver for the linear systems is much higher than in the previous example, the vast majority are done on coarse meshes. The forcing term $f$ and desired state $y_d$ are given by

$$f(\mathbf{x}_1, \mathbf{x}_2) = y_d(\mathbf{x}_1, \mathbf{x}_2) = -|\mathbf{x}_1\mathbf{x}_2 - 0.5| + 0.25.$$

In addition, we set $\alpha$ =5E-2 and $u_0 = 0$.

## 4.1 Conclusions and Outlook

In this paper, we developed a globally convergent solution method in function space for an important class of MPECs via an adaptation of a steepest descent method, which was

| ( iter , mesh ) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 4 | 9 | 7 | 4 | 1 | 2 | 2 |
| 1 | 8 | 7 | 4 | 1 | 1 | 2 | 2 | 2 |
| 2 | 1 | 0 | 0 | | | | | |

Table 4: Behavior of Newton Iteration for finding $d_{\gamma,\varepsilon}$ in Example 4.1. Table entry $(i, j)$ represents number of steps to convergence during $i^{th}$ call to solver on $j^{th}$ mesh with DoF $= (2^{j+1} - 1)^2$.

| DoF | Line Searches | Residual of C-stat. | S-stat. | Lin. Solves | ns | s |
|---|---|---|---|---|---|---|
| 9 | 3 | 1.4228e-08 | Y | 24 | 1 | 2 |
| 49 | 3 | 1.0934e-08 | Y | 33 | 3 | 0 |
| 225 | 2 | 8.9708e-09 | Y | 35 | 2 | 0 |
| 961 | 2 | 7.2331e-09 | Y | 35 | 2 | 0 |
| 3969 | 2 | 5.2003e-08 | Y | 40 | 2 | 0 |
| 16129 | 1 | 8.8768e-07 | Y | 26 | 1 | 0 |
| 65025 | 1 | 8.0202e-07 | Y | 30 | 1 | 0 |
| 261121 | 1 | 7.2865e-07 | Y | 31 | 1 | 0 |

Table 5: Performance of Algorithm 4 in Example 4.2.

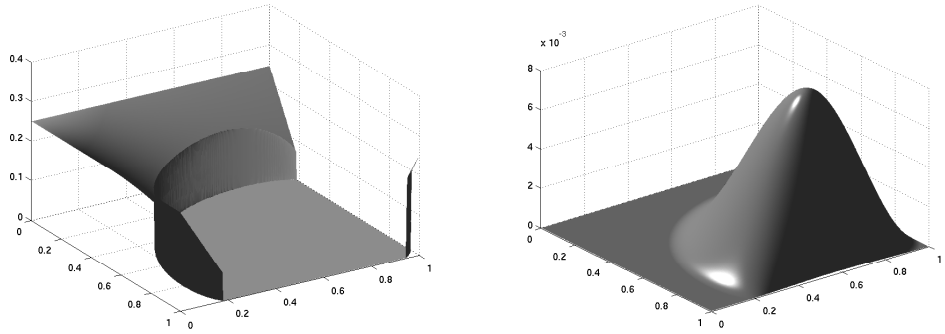


Figure 3: (l.) Slack $\lambda^*$ and (r.) State for Example 4.3

inspired by finite-dimensional results coming from Implicit Programming. By smoothing the directional derivative of the solution operator for the lower-level problem, we were able to avoid possible tractability issues, which would normally be encountered when applying a bundle-method to large-scale nonsmooth optimization problems. As mentioned earlier, the test examples were partly chosen so that we may compare the performance of our algorithm to known results.

Beginning in a qualitative manner, a comparison with the biactive sets from Examples 4.1 and 4.2 to those obtained in [20, 21] shows that by penalizing the original MPEC as in [20, 21], one loses a significant degree of accuracy given the tolerances for relaxation, penalization, and mesh parameters used therein. We also note that whereas the stopping

| ( iter , mesh ) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 0 | | 1 | 7 | 8 | 6 | 3 | 3 | 2 |
| 1 | 1 | 2 | 6 | 5 | 4 | 3 | 2 | 2 |
| 2 | 1 | 1 | 0 | 0 | 2 | | | |
| 3 | | 1 | | | | | | |

Table 6: Behavior of Nonsmooth Newton/Primal-Dual-Active-Set Strategy for Solving the Variational Inequality Example 4.2. Table entry $(i, j)$ represents number of steps to convergence during $i^{th}$ call to solver on $j^{th}$ mesh.

| ( iter , mesh ) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 5 | 5 | 6 | 7 | 7 | 8 | 9 |
| 1 | 3 | 3 | 3 | 2 | 4 | 4 | 5 | 5 |
| 2 | 2 | 3 | 2 | 2 | 3 | | | |
| 3 | 3 | 2 | | | | | | |

Table 7: Behavior of Newton Iteration for finding $d_{\gamma,\varepsilon}$ in Example 4.2. Table entry $(i, j)$ represents number of steps to convergence during $i^{th}$ call to solver on $j^{th}$ mesh.

| DoF | Line Searches | Residual of C-stat. | S-stat. | Lin. Solves | ns | s |
|---|---|---|---|---|---|---|
| 9 | 143 | 9.5975e-07 | Y | 723 | 0 | 143 |
| 49 | 117 | 9.7962e-07 | Y | 593 | 0 | 117 |
| 225 | 92 | 9.9143e-07 | Y | 470 | 0 | 92 |
| 961 | 70 | 9.9389e-07 | Y | 362 | 0 | 70 |
| 3969 | 46 | 9.675e-07 | Y | 240 | 0 | 46 |
| 16129 | 24 | 9.6045e-07 | Y | 130 | 0 | 24 |
| 65025 | 6 | 9.7338e-07 | Y | 42 | 0 | 6 |
| 261121 | 1 | 8.4246e-07 | Y | 17 | 0 | 1 |

Table 8: Performance of Algorithm 4 in Example 4.3

| ( iter , mesh ) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 0 | 3 | 4 | 5 | 5 | 5 | 5 | 6 | 6 |
| 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

Table 9: Behavior of Nonsmooth Newton/Primal-Dual-Active-Set Strategy for Solving the Variational Inequality in Example 4.3. Table entry $(i, j)$ represents number of steps to convergence during $i^{th}$ call to solver on $j^{th}$ mesh. Vertical dots indicate same number of iterations for all $k \geq i$ on mesh $j$.

criteria in [20, 21] are based on the residual of the KKT system for the relaxed problem, ours is based on the actual C/S-stationarity of the current iteration. Nevertheless, the number of

calls to a linear equation solver is comparable. A comprehensive quantitative comparison of all available methods, however, goes beyond the scope of this paper. In this respect, potential future work with respect to the algorithm in this paper would comprise a balancing of the discretization errors with relaxation parameters etc. in a multilevel/multigrid framework.

Finally, we note that the usage the operator $B$ allows for one to consider more general control actions, for example, situations in which the control acts only on the boundary $\partial\Omega$ or on a subset of $\Omega$.

# References


[1] Adams, R. A., and Fournier, J. J.-F. *Sobolev Spaces*, second ed. Elsevier, Amsterdam, 2008.

[2] Aronszajn, N. Differentiability of Lipschitzian mappings between Banach spaces. *Studia Math. 57*, 2 (1976), 147–190.

[3] Attouch, H., Buttazzo, G., and Michaille, G. *Variational Analysis in Sobolev and BV Spaces*. MPS-SIAM Series on Optimization. SIAM and MPS, Philadelphia, 2006.

[4] Barbu, V. *Optimal Control of Variational Inequalities*, vol. 100 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[5] Bergounioux, M. Optimal control of an obstacle problem. *Appl. Math. Optim. 36*, 2 (1997), 147–172.

[6] Bergounioux, M. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim. 36*, 1 (1998), 273–289 (electronic).

[7] Bonnans, J.-F., and Shapiro, A. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, Berlin, 2000.

[8] Brezis, H. R., and Stampacchia, G. Sur la régularité de la solution d'inéquations elliptiques. *Bull. Soc. Math. France 96* (1968), 153–180.

[9] Choquet, G. Theory of capacities. *Ann. Inst. Fourier, Grenoble 5* (1953–1954), 131–295 (1955).

[10] Clarke, F. H. *Optimization and nonsmooth analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1983. A Wiley-Interscience Publication.

[11] De Los Reyes, J. C. Optimal control of a class of variational inequalities of the second kind. *SIAM J. Control Optim. 49*, 4 (2011), 1629–1658.

[12] Fletcher, R., Leyffer, S., Ralph, D., and Scholtes, S. Local convergence of SQP methods for mathematical programs with equilibrium constraints. *SIAM J. Optim. 17*, 1 (2006), 259–286 (electronic).

[13] Fukushima, M., and Pang, J.-S. Convergence of a smoothing continuation method for mathematical programs with complementarity constraints. In *Ill-posed variational problems and regularization techniques (Trier, 1998)*, vol. 477 of *Lecture Notes in Econom. and Math. Systems*. Springer, Berlin, pp. 99–110.

[14] Glowinski, R., Lions, J.-L., and Trémolières, R. *Analyse numérique des inéquations variationnelles. Tome 1*. Dunod, Paris, 1976. Théorie générale premiéres applications, Méthodes Mathématiques de l'Informatique, 5.

[15] Hackbusch, W. *Multigrid methods and applications*, vol. 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[16] Haraux, A. How to differentiate the projection on a convex set in Hilbert space. some applications to variational inequalities. *J. Math. Soc. Japan 29*, 4 (1977), 615–631.

[17] Henrot, A., and Pierre, M. *Variation et optimisation de formes*, vol. 48 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Berlin, 2005.

[18] Herzog, R., Meyer, C., and Wachsmuth, G. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM J. Control Optim. 50*, 5 (2012), 3052–3082.

[19] Hintermüller, M., Ito, K., and Kunisch, K. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim. 13*, 3 (2002).

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

[21] Hintermüller, M., and Kopacka, I. A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs. *Computational Optimization and Applications 50*, 1 (2011), 111–145.

[22] Hintermüller, M., Mordukhovich, B., and Surowiec, T. Several approaches for the derivation of stationarity conditions for elliptic MPECs with upper-level control constraints. *to appear in: Mathematical Programming* (2013).

[23] Hintermüller, M., and Surowiec, T. First order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis. *SIAM J. Optim. 21*, 4 (2012), 1561–1593.

[24] Ito, K., and Kunisch, K. Optimal control of elliptic variational inequalities. *Appl. Math. Optim. 41*, 3 (2000), 343–364.

[25] Kinderlehrer, D., and Stampacchia, G. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York and London, 1980.

[26] Kunisch, K., and Wachsmuth, D. Sufficient optimality conditions and semi-smooth newton methods for optimal control of stationary variational inequalities. *ESAIM: Control, Optimisation and Calculus of Variations 18* (4 2012), 520–547.

[27] Lions, J.-L. Various topics in the theory of optimal control of distributed systems. In *Optimal control theory and its applications (Proc. Fourteenth Biennial Sem. Canad. Math. Congr., Univ. Western Ontario, London, Ont., 1973), Part I*. Springer, Berlin, 1974, pp. 116–309. Lecture Notes in Econom. and Math. Systems, Vol. 105.

[28] Lions, J.-L. *Contrôle des systèmes distribués singuliers*, vol. 13 of *Méthodes Mathématiques de l'Informatique [Mathematical Methods of Information Science]*. Gauthier-Villars, Montrouge, 1983.

[29] Luo, Z.-Q., Pang, J.-S., and Ralph, D. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[30] Mäkelä, M. M., and Neittaanmäki, P. *Nonsmooth optimization*. World Scientific Publishing Co. Inc., River Edge, NJ, 1992. Analysis and algorithms with applications to optimal control.

[31] Mignot, F. Contrôle dans les inéquations variationelles elliptiques. *J. Functional Analysis 22*, 2 (1976), 130–185.

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

[33] Outrata, J., Kočvara, M., and Zowe, J. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, vol. 28 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 1998.

[34] Outrata, J. V., Jařusek, J., and Stará, J. On optimality conditions in control of elliptic variational inequalities. *Set-Valued and Variational Analysis 19* (2011), 23–42.

[35] Preiss, D., and Zajíček, L. Sigma-porous sets in products of metric spaces and sigma-directionally porous sets in Banach spaces. *Real Anal. Exchange 24*, 1 (1998/99), 295–313.

[36] Rodrigues, J.-F. *Obstacle Problems in Mathematical Physics*. No. 134 in North-Holland Mathematics Studies. North-Holland Publishing Co., Amsterdam, 1984.

[37] Scheel, H., and Scholtes, S. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research 25*, 1 (2000), 1–22.

[38] Schiela, A., and Wachsmuth, D. Convergence analysis of smoothing methods for optimal control of stationary variational inequalities. DFG Research Center MATHEON Preprint, Available at: http://opus4.kobv.de/opus4-matheon/frontdoor/index/index/docId/1035, January 2012.

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

[40] Scholtes, S. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim. 11*, 4 (2001), 918–936 (electronic).

[41] Shapiro, A. On concepts of directional differentiability. *JOTA 66*, 3 (1990), 477–487.

[42] Tröltzsch, F. *Optimal control of partial differential equations*, vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[43] Červinka, M. *Hierarchical Structures in Equilibrium Problems*. PhD thesis, Charles University in Prague, 2008.

[44] Wloka, J. *Partielle Differentialgleichungen*. BSB B.G. Teubner Verlagsgesellschaft, Leipzig, 1982.

[45] Yvon, J.-P. *Etude de quelques problèmes de contrôle pour des systèmes distribués*. PhD thesis, 1973. Paris VI.

[46] Yvon, J. P. Optimal control of systems governed by variational inequalities. In *Fifth Conference on Optimization Techniques (Rome, 1973), Part I*. Springer, Berlin, 1973, pp. 265–275. Lecture Notes in Comput. Sci., Vol. 3.

# A Appendix

The capacity of a subset $A \subset \Omega$ can be defined as follows, see e.g. [17] Proposition 3.3.12:

**Definition A.1.**

$$\mathrm{cap}(A) = \inf\{ \int_\Omega |\nabla v|^2 dx, \text{ over } v \in H_0^1(\Omega) \,|\, v \geq 1, \text{ a.e. on an open neighborhood of } A\}.$$

Unlike the Lebesgue measure, $(n-1)$-dimensional manifolds may have positive capacity. The subadditive function $\mathrm{cap}(\cdot)$ is used to define the notion of quasi-continuity: A function $f : \Omega \to \mathbb{R}$ is said to be quasi-continuous provided there exists a nonincreasing sequence of open sets $\Omega_n \subset \Omega$ with $\mathrm{cap}(\Omega_n) \to 0$ such that $f$ is continuous on $\Omega \setminus \Omega_n$ (cf. Chapter 3.3 in [17], Chapter 6.4.3 in [7], or [9]). It is well-known that every $H_0^1(\Omega)$-function $f$ has a unique quasi-continuous representative $\tilde{f}$. Finally, we say that a property holds "quasi-everywhere", denoted here by *q.e.*, provided it holds up to a set of capacity zero. This leads to the following result.

**Lemma A.2.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with $n \geq 1$. Let $\mathcal{A} \subset \Omega$ be compact and assume* (18a) *holds for $\mathcal{A}$. If $d \in H_0^1(\Omega)$ with $d \geq 0$, a.e. $\mathcal{A}$, then $\tilde{d} \geq 0$, q.e. int$(\mathcal{A})$.*

**Remark A.3.** *In other words, in such a situation, there is no difference between almost everywhere and quasi-everywhere on open sets.*

*Proof.* Define $\mathcal{B} \subset \Omega$ by $\mathcal{B} := \left\{ x \in \Omega | \tilde{d}(x) < 0 \right\}$, and note that $\mathcal{B} \cap \mathcal{A}$ has Lebesgue measure zero. By definition of quasi-continuity, there exists a non-increasing sequence of open sets $\Omega_n \subset \Omega$ such that $\tilde{d}|_{\Omega \setminus \Omega_n}$ is continuous and $\mathrm{cap}(\Omega_n) \to 0$ as $n \to +\infty$. It follows then that $\mathcal{B} \setminus \Omega_n$ is open in $\Omega \setminus \Omega_n$ and, consequently, $\mathcal{B} \cup \Omega_n$ is open in $\Omega$. By (18a), $\mathcal{B} \cap \mathrm{int}(\mathcal{A}) \cup \Omega_n$ is open as well. Following the proof of *Lemme* 3.3.30 in [17] or Lemma 6.49 in [7] we deduce $\tilde{d} \geq 0, \ q.e. \, \mathrm{int}(\mathcal{A})$.

□