

## A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs

M. Hintermüller · I. Kopacka

Received: 20 November 2008 / Published online: 4 December 2009  
© Springer Science+Business Media, LLC 2009

**Abstract** An optimal control problem governed by an elliptic variational inequality of the first kind and bilateral control constraints is studied. A smooth penalization technique for the variational inequality is applied and convergence of stationary points of the subproblems to an  $\mathcal{E}$ -almost C-stationary point of the limit problem is shown. The subproblems are solved using a full approximation multigrid scheme (FAS) and alternatively a multigrid method of the second kind for which a convergence result is given. An overall algorithmic concept is provided and its performance is discussed by means of examples.

**Keywords** Constrained optimal control · Elliptic variational inequality · MPEC · Mathematical programs with complementarity constraints · C-stationarity · Nonlinear multigrid

---

M. Hintermüller (✉)

Department of Mathematics, Humboldt-University of Berlin, Unter den Linden 6, 10099 Berlin, Germany

e-mail: [hint@mathematik.hu-berlin.de](mailto:hint@mathematik.hu-berlin.de)

M. Hintermüller · I. Kopacka

Department of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, Heinrichstrasse 36, 8010 Graz, Austria

M. Hintermüller

e-mail: [michael.hintermueller@uni-graz.at](mailto:michael.hintermueller@uni-graz.at)

I. Kopacka

e-mail: [ian.kopacka@uni-graz.at](mailto:ian.kopacka@uni-graz.at)

### 1 Introduction and problem formulation

In this paper we study a new stationarity concept and multigrid algorithms for the following minimization problem  $(\mathcal{P})$ :

$$\begin{aligned} \min \quad & J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2}^2 + \frac{\nu}{2} \|u\|_{L^2}^2 \quad \text{over } y \in V, u \in L^2(\Omega) \\ \text{s.t. (subject to)} \quad & y \in K, \quad a(y, v - y) \geq (f + u, v - y) \quad \forall v \in K, \quad (1.1a) \\ & u \in U_{ad}, \quad (1.1b) \end{aligned}$$

where

$$U_{ad} = \{v \in L^2(\Omega) : a \leq v \leq b \text{ a.e. in } \Omega\},$$

and  $\Omega \subset \mathbb{R}^n, n \leq 3$ , is an open, bounded domain that is either convex and polygonal or has a  $C^{1,1}$ -boundary  $\partial\Omega$ . Further  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ , with  $V = H_0^1(\Omega)$ , denotes the bilinear form

$$a(v, w) = \sum_{i,j=1}^n \int_{\Omega} a_{ij} \frac{\partial v}{\partial x_j} \frac{\partial w}{\partial x_i} dx + \sum_{i=1}^n \int_{\Omega} b_i \frac{\partial v}{\partial x_i} w dx + \int_{\Omega} c v w dx,$$

for all  $v, w \in V$ , where  $b_i, c \in L^\infty(\Omega), a_{ij} \in C^{0,1}(\bar{\Omega})$  and  $c \geq 0$ . We assume that  $a(\cdot, \cdot)$  is bounded and coercive, i.e., there exist constants  $C_B, C_C > 0$  such that for all  $v, w \in V$

$$|a(v, w)| \leq C_B \|v\|_V \|w\|_V \quad \text{and} \quad a(v, v) \geq C_C \|v\|_V^2. \quad (1.2)$$

The variational inequality (VI) constraint (1.1a) involves the cone

$$K = \{v \in V : v \geq 0 \text{ a.e. in } \Omega\}.$$

By  $(\cdot, \cdot)$  and  $\|\cdot\|$  we denote the scalar product and norm, respectively, in  $L^2(\Omega)$  and by  $\langle \cdot, \cdot \rangle$  the duality pairing between  $V$  and  $V^* = H^{-1}(\Omega)$ . Moreover,  $y_d, f \in L^2(\Omega)$  are given,  $\nu > 0$  is fixed, and the bounds on the control variable  $u$  are  $a, b \in L^2(\Omega) \cup \{-\infty, \infty\}$  with  $(b - a) > 0$ . The state variable is denoted by  $y$ . Note that our notation allows us to choose  $a \equiv -\infty, b \equiv \infty$  if no lower or upper bound acts on the control. Further, the cone  $K$  may be replaced by

$$K_\psi = \{v \in V : v \geq \psi \text{ a.e. in } \Omega\},$$

with  $\psi \in H^2(\Omega), \psi|_{\partial\Omega} \leq 0$ , and the subsequent results remain true. For the treatment of the obstacle problem with  $K_\psi$  we refer to [31].

We recall that, given  $u \in L^2(\Omega)$ , the variational inequality (1.1a) admits a unique solution  $y \in V$ . If  $u + f \in L^2(\Omega)$  then  $y \in H^2(\Omega) \cap V$  due to the so-called shift theorem; see, for instance, [13, 15, 25]. In [28] it was proven that  $(\mathcal{P})$  admits a solution.

Our model problem  $(\mathcal{P})$  is a particular instance of a mathematical program with equilibrium constraints (MPEC) in function space. Problems of this type were considered earlier in [1–6, 21, 28, 29] and the references therein. In these papers typically

versions of first order conditions are derived. Let us point out that due to the VI constraint classical constraint qualifications [36] are generically violated by  $(\mathcal{P})$ . Hence, alternative proof concepts for deriving stationarity conditions need to be developed. In the aforementioned papers and monographs mainly two approaches are used: relaxation of a primal-dual form of the VI constraint [2] and regularized penalization ([1, 21, 28], for instance). In [24], as in [2], a relaxation approach was pursued in order to categorize first order conditions similar to finite dimensions [27, 30, 32, 35]. Depending on the realization of the VI constraint and the induced regularity of associated Lagrange multipliers, new versions of stationarity arise in function space; see, e.g., “ $\mathcal{E}$ -almost C-stationarity” (Definition 3.2) and “almost C-stationarity”, concepts which were introduced in [24]. We point out that in finite dimensional space these notions coincide with C-stationarity. Typically, in the aforementioned references, no constraints on the control are considered. Some of the techniques, like the one in [28] for instance which relies on globally optimal solutions of the MPEC (though transferable to local optimality) and conical differentiability of the control-to-state map, do not apply in this situation, in general. Hence, one focus of the present work is to derive first order characterization in the presence of pointwise constraints on the control.

The literature on solution algorithms for  $(\mathcal{P})$  is relatively scarce. In [2] a multiplier method is used but only for the relaxed problems without driving the relaxation parameter to zero. The approach in [21] is based on a nonlinear complementarity (NCP) function based reformulation of the VI constraint and a subsequent least-squares approach to a first order system establishing weak stationarity. Then, in [22] an active-set Newton solver is proposed for obtaining some type of C-stationary points which may be interpreted as a particular version of an implicit programming approach to MPECs; see [30] for the latter. Further, as the proof technique in [24] is constructive an associated solution algorithm is proposed in [24] based on a relaxation scheme and a path-following semismooth Newton solver.

The relaxation approach of [24] is rather general, as it does not rely on the bilevel structure of the underlying problem. Under relatively weak assumptions  $\mathcal{E}$ -almost C-stationarity is derived, and conditions implying stronger optimality conditions, such as C- and strong stationarity, are formulated. While the optimality conditions derived on the basis of a regularized penalty approach as in the present paper are slightly weaker ( $\mathcal{E}$ -almost C-stationarity without further conditions for the stronger concepts), the method however preserves certain problem structures, which might be advantageous in the design of algorithms. For instance,  $(\mathcal{P})$  admits an interpretation as a bilevel optimization problem in function space due to the fact that the VI constraint represents the first order necessary and sufficient condition of a lower level minimization problem. The penalty approach maintains this structure, i.e., after penalizing and smoothing the VI constraint the overall problem still admits a bilevel structure. In addition, the resulting minimization problem resembles an optimal control problem for a semilinear elliptic partial differential equation with point-wise constraints on the control.

Based on [7, 8, 10, 17, 18, 20], in this paper we introduce multigrid solvers for these problems. In fact, we study the full approximation scheme (FAS) [7, 10] and a multigrid method of the second kind [18] and embed these methods in a continuation

scheme for the penalty and/or smoothing parameter. Under a non-degeneracy condition, we establish a convergence result for the version based on the multigrid method of the second kind. To the best of our knowledge these are the first multigrid-based solvers for MPECs in function space.

The rest of the paper is organized as follows. In the next section we introduce our smoothed penalty approach and study some of its properties. Then we derive first order optimality conditions in Sect. 3. Here we focus on the concept of  $\mathcal{E}$ -almost C-stationarity. Section 4 introduces first a FAS-scheme and then a nonlinear multigrid method of the second kind for solving the penalized and smoothed problems. For the latter approach a convergence result is given. The paper ends by a report on numerical tests comparing both variants of multigrid solvers and different types of penalization.

## 2 Penalization of the variational inequality

A common technique for solving a variational inequality of the form (1.1a) is by penalization (see, e.g., [14, 25]). Using this technique the variational inequality is approximated by a sequence of nonlinear boundary value problems. For a given penalty operator  $\pi : V \rightarrow V^*$  with the following properties:

$$\left\{ \begin{array}{l} \pi \text{ is Lipschitz continuous,} \\ \text{Ker}(\pi) = K, \\ \pi \text{ is monotone,} \end{array} \right. \quad (2.1)$$

the VI can be approximated by the penalized equation

$$a(y, v) + \frac{1}{\alpha} \langle \pi(y), v \rangle = (f + u, v) \quad \forall v \in V, \quad (2.2)$$

with  $\alpha > 0$  being the *penalty parameter*. Due to the monotonicity of the nonlinear operator  $\pi$ , (2.2) has a unique solution (see, e.g., [14]) which we denote by  $y_\alpha(u)$ . Furthermore if we denote the solution of (1.1a) by  $y(u)$ , it is well known that

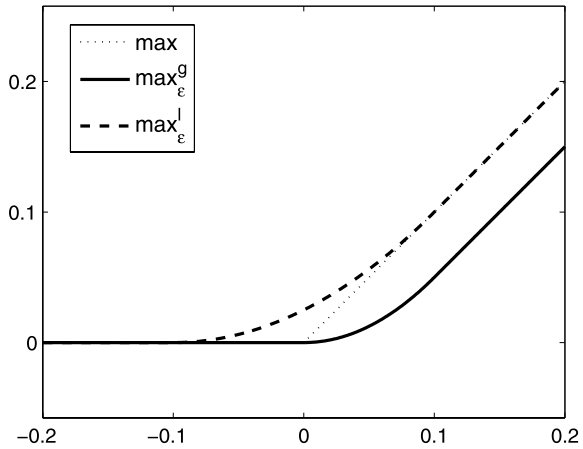
$$y_\alpha(u) \rightarrow y(u) \quad \text{in } V \text{ as } \alpha \downarrow 0 \quad (2.3)$$

(see, e.g., [14]). In this paper we consider the penalty operator

$$\pi(v) := -\max(0, -v) \quad \forall v \in V,$$

where the max-operation is to be understood point-wise. As the  $\max(0, \cdot)$ -function is not differentiable at the origin we introduce different regularizations yielding  $C^1$ -approximations of  $\max(0, \cdot)$ . For a fixed *smoothing parameter*  $\varepsilon > 0$  we define the

**Fig. 1**  $\max(0, \cdot)$  and regularizations  $\max_\varepsilon^g$  and  $\max_\varepsilon^l$



following regularized/smoothed operators:

$$\begin{aligned} \max_\varepsilon^g(0, r) &:= \begin{cases} r - \frac{\varepsilon}{2} & \text{if } r \geq \varepsilon, \\ \frac{r^2}{2\varepsilon} & \text{if } r \in (0, \varepsilon), \\ 0 & \text{if } r \leq 0, \end{cases} \\ \max_\varepsilon^l(0, r) &:= \begin{cases} r & \text{if } r \geq \varepsilon, \\ \frac{r^2}{4\varepsilon} + \frac{r}{2} + \frac{\varepsilon}{4} & \text{if } r \in (-\varepsilon, \varepsilon), \\ 0 & \text{if } r \leq -\varepsilon. \end{cases} \end{aligned} \tag{2.4}$$

Both regularizations are  $C^1$ -functions that smooth the kink of the max-operator. As illustrated in Fig. 1 the regularization  $\max_\varepsilon^l$  alters the function locally around the origin but leaves the operator unchanged outside of this neighborhood, whereas the smoothing of  $\max_\varepsilon^g$  affects the operator on the whole half-plane. We will therefore refer to the two operators as “local” and “global” regularizations, respectively. In the context of penalization we note that the global regularization itself defines a suitable penalty operator for the variational inequality, i.e., it satisfies (2.1), while the local regularization  $\max_\varepsilon^l$  violates the second condition of (2.1) for every  $\varepsilon > 0$ .

In the following lemma useful properties of these operators are provided. We also state some of the differences between the regularizations.

**Lemma 2.1**

- (i) Both  $\max_\varepsilon^g$  and  $\max_\varepsilon^l$  are monotone and convex.
- (ii) The global variant  $\max_\varepsilon^g$  penalizes the set  $K$ , i.e.,  $\max_\varepsilon^g(0, -\cdot)$  satisfies the conditions (2.1). The local variant  $\max_\varepsilon^l$  is not a suitable penalty operator for  $K$ . The violation of (2.1) can be estimated by:

$$\|\max_\varepsilon^l(0, -v)\|_{L^2} \leq \frac{\varepsilon}{4} \sqrt{\text{meas}(\Omega)} \quad \forall v \in K.$$

Furthermore if  $\max_\varepsilon^l(0, -v) = 0$  then  $v \in K$ .

(iii) We have  $(\max_{\varepsilon}^g(0, -v), v) \leq 0$  for all  $v \in L^2(\Omega)$  and the local variant satisfies

$$(\max_{\varepsilon}^l(0, -v), v) \leq \frac{\varepsilon^2}{4} \text{meas}(\Omega) \quad \forall v \in L^2(\Omega).$$

(iv) For all  $r \in \mathbb{R}$  we find that

$$0 \leq \max(0, r) - \max_{\varepsilon}^g(0, r) \leq \frac{\varepsilon}{2},$$

$$0 \leq \max_{\varepsilon}^l(0, r) - \max(0, r) \leq \frac{\varepsilon}{4}.$$

(v) Both  $(\max_{\varepsilon}^g(0, r))'$  and  $(\max_{\varepsilon}^l(0, r))' \in [0, 1]$  for all  $r \in \mathbb{R}$ . As a consequence both regularizations are Lipschitz continuous with Lipschitz constant 1.

(vi) The global regularization has a Lipschitz property with respect to  $\varepsilon$ , i.e.,

$$|\max_{\varepsilon_1}^g(0, r) - \max_{\varepsilon_2}^g(0, r)| \leq \frac{3}{2} |\varepsilon_1 - \varepsilon_2|$$

for all  $r \in \mathbb{R}$  and  $\varepsilon_1, \varepsilon_2 > 0$ .

*Proof* Assertions (i), (iv) and (v) are evident. For the estimate in (ii) assume that  $v \geq 0$  a.e. in  $\Omega$ . Then

$$\|\max_{\varepsilon}^l(0, -v)\|_{L^2}^2 = \int_{\{-\varepsilon < -v \leq 0\}} \max_{\varepsilon}^l(0, -v)^2 \leq \left(\frac{\varepsilon}{4}\right)^2 \int_{\{-\varepsilon < -v \leq 0\}} 1 \leq \left(\frac{\varepsilon}{4}\right)^2 \text{meas}(\Omega),$$

as  $0 \leq \max_{\varepsilon}^l(0, r) \leq \frac{\varepsilon}{4}$  for all  $-\varepsilon \leq r \leq 0$ . The proof of the estimate for  $\max_{\varepsilon}^l$  in (iii) is done analogously. In order to prove (vi) w.l.o.g. let  $\varepsilon_1 < \varepsilon_2$  and  $r \in \mathbb{R}$ . We investigate several cases. If  $r \leq 0$  then  $\max_{\varepsilon_1}^g(0, r) - \max_{\varepsilon_2}^g(0, r) = 0$ . If  $0 < r < \varepsilon_1$  then

$$|\max_{\varepsilon_1}^g(0, r) - \max_{\varepsilon_2}^g(0, r)| = \frac{r^2}{2} \left(\frac{1}{\varepsilon_1} - \frac{1}{\varepsilon_2}\right) = \frac{r^2}{2} \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_1 \varepsilon_2} \leq \frac{1}{2} (\varepsilon_2 - \varepsilon_1).$$

If  $\varepsilon_1 \leq r < \varepsilon_2$  then

$$\begin{aligned} |\max_{\varepsilon_1}^g(0, r) - \max_{\varepsilon_2}^g(0, r)| &= r - \frac{\varepsilon_1}{2} - \frac{1}{2\varepsilon_2} r^2 \leq \varepsilon_2 - \frac{\varepsilon_1}{2} - \frac{\varepsilon_1^2}{2\varepsilon_2} \\ &= \frac{\varepsilon_2}{2} - \frac{\varepsilon_1}{2} + \frac{\varepsilon_2^2}{2\varepsilon_2} - \frac{\varepsilon_1^2}{2\varepsilon_2} = \frac{1}{2} \left(1 + \frac{\varepsilon_2 + \varepsilon_1}{\varepsilon_2}\right) (\varepsilon_2 - \varepsilon_1) \\ &\leq \frac{3}{2} (\varepsilon_2 - \varepsilon_1). \end{aligned}$$

Finally if  $\varepsilon_2 \leq r$  then

$$|\max_{\varepsilon_1}^g(0, r) - \max_{\varepsilon_2}^g(0, r)| = r - \frac{\varepsilon_1}{2} - \left(r - \frac{\varepsilon_2}{2}\right) = \frac{1}{2} (\varepsilon_2 - \varepsilon_1). \quad \square$$

In Sect. 4.1.2 we consider a special multigrid method to solve a penalized version of our MPEC-problem. Its convergence analysis requires additional smoothness of the regularization of the max-operation. We therefore introduce a  $C^2$ -regularization, which we present in its “global” form only, for the sake of simplicity:

$$\max_{\varepsilon}^G(0, r) := \begin{cases} r - \frac{\varepsilon}{2} & \text{if } r \geq \varepsilon, \\ \frac{r^3}{\varepsilon^2} - \frac{r^4}{2\varepsilon^3} & \text{if } r \in (0, \varepsilon), \\ 0 & \text{if } r \leq 0. \end{cases} \tag{2.5}$$

*Remark 2.2* The results of Lemma 2.1 remain true if  $\max_{\varepsilon}^g$  is replaced by  $\max_{\varepsilon}^G$  and the Lipschitz constant in (vi) is replaced by  $7/2$ .

Due to the non-smoothness of the  $\max(0, \cdot)$ -operator we are interested in a result similar to (2.3) but using the smoothed operators. As noted earlier the global regularization satisfies the penalization properties whereas the local regularization violates (2.1) for all  $\varepsilon > 0$ . This is reflected in the fact that the smoothing parameter  $\varepsilon$  has to be driven to zero in order to achieve convergence if the local regularization is used, while the global regularization permits the choice of a fixed positive smoothing parameter. We therefore make the following assumption on the smoothing parameter.

**Assumption 2.1** For each  $\alpha > 0$  let  $\varepsilon(\alpha) > 0$  be given such that

- (i)  $\{\varepsilon(\alpha)\}$  is bounded if  $\max_{\varepsilon}^g$  or  $\max_{\varepsilon}^G$  is used and
- (ii)  $\frac{\varepsilon(\alpha)}{\alpha} \rightarrow 0$  for  $\alpha \rightarrow 0$  if  $\max_{\varepsilon}^l$  is used.

The following theorem describes the approximation properties of the regularized penalized equations.

**Theorem 2.3** Let  $\{u_{\alpha}\}$  be a sequence in  $L^2(\Omega)$ ,  $u \in L^2(\Omega)$  such that  $u_{\alpha} \rightarrow u$  (strongly) in  $V^*$ . Let  $\{\varepsilon(\alpha)\}$  satisfy Assumption 2.1 and  $y_{\alpha} \in V$  denote the solution of the regularized penalized equation

$$\mathcal{A}y_{\alpha} - \frac{1}{\alpha} \max_{\varepsilon(\alpha)}(0, -y_{\alpha}) = u_{\alpha} + f, \tag{2.6}$$

where  $\max_{\varepsilon}$  stands for either of the regularizations introduced in the beginning of this section. Then

$$y_{\alpha} \rightarrow y(u) \quad \text{in } V \text{ if } \alpha \rightarrow 0,$$

where  $y(u)$  denotes the solution of (1.1a).

*Proof* We use similar arguments as in [14], Theorem 3.1, Chap. 1. In the course of this paper we will frequently operate on subsequences which we will, for the sake of readability, not always denote specifically.

In the first part of the proof we consider the global regularizations  $\max_{\varepsilon}^g$  and  $\max_{\varepsilon}^G$ . For the sake of simplicity we give the proof for  $\max_{\varepsilon}^g$  only, but all of our

arguments apply to the  $C^2$ -regularization. Let  $v \in K$  be arbitrarily fixed. Due to the monotonicity of  $\max_{\varepsilon}^g$  and Lemma 2.1(ii) we estimate

$$\begin{aligned} 0 &\leq \left( -\max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}) + \max_{\varepsilon(\alpha)}^g(0, -v), y_{\alpha} - v \right) \\ &= -\left( \max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}), y_{\alpha} - v \right). \end{aligned} \tag{2.7}$$

Using (2.6) we then obtain

$$\langle \mathcal{A}y_{\alpha}, y_{\alpha} \rangle \leq \langle u_{\alpha} + f, y_{\alpha} - v \rangle + \langle \mathcal{A}y_{\alpha}, v \rangle. \tag{2.8}$$

Hence the boundedness and the coercivity of  $\mathcal{A}$  yield the boundedness of  $\{y_{\alpha}\}$  in  $V$ . From (2.6) we infer

$$\| \max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}) \|_{V^*} = \alpha \| \mathcal{A}y_{\alpha} - u_{\alpha} - f \|_{V^*} = \mathcal{O}(\alpha),$$

and therefore

$$\max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}) \rightarrow 0 \quad \text{in } V^* \text{ if } \alpha \rightarrow 0. \tag{2.9}$$

As  $\{y_{\alpha}\}$  is bounded, there exists  $\bar{y} \in V$  such that (on a subsequence denoted the same)  $y_{\alpha} \rightarrow \bar{y}$  weakly in  $V$  and strongly in  $L^2(\Omega)$ . Since  $\{\varepsilon(\alpha)\}$  is a bounded sequence there exists yet another subsequence and a non-negative parameter  $\bar{\varepsilon}$ , such that  $\varepsilon(\alpha) \rightarrow \bar{\varepsilon}$  on that subsequence. Due to the Lipschitz properties (v) and (vi) of Lemma 2.1 we find that

$$\begin{aligned} &\| \max_{\bar{\varepsilon}}^g(0, -\bar{y}) - \max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}) \|_{V^*} \\ &\leq \| \max_{\bar{\varepsilon}}^g(0, -\bar{y}) - \max_{\bar{\varepsilon}}^g(0, -y_{\alpha}) \|_{V^*} + \| \max_{\bar{\varepsilon}}^g(0, -y_{\alpha}) - \max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}) \|_{V^*} \\ &\leq C (\| \bar{y} - y_{\alpha} \| + | \bar{\varepsilon} - \varepsilon(\alpha) |) \rightarrow 0. \end{aligned}$$

This, combined with (2.9), shows that  $\max_{\bar{\varepsilon}}^g(0, -\bar{y}) = 0$  in  $V^*$ . Due to the density of  $V$  in  $L^2(\Omega)$  we find that  $\max_{\bar{\varepsilon}}^g(0, -\bar{y})$  also vanishes in  $L^2(\Omega)$ , hence  $\bar{y} \in K$ . Due to the weak convergence of  $y_{\alpha}$  and the coercivity of  $\mathcal{A}$ , taking the  $\liminf_{\alpha \rightarrow 0}$  in (2.8) yields

$$\langle \mathcal{A}\bar{y}, \bar{y} \rangle \leq \langle u + f, \bar{y} - v \rangle + \langle \mathcal{A}\bar{y}, v \rangle, \quad \forall v \in K.$$

Therefore  $\bar{y} = y(u)$ . Due to the uniqueness of the solution of the variational inequality we conclude that in fact the whole sequence  $\{y_{\alpha}\}$  converges (not just a subsequence). Finally the strong convergence of  $y_{\alpha}$  follows from the coercivity of  $\mathcal{A}$ , (2.6) and (2.7) since

$$\begin{aligned} C_C \| y_{\alpha} - \bar{y} \|_V^2 &\leq \langle \mathcal{A}(y_{\alpha} - \bar{y}), y_{\alpha} - \bar{y} \rangle \\ &= \langle u_{\alpha} + f, y_{\alpha} - \bar{y} \rangle + \frac{1}{\alpha} (\max_{\varepsilon(\alpha)}^g(0, -y_{\alpha}), y_{\alpha} - \bar{y}) - \langle \mathcal{A}\bar{y}, y_{\alpha} - \bar{y} \rangle \\ &\leq \langle u_{\alpha} + f, y_{\alpha} - \bar{y} \rangle - \langle \mathcal{A}\bar{y}, y_{\alpha} - \bar{y} \rangle \rightarrow 0. \end{aligned}$$



The proof for  $\max_{\varepsilon}^l$  uses the same arguments, but replaces the estimate (2.7) by

$$0 \leq -\left(\max_{\varepsilon(\alpha)}^l(0, -y_{\alpha}), y_{\alpha} - v\right) + \frac{1}{4} \text{meas}(\Omega)^{1/2} \varepsilon(\alpha) \|y_{\alpha} - v\|$$

and uses  $\varepsilon(\alpha)/\alpha \rightarrow 0$ . □

### 3 The optimal control problem

In this paper we solve the MPEC ( $\mathcal{P}$ ) by approximating the lower-level problem, i.e., the variational inequality, using the techniques developed in the previous section. For each set of parameters  $(\alpha, \varepsilon) > 0$  we therefore define the smoothed penalized problem ( $\mathcal{P}_{\alpha, \varepsilon}$ ) by

$$\begin{aligned} \min \quad & J(y, u) = \frac{1}{2} \|y - y_d\|_{L^2}^2 + \frac{\nu}{2} \|u\|_{L^2}^2 \quad \text{over } y \in V, u \in L^2(\Omega) \\ \text{s.t.} \quad & \mathcal{A}y - \frac{1}{\alpha} \max_{\varepsilon}(0, -y) = u + f, \\ & a \leq u \leq b \quad \text{a.e. in } \Omega, \end{aligned} \tag{3.1}$$

where  $\max_{\varepsilon}$  is a generic notation that stands for either of the regularizations introduced in Sect. 2. Using arguments similar to those in [24] and considering the properties listed in Lemma 2.1 the existence of a global optimal point for problem ( $\mathcal{P}_{\alpha, \varepsilon}$ ) can be shown. The following result describes the convergence behavior of globally optimal points.

**Theorem 3.1** *Let  $\alpha$  and  $\varepsilon(\alpha)$  be given such that Assumption 2.1 is satisfied and denote by  $(y_{\alpha}, u_{\alpha})$  the global solutions of the smoothed-penalized problem ( $\mathcal{P}_{\alpha, \varepsilon}$ ). Then there exists a global optimal point  $(\bar{y}, \bar{u}) \in V \times L^2(\Omega)$  for the original problem ( $\mathcal{P}$ ) such that*

$$(y_{\alpha}, u_{\alpha}) \rightarrow (\bar{y}, \bar{u}) \quad \text{in } V \times L^2(\Omega).$$

*Proof* Let  $\tilde{u} \in U_{ad}$  be fixed and denote by  $y_{\alpha}(\tilde{u})$  the solution of the regularized penalized equation (2.6) with right hand side  $\tilde{u} + f$ . Using Theorem 2.3 we can then estimate

$$J(y_{\alpha}, u_{\alpha}) \leq J(y_{\alpha}(\tilde{u}), \tilde{u}) = \frac{1}{2} \|y_{\alpha}(\tilde{u}) - y_d\|_{L^2}^2 + \frac{\nu}{2} \|\tilde{u}\|_{L^2}^2 \leq C \quad \forall \alpha > 0,$$

with a constant  $C > 0$  independent of  $\alpha$ . Hence, from the definition of  $J$  we find that  $\{u_{\alpha}\}$  is bounded in  $L^2(\Omega)$  and there exists  $\bar{u} \in L^2(\Omega)$  such that on a subsequence (also denoted by  $\{u_{\alpha}\}$ )  $u_{\alpha}$  converges to  $\bar{u}$  weakly in  $L^2(\Omega)$  and strongly in  $H^{-1}(\Omega)$ . As  $U_{ad}$  is closed and convex it is also weakly closed and  $\bar{u} \in U_{ad}$ . Using Theorem 2.3 again we find that  $y_{\alpha} \rightarrow \bar{y} := y(\bar{u})$  in  $V$ . Therefore,  $(\bar{y}, \bar{u})$  is feasible for the

MPEC  $(\mathcal{P})$ . Now let  $(y^*, u^*) \in V \times L^2(\Omega)$  be a global solution of  $(\mathcal{P})$ . Due to the lower semicontinuity of the involved norm and Theorem 2.3 we find that

$$\begin{aligned} J(y^*, u^*) &\leq J(\bar{y}, \bar{u}) \leq \liminf_{\alpha \rightarrow 0} J(y_\alpha, u_\alpha) \leq \limsup_{\alpha \rightarrow 0} J(y_\alpha, u_\alpha) \\ &\leq \limsup_{\alpha \rightarrow 0} J(y_\alpha(u^*), u^*) = \limsup_{\alpha \rightarrow 0} \left( \frac{1}{2} \|y_\alpha(u^*) - y_d\|_{L^2}^2 + \frac{\nu}{2} \|u^*\|_{L^2}^2 \right) \\ &= J(y^*, u^*). \end{aligned}$$

Therefore,  $(\bar{y}, \bar{u})$  is optimal for  $(\mathcal{P})$ . Furthermore we see that  $J(y_\alpha, u_\alpha) \rightarrow J(\bar{y}, \bar{u})$ . Hence,  $\|u_\alpha\|_{L^2}^2 \rightarrow \|\bar{u}\|_{L^2}^2$ , which, together with the weak convergence of  $\{u_\alpha\}$ , implies strong convergence of  $\{u_\alpha\}$  in  $L^2(\Omega)$ .  $\square$

From an algorithmic point of view, however, finding global solutions of an optimization problem can be a difficult task. It is therefore preferable to use an analysis that relies on stationary points only. In [24] some stationarity concepts resembling those of finite dimensional MPECs (see, e.g., [32, 33, 35]) were introduced for the infinite dimensional problem (1.1).

In this section we derive optimality conditions for the smoothed-penalized problem and show that accumulation points of such stationary points satisfy the conditions of so called  $\mathcal{E}$ -almost C-stationarity for the original problem  $(\mathcal{P})$ ; see Definition 3.2, which generalizes the corresponding definitions to accommodate the bilateral control constraints. For the sake of brevity we set  $\Omega^+ := \{x \in \Omega : y(x) > 0\}$ .

**Definition 3.2** The point  $(y, u) \in V \times L^2(\Omega)$  is called  $\mathcal{E}$ -almost C-stationary for problem (1.1), if there exist  $\xi \in L^2(\Omega)$ ,  $p \in V$  and  $\lambda \in V^*$  such that the following system is satisfied:

$$y - \lambda + \mathcal{A}^* p = y_d, \tag{3.2a}$$

$$\mathcal{A}y - u - \xi = f, \tag{3.2b}$$

$$u \in U_{ad}, \quad (\nu u - p, v - u) \geq 0 \quad \forall v \in U_{ad}, \tag{3.2c}$$

$$\xi \geq 0 \quad \text{a.e.}, \quad y \geq 0, \quad (y, \xi) = 0, \tag{3.2d}$$

$$p = 0 \quad \text{a.e. in } \{\xi > 0\}, \tag{3.2e}$$

$$\langle \lambda, p \rangle \leq 0, \quad \langle \lambda, y \rangle = 0, \tag{3.2f}$$

and furthermore for every  $\tau > 0$  there exists a subset  $E_\tau \subset \Omega^+$  with  $meas(\Omega^+ \setminus E_\tau) \leq \tau$  such that

$$\langle \lambda, \phi \rangle = 0 \quad \forall \phi \in V, \quad \phi = 0 \quad \text{a.e. in } \Omega \setminus E_\tau. \tag{3.3}$$

Note that in the absence of control constraints ( $U_{ad} = L^2(\Omega)$ ) the optimality condition (3.2c) reduces to  $\nu u = p \in V$ . Hence the control  $u$  gains regularity. For this particular problem the concept of strong stationarity was introduced in [24], where individual sign conditions for the multipliers  $p$  and  $\lambda$  are imposed. If  $U_{ad} = L^2(\Omega)$ ,

a point  $(y, u, \xi, p, \lambda)$  is called strongly stationary for the MPEC (1.1) if (3.2a)–(3.2e) hold true and further (3.2f) and (3.3) are replaced by the (stronger) conditions

$$p \leq 0 \quad \text{a.e. in } B, \tag{3.4}$$

$$\langle \lambda, \phi \rangle \geq 0, \quad \forall \phi \in V, \quad \phi \geq 0 \quad \text{a.e. in } B, \quad \phi = 0 \quad \text{a.e. in } \{\xi > 0\}, \tag{3.5}$$

where  $B := \{y = 0\} \cap \{\xi = 0\}$  is the so-called *biactive set*.

As the penalty approach acts as a relaxation of the feasible domain of the MPEC, standard theory can be applied to derive first order optimality conditions for the smoothed penalized problem  $(\mathcal{P}_{\alpha,\varepsilon})$ ; see, e.g., Zowe and Kurcyusz [36]. For details, see [26].

**Theorem 3.3** *Let  $\alpha, \varepsilon > 0$  and  $(y, u) \in V \times L^2(\Omega)$  be an optimal solution of (3.1). Then there exists an adjoint state  $p \in V$  such that*

$$y + \mathcal{A}^* p + \frac{1}{\alpha} \max'_\varepsilon(0, -y)p = y_d, \tag{3.6a}$$

$$\mathcal{A}y - \frac{1}{\alpha} \max_\varepsilon(0, -y) = u + f, \tag{3.6b}$$

$$u \in U_{ad}, \quad (vu - p, v - u) \geq 0 \quad \forall v \in U_{ad}. \tag{3.6c}$$

Note that the variational inequality (3.6c) (or (3.2c)) is equivalent to

$$u = \frac{1}{v} p - \max\left(0, \frac{1}{v} p - b\right) + \max\left(0, -\left(\frac{1}{v} p - a\right)\right), \tag{3.7}$$

which is further equivalent to the projection

$$u = P_{[a,b]}\left(\frac{1}{v} p\right) := \begin{cases} a & \text{where } \frac{1}{v} p < a, \\ \frac{1}{v} p & \text{where } a \leq \frac{1}{v} p \leq b, \\ b & \text{where } \frac{1}{v} p > b, \end{cases} \tag{3.8}$$

see, e.g., [25, 34]. We now state the main result of this paper describing the convergence properties of stationary points of the smoothed penalized problem with respect to the penalty and smoothing parameters.

**Theorem 3.4** *Consider the smooth penalized problem with bounds  $a, b \in L^2(\Omega)$  and a  $C^1$ -regularization of the max-operator. Let  $\{\alpha\}$  and  $\{\varepsilon(\alpha)\}$  satisfy Assumption 2.1. For every  $\alpha > 0$  let  $(y_\alpha, u_\alpha) \in V \times L^2(\Omega)$  be stationary points of the smooth penalized problem (3.1) with corresponding adjoint state  $p_\alpha \in V$ .*

*Then there exist  $(\tilde{y}, \tilde{u}, \tilde{\xi}, \tilde{p}, \tilde{\lambda}) \in V \times L^2(\Omega) \times L^2(\Omega) \times V \times V^*$  and a subsequence (again denoted by  $\{\alpha\}$ ) such that  $y_\alpha \rightarrow \tilde{y}$  in  $V$ ,  $u_\alpha \rightarrow \tilde{u}$  in  $L^2(\Omega)$ ,  $\frac{1}{\alpha} \max_{\varepsilon(\alpha)}(0, -y_\alpha) \rightarrow \tilde{\xi}$  in  $V^*$ ,  $p_\alpha \rightarrow \tilde{p}$  in  $V$  and  $-\frac{1}{\alpha} \max'_{\varepsilon(\alpha)}(0, -y_\alpha)p_\alpha \rightarrow \tilde{\lambda}$  in  $V^*$ , where  $(\tilde{y}, \tilde{u})$  is  $\mathcal{E}$ -almost  $C$ -stationary for the MPEC (1.1).*

*Proof* For the sake of readability we omit the argument of the smoothing parameter and write  $\varepsilon$  instead of  $\varepsilon(\alpha)$ .

(i) *Convergence and feasibility of  $\{(y, u, \xi, p)\}$*  First we show the convergence properties of the sequences  $\{y_\alpha\}$ ,  $\{u_\alpha\}$ ,  $\{p_\alpha\}$ ,  $\{\max_{\varepsilon(\omega)}(0, -y_\alpha)\}$  and verify that the limit elements satisfy equations (3.2b)–(3.2d): From the bilateral constraints  $u_\alpha \in U_{ad}$  it follows that  $\{\|u_\alpha\|_{L^2}\}$  is bounded. Therefore there exists  $\tilde{u} \in L^2(\Omega)$  and a subsequence (denoted the same) such that  $u_\alpha$  converges to  $\tilde{u}$  weakly in  $L^2(\Omega)$  and strongly in  $V^*$ . The sets

$$\{v \in L^2(\Omega) : v \geq a \text{ a.e. in } \Omega\} \quad \text{and} \quad \{v \in L^2(\Omega) : v \leq b \text{ a.e. in } \Omega\}$$

are convex and closed and therefore weakly closed. Hence the limit element satisfies

$$a \leq \tilde{u} \leq b \quad \text{a.e. in } \Omega. \tag{3.9}$$

Using the convergence of  $u_\alpha$  in  $V^*$ , Theorem 2.3 yields the existence of  $\tilde{y} \in V$  with  $y_\alpha \rightarrow \tilde{y}$  in  $V$  and  $(\tilde{y}, \tilde{u})$  solving the variational inequality (1.1a). Note that due to the regularity of the domain  $\Omega$  and the right hand side  $\tilde{u} + f$  the solution  $\tilde{y}$  gains regularity and is in  $H^2(\Omega) \cap V$ . Hence introducing a slack variable (Lagrange multiplier)  $\tilde{\xi} \in L^2(\Omega)$  the VI can equivalently be written as

$$\mathcal{A}\tilde{y} - \tilde{\xi} = \tilde{u} + f, \quad \tilde{y} \geq 0, \quad \tilde{\xi} \geq 0, \quad (\tilde{y}, \tilde{\xi})_{L^2} = 0,$$

see, e.g., [25]. Setting

$$\xi_\alpha := \frac{1}{\alpha} \max_\varepsilon(0, -y_\alpha),$$

we find that

$$\xi_\alpha = \mathcal{A}y_\alpha - u_\alpha - f \rightarrow \mathcal{A}\tilde{y} - \tilde{u} - f = \tilde{\xi} \quad \text{in } V^*.$$

Multiplying the adjoint equation (3.6a) by  $p_\alpha$  and using the coercivity of the operator  $\mathcal{A}$  and the fact that  $\max'_\varepsilon(0, r) \in [0, 1]$  for all  $\varepsilon > 0$  and  $r \in \mathbb{R}$ , we estimate

$$\begin{aligned} C_C \|p_\alpha\|_V^2 &\leq \langle \mathcal{A}^* p_\alpha, p_\alpha \rangle = \langle y_d - y_\alpha, p_\alpha \rangle - \frac{1}{\alpha} \int_\Omega \max'_\varepsilon(0, -y_\alpha) p_\alpha^2 \, dx \\ &\leq \|y_d - y_\alpha\|_{V^*} \|p_\alpha\|_V. \end{aligned}$$

Therefore,  $\{p_\alpha\}$  is bounded in  $V$  and there exists another subsequence (again denoted the same) and an element  $\tilde{p} \in V$  such that  $p_\alpha \rightarrow \tilde{p}$  weakly in  $V$  and strongly in  $L^2(\Omega)$ . Due to the Lipschitz-continuity of the  $\max(\cdot, 0)$ -operator from  $L^2(\Omega)$  to  $L^2(\Omega)$ , (3.7) yields

$$\begin{aligned} u_\alpha &= \frac{1}{\nu} p_\alpha - \max\left(0, \frac{1}{\nu} p_\alpha - b\right) + \max\left(0, -\left(\frac{1}{\nu} p_\alpha - a\right)\right) \\ &\rightarrow \frac{1}{\nu} \tilde{p} - \max\left(0, \frac{1}{\nu} \tilde{p} - b\right) + \max\left(0, -\left(\frac{1}{\nu} \tilde{p} - a\right)\right) = \tilde{u}, \end{aligned}$$

where the convergence is to be understood in  $L^2(\Omega)$ .

(ii) *Convergence and feasibility of  $\lambda$*  In the next step we show the convergence of  $\{-\frac{1}{\alpha} \max'_\varepsilon(0, -y_\alpha)p_\alpha\}$  to  $\tilde{\lambda}$ , satisfying (3.2a), (3.2f) and (3.3). Setting

$$\lambda_\alpha := -\frac{1}{\alpha} \max'_\varepsilon(0, -y_\alpha)p_\alpha,$$

and considering the adjoint equations (3.6a) and (3.2a), we find that  $\lambda_\alpha \rightarrow \tilde{\lambda}$  in  $V^*$ . Next we show that  $\langle \tilde{\lambda}, \tilde{y} \rangle = 0$ .

For  $\max^g_\varepsilon$  we find that  $(\max^g_\varepsilon)'(0, r) = 0$  for all  $r \leq 0$ . Hence

$$\langle \lambda_\alpha, y_\alpha^+ \rangle = \left\langle -\frac{1}{\alpha} (\max^g_\varepsilon)'(0, -y_\alpha)p_\alpha, y_\alpha^+ \right\rangle = 0 \quad \forall \alpha > 0,$$

where  $y_\alpha^+ := \max(0, y_\alpha)$ . As  $y_\alpha \rightarrow \tilde{y}$  in  $V$ , also  $y_\alpha^+ \rightarrow \tilde{y}^+ = \tilde{y}$  in  $V$  (see Lemma A.1 in the Appendix). Therefore,

$$\langle \tilde{\lambda}, \tilde{y} \rangle = \lim_{\alpha \rightarrow 0} \langle \lambda_\alpha, y_\alpha^+ \rangle = \lim_{\alpha \rightarrow 0} 0 = 0.$$

For  $\max^l_\varepsilon$  it is easy to see that  $(\max^l_\varepsilon)'(0, r) \in [0, \frac{1}{2}]$  for all  $r \leq 0$ . Exploiting the fact that  $\varepsilon/\alpha \rightarrow 0$  we find that

$$|\langle \lambda_\alpha, y_\alpha^+ \rangle| \leq \frac{1}{\alpha} \int_{\{0 \leq y_\alpha \leq \varepsilon\}} \left| (\max^l_\varepsilon)'(0, -y_\alpha)p_\alpha y_\alpha \right| dx \leq \frac{\varepsilon}{2\alpha} \int_{\{0 \leq y_\alpha \leq \varepsilon\}} |p_\alpha| dx \rightarrow 0,$$

as  $p_\alpha$  is bounded in  $L^1(\Omega)$ . Hence also in this case  $\langle \tilde{\lambda}, \tilde{y} \rangle = 0$ .

We now show that  $\langle \tilde{\lambda}, \tilde{p} \rangle \leq 0$ . Multiplying the adjoint equation (3.6a) by  $p_\alpha$  yields

$$\langle \mathcal{A}^* p_\alpha, p_\alpha \rangle + \langle y_\alpha - y_d, p_\alpha \rangle = -\frac{1}{\alpha} \langle \max'_\varepsilon(0, -y_\alpha)p_\alpha, p_\alpha \rangle \leq 0 \quad \forall \alpha > 0.$$

Due to boundedness and coercivity of  $\mathcal{A}^*$ , we find that

$$\langle \mathcal{A}^* \tilde{p}, \tilde{p} \rangle + \langle \tilde{y} - y_d, \tilde{p} \rangle \leq \liminf_{\alpha \rightarrow 0} (\langle \mathcal{A}^* p_\alpha, p_\alpha \rangle + \langle y_\alpha - y_d, p_\alpha \rangle) \leq 0.$$

Using (3.2a), we see that this is equivalent to

$$\langle \tilde{\lambda}, \tilde{p} \rangle \leq 0.$$

Next we verify condition (3.3) for  $\tilde{\lambda}$ . Note that  $y_\alpha \rightarrow \tilde{y}$  in  $V$  and hence also in  $L^2(\Omega)$ . Therefore there exists a subsequence (denoted the same) such that  $y_\alpha \rightarrow \tilde{y}$  point-wise a.e. in  $\Omega$ . Now let  $x \in \{\tilde{y} > 0\}$ . Then there exists  $\alpha_1 > 0$  such that

$$y_\alpha(x) > \frac{\tilde{y}(x)}{2} > 0$$

for all  $\alpha < \alpha_1$ . If we use regularization  $\max^g_\varepsilon$ , we immediately see that for all  $\alpha < \alpha_1$  this yields that  $(\max^g_\varepsilon)'(0, -y_\alpha(x)) = 0$  and hence  $\lambda_\alpha(x) = 0$ . If regularization  $\max^l_\varepsilon$

is used, we furthermore know that  $\varepsilon(\alpha) \rightarrow 0$ . Hence there exists  $\alpha_2 > 0$  such that

$$\varepsilon(\alpha) < \frac{\tilde{y}(\mathbf{x})}{2}$$

for all  $\alpha < \alpha_2$ . We therefore find that

$$-y_\alpha(\mathbf{x}) < -\varepsilon(\alpha)$$

for all  $\alpha < \min(\alpha_1, \alpha_2)$ , and hence  $\lambda_\alpha(\mathbf{x}) = 0$  for all  $\alpha < \min(\alpha_1, \alpha_2)$ . We therefore find that in both cases

$$\lambda_\alpha \rightarrow 0 \quad \text{p.w. a.e. in } \{\tilde{y} > 0\}.$$

Using Egorov’s theorem this yields the condition for  $\tilde{\lambda}$  required for  $\mathcal{E}$ -almost C-stationarity.

(iii) *Condition for  $\tilde{p}$*  It remains to show that  $\tilde{p}$  satisfies (3.2e). First we take a look at the global regularization  $\max_\varepsilon^\xi$ . For the sake of brevity we define the sets

$$M_1 = M_1(\alpha) := \{0 < -y_\alpha < \varepsilon\}, \quad M_2 = M_2(\alpha) := \{-y_\alpha \geq \varepsilon\}.$$

Due to the convergence of  $(\xi_\alpha, y_\alpha) \rightarrow 0$  we find that

$$(\xi_\alpha, y_\alpha) = \frac{1}{\alpha} \int_{M_1} \frac{1}{2\varepsilon} y_\alpha^3 + \frac{1}{\alpha} \int_{M_2} \left(-y_\alpha - \frac{\varepsilon}{2}\right) y_\alpha \rightarrow 0. \tag{3.10}$$

As both terms of the sum are non-positive we see that they both individually tend to zero. For the first term on the right hand side of (3.10) this yields

$$\left\| \frac{1}{\sqrt{\alpha\varepsilon}} \chi_{M_1} |y_\alpha|^{3/2} \right\|_{L^2} \rightarrow 0. \tag{3.11}$$

Further note that due to the definition of  $M_2$  we have the estimate

$$\left| y_\alpha(\mathbf{x}) + \frac{\varepsilon}{2} \right| = -y_\alpha(\mathbf{x}) - \frac{\varepsilon}{2} \leq -y_\alpha(\mathbf{x}) = |y_\alpha(\mathbf{x})|$$

for almost every  $\mathbf{x} \in M_2$ . This estimate together with the second term of the sum on the right hand side of (3.10) then yields

$$\left\| \frac{1}{\sqrt{\alpha}} \chi_{M_2} \left(y_\alpha + \frac{\varepsilon}{2}\right) \right\|_{L^2} \rightarrow 0. \tag{3.12}$$

Next recall that  $\{\lambda_\alpha\}$  and  $\{p_\alpha\}$  are bounded in  $V^*$  and  $V$ , respectively. Therefore  $|\langle \lambda_\alpha, p_\alpha \rangle| \leq C$  for some constant  $C > 0$  independent of  $\alpha$ . The definition of  $\lambda_\alpha$  then

yields

$$\begin{aligned}
 |(\lambda_\alpha, p_\alpha)| &= \left| -\frac{1}{\alpha} \int_{M_1} \frac{1}{\varepsilon} (-y_\alpha) p_\alpha^2 - \frac{1}{\alpha} \int_{M_2} p_\alpha^2 \right| \\
 &= \frac{1}{\alpha} \int_{M_1} \frac{1}{\varepsilon} (-y_\alpha) p_\alpha^2 + \frac{1}{\alpha} \int_{M_2} p_\alpha^2 \leq C.
 \end{aligned}
 \tag{3.13}$$

As again both terms of the sum have the same sign, we find that they are both individually bounded by  $C$ , i.e.,

$$\left\| \frac{1}{\sqrt{\alpha\varepsilon}} \chi_{M_1} |y_\alpha|^{1/2} p_\alpha \right\|_{L^2}^2 \leq C \quad \text{and} \quad \left\| \frac{1}{\sqrt{\alpha}} \chi_{M_2} p_\alpha \right\|_{L^2}^2 \leq C.
 \tag{3.14}$$

Furthermore as both integrands are non-negative we find that they remain bounded if we only integrate over a subset. Hence

$$\left\| \frac{1}{\sqrt{\alpha\varepsilon}} \chi_{M_1} |y_\alpha|^{1/2} p_\alpha^+ \right\|_{L^2}^2 \leq C \quad \text{and} \quad \left\| \frac{1}{\sqrt{\alpha}} \chi_{M_2} p_\alpha^+ \right\|_{L^2}^2 \leq C.
 \tag{3.15}$$

Combining (3.11), (3.12) and (3.14), we find that

$$\begin{aligned}
 (\xi_\alpha, p_\alpha) &= \frac{1}{2} \left( \frac{1}{\sqrt{\alpha\varepsilon}} \chi_{M_1} |y_\alpha|^{3/2}, \frac{1}{\sqrt{\alpha\varepsilon}} \chi_{M_1} |y_\alpha|^{1/2} p_\alpha \right) \\
 &\quad - \left( \frac{1}{\sqrt{\alpha}} \chi_{M_2} \left( y_\alpha + \frac{\varepsilon}{2} \right), \frac{1}{\sqrt{\alpha}} \chi_{M_2} p_\alpha \right) \rightarrow 0.
 \end{aligned}
 \tag{3.16}$$

Therefore

$$(\tilde{\xi}, \tilde{p}) = 0.
 \tag{3.17}$$

Note that due to (3.15) the computations in (3.16) also apply if  $p_\alpha$  is replaced by  $p_\alpha^+$ . Therefore using the weak convergence of  $p_\alpha^+$  to  $\tilde{p}^+$  in  $V$  (see Lemma A.1) we find that also  $(\tilde{\xi}, \tilde{p}^+) = 0$ . The non-negativity of  $\tilde{\xi}$  then ensures that  $\tilde{p} = 0$  a.e. in  $\{\tilde{\xi} > 0\}$ .

Finally we show that  $\tilde{p} = 0$  a.e. in  $\{\tilde{\xi} > 0\}$  if regularization  $\max_\varepsilon^l$  is used. The proof is along the same lines as for the global regularization but is somewhat simplified by the fact that we can use the property  $\varepsilon/\alpha \rightarrow 0$ . We find that the part of  $\xi_\alpha$  defined on  $\{|y_\alpha| < \varepsilon\}$  can be neglected. In fact

$$\begin{aligned}
 \|\xi_\alpha\|_{L^2(\{|y_\alpha| < \varepsilon\})}^2 &= \frac{1}{\alpha^2} \int_{\{|y_\alpha| < \varepsilon\}} \left( \frac{1}{4\varepsilon} y_\alpha^2 - \frac{1}{2} y_\alpha + \frac{\varepsilon}{4} \right)^2 \\
 &\leq \left( \frac{\varepsilon}{\alpha} \right)^2 \int_{\{|y_\alpha| < \varepsilon\}} \left( \frac{1}{4} + \frac{1}{2} + \frac{1}{4} \right) \rightarrow 0.
 \end{aligned}$$

Hence we find that

$$\hat{\xi}_\alpha := -\frac{1}{\alpha} \chi_{M_2} y_\alpha \rightarrow \tilde{\xi} \quad \text{in } V^*. \tag{3.18}$$

Using the boundedness of  $\langle \lambda_\alpha, p_\alpha \rangle$  it can be shown that both  $\|\frac{1}{\sqrt{\alpha}} \chi_{M_2} p_\alpha\|_{L^2}$  and  $\|\frac{1}{\sqrt{\alpha}} \chi_{M_2} p_\alpha^+\|_{L^2}$  are bounded. Furthermore it follows from the fact that  $(\hat{\xi}_\alpha, y_\alpha) \rightarrow 0$  that  $\|\frac{1}{\sqrt{\alpha}} \chi_{M_2} y_\alpha\|_{L^2} \rightarrow 0$ . Then

$$(\hat{\xi}_\alpha, p_\alpha) = -\left(\frac{1}{\sqrt{\alpha}} \chi_{M_2} y_\alpha, \frac{1}{\sqrt{\alpha}} \chi_{M_2} p_\alpha\right) \rightarrow 0.$$

This yields  $(\tilde{\xi}, \tilde{p}) = 0$  and similarly  $(\tilde{\xi}, \tilde{p}^+) = 0$  and thus the desired property for  $\tilde{p}$ .  $\square$

We note that the availability of a stronger stationarity principle might depend on the structure of the active and inactive sets of both, the control and the state variable.

*Remark 3.5* (i) The results of Theorem 3.4 remain true if  $\max_\varepsilon^g$  is replaced by  $\max_\varepsilon^G$ . In the proof the specific definition of the operator is only used to show that  $\tilde{p} = 0$  a.e. on  $\{\tilde{\xi} > 0\}$ . For the operator  $\max_\varepsilon^G$  the same proof technique can be employed as for  $\max_\varepsilon^g$ , if the inequality

$$0 \leq \frac{y_\alpha^2}{\varepsilon^2} + \frac{y_\alpha^3}{2\varepsilon^3} \leq \frac{1}{2} \left( \frac{3y_\alpha^2}{\varepsilon^2} + \frac{2y_\alpha^3}{\varepsilon^3} \right) \quad \text{a.e. in } M_1$$

is used in step (3.16).

(ii) If  $a \equiv -\infty$  or  $b \equiv \infty$  is chosen, the results of Theorem 3.4 hold true if additionally the boundedness of  $\{u_\alpha\}$  in  $L^2(\Omega)$  is assumed. In the control-unconstrained case, i.e.,  $a \equiv -\infty, b \equiv \infty$ , the optimality condition (3.6c) is simplified to  $u_\alpha = \frac{1}{v} p_\alpha$ . Therefore  $u_\alpha$  adopts the regularity of  $p_\alpha$  and we find that  $\tilde{u} \in V$  and  $u_\alpha \rightharpoonup \tilde{u}$  in  $V$ .

### 4 The algorithm

The constructive nature of Theorem 3.4 implies an infinite dimensional solution algorithm for the MPEC problem  $(\mathcal{P})$ , where the solution is approximated by the solution of the smoothed penalized problems  $(\mathcal{P}_{\alpha,\varepsilon})$  along a sequence  $\{(\alpha, \varepsilon(\alpha))\}$ . This outer algorithm is described in Algorithm 1. Theorem 3.4 then yields the convergence of the iterates to a point that is  $\varepsilon$ -almost C-stationary for the MPEC problem. The state variable  $y_\alpha$  and the adjoint  $p_\alpha$  converge in  $V$  strongly and weakly, respectively. In the control unconstrained case the control  $u_\alpha$  also converges weakly in  $V$ . Whereas if point-wise bilateral constraints are imposed on the control, then  $u_\alpha$  only converges (strongly) in  $L^2(\Omega)$ . Numerical results illustrating the approximation error of the state  $y_\alpha$  with respect to the penalty parameter are given in Sect. 5.3.



---

**Algorithm 1** (Outer Algorithm)

---

**Data:**  $y_d, f, \alpha_0, \varepsilon_0$ .

- 1: Choose  $(y^0, u^0, p^0)$  and set  $k := 0$ .
  - 2: **repeat**
  - 3:   Solve the first order system (3.6) with  $(\alpha, \varepsilon) = (\alpha_k, \varepsilon_k)$ , to obtain  $(y^{k+1}, u^{k+1}, p^{k+1})$  using initial values  $(y^k, u^k, p^k)$  for solution algorithm.
  - 4:   Choose  $\alpha_{k+1} < \alpha_k$  and  $\varepsilon_{k+1}$ .
  - 5:   Set  $k := k + 1$ .
  - 6: **until** some stopping rule is satisfied.
- 

4.1 Solving the subproblem

For the numerical realization of Algorithm 1 the subproblems are discretized using a uniform grid with mesh size  $h$ . The resulting large nonlinear systems are solved by employing two different multigrid techniques. The first approach uses the Full Approximation Scheme (FAS) due to Brandt [10], to solve the optimality system for the state  $y$  and the adjoint state  $p$  (see also [7–9]). In the second approach the optimality conditions are reduced to a fixed-point equation in the control  $u$  and solved using a nonlinear multigrid method of second kind (see Hackbusch, [16–20]).

4.1.1 FAS multigrid

In the optimality system (3.6) equation (3.7) can be used to eliminate the control  $u$ . The resulting reduced nonlinear system reads

$$\begin{aligned}
 y + \mathcal{A}^* p + \frac{1}{\alpha} \max'_\varepsilon(0, -y) p &= y_d, \\
 \mathcal{A} y - \frac{1}{\alpha} \max_\varepsilon(0, -y) - \frac{1}{\nu} p + \max\left(0, \frac{1}{\nu} p - b\right) & \\
 - \max\left(0, -\frac{1}{\nu} p + a\right) &= f.
 \end{aligned}
 \tag{4.1}$$

This system is subsequently solved using a FAS method [10], which is a natural choice when solving nonlinear PDE-based problems. We briefly outline the method and propose a smoothing technique which is tailored to the specific problem under consideration.

For a given sequence of grids let  $h$  denote the size of the current grid and  $H > h$  denote the mesh size of the next coarser grid. Furthermore let  $U_h$  and  $U_H$  denote the finite dimensional subspaces of  $L^2(\Omega)$  corresponding to the mesh sizes  $h$  and  $H$ . Let  $I_h^H : U_h \rightarrow U_H$  and  $\hat{I}_h^H : U_h \rightarrow U_H$  denote the restriction operators using full weighting and straight injection, respectively, and let  $I_H^h : U_H \rightarrow U_h$  denote the interpolation operator using linear interpolation (see, e.g., [20] for the definition of the operators). Let  $y_d^h, f^h$  denote the discretized data and let  $h_{\max} > 0$  denote the grid size of the coarsest mesh, i.e., the mesh on which the system is solved exactly. Then

**Algorithm 2** (FAS-algorithm)

- Data:**  $(y_d^h, f^h)$ , initial values  $(y^h, p^h)$ .
- 1: **if**  $h = h_{\max}$  **then**
  - 2:   solve (4.1) exactly.
  - 3: **else**
  - 4:   Do  $v_{pre}$  pre-smoothing steps.
  - 5:   Compute residuals  $r_a^h = y_d^h - y^h - (\mathcal{A}^h)^* p^h - \frac{1}{\alpha} \max'_\varepsilon(0, -y^h) p^h$  and  $r_s^h = f^h - \mathcal{A}^h y^h + \frac{1}{\alpha} \max_\varepsilon(0, -y^h) - \frac{1}{\nu} p^h$ .
  - 6:   Restrict residuals to coarser grid using *full weighting*  $r_a^H = I_h^H r_a^h, r_s^H = I_h^H r_s^h$ .
  - 7:   Restrict  $(y^h, p^h)$  to coarser grid using *straight injection*  $y^H = \hat{I}_h^H y^h, p^H = \hat{I}_h^H p^h$ .
  - 8:   Compute right hand sides for coarse problem  $y_d^H = r_a^H + y^H + (\mathcal{A}^H)^* p^H + \frac{1}{\alpha} \max'_\varepsilon(0, -y^H) p^H, f^H = r_s^H + \mathcal{A}^H y^H - \frac{1}{\alpha} \max_\varepsilon(0, -y^H) - \frac{1}{\nu} p^H$ .
  - 9:   Apply  $\gamma$ -times the FAS-algorithm to coarse grid problem with  $h = H$ , right hand side  $(y_d^H, f^H)$  and initial values  $(y^H, p^H)$  to compute new approximation  $(\tilde{y}^H, \tilde{p}^H)$ .
  - 10:   Coarse grid correction:  $y^h = y^h + I_H^h(\tilde{y}^H - y^H), p^h = p^h + I_H^h(\tilde{p}^H - p^H)$ , where  $I_H^h$  denotes the interpolation operator using *linear interpolation*.
  - 11:   Do  $v_{post}$  post-smoothing steps.
  - 12: **end if**

the FAS-algorithm for solving the discrete optimality conditions is given by Algorithm 2 (see, e.g., [8, 12]). Note that in step 9 the algorithm is recursively applied  $\gamma$ -times. Here  $\gamma \in \mathbb{N}$  is called the cycle index. For  $\gamma = 1$  the algorithm defines a V-cycle and for  $\gamma = 2$  a W-cycle. On the coarsest mesh the system is solved exactly. This exact solve is realized using an iterative solver, e.g., a semismooth-Newton method, with a small stopping tolerance (e.g.,  $\sqrt{\epsilon_m}$ , where  $\epsilon_m$  is the machine precision), [23]. The integers  $v_{pre}$  and  $v_{post}$  denote the number of smoothing steps that are applied to the approximations of the solution. These numbers are crucial for the convergence of the algorithm. For well conditioned problems  $v_{pre} = v_{post} = 2$  is usually sufficient but as the conditioning becomes worse (e.g., the cost of the control  $\nu$  becomes small) the number of necessary smoothing steps may increase.

As a smoothing technique we apply a collective nonlinear Gauss-Seidel scheme where the nonlinear scalar equations are solved analytically. For a standard five-point finite difference stencil, the discretized version of system (4.1) is:

$$\begin{aligned}
 & y_{ij} + \frac{4}{h^2} p_{ij} + \frac{1}{\alpha} \max'_\varepsilon(0, -y_{ij}) p_{ij} = B_{ij}, \\
 & \frac{4}{h^2} y_{ij} - \frac{1}{\alpha} \max_\varepsilon(0, -y_{ij}) - \frac{1}{\nu} p_{ij} + \max\left(0, \frac{1}{\nu} p_{ij} - b_{ij}\right) \\
 & \quad + \max\left(0, -\frac{1}{\nu} p_{ij} + a_{ij}\right) = A_{ij}
 \end{aligned} \tag{4.2}$$

with

$$B_{ij} = (y_d)_{ij} + \frac{1}{h^2}(p_{i+1,j} + p_{i-1,j} + p_{i,j+1} + p_{i,j-1}),$$

$$A_{ij} = f_{ij} + \frac{1}{h^2}(y_{i+1,j} + y_{i-1,j} + y_{i,j+1} + y_{i,j-1})$$

for all inner grid points  $(i, j)$ . For given vectors  $y^{(k-1)}$  and  $p^{(k-1)}$  the smoothed approximations  $y^{(k)}$  and  $p^{(k)}$  are obtained by solving (4.2) with right hand sides

$$B_{ij}^{(k)} = (y_d)_{ij} + \frac{1}{h^2}(p_{i+1,j}^{(k-1)} + p_{i-1,j}^{(k)} + p_{i,j+1}^{(k-1)} + p_{i,j-1}^{(k)}),$$

$$A_{ij}^{(k)} = f_{ij} + \frac{1}{h^2}(y_{i+1,j}^{(k-1)} + y_{i-1,j}^{(k)} + y_{i,j+1}^{(k-1)} + y_{i,j-1}^{(k)})$$

for each grid point, looping over the columns and rows. Depending on the realization of  $\max_{\varepsilon}$ , the Gauss-Seidel system (4.2) can be reduced to a cubic equation if a  $C^1$ -regularization is used and to an equation of fifth order if the  $C^2$ -regularization is used. As the system has to be solved for every grid point in every smoothing iteration the use of a cost efficient, yet still accurate solution method is of importance. In the  $C^1$ -case the cubic equation is solved analytically using Cardano’s method. The solution of a polynomial of order five, however, is more intricate and in general finding an analytical solution is a difficult task. In view of the computational cost the use of the  $C^1$ -regularization therefore seems favorable.

For the overall algorithm the FAS-method is incorporated into a nested iteration technique, utilizing a grid hierarchy for the outer loop as well, where the penalty parameter  $\alpha$  is coupled with the mesh size  $h$  and an inexact solution technique is used. In each outer iteration the subproblem is solved on a fixed grid for fixed parameters  $(\alpha, \varepsilon)$  using the FAS-method of Algorithm 2 with a stopping tolerance depending on the current mesh size. After each outer iteration  $\alpha$  is decreased. If the penalty parameter drops below a threshold depending on the mesh size, e.g.,

$$\alpha \leq Ch^{2/s} \tag{4.3}$$

where  $C > 0$  is a suitably chosen constant (e.g., 0.5), and  $s \in (0.5, 1)$ , then the mesh is refined. This condition will be motivated in more detail in Sect. 5.3. The outer algorithm terminates when  $\alpha_k$  satisfies (4.3) for the finest mesh.

The results of our numerical tests runs are given in Sect. 5.5. Although the performance of the method is good, the convergence of the algorithm is still an issue which cannot be guaranteed at this point. We therefore present an alternative multigrid method together with a convergence analysis.

#### 4.1.2 Multigrid method of the second kind

An alternative to the FAS-method is the so called “multigrid iteration of the second kind”, a scheme designed to solve fixed-point equations. It uses the multigrid framework, employing Picard’s iteration as smoothing step. The convergence of the Picard

iteration (contraction of the operator), however, is not necessary for the convergence of the overall algorithm.

In order to apply the multigrid method of second kind we reformulate the optimality conditions of  $(\mathcal{P}_{\alpha,\varepsilon})$  as a fixed point problem. We introduce the control-to-state mapping  $\Lambda_S : L^2(\Omega) \rightarrow Y$  with  $Y := V \cap C(\bar{\Omega})$ , i.e., for a given control  $v \in L^2(\Omega)$  we define  $\Lambda_S(v) = y_\alpha(v)$  as the solution of the state equation

$$\mathcal{A}y_\alpha(v) - \frac{1}{\alpha} \max_{\varepsilon(\alpha)}(0, -y_\alpha(v)) = f + v.$$

We further define the state-to-adjoint mapping  $\Lambda_A : Y \rightarrow V$  in a similar way. For given state  $y$  the corresponding adjoint state  $\Lambda_A(y) = p_\alpha(y)$  is defined as the solution of the adjoint equation

$$\mathcal{A}^* p_\alpha(y) + \frac{1}{\alpha} \max'_{\varepsilon(\alpha)}(0, -y) p_\alpha(y) = y_d - y. \tag{4.4}$$

The optimality conditions for  $(\mathcal{P}_{\alpha,\varepsilon})$  are then equivalent to the fixed-point problem

$$u = F(u), \tag{4.5}$$

with the nonlinear operator  $F : L^2(\Omega) \rightarrow L^2(\Omega)$  given by  $F(v) = P_{[a,b]}(K(v))$  for every  $v \in L^2(\Omega)$ , where  $K = \frac{1}{v} \Lambda_A \circ \Lambda_S$  and  $P$  is the projection defined in (3.8). Note that the evaluation of the operator  $K$  requires the solution of one semilinear and one linear elliptic equation.

For the multigrid framework we denote the discrete problems, obtained by discretizing the differential operators and the data by

$$u^h = F^h(u^h). \tag{4.6}$$

Furthermore we consider perturbed problems of the form

$$u^h = F^h(u^h) + d^h,$$

with sufficiently small perturbations  $d^h$ . The recursive multigrid method is described in Algorithm 3.

---

**Algorithm 3** (MG2: Recursive multigrid algorithm of second kind)

---

**Data:** perturbation  $d^h$ , initial value  $u^h$ .

- 1: **if**  $h = h_{\max}$  **then**
  - 2:   solve (4.6) exactly.
  - 3: **else**
  - 4:    $u^h \leftarrow F^h(u^h) + d^h$  (smoothing).
  - 5:    $d^H := I_h^H(u^h - F^h(u^h) - d^h)$ .
  - 6:   Apply MG2( $d^H, \tilde{u}^H$ ) on the next coarser mesh twice to obtain  $u^H$ . Here  $\tilde{u}^H$  is the solution of (4.6) on the next coarser mesh.
  - 7:    $u^h := u^h - I_H^h(u^H - \tilde{u}^H)$  (coarse grid correction).
  - 8: **end if**
-

Note that in step 5 already the newly assigned  $u^h$  of step 4 is used. Further note that for the application of Algorithm 3 to (4.6) on a given mesh the exact solution of the problem is required on every coarser mesh ( $\tilde{u}^H$  is required in step 7). These values are obtained by using a coarse-to-fine method in an outer loop (see the “for-loop” in Algorithm 4): The optimality system (4.6) is solved on the coarsest given mesh of mesh size  $h_{\max}$  using a non-smooth solver, e.g., a semismooth Newton method, [23]. Using this information, the problem is then solved on the next finer grid using the multigrid method. The exact solutions on the respective next coarser meshes are hence obtained utilizing a nested iteration initialization technique, similar to the full multigrid idea [20]. Solving problem (4.6) on the finest mesh of mesh size  $h_{\min}$  therefore requires several applications of Algorithm 3 on the various meshes, resulting in a relatively large computational cost.

For fixed  $(\alpha, \varepsilon)$  only the non-smooth solver on the coarsest mesh requires a good initialization, as the multigrid algorithms on the various meshes are always initialized by  $\tilde{u}^H$  on the respective next coarser meshes. The continuation method via the penalty and the smoothing parameter (Algorithm 1) described in the previous section is therefore only necessary for the non-smooth solver on the coarsest grid of mesh size  $h_{\max} = h_0$  (see the “while-loop” in Algorithm 4). The complete method is given in Algorithm 4.

---

**Algorithm 4** (Outer loop for a multigrid algorithm of second kind)

---

**Data:**  $(\alpha_1, \varepsilon_1, \alpha_{\min}) > 0$ , initial guess  $u^0$  on mesh  $h_0$ , sequence of grids  $\{h_j\}_{j=0}^M$ . Set  $k = 1$ .

- 1: **while**  $\alpha_k > \alpha_{\min}$  **do**
  - 2:   Compute solution  $u^k$  of (4.6) with  $\alpha = \alpha_k, \varepsilon = \varepsilon_k$  and initial guess  $u^{k-1}$  on grid  $h_0$  using semismooth Newton.
  - 3:   Choose  $\alpha_{k+1}, \varepsilon_{k+1}$  and set  $k := k + 1$ .
  - 4: **end while**
  - 5:   Set  $h := h_0, \tilde{u}^h := u^{k-1}, \alpha := \alpha_{k-1}, \varepsilon := \varepsilon_{k-1}$ .
  - 6: **for**  $j = 1$  to  $M$  **do**
  - 7:   Set  $H := h, h := h_j$ .
  - 8:   Compute  $\tilde{u}^h$ , the result of  $\text{MG2}(0, I_H^h \tilde{u}^H)$ .
  - 9: **end for**
- 

In [18] Hackbusch gives conditions for the convergence of the nonlinear multigrid-method of second kind. For suitable choices of the restriction and the prolongation operators it is sufficient to show that  $F$  (and the discrete operators  $F^h$ ) satisfies a Lipschitz condition of the form

$$F(v) - F(w) = L(v, w)(v - w), \quad \forall v, w \in L^2(\Omega), \tag{4.7}$$

where  $L$  is a bounded linear operator, i.e., there exists  $C_1 > 0$  such that

$$\|L(v, w)\|_{L^2, L^2} \leq C_1, \quad \forall v, w \in L^2(\Omega). \tag{4.8}$$

Furthermore we have to show that  $L(\cdot, \cdot)$  is continuous at the solution  $u^*$  of (4.5), i.e.,

$$\|L(v, w) - L(u^*, u^*)\|_{L^2, L^2} \rightarrow 0, \quad \text{if } v, w \rightarrow u^* \tag{4.9}$$

and that  $I - L(u^*, u^*)$  is invertible, i.e., there exists a constant  $C_2 > 0$ , such that

$$\|(I - L(u^*, u^*))^{-1}\|_{L^2, L^2} \leq C_2. \tag{4.10}$$

If these conditions hold, then the algorithm converges for sufficiently small values of  $h_{\max}$ . We observe that such a property of  $L(\cdot, \cdot)$  can only hold true in non-degenerate situations, i.e., when the set where  $F$  is not differentiable has measure zero.

Motivated by [17] we construct the operator  $L$  using a mean-value representation for which we require the differentiability of the control-to-adjoint mapping  $K$ . The operator  $\Lambda_S$ , mapping the control to the state variable, is Fréchet differentiable from  $L^2(\Omega)$  to  $Y$ , see, e.g., [34]. The adjoint equation, however, contains the term  $\max'_\varepsilon(0, -y)$ . Hence  $\Lambda_A$  need not be differentiable if the regularization  $\max_\varepsilon$  is only  $C^1$ . We will therefore restrict ourselves to the  $C^2$ -regularization for this multi-grid method and show that in this case the state-to-adjoint mapping is differentiable from  $Y$  to  $V$ .

**Lemma 4.1** *Let  $\Lambda_A : Y \rightarrow V$  denote the state-to-adjoint operator defined by (4.4) with  $\max_\varepsilon = \max^G_\varepsilon$ . Then  $\Lambda_A$  is Fréchet differentiable for every  $y \in Y$ .*

The proof of Lemma 4.1 can be found in the appendix.

Using the chain rule we therefore find that the composite mapping  $K = \frac{1}{\nu} \Lambda_A \circ \Lambda_S$  is Fréchet differentiable from  $L^2(\Omega)$  to  $V$ . Setting  $B(v, w) = \int_0^1 K'(w + t(v - w)) dt$  for every  $v, w \in L^2(\Omega)$ , Taylor’s theorem yields

$$K(v) - K(w) = B(v, w)(v - w). \tag{4.11}$$

We can now define the Lipschitz operator  $L : L^2(\Omega) \times L^2(\Omega) \rightarrow \mathcal{L}(L^2(\Omega), L^2(\Omega))$  by setting  $L(v, w) := \theta(v, w)B(v, w)$  with  $\theta : L^2(\Omega) \times L^2(\Omega) \rightarrow L^\infty(\Omega)$  defined by

$$\theta(v, w) := \begin{cases} 0 & \text{in } \{K(v) < a\} \cap \{K(w) < a\}, \\ 1 & \text{in } \{K(v) \in [a, b]\} \cap \{K(w) \in [a, b]\}, \\ 0 & \text{in } \{K(v) > b\} \cap \{K(w) > b\}, \\ (K(v) - a)/(B(v, w)(v - w)) & \text{in } \{K(v) \in [a, b]\} \cap \{K(w) < a\}, \\ (a - K(w))/(B(v, w)(v - w)) & \text{in } \{K(v) < a\} \cap \{K(w) \in [a, b]\}, \\ (K(v) - b)/(B(v, w)(v - w)) & \text{in } \{K(v) \in [a, b]\} \cap \{K(w) > b\}, \\ (b - K(w))/(B(v, w)(v - w)) & \text{in } \{K(v) > b\} \cap \{K(w) \in [a, b]\}, \\ (b - a)/(B(v, w)(v - w)) & \text{in } \{K(v) > b\} \cap \{K(w) < a\}, \\ (a - b)/(B(v, w)(v - w)) & \text{in } \{K(v) < a\} \cap \{K(w) > b\}. \end{cases} \tag{4.12}$$

With this choice

$$F(v) - F(w) = P_{[a,b]}(K(v)) - P_{[a,b]}(K(w)) = L(v, w)(v - w) \quad \forall v, w, \in L^2(\Omega)$$

holds true. We can now formulate the following conditions for the convergence of Algorithm 3.

**Lemma 4.2** *Let  $u^*$  be the solution of the fixed-point problem (4.5) with corresponding state and adjoint state  $(y^*, p^*)$  and assume that*

$$\text{meas}\{\mathbf{x} \in \Omega : |K(u^*) - a| \leq \tau\} \rightarrow 0 \quad \text{as } \tau \rightarrow 0, \tag{4.13a}$$

$$\text{meas}\{\mathbf{x} \in \Omega : |K(u^*) - b| \leq \tau\} \rightarrow 0 \quad \text{as } \tau \rightarrow 0. \tag{4.13b}$$

Further, we assume that (4.13) also holds for the discrete problems. If, moreover,

$$1 - \alpha^{-1} \max_{\varepsilon}''(0, -y^*)p^* \geq 0 \quad \text{a.e. in } \Omega, \tag{4.14}$$

then the method described in Algorithm 3 converges for sufficiently small mesh sizes  $h_{\max}$ .

*Proof* Using Corollary 5.3 of [18], it suffices to show that (4.7)–(4.9) are satisfied for  $F$  and its discretization  $F^h$ . We give the proof for the continuous operator, analogous arguments can be used for the discrete operator.

Equation (4.7) holds true due to the definition of the operator  $L$ . From (4.12) it follows that the operator  $\theta$  takes values only in  $[0, 1]$ . We show this exemplarily for the case where  $K(v)(\mathbf{x}) < a(\mathbf{x})$  and  $K(w)(\mathbf{x}) \in [a(\mathbf{x}), b(\mathbf{x})]$ . Here we estimate

$$\begin{aligned} 0 &\geq a(\mathbf{x}) - K(w)(\mathbf{x}) > a(\mathbf{x}) - K(w)(\mathbf{x}) + (K(v)(\mathbf{x}) - a(\mathbf{x})) \\ &= K(v)(\mathbf{x}) - K(w)(\mathbf{x}) = B(v, w)(v - w)(\mathbf{x}). \end{aligned}$$

Therefore

$$0 \leq (a(\mathbf{x}) - K(w)(\mathbf{x})) / (B(v, w)(v - w)(\mathbf{x})) < 1.$$

For the other cases the proof is similar. As  $K$  is Fréchet differentiable the differential  $K'$  is bounded in a neighborhood of  $u^*$  and hence  $B$  is bounded in a neighborhood of  $(u^*, u^*)$ . As a consequence  $L$  is also bounded in that neighborhood. Therefore it remains to show that (4.9) is satisfied.

For given  $v, w$  we decompose the domain  $\Omega$  into sets according to the feasibility of  $K(v)$  and  $K(w)$ .

$$\begin{aligned} S_1 &:= \{K(u^*) < a, K(v) < a, K(w) < a\}, \\ S_2 &:= \{K(u^*) \in [a, b], K(v) \in [a, b], K(w) \in [a, b]\}, \\ S_3 &:= \{K(u^*) > b, K(v) > b, K(w) > b\}, \\ S_4 &:= \Omega \setminus (S_1 \cup S_2 \cup S_3). \end{aligned}$$

Due to the continuity of  $B(\cdot, \cdot)$  the operator  $L(\cdot, \cdot)$  is continuous on the sets  $S_1, S_2, S_3$ . We therefore only have to show continuity on the set  $S_4$  where  $K(u^*)$  is projected onto a different set than  $K(v)$  or  $K(w)$ . We start by considering the set  $M \subset S_4$  given by

$$M := \{K(u^*) < a \wedge K(v) \in [a, b]\}.$$

Using the continuity of  $B(v, w)$  from  $L^2(\Omega)$  to  $V$ , the fact that  $\|\theta(v, w)\|_{L^\infty} \leq 1$  for all  $v, w \in L^2(\Omega)$  and the information that  $\theta(u^*, u^*)$  vanishes on  $M$ , we can estimate

$$\begin{aligned} \|(L(v, w) - L(u^*, u^*))z\|_{L^2(M)} &= \|\theta(v, w)B(v, w)z\|_{L^2(M)} \\ &\leq \text{meas}(M)^{(q-2)/2q} \|B(v, w)z\|_{L^q} \\ &\leq C \cdot \text{meas}(M)^{(q-2)/2q} \|z\|_{L^2}, \end{aligned} \tag{4.15}$$

for all  $z \in L^2(\Omega)$ , with  $q > 2$  such that  $V \hookrightarrow L^q(\Omega)$ . The expression on the left hand side then goes to zero if the measure of the set  $M$  tends to zero. In order to show that this is the case, we introduce the two sets

$$\begin{aligned} M_1(\tau) &:= \{K(u^*) - a \leq -\tau \wedge K(v) \in [a, b]\}, \\ M_2(\tau) &:= \{|K(u^*) - a| \leq \tau\} \end{aligned}$$

such that  $M \subset M_1(\tau) \cup M_2(\tau)$  for all  $\tau > 0$ . The measure of  $M_1(\tau)$  can be estimated by

$$\begin{aligned} \tau \cdot \text{meas}(M_1(\tau))^{1/2} &\leq \|K(u^*) - K(v)\|_{L^2(M_1(\tau))} = \|B(u^*, v)(u^* - v)\|_{L^2(M_1(\tau))} \\ &\leq C \|u^* - v\|_{L^2(M_1(\tau))}. \end{aligned}$$

Therefore

$$\text{meas}(M) \leq \left( C \frac{1}{\tau} \|u^* - v\|_{L^2} \right)^2 + \text{meas}(M_2(\tau)) \quad \forall \tau > 0.$$

Choosing, e.g.,  $\tau := \|u^* - v\|_{L^2}^\rho$  with  $0 < \rho < 1$  we find that  $\text{meas}(M) \rightarrow 0$  if  $v \rightarrow u^*$  due to (4.13a). Hence  $L(v, w) \rightarrow L(u^*, u^*)$  in  $\mathcal{L}(L^2(\Omega), L^2(M))$ . Similar arguments can also be used for the sets  $\{K(u^*) < a \wedge K(w) \in [a, b]\}$ ,  $\{K(u^*) > b \wedge K(v) \in [a, b]\}$  and  $\{K(u^*) > b \wedge K(w) \in [a, b]\}$ .

Next we show that the operator  $L(\cdot, \cdot)$  is continuous on the sets  $\{K(u^*) \in [a, b] \wedge K(v) < a\}$ ,  $\{K(u^*) \in [a, b] \wedge K(w) < a\}$ ,  $\{K(u^*) \in [a, b] \wedge K(v) > b\}$  and  $\{K(u^*) \in [a, b] \wedge K(w) > b\}$ . We do this exemplarily for the set

$$N := \{K(u^*) \in [a, b] \wedge K(v) < a\}.$$

Let  $z \in L^2(\Omega)$  be given. Again using the fact that  $\|\theta(v, w)\|_{L^\infty} \leq 1$  for all  $v, w \in L^2(\Omega)$ , together with the information that  $\theta(u^*, u^*) \equiv 1$  on  $N$ , we estimate

$$\begin{aligned} \|(L(v, w) - L(u^*, u^*))z\|_{L^2(N)} &= \|(\theta(v, w)B(v, w) - B(u^*, u^*))z\|_{L^2(N)} \\ &\leq \|\theta(v, w)(B(v, w) - B(u^*, u^*))z\|_{L^2(N)} + \|(\theta(v, w) - 1)B(u^*, u^*)z\|_{L^2(N)} \\ &\leq \|(B(v, w) - B(u^*, u^*))z\|_{L^2(N)} + \|B(u^*, u^*)z\|_{L^2(N)} \\ &\leq 3C \cdot \text{meas}(N)^{(q-2)/2q} \|z\|_{L^2}, \end{aligned}$$



with the exponent  $q > 2$  chosen as in (4.15). For each sufficiently small  $\tau > 0$  we introduce the sets

$$N_1(\tau) := \{K(u^*) \in [a + \tau, b] \wedge K(v) < a\},$$

$$N_2(\tau) := \{|K(u^*) - a| \leq \tau\}.$$

Using the same arguments as for the set  $M$  it follows that  $meas(N) \rightarrow 0$  as  $v \rightarrow u^*$ . Hence  $L(v, w) \rightarrow L(u^*, u^*)$  also in  $\mathcal{L}(L^2(\Omega), L^2(N))$  and (4.9) is satisfied.

It remains to show that condition (4.10) holds true. For this purpose, note that  $L(u^*, u^*)$  is given by

$$L(u^*, u^*) = \chi_{\mathcal{I}^*}(v^{-1}(\mathcal{B}(u^*)^{-1})^*(\alpha^{-1} \max''_{\varepsilon}(0, -y^*)p^* - 1)\mathcal{B}(u^*)^{-1}),$$

where  $\chi_{\mathcal{I}^*}$  denotes the characteristic function of the inactive set at  $u^*$  given by  $\mathcal{I}^* := \{a < K(u^*) < b\}$  and  $(\mathcal{B}(u)^{-1})^*$  represents the solution operator of the adjoint equation (4.4) at  $y = y(u)$ . Below, by  $\mathcal{C}^*$  we denote the complement of  $\mathcal{I}^*$  in  $\Omega$ , the so-called coincidence or active set. For the invertibility of  $I - L(u^*, u^*)$  we study the system

$$(I - L(u^*, u^*))\delta u = g \tag{4.16}$$

for arbitrary and fixed  $g \in L^2(\Omega)$ . Let  $E_{\mathcal{C}^*}$  and  $E_{\mathcal{I}^*}$  denote the extension-by-zero operators from  $\mathcal{C}^*$  respectively  $\mathcal{I}^*$  to  $\Omega$ . Their adjoint operators denoted by  $E_{\mathcal{C}^*}^*$  and  $E_{\mathcal{I}^*}^*$  are restrictions to the respective set. Due to the definition of  $\mathcal{I}^*$  and  $\mathcal{C}^*$  we have  $E_{\mathcal{I}^*}^*E_{\mathcal{C}^*} = 0$  and  $E_{\mathcal{I}^*}^*E_{\mathcal{I}^*} = I_{\mathcal{I}^*}$  with the latter being the identity operator on  $\mathcal{I}^*$ . From the structure of (4.16) we get

$$\delta u|_{\mathcal{C}^*} = g|_{\mathcal{C}^*}$$

and further

$$(I_{\mathcal{I}^*} + v^{-1}E_{\mathcal{I}^*}^*(\mathcal{B}(u^*)^{-1})^*(1 - \alpha^{-1} \max''_{\varepsilon}(0, -y^*)p^*)\mathcal{B}(u^*)^{-1})E_{\mathcal{I}^*}\delta u|_{\mathcal{I}^*} = \hat{g}^* \tag{4.17}$$

with

$$\hat{g}^* := g|_{\mathcal{I}^*} + v^{-1}E_{\mathcal{I}^*}^*(\mathcal{B}(u^*)^{-1})^*(\alpha^{-1} \max''_{\varepsilon}(0, -y^*)p^* - 1)\mathcal{B}(u^*)^{-1}E_{\mathcal{C}^*}(g|_{\mathcal{C}^*}).$$

Above,  $g|_{\mathcal{S}}$  is a short hand notation for the restriction of  $g$  to the set  $\mathcal{S} \subset \Omega$ ; similarly for  $\delta u|_{\mathcal{S}}$ . Under the assumptions of this theorem we have that (4.17) is invertible and  $\delta u = E_{\mathcal{C}^*}(g|_{\mathcal{C}^*}) + E_{\mathcal{I}^*}(\delta u|_{\mathcal{I}^*}) \in L^2(\Omega)$  solves (4.16), and the bound for  $\|(I - L(u^*, u^*))^{-1}\|_{L^2, L^2}$  follows readily. Since  $\max''_{\varepsilon}(0, -y^*)p^* \in L^\infty(\Omega)$ , for consistent discretizations the discrete analogue of (4.17) is uniformly invertible for all sufficiently small mesh sizes. This proves the assertion.  $\square$

We mention that the above invertibility result concerning  $I - L(u^*, u^*)$  is tightly related to second order sufficient optimality conditions under a non-degeneracy assumption [34]. In fact, (4.14) is sufficient for second order sufficiency.

## 5 Numerical tests

### 5.1 Examples

For the following test examples the variables are defined on the two dimensional open unit square  $\Omega = (0, 1)^2$ . As a prototype of the elliptic differential operator the negative Laplace-operator, discretized using a standard five-point stencil, is chosen. For all test runs the iterates are initialized by  $y^0 = u^0 = p^0 \equiv 0$ . The penalty parameter is initialized by  $\alpha_0 = 1$  and updated by  $\alpha_{k+1} = c_\alpha \alpha_k$  with  $c_\alpha = 0.25$ . For runs where the smoothing parameter remains constant it is chosen as  $\varepsilon = \varepsilon_0 = 10^{-3}$ . If  $\varepsilon$  is coupled with  $\alpha$ , it is initialized by  $\varepsilon_0 = 1$  and updated by  $\varepsilon_{k+1} = (\alpha_{k+1})^{1+\delta}$  with  $\delta > 0$ .

*Example 5.1* (Lack of strict complementarity) We construct a test problem for which the active set at the solution contains a subset where strict complementarity fails to hold, i.e., where the biactive set  $B := \{y^* = 0\} \cap \{\xi^* = 0\}$  has a positive measure. This situation is challenging, as the active constraint gradients at the solution are linearly dependent. We define

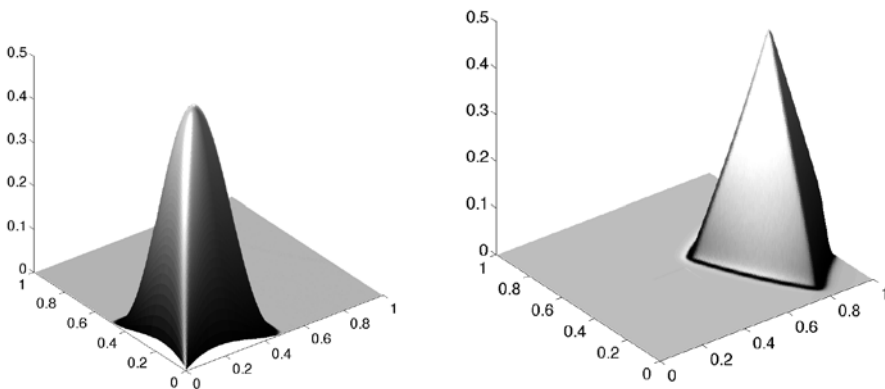
$$y^\dagger(\mathbf{x}_1, \mathbf{x}_2) = \begin{cases} 160(\mathbf{x}_1^3 - \mathbf{x}_1^2 + 0.25\mathbf{x}_1)(\mathbf{x}_2^3 - \mathbf{x}_2^2 + 0.25\mathbf{x}_2) & \text{in } (0, 0.5) \times (0, 0.5), \\ 0 & \text{else,} \end{cases}$$

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

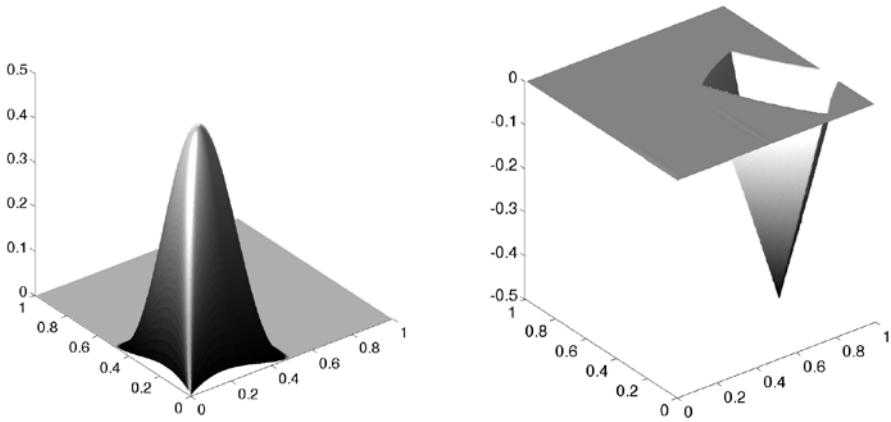
and set

$$\begin{aligned} f &= -\Delta y^\dagger - y^\dagger - \xi^\dagger, \\ y_d &= y^\dagger + \xi^\dagger - \nu \Delta y^\dagger, \end{aligned}$$

with  $\nu = 1$ . We consider no control constraints for this test problem, i.e.,  $a \equiv -\infty$  and  $b \equiv \infty$ . The numerical solutions are displayed in Figs. 2–4. Recall that for the unconstrained case the optimal control  $u^*$  is just a multiple of the adjoint state  $p^*$ , and hence not specifically depicted.

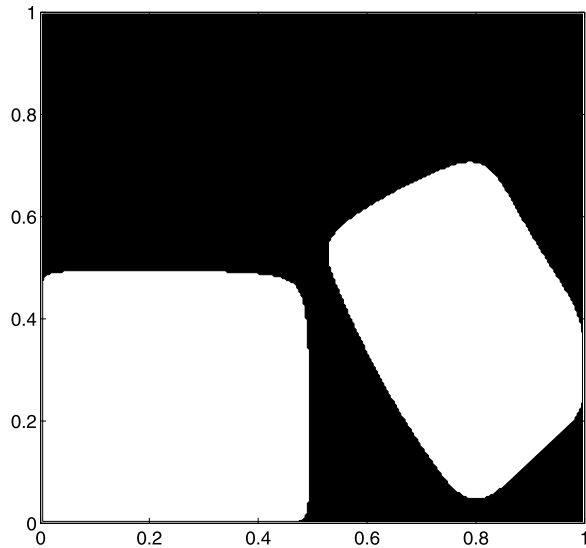


**Fig. 2** Optimal state  $y^*$  (left) and slack variable  $\xi^*$  (right) for Example 5.1



**Fig. 3** Adjoint state  $p^*$  (left) and multiplier  $\lambda^*$  (right) for Example 5.1

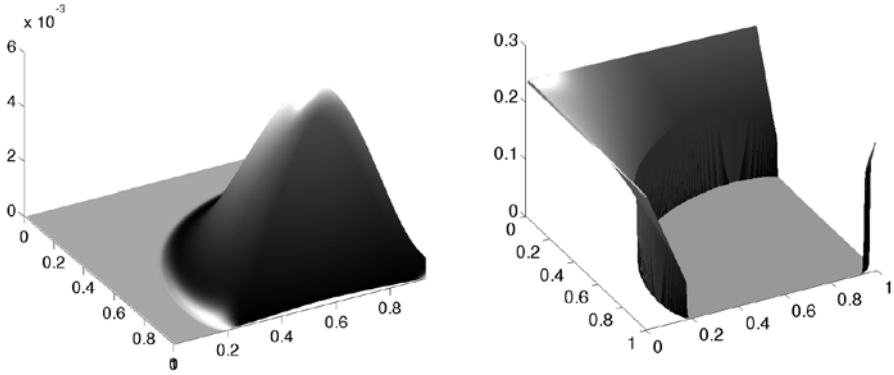
**Fig. 4** Biactive set (black) for Example 5.1



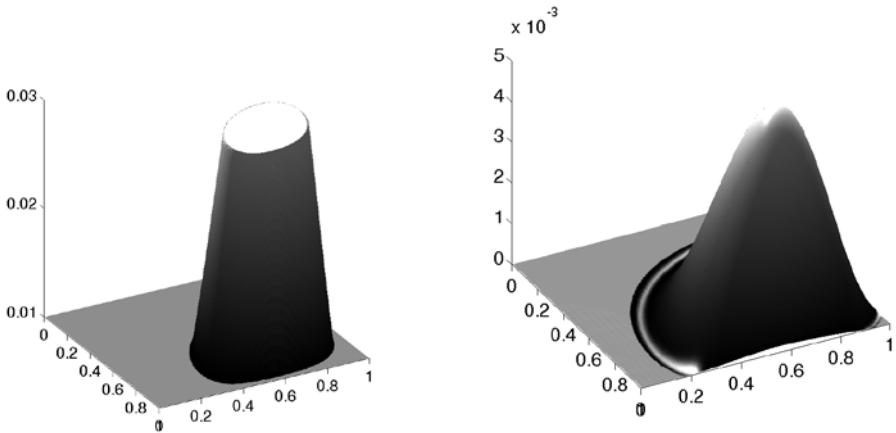
*Example 5.2 (Degenerate solution)* For this example the optimal state  $y^*$  exhibits a very flat transition into the active set (note the scale in Fig. 5). This makes the active set detection challenging. Purely primal active set techniques usually perform poorly in such situations. We set  $\nu = 0.01$ ,  $a = 0.01$  and  $b = 0.03$ . The example is defined by the data

$$f(x_1, x_2) = y_d(x_1, x_2) = -|x_1 x_2 - 0.5| + 0.25.$$

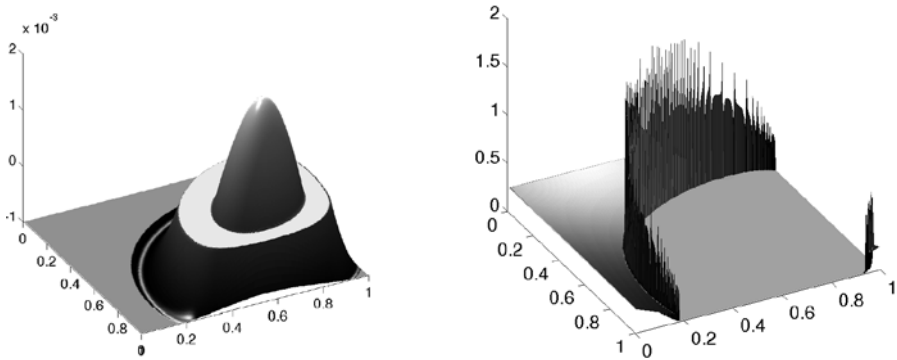
The numerical solution is shown in Figs. 5–7. The multiplier  $\lambda^*$  shows clear signs of low regularity. In order to further illustrate the influence of the pointwise constraints



**Fig. 5** Optimal state  $y^*$  (left) and slack variable  $\xi^*$  (right) for Example 5.2



**Fig. 6** Optimal control  $u^*$  (left) and adjoint state  $p^*$  (right) for Example 5.2



**Fig. 7** Multipliers  $\phi^*$  (left) and  $\lambda^*$  (right) for Example 5.2

on the control we introduce the quantity

$$\phi^* := \max(0, p^* - vb) - \max(0, -p^* + va),$$

which can be interpreted as the Lagrange multiplier corresponding to the constraint (1.1b). The set where  $\phi^*$  is positive corresponds to the set where the upper bound on the control  $u$  is active. Analogously the set where  $\phi^*$  is negative corresponds to the set where the lower bound is active. The zero-level-set of  $\phi^*$  depicts the set where the control constraints are inactive.

### 5.2 Numerical performance

Due to Theorem 3.4 the iterates of Algorithm 1 converge to an  $\mathcal{E}$ -almost C-stationary point of the MPEC. For all our test runs it could be verified by inspection of the multipliers on the biactive set, that the numerical solutions that were obtained even satisfied conditions (3.4) and (3.5) characterizing strong stationarity as  $p^*$  and  $\lambda^*$  (numerically) vanish on the biactive set.

In Table 1 the iteration numbers on the different grids for the FAS-multigrid method using our local and global  $C^1$ -regularizations are displayed. The coupling of  $\alpha$  and  $h$  described in Sect. 4.1.1 was used with a coarsest mesh of 15 inner grid points and a finest mesh of 255 inner grid points per dimension, respectively. Furthermore a grid dependent stopping tolerance ( $tol = 0.5h^2$ ) was employed. The number of iterations are clearly concentrated on the coarse grids, which illustrates the efficiency of the nested iteration.

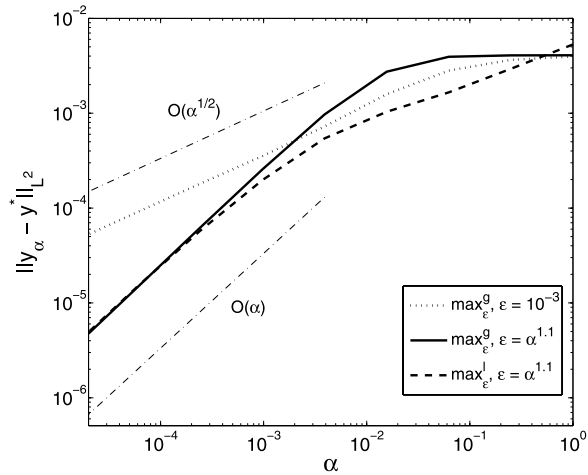
For the multigrid method of second kind discussed in Sect. 4.1.2, the test problems were solved using the  $C^1$ - and the  $C^2$ - $\max_\epsilon$ -regularizations. Although the convergence analysis requires smoothness of the operator the method converged to the same solution for all regularization schemes. On each grid the required accuracy was reached, and even exceeded, after one iteration of the W-cycle. This suggests some kind of “over-solving” due to the requirement of  $\tilde{u}^H$ , the exact solution of (4.6) on the respective coarser mesh.

Table 1 also shows the ratio between the CPU-time of the FAS-method and the multigrid method of second kind (MG2). Every Picard-iteration of MG2 requires the solution of one semilinear and one linear elliptic equation on the fine mesh, the

**Table 1** Numerical performance of the FAS-method with coupling of  $\alpha$  and  $h$

#inner gridpts.	Example 5.1		Example 5.2	
	$\max_\epsilon^g$	$\max_\epsilon^l$	$\max_\epsilon^g$	$\max_\epsilon^l$
15	6	9	5	7
31	3	3	3	2
63	3	3	3	3
127	3	3	5	3
255	2	2	2	2
total	17	20	18	17
CPU ratio FAS/MG2	0.4311	0.3738	0.2324	0.2713

**Fig. 8**  $L^2$ -errors  $\|y_\alpha - y^*\|$  for Example 5.1



smoothing step is therefore quite costly. This is reflected in the small CPU-ratio. The computational effort of MG2, however, can be reduced by employing an inexact version of the multigrid method of second kind, as given in [19].

### 5.3 Approximation properties of the outer algorithm

In order to numerically determine the approximation properties of the continuation method (Algorithm 1), the  $L^2$ -errors  $\|y_\alpha - y^*\|$  were computed for Example 5.1, where  $y_\alpha$  is a stationary point of  $(\mathcal{P}_{\alpha,\varepsilon})$  and  $y^*$  is a (strongly) stationary point of  $(\mathcal{P})$ . The problems were solved on a fixed mesh with  $h = 1/256$  for the global regularization  $\max_\varepsilon^g$ , once with a fixed smoothing parameter  $\varepsilon = 10^{-3}$  and once with  $\varepsilon = \varepsilon(\alpha) = \alpha^{1.1}$ , thereby yielding  $\varepsilon(\alpha)/\alpha \rightarrow 0$ . Furthermore the local regularization  $\max_\varepsilon^l$  of the  $\max(0, \cdot)$ -operator was considered with  $\varepsilon = \varepsilon(\alpha) = \alpha^{1.1}$ . The results are given in Fig. 8 in a log-log-scale. The graphic suggests an approximation order of  $\|y_\alpha - y^*\| = \mathcal{O}(\alpha)$  if  $\varepsilon(\alpha)/\alpha \rightarrow 0$  and  $\|y_\alpha - y^*\| = \mathcal{O}(\alpha^{1/2})$  for fixed  $\varepsilon$  and regularization  $\max_\varepsilon^g$ . A similar behavior could be observed in a number of further numerical tests. In general the approximation properties for  $\varepsilon(\alpha)/\alpha \rightarrow 0$  showed no significant difference for the two types of regularizations considered. Based on these observations we motivate the stopping criterion introduced in Sect. 4.1.1. If we assume a discretization error of order  $h^2$ , i.e.,

$$\|y_h - y^*\|_{L^2} = \mathcal{O}(h^2),$$

where  $y_h$  is the exact solution of the discrete MPEC, as is the case for linear finite elements (see, e.g, [11]), this leads to an estimate of the form

$$\|y_{\alpha,h} - y^*\|_{L^2} = \mathcal{O}(\alpha^s + h^2).$$

Here  $y_{\alpha,h}$  is the solution of the discrete version of the smoothed penalized problem  $(\mathcal{P}_{\alpha,\varepsilon})$  and  $s \in (0.5, 1)$  is the respective approximation order. If the smoothed penalized problem is solved on a fixed grid with mesh size  $h$ , the discretization error dominates the approximation error if  $\alpha \leq h^{2/s}$ .

**Table 2** Number of iterations on the different grids for the FAS-method without coupling of  $\alpha$  and  $h$

#inner gridpts.	Example 5.1		Example 5.2	
	$\max \epsilon^g$	$\max \epsilon^l$	$\max \epsilon^g$	$\max \epsilon^l$
15	20	25	23	26
31	25	29	25	27
63	25	30	26	30
127	28	31	29	29
255	28	32	27	29
511	28	32	27	30

Decreasing the penalty parameter further then leads to no improvement of the numerical solution. On a fixed grid the smoothed penalized problem is therefore solved as long as

$$\alpha \geq Ch^{2/s},$$

where  $C > 0$  is an appropriately chosen constant. This justifies (4.3). If  $\alpha$  drops below that threshold the grid is refined, or if already on the finest mesh, the algorithm is terminated.

### 5.4 Stability under mesh refinement

As Theorem 3.4 proves convergence of the algorithm in a function space setting one expects that the discrete method exhibits some numerical stability under mesh refinements. For a fixed sequence  $\{\alpha_k\}$  (i.e.,  $\alpha_k$  independent of  $h$ ) and a fixed tolerance we therefore solved Examples 5.1 and 5.2 on different meshes, but this time neither coupling  $\alpha$  with  $h$ , nor employing a grid dependent stopping tolerance. For the sake of a better comparison a cycle index of  $\gamma = 1$  (=V-cycle) was used. Table 2 shows the total number of iterations for the FAS-multigrid method with the local and global  $C^1$ -regularizations. The results imply mesh-independent behavior of the method.

### 5.5 Convergence factors for the FAS algorithm

In order to report the numerical behavior of the FAS-scheme, the optimality system (4.1) of a smoothed penalized problem with fixed parameters  $\alpha = \epsilon = 1e-3$  was solved using the FAS-method and the global  $C^1$ -regularization of the max-operator. Both test problems were once solved using a cycle index  $\gamma = 1$  (=V-cycle) and once with  $\gamma = 2$ , resulting in a W-cycle. Furthermore, different values for the cost parameter  $\nu$  in the objective function were used.

For both test problems the control unconstrained case as well as the case with point-wise bilateral control constraints were considered. For Example 5.1 the constraints  $a \equiv 0.1, b \equiv 0.4$  were enforced. The solution of the MPEC with these constraints still showed a biactive set of positive measure. Note that in view of Theorem 3.4 the control constrained case immediately yields convergence of the stationary points of the smoothed penalized problems, whereas the seemingly simpler control-unconstrained case requires the additional assumption on the boundedness of  $\{u_\alpha\}$  in order to guarantee convergence.

**Table 3** Estimated FAS convergence factors for Example 5.1

$\nu$	Constr.		Unconstr.	
	V-cyc.	W-cyc.	V-cyc.	W-cyc.
1	0.0433	0.0219	0.1217	0.0220
1e−1	0.0370	0.0238	0.1615	0.0251
1e−2	0.2777	0.0234	0.7792	0.0724

**Table 4** Estimated FAS convergence factors for Example 5.2

$\nu$	Constr.		Unconstr.	
	V-cyc.	W-cyc.	V-cyc.	W-cyc.
1	0.0387	0.0219	0.0389	0.0218
1e−1	0.0396	0.0217	0.0377	0.0218
1e−2	0.1312	0.0220	0.2425	0.0228

Tables 3 and 4 show the estimated convergence factors for the state variable  $y$ , which are defined as the asymptotic values of the ratios between the discrete  $H^{-1}$ -norms of the residuals of the state equations resulting from two successive multigrid cycles on a given mesh [7, 9].

Tables 3 and 4 suggest that the convergence factor depends on the cost parameter for the control  $\nu$ , i.e., the condition of the problem. If the problem is solved using V-cycles this dependency is quite distinct, whereas the use of W-cycles results in a more stable behavior. Furthermore the presence of control constraints seems to have a stabilizing influence on the performance.

**Acknowledgements** The authors gratefully acknowledge support by the Austrian Federal Ministry of Science and Research (bm.w\_f) and the FWF under START-program Y305 “Interfaces and Free Boundaries”. M.H. further acknowledges support through the DFG Research Center MATHEON under project C28.

### Appendix

The following results are used in this paper. The proof techniques are independent of the MPEC context, nevertheless we provide the proofs for the sake of completeness. For  $v \in L^2(\Omega)$  we use the notation  $v^+ := \max(0, v)$ . The following lemma investigates some properties of the  $\max(0, \cdot)$ -operator.

**Lemma A.1** For  $\{v_\alpha\} \subset V$  and  $\tilde{v} \in V$  the following assertions hold:

- (i) If  $v_\alpha \rightharpoonup \tilde{v}$  in  $V$ , then also  $v_\alpha^+ \rightharpoonup \tilde{v}^+$  in  $V$ .
- (ii) If  $v_\alpha \rightarrow \tilde{v}$  in  $V$ , then also  $v_\alpha^+ \rightarrow \tilde{v}^+$  in  $V$ .

*Proof* Due to the compact embedding of  $V$  into  $L^2(\Omega)$  and the Lipschitz continuity of  $\max(0, \cdot) : L^2(\Omega) \rightarrow L^2(\Omega)$  we find that  $v_\alpha^+$  converges to  $\tilde{v}^+$  strongly in  $L^2(\Omega)$ . Furthermore for every  $\alpha > 0$

$$\|\nabla v_\alpha^+\|_{L^2(\Omega)^n} \leq \|\nabla v_\alpha\|_{L^2(\Omega)^n} \leq C$$



for some  $C > 0$  independent of  $\alpha$ . Hence there exists  $\eta = (\eta_1, \dots, \eta_n) \in L^2(\Omega)^n$  such that on a suitable subsequence  $\frac{\partial v_\alpha^+}{\partial x_i} \rightharpoonup \eta_i$  for all  $1 \leq i \leq n$ . We verify that indeed  $\eta = \nabla \tilde{v}^+$  by multiplying by a test function  $\varphi \in C_0^\infty(\Omega)$ . For every  $1 \leq i \leq n$  we find that

$$(\eta_i, \varphi) = \lim_{\alpha \rightarrow 0} \left( \frac{\partial v_\alpha^+}{\partial x_i}, \varphi \right) = - \lim_{\alpha \rightarrow 0} \left( v_\alpha^+, \frac{\partial \varphi}{\partial x_i} \right) = - \left( \tilde{v}^+, \frac{\partial \varphi}{\partial x_i} \right) = \left( \frac{\partial \tilde{v}^+}{\partial x_i}, \varphi \right).$$

As the weak limit of the subsequence is uniquely determined, we find that in fact  $\nabla v_\alpha^+ \rightharpoonup \nabla \tilde{v}^+$  on the whole sequence. Finally if  $v_\alpha$  converges to  $\tilde{v}$  strongly in  $V$  then  $\nabla v_\alpha$  converges to  $\nabla \tilde{v}$  strongly in  $L^2(\Omega)^n$  and strong convergence of  $\nabla v_\alpha^+$  to  $\nabla \tilde{v}^+$  in  $L^2(\Omega)^n$  follows from

$$\begin{aligned} \|\nabla v_\alpha^+ - \nabla \tilde{v}^+\|_{L^2(\Omega)^n}^2 &= (\nabla v_\alpha^+, \nabla v_\alpha^+) - 2(\nabla v_\alpha^+, \nabla \tilde{v}^+) + (\nabla \tilde{v}^+, \nabla \tilde{v}^+) \\ &= (\nabla v_\alpha^+, \nabla v_\alpha) - 2(\nabla v_\alpha^+, \nabla \tilde{v}^+) + (\nabla \tilde{v}^+, \nabla \tilde{v}^+) \rightarrow 0. \quad \square \end{aligned}$$

We next give the proof of Lemma 4.1

*Proof of Lemma 4.1* Due to the monotonicity of the  $\max_\varepsilon^G$ -operator, the linear operator

$$\mathcal{A}^* + \frac{1}{\alpha} (\max_\varepsilon^G)'(0, -y) : V \rightarrow V^*$$

is bounded and coercive with constants independent of  $y$ . Then from the Lax-Milgram theorem it follows that the operator is invertible and the norm of the inverse is bounded independently of  $y$ . Hence, the operator  $\Lambda_A$  is well defined. The Lipschitz continuity of  $(\max_\varepsilon^G)'(0, \cdot) : Y \rightarrow L^\infty(\Omega)$  furthermore yields the local Lipschitz continuity of  $\Lambda_A$ .

Note that the Nemyzki operator  $y \mapsto \phi(y) := (\max_\varepsilon^G)'(0, -y)$  is Fréchet differentiable from  $Y$  to  $L^\infty(\Omega)$ . Now let  $y, \delta y \in Y$  be given and set  $p_1 = \Lambda_A(y + \delta y)$ ,  $p_2 = \Lambda_A(y)$ . Subtraction of the corresponding adjoint equations (4.4) then yields

$$\begin{aligned} 0 &= \mathcal{A}^*(p_1 - p_2) + \phi(y + \delta y)p_1 - \phi(y)p_2 + \delta y \\ &= (\mathcal{A}^* + \phi(y))(p_1 - p_2) + (\phi(y + \delta y) - \phi(y))p_1 + \delta y \\ &= (\mathcal{A}^* + \phi(y))(p_1 - p_2) + (\phi'(y)\delta y + r_y(y, \delta y))p_1 + \delta y, \end{aligned} \tag{A.1}$$

where  $\|r_y(y, \delta y)\|_{L^\infty} = \mathcal{O}(\|\delta y\|_Y)$ . Setting

$$\begin{aligned} D\Lambda_A(y) &= -(\mathcal{A}^* + \phi(y))^{-1}(\phi'(y)p_2 + I), \\ r_\Lambda(y, \delta y) &= -(\mathcal{A}^* + \phi(y))^{-1}(\phi'(y)\delta y(p_1 - p_2) + r_y(y, \delta y)p_1), \end{aligned}$$

(A.1) can be reformulated as

$$\Lambda_A(y + \delta y) - \Lambda_A(y) = D\Lambda_A(y)\delta y + r_\Lambda(y, \delta y).$$

Due to the boundedness of  $(\mathcal{A}^* + \phi(y))^{-1}$  in  $\mathcal{L}(V^*, V)$  and the boundedness of  $\phi'$  in  $L^\infty(\Omega)$  we find that  $D\Lambda_A(y)$  is linear and bounded. The Lipschitz continuity of  $\Lambda_A$  furthermore yields that  $\|r_\Lambda(y, \delta y)\|_V = o(\|\delta y\|_Y)$ , hence  $\Lambda_A$  is Fréchet differentiable.  $\square$

## References

1. Barbu, V.: Optimal Control of Variational Inequalities. Research Notes in Mathematics, vol. 100. Pitman (Advanced Publishing Program), Boston (1984)
2. Bergounioux, M.: Optimal control of an obstacle problem. *Appl. Math. Optim.* **36**(2), 147–172 (1997)
3. Bergounioux, M.: Use of augmented Lagrangian methods for the optimal control of obstacle problems. *J. Optim. Theory Appl.* **95**(1), 101–126 (1997)
4. Bergounioux, M.: Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim.* **36**(1), 273–289 (1998) (electronic)
5. Bergounioux, M., Dietrich, H.: Optimal control of problems governed by obstacle type variational inequalities: a dual regularization-penalization approach. *J. Convex Anal.* **5**(2), 329–351 (1998)
6. Bergounioux, M., Mignot, F.: Optimal control of obstacle problems: existence of Lagrange multipliers. *ESAIM Control Optim. Calc. Var.* **5**, 45–70 (2000) (electronic)
7. Borzi, A.: Multigrid methods for parabolic distributed optimal control problems. *J. Comput. Appl. Math.* **157**(2), 365–382 (2003)
8. Borzi, A.: High-order discretization and multigrid solution of elliptic nonlinear constrained optimal control problems. *J. Comput. Appl. Math.* **200**(1), 67–85 (2007)
9. Borzi, A., Kunisch, K.: A multigrid scheme for elliptic constrained optimal control problems. *Comput. Optim. Appl.* **31**(3), 309–333 (2005)
10. Brandt, A.: Multi-level adaptive solutions to boundary-value problems. *Math. Comput.* **31**(138), 333–390 (1977)
11. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, vol. 15. Springer, New York (1994)
12. Briggs, W.L., Henson, V.E., McCormick, S.F.: A Multigrid Tutorial, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2000)
13. Friedman, A.: Variational Principles and Free-Boundary Problems. Pure and Applied Mathematics. Wiley, New York (1982). A Wiley-Interscience Publication
14. Glowinski, R., Lions, J.-L., Trémolières, R.: Numerical Analysis of Variational Inequalities. Studies in Mathematics and its Applications, vol. 8. North-Holland, Amsterdam (1981). Translated from the French
15. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Monographs and Studies in Mathematics, vol. 24. Pitman (Advanced Publishing Program), Boston (1985)
16. Hackbusch, W.: On the fast solutions of nonlinear elliptic equations. *Numer. Math.* **32**(1), 83–95 (1979)
17. Hackbusch, W.: On the fast solving of parabolic boundary control problems. *SIAM J. Control Optim.* **17**(2), 231–244 (1979)
18. Hackbusch, W.: Die schnelle Auflösung der Fredholmschen Integral gleichung zweiter Art. *Beitr. Numer. Math.* **9**(1), 47–62 (1981)
19. Hackbusch, W.: Error analysis of the nonlinear multigrid method of the second kind. *Apl. Mat.* **26**(1), 18–29 (1981). With a loose Russian summary
20. Hackbusch, W.: Multigrid Methods and Applications. Springer Series in Computational Mathematics, vol. 4. Springer, Berlin (1985)
21. Hintermüller, M.: Inverse coefficient problems for variational inequalities: optimality conditions and numerical realization. *M2AN Math. Model. Numer. Anal.* **35**(1), 129–152 (2001)
22. Hintermüller, M.: An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities. *Inverse Probl.* **24**(3), 034017, 23 (2008)
23. Hintermüller, M., Ito, K., Kunisch, K.: The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13**(3), 865–888 (2003), 2002 (electronic)
24. Hintermüller, M., Kopacka, I.: Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM J. Optim.* **20**(2), 868–902 (2009) (electronic)

25. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Classics in Applied Mathematics, vol. 31. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2000). Reprint of the 1980 original
26. Kopacka, I.: MPECs/MPCCs in function space: first order optimality concepts, path-following, and multilevel algorithms. PhD thesis, Department of Mathematics and Scientific Computing, University of Graz (2009)
27. Luo, Z.-Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
28. Mignot, F., Puel, J.-P.: Optimal control in some variational inequalities. *SIAM J. Control Optim.* **22**(3), 466–476 (1984)
29. Neittaanmaki, P., Sprekels, J., Tiba, D.: Optimization of Elliptic Systems. Springer Monographs in Mathematics. Springer, New York (2006). Theory and applications
30. Outrata, J., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. *Nonconvex Optimization and Its Applications*, vol. 28. Kluwer Academic, Dordrecht (1998). Theory, applications and numerical results
31. Rodrigues, J.-F.: Obstacle Problems in Mathematical Physics. North-Holland Mathematics Studies, vol. 134. North-Holland, Amsterdam (1987). *Notas de Matemática [Mathematical Notes]*, 114
32. Scheel, H., Scholtes, S.: Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. *Math. Oper. Res.* **25**(1), 1–22 (2000)
33. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim.* **11**(4), 918–936 (2001) (electronic)
34. Tröltzsch, F.: *Optimale Steuerung partieller Differentialgleichungen*. Vieweg, Wiesbaden (2005)
35. Ye, J.J.: Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *J. Math. Anal. Appl.* **307**(1), 350–369 (2005)
36. Zowe, J., Kurcyusz, S.: Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.* **5**(1), 49–62 (1979)