

INVERSE COEFFICIENT PROBLEMS FOR VARIATIONAL INEQUALITIES: OPTIMALITY CONDITIONS AND NUMERICAL REALIZATION*

MICHAEL HINTERMÜLLER¹

Abstract. We consider the identification of a distributed parameter in an elliptic variational inequality. On the basis of an optimal control problem formulation, the application of a primal-dual penalization technique enables us to prove the existence of multipliers giving a first order characterization of the optimal solution. Concerning the parameter we consider different regularity requirements. For the numerical realization we utilize a complementarity function, which allows us to rewrite the optimality conditions as a set of equalities. Finally, numerical results obtained from a least squares type algorithm emphasize the feasibility of our approach.

Mathematics Subject Classification. 49N50, 35R30, 35J85.

Received: December 27, 1999. Revised: November 9, 2000.

1. INTRODUCTION

In this paper, we focus on the identification of a distributed parameter in a variational inequality. Specifically, we consider the problem of identifying $u \in U := \{w \in H^2(\Omega) | w \geq \underline{\epsilon} > 0 \text{ a.e. in } \Omega\}$ in

$$\int_{\Omega} e(u) \nabla y \nabla (v - y) dx \geq \langle f(u), v - y \rangle \quad \text{for all } v \in K, \quad y \in K, \quad (1.1)$$

from given data $y_d \in L^2(\Omega)$, where $K = \{y \in H_o^1(\Omega) | y \geq 0\}$. We assume that $e \in C^2(\mathbb{R}_+; \mathbb{R}_+)$, with $e(z) \geq \epsilon > 0$ for all $z \geq \underline{\epsilon} > 0$, and that $e^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ exists. Moreover, $f(u) = Fu + g$, with $g \in H^{-1}(\Omega)$ and $F \in L(U, H^{-1}(\Omega))$ completely continuous, *i.e.* $\{u_n\} \subset U$ converging to u weakly in U implies that $\{Fu_n\}$ converges to Fu strongly in $H^{-1}(\Omega)$. Further, $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{H^{-1}, H_o^1}$ denotes the duality pairing between $H_o^1(\Omega)$ and its dual $H^{-1}(\Omega)$. The domain Ω is a bounded subset of \mathbb{R}^d , with $1 \leq d \leq 3$ and a sufficiently smooth boundary Γ .

Identification problems for variational inequalities of type (1.1) frequently occur in practical applications. One instance is the elasto-hydrodynamic lubrication problem in a journal bearing, where $e(z) = \mu^{-1} z^3$, $f(u) = -c \frac{\partial u}{\partial x_2}$, with μ the constant viscosity coefficient and c a constant relative velocity; see for instance [2, 10, 15, 16]. The coefficient depends on the distributed height function u between two rotating surfaces, and y represents the pressure in the lubricant which fills the gap between the surfaces. The pressure y must satisfy $y \geq 0$, and

Keywords and phrases. Bilevel problem, complementarity function, inverse problem, optimal control, variational inequality.

* The work was supported by the Fonds zur Förderung der wissenschaftlichen Forschung under "Spezialforschungsbereich Optimierung und Kontrolle", SFB03.

¹ Karl-Franzens University of Graz, Department of Mathematics, Heinrichstraße 36, 8010 Graz, Austria.
e-mail: michael.hintermueller@kfunigraz.ac.at

u has to be strictly positive in order to avoid damage of the surfaces. Since only the pressure is accessible *via* measurements y_d , the task is to recover the distributed height function u from y_d .

A commonly used technique to identify the parameter u in (1.1) from measurements y_d is to use a least squares formulation which, in our case, results in the following *bilevel* optimal control problem where we consider

$$(P) \quad \begin{aligned} & \text{minimize} && \frac{1}{2} \|y(u) - y_d\|_{L^2}^2 + \frac{\alpha}{2} \|u\|_U^2 && (1.2) \\ & \text{subject to} && u \in U, && (1.3) \end{aligned}$$

$$y(u) = \operatorname{argmin} \left\{ \frac{1}{2} \int_{\Omega} e(u) |\nabla y|^2 dx - \langle f(u), y \rangle \mid y \geq 0 \right\}, \quad (1.4)$$

with $\|\cdot\|_{L^2}$ denoting the L^2 -norm in Ω . Moreover, U is endowed with the norm $\|\cdot\|_U = \|\cdot\|_{H^2}$. We remark that in Section 4.3 we also address the case of rough coefficients, *i.e.* we reduce the regularity requirements for the parameter u . The term $\frac{\alpha}{2} \|u\|_U^2$ in (1.2) corresponds to Tikhonov's regularization with parameter $\alpha > 0$. Generalizations of (1.2), like *minimize* $h_1(y) + h_2(u)$, are possible which (under appropriate assumptions on h_1 and h_2) do not pose additional difficulties; see [1, 4, 18].

The term *bilevel* refers to the fact that in the optimal control problem (P) (1.4) again is a lower level infinite dimensional optimization problem. Compared to a standard (constrained) optimal control problem, the bilevel character of (P) poses additional difficulties. In fact, if (1.4) in (P) is replaced by its optimality system (as a so called *equilibrium constraint* for (P)), *i.e.*

$$\begin{aligned} \int_{\Omega} e(u) \nabla y(u) \nabla v dx - \langle f(u), v \rangle - \int_{\Omega} \lambda v dx &= 0 \quad \text{for all } v \in H_o^1(\Omega), \\ \lambda \geq 0, \quad y(u) \geq 0, \quad \int_{\Omega} \lambda y(u) dx &= 0, \end{aligned} \quad (1.5)$$

then the existence of multipliers for the upper level problem may fail. See for instance [7, 18] for a discussion in the case where u enters the variational inequality in an affine way. Note also that the multiplier of the lower level problem, *i.e.* λ , appears as a primal variable in the upper level problem.

In this paper we guarantee existence of multipliers for (P) where classical (Lagrange) approaches based on (1.5) fail; see for instance [7]. This is achieved by a primal-dual penalization technique. In addition, the optimality system for (P), which is derived on the basis of this penalization technique and the utilization of the concept of complementarity functions, is amenable to numerical realization. We shall mention that our approach extending a technique used in [18] differs significantly from relaxation and/or dualization techniques like in [4, 5, 19], regularization techniques like in [1], and techniques based on the conical derivative as in [22]. Besides the nonlinearity considered here, the first order characterization derived subsequently is more general than the one in [8], where the optimality system is based on two special directions in control space. Moreover, in contrast to the newly derived system many of the aforementioned first order characterizations cannot be used for numerical realization.

In order to find numerically a stationary point for problem (P), the discretized first order system is solved by a stabilized Gauss Newton method; see [12]. The choice of algorithm together with its globalization strategy is based on smoothness properties of the reformulation of the complementarity condition. Analogous ideas are developed in [13] and [21] but in a different context. Moreover, we take care of the fact that without further assumptions the parameter u cannot be estimated from y on the singular set $S_o = \{x \in \Omega \mid \nabla y(x) = 0\}$; see [17] in the case of variational equalities.

The paper is organized as follows: In Section 2 we prove the existence of a solution of (P). The primal-dual reformulation of the lower level problem eliminating $y \geq 0$ from the set of explicit constraints is introduced in Section 3. Moreover, the complementarity condition is reformulated by means of a complementarity function. This technique results in an equivalent formulation of (P) which is well suited for numerical realization. Section 4 is concerned with the derivation of first order conditions for (P). This is done by regularization and passage

to the limit. Another aspect addressed in Section 4 is the reduction of regularity requirements for u . In Section 5 we describe a Gauss Newton based method to solve the discretized first order system. Finally, numerical results shall emphasize the feasibility of our approach.

Throughout the paper we shall invoke the following notation: The norm in a space $S(\Omega)$ is denoted by $|\cdot|_S$, (\cdot, \cdot) stands for the L^2 -inner product. By $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{H^{-1}, H_0^1}$ we indicate the duality pairing between $H_0^1(\Omega)$ and its dual $H^{-1}(\Omega)$. The relations “ \geq ” and “ $=$ ” in function spaces are understood in the pointwise almost everywhere sense. Moreover, C, C_1, C_2 shall denote generic positive constants which can take different values on different occasions. We use “ \rightarrow ” for convergence in the strong sense, “ \rightharpoonup ” for convergence in the weak sense, and “ $\overset{*}{\rightharpoonup}$ ” for weak* convergence.

2. EXISTENCE OF A SOLUTION OF PROBLEM (P)

In this section, we will prove that the bilevel optimal control problem (P) admits a solution $(y^*, u^*) \in H_0^1(\Omega) \times U$. First note that for fixed $u \in U$ the functional

$$J^u(y) = \frac{1}{2} \int_{\Omega} e(u) |\nabla y|^2 dx - \langle f(u), y \rangle, \quad J^u : H_0^1(\Omega) \rightarrow \mathbb{R}$$

is Gateaux differentiable and, due to $e(u) \geq \epsilon > 0$ a.e. in Ω , strictly convex. The set $K = \{y \in H_0^1(\Omega) | y \geq 0\}$ is convex and closed. Thus by standard arguments it is seen that (1.4) admits a unique solution $y(u) \in H_0^1(\Omega)$. The optimal $y(u)$ is characterized by

$$\langle J^{u'}(y(u)), y - y(u) \rangle \geq 0 \quad \text{for all } y \in K \tag{2.1}$$

where $J^{u'}$ denotes the Gateaux derivative of J^u . Note that for $v \in H_0^1(\Omega)$

$$\langle J^{u'}(y), v \rangle = a^u(y, v) - \langle f(u), v \rangle, \tag{2.2}$$

with $a^u(y, v) = (e(u) \nabla y, \nabla v)$, and that due to our assumptions on U and e the bilinear form $a^u : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ is symmetric, bounded and $H_0^1(\Omega)$ -elliptic. The boundedness ensues from the compact embedding of $H^2(\Omega)$ in $C^{0,\nu}(\bar{\Omega})$, with $0 < \nu < \frac{1}{2}$, and the regularity of e . If we choose $y = 0 \in K$ in (2.1), then

$$C|y(u)|_{H_0^1}^2 \leq a^u(y(u), y(u)) \leq \langle f(u), y(u) \rangle \leq |f(u)|_{H^{-1}} |y(u)|_{H_0^1}$$

implying $|y(u)|_{H_0^1} \leq C|f(u)|_{H^{-1}}$.

For the mapping $u \mapsto y(u)$ the following continuity property holds.

Lemma 2.1. *Let $\Phi : U \rightarrow H_0^1(\Omega)$ be given by $\Phi(u) = y(u)$ and $\{u_n\} \subset U$ be a sequence converging to u weakly in U . Then $\{\Phi(u_n)\}$ converges to $y(u)$ strongly in $H_0^1(\Omega)$, i.e. the mapping Φ is completely continuous.*

Proof. For $u \in U$ define the operator $A^u \in L(H_0^1(\Omega), H^{-1}(\Omega))$ by $\langle A^u y, v \rangle = a^u(y, v)$ for all $y, v \in H_0^1(\Omega)$. Since $H^2(\Omega)$ is compactly embedded in $C^0(\bar{\Omega})$ and $e \in C^2(\mathbb{R}_+; \mathbb{R}_+)$, we obtain

$$|A^{u_n} - A^u|_{H^{-1} \leftarrow H_0^1} \rightarrow 0 \quad \text{for } u_n \rightharpoonup u \text{ in } U. \tag{2.3}$$

From (2.1) and (2.2) we obtain

$$\langle A^{u_n} y(u_n), y(u_n) - y(u_m) \rangle - \langle f(u_n), y(u_n) - y(u_m) \rangle \leq 0$$

and (by interchanging the role of $y(u_n)$ and $y(u_m)$)

$$\langle A^{u_m} y(u_m), y(u_m) - y(u_n) \rangle - \langle f(u_m), y(u_m) - y(u_n) \rangle \leq 0.$$

Addition of the previous two equations and estimation yield

$$\begin{aligned} 0 &\geq \langle A^{u_n}(y(u_n) - y(u_m)) + (A^{u_n} - A^{u_m})y(u_m), y(u_n) - y(u_m) \rangle \\ &\quad + \langle f(u_m) - f(u_n), y(u_n) - y(u_m) \rangle \\ &\geq C|y(u_n) - y(u_m)|_{H_o^1}^2 - |A^{u_n} - A^{u_m}|_{H^{-1} \leftarrow H_o^1} |y(u_m)|_{H_o^1} |y(u_n) - y(u_m)|_{H_o^1} \\ &\quad - |f(u_m) - f(u_n)|_{H^{-1}} |y(u_n) - y(u_m)|_{H_o^1}. \end{aligned}$$

Hence, we obtain

$$|y(u_n) - y(u_m)|_{H_o^1} \leq C(|A^{u_n} - A^{u_m}|_{H^{-1} \leftarrow H_o^1} |f(u_m)|_{H^{-1}} + |f(u_m) - f(u_n)|_{H^{-1}}).$$

Since $\{u_n\}$ converges to u weakly in U (by assumption), the previous estimate together with (2.3) and the complete continuity of f yield the existence of $\bar{y} \in H_o^1(\Omega)$ such that $\{y(u_n)\}$ converges to \bar{y} strongly in $H_o^1(\Omega)$. Letting n tend to ∞ in $a^{u_n}(y(u_n), y - y(u_n)) \geq \langle f(u_n), y - y(u_n) \rangle$ results in $a^u(\bar{y}, y - \bar{y}) \geq \langle f(u), y - \bar{y} \rangle$ for all $y \in K$. Then $\bar{y} = y(u)$ follows from the uniqueness of $y(u)$. \square

Next we can establish the main result of this section.

Theorem 2.2. *There exists an optimal solution $(y^*, u^*) \in H_o^1(\Omega) \times U$ of (P).*

Proof. Let $\kappa = \inf\{\frac{1}{2}|y(u) - y_d|_{L^2}^2 + \frac{\alpha}{2}|u|_U^2 | u \in U\}$, and let $\{u_n\} \subset U$ be a sequence satisfying

$$\kappa \leq \frac{1}{2}|y(u_n) - y_d|_{L^2}^2 + \frac{\alpha}{2}|u_n|_U^2 \leq \kappa + \frac{1}{n}.$$

Then $\{u_n\}$ is weakly compact in U . Hence, there exists a subsequence $\{u_{n(k)}\}$ such that $u_{n(k)} \rightharpoonup u^*$ in U and $y(u_{n(k)}) \rightarrow y(u^*) =: y^*$ in $H_o^1(\Omega)$, where the second assertion holds due to Lemma 2.1. The weak lower semicontinuity of norms yields

$$\frac{1}{2}|y^* - y_d|_{L^2}^2 + \frac{\alpha}{2}|u^*|_U^2 = \kappa,$$

and hence (y^*, u^*) is optimal solution of (P). \square

3. REFORMULATION OF THE LOWER LEVEL PROBLEM

In this section, we shall utilize a primal-dual reformulation technique for (1.4) in order to eliminate $y \in K$ from the set of explicit constraints. In the spirit of penalizing violations of $y \geq 0$, let us consider the problem

$$\text{minimize } J_c^u(y) := J^u(y) + \frac{1}{2c} |\max\{\bar{\lambda} - cy, 0\}|_{L^2}^2 \quad \text{over } y \in H_o^1(\Omega), \quad (3.1)$$

where $\bar{\lambda} \in L^2(\Omega)$, with $\bar{\lambda} \geq 0$, is arbitrarily fixed, and $c > 0$. The role of $\bar{\lambda}$ is discussed at the end of this section. Note that $J^u(\cdot)$ and $|\max\{\bar{\lambda} - c \cdot, 0\}|_{L^2}^2$ are Gateaux-differentiable and strictly convex and convex, respectively. This together with boundedness from below, radial unboundedness and semi continuity of $J_c^u(\cdot)$ guarantee the existence and uniqueness of the solution $y_c(u) \in H_o^1(\Omega)$ of (3.1). It is readily checked (by differentiation of J_c^u) that it satisfies

$$a^u(y_c(u), d_y) - \langle f(u), d_y \rangle - (\max\{\bar{\lambda} - cy_c(u), 0\}, d_y) = 0 \quad (3.2)$$

for all $d_y = y - y_c(u)$ with $y \in H_o^1(\Omega)$. Observe that for $d_y = -y_c(u)$ condition (3.2) becomes

$$a^u(y_c(u), y_c(u)) = \langle f(u), y_c(u) \rangle + (\max\{\bar{\lambda} - cy_c(u), 0\}, y_c(u)).$$

Since $(\max\{\bar{\lambda} - cy_c(u), 0\}, y_c(u)) \leq |\bar{\lambda}|_{L^2} |y_c(u)|_{H_o^1}$, it follows that $|y_c(u)|_{H_o^1} \leq C(|f(u)|_{H^{-1}} + |\bar{\lambda}|_{L^2})$. Note that due to our definition of $A^u \in L(H_o^1(\Omega), H^{-1}(\Omega))$ the first order optimality condition (3.2) can equivalently be written as

$$A^u y_c(u) - f(u) - \max\{\bar{\lambda} - cy_c(u), 0\} = 0 \quad \text{in } H^{-1}(\Omega). \quad (3.3)$$

Next we shall study the asymptotic behaviour of $\{y_{c_l}(u)\}$ for $c_l \rightarrow \infty$.

Lemma 3.1. *Let $\{c_l\} \subset \mathbb{R}_+$ be a sequence of penalty parameters satisfying $c_l \rightarrow \infty$ for $l \rightarrow \infty$. Then the sequence $\{y_{c_l}(u)\}$ converges to $y(u)$ strongly in $H_o^1(\Omega)$.*

Proof. First observe that

$$J^u(y_{c_l}(u)) \leq J_{c_l}^u(y_{c_l}(u)) \leq J_{c_l}^u(y) \leq J^u(y) + \frac{1}{2c_l} |\bar{\lambda}|_{L^2}^2 \quad \text{for all } y \in K. \quad (3.4)$$

For the specific choice $y = 0 \in K$ we obtain

$$\begin{aligned} \frac{1}{2c_l} |\bar{\lambda}|_{L^2}^2 &\geq \frac{1}{2c_l} |\max\{\bar{\lambda} - c_l y_{c_l}(u), 0\}|_{L^2}^2 + J^u(y_{c_l}(u)) \\ &\geq \frac{1}{2c_l} |\max\{\bar{\lambda} - c_l y_{c_l}(u), 0\}|_{L^2}^2 - \langle f(u), y_{c_l}(u) \rangle + C|y_{c_l}(u)|_{H_o^1}^2. \end{aligned} \quad (3.5)$$

This yields $|y_{c_l}(u)|_{H_o^1} \leq C$, where the constant C is independent of $\{c_l\}$. Hence, there exists a subsequence $\{c_{l(k)}\}$ such that $y_{c_{l(k)}}(u) \rightharpoonup \bar{y}$ in $H_o^1(\Omega)$. For $k \rightarrow \infty$ (3.5) yields

$$|\max\{\frac{\bar{\lambda}}{c_{l(k)}} - y_{c_{l(k)}}(u), 0\}|_{L^2}^2 \rightarrow 0$$

implying $\bar{y} \in K$, and cf.(3.4)

$$J^u(\bar{y}) = \frac{1}{2} a^u(\bar{y}, \bar{y}) - \langle f(u), \bar{y} \rangle \leq \frac{1}{2} a^u(y, y) - \langle f(u), y \rangle = J^u(y) \quad \text{for all } y \in K.$$

From the uniqueness of the solution of (1.4) we obtain $\bar{y} = y(u)$, and thus $y_{c_l}(u) \rightarrow y(u)$ in $H_o^1(\Omega)$.

Moreover, we deduce $\lim a^u(y_{c_l}(u), y_{c_l}(u)) = a^u(y(u), y(u))$. For sufficiently small $\gamma > 0$ the functional $\Psi(z) = a^u(z, z) - \gamma|z|_{H_o^1}^2$ is weakly lower semicontinuous. Hence,

$$\begin{aligned} a^u(y(u), y(u)) - \gamma|y(u)|_{H_o^1}^2 &\leq \liminf_{l \rightarrow \infty} \left(a^u(y_{c_l}(u), y_{c_l}(u)) - \gamma|y_{c_l}(u)|_{H_o^1}^2 \right) \\ &\leq a^u(y(u), y(u)) - \limsup_{l \rightarrow \infty} \gamma|y_{c_l}(u)|_{H_o^1}^2 \end{aligned}$$

implying $\limsup \gamma|y_{c_l}(u)|_{H_o^1}^2 \leq \gamma|y(u)|_{H_o^1}^2$. This together with the weak lower semicontinuity yields

$$\lim \gamma|y_{c_l}(u)|_{H_o^1}^2 = \gamma|y(u)|_{H_o^1}^2,$$

and thus $y_{c_l}(u) \rightarrow y(u)$ in $H_o^1(\Omega)$. □

Based on the reformulation (3.1) we shall now consider the penalized version of (P) which is to

$$\begin{aligned} (\text{P}_c) \quad & \text{minimize} && \frac{1}{2} |y_c(u) - y_d|_{L^2}^2 + \frac{\alpha}{2} |u|_U^2 \\ & \text{subject to} && u \in U, \\ & && y_c(u) = \operatorname{argmin} \{J_c^u(y) | y \in H_o^1(\Omega)\} \end{aligned} \quad (3.6)$$

Before we prove existence of a solution of (P_c) , we state a continuity result which is analogous to the complete continuity of Φ (see Lem. 2.1). For this purpose define

$$\Phi_c : U \rightarrow H_o^1(\Omega), \quad \Phi_c(u) = y_c(u).$$

Lemma 3.2. *The mapping Φ_c is completely continuous uniformly in $c \in \mathbb{R}^+$.*

Proof. Consider the first order condition (3.2) first with $(u, d_y) = (u_n, y_c(u_m) - y_c(u_n)) \in U \times H_o^1(\Omega)$ and then with $(u, d_y) = (u_m, y_c(u_n) - y_c(u_m)) \in U \times H_o^1(\Omega)$. Adding both resulting equalities yields

$$\begin{aligned} 0 &= a^{u_n}(y_c(u_n), y_c(u_m) - y_c(u_n)) - a^{u_m}(y_c(u_m), y_c(u_m) - y_c(u_n)) \\ &\quad - (\max\{\bar{\lambda} - cy_c(u_n), 0\} - \max\{\bar{\lambda} - cy_c(u_m), 0\}, y_c(u_m) - y_c(u_n)) \\ &\quad + \langle f(u_m) - f(u_n), y_c(u_m) - y_c(u_n) \rangle, \end{aligned}$$

which is equivalent to

$$\begin{aligned} 0 &= \langle A^{u_n}(y_c(u_m) - y_c(u_n)) + (A^{u_m} - A^{u_n})y_c(u_m), y_c(u_m) - y_c(u_n) \rangle \\ &\quad + (\max\{\bar{\lambda} - cy_c(u_n), 0\} - \max\{\bar{\lambda} - cy_c(u_m), 0\}, y_c(u_m) - y_c(u_n)) \\ &\quad - \langle f(u_m) - f(u_n), y_c(u_m) - y_c(u_n) \rangle. \end{aligned}$$

There obviously holds

$$(\max\{\bar{\lambda} - cy_c(u_n), 0\} - \max\{\bar{\lambda} - cy_c(u_m), 0\}, y_c(u_m) - y_c(u_n)) \geq 0.$$

Thus we obtain

$$|y_c(u_m) - y_c(u_n)|_{H_o^1} \leq C (|A^{u_n} - A^{u_m}|_{H^{-1} \leftarrow H_o^1} (|f(u_m)|_{H^{-1}} + |\bar{\lambda}|_{L^2}) + |f(u_m) - f(u_n)|_{H^{-1}}). \quad (3.7)$$

From $|A^{u_n} - A^u|_{H^{-1} \leftarrow H_o^1} \rightarrow 0$ for $u_n \rightarrow u$ in U , the complete continuity of f and the uniqueness of $y_c(u)$ we deduce the complete continuity of Φ_c . Uniformity w.r.t. c follows from $|y_c(u)|_{H_o^1} \leq C(|f(u)|_{H^{-1}} + |\bar{\lambda}|_{L^2})$ and (3.7) where in both cases the upper bounds do not depend on c . \square

Concerning the existence of a solution of (P_c) , the following result holds.

Theorem 3.3. *There exists an optimal solution $(y_c^*, u_c^*) \in H_o^1(\Omega) \times U$ of (P_c) .*

Proof. The proof follows the lines for that of Theorem 2.2. We only have to consider $y_c(u)$ and Lemma 3.2 instead of $y(u)$ and Lemma 2.1. \square

In the sequel, we shall call $(y, u) \in H_o^1(\Omega) \times U$ a strong-weak accumulation point of a sequence $\{(y_n, u_n)\}$ if there exists a subsequence $\{n(k)\}$ such that $y_{n(k)} \rightarrow y$ in $H_o^1(\Omega)$ and $u_{n(k)} \rightarrow u$ in U , i.e. $\{(y_{n(k)}, u_{n(k)})\}$ converges strongly-weakly to (y, u) .

Lemma 3.2 and Theorem 3.3 yield the following result.

Theorem 3.4. *Let $\{c_n\} \subset \mathbb{R}_+$ be a sequence of penalty parameters satisfying $c_n \rightarrow \infty$ for $n \rightarrow \infty$, and let $(y_{c_n}(u_{c_n}), u_{c_n})$ be a solution of (P_{c_n}) . Then a strong-weak accumulation point $(y^*, u^*) \in H_o^1(\Omega) \times U$ of the sequence $\{(y_{c_n}(u_{c_n}), u_{c_n})\}$ as $c_n \rightarrow \infty$ exists, and every such accumulation point is a solution of (P) .*

Proof. Let $\{(y_{c_n}(u_{c_n}), u_{c_n})\}$ denote a sequence of optimal solutions to (P_c) with c replaced by the sequence $\{c_n\}$. For arbitrary $u \in U$ let $y_{c_n}(u)$ be a solution of (3.2) for $c = c_n$. Then

$$\frac{1}{2}|y_{c_n}(u_{c_n}) - y_d|_{L^2}^2 + \frac{\alpha}{2}|u_{c_n}|_U^2 \leq \frac{1}{2}|y_{c_n}(u) - y_d|_{L^2}^2 + \frac{\alpha}{2}|u|_U^2. \quad (3.8)$$

From $|y_{c_n}(u)| \leq C(|f(u)|_{H^{-1}} + |\bar{\lambda}|_{L^2})$ we infer that $|u_{c_n}|_U$ is bounded independently of $\{c_n\}$. Hence, there exist a subsequence $\{n(k)\}$ and $u^* \in U$ such that $u_{c_{n(k)}} \rightarrow u^*$ in U . Moreover, we have

$$|y_{c_{n(k)}}(u_{c_{n(k)}}) - y(u^*)|_{H_o^1} \leq |y_{c_{n(k)}}(u_{c_{n(k)}}) - y_{c_{n(k)}}(u^*)|_{H_o^1} + |y_{c_{n(k)}}(u^*) - y(u^*)|_{H_o^1},$$

where the first term on the right hand side vanishes as $k \rightarrow \infty$ by Lemma 3.2, and the second term becomes zero as $k \rightarrow \infty$ by Lemma 3.1. Thus, $y_{c_{n(k)}}(u_{c_{n(k)}}) \rightarrow y(u^*) =: y^*$ in $H_o^1(\Omega)$. We further deduce that

$$a^{u_{c_{n(k)}}}(y_{c_{n(k)}}(u_{c_{n(k}}), y_{c_{n(k)}}(u_{c_{n(k}}))) \rightarrow a^{u^*}(y^*, y^*).$$

From Lemma 3.1 it follows that $y_{c_{n(k)}}(u) \rightarrow y(u)$ in $H_o^1(\Omega)$. Therefore, from (3.8) we infer

$$\frac{1}{2}|y^* - y_d|_{L^2}^2 + \frac{\alpha}{2}|u^*|_U^2 \leq \frac{1}{2}|y(u) - y_d|_{L^2}^2 + \frac{\alpha}{2}|u|_U^2$$

for all solutions $(y(u), u)$ of (1.4). This proves the assertion. □

The remainder of this section clarifies the role of $\bar{\lambda}$. In fact, for a specific choice we derive important properties of the solution of (3.2). But before we state the corresponding result, we shall introduce the notion of a *complementarity function*. For optimization problems with inequality constraints the first order conditions typically include a complementarity condition. For instance, in the case of the simple inequality constraint $y \geq 0$ the corresponding complementarity condition is

$$y \geq 0, \quad \lambda \geq 0, \quad (y, \lambda) = 0,$$

where λ denotes the pertinent Lagrange multiplier. This condition is not immediately amenable to numerical realization, and in the context of bilevel problems (where the complementarity condition becomes a constraint in the upper level problem) the existence of Lagrange multipliers may fail; see [7, 14, 18, 20]. A possible remedy is based on a reformulation with a complementarity function. A function $\Theta : \mathbb{R}^2 \rightarrow \mathbb{R}$ is called complementarity function iff the relation

$$\Theta(a, b) = 0 \iff a \geq 0, \quad b \geq 0, \quad ab = 0$$

is satisfied. There exist many instances in the literature like the Fischer-Burmeister function $\Theta_{\text{FB}}(a, b) = \sqrt{a^2 + b^2} - (a + b)$ (see [13]), or the Moreau-Yosida based function $\Theta_{\text{MY}}(a, b) = a - \max\{a - cb, 0\}$, with $c > 0$ arbitrarily fixed (see [6]). For other choices (for more general complementarity problems) and references we refer to [21].

This concept of reformulation by means of complementarity functions is used in the next result. From now on, we invoke the following assumption:

$$f : U \rightarrow L^2(\Omega). \tag{A}$$

Theorem 3.5. *Let (A) be fulfilled, and let $\{c_n\} \subset \mathbb{R}_+$ satisfy $c_n \rightarrow \infty$ for $n \rightarrow \infty$ and Θ be a complementarity function. For $\bar{\lambda} := \bar{\lambda}(u) = \max\{-f(u), 0\}$ we have $y_{c_n}(u) \geq 0$. Moreover $\{\lambda_{c_n}(u)\}_{n=1}^\infty = \{\max\{\bar{\lambda}(u) - c_n y_{c_n}(u), 0\}\}_{n=1}^\infty$ is uniformly bounded in $L^2(\Omega)$ and converges weakly to $\lambda(u) \in L^2(\Omega)$ satisfying*

$$A^u y(u) = f(u) + \lambda(u), \quad \text{and} \quad \Theta(\lambda(u), y(u)) = 0 \quad \text{a.e. in } \Omega. \tag{3.9}$$

Proof. Multiplying (3.3) by $v = \max\{-y_{c_n}(u), 0\}$ yields

$$0 = \langle A^u y_{c_n}(u), v \rangle - \langle f(u) + \bar{\lambda}(u), v \rangle - c_n \langle v, v \rangle \leq \langle A^u y_{c_n}(u), v \rangle \leq -C|v|_{H_o^1}^2 \leq 0.$$

Assume that $\Omega_{c_n}^- = \{y_{c_n}(u) < 0\} \neq \emptyset$. From the above relation we deduce that $y_{c_n}(u)|_{\Omega_{c_n}^-} = 0$ which is a contradiction. Hence, we have $y_{c_n}(u) \geq 0$ a.e. in Ω . From the definition of $\lambda_{c_n}(u)$ we immediately obtain

$$0 \leq \lambda_{c_n}(u) \leq \bar{\lambda}(u) \text{ a.e. in } \Omega.$$

This implies that there exists a weakly convergent subsequence $\{\lambda_{c_{n(k)}}(u)\}$. Let $\lambda(u) \in L^2(\Omega)$ denote the weak limit point. For $c_{n(k)} \rightarrow \infty$ we obtain from Lemma 3.1 that $y_{c_{n(k)}}(u)$ converges to $y(u)$ strongly in $H_o^1(\Omega)$. Thus we have $\langle A^u y_{c_{n(k)}}(u), y_{c_{n(k)}}(u) \rangle \rightarrow \langle f(u) + \lambda(u), y(u) \rangle$. Since

$$\langle A^u y - A^u z, y - z \rangle = a^u(y - z, y - z) \geq C|y - z|_{H_o^1}^2 \quad \text{for all } y, z \in H_o^1(\Omega),$$

the operator A^u is monotone. Moreover A^u is hemicontinuous, and thus maximal monotone [1]. Therefore, we have

$$\langle A^u y_{c_{n(k)}}(u) - A^u z, y_{c_{n(k)}}(u) - z \rangle \geq 0$$

implying for $c_{n(k)} \rightarrow \infty$

$$\langle f(u) + \lambda(u) - A^u z, y(u) - z \rangle \geq 0 \quad \text{for all } z \in H_o^1(\Omega).$$

The maximal monotonicity of A^u then yields $A^u y(u) = f(u) + \lambda(u)$. This shows the uniqueness of the weak limit $\lambda(u)$ of $\{\lambda_{c_{n(k)}}(u)\}$. Moreover, for all $c_n > 0$ we have $\lambda_{c_n}(u) \in \{\lambda \in L^2(\Omega) | \lambda \geq 0\}$ implying $\lambda_{c_n}(u) \geq 0$ and

$$-\frac{\bar{\lambda}(u)}{c_n} < -y_{c_n}(u) \leq 0 \quad \text{a.e. in } \{x \in \Omega | \lambda_{c_n}(u)(x) > 0\}.$$

Hence, $y_{c_n}(u)\lambda_{c_n}(u) \rightarrow 0$ in $L^1(\Omega)$ for $c_n \rightarrow \infty$. From $y_{c_n}(u)\lambda_{c_n}(u) \xrightarrow{*} y(u)\lambda(u)$ in $L^\infty(\Omega)^*$ we deduce $y(u)\lambda(u) = 0$ a.e. in Ω , and it follows that $\Theta(\lambda(u), y(u)) = 0$ a.e. in Ω . This completes the proof. \square

4. FIRST ORDER CONDITIONS

This section is devoted to the development of first order necessary optimality conditions for the bilevel problem (P). First we regularize the non differentiable max-operation which appears in (3.2).

4.1. Regularization

Consider the following C^1 -regularization of $x \mapsto \max\{x, 0\}$:

$$\max_c\{x, 0\} = \begin{cases} x & \text{for } x \geq \frac{1}{2c}, \\ \frac{c}{2}(x + \frac{1}{2c})^2 & \text{for } |x| \leq \frac{1}{2c}, \\ 0 & \text{for } x \leq -\frac{1}{2c}. \end{cases}$$

Then there obviously holds $\max_c\{x, 0\} = \int_{-\infty}^x \text{sgn}_c(t) dt$, with

$$\text{sgn}_c(x) = \begin{cases} 1 & \text{for } x \geq \frac{1}{2c}, \\ c(x + \frac{1}{2c}) & \text{for } |x| \leq \frac{1}{2c}, \\ 0 & \text{for } x \leq -\frac{1}{2c}. \end{cases}$$

Note that $\text{sgn}_c(x) \geq 0$ for all x .

Next consider the regularized version of (P_c) which is to

$$\begin{aligned} (\tilde{P}_c) \quad & \text{minimize} && \frac{1}{2}|y - y_d|_{L^2}^2 + \frac{\alpha}{2}|u|_U^2 \\ & \text{subject to} && u \in U, \\ & && A^u y - f(u) - \max_c\{\bar{\lambda} - cy, 0\} = 0, \end{aligned}$$

where $c > 0$ and $\bar{\lambda} \in L^2(\Omega)$, with $\bar{\lambda} \geq 0$. Define $\Psi_c(x) = \int_0^x \max_c\{t, 0\}dt$, and consider

$$\text{minimize } \tilde{J}_c^u(y) := J^u(y) + \int_{\Omega} \frac{1}{c} \Psi_c(\bar{\lambda} - cy) dx \quad \text{over } y \in H_o^1(\Omega). \quad (4.1)$$

Some simple manipulations show that

$$0 \leq \Psi_c(x) \leq \max\left(\frac{7}{48c^2}, \frac{x^2}{2} + \frac{1}{48c^2}\right) \quad \text{for } x \geq 0, \quad -\frac{1}{48c^2} \leq \Psi_c(x) \quad \text{for } x < 0.$$

Moreover, from its definition it is easily seen that Ψ_c is monotonically increasing, and

$$\int_{\Omega} \frac{1}{c} \Psi_c(\bar{\lambda} - cv) dx \leq \int_{\Omega} \frac{1}{c} \Psi_c(\bar{\lambda}) dx < \infty \quad \text{for all } v \in K. \quad (4.2)$$

Concerning the existence of a solution of (4.1) the following result holds.

Lemma 4.1. (a) For all $c > 0$ and $u \in U$ problem (4.1) admits a unique solution $\tilde{y}_c(u) \in H_o^1(\Omega)$.
 (b) For $c_n \rightarrow \infty$ the sequence $\{\tilde{y}_{c_n}(u)\}$ converges to $y(u)$, the unique solution of (1.4), strongly in $H_o^1(\Omega)$.

Proof. (a) The existence of the unique solution $\tilde{y}_c(u)$ of (4.1) for some $c > 0$ follows from the monotonicity of Ψ_c , the property (4.2) and the properties of $J^u(y)$ (see Sect. 2).

(b) Due to the continuous differentiability of Ψ_c the unique solution satisfies the first order conditions

$$A^u \tilde{y}_c(u) - f(u) - \max_c\{\bar{\lambda} - c\tilde{y}_c(u), 0\} = 0. \quad (4.3)$$

Now consider (4.3) with $c = c_n$. Multiplication by $\tilde{y}_{c_n}(u)$ yields

$$C_1 |\tilde{y}_{c_n}(u)|_{H_o^1}^2 \leq |f(u)|_{L^2} |\tilde{y}_{c_n}(u)|_{H_o^1} + C_2 |\tilde{y}_{c_n}(u)|_{H_o^1}$$

implying $|\tilde{y}_{c_n}(u)|_{H_o^1} \leq C$, where the constant C is independent of c_n . Thus, there exists a subsequence $\{\tilde{y}_{c_{n(k)}}(u)\}$ converging to some \hat{y} weakly in $H_o^1(\Omega)$.

Next we show that $\hat{y} \in K$. For this purpose define

$$\Omega_{c_{n(k)}}^- = \{\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u) < 0\} \quad \text{and} \quad \Omega_{c_{n(k)}}^+ = \{\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u) \geq 0\},$$

and observe that

$$\tilde{J}_{c_n}^u(\tilde{y}_{c_n}(u)) \leq J^u(y) + \int_{\Omega} \frac{1}{c_n} \Psi_{c_n}(\bar{\lambda}) dx \quad \text{for all } y \in K.$$

For $y := 0 \in K$ this results in

$$\int_{\Omega} \frac{1}{c_n} \Psi_{c_n}(\bar{\lambda}) dx + C |f(u)|_{L^2} \geq \int_{\Omega} \frac{1}{c_n} \Psi_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}(u)) dx.$$

Since $\Psi_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}(u)) \geq -1/(48c_n^2)$, the sequence $\{\int_{\Omega} c_n^{-1} \Psi_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}(u)) dx\}$ is uniformly bounded. This yields the uniform boundedness of

$$\left\{ \int_{\Omega_{c_{n(k)}}^-} \frac{1}{c_{n(k)}} \Psi_{c_{n(k)}}(\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u)) dx \right\},$$

which tends to 0 as $c_{n(k)} \rightarrow \infty$, and the uniform boundedness of

$$\left\{ \int_{\Omega_{c_{n(k)}^+}} \frac{1}{c_{n(k)}} \Psi_{c_{n(k)}}(\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u)) dx \right\}.$$

From $\Psi_{c_{n(k)}}(x) \geq (x^2/2)$ for all $x \geq 0$ it follows that $c_{n(k)}^{-1} |\max\{\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u), 0\}|_{L^2}^2$ is uniformly bounded and

$$\lim_{k \rightarrow \infty} |\max\{\frac{1}{c_{n(k)}} \bar{\lambda} - \tilde{y}_{c_{n(k)}}(u), 0\}|_{L^2}^2 = 0.$$

Since $\bar{\lambda} \geq 0$, the last equation implies $\hat{y} \geq 0$.

Finally, let $v \in K$ be arbitrarily fixed. Then for some $0 < \eta < 1$

$$\begin{aligned} \frac{1}{\eta} \left(J^u(\tilde{y}_{c_{n(k)}}(u)) + \int_{\Omega} \frac{1}{c_{n(k)}} \Psi_{c_{n(k)}}(\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u)) dx \right) &\leq \frac{1}{\eta} \left(J^u(\tilde{y}_{c_{n(k)}}(u) + \eta(v - \tilde{y}_{c_{n(k)}}(u))) \right. \\ &\quad \left. + \int_{\Omega} \frac{1}{c_{n(k)}} \Psi_{c_{n(k)}}(\bar{\lambda} - c_{n(k)}(\tilde{y}_{c_{n(k)}}(u) + \eta(v - \tilde{y}_{c_{n(k)}}(u)))) dx \right) \end{aligned}$$

which is equivalent to

$$\begin{aligned} a^u(\tilde{y}_{c_{n(k)}}(u), v - \tilde{y}_{c_{n(k)}}(u)) + \frac{\eta}{2} a^u(v - \tilde{y}_{c_{n(k)}}(u), v - \tilde{y}_{c_{n(k)}}(u)) - \langle f(u), v - \tilde{y}_{c_{n(k)}}(u) \rangle \\ + \frac{1}{\eta} \int_{\Omega} \frac{1}{c_{n(k)}} \Psi_{c_{n(k)}}(\bar{\lambda} - c_{n(k)}(\tilde{y}_{c_{n(k)}}(u) + \eta(v - \tilde{y}_{c_{n(k)}}(u)))) dx \\ - \frac{1}{\eta} \int_{\Omega} \frac{1}{c_{n(k)}} \Psi_{c_{n(k)}}(\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}(u)) dx \geq 0. \end{aligned}$$

For $c_{n(k)} \rightarrow \infty$ it follows that

$$a^u(\hat{y}, v - \hat{y}) + \frac{\eta}{2} a^u(v - \hat{y}, v - \hat{y}) - \langle f(u), v - \hat{y} \rangle \geq 0 \quad \text{for all } v \in K.$$

Then $\eta \rightarrow 0^+$ yields

$$a^u(\hat{y}, v - \hat{y}) \geq \langle f(u), v - \hat{y} \rangle \quad \text{for all } v \in K,$$

which coincides with (2.1). Thus, we have $\hat{y} = y(u)$.

The strong convergence of $\{\tilde{y}_{c_n}(u)\}$ towards $y(u)$ ensues from arguments similar to those in the proof of Lemma 3.1 with $y_{c_l}(u)$ and J^u replaced by $\tilde{y}_{c_n}(u)$ and \tilde{J}^u . \square

The next theorem is an analogue to Theorem 3.5.

Theorem 4.2. *Let (A) be fulfilled, and let $\{c_n\} \subset \mathbb{R}_+$ satisfy $c_n \rightarrow \infty$ as $n \rightarrow \infty$ and Θ be a complementarity function. Then for $\bar{\lambda} := \bar{\lambda}(u) = \max\{-f(u), 0\}$ we have $\tilde{y}_{c_n}(u) \geq 0$. Moreover $\{\bar{\lambda}_{c_n}(u)\}_{n=1}^{\infty} = \{\max_{c_n} \{\bar{\lambda}(u) - c_n \tilde{y}_{c_n}(u), 0\}\}_{n=1}^{\infty}$ is uniformly bounded in $L^2(\Omega)$ and converges weakly to $\lambda(u) \in L^2(\Omega)$ satisfying*

$$A^u y(u) = f(u) + \lambda(u), \quad \text{and} \quad \Theta(\lambda(u), y(u)) = 0 \quad \text{a.e. in } \Omega.$$

Proof. From the definition of $\max_c\{x, 0\}$ we immediately obtain that $\max_c\{x, 0\} \geq \max\{x, 0\} \geq x$. Using this fact and multiplying (4.3) by $v = \max\{-\tilde{y}_{c_n}(u), 0\}$ we obtain

$$0 \leq \langle A^u \tilde{y}_{c_n}(u), v \rangle - \langle f(u) + \max\{-f(u), 0\}, v \rangle - c_n |v|_{L^2}^2 \quad (4.4)$$

which implies $\langle A^u \tilde{y}_{c_n}(u), v \rangle \geq 0$. Assume that $\Omega_{c_n}^- = \{\tilde{y}_{c_n}(u) < 0\} \neq \emptyset$. From (4.4) it follows that

$$0 \leq \langle A^u \tilde{y}_{c_n}(u), v \rangle \leq -C|v|_{H_o^1}^2 \leq 0,$$

and thus $\tilde{y}_{c_n}(u)|_{\Omega_{c_n}^-} = 0$ which is a contradiction. Hence, we have $\tilde{y}_{c_n}(u) \geq 0$ a.e. in Ω .

The definition of \max_c yields

$$0 \leq \tilde{\lambda}_{c_n}(u) \leq \max\{\bar{\lambda}(u), \frac{1}{2c_n}\}.$$

This implies that there exists a weakly convergent subsequence $\{\tilde{\lambda}_{c_{n(k)}}(u)\}$. Let $\lambda(u) \in L^2(\Omega)$ denote the weak limit point. From the previous lemma we know that $\tilde{y}_{c_n}(u)$ converges to $y(u)$ strongly in $H_o^1(\Omega)$ as $c_n \rightarrow \infty$. Thus we have

$$\langle A^u \tilde{y}_{c_{n(k)}}(u), \tilde{y}_{c_{n(k)}}(u) \rangle \rightarrow \langle f(u) + \lambda(u), y(u) \rangle.$$

The maximal monotonicity of A^u (see proof of Th. 3.5) then yields $A^u y(u) = f(u) + \lambda(u)$ implying the uniqueness of the weak limit $\lambda(u)$. For all $c_n > 0$ we have $\tilde{\lambda}_{c_n}(u) \geq 0$ by similar arguments as in the proof of Theorem 3.5. Moreover,

$$-\frac{\bar{\lambda}(u)}{c_n} - \frac{1}{2c_n^2} < -\tilde{y}_{c_n}(u) \leq 0 \quad \text{a.e. in } \{x \in \Omega | \tilde{\lambda}_{c_n}(u)(x) > 0\}.$$

The assertion $\Theta(\lambda(u), y(u)) = 0$ a.e. in Ω then ensues from the same arguments as in the proof of Theorem 3.5. \square

The complete continuity of the mapping $u \mapsto \tilde{\Phi}_c(u) = \tilde{y}_c(u)$ is established next.

Lemma 4.3. *The mapping $\tilde{\Phi}_c$ is completely continuous uniformly in $c \in \mathbb{R}^+$.*

Proof. The proof essentially follows that of Lemma 3.2. Instead of (3.2) consider now (4.3), and observe that \max_c satisfies

$$(\max_c\{\bar{\lambda} - c\tilde{y}_c(u_n), 0\} - \max_c\{\bar{\lambda} - c\tilde{y}_c(u_m), 0\}, \tilde{y}_c(u_m) - \tilde{y}_c(u_n)) \geq 0. \quad \square$$

Finally, we can guarantee that a sequence of optimal solutions of the regularized problems (\tilde{P}_{c_n}) tends (in the strong-weak sense) towards an optimal solution of the original problem (P).

Theorem 4.4. (a) *For $c > 0$ there exists an optimal solution $(\tilde{y}_c^*, \tilde{u}_c^*) \in H_o^1(\Omega) \times U$ of (\tilde{P}_c) .*

(b) *Let $\{c_n\} \subset \mathbb{R}_+$ satisfy $c_n \rightarrow \infty$ as $n \rightarrow \infty$. Then a strong-weak accumulation point $(y^*, u^*) \in H_o^1(\Omega) \times U$ of the sequence $\{(\tilde{y}_{c_n}^*, \tilde{u}_{c_n}^*)\}$ of optimal solutions to $\{(\tilde{P}_{c_n})\}$ as $n \rightarrow \infty$ exists, and every such accumulation point is a solution of (P).*

Proof. The proofs of (a) and (b) are similar to the proofs of Theorem 3.3 and Theorem 3.4, respectively. \square

4.2. First order necessary conditions

Consider the regularized equilibrium constraint of problem (\tilde{P}_c) , i.e.

$$\mathcal{E}(y, u) = A^u y - f(u) - \max_c\{\bar{\lambda} - cy, 0\}, \quad \mathcal{E} : H_o^1(\Omega) \times U \rightarrow H^{-1}(\Omega).$$

Let $E = \mathcal{E}'$ denote the Gateaux-derivative of \mathcal{E} , and recall that $f(u) = Fu + g$ (note that the analysis so far did not require the affine nature of $f(u)$). Hence, for $(d_y, d_u) \in H_o^1(\Omega) \times U$ we have

$$E(y, u)(d_y, d_u) = A^u d_y + A^{u'}(y)d_u - Fd_u + c \operatorname{sgn}_c(\bar{\lambda} - cy)d_y,$$

with $\langle A^{u'}(y)d_u, v \rangle_{U^*, U} = (e'(u)d_u \nabla y, \nabla v)$, where U^* is the dual of U . Define the bilinear form $\tilde{a}^{y,u} : H_o^1(\Omega) \times H_o^1(\Omega) \rightarrow \mathbb{R}$ by

$$\tilde{a}^{y,u}(d_y, v) = (e(u) \nabla d_y, \nabla v) + c(\operatorname{sgn}_c(\bar{\lambda} - cy)d_y, v).$$

Let F^* denote the adjoint operator of F , and define the linear functional $\tilde{f}^{y,u} : U \rightarrow H^{-1}$ by

$$\langle \tilde{f}^{y,u}(d_u), v \rangle = (e'(u)d_u \nabla y, \nabla v) - \langle F^*v, d_u \rangle$$

for all $v \in H_o^1(\Omega)$. From $\text{sgn}_c(x) \geq 0$ for all x it immediately ensues that $\tilde{a}^{y,u}(v, v) \geq C|v|_{H_o^1}^2$ for all $v \in H_o^1(\Omega)$. The boundedness of $\tilde{a}^{y,u}$ in $H_o^1(\Omega)$, i.e. $\tilde{a}^{y,u}(v, w) \leq C|v|_{H_o^1}|w|_{H_o^1}$, is a consequence of $\text{sgn}_c(x) \leq 1$ for all x . The Lax-Milgram theorem then ensures the surjectivity of E .

Now it is an easy exercise to derive first order optimality conditions characterizing an optimal solution $(\tilde{y}_c^*, \tilde{u}_c^*)$ of the regularized problem (\tilde{P}_c) . For this purpose, let B denote the representer of $|\cdot|_U$ and B^* its dual. Then we have

$$\tilde{y}_c^* - y_d + A^{\tilde{u}_c^*} \tilde{p}_c^* + c \text{sgn}_c(\bar{\lambda} - c\tilde{y}_c^*) \tilde{p}_c^* = 0, \quad (4.5)$$

$$(\alpha B^* B \tilde{u}_c^* + e'(u^*) \nabla \tilde{y}_c^* \cdot \nabla \tilde{p}_c^* - F^* \tilde{p}_c^*, u - \tilde{u}_c^*) \geq 0 \quad \text{for all } u \in U, \quad (4.6)$$

$$A^{\tilde{u}_c^*} \tilde{y}_c^* - \max_c \{ \bar{\lambda} - c\tilde{y}_c^*, 0 \} = f(\tilde{u}_c^*), \quad (4.7)$$

where $\tilde{p}_c^* \in H_o^1(\Omega)$ denotes the adjoint state.

We will use the above first order conditions as c is replaced by a sequence $\{c_n\}$ with $c_n \rightarrow \infty$ in order to derive first order necessary conditions for the bilevel control problem (P).

Theorem 4.5. *Assume that (A) is satisfied, $F \in L(U, L^\infty(\Omega))$, and let $\{c_n\} \subset \mathbb{R}_+$ satisfy $c_n \rightarrow \infty$ for $n \rightarrow \infty$. Moreover, let Θ be a complementarity function. Then for each sequence $\{(\tilde{y}_{c_n}^*, \tilde{u}_{c_n}^*)\}$ of solutions of the problems (\tilde{P}_{c_n}) converging strongly-weakly to (y^*, u^*) , a solution of (P), there exist a subsequence $\{n(k)\}$ and multipliers $(p^*, \mu^*) \in H_o^1(\Omega) \times L^\infty(\Omega)^*$ such that $\tilde{p}_{c_{n(k)}}^*$ converges to p^* weakly in $H_o^1(\Omega)$, and $\tilde{\mu}_{c_{n(k)}}^* := c_{n(k)} \text{sgn}_{c_{n(k)}}(\bar{\lambda} - c_{n(k)} \tilde{y}_{c_{n(k)}}^*) \tilde{p}_{c_{n(k)}}^*$, with $0 \leq \bar{\lambda} \in L^2(\Omega)$ suitably chosen, converges to μ^* weakly* in $L^\infty(\Omega)^*$. Moreover, p^* and μ^* satisfy the first order conditions*

$$A^{u^*} p^* + \mu^* + y^* - y_d = 0, \quad (4.8)$$

$$(\alpha B^* B u^* + e'(u^*) \nabla y^* \cdot \nabla p^* - F^* p^*, u - u^*) \geq 0 \quad \text{for all } u \in U, \quad (4.9)$$

$$p^* \lambda^* = 0, \quad \mu^* y^* = 0, \quad (4.10)$$

$$A^{u^*} y^* - \lambda^* = f(u^*), \quad (4.11)$$

$$\Theta(\lambda^*, y^*) = 0 \quad \text{a.e. in } \Omega, \quad (4.12)$$

where $\lambda^* \in L^2(\Omega)$ is the weak limit of $\{\tilde{\lambda}_{c_n}^*\}$, with $\tilde{\lambda}_{c_n}^* = \max_{c_n} \{ \bar{\lambda} - c_n \tilde{y}_{c_n}^*, 0 \}$. Moreover, we have

$$\alpha^{u^*} (p^*, \chi y^*) + (y^* - y_d, \chi y^*) = 0 \quad \text{for all } \chi \in C^1(\Omega), \quad (4.13)$$

$$(y^* - y_d, p^*) + \alpha^{u^*} (p^*, p^*) \leq 0. \quad (4.14)$$

Proof. Let $\{(\tilde{y}_{c_n}^*, \tilde{u}_{c_n}^*)\}$ denote a sequence of solutions to (\tilde{P}_{c_n}) which converges to (y^*, u^*) , a solution of (P), strongly-weakly in $H_o^1(\Omega) \times U$. Multiplying (4.5) by $\tilde{p}_{c_n}^*$ yields

$$\begin{aligned} 0 &= (\tilde{y}_{c_n}^* - y_d, \tilde{p}_{c_n}^*) + \alpha^{\tilde{u}_{c_n}^*} (\tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*) + c_n (\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*) \\ &\geq (\tilde{y}_{c_n}^* - y_d, \tilde{p}_{c_n}^*) + C_1 |\tilde{p}_{c_n}^*|_{H_o^1}^2 + c_n (\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*). \end{aligned} \quad (4.15)$$

From $\text{sgn}_{c_n} \geq 0$ we obtain

$$C_1 |\tilde{p}_{c_n}^*|_{H_o^1}^2 + c_n (\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*) \leq C, \quad (4.16)$$

where the constant C is independent of $\{c_n\}$. Next define

$$\rho_\delta(x) = \begin{cases} 1 & \text{for } x \geq \delta, \\ \frac{x}{\delta} & \text{for } |x| \leq \delta, \\ -1 & \text{for } x \leq -\delta. \end{cases}$$

From this definition it follows that

$$\begin{aligned} 0 &= (\tilde{y}_{c_n}^* - y_d, \rho_\delta(\tilde{p}_{c_n}^*)) + (\rho_\delta'(\tilde{p}_{c_n}^*) \nabla \tilde{p}_{c_n}^*, e(\tilde{u}_{c_n}^*) \nabla \tilde{p}_{c_n}^*) \\ &\quad + (c_n \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \rho_\delta(\tilde{p}_{c_n}^*)) \\ &\geq -|\tilde{y}_{c_n}^* - y_d|_{L^1} + (c_n \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \rho_\delta(\tilde{p}_{c_n}^*)) \end{aligned} \quad (4.17)$$

where inequality holds due to $|\rho_\delta(\tilde{p}_{c_n}^*)|_{L^\infty} \leq 1$ and $\rho_\delta'(x) \geq 0$ for all x . Since $\operatorname{sgn}_{c_n}(x) \leq 1$, $c_n \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \geq 0$ and $\tilde{\mu}_{c_n}^* = c_n \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*$, relation (4.17) yields for $\delta \rightarrow 0$

$$|\tilde{\mu}_{c_n}^*|_{L^1} \leq |\tilde{y}_{c_n}^* - y_d|_{L^1}, \quad (4.18)$$

where we additionally used $\rho_\delta(\tilde{p}_{c_n}^*) \tilde{p}_{c_n}^* \rightarrow |\tilde{p}_{c_n}^*|$ as $\delta \rightarrow 0$. Note that the right hand side above is uniformly bounded (with respect to $\{c_n\}$). Thus, due to (4.16) there exist a subsequence $\{n(k)\}$, $p^* \in H_o^1(\Omega)$ and $\mu^* \in L^\infty(\Omega)^*$ such that $\{\tilde{p}_{c_n(k)}^*\}$ converges to p^* weakly in $H_o^1(\Omega)$, and $\{\tilde{\mu}_{c_n(k)}^*\}$ converges to μ^* weakly* in $L^\infty(\Omega)^*$. Now, (4.8) and (4.9) immediately follow from (4.5) and (4.6), respectively. Moreover, (4.14) holds due to (4.15).

Due to the assumption $F \in L(U, L^\infty(\Omega))$ and the weak convergence of $\{\tilde{u}_{c_n}^*\}$ in U , there exists a constant M independent of $\{c_n\}$ such that $|F\tilde{u}_{c_n}^*|_{L^\infty} \leq M$. Thus, for $\bar{\lambda} \geq \max\{-g, 0\} + M$ we have $y^* \geq 0$ (see Theorem 4.2). For $\tilde{\lambda}_{c_n}^* = \max_{c_n}\{\bar{\lambda} - c_n \tilde{y}_{c_n}^*, 0\}$ it follows that $\tilde{\lambda}_{c_n}^* \leq \max\{\bar{\lambda}, (2c_n)^{-1}\}$. Hence, for a subsequence $\{n(k)\}$ we have that $\{\tilde{\lambda}_{c_n(k)}^*\}$ converges to λ^* weakly in $L^2(\Omega)$ and $A^{u^*} y^* - \lambda^* = f(u^*)$, $\Theta(\lambda^*, y^*) = 0$ a.e. in Ω . This proves (4.11) and (4.12).

In order to establish the first part of (4.10) define $S = \{x \in \Omega | 0 < \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*)(x) < 1\}$. On S we obtain

$$\begin{aligned} \tilde{\lambda}_{c_n}^* &= \max_{c_n}\{\bar{\lambda} - c_n \tilde{y}_{c_n}^*, 0\} = \int_{-\infty}^{\bar{\lambda} - c_n \tilde{y}_{c_n}^*} \operatorname{sgn}_{c_n}(t) dt = \int_{-1/(2c_n)}^{\bar{\lambda} - c_n \tilde{y}_{c_n}^*} c_n \left(t + \frac{1}{2c_n}\right) dt \\ &= \frac{c_n}{2} \left[(\bar{\lambda} - c_n \tilde{y}_{c_n}^*)^2 + \frac{1}{c_n} (\bar{\lambda} - c_n \tilde{y}_{c_n}^*) + \frac{1}{4c_n^2} \right] \\ &< \frac{1}{2c_n} c_n \left((\bar{\lambda} - c_n \tilde{y}_{c_n}^*) + \frac{1}{2c_n} \right) = \frac{1}{2c_n} \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*). \end{aligned}$$

Thus, we can infer that

$$\begin{aligned} \int_{\Omega} |\tilde{p}_{c_n}^* \tilde{\lambda}_{c_n}^*| dx &= \int_S |\tilde{p}_{c_n}^* \tilde{\lambda}_{c_n}^*| dx + \int_{\Omega \setminus S} |\tilde{p}_{c_n}^* \tilde{\lambda}_{c_n}^*| dx \\ &\leq \frac{1}{2c_n} \int_S |\tilde{p}_{c_n}^* \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*)| dx + |\tilde{p}_{c_n}^*|_{L^2(\Omega \setminus S)} |\tilde{\lambda}_{c_n}^*|_{L^2} \\ &\leq \frac{1}{2c_n^2} |\tilde{\mu}_{c_n}^*|_{L^1} + \left(\int_{\Omega} \operatorname{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^* \tilde{p}_{c_n}^* dx \right)^{1/2} |\tilde{\lambda}_{c_n}^*|_{L^2} \\ &\leq \frac{1}{2c_n^2} |\tilde{\mu}_{c_n}^*|_{L^1} + \left(\frac{C}{c_n} \right)^{1/2} |\tilde{\lambda}_{c_n}^*|_{L^2} \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

where the next to the last inequality comes from the fact that $\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) = 0$ implies $\max_{c_n} \{\bar{\lambda} - c_n \tilde{y}_{c_n}^*, 0\} = 0$ which in turn implies $\tilde{\lambda}_{c_n}^* = 0$, and $0 \leq \text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \leq 1$. The last inequality uses (4.16). From

$$\max\{|\tilde{\mu}_{c_n}^*|_{L^1}, |\tilde{\lambda}_{c_n}^*|_{L^2}, |\tilde{p}_{c_n}^* \tilde{\lambda}_{c_n}^*|_{L^1}\} \leq C,$$

with C independent of $\{c_n\}$, it follows that $\tilde{p}_{c_n}^* \tilde{\lambda}_{c_n}^*$ converges to $p^* \lambda^*$ weakly* in $L^\infty(\Omega)^*$. Thus we derive $p^* \lambda^* = 0$ a.e. in Ω .

We further have

$$-\tilde{\mu}_{c_n}^* \tilde{y}_{c_n}^* = -c_n \text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^* \tilde{y}_{c_n}^* \xrightarrow{n \rightarrow \infty} 0 \quad \text{in } L^1(\Omega). \quad (4.19)$$

In fact, we define $T = \{x \in \Omega \mid \text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*)(x) > 0\}$, and consider

$$\begin{aligned} \int_{\Omega} |\tilde{\mu}_{c_n}^* \tilde{y}_{c_n}^*| dx &= c_n \int_{\Omega} |\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^* \tilde{y}_{c_n}^*| dx \\ &\leq c_n (\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*)^{1/2} \left(\int_T \text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) |\tilde{y}_{c_n}^*|^2 dx \right)^{1/2} \\ &\leq (\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*)^{1/2} \left(\int_{\Omega} \left(\frac{1}{2c_n} + \bar{\lambda} \right)^2 dx \right)^{1/2}. \end{aligned}$$

Finally, note that due to $(\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*) \leq (C/c_n)$ (cf. (4.16)) it follows that

$$(\text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \tilde{p}_{c_n}^*)^{1/2} \left(\int_{\Omega} \left(\frac{1}{2c_n} + \bar{\lambda} \right)^2 dx \right)^{1/2} \xrightarrow{n \rightarrow \infty} 0.$$

From (4.7) we derive that $\tilde{y}_{c_n}^*$ is bounded in $H^2(\Omega)$, and since $H^2(\Omega)$ is compactly embedded in $C^0(\bar{\Omega})$, equation (4.19) yields $\mu^* y^* = 0$ a.e. in Ω , which constitutes the second part of (4.10).

Next let $\chi \in C^1(\Omega)$. Then it ensues that

$$(\tilde{y}_{c_n}^* - y_d, \chi \tilde{y}_{c_n}^*) + a^{\tilde{u}_{c_n}}(\tilde{p}_{c_n}^*, \chi \tilde{y}_{c_n}^*) + (c_n \text{sgn}_{c_n}(\bar{\lambda} - c_n \tilde{y}_{c_n}^*) \tilde{p}_{c_n}^*, \chi \tilde{y}_{c_n}^*) = 0.$$

As $c_n \rightarrow \infty$ this yields (4.13). \square

To interpret μ^* we assume that multipliers in the classical sense exist, *i.e.* we can set up the Lagrangian for (P) with (1.4) replaced by (1.5), which is

$$\mathcal{L}(y, u, \lambda, p, \nu_1, \nu_2, \nu_3) = J^u(y) + a^u(y, p) - (f(u) + \lambda, p) - (\nu_1, \lambda) - (\nu_2, y) + \nu_3(\lambda, y),$$

and consider the corresponding first order system

$$0 \leq (\nabla_u \mathcal{L}, u - u^*) = (\alpha B^* B u^* - e'(u^*) \nabla y^* \cdot \nabla p^* - F^* p^*, u - u^*) \quad \text{for all } u \in U, \quad (4.20)$$

$$0 = \nabla_y \mathcal{L} = y^* - y_d + A^{u^*} p^* - \nu_2^* + \nu_3^* \lambda^*, \quad (4.21)$$

$$0 = \nabla_\lambda \mathcal{L} = -p^* - \nu_1^* + \nu_3^* y^*, \quad (4.22)$$

$$0 = \nabla_p \mathcal{L} = A^{u^*} y^* - f(u^*) - \lambda^*, \quad (4.23)$$

$$0 = \nabla_{\nu_3} \mathcal{L} = (\lambda^*, y^*), \quad (4.24)$$

$$0 = (\lambda^*, \nu_1^*), \quad 0 \leq \lambda^*, \quad 0 \leq \nu_1^*, \quad (4.25)$$

$$0 = (y^*, \nu_2^*), \quad 0 \leq y^*, \quad 0 \leq \nu_2^*. \quad (4.25)$$

Equation (4.20) suggests to set $\mu^* = \nu_3^* \lambda^* - \nu_2^*$. Then (4.20) formally corresponds to (4.8). From the definition of μ^* and (4.21) we obtain (4.10) by pointwise multiplication with y^* and λ^* , respectively. Finally, (4.23)–(4.25) yield (4.12).

4.3. Rough coefficients

Up to now we have assumed that the parameter u satisfies $u \in \{H^2(\Omega) | u \geq \underline{\epsilon} > 0\}$. In this section we reduce the regularity requirements, *i.e.* we consider the case of

$$u \in U := \{w \in L^\infty(\Omega) \cap \mathcal{U} | 0 < \underline{\epsilon} \leq w \leq \bar{\epsilon} < +\infty\},$$

where the Hilbert space \mathcal{U} is compactly embedded in some $L^r(\Omega)$, with $r \geq 2$ sufficiently large (as specified in the proof of Lem. 4.6 below). Moreover, U is endowed with the norm $|\cdot|_U = |\cdot|_{\mathcal{U}}$. In this situation, $|A^{u_n} - A^u|_{H^{-1} \leftarrow H^1_0} \rightarrow 0$ cannot be derived from the same arguments as in the proof of Lemma 2.1. Hence, we need to establish the complete continuity of $u \mapsto y(u)$ in the present case in order to be able to apply the techniques of the previous sections. Throughout this section we shall assume that $\Omega \subset \mathbb{R}^d$, with $d \in \{1, 2\}$, is bounded with sufficiently smooth boundary, and $f(u) \in L^2(\Omega)$ for all $u \in U$. Then, for instance, the choice $\mathcal{U} = H^1(\Omega)$, $|\cdot|_{\mathcal{U}} = |\cdot|_{H^1}$, frequently arising in applications is covered by the above requirements.

First note that for every $u \in U$ by standard arguments the variational inequality

$$a^u(y, v - y) - (f(u), v - y) \geq 0 \quad \text{for all } v \in K, \quad y \in K, \tag{4.26}$$

admits a unique solution $y(u) \in K$, which satisfies $|y(u)|_{H^1_0} \leq C|f(u)|_{H^{-1}}$ for some constant C . Moreover, from [3, Chap. 1, Th. 4.2] it follows that there exists $q > 2$ such that for $f(u) \in W^{-1,q}(\Omega)$ the estimate

$$|y(u)|_{W^{1,q}} \leq C|f(u)|_{W^{-1,q}} \tag{4.27}$$

holds. Note that $L^2(\Omega)$ is continuously and densely embedded in $W^{-1,q}(\Omega)$ for $q > 2$.

Lemma 4.6. *The mapping $\Phi : U \rightarrow H^1_0(\Omega)$, $\Phi(u) = y(u)$, is completely continuous.*

Proof. Observe that due to (4.27) there exists a subsequence $\{n(k)\}$ such that $\{y(u_{n(k)})\}$ converges to \bar{y} weakly in $H^1_0(\Omega)$. Consider

$$\begin{aligned} |a^{u_{n(k)}}(y(u_{n(k)}), v) - a^u(\bar{y}, v)| &\leq |a^{u_{n(k)}}(y(u_{n(k)}), v) - a^u(y(u_{n(k)}), v)| + |a^u(\bar{y} - y(u_{n(k)}), v)| \\ &\leq |e(u_{n(k)}) - e(u)|_{L^{q'}} | \nabla y(u_{n(k)}) |_{L^q} | \nabla v |_{L^2} \\ &\quad + |a^u(\bar{y} - y(u_{n(k)}), v)|. \end{aligned} \tag{4.28}$$

$$+ |a^u(\bar{y} - y(u_{n(k)}), v)|. \tag{4.29}$$

for all $v \in H^1_0(\Omega)$. For the last inequality above we used (4.27) and Hölders inequality with $q' = 2q/(q-2)$. Since $u_{n(k)} \in U$ and thus $0 < \underline{\epsilon} \leq u_{n(k)} \leq \bar{\epsilon} < \infty$ for all k , $u_n \rightharpoonup u$ in \mathcal{U} , $y(u_{n(k)}) \rightharpoonup \bar{y}$ in $H^1_0(\Omega)$, and $e \in C^2(\mathbb{R}_+; \mathbb{R}_+)$, we have $\lim_k |a^u(\bar{y} - y(u_{n(k)}), v)| = 0$ for all $v \in H^1_0(\Omega)$. Hence, the term in (4.29) vanishes as $k \rightarrow \infty$. For (4.28) consider the fact that \mathcal{U} is compactly embedded in $L^{q'}(\Omega)$ by assumption, and that due to $u, u_{n(k)} \in U$ for all k and $e \in C^2(\mathbb{R}_+; \mathbb{R}_+)$ there exists a positive constant L such that

$$|e(u_{n(k)}) - e(u)|_{L^{q'}} \leq L|u_{n(k)} - u|_{L^{q'}} \rightarrow 0 \quad \text{for } k \rightarrow \infty.$$

Thus we obtain

$$\lim_{k \rightarrow \infty} (a^{u_{n(k)}}(y(u_{n(k)}), v) - (f(u_{n(k)}), v)) = a^u(\bar{y}, v) - (f(u), v) \tag{4.30}$$

for all $v \in H^1_o(\Omega)$. Together with (4.27) we further infer

$$\int_{\Omega} |(e(u) - e(u_{n(k)})) \nabla y(u_{n(k)}) \nabla y(u_{n(k)})| dx \leq C |u - u_{n(k)}|_{L^{q'}} |f(u_{n(k)})|_{W^{-1,q}}^2.$$

Note that due to $u_n \rightharpoonup u$ in U , the sequence $\{u_n\}$ is uniformly bounded. Therefore we have

$$\lim_{k \rightarrow \infty} \int_{\Omega} |(e(u) - e(u_{n(k)})) \nabla y(u_{n(k)}) \nabla y(u_{n(k)})| dx = 0.$$

The convexity of $v \mapsto \int_{\Omega} e(u) \nabla v \nabla v dx$ implies $\liminf_{k \rightarrow \infty} a^u(y(u_{n(k)}), y(u_{n(k)})) \geq a^u(\bar{y}, \bar{y})$. Hence, from (4.30) it follows that

$$a^u(\bar{y}, v - \bar{y}) = \int_{\Omega} e(u) \nabla \bar{y} \nabla (v - \bar{y}) dx \geq \int_{\Omega} f(u)(v - \bar{y}) dx = (f(u), v - \bar{y}) \quad \text{for all } v \in K.$$

Due to the uniqueness of the solution of the variational inequality (4.26) we have $\bar{y} = y(u)$. Since $u \in U$ we infer

$$\int_{\Omega} e(u) \nabla (y(u_{n(k)}) - \bar{y}) \nabla (y(u_{n(k)}) - \bar{y}) dx \geq C |y(u_{n(k)}) - \bar{y}|_{H^1_o}^2 \geq 0.$$

The fact that $y(u_{n(k)})$ is the unique solution of (4.26) with u replaced by $u_{n(k)}$ implies

$$\begin{aligned} a^u(y(u_{n(k)}) - \bar{y}, y(u_{n(k)}) - \bar{y}) &\leq a^u(y(u_{n(k)}), y(u_{n(k)})) - 2a^u(y(u_{n(k)}), \bar{y}) + a^u(\bar{y}, \bar{y}) \\ &\quad + a^{u_{n(k)}}(y(u_{n(k)}), v - y(u_{n(k)})) - (f(u_{n(k)}), v - y(u_{n(k)})) \\ &= a^{u_{n(k)}}(y(u_{n(k)}), v) - (f(u_{n(k)}), v - y(u_{n(k)})) + a^u(\bar{y}, \bar{y}) \\ &\quad + a^u(y(u_{n(k)}), y(u_{n(k)})) - a^{u_{n(k)}}(y(u_{n(k)}), y(u_{n(k)})) - 2a^u(y(u_{n(k)}), \bar{y}) \end{aligned}$$

for all $v \in K$. Hence, we obtain

$$\begin{aligned} \liminf_{k \rightarrow \infty} a^u(y(u_{n(k)}) - \bar{y}, y(u_{n(k)}) - \bar{y}) &\leq \limsup_{k \rightarrow \infty} a^u(y(u_{n(k)}) - \bar{y}, y(u_{n(k)}) - \bar{y}) \\ &\leq a^u(y(u), v - y(u)) - (f(u), v - y(u)). \end{aligned}$$

For $v = y(u) \in K$ the previous computations result in

$$\lim_{k \rightarrow \infty} |y(u_{n(k)}) - y(u)|_{H^1_o} = 0. \quad \square$$

The proof of $\mu^* y^* = 0$ a.e. in Ω (see Th. 4.5) is now based on the fact that $L^2(\Omega)$ is continuously embedded in $W^{-1,q}(\Omega)$, with $q > 2$, and hence $\{\tilde{y}_{c_n}^*\}$ is uniformly bounded in $W^{1,q}_o(\Omega)$. Since $W^{1,q}_o(\Omega)$ is compactly embedded in $C^o(\bar{\Omega})$, and thus $\{\tilde{y}_{c_n}^*\}$ converges to y^* strongly in $C^o(\bar{\Omega})$, the claim $\mu^* y^* = 0$ a.e. in Ω follows essentially from the same arguments as in the proof of Theorem 4.5.

Note that for $\mathcal{U} = H^1(\Omega)$, $|\cdot|_{\mathcal{U}} = |\cdot|_{H^1}$, in the case of $d = 1$ the Sobolev space $H^1(\Omega)$ is compactly embedded in $C^{o,\nu}(\bar{\Omega})$, with $0 < \nu < 1/2$, and thus Lemma 2.1 can be applied instead of Lemma 4.6.

5. NUMERICAL REALIZATION

We shall now discuss some issues concerning the discretization of the first order system (4.8)–(4.12), and suggest an algorithm for solving the discretized system. Finally, a brief report on numerical results will end the section.

5.1. Discretization

For the discretization of the function spaces U , $H_o^1(\Omega)$, $L^2(\Omega)$ and $L^\infty(\Omega)^*$ we use finite dimensional subspaces U_h , W_h , L_h and M_h , respectively. Let $u_i \geq 0$, $i = 1, \dots, N_{U_h}$, w_i , $i = 1, \dots, N_{W_h}$, l_i , $i = 1, \dots, N_{L_h}$ and m_i , $i = 1, \dots, N_{M_h}$, denote the linearly independent basis functions of the respective finite dimensional subspace. Then, in terms of the basis functions, we obtain the following finite dimensional approximations of $u \in U$, $y, p \in H_o^1(\Omega)$, $\lambda \in L^2(\Omega)$, $\mu \in L^\infty(\Omega)^*$:

$$\begin{aligned} u_h(x) &= \sum_{i=1}^{N_{U_h}} U_i u_i(x), & y_h(x) &= \sum_{i=1}^{N_{W_h}} Y_i w_i(x), & p_h(x) &= \sum_{i=1}^{N_{W_h}} P_i w_i(x), \\ \lambda_h(x) &= \sum_{i=1}^{N_{L_h}} \Lambda_i l_i(x), & \mu_h(x) &= \sum_{i=1}^{N_{M_h}} M_i m_i(x). \end{aligned}$$

Let $\underline{U} = (U_1, \dots, U_{N_{U_h}})^T \in \mathbb{R}^{N_{U_h}}$, and analogously for \underline{Y} , \underline{P} , $\underline{\Lambda}$ and \underline{M} . Next define $A_h(\underline{U}) \in \mathbb{R}^{N_{W_h} \times N_{W_h}}$ and $C_h(\underline{Y}) \in \mathbb{R}^{N_{W_h} \times N_{U_h}}$ by

$$\begin{aligned} \underline{W}^T A_h(\underline{U}) \underline{Y} &= (e_h(u_h) \nabla y_h, \nabla w_h) \quad \text{for all } w_h \in W_h, \\ \underline{V}^T C_h(\underline{Y})^T \underline{P} &= (e'_h(u_h) v_h, \nabla y_h \cdot \nabla p_h) \quad \text{for all } v_h \in U_h, \end{aligned}$$

where $e_h(v_h)(x) = \sum_{i=1}^{N_{U_h}} e(V_i) u_i(x)$ and $e'_h(v_h)(x) = \sum_{i=1}^{N_{U_h}} e'(V_i) u_i(x)$ for all $v_h \in U_h$. Note that $\underline{V}^T C_h(\underline{Y})^T \underline{P} = \underline{V}^T C_h(\underline{P})^T \underline{Y}$. Moreover, we define $B_h \in \mathbb{R}^{N_{U_h} \times N_{U_h}}$ by $\underline{U}^T B_h \underline{Q} = (u_h, q_h)_U$ for $u_h, q_h \in U_h$, with $(\cdot, \cdot)_U$ denoting the inner product of U , and $F_h \in \mathbb{R}^{N_{W_h} \times N_{U_h}}$ by $\underline{Q}^T F_h^T \underline{P} = (F^* p_h, q_h)$. Further we use the following notation for mass matrices: $M_{W_h}^{U_h} \in \mathbb{R}^{N_{W_h} \times N_{W_h}}$ is defined by $(M_{W_h}^{U_h})_{i,j} = (u_i, w_j)$.

With these definitions the discretizations of (4.8), (4.9) and (4.11) become

$$A_h(\underline{U}^*) \underline{P}^* + M_{M_h}^{W_h} \underline{M}^* + M_{W_h}^{W_h} \underline{Y}^* - M_{L_h}^{W_h} \underline{Y}_d = 0_{N_{W_h}}, \quad (5.1)$$

$$(\alpha B_h \underline{U}^* + C_h(\underline{Y}^*)^T \underline{P}^* - F_h^T \underline{P}^*)^T (\underline{U} - \underline{U}^*) \geq 0 \quad \text{for all } \underline{U} \in \mathcal{V}, \quad (5.2)$$

$$A_h(\underline{U}^*) \underline{Y}^* - M_{L_h}^{W_h} \underline{\Lambda}^* - F_h \underline{U}^* - M_{L_h}^{W_h} \underline{G} = 0_{N_{W_h}}, \quad (5.3)$$

with $\mathcal{V} = \{\underline{U} \in \mathbb{R}^{N_{U_h}} | u_h \in U_h\}$ the set of feasible coefficient vectors. The pointwise almost everywhere conditions (4.10) and (4.12) are enforced at the nodal points. For this purpose we assume that $N = N_{W_h} = N_{L_h} = N_{M_h}$, and $w_i \geq 0$, $l_i \geq 0$ for all $i = 1, \dots, N$. The complementarity condition (4.12) is then defined in a (vector) componentwise sense, *i.e.* $\Theta_N : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^N$,

$$\Theta_N(\underline{\Lambda}, \underline{Y}) = (\phi(\Lambda_1, Y_1), \dots, \phi(\Lambda_N, Y_N))^T,$$

with $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ denoting a complementarity function. One option for ϕ is the Fischer-Burmeister function $\phi(a, b) = \phi_{FB}(a, b) = \sqrt{a^2 + b^2} - (a + b)$; see [13]. Let \bullet denote the componentwise product of vectors. Then the discretization of (4.10) and (4.12) becomes

$$\underline{P}^* \bullet \underline{\Lambda}^* = 0_N, \quad (5.4)$$

$$\underline{M}^* \bullet \underline{Y}^* = 0_N, \quad (5.5)$$

$$\Theta_N(\underline{\Lambda}^*, \underline{Y}^*) = 0_N. \quad (5.6)$$

An alternative representation of the discretized optimality system is obtained by using $\underline{M}^* = \underline{R}^* \bullet \underline{\Lambda}^*$ with $\underline{R}^* \in \mathbb{R}^N$. If strict complementarity is satisfied, *i.e.* $Y_i^* = 0$ implies $\Lambda_i^* > 0$, then (5.1)–(5.6) are equivalent to

$$\begin{aligned} A_h(\underline{U}^*)\underline{P}^* + M_{M_h}^{W_h}(\underline{R}^* \bullet \underline{\Lambda}^*) + M_{W_h}^{W_h}\underline{Y}^* - M_{L_h}^{W_h}\underline{Y}_d &= 0_{N_{W_h}}, \\ (\alpha B_h \underline{U}^* + C_h(\underline{Y}^*)^T \underline{P}^* - F_h^T \underline{P}^*)^T (\underline{U} - \underline{U}^*) &\geq 0 \quad \text{for all } \underline{U} \in \mathcal{V}, \\ A_h(\underline{U}^*)\underline{Y}^* - M_{L_h}^{W_h} \underline{\Lambda}^* - F_h \underline{U}^* - M_{L_h}^{W_h} \underline{G} &= 0_{N_{W_h}}, \\ \Theta_N(\underline{\Lambda}^*, \underline{Y}^*) &= 0_N, \\ \underline{P}^* - \underline{R}^* \bullet \underline{Y}^* &= 0_N. \end{aligned} \tag{5.7}$$

Note that the number of equations is reduced, since (5.7) replaces (5.4) and (5.5). However, we prefer to use (5.1)–(5.6) instead of the above smaller system due to the fact that the latter system requires strict complementarity to be satisfied. Although this might be the case at an optimal solution of the discrete problem to (P), during the iterations of the algorithm, which will be used, the iterates may (nearly) lack strict complementarity, and hence (5.7) causes numerical instabilities.

5.2. Algorithmic issues

Before we shall discuss the algorithm, another important property already briefly alluded to in the introductory section is addressed. Due to the inverse nature of (1.1), without further assumptions the parameter u cannot be identified on the singular set $S_o = \{x \in \Omega \mid \nabla y(x) = 0\}$; see [17] for a discussion on this issue in the case of variational equalities. On the other hand, as it is seen from its definition $A_h(\underline{U}^*)$ depends on $\underline{U}^*|_{S_{o,h}}$, with $S_{o,h}$ the discrete counterpart of S_o . In order to cope with this difficulty, from now on we assume that a (sufficiently good) approximation of $u|_S$, with $S_o \subset S$, is available. Let $\underline{U}_a^* \in \mathbb{R}^{|S_h|}$, with $S_{o,h} \subset S_h$, denote the discretized available information on the parameter, and $\underline{U}^n \in \mathcal{V}$ an iterate of the subsequent algorithm. Then we fix \underline{U}_a^* in \underline{U}^n , *i.e.* $\underline{U}_{|S_h}^n = \underline{U}_a^*$, in all iterations. Hence, compared to the original finite dimensional system (5.1)–(5.6) the number of unknowns is reduced by $|S_h| \leq N$.

In order to incorporate the inequality condition (5.2), the iterates of the subsequent algorithm satisfy $\underline{U}^n \in \text{int}\mathcal{V}$, the (strict) interior of \mathcal{V} , for all iterations n . Hence, (5.2) becomes

$$\alpha B_h \underline{U} + C_h(\underline{Y})^T \underline{P} - F_h^T \underline{P} = 0_{N_{U_h}}. \tag{5.2'}$$

during the iteration. Due to the fact that (5.1),(5.2'), (5.3)–(5.6) is an overdetermined system, *i.e.* $5N + N_{U_h}$ equations have to be satisfied by $4N + N_{U_h} - |S_h|$ unknowns, the computation of a solution is realized by a least squares technique. For this purpose, we assume that the complementarity function ϕ is chosen such that

$$\Psi_N(\underline{\Lambda}, \underline{Y}) = \Theta_N(\underline{\Lambda}, \underline{Y})^T \Theta_N(\underline{\Lambda}, \underline{Y}) \in C^1(\mathbb{R}^{2N}; \mathbb{R}). \tag{B}$$

The Fischer-Burmeister function introduced in Section 5.1 fulfills (B). For a corresponding proof we refer to [13]. The Moreau-Yosida based function $\phi_{MY}(a, b) = a - \max\{a - cb, 0\}$ with $c > 0$, on the other hand, does not satisfy (B).

Let $\underline{X} = (\underline{Y}, \hat{\underline{U}}, \underline{P}, \underline{\Lambda}, \underline{M}) \in \mathbb{R}^{N_v}$, with $N_v = 4N + N_{U_h} - |S_h|$ and $\hat{\underline{U}}$ the unknown part of \underline{U} , and let $f_i(\underline{X})$, with $i = 1, \dots, 6$, denote the left hand sides of equations (5.1), (5.2'), (5.3)–(5.6) at \underline{X} with $\underline{U}_{|S_h}^* = \underline{U}_a^*$ fixed, and define

$$\mathcal{F}(\underline{X}) = \frac{1}{2} \sum_{i=1}^6 f_i(\underline{X})^T f_i(\underline{X}). \tag{5.8}$$

Note that \mathcal{F} is continuously differentiable, and for $\phi = \phi_{FB}$ we have

$$\nabla_{(\underline{\Lambda}, \underline{Y})} (f_6(\underline{X})^T f_6(\underline{X})) = \nabla_{(\underline{\Lambda}, \underline{Y})} \Psi_N(\underline{\Lambda}, \underline{Y}) = 2\partial\Theta_N(\underline{\Lambda}, \underline{Y})^T \Theta_N(\underline{\Lambda}, \underline{Y}),$$

with $\partial\Theta_N(\underline{\Lambda}, \underline{Y})$ denoting the generalized Jacobian of Θ_N in Clarke's sense [11], which exists by convexity of ϕ . Clearly, we wish to compute \underline{X}^* such that $\mathcal{F}(\underline{X}^*) = 0$. Hence,

$$\underline{X}^* = \operatorname{argmin}\{\mathcal{F}(\underline{X}) \mid \underline{X} = (\underline{Y}, \underline{U}, \underline{P}, \underline{\Lambda}, \underline{M}) \in \mathbb{R}^{N_v}, \underline{U} \in \mathcal{V}\}.$$

Since we expect that (5.8) results in either a zero residual or a small residual problem, a Gauss-Newton-type method [12] for the iterative solution of the discretized first order conditions is applied.

Let us address some details of our implementation: We use the forcing function $\rho : \mathbb{R} \rightarrow \mathbb{R}_+$, with $\rho(z) = z^3$ for $z \geq \epsilon > 0$, where $\epsilon \ll 1$, and $\rho(z) = 0$ otherwise, to modify the Gauss-Newton iteration matrix $J(\underline{X}^n)^T J(\underline{X}^n)$ by adding $\rho(\mathcal{F}(\underline{X}^n))I$, with $I \in \mathbb{R}^{N_v \times N_v}$ the unit matrix. Here $J(\underline{X}) \in \mathbb{R}^{N_e \times N_v}$ denotes the Jacobian of $(f_1, f_2, f_3, f_4, f_5, f_6)^T(\underline{X}) \in \mathbb{R}^{N_e}$, with $N_e = 5N + N_{U_h}$. This technique enhances robustness of the Gauss-Newton method and corresponds to a Levenberg-Marquardt-type approach. Globalization, *i.e.* damping of the full stabilized Gauss-Newton method for convergence from an arbitrary starting point, is achieved by a Wolfe type line search procedure based on quadratic interpolation. Whenever the stabilized Gauss-Newton direction \underline{D}_{sGN}^n is not sufficiently descent, *i.e.*

$$\nabla\mathcal{F}(\underline{X}^n)^T \underline{D}_{sGN}^n > -\nu |\underline{D}_{sGN}^n|_{l^2(\mathbb{R}^{N_v})} |\nabla\mathcal{F}(\underline{X}^n)|_{l^2(\mathbb{R}^{N_v})},$$

with $\nu > 0$ small, then the steepest descent direction $\underline{D}_S^n = -\nabla\mathcal{F}(\underline{X}^n)$ is used with a safeguarded Armijo type line search such that strict feasibility of \underline{U}^{n+1} is guaranteed. In order to maintain the requirement $\underline{U}^n \in \operatorname{int}\mathcal{V}$ we choose the initial data such that $\underline{U}^o \in \operatorname{int}\mathcal{V}$, and restrict, if necessary, the length of the search direction for \underline{U} such that strict feasibility for the full step is conserved. This is done by means of a trust region technique.

5.3. Results

Subsequently we report on some results of our test runs. In all examples listed below the domain was chosen to be $\Omega = (0, 1)^2$. For our finite element discretization we use a *classical* triangulation with mesh size $h = 2^{-5}$ and piecewise linear elements. Within this feasibility study we restrict ourselves to this simple discretization, although the different regularity properties of the variables suggest to use different elements. We believe that this issue requires further investigation which is not the focus of the present research. Typically in our tests, we first fix the control $u = \bar{u}$ and solve the forward problem, *i.e.* the variational inequality resulting from (1.4), on a fine grid. Afterwards we use the restriction (to the coarser grid) of the solution (state) of the forward problem, y_{h,V_I}^* , as desired state y_{dh} and try to recover u_h on the observation part $\hat{\Omega} = \Omega \setminus S$ of Ω . On S we fix $u_{h|S} = \bar{u}_{h|S}$. We shall also report on test runs where we impose noise on y_{dh} .

Let us point out some of the properties of the examples considered below. From the solutions to the discretized problems – see Figures 1–5 – one can conclude that the example in Section 5.3.1 safely satisfies the strict complementarity condition, thus allowing to apply the classical (Lagrangian) theory for deriving a first order optimality condition. The examples in Section 5.3.2–5.3.3 exhibit a certain degree of degeneracy, *i.e.*, typically y_h is close to zero on the part of the inactive set $\mathcal{I}_h = \{x \in \Omega \mid y_h(x) > 0\}$ close to the active (or coincidence) set $\mathcal{A}_h = \Omega \setminus \mathcal{I}_h$. Although the classical theory could be applied, typically numerical algorithms based on the classical first order system exhibit difficulties in locating \mathcal{I}_h and \mathcal{A}_h . The final example in Section 5.3.4 is constructed in a way that the classical theory fails.

5.3.1. Example

We consider the variational inequality arising from the Reynolds lubrication equation; see for instance [2, 9, 10, 15, 16] for details. The control u has the meaning of the height of the gap between two rotating surfaces, and the state y corresponds to the pressure in the lubricant, which fills the gap. Here

$$e(z) = z^3 \quad \text{and} \quad f(u) = -\frac{\partial u}{\partial x_2}.$$

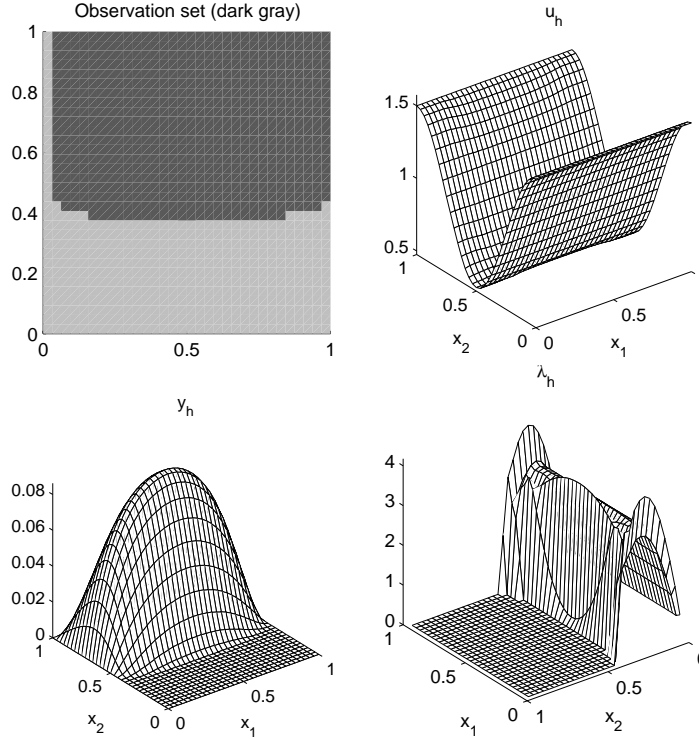
FIGURE 1. Observation part $\hat{\Omega}_h$ and optimal primal variables for Example 5.3.1.

TABLE 1. Results for Example 5.3.1.

#it	\mathcal{F}^*	#feval	$\text{rel}(\underline{Y}^*)$	$\text{rel}(\underline{U}^*)$	$\text{rel}(\underline{\Lambda}^*)$
35	8.39E-8	293	0.142	0.074	0.634

We use $\bar{u} = 1 + 0.5 \cos(2\pi x_2)$, and the regularization parameter $\alpha = 10^{-3}$. The graph in the upper left corner of Figure 1 displays the observation part $\hat{\Omega}_h$. For the inverse coefficient problem the following initial values were used:

$$\begin{aligned} \underline{Y}^o &= \underline{Y}_d, \quad \hat{\underline{U}}_{|\hat{\Omega}_h}^o \equiv 2, \quad \underline{M}_{|\Omega_{\underline{Y}_d}^o} \equiv 0.1, \quad \underline{M}_{|\Omega_{\underline{Y}_d}^+} \equiv 0, \\ \underline{\Delta}_{|\Omega_{\underline{Y}_d}^o} &\equiv 1.5, \quad \underline{\Delta}_{|\Omega_{\underline{Y}_d}^+} \equiv 0, \quad \underline{P}_{|\Omega_{\underline{Y}_d}^o} \equiv 0, \quad \underline{P}_{|\Omega_{\underline{Y}_d}^+} \equiv 0.001, \end{aligned}$$

where $\Omega_{\underline{Y}_d}^o = \{i | (\underline{Y}_d)_i \leq 0\}$ and $\Omega_{\underline{Y}_d}^+ = \{i | (\underline{Y}_d)_i > 0\}$.

In Table 1, #it denotes the number of iterations, \mathcal{F}^* is the function value at termination, #feval is the number of function evaluations and

$$\text{rel}(\underline{Y}^*) = \frac{|\underline{Y}^* - \underline{Y}_{VI}^*|_\infty}{|\underline{Y}_{VI}^*|_\infty}, \quad \text{rel}(\underline{U}^*) = \frac{|\underline{U}^* - \bar{\underline{U}}|_\infty}{|\bar{\underline{U}}|_\infty}, \quad \text{rel}(\underline{\Lambda}^*) = \frac{|\underline{\Lambda}^* - \underline{\Lambda}_{VI}^*|_\infty}{|\underline{\Lambda}_{VI}^*|_\infty}$$

denote relative errors, where $(\underline{Y}^*, \underline{U}^*, \underline{\Lambda}^*)$ are the values at the termination of the algorithm. Moreover, $\underline{\Lambda}_{VI}^*$ represents the restricted multiplier obtained from solving the forward problem on a fine grid.

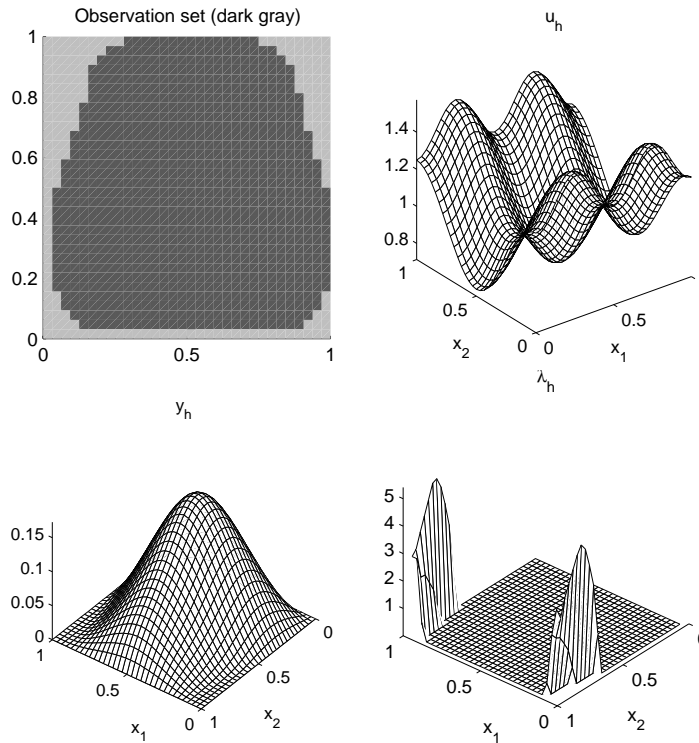


FIGURE 2. Observation part $\hat{\Omega}_h$ and optimal primal variables for Example 5.3.2.

TABLE 2. Results for Example 5.3.2.

η	#it	\mathcal{F}^*	#feval	$\text{rel}(\underline{Y}^*)$	$\text{rel}(\underline{U}^*)$	$\text{rel}(\underline{\Lambda}^*)$
η_{00}	17	5.00E-10	42	0.056	0.168	0.148
η_{01}	43	5.00E-10	200	0.073	0.200	0.128
η_{10}	44	7.07E-10	385	0.061	0.157	0.618

The final iterations primarily reduce the relative gradient norm to the accuracy demanded by the stopping rule, while the function values are only slightly decreasing. All iterations accept the stabilized Gauss-Newton direction, *i.e.* $\underline{U}^n \in \text{int}\mathcal{V}$ without invoking the trust region modification.

5.3.2. Example

For the forward problem we use $\bar{u} = 0.25(\sin^2(2\pi x_1) + \cos(2\pi x_2)) + 1$. The forcing term is $f(u) = u - \pi \cos(2\pi x_1) + 0.5\pi \sin(2\pi x_2)$, and $e(z) = (z + 0.01)^2$. The regularization parameter is chosen to be $\alpha = 10^{-4}$. The same starting values like in Example 5.3.1 were chosen. Figure 2 displays the observation part $\hat{\Omega}_h$ and the optimal primal variables. In Table 2 the quantities η_{10} and η_{01} denote uniformly distributed random noise in $[0; 0.01]$ and $[0; 0.001]$, respectively, and η_{00} represents the noise-free case.

The number of iterations, the number of function evaluations and the average relative error, *i.e.* $(\text{rel}(\underline{Y}^*) + \text{rel}(\underline{U}^*) + \text{rel}(\underline{\Lambda}^*))/3$, increase with increasing noise level. This dependence on the noise level is typical within a range of test examples.

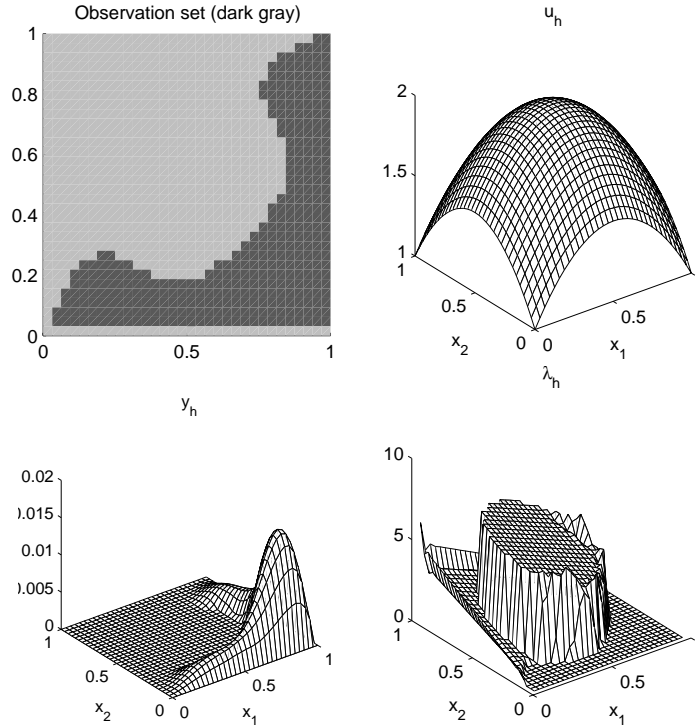


FIGURE 3. Observation part $\hat{\Omega}_h$ and optimal primal variables for Example 5.3.3.

TABLE 3. Results for Example 5.3.3.

η	#it	\mathcal{F}^*	#feval	rel(\underline{Y}^*)	rel(\underline{U}^*)	rel($\underline{\Lambda}^*$)
η_{00}	38	1.78E-9	244	0.149	0.044	0.536
η_{01}	40	1.80E-9	318	0.789	0.127	0.536
η_{10}	167	3.70E-9	1048	0.825	0.170	0.425

5.3.3. Example

We use $\bar{u} = -2((x_1 - 0.5)^2 + (x_2 - 0.5)^2) + 2$ in the forward problem. The forcing term is chosen to be $f(u) = \frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} + g$ with $g \equiv -5$ on Ω^+ and $g \equiv 0$ on $\Omega \setminus \Omega^+$, where $\Omega^+ = \{x \in \Omega | \bar{u} \geq 1.75\}$. We further have $e(z) = z$ and $\alpha = 10^{-4}$. Again, we use the same start-up values as for Example 5.3.1. Figure 3 contains the observation part and the optimal primal variables. In Table 3, η_{01} and η_{10} denote uniformly distributed random noise in $[0; 0.0001]$ and $[0; 0.001]$, respectively.

Concerning the number of iterations, the number of function evaluations and the average relative errors the same conclusions as for Example 5.3.2 can be drawn. The significant change for η_{01} and η_{10} can be explained by the fact that y_h^* is less than $2E-3$ on large parts of $\hat{\Omega}_h$, and thus the relative noise level in y_{dh} is large. Therefore, u_h has to be recovered from very poor data.

5.3.4. Example

The data are as follows: $\bar{u} = 1 + x_1x_2$, $F \equiv 0$, and $g = A(\bar{u})y(\bar{u}) - \lambda(\bar{u})$ with

$$\begin{aligned}
 y(\bar{u}) &= 100 (\max\{0, x_1 - 0.4\}^2 \max\{0, 0.6 - x_1\}^2 \max\{0, x_2 - 0.25\}^2 \max\{0, 0.75 - x_2\}^2), \\
 \lambda(\bar{u}) &= \max\{0, 0.25 - x_1\}^2 + \max\{0, x_1 - 0.75\}^2.
 \end{aligned}$$

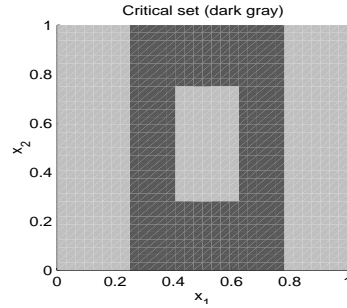


FIGURE 4. Critical set (dark gray) of lack of strict complementarity.

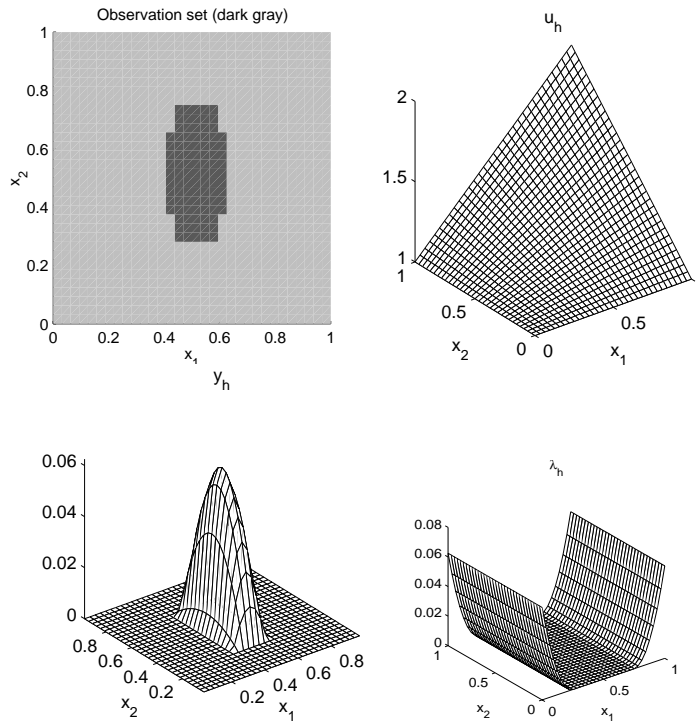


FIGURE 5. Observation part $\hat{\Omega}_h$ and optimal primal variables for Example 5.3.4.

Note that $y(\bar{u})$ with corresponding multiplier $\lambda(\bar{u})$ is the optimal solution of the lower level problem for $u = \bar{u}$. The dark gray region in Figure 4 corresponds to the discrete analogue of the set $\{x \in \Omega | y(\bar{u})(x) = 0 \text{ and } \lambda(\bar{u})(x) = 0\}$, *i.e.* the set of lack of strict complementarity. From this figure we can see that the classical theory cannot be applied. In Figure 5 the observation set and the optimal primal solution are shown. The

results in Table 4 of a run of our algorithm show that the new first order characterization is effective, and that the algorithm is *not* affected by lack of strict complementarity.

TABLE 4. Results for Example 5.3.4.

#it	\mathcal{F}^*	#feval	$\text{rel}(\underline{Y}^*)$	$\text{rel}(\underline{U}^*)$	$\text{rel}(\underline{\Lambda}^*)$
20	7.98E-9	74	1.3E-5	2.0E-4	0.13

The following starting values were used:

$$\begin{aligned} \underline{Y}^o &= \underline{Y}_d, \quad \hat{U}_{|\Omega_h}^o \equiv 1.5, \quad \underline{M}_{|\Omega_{\underline{Y}_d}^o} \equiv 1.0, \quad \underline{M}_{|\Omega_{\underline{Y}_d}^+} \equiv 0, \\ \underline{\Lambda}_{|\Omega_{\underline{Y}_d}^o} &\equiv 0.5, \quad \underline{\Lambda}_{|\Omega_{\underline{Y}_d}^+} \equiv 0, \quad \underline{P}_{|\Omega_{\underline{Y}_d}^o} \equiv 0, \quad \underline{P}_{|\Omega_{\underline{Y}_d}^+} \equiv 0.1, \end{aligned}$$

REFERENCES

- [1] V. Barbu, *Optimal Control of Variational Inequalities. Res. Notes Math.*, Pitman, **100** (1984).
- [2] B. Bayada and M. El Aalaoui Talibi, *Control by the coefficients in a variational inequality: the inverse elastohydrodynamic lubrication problem*. Report no. 173, I.N.S.A. Lyon (1994).
- [3] A. Bensoussan, J.L. Lions and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam (1978).
- [4] M. Bergounioux, Optimal control problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim.* **36** (1998) 273–289.
- [5] M. Bergounioux and H. Dietrich, Optimal control problems governed by obstacle type variational inequalities: a dual regularization penalization approach. *J. Convex Anal.* **5** (1998) 329–351.
- [6] M. Bergounioux, K. Ito and K. Kunisch, Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **37** (1999) 1176–1194.
- [7] M. Bergounioux and F. Mignot, Optimal control of obstacle problems: existence of Lagrange multipliers. *ESAIM: COCV* **5** (2000) 45–70.
- [8] A. Bermudez and C. Saguez, Optimality conditions for optimal control problems of variational inequalities, in: Control problems for systems described by partial differential equations and applications. I. Lasiecka and R. Triggiani Eds., *Lect. Notes Control and Information Sciences*, Springer, Berlin (1987).
- [9] G. Capriz and G. Cimatti, Free boundary problems in the theory of hydrodynamic lubrication: a survey, in: Free Boundary Problems: Theory and Applications, Vol. II, A. Fasano and M. Primicerio Eds., *Res. Notes Math.*, Pitman, **79** (1983).
- [10] G. Cimatti, On a problem of the theory of lubrication governed by a variational inequality. *Appl. Math. Optim.* **3** (1977) 227–242.
- [11] F. Clarke, *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York (1983).
- [12] J. Dennis and R. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Series in Computational Mathematics, Prentice-Hall, Englewood Cliffs, New Jersey (1983).
- [13] F. Facchinei, A. Fischer and C. Kanzow, A semismooth Newton method for variational inequalities: the case of box constraints, in *Complementarity and Variational Problems, State of the Art*, M. Ferris and J. Pang Eds., SIAM, Philadelphia (1997).
- [14] F. Facchinei, H. Jiang and L. Qi, A smoothing method for mathematical programs with equilibrium constraints. *Math. Prog.* **85** (1999) 107–134.
- [15] J. Guo, *A variational inequality associated with a lubrication problem*, IMA Preprint Series, no. 530 (1989).
- [16] B. Hu, A quasi-variational inequality arising in elastohydrodynamics. *SIAM J. Math. Anal.* **21** (1990) 18–36.
- [17] K. Ito and K. Kunisch, On the injectivity and linearization of the coefficient-to-solution mapping for elliptic boundary value problems. *J. Math. Anal. Appl.* **188** (1994) 1040–1066.
- [18] K. Ito and K. Kunisch, Optimal control of elliptic variational inequalities. *Appl. Math. Optim.* **41** (2000) 343–364.
- [19] W. Liu and J. Rubio, Optimality conditions for strongly monotone variational inequalities. *Appl. Math. Optim.* **27** (1993) 291–312.
- [20] Z. Luo, J. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, New York (1996).
- [21] Z. Luo and P. Tseng, A new class of merit functions for the nonlinear complementarity problem, in *Complementarity and Variational Problems, State of the Art*, M. Ferris and J. Pang Eds., SIAM, Philadelphia (1997).
- [22] F. Mignot and J.P. Puel, Optimal control in some variational inequalities. *SIAM J. Control Optim.* **22** (1984) 466–476.