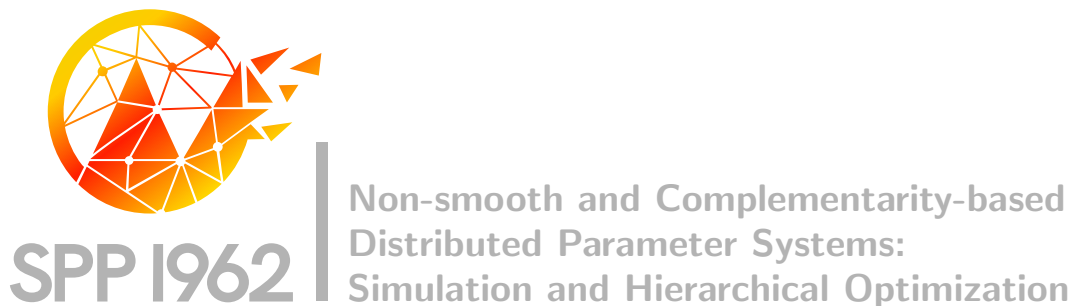


**DFG** Deutsche  
Forschungsgemeinschaft  
Priority Programme 1962

*M-stationarity for a class of MPCCs in Lebesgue  
spaces*

Felix Harder, Gerd Wachsmuth



Preprint Number SPP1962-180

received on October 29, 2021

Edited by  
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)  
Leibniz Institute in the Forschungsverbund Berlin e.V.  
Mohrenstraße 39, 10117 Berlin, Germany  
E-Mail: [spp1962@wias-berlin.de](mailto:spp1962@wias-berlin.de)

World Wide Web: <http://spp1962.wias-berlin.de/>

# M-stationarity for a class of MPCCs in Lebesgue spaces

Felix Harder      Gerd Wachsmuth

October 21, 2021

Chair of Optimal Control, BTU Cottbus-Senftenberg

## Abstract

We show that an optimality condition of M-stationarity type holds for minimizers of a class of mathematical programs with complementarity constraints (MPCCs) in Lebesgue spaces. We apply these results also to local minimizers of an inverse optimal control problem (which is an instance of an infinite-dimensional bilevel optimization problem). The multipliers for the M-stationarity system can be constructed via convex combinations of various multipliers to auxiliary, linear problems. However, proving the existence of the multipliers to these auxiliary problems is difficult and only possible in some situations.

**Keywords:** Mathematical program with complementarity constraints, Necessary optimality conditions, M-stationarity, inverse optimal control

**MSC (2020):** [90C33](#), [49K20](#), [49K21](#)

## 1 Introduction

In this article, we consider so-called mathematical programs with complementarity constraints (MPCCs for short) in Lebesgue spaces, see [\(MPCC\)](#) for such an optimization problem. For this type of problems, it is characteristic that they contain a complementarity constraint, i.e. a constraint of the form

$$0 \leq G(x) \quad \wedge \quad G(x)H(x) = 0 \quad \wedge \quad H(x) \geq 0 \quad \text{a.e. on } \Omega,$$

where  $G : X \rightarrow L^p(\Omega)$ ,  $H : X \rightarrow L^q(\Omega)$  are functions,  $X$  is a Banach space and  $L^p(\Omega)$ ,  $L^q(\Omega)$  are Lebesgue spaces over some domain  $\Omega$ . Other (non-complementarity) constraints can involve Banach spaces which are not necessarily Lebesgue spaces.

We are interested in first-order necessary stationarity conditions for such a class of optimization problems.

In finite-dimensional spaces, MPCCs have been studied frequently, both from a numerical and theoretical perspective. Even in finite-dimensional spaces, MPCCs form a challenging class of optimization problems. Several problem-tailored stationarity conditions have been developed. One such stationarity condition is the so-called strong stationarity (which is equivalent to the KKT conditions). However, strong stationarity does not need to hold for local minimizers even if the data is linear. A stationarity condition which holds under the very weak constraint qualification MPCC-GCQ (which, e.g., holds if the data is linear) is M-stationarity, see [Flegel, Kanzow, Outrata, 2007, Theorem 14]. Therein, this result was proved by using the so-called limiting normal cone and its calculus rules.

Recently, a simpler proof of M-stationarity for finite-dimensional MPCCs has been found, see [Harder, 2021a]. In particular, the concept of the limiting normal cone was not required. Instead, the multipliers to the system of M-stationarity were constructed as a convex combination of several multipliers which satisfy a system of  $A_\beta$ -stationarity (where we use  $A_\beta$ -stationarity as defined in Definition 2.1). These  $A_\beta$ -stationarity systems have already been observed in [Flegel, Kanzow, 2005, Theorem 3.4].

A first idea to prove M-stationarity for MPCCs in Lebesgue spaces is to utilize the calculus for the limiting normal cone. This fails for two reasons. First, it was shown in [Mehlitz, Wachsmuth, 2018; 2019] that the limiting normal cone to a complementarity set in  $L^2(\mu)$  is always convex and corresponds to weak stationarity. Second, the calculus rules for the limiting normal cone typically utilize the notion of *sequential normal compactness*, but this fails for many subsets of Lebesgue spaces, see [Mehlitz, 2018].

The situation is slightly different for complementarity sets in Sobolev spaces. In the one-dimensional situation, it can be shown that the limiting normal cone is small enough and, thus, it is not too difficult to prove necessary optimality conditions of M-stationary type, see [Jarušek, Outrata, 2007; Wachsmuth, 2016]. In dimensions  $n \geq 2$ , these arguments cannot be transferred, see [Hintermüller, Mordukhovich, Surowiec, 2014; Wachsmuth, 2016]. The main reason was provided later in the paper [Harder, Wachsmuth, 2018b], which shows that again the limiting normal cone of a complementarity set in Sobolev spaces is too big if the dimension is bigger than 1.

In the present paper, we are going to transfer the approach of [Harder, 2021a]. The idea is to take convex combinations of so-called  $A_\beta$ -stationary multipliers, where  $\beta$  ranges over the measurable subsets of the biactive set, see Definition 2.1. Under some regularity assumptions on these multipliers, we were indeed able to prove the existence of an M-stationary multiplier. A very crucial difference to the finite-dimensional situation is that the existence of the  $A_\beta$ -stationary multipliers does not come for free. To explain this further, the associated system of  $A_\beta$ -stationarity is precisely the KKT system of a tightened and linearized problem (see  $(LP(\beta))$  for such a problem). In finite dimensions, the existence of multipliers follows directly from the linearity of the problem, but this is

not the case in infinite dimensions. Quite surprisingly, there exist situations in which solutions of this problem fail to possess multipliers, see [Example 4.8](#).

Therefore, the main contributions of this paper are the following.

- (i) Combining  $A_\beta$ -stationary multipliers into M-stationary multipliers is possible ([Theorem 3.1](#)).
- (ii) We prove the existence of  $A_\beta$ -stationary multipliers for a linear MPCC ([Proposition 4.7](#)).
- (iii) By means of an example, we show that nonlinear MPCCs can be linearized ([Lemma 5.8](#)) in order to apply the previous results ([Theorem 5.9](#)).

Thus, we give the first result which provides a system of M-stationarity for an MPCC in Lebesgue spaces. Previously, it was unknown whether M-stationarity holds for instances of MPCCs and bilevel problems in Lebesgue spaces, see the open questions in [[Harder, Wachsmuth, 2018a](#), p. 642], [[Harder, 2021c](#), p. 224].

In [Section 2](#), we formally introduce MPCCs in Lebesgue spaces and define the corresponding stationarity conditions. [Section 3](#) deals with the combination of multipliers to the  $A_\beta$ -stationarity system into multipliers of the M-stationarity system. In [Section 4](#) we apply this to a class of linear MPCCs. It turns out that showing  $A_\beta$ -stationarity of local minimizers is difficult and is only possible under some further assumptions on the structure of the MPCC. In [Section 5](#) we study an inverse optimal control problem, which is nonlinear and more natural looking than a linear MPCC. The KKT reformulation of the inverse optimal control problem is an MPCC problem in Lebesgue spaces and thus we can apply the previous results. However, the linearization of the inverse optimal control problem requires some effort, including regularization methods. Finally, in [Section 6](#) we give an outlook and conclude.

## 2 Stationarity concepts for MPCCs in Lebesgue spaces

We first consider an abstract class of MPCCs in Lebesgue spaces. This class is broad enough such that all MPCCs from this article fit into the setting.

Let  $(\Omega_i, \mathcal{A}_i, m_i)$  be a measure space for all  $i \in \{1, 2\}$  and let  $X, Y$  be Banach spaces. We consider the class of MPCCs given by

$$\begin{aligned}
 & \min_{x \in X} F(x) \\
 & \text{s.t. } h(x) = 0, \\
 & \quad g(x) \geq 0 \quad \text{a.e. on } \Omega_1, \\
 & \quad 0 \leq G(x) \perp H(x) \geq 0 \quad \text{a.e. on } \Omega_2,
 \end{aligned} \tag{MPCC}$$

where  $F: X \rightarrow \mathbb{R}$ ,  $h: X \rightarrow Y$ ,  $g: X \rightarrow L^2(\Omega_1)$ ,  $G, H: X \rightarrow L^2(\Omega_2)$  are continuously Fréchet differentiable functions. The last constraint in (MPCC) is an abbreviation of the pointwise complementarity

$$0 \leq G(x) \quad \wedge \quad G(x)H(x) = 0 \quad \wedge \quad H(x) \geq 0 \quad \text{a.e. on } \Omega_2.$$

For a measurable function  $v: \Omega_i \rightarrow \mathbb{R}$ , we will use the notation  $\{v > 0\}$  to refer to the set  $\{\omega \in \Omega_i \mid v(\omega) > 0\}$ , and similar notations for other pointwise relations will be used throughout this paper.

Let us define stationarity conditions for this MPCC.

**Definition 2.1.** Let  $\bar{x}$  be a feasible point of (MPCC). We call  $\bar{x}$  a *weakly stationary point* if there exist multipliers  $\bar{\eta} \in Y^*$ ,  $\bar{\rho} \in L^2(\Omega_1)$ ,  $\bar{\mu}, \bar{\nu} \in L^2(\Omega_2)$  such that

$$\begin{aligned} F'(\bar{x}) + h'(\bar{x})^* \bar{\eta} + g'(\bar{x})^* \bar{\rho} + G'(\bar{x})^* \bar{\mu} + H'(\bar{x})^* \bar{\nu} &= 0, \\ \bar{\rho} &\leq 0 \quad \text{a.e. on } \Omega_1, \\ \bar{\rho} &= 0 \quad \text{a.e. on } \{g(\bar{x}) > 0\}, \\ \bar{\mu} &= 0 \quad \text{a.e. on } \{G(\bar{x}) > 0\}, \\ \bar{\nu} &= 0 \quad \text{a.e. on } \{H(\bar{x}) > 0\}. \end{aligned}$$

If we have the additional condition

$$(\bar{\mu} < 0 \wedge \bar{\nu} < 0) \vee \bar{\mu}\bar{\nu} = 0 \quad \text{a.e. on } \{G(\bar{x}) = H(\bar{x}) = 0\}$$

then  $\bar{x}$  is an *M-stationary point*. We say that  $\bar{x}$  is a *strongly stationary* (or *S-stationary*) point if  $\bar{x}$  is weakly stationary and

$$\bar{\mu} \leq 0 \wedge \bar{\nu} \leq 0 \quad \text{a.e. on } \{G(\bar{x}) = H(\bar{x}) = 0\}$$

holds. We say that  $\bar{x}$  is *A $_\beta$ -stationary*, where  $\beta \subset \{G(\bar{x}) = H(\bar{x}) = 0\}$  is measurable, if there exist weakly stationary multipliers which additionally satisfy

$$\bar{\mu} \leq 0 \quad \text{a.e. on } \{G(\bar{x}) = H(\bar{x}) = 0\} \setminus \beta \quad \text{and} \quad \bar{\nu} \leq 0 \quad \text{a.e. on } \beta.$$

If there exists a measurable  $\beta \subset \{G(\bar{x}) = H(\bar{x}) = 0\}$  such that  $\bar{x}$  is *A $_\beta$ -stationary*, then  $\bar{x}$  is called *A-stationary*. If  $\bar{x}$  is *A $_\beta$ -stationary* for all measurable  $\beta \subset \{G(\bar{x}) = H(\bar{x}) = 0\}$ , we say that  $\bar{x}$  is *A $_\forall$ -stationary*.

In the case that  $\Omega_i$  is finite (and equipped with the counting measure  $m_i$ ) and  $X, Y$  are finite dimensional, Definition 2.1 coincides with the definitions for finite-dimensional MPCCs in the literature. An exception is *A $_\forall$ -stationarity*, which was not yet defined for MPCCs, but is equivalent to linearized B-stationarity, see [Harder, 2021b, Proposition 3.1 (b)]. We further note that (as far as we know) *A $_\beta$ -stationarity* was only recently defined in finite dimensions in [Harder, 2021b, Definition 2.5].

We also mention that in an infinite-dimensional setting, the name ‘‘M-stationarity’’ can sometimes refer to a condition involving the limiting normal cone (see, e.g., [Mehlitz, 2017, Definition 3.2]), which would be different from our pointwise a.e. definition.

### 3 From $A_{\forall}$ -stationarity to M-stationarity

In [Harder, 2021a, Lemma 3.2], it was shown for finite dimensional MPCCs that one can construct multipliers for the system of M-stationarity as a convex combination of the collection of multipliers which satisfy a system of  $A_{\beta}$ -stationarity for some  $\beta$ . In this section, we want to make a similar statement in an infinite-dimensional setting. As we do not have finitely many multipliers anymore, this requires some additional assumptions, and also the proof is much more complicated. We formulate the main theorem of this section. Its proof is given after some auxiliary lemmas.

**Theorem 3.1.** Let  $(\Omega, \mathcal{A}, m)$  be a measure space and  $1 < p, q < \infty$ . Suppose that for all  $\alpha \in \mathcal{A}$  there exist  $\mu^\alpha \in L^p(\Omega), \nu^\alpha \in L^q(\Omega)$  with

$$\mu^\alpha \leq 0 \quad \text{a.e. on } \Omega \setminus \alpha, \quad \nu^\alpha \leq 0 \quad \text{a.e. on } \alpha. \quad (3.1)$$

Furthermore, we assume the existence of a measurable function  $c_0 : \Omega \rightarrow \mathbb{R}$  and a constant  $c_1 > 0$  with

$$\max(\mu^\alpha, \nu^\alpha) \leq c_0 \quad \text{a.e. on } \Omega, \quad \forall \alpha \in \mathcal{A}, \quad (3.2a)$$

$$\|\mu^\alpha\|_{L^p(\Omega)} + \|\nu^\alpha\|_{L^q(\Omega)} \leq c_1 \quad \forall \alpha \in \mathcal{A}. \quad (3.2b)$$

Then, there exists a point

$$(\bar{\mu}, \bar{\nu}) \in B := \overline{\text{conv}}\{(\mu^\alpha, \nu^\alpha) \mid \alpha \in \mathcal{A}\} \subset L^p(\Omega) \times L^q(\Omega)$$

which satisfies

$$(\bar{\mu} < 0 \wedge \bar{\nu} < 0) \vee \bar{\mu}\bar{\nu} = 0 \quad \text{a.e. on } \Omega. \quad (3.3)$$

Here and in the sequel,  $\overline{\text{conv}} D$  denotes the closed convex hull of a set  $D$ .

If  $\Omega$  is finite (and if  $\mathcal{A}$  is the power set of  $\Omega$ ), then (3.2) is automatically satisfied, and the statement is the same as [Harder, 2021a, Lemma 3.2].

Later, we are going to apply this theorem to an instance of (MPCC) with the setting  $\Omega = \{G(\bar{x}) = H(\bar{x}) = 0\}$ . Then, required existence of multipliers satisfying (3.1) is basically the  $A_{\forall}$ -stationarity of  $\bar{x}$  (with the additional regularity (3.2)), whereas the constructed multipliers  $(\bar{\mu}, \bar{\nu})$  will yield the M-stationarity of  $\bar{x}$ , see Theorem 5.9.

The proof requires several steps. First, for all  $\alpha \in \mathcal{A}$  we define the sets

$$A^\alpha := \{(\mu, \nu) \in L^p(\Omega) \times L^q(\Omega) \mid \mu \leq 0 \text{ a.e. on } \Omega \setminus \alpha, \nu \leq 0 \text{ a.e. on } \alpha\}$$

which describe (3.1). Clearly, we have  $(\mu^\alpha, \nu^\alpha) \in B \cap A^\alpha$  for all  $\alpha \in \mathcal{A}$ . We also define the set of sets

$$\mathcal{C} := \left\{ \overline{\text{conv}}\{x^\alpha \mid \alpha \in \mathcal{A}\} \mid \{x^\alpha\}_{\alpha \in \mathcal{A}} \subset B : \forall \alpha \in \mathcal{A} : x^\alpha \in A^\alpha \right\} \subset \mathcal{P}(L^p(\Omega) \times L^q(\Omega)).$$

Clearly we have  $B \in \mathcal{C}$ . Note that all sets in  $\mathcal{C}$  are closed, convex and nonempty subsets of  $B$ . As  $B$  is bounded due to (3.2b), all sets in  $\mathcal{C}$  are bounded. We are interested in a set in  $\mathcal{C}$  which is minimal with respect to set inclusion. The next two lemmas prove the existence of such a minimal set.

**Lemma 3.2.** Let  $I$  be a nonempty set and let  $Y$  be a topological space. Suppose there exists a nonempty compact set  $D_i \subset Y$  for each  $i \in I$  and that these sets satisfy

$$D_i \subset D_j \vee D_j \subset D_i \quad \forall i, j \in I. \quad (3.4)$$

Then  $\bigcap_{i \in I} D_i \neq \emptyset$  holds.

*Proof.* Pick  $i_0 \in I$  arbitrary. Suppose that  $\bigcap_{i \in I} D_i = \emptyset$ . Then the family  $\{Y \setminus D_i\}_{i \in I}$  of open sets covers  $Y$ , and in particular  $D_{i_0}$ . Thus, we can select a finite subcover  $\{Y \setminus D_{i_j}\}_{j=1}^n$  which covers  $D_{i_0}$ . Due to (3.4) we can without loss of generality assume that  $D_{i_1} \subset D_{i_2} \subset \dots \subset D_{i_n}$ . Thus, the single open set  $Y \setminus D_{i_1}$  is a cover of  $D_{i_0}$ . This implies  $D_{i_1} \cap D_{i_0} = \emptyset$ , which contradicts (3.4).

**Lemma 3.3.** There exists  $C_{\min} \in \mathcal{C}$  which is minimal with respect to set inclusion.

*Proof.* We will use Zorn's Lemma on  $\mathcal{C}$  with the partial ordering " $\subset$ ". Clearly,  $\mathcal{C}$  is nonempty. Let  $\{C_i\}_{i \in I} \subset \mathcal{C}$  be a nonempty totally ordered chain. We want to show that this chain has a lower bound. It follows that  $\{A^\alpha \cap C_i\}_{i \in I}$  is a totally ordered chain for each  $\alpha \in \mathcal{A}$ . Because  $A^\alpha \cap C_i \subset B$  is a bounded and closed set, it is weakly compact for all  $\alpha \in \mathcal{A}$ ,  $i \in I$ . Moreover, these sets are nonempty. Thus, for each  $\alpha$ , we can apply Lemma 3.2 for the weak topology on  $L^p(\Omega) \times L^q(\Omega)$  to the family of sets  $\{A^\alpha \cap C_i\}_{i \in I}$ . This yields that the intersection is nonempty, i.e. there exists a point  $\hat{x}^\alpha$  with  $\hat{x}^\alpha \in A^\alpha \cap C_i$  for all  $i \in I$ . Now we set  $C_0 := \overline{\text{conv}}\{\hat{x}^\alpha \mid \alpha \in \mathcal{A}\}$ . Clearly,  $C_0 \in \mathcal{C}$ . Since the points  $\hat{x}^\alpha$  are contained in  $C_i$  for all  $i \in I$ , it follows that  $C_0 \subset C_i$  for all  $i \in I$ . Thus, the totally ordered chain  $\{C_i\}_{i \in I}$  has a lower bound in  $\mathcal{C}$ .

By Zorn's Lemma, there exists a set  $C_{\min} \in \mathcal{C}$  which is minimal with respect to the relation " $\subset$ ".

It is clear that  $C_{\min}$  does not have to be unique. By construction of  $\mathcal{C}$ , the set  $C_{\min}$  is non-empty, closed and convex and it is bounded due to (3.2b). Owing to the Krein–Milman theorem, we know that  $C_{\min}$  possesses extreme points. For our proof of Theorem 3.1, we need the stronger notion of so-called *strongly exposed points* of  $C_{\min}$ . We start with the definition.



**Definition 3.4.** Let  $D \subset X$ , and  $X$  a normed space. Then  $x \in D$  is a strongly exposed point if there exists  $x^* \in X^*$  such that  $x$  maximizes  $x^*$  over  $D$  and

$$\langle x^*, x_k \rangle \rightarrow \langle x^*, x \rangle \implies x_k \rightarrow x$$

holds for all sequences  $(x_k)_{k \in \mathbb{N}} \subset D$ .

It can be seen that a strongly exposed point is always an extreme point. A non-empty, convex, closed and bounded set in reflexive spaces always has a strongly exposed point, see [Fabian et al., 2001, Theorem 8.28]. In particular,  $C_{\min}$  possesses strongly exposed points. Our next goal is to show that such a strongly exposed point of  $C_{\min}$  satisfies (3.3), which would prove Theorem 3.1. We give a preparatory lemma.

**Lemma 3.5.** Let  $\xi \in (L^p(\Omega) \times L^q(\Omega))^*$  be a bounded linear functional on  $L^p(\Omega) \times L^q(\Omega)$ . Then there exists a collection  $\{\hat{x}^\alpha\}_{\alpha \in \mathcal{A}}$  such that

$$\hat{x}^\alpha \in C_{\min} \cap A^\alpha \quad \forall \alpha \in \mathcal{A}, \tag{3.5}$$

$$\langle \xi, \hat{x}^\alpha \rangle \leq \langle \xi, y^\alpha \rangle \quad \forall y^\alpha \in C_{\min} \cap A^\alpha \quad \forall \alpha \in \mathcal{A}, \tag{3.6}$$

$$\overline{\text{conv}}\{\hat{x}^\beta \mid \beta \in \mathcal{A}\} = C_{\min}. \tag{3.7}$$

*Proof.* We choose  $\hat{x}^\alpha$  as a minimizer of  $\xi$  over  $C_{\min} \cap A^\alpha$ . Such a minimizer exists because  $C_{\min} \cap A^\alpha$  is weakly compact and  $\xi$  is weakly continuous.

The condition (3.7) follows because all points  $\hat{x}^\beta$  are in the closed and convex set  $C_{\min}$  and because  $C_{\min}$  was chosen as a minimal set in  $\mathcal{C}$  with respect to set inclusion.

**Lemma 3.6.** Every strongly exposed point  $(\bar{\mu}, \bar{\nu})$  of  $C_{\min}$  satisfies (3.3).

*Proof.* Let  $\bar{x} = (\bar{\mu}, \bar{\nu})$  be a strongly exposed point of  $C_{\min}$  and let  $\xi \in (L^p(\Omega) \times L^q(\Omega))^*$  be the corresponding multiplier. We choose  $\{\hat{x}^\alpha\}_{\alpha \in \mathcal{A}}$  according to Lemma 3.5.

Let  $k \in \mathbb{N}$  be given and suppose that  $\langle \xi, \bar{x} \rangle \geq \langle \xi, \hat{x}^\alpha \rangle + \frac{1}{k}$  holds for all  $\alpha \in \mathcal{A}$ . Then it would follow from (3.7) that  $\xi$  separates the closed convex set  $C_{\min}$  from  $\bar{x}$  and therefore  $\bar{x} \notin C_{\min}$ , which is not possible since  $\bar{x}$  is an extremal point of  $C_{\min}$ . Thus, there exists a set  $\alpha_k \in \mathcal{A}$  such that

$$\langle \xi, \bar{x} \rangle - \frac{1}{k} \leq \langle \xi, \hat{x}^{\alpha_k} \rangle \leq \langle \xi, \bar{x} \rangle$$

holds. Let  $\ell \in \mathbb{N}$  with  $\ell \geq 2$  be fixed. We make the definitions

$$D_\ell := \{(a_1, a_2) \in \mathbb{R}^2 \mid \min(|x_1|, |x_2|) > 1/\ell \wedge x_1 x_2 < 0\},$$

$$\beta_k := \{\omega \in \Omega \mid \hat{x}^{\alpha_k}(\omega) \in D_\ell \wedge c_0(\omega) \leq \ell\} \in \mathcal{A},$$

$$\gamma_k := \alpha_k \Delta \beta_k \in \mathcal{A}$$

for all  $k \in \mathbb{N}$ , where  $\Delta$  denotes the symmetric difference. We claim that

$$y_k := (1 - \ell^{-3})\hat{x}^{\alpha_k} + \ell^{-3}\hat{x}^{\gamma_k} \in A^{\alpha_k} \cap C_{\min}$$

holds for all  $k \in \mathbb{N}$ . Indeed, for  $\omega \in \alpha_k \cap \beta_k$  we have  $y_k(\omega)_2 \leq (1 - \ell^{-3})\hat{x}^{\alpha_k}(\omega)_2 + \ell^{-3}c_0(\omega) \leq (1 - \ell^{-3})(-\ell^{-1}) + \ell^{-3}\ell \leq 0$  a.e. on  $\alpha_k \cap \beta_k$ . Similarly, for  $\omega \in \beta_k \setminus \alpha_k$  we have  $y_k(\omega)_1 \leq (1 - \ell^{-3})\hat{x}^{\alpha_k}(\omega)_1 + \ell^{-3}c_0(\omega) \leq (1 - \ell^{-3})(-\ell^{-1}) + \ell^{-3}\ell \leq 0$  a.e. on  $\beta_k \setminus \alpha_k$ . If  $\omega \notin \beta_k$ , then we have  $\omega \in \alpha_k \Leftrightarrow \omega \in \gamma_k$ , which implies  $y_k(\omega)_1 \leq 0$  a.e. on  $\Omega \setminus (\alpha_k \cup \beta_k)$  and  $y_k(\omega)_2 \leq 0$  a.e. on  $\alpha_k \setminus \beta_k$  by convexity and  $\hat{x}^{\gamma_k} \in A^{\gamma_k}$ . This shows  $y_k \in A^{\alpha_k}$ , and  $y_k \in C_{\min}$  follows by convexity.

According to (3.6) we then have  $\langle \xi, \hat{x}^{\alpha_k} \rangle \leq \langle \xi, y^k \rangle$ , which implies

$$\langle \xi, \bar{x} \rangle - \frac{1}{k} \leq \langle \xi, \hat{x}^{\alpha_k} \rangle \leq \langle \xi, \hat{x}^{\gamma_k} \rangle \leq \langle \xi, \bar{x} \rangle.$$

Then the fact that  $\bar{x}$  is a strongly exposed point of  $C_{\min}$  implies the convergences  $\hat{x}^{\alpha_k} \rightarrow \bar{x}$  and  $\hat{x}^{\gamma_k} \rightarrow \bar{x}$ . Next, using  $\hat{x}^{\alpha_k} \in A^{\alpha_k}$ ,  $\hat{x}^{\gamma_k} \in A^{\gamma_k}$ , and  $\hat{x}^{\alpha_k} \in D_\ell$  a.e. on  $\beta_k$ , we observe that

$$\begin{aligned} \hat{x}_2^{\alpha_k} \leq 0 \wedge \hat{x}_1^{\alpha_k} > 1/\ell \wedge \hat{x}_1^{\gamma_k} \leq 0 & \quad \text{a.e. on } \beta_k \cap \alpha_k, \\ \hat{x}_1^{\alpha_k} \leq 0 \wedge \hat{x}_2^{\alpha_k} > 1/\ell \wedge \hat{x}_2^{\gamma_k} \leq 0 & \quad \text{a.e. on } \beta_k \setminus \alpha_k \end{aligned}$$

holds. This implies

$$\ell \max(|\hat{x}^{\alpha_k}(\omega)_1 - \hat{x}^{\gamma_k}(\omega)_1|, |\hat{x}^{\alpha_k}(\omega)_2 - \hat{x}^{\gamma_k}(\omega)_2|) \geq 1 \quad \text{a.e. on } \beta_k.$$

By integrating, we obtain

$$\begin{aligned} m(\beta_k) &= \int_{\beta_k} 1 \, dm \leq \int_{\beta_k} \ell^p |\hat{x}_1^{\alpha_k} - \hat{x}_1^{\gamma_k}|^p \, dm + \int_{\beta_k} \ell^q |\hat{x}_2^{\alpha_k} - \hat{x}_2^{\gamma_k}|^q \, dm \\ &\leq \ell^p \|\hat{x}_1^{\alpha_k} - \hat{x}_1^{\gamma_k}\|_{L^p(\Omega)}^p + \ell^q \|\hat{x}_2^{\alpha_k} - \hat{x}_2^{\gamma_k}\|_{L^q(\Omega)}^q \rightarrow 0 \quad \text{as } k \rightarrow \infty, \end{aligned}$$

where the last convergence follows from  $\hat{x}^{\alpha_k} \rightarrow \bar{x}$  and  $\hat{x}^{\gamma_k} \rightarrow \bar{x}$ . Without loss of generality we can say that the convergence  $\hat{x}^{\alpha_k} \rightarrow \bar{x}$  also holds pointwise a.e. (otherwise one can choose a subsequence with this property). Since  $D_\ell$  is open,  $\chi_{\{\bar{x} \in D_\ell\}} \leq \liminf_{k \rightarrow \infty} \chi_{\{\hat{x}^{\alpha_k} \in D_\ell\}}$  and therefore  $\chi_{\{\bar{x} \in D_\ell\} \cap \{c_0 \leq \ell\}} \leq \liminf_{k \rightarrow \infty} \chi_{\beta_k}$  hold a.e. in  $\Omega$ . Applying Fatou's Lemma to the last inequality yields

$$m(\{\bar{x} \in D_\ell\} \cap \{c_0 \leq \ell\}) \leq \liminf_{k \rightarrow \infty} m(\beta_k) = 0.$$

Since  $\ell \in \mathbb{N}$ ,  $\ell \geq 2$  was arbitrary, this implies

$$\begin{aligned} 0 &= m\left(\bigcup_{\ell=2}^{\infty} \{\bar{x} \in D_\ell\} \cap \{c_0 \leq \ell\}\right) = m\left(\left(\bigcup_{\ell=2}^{\infty} \{\bar{x} \in D_\ell\}\right) \cap \left(\bigcup_{\ell=2}^{\infty} \{c_0 \leq \ell\}\right)\right) \\ &= m\left(\left(\bigcup_{\ell=2}^{\infty} \{\bar{x} \in D_\ell\} \cap \Omega\right)\right) = m\left(\bigcup_{\ell=2}^{\infty} \{\bar{x} \in D_\ell\}\right) = m(\{\bar{x}_1 \bar{x}_2 < 0\}), \end{aligned}$$

where we used that  $\{\{\bar{x} \in D_\ell\}\}_{\ell \in \mathbb{N}}$  and  $\{\{c_0 \leq \ell\}\}_{\ell \in \mathbb{N}}$  are monotonically increasing sequences. Thus, we have  $\bar{x}_1 \bar{x}_2 \geq 0$  a.e. on  $\Omega$ . Due to  $\hat{x}^{\alpha_k} \in A^{\alpha_k}$  we have  $\min(\hat{x}_1^{\alpha_k}, \hat{x}_2^{\alpha_k}) \leq 0$  a.e. on  $\Omega$ , and by considering a pointwise a.e. converging subsequence of  $\{\hat{x}^{\alpha_k}\}_{k \in \mathbb{N}}$ , we obtain  $\min(\bar{x}_1, \bar{x}_2) \leq 0$  a.e. on  $\Omega$ . In summary, we have  $\min(\bar{\mu}, \bar{\nu}) \leq 0$  and  $\bar{\mu} \bar{\nu} \geq 0$  a.e. on  $\Omega$ . However, this is equivalent to (3.3).

Now we have all the necessary ingredients to prove [Theorem 3.1](#).

*Proof of [Theorem 3.1](#).* As the space  $L^p(\Omega) \times L^q(\Omega)$  is reflexive, the set  $C_{\min}$  is a weakly compact subset of  $L^p(\Omega) \times L^q(\Omega)$ . Thus, according to [[Fabian et al., 2001](#), Theorem 8.28] the set  $C_{\min}$  is the closed convex hull of its strongly exposed points. Since  $C_{\min}$  is nonempty, this implies that  $C_{\min}$  has at least one strongly exposed point. By [Lemma 3.6](#), this point satisfies (3.3) and our proof of [Theorem 3.1](#) is complete.

**Remark 3.7.** We mention that [Theorem 3.1](#) not only yields the existence of a single multiplier pair satisfying (3.3). Indeed, every strongly exposed point of  $C_{\min}$  satisfies (3.3) and  $C_{\min}$  is the closed convex hull of its strongly exposed points. Moreover,  $C_{\min}$  has a non-empty intersection with all  $A^\alpha$ .

To summarize, the closed convex hull of all pairs  $(\hat{\mu}, \hat{\nu}) \in B$  satisfying (3.3) has a non-empty intersection with  $A^\alpha$  for all  $\alpha \in \mathcal{A}$ .

Furthermore, suppose that there exists no point  $(\bar{\mu}, \bar{\nu}) \in B$  with  $\bar{\mu} \leq 0$  and  $\bar{\nu} \leq 0$  a.e. in  $\Omega$  (which corresponds to strong stationarity). This means  $\bigcap_{\alpha \in \mathcal{A}} A^\alpha \cap B = \emptyset$ , and therefore there must exist at least two points  $(\bar{\mu}, \bar{\nu}) \in B$  that satisfy (3.3).

The assumptions in [Theorem 3.1](#) require that there is a bound on  $\|\mu^\alpha\|_{L^p(\Omega)}, \|\nu^\alpha\|_{L^q(\Omega)}$ . The next result is based on [Theorem 3.1](#) but has different assumptions. In particular, it could also be applied in situations where  $\|\mu^\alpha\|_{L^p(\Omega)}, \|\nu^\alpha\|_{L^q(\Omega)}$  might be unbounded, provided that some other assumptions are satisfied.

**Lemma 3.8.** Let  $(\Omega, \mathcal{A}, m)$  be a measure space and  $Y$  be a reflexive Banach space. Suppose that for all  $\alpha \in \mathcal{A}$  there exist  $\mu^\alpha, \nu^\alpha \in L^2(\Omega)$ ,  $\rho^\alpha \in Y$  such that

$$\mu^\alpha \leq 0 \quad \text{a.e. on } \Omega \setminus \alpha, \quad \nu^\alpha \leq 0 \quad \text{a.e. on } \alpha. \quad (3.8)$$

Furthermore, we assume that there exist functions  $c_0 \in L^2(\Omega)$ ,  $c_1: \mathcal{A} \rightarrow (0, \infty)$  with

$$\max(|\mu^\alpha + \nu^\alpha|, \nu^\alpha) \leq c_0 \quad \text{a.e. on } \Omega \quad \forall \alpha \in \mathcal{A}, \quad (3.9)$$

$$|\nu^\alpha| \leq c_1(\alpha)c_0 \quad \text{a.e. on } \Omega \quad \forall \alpha \in \mathcal{A}, \quad (3.10)$$

and a bounded linear operator  $T: L^2(\Omega) \rightarrow Y$  and constant  $c_2 > 0$  with

$$\|\rho^\alpha - T\nu^\alpha\|_Y \leq c_2 \quad \forall \alpha \in \mathcal{A}. \quad (3.11)$$

Then there exists a point

$$(\bar{\mu}, \bar{\nu}, \bar{\rho}) \in B := \overline{\text{conv}}\{(\mu^\alpha, \nu^\alpha, \rho^\alpha) \mid \alpha \in \mathcal{A}\}$$

such that  $\bar{\mu}, \bar{\nu}$  satisfy (3.3).

*Proof.* We want to apply [Theorem 3.1](#). For that purpose, we want to modify  $\mu^\alpha, \nu^\alpha$  in such a way that they are bounded in  $L^2(\Omega)$ . We define the sets

$$\begin{aligned} A^\alpha &:= \{(\mu, \nu, \rho) \in L^2(\Omega)^2 \times Y \mid \mu \leq 0 \text{ a.e. on } \Omega \setminus \alpha, \nu \leq 0 \text{ a.e. on } \alpha\}, \\ B &:= \overline{\text{conv}}\{(\mu^\alpha, \nu^\alpha, \rho^\alpha) \mid \alpha \in \mathcal{A}\}. \end{aligned}$$

For convexity and continuity reasons we also have the properties

$$\max(|\mu + \nu|, \nu) \leq c_0 \quad \text{a.e. on } \Omega \quad \forall (\mu, \nu, \rho) \in B, \quad (3.12)$$

$$\|\rho - T\nu\| \leq c_2 \quad \forall (\mu, \nu, \rho) \in B. \quad (3.13)$$

Let  $\alpha \in \mathcal{A}$  be given. We define

$$c_3(\alpha) := \inf\{d > 0 \mid \exists (\mu, \nu, \rho) \in B \cap A^\alpha : |\nu| \leq dc_0 \text{ a.e. on } \Omega\}. \quad (3.14)$$

We claim that there exists  $(\hat{\mu}^\alpha, \hat{\nu}^\alpha, \hat{\rho}^\alpha) \in B \cap A^\alpha$  with

$$|\hat{\nu}^\alpha| \leq c_3(\alpha)c_0 \quad \text{a.e. on } \Omega. \quad (3.15)$$

Indeed, since  $c_3(\alpha) \leq c_1(\alpha) < \infty$  we find sequences  $\{d_k\}_{k \in \mathbb{N}} \subset (0, \infty)$ ,  $\{(\mu^k, \nu^k, \rho^k)\}_{k \in \mathbb{N}} \subset B \cap A^\alpha$  with  $d_k \downarrow c_3(\alpha)$  and  $|\nu^k| \leq d_k c_0$  a.e. on  $\Omega$ . Then  $\{\nu^k\}_{k \in \mathbb{N}}$  is a bounded sequence. By (3.12) and (3.13) it follows that  $\{(\mu^k, \nu^k, \rho^k)\}_{k \in \mathbb{N}}$  is also a bounded sequence and therefore has a weakly convergent subsequence. Let  $(\hat{\mu}^\alpha, \hat{\nu}^\alpha, \hat{\rho}^\alpha)$  denote a weak limit of a subsequence. Then  $(\hat{\mu}^\alpha, \hat{\nu}^\alpha, \hat{\rho}^\alpha) \in B \cap A^\alpha$  holds since  $B \cap A^\alpha$  is a closed convex set. Similarly, for all  $k \in \mathbb{N}$  the condition  $|\hat{\nu}^\alpha| \leq d_k c_0$  holds a.e. on  $\Omega$ . Then (3.15) follows from the convergence  $d_k \downarrow c_3(\alpha)$ .

Our next goal will be to show that  $c_3(\alpha) \leq 1$  holds. Suppose, by contradiction, that  $c_3(\alpha) > 1$  holds. Combining (3.12) with  $(\hat{\mu}^\alpha, \hat{\nu}^\alpha, \hat{\rho}^\alpha) \in B \cap A^\alpha$  yields the conditions

$$\begin{aligned} \hat{\nu}^\alpha &\geq -\hat{\mu}^\alpha - c_0 \geq -c_0 \quad \text{a.e. on } \Omega \setminus \alpha, \\ \hat{\nu}^\alpha &\leq c_0 \quad \text{a.e. on } \Omega. \end{aligned}$$

For  $t \in [0, 1]$  we make the definitions

$$\begin{aligned} \beta &:= \alpha \cap \{\hat{\nu}^\alpha \geq -c_0\} \in \mathcal{A}, \\ (\mu^t, \nu^t, \rho^t) &:= (1-t)(\hat{\mu}^\alpha, \hat{\nu}^\alpha, \hat{\rho}^\alpha) + t(\mu^\beta, \nu^\beta, \rho^\beta) \in B. \end{aligned}$$

We claim that  $(\mu^t, \nu^t, \rho^t) \in A^\alpha$  holds for all  $t \in [0, 1/2]$ . Indeed, the condition  $\nu^t \leq 0$  a.e. on  $\alpha$  follows from  $\nu^\beta \leq 0$  a.e. on  $\beta \subset \alpha$  and  $\nu^\beta \leq c_0$  a.e. on  $\alpha \setminus \beta = \{\hat{\nu}^\alpha < -c_0\}$ , and the condition  $\mu^t \leq 0$  a.e. on  $\Omega \setminus \alpha$  follows from  $\mu^\beta \leq 0$  a.e. on  $\Omega \setminus \alpha \subset \Omega \setminus \beta$ .

We are interested in lower bounds on  $\nu^t$ . Combining (3.8) and (3.9) yields the estimate  $\nu^\beta \geq -\mu^\beta - c_0 \geq -c_0$  a.e. on  $\Omega \setminus \beta$ . Now we have

$$\begin{aligned} \nu^t &= (1-t)\hat{\nu}^\alpha + t\nu^\beta \geq (1-t)(-c_3(\alpha)c_0) - tc_0 && \text{a.e. on } \Omega \setminus \beta, \\ \nu^t &= (1-t)\hat{\nu}^\alpha + t\nu^\beta \geq (1-t)(-c_0) + t(-c_1(\beta)c_0) && \text{a.e. on } \beta, \\ \nu^t &= (1-t)\hat{\nu}^\alpha + t\nu^\beta \leq (1-t)c_0 + tc_0 = c_0 && \text{a.e. on } \Omega. \end{aligned}$$

This implies

$$|\nu^t| \leq \max((1-t)c_3(\alpha) + t, 1-t + tc_1(\beta), 1)c_0 \quad \text{a.e. on } \Omega.$$

We set  $\hat{t} \in (0, 1/2)$  such that  $(1-\hat{t})c_3(\alpha) + \hat{t} > 1 - \hat{t} + \hat{t}c_1(\beta)$  holds (which is possible due to  $c_3(\alpha) > 1$ ). This results in  $\max((1-\hat{t})c_3(\alpha) + \hat{t}, 1) < c_3(\alpha)$  and

$$|\nu^{\hat{t}}| \leq \max((1-\hat{t})c_3(\alpha) + \hat{t}, 1)c_0 < c_3(\alpha)c_0 \quad \text{a.e. on } \Omega.$$

Due to  $(\mu^{\hat{t}}, \nu^{\hat{t}}, \rho^{\hat{t}}) \in B \cap A^\alpha$  this is a contradiction to (3.14). Thus, our assumption  $c_3(\alpha) > 1$  is false. Using (3.15) results in  $|\hat{\nu}^\alpha| \leq c_0$  a.e. on  $\Omega$ . Then we also have

$$\max(\hat{\nu}^\alpha, \hat{\mu}^\alpha) \leq 2c_0 \text{ a.e. on } \Omega$$

and

$$\|\hat{\nu}^\alpha\|_{L^2(\Omega)} + \|\hat{\mu}^\alpha\|_{L^2(\Omega)} \leq 3\|c_0\|_{L^2(\Omega)}.$$

Since the bounds no longer depend on  $\alpha$ , we can apply Theorem 3.1, which yields a point  $(\bar{\mu}, \bar{\nu}) \in \overline{\text{conv}}\{(\hat{\mu}^\alpha, \hat{\nu}^\alpha) \mid \alpha \in \mathcal{A}\}$  such that  $\bar{\mu}, \bar{\nu}$  satisfy (3.3). By the definition of  $\overline{\text{conv}}$ , there exists a sequence  $\{(\mu_k, \nu_k)\}_{k \in \mathbb{N}} \subset \text{conv}\{(\hat{\mu}^\alpha, \hat{\nu}^\alpha) \mid \alpha \in \mathcal{A}\}$  such that  $(\mu_k, \nu_k) \rightarrow (\bar{\mu}, \bar{\nu})$ . Let  $\{\rho_k\}_{k \in \mathbb{N}} \subset Y$  be a sequence such that  $(\mu_k, \nu_k, \rho_k) \in \text{conv}\{(\hat{\mu}^\alpha, \hat{\nu}^\alpha, \hat{\rho}^\alpha) \mid \alpha \in \mathcal{A}\} \subset B$  holds for all  $k \in \mathbb{N}$ . Due to (3.13), the sequence  $\{\rho_k\}_{k \in \mathbb{N}}$  is a bounded sequence and, thus, possesses a weak accumulation point  $\bar{\rho}$ . Since  $B$  is weakly sequentially closed, this yields  $(\bar{\mu}, \bar{\nu}, \bar{\rho}) \in B$ , which completes the proof.

## 4 A class of linear MPCCs in Lebesgue spaces

We consider the linear MPCC

$$\begin{aligned}
& \min_{u,w,\xi \in L^2(\Omega)} \langle F_u, u \rangle + \langle F_w, w \rangle + \langle F_\xi, \xi \rangle \\
& \text{s.t.} \quad Au - w - \xi = 0, \\
& \quad \quad u = 0 \quad \text{a.e. on } \Omega^{0+}, \\
& \quad \quad 0 \leq u \perp \xi \geq 0 \quad \text{a.e. on } \Omega^{00}, \\
& \quad \quad \xi = 0 \quad \text{a.e. on } \Omega^{+0}, \\
& \quad \quad w \geq 0 \quad \text{a.e. on } \Omega_w,
\end{aligned} \tag{MPCC}_{\text{lin}}$$

where  $A: L^2(\Omega) \rightarrow L^2(\Omega)$  is a linear and continuous operator,  $(\Omega, \mathcal{A}, m)$  is a measure space, and  $F_u, F_w, F_\xi \in L^2(\Omega)$  are fixed. The sets  $\Omega^{0+}, \Omega^{00}, \Omega^{+0}$  are a measurable partition of  $\Omega$ , and  $\Omega_w \subset \Omega$  is measurable.

Since  $(\text{MPCC}_{\text{lin}})$  will appear as a linearized problem, see [Lemma 5.8](#) below, we are only interested in the feasible point  $(0, 0, 0)$ . In the case that  $(0, 0, 0)$  is a minimizer, we want to show that M-stationarity holds (under some additional assumptions on the structure of  $(\text{MPCC}_{\text{lin}})$ ). As we are going to apply the methods in [Section 3](#), we first need to show that the multipliers to the system of  $A_\beta$ -stationarity exist for all measurable  $\beta \subset \Omega^{00}$  and that they satisfy the regularity requirements.

### 4.1 $A_\forall$ -stationarity

In this section we show that  $A_\forall$ -stationarity is satisfied. To this end, let  $\beta \subset \Omega^{00}$  be measurable. The  $A_\beta$ -stationary multipliers will be constructed via the optimality system of the tightened linear optimization problem

$$\begin{aligned}
& \min_{u,w,\xi \in L^2(\Omega)} \langle F_u, u \rangle + \langle F_w, w \rangle + \langle F_\xi, \xi \rangle \\
& \text{s.t.} \quad Au - w - \xi = 0, \\
& \quad \quad u \geq 0 \quad \text{a.e. on } \Omega^{00} \setminus \beta, \quad u = 0 \quad \text{a.e. on } \Omega^{0+} \cup \beta, \\
& \quad \quad \xi \geq 0 \quad \text{a.e. on } \beta, \quad \xi = 0 \quad \text{a.e. on } \Omega^{+0} \cup (\Omega^{00} \setminus \beta), \\
& \quad \quad w \geq 0 \quad \text{a.e. on } \Omega_w.
\end{aligned} \tag{LP}(\beta)$$

Note that if  $(0, 0, 0)$  is a minimizer of  $(\text{MPCC}_{\text{lin}})$ , then it is also a minimizer of  $(\text{LP}(\beta))$ . In the sequel, we show that  $(0, 0, 0)$  is a KKT point of  $(\text{LP}(\beta))$  if it is a minimizer. Although the problem  $(\text{LP}(\beta))$  is linear and seems to be quite innocent, this can be quite complicated and only works for some instances of  $(\text{MPCC}_{\text{lin}})$ . We provide a counterexample ([Example 4.8](#)) which fails to satisfy the KKT conditions at the local minimizer  $(0, 0, 0)$  of  $(\text{LP}(\beta))$ .

The KKT system for  $(\text{LP}(\beta))$  at  $(0, 0, 0)$  is given by

$$\begin{aligned} F_u + A^*p + \mu &= 0, & F_w - p + \lambda &= 0, & F_\xi - p + \nu &= 0, \\ \mu \leq 0 & \text{ a.e. on } \Omega^{00} \setminus \beta, & \mu = 0 & \text{ a.e. on } \Omega^{+0}, & & \\ \nu \leq 0 & \text{ a.e. on } \beta, & \nu = 0 & \text{ a.e. on } \Omega^{0+}, & & \\ \lambda \leq 0 & \text{ a.e. on } \Omega_w, & \lambda = 0 & \text{ a.e. on } \Omega \setminus \Omega_w, & & \end{aligned} \tag{KKT}(\beta)$$

where  $p, \nu, \lambda, \mu \in L^2(\Omega)$  are multipliers. Note that this system coincides with the system of  $A_\beta$ -stationarity for  $(\text{MPCC}_{\text{lin}})$ .

We formulate scenarios under which we are able to prove  $A_\nabla$ -stationarity.

**Assumption 4.1.** (a) The operator  $A$  satisfies  $A^*v \geq 0$  a.e. on  $\Omega$  for all  $v \in L^2(\Omega)$  with  $v \geq 0$  a.e. on  $\Omega$ , and  $m(\Omega^{+0}) = 0$  holds.

(b) The set  $\Omega$  has finite measure and the operator  $A$  is given by  $Av := v + \langle 1, v \rangle$  for all  $v \in L^2(\Omega)$ .

(c) The set  $\Omega$  has finite measure and the operator  $A$  is given by  $Av := d_1v + d_2\langle 1, v \rangle$  for all  $v \in L^2(\Omega)$ , where  $d_1, d_2 > 0$  are fixed.

Note that the operator given in part (b) also satisfies the nonnegativity condition for  $A^*$  in part (a). Clearly, part (b) is only a special case of part (c).

In the following, we define the function

$$c_0 := A^*(|F_u| + |F_w| + |F_\xi|) + |F_u| + |F_w| + |F_\xi| \in L^2(\Omega). \tag{4.1}$$

Under [Assumption 4.1](#) we have  $c_0 \geq 0$  a.e. due to the nonnegativity of  $A^*$ . This function (together with some constants) will be used as a pointwise a.e. upper bound on the multipliers in the next lemmas. We show that if multipliers exist, that one can also find multipliers that can be bounded by  $c_0$  (and some constant, which can depend on  $\beta$  in the case of [Assumption 4.1\(b\)](#)). This serves two purposes: First, this can be used to show that the set of functions  $F_u, F_w, F_\xi$  for which multipliers exist is closed, which is sufficient to show the existence of KKT multipliers if  $(0, 0, 0)$  is a minimizer, see the proof of [Proposition 4.7](#). Second, the pointwise bounds are required to apply the theory from [Section 3](#).

Let us start with the easier case of [Assumption 4.1\(a\)](#).

**Lemma 4.2.** Suppose that [Assumption 4.1\(a\)](#) holds. Let  $p_0, \mu_0, \nu_0, \lambda_0$  be multipliers that satisfy  $(\text{KKT}(\beta))$ . Then one can find multipliers  $p_1, \mu_1, \nu_1, \lambda_1$  that satisfy  $(\text{KKT}(\beta))$  such that

$$\max(|p_1|, |\nu_1|, |\lambda_1|, |\mu_1|) \leq 2c_0 \quad \text{a.e. on } \Omega.$$

*Proof.* First, we introduce the function

$$a_0 := \max(p_0, \lambda_0, \nu_0)^- \geq 0.$$

Here,  $v^- := \max(-v, 0)$  in a pointwise sense. Then we define the new multipliers via

$$p_1 := p_0 + a_0, \quad \lambda_1 := \lambda_0 + a_0, \quad \nu_1 := \nu_0 + a_0, \quad \mu_1 := \mu_0 - A^*a_0.$$

We claim that  $p_1, \mu_1, \nu_1, \lambda_1$  still satisfies [\(KKT\( \$\beta\$ \)\)](#). Indeed, the first three equations of [\(KKT\( \$\beta\$ \)\)](#) are true due to linearity. The conditions involving  $\mu_1$  are true due to  $A^*a_0 \geq 0$  a.e. on  $\Omega$  and  $m(\Omega^{+0}) = 0$ . On  $\{\nu_0 \leq 0\}$  we observe that  $\nu_0 \leq -a_0$  and therefore  $\nu_0 \leq \nu_1 \leq 0$  holds, which implies the conditions involving  $\nu_1$ . Similarly, for the conditions involving  $\lambda_1$  we can use  $\lambda_0 \leq \lambda_1 \leq 0$  on  $\{\lambda_0 \leq 0\}$ .

It remains to verify the pointwise bound. We start with bounds for  $p_1$ . The upper bound follows from

$$p_1 = F_w + \lambda_1 \leq F_w \quad \text{a.e. on } \Omega.$$

We also have the lower bounds

$$\begin{aligned} p_1 &\geq 0 && \text{a.e. on } \{p_1 \geq 0\}, \\ p_1 = F_w + \lambda_1 &= F_w && \text{a.e. on } \{\lambda_1 = 0\}, \\ p_1 = F_\xi + \nu_1 &\geq F_\xi && \text{a.e. on } \{\nu_1 \geq 0\}. \end{aligned}$$

Due to the construction of  $a_0$ , we have  $\Omega = \{p_1 \geq 0\} \cup \{\lambda_1 = 0\} \cup \{\nu_1 \geq 0\}$ , which leads to the lower bound  $p_1 \geq \min(F_w, F_\xi, 0)$  a.e. on  $\Omega$ . Combining the above shows  $|p_1| \leq c_0$  a.e. on  $\Omega$ . Then  $\max(|\nu_1|, |\lambda_1|) \leq 2c_0$  a.e. on  $\Omega$  follows. Finally, using the sign condition of  $A^*$  we get

$$|\mu_1| \leq |F_u| + |A^*p_1| \leq |F_u| + A^*|p_1| \leq c_0 \quad \text{a.e. on } \Omega.$$

The desired pointwise a.e. inequality follows by taking the maximum.

The case of [Assumption 4.1\(b\)](#) is more complicated. We require several lemmas to investigate different cases.

**Lemma 4.3.** Suppose that [Assumption 4.1\(b\)](#) holds. Additionally, we assume  $m(\Omega^{+0}) > 0$ . Then

$$\langle 1, |p| \rangle \leq (2 + m(\Omega^{+0})^{-1})c_0$$

holds for all  $(p, \nu, \mu, \lambda)$  that satisfy [\(KKT\( \$\beta\$ \)\)](#).

*Proof.* First, we have  $p = F_w + \lambda \leq F_w$  a.e. on  $\Omega$ . This implies  $p^+ \leq F_w^+$  a.e. on  $\Omega$  and also  $\langle 1, p^+ \rangle \leq c_0$ , where  $v^+ := \max(v, 0)$ . From  $\mu = 0$  a.e. on  $\Omega^{+0}$  we get  $F_u + p + \langle 1, p \rangle = 0$  a.e. on  $\Omega^{+0}$ , or

$$\langle 1, p \rangle = -p - F_u \geq -F_w - F_u \quad \text{a.e. on } \Omega^{+0}.$$



Integrating over  $\Omega^{+0}$  yields

$$\begin{aligned} m(\Omega^{+0})\langle 1, p \rangle &\geq \langle \chi_{\Omega^{+0}}, -F_w - F_u \rangle \geq -\langle \chi_{\Omega^{+0}}, |F_w| + |F_u| \rangle \\ &\geq -\langle 1, |F_w| + |F_u| \rangle \geq -A^*(|F_w| + |F_u|) \geq -c_0. \end{aligned}$$

Since  $m(\Omega^{+0}) > 0$  holds, this implies  $-\langle 1, p \rangle \leq m(\Omega^{+0})^{-1}c_0$ , and combined with the estimate for  $\langle 1, p^+ \rangle$  this yields

$$\langle 1, |p| \rangle = 2\langle 1, p^+ \rangle - \langle 1, p \rangle \leq 2c_0 + m(\Omega^{+0})^{-1}c_0.$$

**Lemma 4.4.** Suppose that [Assumption 4.1\(b\)](#) holds. Then we have

$$|p| \leq 2c_0 \quad \text{a.e. on } \Omega \setminus (\Omega_w \cap \beta)$$

for all  $(p, \nu, \mu, \lambda)$  that satisfy [\(KKT\( \$\beta\$ \)\)](#). In particular,  $|p| \leq 2c_0$  holds a.e. on  $\Omega$  if  $m(\Omega_w \cap \beta) = 0$ .

*Proof.* Again, we have  $p = F_w + \lambda \leq F_w \leq c_0$  a.e. on  $\Omega$  and  $\langle 1, p \rangle \leq c_0$ . We also have the lower bounds

$$\begin{aligned} p &= -F_u - \mu - \langle 1, p \rangle \geq -F_u - c_0 \geq -2c_0 \quad \text{a.e. on } \Omega^{+0} \cup (\Omega^{00} \setminus \beta), \\ p &= F_\xi + \nu = F_\xi \geq -2c_0 \quad \text{a.e. on } \Omega^{0+}, \\ p &= F_w + \lambda = F_w \geq -2c_0 \quad \text{a.e. on } \Omega \setminus \Omega_w. \end{aligned}$$

The result then follows from  $\Omega^{+0} \cup (\Omega^{00} \setminus \beta) \cup \Omega^{0+} \cup (\Omega \setminus \Omega_w) = \Omega \setminus (\Omega_w \cap \beta)$ .

**Lemma 4.5.** Suppose that [Assumption 4.1\(b\)](#) holds. Additionally, we assume  $m(\Omega^{+0}) > 0$  and  $m(\Omega_w \cap \beta) > 0$ . Let  $p_0, \mu_0, \nu_0, \lambda_0$  be multipliers that satisfy [\(KKT\( \$\beta\$ \)\)](#). Then one can find multipliers  $p_1, \mu_1, \nu_1, \lambda_1$  satisfying [\(KKT\( \$\beta\$ \)\)](#) such that

$$|p_1| \leq \left(2 + m(\Omega_w \cap \beta)^{-1}(2 + m(\Omega^{+0})^{-1})\right)c_0 \quad \text{a.e. on } \Omega.$$

*Proof.* We define the function

$$a_0 := \chi_{\Omega_w \cap \beta} \max(p_0, \lambda_0, \nu_0)^- \geq 0$$

and subsequently define

$$p_2 := p_0 + a_0, \quad \lambda_2 := \lambda_0 + a_0, \quad \nu_2 := \nu_0 + a_0.$$

We observe that the conditions

$$\begin{aligned} F_w - p_2 + \lambda_2 &= 0, & F_\xi - p_2 + \nu_2 &= 0, \\ \nu_2 &\leq 0 \quad \text{a.e. on } \beta, & \lambda_2 &\leq 0 \quad \text{a.e. on } \Omega \end{aligned}$$

still hold for these new multipliers. Next, we want to find a bound on  $|p_2|$ . We already have the bound  $|p_2| = |p_0| \leq 2c_0$  a.e. on  $\Omega \setminus (\Omega_w \cap \beta)$  due to [Lemma 4.4](#). Again,  $p_2 = F_w + \lambda_2 \leq F_2 \leq c_0$  yields an upper bound. We also have the lower bounds

$$\begin{aligned} p_2 &\geq 0 && \text{a.e. on } \{p_2 \geq 0\} \cap \Omega_w \cap \beta, \\ p_2 = F_w + \lambda_2 &= F_w && \text{a.e. on } \{\lambda_2 = 0\} \cap \Omega_w \cap \beta, \\ p_2 = \nu_2 + F_\xi &\geq F_\xi && \text{a.e. on } \{\nu_2 \geq 0\} \cap \Omega_w \cap \beta. \end{aligned}$$

Using the definition of  $a_0$ , one can show that  $\Omega_w \cap \beta \subset \{p_2 \geq 0\} \cup \{\lambda_2 = 0\} \cup \{\nu_2 \geq 0\}$  holds, which leads to the lower bound  $p_2 \geq -c_0$  a.e. on  $\Omega_w \cap \beta$ . In summary, we have

$$|p_2| \leq 2c_0 \quad \text{a.e. on } \Omega.$$

In general,  $\mu_2 := \mu_0 - A^*a_0$  does not satisfy  $\mu_2 = 0$  a.e. on  $\Omega^{+0}$ . Thus, we need to make further modifications. For this purpose, we first define

$$a_2 := -m(\Omega_w \cap \beta)^{-1} \langle 1, a_0 \rangle \chi_{\Omega_w \cap \beta} \leq 0.$$

Then we have  $\langle 1, a_2 \rangle = -\langle 1, a_0 \rangle$  and

$$\begin{aligned} |a_2| &= -a_2 \leq m(\Omega_w \cap \beta)^{-1} \langle 1, a_0 \rangle \leq m(\Omega_w \cap \beta)^{-1} \langle 1, p_0^- \rangle \\ &\leq m(\Omega_w \cap \beta)^{-1} (2 + m(\Omega^{+0})^{-1}) c_0, \end{aligned}$$

where we used [Lemma 4.3](#) in the last inequality. Finally, we define

$$p_1 := p_2 + a_2, \quad \nu_1 := \nu_2 + a_2, \quad \lambda_1 := \lambda_2 + a_2, \quad \mu_1 := \mu_0 - A^*a_0 - A^*a_2.$$

We claim that these functions satisfy  $(\text{KKT}(\beta))$ . That the three equations hold can be seen directly. For  $\mu_1$  we observe that

$$\mu_1 = \mu_0 - a_0 - a_2 - \langle 1, a_0 + a_2 \rangle = \mu_0 - a_0 - a_2$$

holds. In particular, we have  $\mu_1 = \mu_0$  a.e. on  $\Omega^{+0} \cup (\Omega^{00} \setminus \beta)$  and thus the conditions for  $\mu_1$  in  $(\text{KKT}(\beta))$  are satisfied. For  $\nu_1$ , we observe  $\nu_1 = \nu_2 + a_2 \leq \nu_2 \leq 0$  a.e. on  $\beta$ , and because of  $a_0 = a_2 = 0$  a.e. on  $\Omega^{0+}$ , we also have  $\nu_1 = 0$  a.e. on  $\Omega^{0+}$ . Similarly, for  $\lambda_1$ , we observe  $\lambda_1 = \lambda_2 + a_2 \leq \lambda_2 \leq 0$  a.e. on  $\Omega_w$ , and because of  $a_0 = a_2 = 0$  a.e. on  $\Omega \setminus \Omega_w$  we also have  $\lambda_1 = 0$  a.e. on  $\Omega \setminus \Omega_w$ .

To summarize, the multipliers  $p_1, \mu_1, \nu_1, \lambda_1$  satisfy  $(\text{KKT}(\beta))$ . The estimate for  $p_1$  follows by combining the estimates for  $p_2$  and  $a_2$ .

Finally, we can make an analogous statement to [Lemma 4.2](#), but for the case of [Assumption 4.1\(b\)](#). A notable difference is that the bound depends on  $\beta$ .

**Lemma 4.6.** Suppose that [Assumption 4.1\(b\)](#) holds. Let  $p_0, \mu_0, \nu_0, \lambda_0$  be multipliers that satisfy  $(\text{KKT}(\beta))$ . Then there exist multipliers  $p_1, \mu_1, \nu_1, \lambda_1$  that satisfy  $(\text{KKT}(\beta))$

such that the pointwise a.e. inequality

$$\max(|p_1|, |\nu_1|, |\lambda_1|, |\mu_1|) \leq C_\beta c_0$$

holds, where the constant  $C_\beta > 0$  can be chosen as  $C_\beta = 2$  if  $m(\Omega^{+0}) = 0$ ,  $C_\beta = 5 + 2m(\Omega)$  if  $m(\Omega_w \cap \beta) = 0$ , and  $C_\beta = 1 + (2 + m(\Omega_w \cap \beta))^{-1}(2 + m(\Omega^{+0}))^{-1}(2 + m(\Omega))$  if  $m(\Omega^{+0}) > 0$  and  $m(\Omega_w \cap \beta) > 0$ .

*Proof.* If  $m(\Omega^{+0}) = 0$ , then the claim follows from [Lemma 4.2](#). If  $m(\Omega_w \cap \beta) = 0$ , then we choose  $(p_1, \mu_1, \nu_1, \lambda_1) := (p_0, \mu_0, \nu_0, \lambda_0)$ , and [Lemma 4.4](#) yields a pointwise a.e. estimate for  $|p_1|$ . In the remaining case of  $m(\Omega^{+0}) > 0$  and  $m(\Omega_w \cap \beta) > 0$  we apply [Lemma 4.5](#), which also yields an estimate for  $|p_1|$ . In any case, we have an estimate of the form  $|p_1| \leq \hat{C}_\beta c_0$  a.e. in  $\Omega$ , where  $\hat{C}_\beta := 2 + m(\Omega_w \cap \beta)^{-1}(2 + m(\Omega^{+0}))^{-1}$  if  $m(\Omega_w \cap \beta) > 0$  or  $\hat{C}_\beta := 2$  if  $m(\Omega_w \cap \beta) = 0$ . From [\(KKT\( \$\beta\$ \)\)](#), we can obtain the estimates  $\max(|\nu_1|, |\lambda_1|) \leq (\hat{C}_\beta + 1)c_0$  a.e. in  $\Omega$ . Finally, we have

$$\begin{aligned} |\mu_1| &= |-F_u - A^*p_1| \leq |F_u| + |A^*p_1| \leq c_0 + |p_1| + |\langle 1, p_1 \rangle| \\ &\leq (1 + \hat{C}_\beta)c_0 + \langle 1, p_1 \rangle \leq c_0 + \hat{C}_\beta c_0 + \hat{C}_\beta \langle 1, c_0 \rangle \leq c_0 + \hat{C}_\beta c_0 + \hat{C}_\beta (2 + m(\Omega))c_0 \\ &= (1 + \hat{C}_\beta(2 + m(\Omega)))c_0, \end{aligned}$$

where the estimate  $\langle 1, c_0 \rangle \leq (2 + m(\Omega))c_0$  that was used in the last inequality can be obtained from the definition of  $c_0$ .

These lemmas enable us to verify the existence of KKT multipliers if  $(0, 0, 0)$  is a minimizer of [\(LP\( \$\beta\$ \)\)](#).

**Proposition 4.7.** Suppose that [Assumption 4.1\(a\)](#) or [Assumption 4.1\(b\)](#) holds. Further suppose that  $(0, 0, 0)$  is a minimizer of [\(LP\( \$\beta\$ \)\)](#). Then there exist multipliers  $(p, \mu, \nu, \lambda)$  that satisfy [\(KKT\( \$\beta\$ \)\)](#).

*Proof.* We introduce the notation  $T: L^2(\Omega)^3 \rightarrow L^2(\Omega)$ ,  $T(u, w, \xi) := Au - w - \xi$ ,

$$P := \left\{ (u, w, \xi) \in L^2(\Omega) \left| \begin{array}{ll} u \geq 0 \text{ a.e. on } \Omega^{00} \setminus \beta, & u = 0 \text{ a.e. on } \Omega^{0+} \cup \beta, \\ w \geq 0 \text{ a.e. on } \Omega_w, & \\ \xi \geq 0 \text{ a.e. on } \beta, & \xi = 0 \text{ a.e. on } \Omega^{+0} \cup (\Omega^{00} \setminus \beta) \end{array} \right. \right\}$$

and  $F := (F_u, F_w, F_\xi) \in L^2(\Omega)^3$ . Then, [\(LP\( \$\beta\$ \)\)](#) reads

$$\text{Minimize } \langle F, x \rangle \text{ such that } Tx = 0, x \in P.$$

The optimality of  $x = (0, 0, 0)$  yields

$$-F \in (T^{-1}(\{0\}) \cap P)^\circ,$$

while  $(\text{KKT}(\beta))$  reads

$$-F \in T^*L^2(\Omega) + P^\circ.$$

Thus, it remains to show that the right-hand sides are equal and this holds if (and only if) the set

$$B := T^*L^2(\Omega) + P^\circ = \left\{ \left( \begin{array}{l} A^*p + \mu \\ -p + \lambda \\ -p + \nu \end{array} \right) \left| \begin{array}{ll} \mu \leq 0 \text{ a.e. on } \Omega^{00} \setminus \beta, & \mu = 0 \text{ a.e. on } \Omega^{+0}, \\ \lambda \leq 0 \text{ a.e. on } \Omega_w, & \lambda = 0 \text{ a.e. on } \Omega \setminus \Omega_w, \\ \nu \leq 0 \text{ a.e. on } \beta, & \nu = 0 \text{ a.e. on } \Omega^{0+} \end{array} \right. \right\}$$

is closed, see [Schlotzke, 2007, Lemma 2.4.1].

Let  $\{F_k\}_k = \{(F_{u,k}, F_{w,k}, F_{\xi,k})\}_k \subset -B \subset L^2(\Omega)^3$  be a sequence with  $F_k \rightarrow F_0$  in  $L^2(\Omega)^3$ . Hence, there exists a sequence  $\{(p_k, \lambda_k, \nu_k, \mu_k)\}_k \subset L^2(\Omega) \times P^\circ$  such that  $(p_k, \lambda_k, \nu_k, \mu_k)$  satisfies  $(\text{KKT}(\beta))$  with  $(F_u, F_w, F_\xi)$  replaced by  $(F_{u,k}, F_{w,k}, F_{\xi,k})$ . Due to Assumption 4.1, we can apply Lemma 4.2 or Lemma 4.6 to obtain  $(\hat{p}_k, \hat{\lambda}_k, \hat{\nu}_k, \hat{\mu}_k) \in L^2(\Omega) \times P^\circ$  with

$$F_k + T^*\hat{p}_k + (\hat{\lambda}_k, \hat{\nu}_k, \hat{\mu}_k) = 0 \tag{4.2}$$

and  $\max(|\hat{p}_k|, |\hat{\nu}_k|, |\hat{\lambda}_k|, |\hat{\mu}_k|) \leq C_\beta c_{0,k}$ , where  $c_{0,k}$  is defined analogously to (4.1). Note that  $c_{0,k}$  is bounded in  $L^2(\Omega)$  due to the boundedness of  $\{F_k\}_k$  in  $L^2(\Omega)^3$ . Therefore, the sequence  $\{(\hat{p}_k, \hat{\lambda}_k, \hat{\nu}_k, \hat{\mu}_k)\}_k$  has a weakly convergent subsequence. By passing to the limit (along this subsequence) in (4.2), it follows that  $F_0 \in -B$  due to the weak closeness of  $P^\circ$ . Thus,  $B$  is closed and this yields the claim.

Finally, we give an example to show that KKT multipliers for the local minimizer  $(0, 0, 0)$  of  $(\text{LP}(\beta))$  may fail to exist.

**Example 4.8.** We consider the case that  $\Omega = (-1, 1)$  and  $S := (-\Delta_0)^{-1}: L^2(\Omega) \rightarrow L^2(\Omega)$ , where  $\Delta_0$  is the Laplacian equipped with homogeneous Dirichlet boundary conditions. We further choose

$$\begin{aligned} A &:= \text{id} + S^*S: L^2(\Omega) \rightarrow L^2(\Omega), \\ \Omega_1 &:= \bigcup_{k \in \mathbb{N}} [2^{-2k-1}, 2^{-2k}], \\ \Omega_2 &:= \Omega \setminus \Omega_1 = (-1, 0] \cup \bigcup_{k \in \mathbb{N}} (2^{-2k}, 2^{-2k+1}) \cup [1/2, 1), \\ \Omega_w &:= \Omega_1, \quad \Omega^{+0} := \Omega_2, \quad \Omega^{00} := \beta := \Omega_1, \quad \Omega^{0+} := \emptyset, \\ F_w &:= F_\xi := 0, \quad F_u(\omega) := \chi_{\Omega_2}(\frac{1}{6}|\omega|^3 - \frac{1}{2}\omega^2 + \frac{1}{3}) \quad \forall \omega \in \Omega. \end{aligned}$$

Then  $(0, 0, 0)$  is a (global) minimizer of  $(\text{LP}(\beta))$ , but no multipliers exist such that  $(\text{KKT}(\beta))$  is satisfied.

Moreover, the set of functions  $F_u, F_w, F_\xi$  for which multipliers exist such that  $(\text{KKT}(\beta))$

is satisfied, is nonempty and not closed.

*Proof.* We start with showing that the set of functions  $F_u, F_w, F_\xi$  for which multipliers exist such that  $(\text{KKT}(\beta))$  is satisfied is nonempty and not closed. For  $k \in \mathbb{N}$ , we define

$$\begin{aligned} p_k &:= -2^{2k+2} \chi_{(2^{-2k-1}, 2^{-2k})}, \\ \lambda_k &:= \nu_k := p_k, \\ \mu_k &:= \chi_{\Omega_1}(-A^* p_k), \\ F_u^k &:= \chi_{\Omega_2}(-A^* p_k). \end{aligned}$$

Then one can check that  $(\text{KKT}(\beta))$  is satisfied (by using  $p_k, \lambda_k, \nu_k, \mu_k, F_u^k$  instead of  $p, \lambda, \nu, \mu, F_u$ ). One can also calculate  $Sp_k$  and confirm that it satisfies

$$v_k(\omega) := (Sp_k)(\omega) = \begin{cases} c_k(\omega + 1) & \text{if } \omega \in (-1, 2^{-2k-1}], \\ c_k + 2^{-2k-1} + (c_k - 2)\omega + 2^{2k+1}\omega^2 & \text{if } \omega \in (2^{-2k-1}, 2^{-2k}), \\ (c_k + 2)(\omega - 1) & \text{if } \omega \in [2^{-2k}, 1), \end{cases}$$

where  $c_k = -1 + 3 \cdot 2^{-2k-2}$ . We also observe that  $v_k$  converges in  $L^2(\Omega)$  with limit  $v_0(\omega) := |\omega| - 1$ , and that  $(S^* v_0)(\omega) = (Sv_0)(\omega) = -(\frac{1}{6}|\omega|^3 - \frac{1}{2}\omega^2 + \frac{1}{3})$  holds. Thus, we have the convergence

$$\begin{aligned} F_u^k &= \chi_{\Omega_2}(-A^* p_k) = \chi_{\Omega_2}(-p_k - S^* Sp_k) = \chi_{\Omega_2}(-S^* Sp_k) = \chi_{\Omega_2}(-S^* v_k) \\ &\rightarrow \chi_{\Omega_2}(-S^* v_0) = F_u \quad \text{as } k \rightarrow \infty \end{aligned}$$

in  $L^2(\Omega)$ . Suppose that  $p, \mu, \nu, \lambda \in L^2(\Omega)$  satisfy  $(\text{KKT}(\beta))$ . Then we have  $p = \lambda = \nu$ ,  $p = 0$  a.e. on  $\Omega_2$ , and  $F_u + A^* p = -\mu = 0$  a.e. on  $\Omega_2$ . Thus,

$$0 = F_u + A^* p = -S^* v_0 + S^* Sp = S^*(Sp - v_0) \quad \text{a.e. on } \Omega_2.$$

Since  $Sp$  and  $v_0$  are continuous, taking second derivatives in  $\Omega_2$  in the above equation yields  $Sp = v_0$  on  $\Omega_2$ . We also have equality of derivatives of  $Sp$  and  $v_0$  in the interior of  $\Omega_2$ . This shows that the derivative of  $Sp$  is discontinuous at 0, which is a contradiction to its regularity  $(Sp)' \in H^1(\Omega) \hookrightarrow C(\bar{\Omega})$ . Thus, there exist no multipliers that satisfy  $(\text{KKT}(\beta))$  with our choices for  $F_u, F_w, F_\xi$ . Due to  $F_u^k \rightarrow F_u$ , the set of functions  $F_u, F_w, F_\xi$  for which multipliers exist is nonempty, but not closed.

It remains to show that  $(0, 0, 0)$  is a minimizer of  $(\text{LP}(\beta))$ . Let  $(u, w, \xi) \in L^2(\Omega)^3$  be another feasible point of  $(\text{LP}(\beta))$ . Note that  $u = 0$  a.e. on  $\Omega_1$  and  $Au = w + \xi \geq 0$  a.e.

on  $\Omega_1$ . For the objective function we have

$$\begin{aligned} \langle F_u, u \rangle + \langle F_w, w \rangle + \langle F_\xi, \xi \rangle &= \langle F_u, u \rangle = \lim_{k \rightarrow \infty} \langle F_u^k, u \rangle = \lim_{k \rightarrow \infty} \langle \chi_{\Omega_2}(-A^* p_k), u \rangle \\ &= \lim_{k \rightarrow \infty} \langle -p_k, A \chi_{\Omega_2} u \rangle = \lim_{k \rightarrow \infty} \langle -\chi_{\Omega_1} p_k, \chi_{\Omega_1} A u \rangle \\ &\geq 0 = \langle F_u, 0 \rangle + \langle F_w, 0 \rangle + \langle F_\xi, 0 \rangle \end{aligned}$$

where we used  $-p_k \geq 0$  and  $\chi_{\Omega_1} A u \geq 0$  a.e. on  $\Omega$  in the last step. Since  $(u, w, \xi) \in L^2(\Omega)^3$  was an arbitrary feasible point of  $(\text{LP}(\beta))$ ,  $(0, 0, 0)$  is a (global) minimizer of this optimization problem.

## 4.2 M-stationarity

We combine the results of [Section 3](#) with those of [Section 4.1](#) to show that  $(0, 0, 0)$  is M-stationary if it is a minimizer of  $(\text{MPCC}_{\text{lin}})$ . Using [Definition 2.1](#), the system of M-stationarity for  $(\text{MPCC}_{\text{lin}})$  at  $(0, 0, 0)$  is

$$\begin{aligned} F_u + A^* \bar{p} + \bar{\mu} &= 0 & F_w - \bar{p} + \bar{\lambda} &= 0 & F_\xi - \bar{p} + \bar{\nu} &= 0, \\ \bar{\lambda} &\leq 0 \quad \text{a.e. on } \Omega_w, & \bar{\lambda} &= 0 \quad \text{a.e. on } \Omega \setminus \Omega_w, \\ \bar{\mu} &= 0 \quad \text{a.e. on } \Omega^{+0}, & \bar{\nu} &= 0 \quad \text{a.e. on } \Omega^{0+}, \\ (\bar{\mu} < 0 \wedge \bar{\nu} < 0) &\vee \bar{\mu} \bar{\nu} = 0 & & \text{a.e. on } \Omega^{00}, \end{aligned} \tag{4.3}$$

where  $\bar{p}, \bar{\nu}, \bar{\lambda}, \bar{\mu} \in L^2(\Omega)$  are the multipliers. First, we check that M-stationarity holds under [Assumption 4.1\(a\)](#) for a local minimizer  $(0, 0, 0)$ .

**Theorem 4.9.** Suppose that [Assumption 4.1\(a\)](#) holds. If  $(0, 0, 0)$  is a (global) minimizer of  $(\text{MPCC}_{\text{lin}})$ , then it is an M-stationary point.

*Proof.* For each measurable  $\beta \subset \Omega^{00}$ ,  $(0, 0, 0)$  is also a minimizer of  $(\text{LP}(\beta))$  because this problem has a smaller feasible set. By [Proposition 4.7](#) there exist multipliers  $(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta) \in L^2(\Omega)^4$  that satisfy  $(\text{KKT}(\beta))$ .

We want to apply [Theorem 3.1](#) on  $\Omega^{00}$  (instead of  $\Omega$ ). By [Lemma 4.2](#), we can choose  $(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta) \in L^2(\Omega)^4$  in such a way that there is a common  $L^2(\Omega)$ -function that bounds these multipliers. This implies that [\(3.2a\)](#) and [\(3.2b\)](#) are satisfied. Applying [Theorem 3.1](#) yields a point  $(\bar{\mu}^{00}, \bar{\nu}^{00}) \in \overline{\text{conv}}\{(\chi_{\Omega^{00}} \mu^\beta, \chi_{\Omega^{00}} \nu^\beta) \mid \beta \subset \Omega^{00} \text{ measurable}\}$  which satisfies  $(\bar{\mu}^{00} < 0 \wedge \bar{\nu}^{00} < 0) \vee \bar{\mu}^{00} \bar{\nu}^{00} = 0$  a.e. on  $\Omega^{00}$ . By the definition of  $\overline{\text{conv}}$ , there exists a sequence  $\{(\mu_k^{00}, \nu_k^{00})\}_{k \in \mathbb{N}} \subset \text{conv}\{(\chi_{\Omega^{00}} \mu^\beta, \chi_{\Omega^{00}} \nu^\beta) \mid \beta \subset \Omega^{00} \text{ measurable}\}$  with  $\mu_k^{00} \rightarrow \bar{\mu}^{00}$ ,  $\nu_k^{00} \rightarrow \bar{\nu}^{00}$ . Using the same convex combination, one can find multipliers  $(p_k, \mu_k, \nu_k, \lambda_k) \in \text{conv}\{(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta) \mid \beta \subset \Omega^{00} \text{ measurable}\}$  with  $\mu_k^{00} = \chi_{\Omega^{00}} \mu_k$  and  $\nu_k^{00} = \chi_{\Omega^{00}} \nu_k$  for all  $k \in \mathbb{N}$ . Due to convexity, these multipliers satisfy the conditions of  $(\text{KKT}(\beta))$  that do not depend on  $\beta$ . As  $(p_k, \mu_k, \nu_k, \lambda_k)$  is bounded, there exists a weakly convergent subsequence whose weak limit we will denote by  $(\bar{p}, \bar{\mu}, \bar{\nu}, \bar{\lambda})$ . These multipliers

also satisfy the conditions of  $(\text{KKT}(\beta))$  that do not depend on  $\beta$ . It remains to show that the last condition of (4.3) holds for  $\bar{\mu}$  and  $\bar{\nu}$ . This, however, is true due to  $\bar{\mu} = \bar{\mu}^{00}$  and  $\bar{\nu} = \bar{\nu}^{00}$  a.e. on  $\Omega^{00}$ .

For the case of Assumption 4.1(b), the proof is a bit more complicated, as it is based on Lemma 3.8, which has other requirements than Theorem 3.1.

**Theorem 4.10.** Suppose that Assumption 4.1(b) or Assumption 4.1(c) holds. If  $(0, 0, 0)$  is a (global) minimizer of  $(\text{MPCC}_{\text{lin}})$ , then it is an M-stationary point.

*Proof.* We start with the case that Assumption 4.1(b) holds. For each measurable  $\beta \subset \Omega^{00}$ ,  $(0, 0, 0)$  is also a minimizer of  $(\text{LP}(\beta))$  because it has a smaller feasible set. By Proposition 4.7 there exist multipliers  $(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta) \in L^2(\Omega)^4$  that satisfy  $(\text{KKT}(\beta))$ .

Let us argue that it is possible to choose the multipliers  $(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta)$  in such a way that they satisfy  $(\text{KKT}(\beta))$  and

$$\langle 1, |p^\beta| \rangle \leq Cc_0 \tag{4.4}$$

$$|\nu^\beta| \leq C_\beta c_0 \tag{4.5}$$

hold a.e. in  $\Omega$ , where  $C > 0$  is a constant which does not depend on  $\beta$  and  $C_\beta > 0$  is a constant which can depend on  $\beta$ .

Indeed, if  $\Omega^{+0}$  has measure zero, then we can use Lemma 4.2 to redefine the multipliers  $(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta)$  in such a way that  $\max(|p^\beta|, |\nu^\beta|) \leq 2c_0$  is satisfied. In this case (4.4) is satisfied with  $C = 2m(\Omega)$  and (4.5) is satisfied with  $C_\beta = 2$ . If, on the other hand,  $\Omega^{+0}$  has positive measure, then we can use Lemma 4.6 to redefine the multipliers  $(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta)$  such that (4.5) is satisfied. In this case, (4.4) follows from Lemma 4.3.

We want to apply Lemma 3.8 on  $\Omega^{00}$  (instead of  $\Omega$ ), where we consider the restrictions of  $\mu^\beta, \nu^\beta$  on  $\Omega^{00}$ , and with

$$\rho^\beta := (p^\beta, \chi_{\Omega^{0+}}\mu^\beta, \chi_{\Omega^{0+}}\nu^\beta, \lambda^\beta) \in L^2(\Omega) \times L^2(\Omega^{0+}) \times L^2(\Omega^{+0}) \times L^2(\Omega) =: Y.$$

The conditions (3.8) follow directly from  $(\text{KKT}(\beta))$ . Next, we will show

$$\max(|\mu^\beta + \nu^\beta|, |\nu^\beta|) \leq Cc_0 \tag{4.6}$$

for all measurable  $\beta \subset \Omega^{00}$ , where  $C > 0$  is a constant which does not depend on  $\beta$ . From  $(\text{KKT}(\beta))$  we obtain

$$\nu^\beta = -F_\xi + p^\beta = -F_\xi + F_w + \lambda^\beta \leq -F_\xi + F_w \leq c_0 \quad \text{a.e. on } \Omega.$$

From  $(\text{KKT}(\beta))$  we also obtain

$$\mu^\beta + \nu^\beta = -F_\xi + p^\beta - F_u - A^*p^\beta = -F_\xi + -F_u - \langle 1, p^\beta \rangle.$$

This implies  $|\mu^\beta + \nu^\beta| \leq c_0 + \langle 1, |p^\beta| \rangle \leq (1 + C)c_0$  a.e. in  $\Omega$  due to (4.4). In summary, (4.6) holds. The assumptions (3.9) and (3.10) follow from (4.5) and (4.6) with suitable constants.

Next, we will show (3.11) with the bounded operator  $T: L^2(\Omega^{00}) \rightarrow Y$  given by  $T\nu = (\nu, 0, 0, \nu)$ . We observe that we have  $|p^\beta| \leq 2c_0$  a.e. on  $\Omega \setminus \Omega^{00}$  by Lemma 4.4. This can be combined with (4.4) to yield  $|A^*p^\beta| \leq |p^\beta| + \langle 1, |p^\beta| \rangle \leq (2 + C)c_0$  a.e. on  $\Omega \setminus \Omega^{00}$ . Using these properties and (KKT( $\beta$ )) one can obtain estimates of the form

$$\|p^\beta\|_{L^2(\Omega \setminus \Omega^{00})} + \|\nu^\beta\|_{L^2(\Omega \setminus \Omega^{00})} + \|\mu^\beta\|_{L^2(\Omega \setminus \Omega^{00})} + \|\lambda^\beta\|_{L^2(\Omega \setminus \Omega^{00})} \leq (11 + C)\|c_0\|_{L^2(\Omega)}.$$

Then (3.11) follows by the estimates

$$\begin{aligned} \|\rho^\beta - T(\chi_{\Omega^{00}}\nu^\beta)\|_Y &= \|p^\beta - \chi_{\Omega^{00}}\nu^\beta\|_{L^2(\Omega)} + \|\mu^\beta\|_{L^2(\Omega^{0+})} \\ &\quad + \|\nu^\beta\|_{L^2(\Omega^{0+})} + \|\lambda^\beta - \chi_{\Omega^{00}}\nu^\beta\|_{L^2(\Omega)} \\ &\leq \|p^\beta\|_{L^2(\Omega \setminus \Omega^{00})} + \|p^\beta - \nu^\beta\|_{L^2(\Omega^{00})} + \|\mu^\beta\|_{L^2(\Omega^{0+})} \\ &\quad + \|\nu^\beta\|_{L^2(\Omega^{0+})} + \|\lambda^\beta\|_{L^2(\Omega \setminus \Omega^{00})} + \|\lambda^\beta - \nu^\beta\|_{L^2(\Omega^{00})} \\ &\leq (11 + C)\|c_0\|_{L^2(\Omega)} + \|F_\xi\|_{L^2(\Omega^{00})} + \|p^\beta - F_w - p^\beta + F_\xi\|_{L^2(\Omega^{00})} \\ &\leq (14 + C)\|c_0\|_{L^2(\Omega)}. \end{aligned}$$

In summary, Lemma 3.8 can be applied. Thus, there exists a point  $(\bar{p}, \bar{\mu}, \bar{\nu}, \bar{\lambda})$  in the closure of the convex hull of  $\{(p^\beta, \mu^\beta, \nu^\beta, \lambda^\beta) \mid \beta \subset \Omega^{00} \text{ measurable}\}$  which satisfies (3.3). Due to convexity, the other conditions of (4.3) follow from (KKT( $\beta$ )).

For the case of Assumption 4.1(c), the result follows from the result for the case of Assumption 4.1(c) by rescaling. First, one can scale  $A$  (and rescale the data, variables and multipliers accordingly) such that the result applies to the case that  $A$  is given by  $Av = d_1v + d_1\langle 1, v \rangle$  for all  $v \in L^2(\Omega)$ . Then, one can scale the measure  $m$  by  $d_2/d_1$  to allow for an operator  $A$  with  $Av = d_1v + d_2\langle 1, v \rangle$  for all  $v \in L^2(\Omega)$ .

## 5 An inverse optimal control problem

In this section, we consider a class of inverse optimal control problems. We will see that our M-stationarity results from Sections 3 and 4 can be applied to the KKT reformulation of this inverse optimal control problem under some assumptions.



### 5.1 Problem setting and existence of solutions

Inverse optimal control problems can be modeled as bilevel optimization problems. We are interested in the following problem class. The upper-level problem is given by

$$\begin{aligned} \min_{u \in L^2(\Omega), w \in H_0^1(\Omega)} \quad & F(u, w) := f(u) + \frac{1}{2} \|w\|_{H_0^1(\Omega)}^2 + \langle \zeta, w \rangle_{L^2(\Omega)} \\ \text{s.t.} \quad & u \text{ solves } \text{(OC}(w)), \\ & w \geq w_a \quad \text{a.e. on } \Omega, \end{aligned} \tag{IOC}$$

while for each  $w \in L^2(\Omega)$  the lower-level problem is the parametrized optimal control problem

$$\begin{aligned} \min_{u \in L^2(\Omega)} \quad & \frac{1}{2} \|Su - y_d\|_Y^2 + \frac{\alpha}{2} \|u - w\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad & u \geq u_a \quad \text{a.e. on } \Omega. \end{aligned} \tag{OC}(w)$$

We fix the assumptions concerning the data in both problems. Throughout this section,  $\Omega \subset \mathbb{R}^n$  is open and bounded (and equipped with the Lebesgue measure),  $n \in \mathbb{N}$ ,  $Y$  is a (real) Hilbert space,  $S \in \mathcal{L}(L^2(\Omega), Y)$ ,  $\alpha > 0$ ,  $u_a \in L^2(\Omega)$ . For the upper-level problem, we assume  $w_a \in H^1(\Omega)$  such that  $-\Delta w_a \in L^2(\Omega)$  and  $w_a \leq 0$  on  $\partial\Omega$  in the sense  $\max(w_a, 0) \in H_0^1(\Omega)$ ,  $\zeta \in L^2(\Omega)$  and  $f: L^2(\Omega) \rightarrow \mathbb{R}$  is assumed to be continuously Fréchet differentiable and is bounded from below. We further require some regularity, i.e., we assume that there exists  $p > 2$  such that  $S^*S \in \mathcal{L}(L^2(\Omega), L^p(\Omega))$ ,  $S^*y_d, u_a \in L^p(\Omega)$  and  $H_0^1(\Omega) \hookrightarrow L^p(\Omega)$ . Finally, we mention that we use

$$\|w\|_{H_0^1(\Omega)} := \|\nabla w\|_{L^2(\Omega)^n},$$

i.e., we equip  $H_0^1(\Omega)$  with the  $H^1(\Omega)$ -seminorm, and the associated inner product is denoted by

$$\langle v, w \rangle_{H_0^1(\Omega)} := \langle \nabla v, \nabla w \rangle_{L^2(\Omega)^n}.$$

We note that the upper-level variable  $w$  enters the lower-level problem  $\text{(OC}(w))$  as the desired control (sometimes also denoted reference control).

We define the lower-level and upper-level feasible sets by

$$U_{\text{ad}} := \{v \in L^2(\Omega) \mid v \geq u_a \text{ a.e.}\}, \quad W_{\text{ad}} := \{v \in H_0^1(\Omega) \mid v \geq w_a \text{ a.e.}\},$$

respectively. It is well known that for fixed  $w \in L^2(\Omega)$ , the unique solution  $u \in U_{\text{ad}}$  of  $\text{(OC}(w))$  is characterized by the variational inequality (VI)

$$\langle S^*(Su - y_d) + \alpha(u - w), v - u \rangle_{L^2(\Omega)} \geq 0 \quad \forall v \in U_{\text{ad}}. \tag{5.1}$$

For later reference, the solution operator (mapping  $w \mapsto u$ ) of  $\text{(OC}(w))$  (or, equivalently, (5.1)) is denoted by  $T: L^2(\Omega) \rightarrow L^2(\Omega)$ . Using the continuity of  $T$  and the compact embedding  $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ , see [Troianiello, 1987, Theorem 1.34], we can show the existence of solutions of (IOC).

**Lemma 5.1.** The inverse optimal control problem (IOC) possesses a global solution.

*Proof.* The solution mapping  $w \mapsto u$  of the lower-level problem (OC( $w$ )) is continuous from  $L^2(\Omega)$  to  $L^2(\Omega)$ , which can be seen from the projection formula (5.4). Since the objective of (IOC) is coercive w.r.t.  $w \in H_0^1(\Omega)$ , existence of solutions follows by the direct method of calculus of variations.

Due to the non-convexity of (IOC), one cannot expect uniqueness of solutions.

## 5.2 Regularity of solutions

In this section, we show that every locally optimal  $\bar{w} \in H_0^1(\Omega)$  of (IOC) enjoys the additional regularity  $-\Delta \bar{w} \in L^2(\Omega)$ . This will become important for a density argument in the B-stationarity system, see Lemma 5.6 below.

We mention that a similar regularity effect is well known for the optimal control of the obstacle problem. Using a regularization approach, one can derive a system of C-stationarity and this can be used to show  $H^1$ -regularity of the control using the projection formula. We further note, that such a regularization procedure is not needed in the unconstrained setting, since the additional regularity follows directly from the system of B-stationarity, see [Mignot, 1976, Proof of Proposition 4.1], [Wachsmuth, 2019, Proof of Theorem 5.3].

We also mention that we do not need to derive a system of C-stationarity in its full glory, since the additional regularity of  $\bar{w}$  is sufficient in the next section.

We start by regularizing the VI (5.1). To this end, we introduce the monotone penalty function  $\pi: \mathbb{R} \rightarrow \mathbb{R}$  via

$$\pi(s) := \begin{cases} s + \frac{1}{2} & \text{if } s \leq -1, \\ -\frac{1}{2}s^2 & \text{if } -1 < s < 0, \\ 0 & \text{if } s \geq 0. \end{cases}$$

Note that  $\pi$  is continuously differentiable with a bounded derivative. The Nemytskii operators associated with  $\pi$  and  $\pi'$  are denoted by the same symbols, respectively. We approximate (5.1) by the equation

$$e_\gamma(u, w) := S^*(Su - y_d) + \alpha(u - w) + \gamma\pi(u - u_a) = 0, \quad (\text{OC}_\gamma(w))$$

where  $\gamma > 0$  is the regularization parameter. The first lemma provides the differentiability of this regularization.

**Lemma 5.2.** For every  $w \in L^p(\Omega)$ , there exists a unique solution  $u =: T_\gamma(w) \in L^p(\Omega)$  of (OC $_\gamma(w)$ ). The solution operator  $T_\gamma: L^p(\Omega) \rightarrow L^2(\Omega)$  is Fréchet differentiable. For all  $w, h \in L^p(\Omega)$ , the derivative  $v = T'_\gamma(w)h$  satisfies

$$S^*Sv + \alpha(v - h) + \gamma\pi'(u - u_a)v = 0, \quad (5.2)$$

where  $u = T_\gamma(w)$ .

*Proof.* By monotone operator theory, e.g. [Troianiello, 1987, Theorem 4.17], the existence of a solution  $u \in L^2(\Omega)$  of  $(\text{OC}_\gamma(w))$  for all  $w \in L^2(\Omega)$  follows. Uniqueness and Lipschitz continuity of  $T_\gamma: L^2(\Omega) \rightarrow L^2(\Omega)$  follows by a standard argument: if  $u_i$  is a solution for  $w_i$ ,  $i = 1, 2$ , we get

$$\begin{aligned} 0 &= \langle e_\gamma(u_2, w_2) - e_\gamma(u_1, w_1), u_2 - u_1 \rangle_{L^2(\Omega)} \\ &\geq \alpha \|u_2 - u_1\|_{L^2(\Omega)}^2 - \alpha \langle w_2 - w_1, u_2 - u_1 \rangle_{L^2(\Omega)} \end{aligned}$$

and this yields  $\|u_2 - u_1\|_{L^2(\Omega)} \leq \|w_2 - w_1\|_{L^2(\Omega)}$ .

In case  $w \in L^p(\Omega)$ , the additional regularity of  $u$  follows by a bootstrapping argument: By rearranging the equation, we find

$$u - u_a = (\alpha \text{id} + \gamma\pi)^{-1} (S^*(y_d - Su) + \alpha(w - u_a)),$$

where  $(\alpha \text{id} + \gamma\pi)^{-1}: \mathbb{R} \rightarrow \mathbb{R}$  is the Lipschitz continuous inverse of  $\alpha \text{id} + \gamma\pi: \mathbb{R} \rightarrow \mathbb{R}$ . Due to the assumed regularities  $S^*S \in \mathcal{L}(L^2(\Omega), L^p(\Omega))$ ,  $S^*y_d, u_a \in L^p(\Omega)$ , the Lipschitz regularity  $T_\gamma: L^p(\Omega) \rightarrow L^2(\Omega)$  is improved to the spaces  $T_\gamma: L^p(\Omega) \rightarrow L^p(\Omega)$ .

Using [Goldberg, Kampowsky, Tröltzsch, 1992, Theorem 7], we find that  $e_\gamma: L^p(\Omega) \times L^p(\Omega) \rightarrow L^2(\Omega)$  is Fréchet differentiable and the partial Fréchet derivative  $e_{\gamma,u}(u, w) \in \mathcal{L}(L^p(\Omega), L^2(\Omega))$  is given by

$$e_{\gamma,u}(u, w)v = S^*Sv + \alpha v + \gamma\pi'(u - u_a)v \quad \forall v \in L^p(\Omega).$$

This operator can be extended to  $e_{\gamma,u}(u, w) \in \mathcal{L}(L^2(\Omega), L^2(\Omega))$  and, via the Lemma of Lax–Milgram, this extension is bijective.

Now, we can apply the differentiability result [Wachsmuth, 2014, Theorem 2.1] with the setting

$$Y^+ = U = L^p(\Omega), \quad Y^0 = Z^0 = L^2(\Omega).$$

This shows the Fréchet differentiability of  $T_\gamma: L^p(\Omega) \rightarrow L^2(\Omega)$  and the stated equation for the derivative.

Next, we address the convergence of the regularization.

**Lemma 5.3.** Let  $(w_k)_{k \in \mathbb{N}} \subset L^2(\Omega)$  be a sequence with  $w_k \rightarrow w$  in  $L^2(\Omega)$ . Further, let  $(\gamma_k)_{k \in \mathbb{N}} \subset \mathbb{R}^+$  be given such that  $\gamma_k \rightarrow \infty$ . We denote by  $u_k = T_{\gamma_k}(w_k) \in L^2(\Omega)$  the corresponding unique solution of  $(\text{OC}_{\gamma_k}(w))$ , i.e.,  $e_{\gamma_k}(u_k, w_k) = 0$ . Then,  $u_k \rightarrow u$  in  $L^2(\Omega)$ , where  $u$  solves  $(\text{OC}(w))$ .

*Proof.* We start by multiplying the equation  $e_{\gamma_k}(u_k, w_k) = 0$  by  $u_k - u_a$ . This yields

$$\begin{aligned} & \|Su_k\|_Y^2 + \alpha \|u_k\|_{L^2(\Omega)}^2 + \gamma_k \langle \pi(u_k - u_a), u_k - u_a \rangle_{L^2(\Omega)} \\ &= \langle y_d + Su_a, Su_k \rangle_Y - \langle y_d, Su_a \rangle_Y + \alpha \langle w_k + u_a, u_k \rangle_{L^2(\Omega)} - \alpha \langle w_k, u_a \rangle_{L^2(\Omega)}. \end{aligned}$$

On the right-hand side, we can use Young's inequality to obtain

$$\frac{1}{2} \|Su_k\|_Y^2 + \frac{\alpha}{2} \|u_k\|_{L^2(\Omega)}^2 + \gamma_k \langle \pi(u_k - u_a), u_k - u_a \rangle_{L^2(\Omega)} \leq C,$$

where the constant  $C \geq 0$  is independent of  $k \in \mathbb{N}$ . Thus, there exists a subsequence (denoted by the same symbol) with  $u_k \rightharpoonup \tilde{u}$  in  $L^2(\Omega)$ . Due to convexity of the mapping  $v \mapsto \langle \pi(v), v \rangle_{L^2(\Omega)}$ , we find

$$0 \leq \langle \pi(\tilde{u} - u_a), \tilde{u} - u_a \rangle_{L^2(\Omega)} \leq \liminf_{k \rightarrow \infty} \langle \pi(u_k - u_a), u_k - u_a \rangle_{L^2(\Omega)} \leq \liminf_{k \rightarrow \infty} \frac{C}{\gamma_k} = 0.$$

Thus,  $\tilde{u} \in U_{\text{ad}}$ . Hence, we can use  $v = \tilde{u}$  in (5.1). Similarly, we test  $e_{\gamma_k}(u_k, w_k) = 0$  by  $u_k - u$ . Adding both relations yield

$$\begin{aligned} 0 &\geq \langle Su - y_d, S(u - \tilde{u}) \rangle_Y + \langle Su_k - y_d, S(u_k - u) \rangle_Y \\ &\quad + \alpha \langle u - w, u - \tilde{u} \rangle_{L^2(\Omega)} + \alpha \langle u_k - w_k, u_k - u \rangle_{L^2(\Omega)} + \gamma_k \langle \pi(u_k - u_a), u_k - u \rangle_{L^2(\Omega)} \\ &\geq \|S(u_k - u)\|_Y^2 + \langle Su - y_d, S(u_k - \tilde{u}) \rangle_Y \\ &\quad + \alpha \|u_k - u\|_{L^2(\Omega)}^2 + \alpha \langle u - w, u_k - \tilde{u} \rangle_{L^2(\Omega)} + \alpha \langle w - w_k, u_k - u \rangle_{L^2(\Omega)}. \end{aligned}$$

Here, we have used  $\langle \pi(u_k - u_a), u_k - u \rangle_{L^2(\Omega)} \geq 0$ , which follows from  $\pi(u_k - u_a) = 0$  a.e. on  $\{u_k \geq u_a\}$ , whereas  $\pi(u_k - u_a), u_k - u \leq 0$  a.e. on  $\{u_k < u_a\}$ . Together with the weak convergence  $u_k \rightharpoonup \tilde{u}$  this shows

$$0 \geq \limsup_{k \rightarrow \infty} \left( \|S(u_k - u)\|_Y^2 + \alpha \|u_k - u\|_{L^2(\Omega)}^2 \right) \geq 0,$$

i.e.,  $u_k \rightarrow u$  in  $L^2(\Omega)$ . Since the limit  $u = \tilde{u}$  is independent of the chosen subsequence, a usual subsequence-subsequence argument shows the convergence of the entire sequence.

Finally, we can approximate (IOC) by a regularized problem to prove enhanced regularity of local minimizers.

**Theorem 5.4.** Let  $(\bar{u}, \bar{w}) \in L^2(\Omega) \times H_0^1(\Omega)$  be a local minimizer of (IOC). Then,  $-\Delta \bar{w} \in L^2(\Omega)$ .

*Proof.* We denote by  $r > 0$  the radius of optimality and consider the regularized problem

$$\begin{aligned} \min_{u \in L^2(\Omega), w \in H_0^1(\Omega)} \quad & F(u, w) + \frac{1}{2} \|w - \bar{w}\|_{H_0^1(\Omega)}^2 \\ \text{s.t.} \quad & e_{\gamma_k}(u, w) = 0, \\ & w \geq w_a, \\ & (u, w) \in B_r((\bar{u}, \bar{w})). \end{aligned} \tag{IOC}_k$$

Here,  $(\gamma_k)_{k \in \mathbb{N}} \subset \mathbb{R}^+$  with  $\gamma_k \rightarrow \infty$  is arbitrary. By arguing as in [Lemma 5.1](#), we obtain the existence of global solutions  $(u_k, w_k)$  for this problem. Due to boundedness, we can extract a weakly convergent subsequence (without relabeling) with  $(u_k, w_k) \rightharpoonup (\tilde{u}, \tilde{w})$  in  $L^2(\Omega) \times H_0^1(\Omega)$ . Due to the compact embedding  $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$  and [Lemma 5.3](#), we find  $\tilde{u} = T(\tilde{w})$  and  $u_k \rightarrow \tilde{u}$  in  $L^2(\Omega)$ . For  $k$  large enough,  $(T_{\gamma_k}(\bar{w}), \bar{w})$  is feasible for  $(\text{IOC}_k)$ , see [Lemma 5.3](#). Thus,

$$F(u_k, w_k) + \frac{1}{2} \|w_k - \bar{w}\|_{H_0^1(\Omega)}^2 \leq F(T_{\gamma_k}(\bar{w}), \bar{w}).$$

Using the sequential weak lower semicontinuity of  $F$  and again [Lemma 5.3](#) yields

$$\begin{aligned} F(\tilde{u}, \tilde{w}) + \frac{1}{2} \|\tilde{w} - \bar{w}\|_{H_0^1(\Omega)}^2 &\leq \liminf_{k \rightarrow \infty} F(u_k, w_k) + \frac{1}{2} \limsup_{k \rightarrow \infty} \|w_k - \bar{w}\|_{H_0^1(\Omega)}^2 \\ &\leq \limsup_{k \rightarrow \infty} \left( F(u_k, w_k) + \frac{1}{2} \|w_k - \bar{w}\|_{H_0^1(\Omega)}^2 \right) \leq F(\bar{u}, \bar{w}). \end{aligned}$$

Since  $(\bar{u}, \bar{w})$  is a local minimizer and  $(\tilde{u}, \tilde{w}) = (T(\tilde{w}), \tilde{w}) \in B_r((\bar{u}, \bar{w}))$ ,  $(\tilde{u}, \tilde{w}) = (\bar{u}, \bar{w})$  as well as  $w_k \rightarrow \bar{w}$  in  $H_0^1(\Omega)$  follows. Thus, for  $k$  large enough, the last constraint in  $(\text{IOC}_k)$  is not binding and, consequently,  $w_k$  is a local solution of the reduced problem

$$\min_{w \in W_{\text{ad}}} F(T_{\gamma_k}(w), w) + \frac{1}{2} \|w - \bar{w}\|_{H_0^1(\Omega)}^2.$$

Due to the Fréchet differentiability of  $T_{\gamma_k}$  from [Lemma 5.2](#),  $w_k$  satisfies the necessary optimality condition

$$F'(u_k, w_k)(T'_{\gamma_k}(w_k)h, h) + \langle w_k - \bar{w}, h \rangle_{H_0^1(\Omega)} \geq 0 \quad \forall h \in \mathcal{T}_{W_{\text{ad}}}(w_k).$$

Now, we introduce the adjoint  $p_k \in L^2(\Omega)$  as the unique solution of

$$(\alpha + S^*S)p_k + \gamma_k \pi'(u_k - u_a)p_k = -\alpha f'(u_k).$$

Testing with  $p_k$  shows that  $p_k$  is bounded in  $L^2(\Omega)$  uniformly w.r.t.  $k \in \mathbb{N}$ , since  $f'(u_k) \rightarrow f'(\bar{u})$ . Next, we test the equation for  $p_k$  with  $v := T'_{\gamma_k}(w_k)h$  and (5.2) with  $p_k$ . This gives

$$\begin{aligned} -\alpha \langle f'(u_k), T'_{\gamma_k}(w_k)h \rangle_{L^2(\Omega)} &= \langle (\alpha + S^*S)p_k + \gamma_k \pi'(u_k - u_a)p_k, v \rangle_{L^2(\Omega)} \\ &= \langle p_k, (\alpha + S^*S)v + \gamma_k \pi'(u_k - u_a)v \rangle_{L^2(\Omega)} \\ &= \alpha \langle p_k, h \rangle_{L^2(\Omega)}. \end{aligned}$$

Thus, the above optimality condition becomes

$$\langle \zeta - p_k, h \rangle_{L^2(\Omega)} + \langle 2w_k - \bar{w}, h \rangle_{H_0^1(\Omega)} \geq 0 \quad \forall h \in \mathcal{T}_{W_{\text{ad}}}(w_k).$$

Up to a subsequence, we have  $p_k \rightharpoonup p$  in  $L^2(\Omega)$ . Due to  $w_k \rightarrow \bar{w}$  in  $H_0^1(\Omega)$ , we can pass to the limit and obtain

$$\langle \zeta - p, h \rangle_{L^2(\Omega)} + \langle \bar{w}, h \rangle_{H_0^1(\Omega)} \geq 0 \quad \forall h \in \mathcal{T}_{W_{\text{ad}}}(\bar{w}).$$

This is an obstacle problem for  $\bar{w}$  in  $H_0^1(\Omega)$ . We are going to apply the regularity result [Troianiello, 1987, Theorem 4.32]. Since we are dealing with the space  $H_0^1(\Omega)$ , we can choose  $\Gamma = \emptyset$  therein. Thus, the requirement that  $\Gamma$  is of class  $C^1$  is trivially satisfied. Since  $\zeta - p \in L^2(\Omega)$  and  $-\Delta w_a \in L^2(\Omega)$ , the application of [Troianiello, 1987, Theorem 4.32] yields

$$\langle p - \zeta, h \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} \leq \langle -\Delta \bar{w}, h \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} \leq \langle \max(p - \zeta, -\Delta w_a), h \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)}$$

for all  $h \in H_0^1(\Omega)$  with  $h \geq 0$ . Due to  $p - \zeta, -\Delta w_a \in L^2(\Omega)$  this implies  $-\Delta \bar{w} \in L^2(\Omega)$ .

We remark that  $-\Delta \bar{w} \in L^2(\Omega)$  has to be understood in the distributional sense, i.e.,

$$\langle -\Delta \bar{w}, \varphi \rangle_{L^2(\Omega)} = \langle \bar{w}, \varphi \rangle_{H_0^1(\Omega)} \tag{5.3}$$

holds for all  $\varphi \in C_c^\infty(\Omega)$ . Since both sides of the equation are continuous w.r.t. the  $H_0^1(\Omega)$ -norm of  $\varphi$ , it even holds for all  $\varphi \in H_0^1(\Omega) = \text{cl}_{H_0^1(\Omega)}(C_c^\infty(\Omega))$ .

We further note that we do not derive an equation for the limit  $p$  in the proof of Theorem 5.4. We expect that this would lead to a system of C-stationarity.

### 5.3 Linearization and M-stationarity

After we have proved the regularity  $-\Delta \bar{w} \in L^2(\Omega)$  in Theorem 5.4, we investigate optimality conditions for (IOC). We start by rewriting the VI (5.1), which characterizes the solution mapping  $T$  of (OC( $w$ )). To this end, we define the coercive operator  $A \in \mathcal{L}(L^2(\Omega), L^2(\Omega))$ ,

$$A := \alpha \text{id} + S^*S,$$

and, consequently, the associated inner product  $\langle \cdot, \cdot \rangle_A := \langle A \cdot, \cdot \rangle_{L^2(\Omega)}$  is equivalent to the original inner product on  $L^2(\Omega)$ . Using

$$S^*(Su - y_d) + \alpha(u - w) = A\left(u - A^{-1}(S^*y_d + \alpha w)\right),$$

the VI (5.1) can be written as

$$\langle u - A^{-1}(S^*y_d + \alpha w), v - u \rangle_A \geq 0 \quad \forall v \in U_{\text{ad}}.$$

This is, in turn, equivalent to

$$u = \text{Proj}_{U_{\text{ad}}}^A\left(A^{-1}(S^*y_d + \alpha w)\right), \tag{5.4}$$

where  $\text{Proj}_{U_{\text{ad}}}^A$  denotes the orthogonal projection onto  $U_{\text{ad}}$  w.r.t. the inner product  $A$ . This representation of  $T$  can be used to obtain the directional differentiability of  $T$  in a very convenient way.

**Lemma 5.5.** The solution mapping  $T: L^2(\Omega) \rightarrow L^2(\Omega)$  of (OC( $w$ )) is directionally differentiable. For  $\bar{w}, w \in L^2(\Omega)$ , the directional derivative  $z = T'(\bar{w}; w)$  is characterized by the variational inequality

$$z \in \mathcal{K}(\bar{w}), \quad \langle S^*Sz + \alpha(z - w), z - v \rangle_{L^2(\Omega)} \leq 0 \quad \forall v \in \mathcal{K}(\bar{w}). \tag{5.5}$$

Here,

$$\mathcal{K}(\bar{w}) := \{v \in \mathcal{T}_{U_{\text{ad}}}(\bar{u}) \mid \langle S^*(S\bar{u} - y_d) + \alpha(\bar{u} - \bar{w}), v \rangle_{L^2(\Omega)} = 0\}$$

is the critical cone with  $\bar{u} = T(\bar{w})$ . The mapping  $L^2(\Omega) \ni w \mapsto T'(\bar{w}; w) \in L^2(\Omega)$  is Lipschitz with constant 1.

*Proof.* It is well known that the set  $U_{\text{ad}}$  is a polyhedric subset of  $L^2(\Omega)$ , see, e.g., [Wachsmuth, 2019, Example 4.21 (1)]. It follows, that the solution mapping  $T$  of (5.4) is directionally differentiable and the directional derivative  $z = T'(\bar{w}; w)$  satisfies

$$z \in \mathcal{K}(\bar{w}), \quad \langle z - A^{-1}(\alpha w), z - v \rangle_A \leq 0 \quad \forall v \in \mathcal{K}(\bar{w}),$$

where the critical cone  $\mathcal{K}(\bar{w})$  is given by

$$\mathcal{K}(\bar{w}) := \{v \in \mathcal{T}_{U_{\text{ad}}}(\bar{u}) \mid \langle \bar{u} - A^{-1}(S^*y_d + \alpha\bar{w}), v \rangle_A = 0\},$$

see [Mignot, 1976, Théorème 2.1], [Haraux, 1977, Theorem 2], [Wachsmuth, 2019, Theorem 5.2]. Utilizing the definition of  $A$ , the VI (5.5) follows. The Lipschitz continuity follows by a suitable testing of (5.5).

Using the extra regularity of local minimizers from Theorem 5.4 together with the differentiability of  $T$  from the previous lemma, we are in position to prove an optimality system of B-stationarity.

**Lemma 5.6.** Let  $(\bar{u}, \bar{w}) \in L^2(\Omega) \times H_0^1(\Omega)$  be a local minimizer of (IOC). Then,

$$\langle f'(\bar{u}), T'(\bar{w}; w) \rangle_{L^2(\Omega)} + \langle -\Delta \bar{w} + \zeta, w \rangle_{L^2(\Omega)} \geq 0 \quad \forall w \in \text{cl}_{L^2(\Omega)} \mathcal{T}_{W_{\text{ad}}}(\bar{w}). \quad (5.6)$$

Here,  $\mathcal{T}_{W_{\text{ad}}}(\bar{w})$  is the tangent cone of  $W_{\text{ad}} \subset H_0^1(\Omega)$  at  $\bar{w}$  w.r.t. the  $H_0^1(\Omega)$ -topology.

*Proof.* Using the continuous solution operator  $T$  of (OC( $w$ )),  $\bar{w}$  is a local solution to the reduced problem

$$\min_{w \in W_{\text{ad}}} F(T(w), w) = f(T(w)) + \frac{1}{2} \|w\|_{H_0^1(\Omega)}^2 + \langle \zeta, w \rangle_{L^2(\Omega)}.$$

Using the differentiability and Lipschitz continuity of  $T$ , this implies

$$\langle f'(\bar{u}), T'(\bar{w}; w) \rangle_{L^2(\Omega)} + \langle \bar{w}, w \rangle_{H_0^1(\Omega)} + \langle \zeta, w \rangle_{L^2(\Omega)} \geq 0 \quad \forall w \in \mathcal{T}_{W_{\text{ad}}}(\bar{w}).$$

Finally, we use (5.3) and the continuity of  $L^2(\Omega) \ni w \mapsto T'(\bar{w}; w) \in L^2(\Omega)$  (see Lemma 5.5) to arrive at (5.6).

The next lemma characterizes the  $L^2(\Omega)$ -closure of the  $H_0^1(\Omega)$ -tangent cone  $\mathcal{T}_{W_{\text{ad}}}(\bar{w})$ .

**Lemma 5.7.** For all  $\bar{w} \in W_{\text{ad}}$  we have

$$\text{cl}_{L^2} \mathcal{T}_{W_{\text{ad}}}(\bar{w}) = \{v \in L^2(\Omega) \mid v \geq 0 \text{ a.e. on } \{\bar{w} = w_a\}\}.$$

*Proof.* We denote the set on the right-hand side by  $L$ . Due to  $\mathcal{T}_{W_{\text{ad}}}(\bar{w}) = \text{cl}_{H_0^1(\Omega)} \text{cone}(W_{\text{ad}} - \bar{w})$ , the inclusion “ $\subset$ ” follows from  $W_{\text{ad}} - \bar{w} \subset L$ , since  $L$  is a closed convex cone in  $L^2(\Omega)$ .

Now, let  $v \in L$  be given. Note that  $L$  is the  $L^2(\Omega)$ -tangent cone of

$$\hat{W}_{\text{ad}} := \{w \in L^2(\Omega) \mid w \geq w_a \text{ a.e. on } \Omega\}.$$

Thus, for each  $\varepsilon > 0$ , there is  $t > 0$  and  $w \in \hat{W}_{\text{ad}}$ , such that  $\|v - t(w - \bar{w})\|_{L^2(\Omega)} \leq \varepsilon$ . Due to the density of the embedding  $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ , there exists  $\psi \in H_0^1(\Omega)$  with  $\|\psi - w\|_{L^2(\Omega)} \leq \varepsilon/t$ . We define  $\tilde{w} := \max(\psi, w_a)$ . Then, we have  $\tilde{w} \in H_0^1(\Omega)$ , since  $\max(0, w_a) \leq \max(\psi, w_a) \leq |\psi| + \max(0, w_a)$  and  $\psi, \max(0, w_a) \in H_0^1(\Omega)$ . In particular,  $\tilde{w} \in W_{\text{ad}}$ . Moreover,  $|\tilde{w} - w| = |\max(\psi - w, w_a - w)| \leq |\psi - w|$  implies  $\|\tilde{w} - w\|_{L^2(\Omega)} \leq \|\psi - w\|_{L^2(\Omega)} \leq \varepsilon/t$ . Thus,  $\|v - t(\tilde{w} - \bar{w})\|_{L^2(\Omega)} \leq 2\varepsilon$ . Since  $\varepsilon > 0$  was arbitrary, this shows  $v \in \text{cl}_{L^2(\Omega)} \mathcal{T}_{W_{\text{ad}}}(\bar{w})$ .

We remark that one has

$$\mathcal{T}_{W_{\text{ad}}}(\bar{w}) = \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ q.e. on } \{\bar{w} = w_a\}\}$$



and this characterization requires the tools of capacity theory, see [Mignot, 1976, Lemma 3.2]. Note that the above proof does not use this form of  $\mathcal{T}_{W_{\text{ad}}}(\bar{w})$ .

The next lemma shows that local optimality implies that  $(0, 0, 0)$  is a local minimizer of a linearized problem similar to (MPCC<sub>lin</sub>).

**Lemma 5.8.** Let  $(\bar{u}, \bar{w}) \in L^2(\Omega) \times H_0^1(\Omega)$  be a local minimizer of (IOC) and let  $\bar{\xi} := S^*(S\bar{u} - y_d) + \alpha(\bar{u} - \bar{w})$ . Then,  $(0, 0, 0) \in L^2(\Omega)^3$  is a (global) minimizer of

$$\begin{aligned} \min_{u, w, \xi \in L^2(\Omega)} \quad & \langle f'(\bar{u}), u \rangle_{L^2(\Omega)} + \langle -\Delta \bar{w} + \zeta, w \rangle_{L^2(\Omega)} \\ \text{s.t.} \quad & (\alpha + S^*S)u - \alpha w - \xi = 0, \\ & w \geq 0 \text{ a.e. on } \{\bar{w} = w_a\}, \\ & u = 0 \text{ a.e. on } \{\bar{\xi} > 0\}, \\ & \xi = 0 \text{ a.e. on } \{\bar{u} > u_a\}, \\ & 0 \leq u \perp \xi \geq 0 \text{ a.e. on } \{\bar{u} = u_a, \bar{\xi} = 0\}. \end{aligned} \tag{5.7}$$

*Proof.* Using the lower-level optimality conditions, one can see that  $0 \leq \bar{u} - u_a \perp \bar{\xi} \geq 0$  holds a.e. in  $\Omega$ . The claim follows from Lemmas 5.6 and 5.7 by using that  $u = T'(\bar{w}; w)$  can be characterized by the VI (5.5) and

$$\begin{aligned} \mathcal{K}(\bar{w}) &= \{u \in L^2(\Omega) \mid u \geq 0 \text{ a.e. on } \{\bar{u} = u_a, \bar{\xi} = 0\}, u = 0 \text{ a.e. on } \{\bar{\xi} > 0\}\}, \\ -\mathcal{K}(\bar{w})^\circ &= \{\xi \in L^2(\Omega) \mid \xi \geq 0 \text{ a.e. on } \{\bar{u} = u_a, \bar{\xi} = 0\}, \xi = 0 \text{ a.e. on } \{\bar{u} > u_a\}\}. \end{aligned}$$

We want to consider a system for M-stationarity and other stationarity systems for (IOC). As we have only defined these stationarity systems for MPCCs, we will consider the KKT reformulation of (IOC), which is obtained by replacing the lower-level by its KKT conditions.

$$\begin{aligned} \min_{u, \xi \in L^2(\Omega), w \in H_0^1(\Omega)} \quad & F(u, w) \\ \text{s.t.} \quad & w \geq w_a \quad \text{a.e. on } \Omega, \\ & S^*(Su - y_d) + \alpha(u - w) - \xi = 0, \\ & 0 \leq u - u_a \perp \xi \geq 0 \quad \text{a.e. on } \Omega. \end{aligned} \tag{KKTR}$$

This optimization problem is an instance of (MPCC). Note that replacing the lower-level by its KKT condition leads to an equivalent problem because the multiplier  $\xi$  depends continuously on  $u$  and  $w$ .

Finally, we combine the linearization results with the results of Section 4 to obtain M-stationarity of local minimizers of (IOC) under some assumptions.

**Theorem 5.9.** Let  $(\bar{u}, \bar{w}) \in L^2(\Omega) \times H_0^1(\Omega)$  be a local minimizer of (IOC) and let  $\bar{\xi} := S^*(S\bar{u} - y_d) + \alpha(\bar{u} - \bar{w})$ . We require that one of the following assumptions is satisfied.

- (a) The operator  $S^*S$  preserves non-negativity, i.e.,  $u \geq 0$  a.e. implies  $S^*Su \geq 0$  a.e., and  $\bar{u} = u_a$  holds a.e. in  $\Omega$ .
  - (b) It holds  $Y = \mathbb{R}$  and  $Sv = \beta \langle 1, v \rangle_{L^2(\Omega)}$  for all  $v \in L^2(\Omega)$ , where  $\beta > 0$  is a constant.
- Then  $(\bar{u}, \bar{w}, \bar{\xi})$  is an M-stationary point of (KKTR).

*Proof.* According to Lemma 5.8,  $(0, 0, 0)$  is a minimizer of the linearized optimization problem (5.7). This optimization problem is an instance of (MPCC<sub>lin</sub>) (but one does need to rescale  $w$  by  $\alpha$  such that  $\alpha w$  in the first constraint of (5.7) becomes  $w$ ). In case (a), Assumption 4.1(a) is satisfied, whereas in case (b), Assumption 4.1(c) is satisfied. Thus, according to Theorems 4.9 and 4.10,  $(0, 0, 0)$  is an M-stationary point of (5.7).

When one writes down the M-stationarity systems of (KKTR) at  $(\bar{u}, \bar{w}, \bar{\xi})$  and of (5.7) at  $(0, 0, 0)$ , it is easy to check that they are equivalent. Thus,  $(\bar{u}, \bar{w}, \bar{\xi})$  is an M-stationary point of (KKTR).

### 5.4 Example without strong stationarity

Finally, we present a class of problems for which the (global) minimizer fails to be strongly stationary. In addition to the general setting from Section 5.1, we specify

$$y_d = 0 \in Y, \quad u_a = w_a \equiv 0, \quad \zeta \equiv 1, \quad f(u) = - \int_{\Omega} u \, d\omega.$$

Further, we require that  $S^*S$  preserves non-negativity, i.e.,  $u \geq 0$  implies  $S^*Su \geq 0$ .

For  $w \in L^2(\Omega)$  with  $w \geq 0$ , the solution  $u = T(w)$  of (OC(w)) satisfies the VI (5.1), i.e.,

$$u = \text{Proj}_{U_{\text{ad}}}^{L^2(\Omega)}(w - \alpha^{-1}S^*Su) = \max\left(0, w - \alpha^{-1}S^*Su\right).$$

From  $u \geq 0$  we infer  $S^*Su \geq 0$ , thus,  $u \leq w$ . Hence, we have can estimate the objective for every feasible pair  $(u, w) = (T(u), w)$  of (IOC) via

$$F(u, w) = \int_{\Omega} (w - u) \, d\omega + \frac{1}{2} \|w\|_{H_0^1(\Omega)}^2 \geq 0.$$

Together with  $0 = T(0)$  and  $F(0, 0) = 0$ , it follows that  $(\bar{u}, \bar{w}) = (0, 0)$  is the unique global minimizer of (IOC) in this setting. Thus,  $(\bar{u}, \bar{w}, \bar{\xi}) = (0, 0, 0)$  is the unique global minimizer of the equivalent (KKTR).

We check that this global minimizer is not strongly stationary. If one applies Definition 2.1 to (KKTR), the resulting system of strong stationarity is

$$\begin{aligned} -1 + (\alpha \text{id} + S^*S)^*p + \mu &= 0, \\ 1 - p + \lambda &= 0, \end{aligned}$$

$$\begin{aligned}
-p + \nu &= 0, \\
\lambda &\leq 0 \quad \text{a.e. on } \Omega, \\
\mu &\leq 0 \quad \text{a.e. on } \Omega, \\
\nu &\leq 0 \quad \text{a.e. on } \Omega.
\end{aligned}$$

The third equation gives  $p = \nu \leq 0$  a.e. in  $\Omega$ , which implies  $(\alpha \text{id} + S^*S)^*p \leq 0$  a.e. in  $\Omega$ . Then

$$0 = -1 + (\alpha \text{id} + S^*S)^*p + \mu \leq -1 + \mu \leq -1 \quad \text{a.e. on } \Omega$$

follows, which is false. Therefore,  $(\bar{u}, \bar{w}, \bar{\xi}) = (0, 0, 0)$  is not strongly stationary.

However, [Theorem 5.9\(a\)](#) can be applied, which tells us that  $(\bar{u}, \bar{w}, \bar{\xi}) = (0, 0, 0)$  is an M-stationary point. One possibility for the multipliers to the system of M-stationarity can be found by setting  $\nu = 0$ , which implies  $p = \nu = 0$ ,  $\mu = 1$ ,  $\lambda = -1$ . It is natural to ask, whether multipliers to the system of M-stationarity can also be found with  $\mu = 0$ . This would imply  $p = (\alpha + S^*S)^{-1}(1)$  and  $p = 1 + \lambda \leq 1$ . However, if  $\Omega = (0, 1)$ ,  $Sv = \alpha \langle 1, v \rangle_{L^2(\Omega)}$  for all  $v \in L^2(\Omega)$ , and  $\alpha$  is sufficiently small, we have  $p = 1/(\alpha + \alpha^2) > 1$ . Thus, multipliers to the system of M-stationarity with  $\mu = 0$  cannot always be found. However, since [Theorem 5.9\(a\)](#) is based on [Theorem 3.1](#) and strong stationarity does not hold, we know from [Remark 3.7](#) that there must be more than one choice for the multipliers  $\bar{\mu}, \bar{\nu}$  such that the system of M-stationarity can be satisfied.

## 6 Conclusion and outlook

Generalizing the approach from [\[Harder, 2021a\]](#) to Lebesgue spaces proved to be no simple task. Interestingly, the  $A_{\forall}$ -stationarity seems to be a bigger problem than going from there to M-stationarity. The  $A_{\forall}$ -stationarity requires to find KKT multipliers for [\(LP\( \$\beta\$ \)\)](#) for all measurable  $\beta$ , but even though these are convex and linear optimization problems, finding KKT multipliers is not always possible, as shown by [Example 4.8](#).

Nonetheless, in [Section 5](#), we demonstrated that our results can be applied for a class of inverse optimal control problems. Thus, we were able to show that local minimizers of these problems are M-stationarity points, and this result is also applicable in situations where the local minimizer is not strongly stationary.

To the best of our knowledge, this was the first instance that M-stationarity could be shown for MPCC-type problems in Lebesgue spaces. It would be interesting to know, if these methods could be improved to show M-stationarity for other MPCCs in Lebesgue spaces. Furthermore, some ideas may be used to improve stationarity conditions for MPCCs in the Sobolev space  $H_0^1(\Omega)$ , e.g., for the optimal control of the obstacle problem.

## Acknowledgment

This work is supported by the DFG Grant *Bilevel Optimal Control: Theory, Algorithms, and Applications* (Grant No. WA 3636/4-2) within the Priority Program SPP 1962 (Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization).

## References

- Fabian, M.; Habala, P.; Hájek, P.; Santalucía, V. M.; Pelant, J.; Zizler, V. (2001). *Functional Analysis and Infinite-Dimensional Geometry*. Springer New York. DOI: [10.1007/978-1-4757-3480-5](https://doi.org/10.1007/978-1-4757-3480-5).
- Flegel, M. L.; Kanzow, C. (2005). “Abadie-type constraint qualification for mathematical programs with equilibrium constraints”. *Journal of Optimization Theory and Applications* 124.3, pp. 595–614. DOI: [10.1007/s10957-004-1176-x](https://doi.org/10.1007/s10957-004-1176-x).
- Flegel, M. L.; Kanzow, C.; Outrata, J. V. (2007). “Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints”. *Set-Valued Analysis* 15.2, pp. 139–162. DOI: [10.1007/s11228-006-0033-5](https://doi.org/10.1007/s11228-006-0033-5).
- Goldberg, H.; Kampowsky, W.; Tröltzsch, F. (1992). “On Nemytskij operators in  $L_p$ -spaces of abstract functions”. *Mathematische Nachrichten* 155, pp. 127–140. DOI: [10.1002/mana.19921550110](https://doi.org/10.1002/mana.19921550110).
- Haraux, A. (1977). “How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities”. *Journal of the Mathematical Society of Japan* 29.4, pp. 615–631. DOI: [10.2969/jmsj/02940615](https://doi.org/10.2969/jmsj/02940615).
- Harder, F. (2021a). “A new elementary proof for M-stationarity under MPCC-GCQ for mathematical programs with complementarity constraints”. *Journal of Nonsmooth Analysis and Optimization* 2. DOI: [10.46298/jnsao-2021-6903](https://doi.org/10.46298/jnsao-2021-6903).
- Harder, F. (2021b). *New stationarity conditions between strong and M-stationarity for mathematical programs with complementarity constraints*. arXiv: [2109.01623](https://arxiv.org/abs/2109.01623).
- Harder, F. (2021c). “On bilevel optimization problems in infinite-dimensional spaces”. PhD thesis. BTU Cottbus-Senftenberg. DOI: [10.26127/BTUOPEN-5375](https://doi.org/10.26127/BTUOPEN-5375).
- Harder, F.; Wachsmuth, G. (2018a). “Optimality conditions for a class of inverse optimal control problems with partial differential equations”. *Optimization* 68.2-3, pp. 615–643. DOI: [10.1080/02331934.2018.1495205](https://doi.org/10.1080/02331934.2018.1495205).
- Harder, F.; Wachsmuth, G. (2018b). “The limiting normal cone of a complementarity set in Sobolev spaces”. *Optimization* 67.10, pp. 1579–1603. DOI: [10.1080/02331934.2018.1484467](https://doi.org/10.1080/02331934.2018.1484467).
- Hintermüller, M.; Mordukhovich, B. S.; Surowiec, T. (2014). “Several approaches for the derivation of stationarity conditions for elliptic MPECs with upper-level control constraints”. *Mathematical Programming* 146.1-2, Ser. A, pp. 555–582. DOI: [10.1007/s10107-013-0704-6](https://doi.org/10.1007/s10107-013-0704-6).

- Jarušek, J.; Outrata, J. V. (2007). “On sharp necessary optimality conditions in control of contact problems with strings”. *Nonlinear Analysis* 67.4, pp. 1117–1128. DOI: [10.1016/j.na.2006.05.021](https://doi.org/10.1016/j.na.2006.05.021).
- Mehlitz, P. (2017). “Contributions to complementarity and bilevel programming in Banach spaces”. PhD thesis. Technische Universität Bergakademie Freiberg. URN: [urn:nbn:de:bsz:105-qucosa-227091](https://nbn-resolving.org/urn:nbn:de:bsz:105-qucosa-227091).
- Mehlitz, P. (2018). “On the sequential normal compactness condition and its restrictiveness in selected function spaces”. *Set-Valued and Variational Analysis* 27.3, pp. 763–782. DOI: [10.1007/s11228-018-0475-6](https://doi.org/10.1007/s11228-018-0475-6).
- Mehlitz, P.; Wachsmuth, G. (2018). “The limiting normal cone to pointwise defined sets in Lebesgue spaces”. *Set-Valued and Variational Analysis* 26.3, pp. 449–467. DOI: [10.1007/s11228-016-0393-4](https://doi.org/10.1007/s11228-016-0393-4).
- Mehlitz, P.; Wachsmuth, G. (2019). “The weak sequential closure of decomposable sets in Lebesgue spaces and its application to variational geometry”. *Set-Valued and Variational Analysis* 27.1, pp. 265–294. DOI: [10.1007/s11228-017-0464-1](https://doi.org/10.1007/s11228-017-0464-1).
- Mignot, F. (1976). “Contrôle dans les inéquations variationnelles elliptiques”. *Journal of Functional Analysis* 22.2, pp. 130–185. DOI: [10.1016/0022-1236\(76\)90017-3](https://doi.org/10.1016/0022-1236(76)90017-3).
- Schirotzek, W. (2007). *Nonsmooth analysis*. Universitext. Springer, Berlin. DOI: [10.1007/978-3-540-71333-3](https://doi.org/10.1007/978-3-540-71333-3).
- Troianiello, G. M. (1987). *Elliptic differential equations and obstacle problems*. The University Series in Mathematics. New York: Plenum Press, pp. xiv+353.
- Wachsmuth, G. (2014). “Differentiability of implicit functions: Beyond the implicit function theorem”. *Journal of Mathematical Analysis and Applications* 414.1, pp. 259–272. DOI: [10.1016/j.jmaa.2014.01.007](https://doi.org/10.1016/j.jmaa.2014.01.007).
- Wachsmuth, G. (2016). “Towards M-stationarity for optimal control of the obstacle problem with control constraints”. *SIAM Journal on Control and Optimization* 54.2, pp. 964–986. DOI: [10.1137/140980582](https://doi.org/10.1137/140980582).
- Wachsmuth, G. (2019). “A guided tour of polyhedral sets”. *Journal of Convex Analysis* 26.1, pp. 153–188.