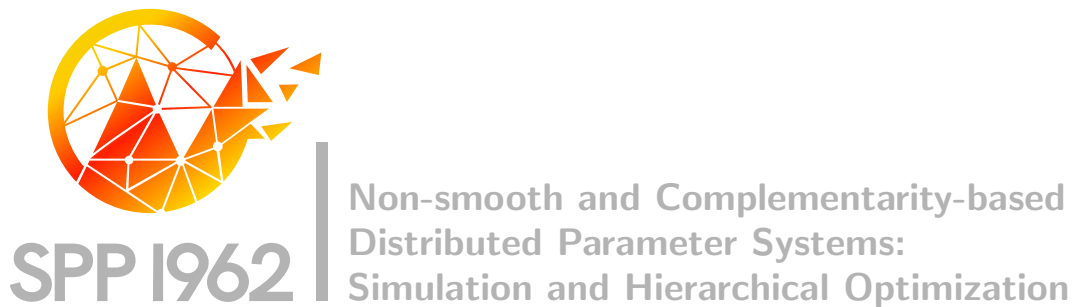


**DFG** Deutsche  
Forschungsgemeinschaft  
Priority Programme 1962

*Optimality Conditions for a Class of Inverse Optimal Control Problems  
with Partial Differential Equations*

Felix Harder, Gerd Wachsmuth



Preprint Number SPP1962-048

received on January 27, 2018

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/>

# OPTIMALITY CONDITIONS FOR A CLASS OF INVERSE OPTIMAL CONTROL PROBLEMS WITH PARTIAL DIFFERENTIAL EQUATIONS

Felix Harder\*      Gerd Wachsmuth<sup>†</sup> 

January 24, 2018

Research Group

Numerical Mathematics

(Partial Differential Equations)

We consider bilevel optimization problems which can be interpreted as inverse optimal control problems. The lower-level problem is an optimal control problem with a parametrized objective function. The upper-level problem is used to identify the parameters of the lower-level problem. Our main focus is the derivation of first-order necessary optimality conditions. We prove C-stationarity of local solutions of the inverse optimal control problem and give a counterexample to show that strong stationarity might be violated at a local minimizer.

**KEYWORDS:** inverse optimal control problem, optimality condition, C-stationarity, strong stationarity

**MSC:** [49K20](#), [90C31](#)

## 1 INTRODUCTION

Nowadays, the optimal control of partial differential equations (PDEs) is well understood, both theoretically and numerically. In the last decade, the interest in inverse optimal

---

\*Technische Universität Chemnitz, Faculty of Mathematics, Professorship Numerical Mathematics (Partial Differential Equations), 09107 Chemnitz, Germany, [felix.harder@mathematik.tu-chemnitz.de](mailto:felix.harder@mathematik.tu-chemnitz.de), [https://www.tu-chemnitz.de/mathematik/part\\_dgl/people/harder/index.php](https://www.tu-chemnitz.de/mathematik/part_dgl/people/harder/index.php)

<sup>†</sup>Technische Universität Chemnitz, Faculty of Mathematics, Professorship Numerical Mathematics (Partial Differential Equations), 09107 Chemnitz, Germany, [gerd.wachsmuth@mathematik.tu-chemnitz.de](mailto:gerd.wachsmuth@mathematik.tu-chemnitz.de), [https://www.tu-chemnitz.de/mathematik/part\\_dgl/people/wachsmuth/index.php](https://www.tu-chemnitz.de/mathematik/part_dgl/people/wachsmuth/index.php)

control problems emerged. That is, one wants to identify parameters in an optimal control problem by measurements of the optimal control or state. We refer to [Albrecht, Leibold, Ulbrich, 2012](#); [Albrecht, Ulbrich, 2017](#); [Mombaur, Truong, Laumond, 2010](#); [Hatz, Schlöder, Bock, 2012](#) for inverse optimal control problems with ordinary differential equations (ODEs). In these contributions, the first step is the discretization of the inverse optimal control problem. Afterwards, the resulting finite-dimensional problem is analyzed.

In our work, we consider the inverse optimal control problem governed by a linear elliptic partial differential equation. For a finite number of parameters  $\alpha = (\alpha_i)_{i=1}^n$  we consider the parametrized optimal control problem

$$\begin{aligned} \min_{y,u} \quad & \sum_{i=1}^n \alpha_i f_i(y, u) + \frac{\alpha_0}{2} \|u\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad & -\Delta y = u \quad \text{in } H^{-1}(\Omega), \\ & y \in H_0^1(\Omega), \\ & u \in U_{\text{ad}} \subset L^2(\Omega). \end{aligned} \tag{P(\alpha)}$$

For the precise assumptions on the data of  $(P(\alpha))$ , we refer to [Assumption 2.1](#) below. We just mention that  $\alpha_i \geq 0$ ,  $\alpha_0 > 0$ , and that the functions  $f_i$  are assumed to be jointly convex in  $(y, u)$ . Therefore, problem  $(P(\alpha))$  has a unique solution for all choices of the parameter vector  $\alpha$ .

Given a measured state  $y_d \in L^2(\Omega)$ , we consider the inverse optimal control problem

$$\begin{aligned} \min_{\alpha, y, u} \quad & \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad & (y, u) \text{ solves } (P(\alpha)), \\ & \alpha \geq 0 \text{ and } \sum_{i=1}^n \alpha_i = 1. \end{aligned} \tag{IOP}$$

That is, we try to find the parameter vector  $\alpha$  which reproduces the measurement  $y_d$  as good as possible. Since the solution mapping of  $(P(\alpha))$  is a constraint in  $(\text{IOP})$ , this problem becomes a bilevel optimal control problem. As already mentioned, the pair  $(y, u)$  is uniquely determined by  $\alpha$ . Hence, we do not have to distinguish between pessimistic and optimistic formulations of the bilevel problem  $(\text{IOP})$ , see [Dempe, 2002](#), Section 5.1. Nevertheless, the mapping from  $\alpha$  to the solution  $(y, u)$  of  $(P(\alpha))$  is, in general, not differentiable. Thus, we cannot employ standard methods for the analysis of  $(\text{IOP})$ .

One possibility for the derivation of optimality conditions for  $(\text{IOP})$  is to replace  $(P(\alpha))$  by its (necessary and sufficient) optimality conditions. The resulting problem is similar to a mathematical program with complementarity constraints (MPCCs) in an infinite-dimensional space. This class of problems has recently been studied in [Wachsmuth, 2015](#); [Mehlitz, Wachsmuth, 2016](#). However, the constraint qualifications derived in these papers are not applicable for reformulations of  $(\text{IOP})$ .

The main goal of our work is the derivation of necessary optimality conditions for problem (IOP). As already mentioned, this is not possible by standard methods. We are going to penalize the control constraints in the optimal control problem  $(P(\alpha))$ . Consequently, the solution map becomes differentiable and we can derive optimality conditions for the regularized version of (IOP) by an adjoint calculus. Passing to the limit in these optimality systems, we arrive at an optimality system for the original problem.

The structure of our paper is as follows. In Section 2, we discuss assumptions and basic properties of the lower-level problem and investigate properties of the solution map  $\alpha \mapsto u$  of the lower-level problem. In Section 3, we turn our attention to the upper-level problem. We prove strong stationarity in the unconstrained case  $U_{\text{ad}} = L^2(\Omega)$  and define the stationarity systems for (IOP). Additionally, we provide a counterexample in Section 3.2 which shows that strong stationarity may not hold in the general case. Finally, Section 4 is dedicated to proving C-stationarity for a local minimizer of (IOP).

## 2 LOWER-LEVEL PROBLEM

In this section, we discuss properties of the lower-level problem  $(P(\alpha))$ . First, we introduce the precise assumptions on the data of  $(P(\alpha))$ .

### Assumption 2.1.

1. The vector of parameters  $\alpha = (\alpha_i)_{i=1}^n$  satisfies  $0 \leq \alpha_i \leq 2$  for each  $i = 1, \dots, n$ , whereas  $\alpha_0$  is a fixed positive constant.
2. The set  $\Omega \subset \mathbb{R}^d$  is open and bounded, and  $1 \leq d \leq 3$ .
3. The set of admissible controls  $U_{\text{ad}}$  satisfies

$$U_{\text{ad}} := \{u \in L^2(\Omega) \mid u_a \leq u \leq u_b \text{ a.e. in } \Omega\}, \quad (2.1)$$

where  $u_a : \Omega \rightarrow [-\infty, \infty)$  and  $u_b : \Omega \rightarrow (-\infty, \infty]$  are measurable with  $u_a < u_b$  a.e. in  $\Omega$ . Moreover, there exists  $u_0 \in U_{\text{ad}} \cap L^\infty(\Omega)$ .

4. The functions  $f_i : L^\infty(\Omega) \times L^2(\Omega) \rightarrow \mathbb{R}$  are given by

$$f_i(y, u) := \int_{\Omega} \phi_i(x, y(x), u(x)) \, dx,$$

where  $\phi_i : \Omega \times \mathbb{R} \times \mathbb{R}$  are Carathéodory functions for all  $i = 1, \dots, n$  (i.e.,  $\phi(x, \cdot, \cdot)$  is continuous for a.a.  $x \in \Omega$  and  $\phi(\cdot, y, u)$  is measurable for all  $(y, u) \in \mathbb{R}^2$ ). Moreover,  $\phi_i(x, \cdot, \cdot)$  is assumed to be  $C^2$  and convex for a.a.  $x \in \Omega$ .

5. We have  $\phi_i(\cdot, 0, 0) \in L^1(\Omega)$ . Finally, for every  $M > 0$ , there exist constants  $C_M, \hat{C}_M > 0$ ,  $\gamma_M \in L^1(\Omega)$ ,  $\mu_M \in L^{\hat{p}}(\Omega)$ ,  $\hat{p} \geq 1$  and  $\hat{p} > d/2$ , such that

$$|D_u \phi_i(x, y, u)| \leq C_M \quad (2.2a)$$

$$|D_{yy} \phi_i(x, y, u)| + |D_{yu} \phi_i(x, y, u)| + |D_{uu} \phi_i(x, y, u)| \leq C_M \quad (2.2b)$$

hold for a.a.  $x \in \Omega$  and all  $y, u \in \mathbb{R}$  with  $|y|, |u| \leq M$  and such that

$$|\phi_i(x, y, u)| \leq \gamma_M(x) + C_M |u|^2 \quad (2.2c)$$

$$|D_y \phi_i(x, y, u)| \leq \mu_M(x) + \hat{C}_M |u|^2 \quad (2.2d)$$

holds for a.a.  $x \in \Omega$  and all  $y, u \in \mathbb{R}$  with  $|y| \leq M$ .

For the upper bound on  $\alpha$  we remark that by scaling  $\alpha$  and  $\alpha_0$  appropriately, the problem can be equivalently reformulated for arbitrary upper bounds on  $\alpha$ .

Further, we mention that the admissible set  $U_{\text{ad}}$  might not be bounded neither in  $L^2(\Omega)$  nor in  $L^\infty(\Omega)$ . Since we will consider a version of  $(P(\alpha))$  in which the control constraints are penalized in [Section 4](#), we need to allow this generality. Together with the mild requirements on the functions  $\phi_i$ , we cannot use the standard theory from, e.g., [Tröltzsch, 2010](#), to infer existence and regularity of an optimal control.

Existence of controls will be ensured by the regularization term. We provide a pointwise optimality condition and this can be used to check that the optimal controls are uniformly bounded in  $L^\infty(\Omega)$ . Next, we verify that the mapping from  $\alpha$  to  $u$  is Lipschitz continuous into  $L^\infty(\Omega)$ . This Lipschitz continuity allows us to apply a differentiability result to the mapping  $\alpha \mapsto u$ .

For convenience, we will use

$$f_0(x, y, u) = \frac{1}{2} \|u\|_{L^2(\Omega)}^2, \quad \phi_0(x, y, u) = \frac{1}{2} |u|^2.$$

It is clear that  $f_0$  satisfies [Assumption 2.1](#).

## 2.1 EXISTENCE AND REGULARITY OF OPTIMAL CONTROLS

Our first goal is to prove existence of an optimal control. This is not immediate, since the admissible set  $U_{\text{ad}}$  may not be bounded in  $L^2(\Omega)$ . Since the PDE is linear and the objective is assumed to be strictly convex, the optimal control is automatically unique.

As an auxiliary result, we check that the functions  $f_i$  are well defined in Lebesgue spaces.

**Lemma 2.2.** For all  $i = 1, \dots, n$ , the function  $f_i$  maps  $L^\infty(\Omega) \times L^2(\Omega)$  to  $\mathbb{R}$  and is continuous in these spaces.

*Proof.* For every  $y \in L^\infty(\Omega)$  and  $u \in L^2(\Omega)$  we have  $|y| \leq M$  for some  $M > 0$  and [\(2.2c\)](#) implies

$$\begin{aligned} |f_i(y, u)| &\leq \int_{\Omega} |\phi_i(x, y(x), u(x))| \, dx \\ &\leq \int_{\Omega} \gamma_M(x) + C_M |u(x)|^2 \, dx \leq \|\gamma_M\|_{L^1(\Omega)} + C_M \|u\|_{L^2(\Omega)}^2 < \infty. \end{aligned}$$

To prove the continuity of  $f_i$ , let  $y_k \rightarrow y$  in  $L^\infty(\Omega)$  and  $u_k \rightarrow u$  in  $L^2(\Omega)$ . W.l.o.g. (otherwise, pick a subsequence and use uniqueness of the limit  $f(y, u)$ ), we assume that  $\|u_k - u\|_{L^2(\Omega)} \leq 2^{-k}$  and  $u_k \rightarrow u$  pointwise a.e. in  $\Omega$ . In particular, this yields

$$\phi_i(x, y_k(x), u_k(x)) \rightarrow \phi_i(x, y(x), u(x))$$

for a.a.  $x \in \Omega$ . We define

$$\bar{u} := |u| + \sum_{k=1}^{\infty} |u_k - u| \in L^2(\Omega).$$

This yields  $|u_k| \leq |\bar{u}|$  a.e. in  $\Omega$ . Again, there exists  $M > 0$  with  $|y_k| \leq M$  for all  $k$ . Hence,

$$|\phi_i(x, y_k(x), u_k(x))| \leq \gamma_M(x) + C_M |\bar{u}(x)|^2$$

and the right-hand side is integrable. The dominated convergence theorem implies  $f_i(y_k, u_k) \rightarrow f_i(y, u)$ .

Let us denote by  $G = (-\Delta)^{-1} : L^2(\Omega) \rightarrow H_0^1(\Omega)$  the solution operator associated to the PDE in  $(P(\alpha))$ . Then, the classical result [Stampacchia, 1960–1961](#), Théorème 1 implies the existence of  $C > 0$  with  $\|Gu\|_{L^\infty(\Omega)} \leq C\|u\|_{L^2(\Omega)}$ .

For the moment, the coefficient vector  $\alpha$  will be fixed and we abbreviate

$$f(y, u) := \sum_{i=0}^n \alpha_i f_i(y, u), \quad \phi(x, y, u) := \sum_{i=0}^n \alpha_i \phi_i(x, y, u).$$

Now, we are in position to check the solvability of  $(P(\alpha))$ .

**Theorem 2.3.** For all  $\alpha$  satisfying [Assumption 2.1 1](#), there exists a unique global minimizer  $(\bar{y}, \bar{u}) \in L^\infty(\Omega) \times L^2(\Omega)$  of  $(P(\alpha))$ . Moreover,

$$\|\bar{u}\|_{L^2(\Omega)} \leq \|u_0\|_{L^2(\Omega)} + \frac{2}{\alpha_0} (\|D_y f(y_0, u_0)\|_{L^1(\Omega)} \|G\|_{\mathcal{L}(L^2(\Omega), L^\infty(\Omega))} + \|D_u f(y_0, u_0)\|_{L^2(\Omega)}) \quad (2.3)$$

where  $y_0 := Gu_0$ .

*Proof.* If we consider  $f_i$  as a function from  $L^\infty(\Omega)^2$  to  $\mathbb{R}$ , assumption [\(2.2\)](#) ensures that  $f_i$  is Gâteaux differentiable by using standard arguments (pointwise convergence of the difference quotient associated to  $\phi_i$  and the dominated convergence theorem).

Note that the partial derivatives of  $f_i$  are given by

$$\begin{aligned} D_y f_i(y_0, u_0) &= D_y \phi_i(\cdot, y_0, u_0) \in L^1(\Omega), \\ D_u f_i(y_0, u_0) &= D_u \phi_i(\cdot, y_0, u_0) \in L^\infty(\Omega). \end{aligned}$$

Together with the convexity of  $f_i$  and  $u_0 \in U_{\text{ad}} \cap L^\infty(\Omega)$ , we find

$$f_i(y, u) \geq f_i(y_0, u_0) + D_y f_i(y_0, u_0)(y - y_0) + D_u f_i(y_0, u_0)(u - u_0) \quad \forall y, u \in L^\infty(\Omega),$$

where  $y_0 = Gu_0$ . Using [Lemma 2.2](#), this inequality even holds for all  $u \in L^2(\Omega)$ . Moreover, for  $f_0$  we even have

$$\frac{1}{2} \|u\|_{L^2(\Omega)}^2 = \frac{1}{2} \|u_0\|_{L^2(\Omega)}^2 + (u_0, u - u_0)_{L^2(\Omega)} + \frac{1}{2} \|u - u_0\|_{L^2(\Omega)}^2.$$

Hence,

$$f(y, u) \geq f(y_0, u_0) + D_y f(y_0, u_0)(y - y_0) + D_u f(y_0, u_0)(u - u_0) + \frac{\alpha_0}{2} \|u - u_0\|_{L^2(\Omega)}^2$$

for all  $y \in L^\infty(\Omega)$  and  $u \in L^2(\Omega)$ . This inequality implies that minimizing sequences for  $(P(\alpha))$  are bounded in  $L^\infty(\Omega) \times L^2(\Omega)$ . Now, we can proceed as usual using that  $f$  is weakly lower semicontinuous and  $U_{\text{ad}}$  is weakly closed.

Inequality [\(2.3\)](#) follows from the above estimate with  $y = \bar{y}$ ,  $u = \bar{u}$  and using  $f(y_0, u_0) \geq f(\bar{y}, \bar{u})$ .

Note that the solution operator  $G = (-\Delta)^{-1}$  maps even  $L^{3/2+\varepsilon}(\Omega)$  continuously to  $L^\infty(\Omega)$  for all  $\varepsilon > 0$ , see [Stampacchia, 1960–1961](#), Théorème 1. Hence, the adjoint  $G^*$  maps  $(L^\infty(\Omega))^*$  to  $L^{3-\varepsilon}(\Omega)$  for all  $\varepsilon \in (0, 2]$ . In particular,  $G^*r \in L^{3-\varepsilon}(\Omega)$  for  $r \in L^1(\Omega)$  and  $\varepsilon \in (0, 2]$ . Now, we are in position to provide optimality conditions for  $(P(\alpha))$ . However, our assumptions on  $D_u \phi_i$  are too weak to conclude any regularity of  $D_u \phi_i(\cdot, \bar{y}, \bar{u})$  at this point. Hence, we write down the optimality conditions in a pointwise sense. After we have established more regularity of  $\bar{u}$ , a global variational inequality is given in [Corollary 2.8](#).

**Theorem 2.4.** Let  $(\bar{y}, \bar{u}) \in L^\infty(\Omega) \times L^2(\Omega)$  be the solution of  $(P(\alpha))$ . We define

$$\bar{p} := G^* D_y \phi(\cdot, \bar{y}, \bar{u}) = G^* \sum_{i=1}^n \alpha_i D_y \phi_i(\cdot, \bar{y}, \bar{u}). \quad (2.4)$$

Then,  $\bar{p} \in L^{3-\varepsilon}(\Omega)$  for all  $\varepsilon \in (0, 2]$  and

$$\left( \bar{p}(x) + \sum_{i=0}^n \alpha_i D_u \phi_i(x, \bar{y}(x), \bar{u}(x)) \right) (u - \bar{u}(x)) \geq 0 \quad \forall u \in [u_a(x), u_b(x)], \quad (2.5)$$

holds for a.a.  $x \in \Omega$ .

*Proof.* We fix  $M > \|\bar{y}\|_{L^\infty(\Omega)}$  and choose  $h \in L^\infty(\Omega)$  with  $h = 0$  on  $\{|\bar{u}| > M\}$  and  $\bar{u} + h \in U_{\text{ad}}$ . We set  $y_h := Gh$ .



For  $t \in (0, 1)$  we have

$$0 \leq \frac{f(\bar{y} + t y_h, \bar{u} + t h) - f(\bar{y}, \bar{u})}{t} = \int_{\{|u| \leq M\}} \frac{\phi(\cdot, \bar{y} + t y_h, \bar{u} + t h) - \phi(\cdot, \bar{y}, \bar{u})}{t} dx \\ + \int_{\{|u| > M\}} \frac{\phi(\cdot, \bar{y} + t y_h, \bar{u}) - \phi(\cdot, \bar{y}, \bar{u})}{t} dx$$

The differentiability of  $\phi$  implies the pointwise convergence of the integrands as  $t \searrow 0$ . Assumptions (2.2a) and (2.2d) yield an integrable upper bound. Thus, the dominated convergence theorem gives

$$0 \leq \lim_{t \searrow 0} \frac{f(\bar{y} + t y_h, \bar{u} + t h) - f(\bar{y}, \bar{u})}{t} = \int_{\Omega} D_y \phi(\cdot, \bar{y}, \bar{u}) y_h + D_u \phi(\cdot, \bar{y}, \bar{u}) h dx.$$

Since (2.2d) implies  $D_y \phi(\cdot, \bar{y}, \bar{u}) \in L^1(\Omega)$ , the first addend under the integral can be written as

$$\int_{\Omega} D_y \phi(\cdot, \bar{y}, \bar{u}) y_h dx = \langle D_y \phi(\cdot, \bar{y}, \bar{u}), Gh \rangle_{L^1(\Omega), L^\infty(\Omega)} \\ = \langle D_y \phi(\cdot, \bar{y}, \bar{u}), Gh \rangle_{(L^\infty(\Omega))^*, L^\infty(\Omega)} \\ = \langle G^* D_y \phi(\cdot, \bar{y}, \bar{u}), h \rangle_{L^{3/2-\varepsilon}(\Omega), L^{3+\varepsilon}(\Omega)}$$

where  $\tilde{\varepsilon} > 0$  is arbitrary and  $\varepsilon$  given implicitly by the requirement that both exponents are conjugate. Thus,

$$\int_{\Omega} (\bar{p} + D_u \phi(\cdot, \bar{y}, \bar{u})) h dx \geq 0.$$

Now, the pointwise variational inequality can be derived by the usual arguments.

The next goal is to prove  $\bar{u} \in L^\infty(\Omega)$  and to provide uniform bounds for  $\|\bar{u}\|_{L^\infty(\Omega)}$ . Since  $U_{\text{ad}}$  might not be bounded in  $L^\infty(\Omega)$ , this is not obvious. We are going to use a bootstrap argument to increase the regularity of  $\bar{u}$ . In fact, the regularity of  $\bar{u}$  implies regularity of  $\bar{p}$  via adjoint equation (2.4) and, conversely, regularity of  $\bar{p}$  gives regularity of  $\bar{u}$  via the pointwise variational inequality (2.5). This is covered in the following two lemmas.

**Lemma 2.5.** Let  $(\bar{y}, \bar{u}) \in L^\infty(\Omega) \times L^2(\Omega)$  be the solution of  $(P(\alpha))$  and denote by  $\bar{p}$  the adjoint state given by (2.4). We further assume that  $\bar{u} \in L^s(\Omega)$  for some  $s \in [2, \infty]$ . Then,  $\bar{p} \in L^r(\Omega)$  for all  $r \in [2, \infty]$  satisfying  $1/r > 2/s - 2/d$  and

$$\|\bar{p}\|_{L^r(\Omega)} \leq C (\|\bar{u}\|_{L^s(\Omega)}^2 + 1). \quad (2.6)$$

Here, the constant  $C$  depends on  $\Omega$ ,  $n$ , the exponents  $r, s$  and on  $\mu_M, \hat{C}_M$  from (2.2d) with  $M = \|\bar{y}\|_{L^\infty(\Omega)}$ .

*Proof.* We first provide an estimate for the right-hand side of the adjoint equation (2.4). By using (2.2d) with  $M = \|\bar{y}\|_{L^\infty(\Omega)}$ , we find

$$|D_y \phi_i(x, \bar{y}(x), \bar{u}(x))| \leq \mu_M(x) + \hat{C}_M |\bar{u}(x)|^2$$

for a.a.  $x \in \Omega$ . Hence, with  $t := \min(\hat{p}, s/2) \geq 1$  we have

$$\|D_y \phi_i(\cdot, \bar{y}, \bar{u})\|_{L^t(\Omega)} \leq \|\mu_M\|_{L^t(\Omega)} + \hat{C}_M \|\bar{u}\|_{L^{2t}(\Omega)}^2 \leq C_{s,t} (\|\mu_M\|_{L^{\hat{p}}(\Omega)} + \hat{C}_M \|\bar{u}\|_{L^s(\Omega)}^2),$$

where the constant  $C_{s,t}$  only depends on the exponents  $s, t$  and on the measure of  $\Omega$ . Consequently, by using  $\sum_{i=1}^n \alpha_i \leq 2n$ , we infer

$$\left\| \sum_{i=1}^n \alpha_i D_y \phi_i(\cdot, \bar{y}, \bar{u}) \right\|_{L^t(\Omega)} \leq C_{s,t} 2n (\|\mu_M\|_{L^{\hat{p}}(\Omega)} + \hat{C}_M \|\bar{u}\|_{L^s(\Omega)}^2).$$

Let us check that  $1/r > 1/t - 2/d$ . If  $t = s/2$ , this is the assumption. Otherwise, we have  $t = \hat{p} > d/2$ , thus  $1/t - 2/d < 0 \leq 1/r$ . Thus, we can invoke the regularity result Stampacchia, 1960–1961, Théorème 1 (and a duality argument) to obtain  $\bar{p} \in L^r(\Omega)$  with

$$\|\bar{p}\|_{L^r(\Omega)} \leq \hat{C}_{r,t} \left\| \sum_{i=1}^n \alpha_i D_y \phi_i(\cdot, \bar{y}, \bar{u}) \right\|_{L^t(\Omega)} \leq \hat{C}_{r,t} C_{s,t} 2n (\|\mu_M\|_{L^{\hat{p}}(\Omega)} + \hat{C}_M \|\bar{u}\|_{L^s(\Omega)}^2),$$

where  $\hat{C}_{r,t}$  depends only on the exponents  $r, t$  and on  $\Omega$ .

**Lemma 2.6.** Let  $(\bar{y}, \bar{u}) \in L^\infty(\Omega) \times L^2(\Omega)$  be the solution of  $(P(\alpha))$  and denote by  $\bar{p}$  the adjoint state given by (2.4). We further assume that  $\bar{p} \in L^r(\Omega)$  for some  $r \in [0, \infty]$ . Then,  $\bar{u} \in L^r(\Omega)$  with

$$\|\bar{u}\|_{L^r(\Omega)} \leq \|u_0\|_{L^r(\Omega)} + \frac{1}{\alpha_0} \|\bar{p}\|_{L^r(\Omega)} + \frac{1}{\alpha_0} \|D_u f(\bar{y}, u_0)\|_{L^r(\Omega)}. \quad (2.7)$$

*Proof.* We use the monotonicity of  $D_u \phi_i(x, y, \cdot)$  to obtain

$$\alpha_0 (u_0(x) - \bar{u}(x))^2 \leq [D_u \phi(x, \bar{y}(x), u_0(x)) - D_u \phi(x, \bar{y}(x), \bar{u}(x))] (u_0(x) - \bar{u}(x)).$$

Now, we use (2.5) with  $u = u_0(x)$  and obtain

$$\alpha_0 (u_0(x) - \bar{u}(x))^2 \leq [\bar{p}(x) + D_u \phi(x, \bar{y}(x), u_0(x))] (u_0(x) - \bar{u}(x)).$$

Now, (2.7) follows from division by  $u_0(x) - \bar{u}(x)$  and taking norms.

It remains to apply these two lemmas alternately to obtain the desired regularity of  $\bar{u}$ .

**Theorem 2.7.** Let  $(\bar{y}, \bar{u}) \in L^\infty(\Omega) \times L^2(\Omega)$  be the solution of  $(P(\alpha))$ . Then,

$$\|\bar{u}\|_{L^\infty(\Omega)} \leq C, \quad (2.8)$$

where  $C$  only depends on the following quantities:  $\Omega$ ,  $n$ ,  $\|D_y f(Gu_0, u_0)\|_{L^1(\Omega)}$ ,  $\|D_u f(Gu_0, u_0)\|_{L^2(\Omega)}$ ,  $\mu_M, \hat{C}_M$  from (2.2d) with  $M \geq \|\bar{y}\|_{L^\infty(\Omega)}$ ,  $\|D_u f(\bar{y}, u_0)\|_{L^\infty(\Omega)}$ . In particular,

$$\hat{M} := \sup\{\|u\|_{L^\infty(\Omega)} + \|y\|_{L^\infty(\Omega)} \mid (y, u) \text{ solves } (P(\alpha)), \alpha \in [0, 2]^n\} < \infty. \quad (2.9)$$

*Proof.* From (2.3) we get  $\|\bar{u}\|_{L^2(\Omega)} \leq C$ . Now, we apply Lemmas 2.5 and 2.6 thrice. The first application of Lemma 2.5 proves  $\bar{p} \in L^{5/2}(\Omega)$  since  $2/5 > 2/2 - 2/d$ . Lemma 2.6 gives  $\bar{u} \in L^{5/2}(\Omega)$ . Then,  $\bar{p}, \bar{u} \in L^7(\Omega)$  since  $1/7 > 4/5 - 2/d$ . A last application of the lemmas yields  $\bar{p}, \bar{u} \in L^\infty(\Omega)$  since  $0 = 1/\infty > 2/7 - 2/d$ .

The final constant in (2.8) depends on all quantities in (2.3), (2.6) and (2.7). This yields the list of dependencies in the assertion. We emphasize that  $M$  can be chosen independent of  $\bar{u}$  or  $\bar{y}$ , because as a consequence of (2.3) we obtain a uniform bound on  $\|\bar{y}\|_{L^\infty(\Omega)}$ .

As a corollary, we give a global variant of the pointwise VI (2.5).

**Corollary 2.8.** Let  $(\bar{y}, \bar{u}) \in L^\infty(\Omega) \times L^2(\Omega)$  be given and denote by  $\bar{p}$  the adjoint state, given by (2.4). Then,  $(\bar{y}, \bar{u})$  is the solution of  $(P(\alpha))$  if and only if  $\bar{u} \in L^\infty(\Omega)$  and

$$\left( \bar{p} + \sum_{i=0}^n \alpha_i D_u \phi_i(\cdot, \bar{y}, \bar{u}), u - \bar{u} \right)_{L^2(\Omega)} \geq 0 \quad \forall u \in U_{\text{ad}}. \quad (2.10)$$

*Proof.* If the tuple is optimal, then  $\bar{u} \in L^\infty(\Omega)$  by Theorem 2.7. By the boundedness of  $(\bar{y}, \bar{u})$ , the functions  $D_u \phi_i(\cdot, \bar{y}, \bar{u})$  belong to  $L^\infty(\Omega)$ . The claim now follows from integrating (2.5).

The converse direction follows from convexity arguments and  $\alpha_0 > 0$ .

## 2.2 LIPSCHITZ DEPENDENCE ON THE PARAMETERS $\alpha$

In this subsection our goal is to show that the solution operator

$$S : [0, 1]^n \rightarrow L^\infty(\Omega), \quad \alpha \mapsto u, \quad (y, u) \text{ solves } (P(\alpha)) \quad (2.11)$$

is Lipschitz continuous. Note that the solution  $u$  is in  $L^\infty(\Omega)$  due to Theorem 2.7. Such a Lipschitz estimate has previously been shown in Unger, 1997; Tröltzsch, 2000. Since

these references dealt with a slightly different setting, we cannot directly apply their results.

We start with showing the Lipschitzian dependence on  $\alpha$  for the  $L^2(\Omega)$ -norm and will later show the same for the  $L^\infty(\Omega)$  norm.

**Lemma 2.9.** There exists a constant  $C > 0$  such that for all  $\alpha^1, \alpha^2 \in [0, 2]^n$  we have

$$\|u_1 - u_2\|_{L^2(\Omega)} \leq C|\alpha^1 - \alpha^2|, \quad (2.12)$$

where  $(y_1, u_1)$  is a solution of  $(P(\alpha^1))$  and  $(y_2, u_2)$  is a solution of  $(P(\alpha^2))$ . Here,  $|\cdot|$  denotes the  $\ell^1$ -norm on  $\mathbb{R}^n$ . That is, the solution mapping  $S : [0, 2]^n \rightarrow L^2(\Omega)$  is (globally) Lipschitz continuous.

*Proof.* Let  $j \in \{1, 2\}$ . Due to the term  $\frac{\alpha_0}{2}\|u\|_{L^2(\Omega)}^2$  in  $(P(\alpha))$  it can be shown that the quadratic growth condition

$$\sum_{i=0}^n \alpha_i^j f_i(y, u) - \sum_{i=0}^n \alpha_i^j f_i(y_j, u_j) \geq \frac{\alpha_0}{2} \|u - u_j\|_{L^2(\Omega)}^2$$

holds for all feasible  $(y, u)$ . Adding the inequalities for the cases  $j = 1, 2$  with the choice  $(y, u) = (y_{3-j}, u_{3-j})$  yields

$$\begin{aligned} \alpha_0 \|u_1 - u_2\|_{L^2(\Omega)}^2 &\leq \sum_{i=0}^n (\alpha_i^1 - \alpha_i^2) (f_i(y_2, u_2) - f_i(y_1, u_1)) \\ &\leq |\alpha^1 - \alpha^2| \max_{i \in \{0, \dots, n\}} |f_i(y_2, u_2) - f_i(y_1, u_1)|. \end{aligned}$$

Now we can use [Theorem 2.7](#) and our assumptions [\(2.2a\)](#), [\(2.2d\)](#) with  $M = \hat{M}$  from [\(2.9\)](#). Therefore, we have

$$|f_i(y_2, u_2) - f_i(y_1, u_2)| \leq \|y_1 - y_2\|_{L^\infty(\Omega)} \|\mu_M + \hat{C}_M |u_2|^2\|_{L^1(\Omega)} \leq C \|u_1 - u_2\|_{L^2(\Omega)}$$

and

$$|f_i(y_1, u_2) - f_i(y_1, u_1)| \leq \text{meas}(\Omega)^{1/2} C_M \|u_2 - u_1\|_{L^2(\Omega)}.$$

If we combine this with the inequality above we have

$$\alpha_0 \|u_1 - u_2\|_{L^2(\Omega)}^2 \leq C |\alpha^1 - \alpha^2| \|u_1 - u_2\|_{L^2(\Omega)}$$

and dividing by  $\|u_1 - u_2\|_{L^2(\Omega)}$  completes the proof.

Now, we can use similar arguments as in the previous section in order to obtain the Lipschitz stability in  $L^\infty(\Omega)$ .

**Lemma 2.10.** For  $j = 1, 2$  let  $(y_j, u_j)$  be a solution of  $(P(\alpha^j))$  and  $p_j$  be defined like in (2.4). Furthermore, let  $s, r \in [2, \infty]$  be exponents that satisfy  $1/r > 1/s - 2/d$ . Then we have

$$\|p_1 - p_2\|_{L^r(\Omega)} \leq C(\|u_1 - u_2\|_{L^s(\Omega)} + |\alpha^1 - \alpha^2|).$$

for a constant  $C > 0$  which does not depend on  $\alpha^1, \alpha^2$ .

*Proof.* We start with

$$\begin{aligned} \|p_1 - p_2\|_{L^r(\Omega)} &= \left\| G^* \sum_{i=1}^n (\alpha_i^1 D_y \phi_i(\cdot, y_1, u_1) - \alpha_i^2 D_y \phi_i(\cdot, y_2, u_2)) \right\|_{L^r(\Omega)} \\ &\leq C \max_{i \in \{1, \dots, n\}} \|\alpha_i^1 D_y \phi_i(\cdot, y_1, u_1) - \alpha_i^2 D_y \phi_i(\cdot, y_2, u_2)\|_{L^s(\Omega)} \end{aligned}$$

where we used our conditions for  $r, s$  and the regularity of  $G^*$  that result from Stampacchia, 1960–1961, Théorème 1, cf. the proof of Lemma 2.5.

Because  $y_j, u_j$  are bounded in  $L^\infty(\Omega)$  according to Theorem 2.7, we know by (2.2b) that  $D_y \phi_i(x, \cdot, \cdot)$  is Lipschitz continuous for a.a.  $x \in \Omega$  with Lipschitz constant  $C_M$ , where  $M = \hat{M}$  from (2.9). This yields

$$\begin{aligned} &\|\alpha_i^1 D_y \phi_i(\cdot, y_1, u_1) - \alpha_i^2 D_y \phi_i(\cdot, y_2, u_2)\|_{L^s(\Omega)} \\ &\leq |\alpha^1 - \alpha^2| \|D_y \phi_i(\cdot, y_1, u_1)\|_{L^s(\Omega)} + C_M |\alpha^2| (\|y_1 - y_2\|_{L^s(\Omega)} + \|u_1 - u_2\|_{L^s(\Omega)}) \\ &\leq C(|\alpha^1 - \alpha^2| + \|u_1 - u_2\|_{L^s(\Omega)}) \end{aligned}$$

which completes the proof.

**Lemma 2.11.** For  $j = 1, 2$  let  $(y_j, u_j)$  be a solution of  $(P(\alpha^j))$  and  $p_j$  be defined like in (2.4). Then for  $r \in [2, \infty]$  we have

$$\|u_1 - u_2\|_{L^r(\Omega)} \leq C(\|p_1 - p_2\|_{L^r(\Omega)} + |\alpha^1 - \alpha^2| + \|y_1 - y_2\|_{L^r(\Omega)})$$

for a constant  $C > 0$ .

*Proof.* By using the monotonicity of  $D_u \phi_i(x, y, \cdot)$  and applying (2.5) twice we obtain

$$\begin{aligned} \alpha_0 |u_1 - u_2|^2 &\leq \sum_{i=0}^n \alpha_i^1 (D_u \phi_i(\cdot, y_1, u_1) - D_u \phi_i(\cdot, y_1, u_2))(u_1 - u_2) \\ &\leq \left( -p_1 - \sum_{i=0}^n \alpha_i^1 D_u \phi_i(\cdot, y_1, u_2) \right) (u_1 - u_2) \\ &\leq \left( p_2 - p_1 + \sum_{i=0}^n \alpha_i^2 D_u \phi_i(\cdot, y_2, u_2) - \sum_{i=0}^n \alpha_i^1 D_u \phi_i(\cdot, y_1, u_2) \right) (u_1 - u_2). \end{aligned}$$

This yields the estimate

$$\alpha_0 \|u_1 - u_2\|_{L^r(\Omega)} \leq \|p_1 - p_2\|_{L^r(\Omega)} + \sum_{i=1}^n \|\alpha_i^2 D_u \phi_i(\cdot, y_2, u_2) - \alpha_i^1 D_u \phi_i(\cdot, y_1, u_2)\|_{L^r(\Omega)}.$$

Now let  $M > 0$  be such that  $M > |u_j|, M > |y_j|, j = 1, 2$  and let  $i \in \{1, \dots, n\}$  be given. Using the triangle inequality for the last term, we have

$$|\alpha_i^2 D_u \phi_i(\cdot, y_2, u_2) - \alpha_i^1 D_u \phi_i(\cdot, y_1, u_2)| \leq |\alpha_i^1 - \alpha_i^2| C_M + \alpha_i^2 |y_1 - y_2| C_M$$

where we used (2.2a) and  $|D_{yu} \phi_i(x, y_2, u_2)| \leq C_M$ . Combined with the estimate above this completes the proof.

Using these two lemmas, we obtain the Lipschitz stability in  $L^\infty(\Omega)$ .

**Theorem 2.12.** The solution operator  $S : [0, 2]^n \rightarrow L^\infty(\Omega)$  given by (2.11) is (globally) Lipschitz continuous.

*Proof.* We proceed similarly to the proof of Theorem 2.7. By Lemma 2.9 we have local Lipschitz continuity with respect to the  $L^2(\Omega)$  norm. Then, we can apply Lemmas 2.10 and 2.11 twice to repeatedly prove Lipschitz continuity of the maps  $\alpha \mapsto p$  and  $\alpha \mapsto u$  for the norms  $L^5(\Omega), L^\infty(\Omega)$ , respectively.

### 2.3 DIRECTIONAL DIFFERENTIABILITY OF THE SOLUTION OPERATOR

Our next goal is to provide the directional differentiability of the solution operator  $S$  of (2.11). Traditionally, this is established by first proving the directional differentiability of an auxiliary, linearized problem and then applying Donchev's implicit function theorem. We refer exemplarily to Malanowski, 2002 for this approach. We are going to apply the differentiability results of Christof, Wachsmuth, 2017, Theorem 2.13.

In a first step, we characterize the solution of (2.11) by a variational inequality. Therefore, let  $\alpha^0 \in [0, 1]^n$  be fixed and we define  $\bar{u}_0 := S(\alpha^0)$ . Then, there is a constant  $C > 0$ , such that the solutions of (2.11) are bounded by  $C$  in  $L^\infty(\Omega)$  for all  $\alpha \in [0, 2]^n$ . We define

$$U_{\text{ad}}^C := U_{\text{ad}} \cap \{u \in L^\infty(\Omega) \mid -C - 1 \leq u \leq C + 1\}$$

and the mapping  $A : \mathbb{R}^n \times L^2(\Omega) \rightarrow L^2(\Omega)$  via

$$A(\alpha, u) = \begin{cases} G^* \sum_{i=1}^n \alpha_i D_y \phi_i(Gu, u) + \sum_{i=0}^n \alpha_i D_u \phi_i(Gu, u) & \text{if } u \in U_{\text{ad}}^C, \\ 0 & \text{else.} \end{cases}$$

Then for  $\alpha \in [0, 2]^n$ , it is clear from Corollary 2.8 that  $u_\alpha = S(\alpha)$  if and only if

$$u_\alpha \in U_{\text{ad}}^C \quad \text{and} \quad (A(\alpha, u_\alpha), u - u_\alpha) \geq 0 \quad \forall u \in U_{\text{ad}}^C. \quad (2.13)$$

Moreover, the solution operator  $S$  is Lipschitz continuous from  $[0, 2]^n$  to  $L^\infty(\Omega)$ . Further,  $A : \mathbb{R}^n \times L^\infty(\Omega) \rightarrow L^2(\Omega)$  is Fréchet differentiable in an  $\mathbb{R}^n \times L^\infty(\Omega)$  neighborhood of  $(\alpha^0, \bar{u}_0)$  and we denote the partial derivatives in  $(\alpha^0, \bar{u}_0)$  by  $A_\alpha = D_\alpha A(\alpha^0, \bar{u}_0)$  and  $A_u = D_u A(\alpha^0, \bar{u}_0)$ .

**Theorem 2.13.** With the notation given above, the solution map  $S : [0, 2]^n \rightarrow L^2(\Omega)$  is directionally differentiable in  $\alpha^0 \in [0, 1]^n$  for all feasible directions  $\beta \in \mathcal{T}_{[0,2]^n}(\alpha^0)$  and the directional derivative  $h = S'(\alpha^0; \beta)$  is given by the unique solution of the VI

$$h \in \mathcal{C}_{U_{\text{ad}}}(\bar{u}_0, A(\alpha^0, \bar{u}_0)) \quad \text{and} \quad (A_\alpha \beta + A_u h, v - h) \geq 0 \quad \forall v \in \mathcal{C}_{U_{\text{ad}}}(\bar{u}_0, A(\alpha^0, \bar{u}_0)). \quad (2.14)$$

Here,

$$\mathcal{C}_{U_{\text{ad}}}(\bar{u}_0, A(\alpha^0, \bar{u}_0)) = \mathcal{T}_{U_{\text{ad}}}(\bar{u}_0) \cap A(\alpha^0, \bar{u}_0)^\perp$$

denotes the critical cone and  $\mathcal{T}_{[0,2]^n}(\alpha^0)$ ,  $\mathcal{T}_{U_{\text{ad}}}(\bar{u}_0)$  are tangent cones in the sense of convex analysis.

*Proof.* In order to apply [Christof, Wachsmuth, 2017](#), Theorem 2.13 under its condition (ii) with the setting  $X = Y = L^2(\Omega)$  and  $j = \delta_{U_{\text{ad}}^C}$  (indicator function of  $U_{\text{ad}}^C$  in the sense of convex analysis), we have to check the following facts:

- Point (i) in [Christof, Wachsmuth, 2017](#), Assumption 2.2 is void since we only consider directional differentiability. The Lipschitz estimate in point (ii) follows from [Theorem 2.12](#). Finally, the differentiability assumption (iii) follows from the Lipschitz continuity of  $S$  in  $L^\infty(\Omega)$  and the Fréchet differentiability of  $A$  w.r.t.  $L^\infty(\Omega)$ .
- Since the set  $U_{\text{ad}}^C$  is polyhedral in  $L^2(\Omega)$ , the function  $j = \delta_{U_{\text{ad}}^C}$  is strongly twice epi-differentiable in  $\bar{u}_0$  w.r.t.  $-A(\alpha^0, \bar{u}_0)$  due to [Christof, Wachsmuth, 2017](#), Corollary 3.3. Moreover, the second subderivative of  $j$  in  $\bar{u}_0$  w.r.t.  $-A(\alpha^0, \bar{u}_0)$  is given by the indicator function of  $\mathcal{C}_{U_{\text{ad}}^C}(\bar{u}_0, A(\alpha^0, \bar{u}_0))$  and this set coincides with  $\mathcal{C}_{U_{\text{ad}}}(\bar{u}_0, A(\alpha^0, \bar{u}_0))$  since  $\|\bar{u}_0\|_{L^\infty(\Omega)} \leq C$ .
- One can check that  $(v, A_u v) \geq \alpha_0 \|v\|_{L^2(\Omega)}^2$ . This implies the weak lower semicontinuity of  $v \mapsto (v, A_u v)$  and that (2.14) admits at most one solution.

The application of [Christof, Wachsmuth, 2017](#), Theorem 2.13 yields the claim.

Due to the Lipschitz continuity of  $S : [0, 2]^n \rightarrow L^\infty(\Omega)$  one even has the convergence of

$$\frac{S(\alpha^0 + t\beta) - S(\alpha^0)}{t}$$

to  $S'(\alpha^0; \beta)$  strongly in all  $L^p(\Omega)$ ,  $p \in [1, \infty)$  and weak- $\star$  in  $L^\infty(\Omega)$ .

We remark that due to the fact that  $S$  is Lipschitz continuous, we even have Hadamard directional differentiability for  $S$ .

### 3 UPPER-LEVEL PROBLEM

In this section we discuss some basic properties of the upper-level problem (IOP). Let us introduce the reduced objective function

$$\psi : [0, 2]^n \rightarrow \mathbb{R}, \quad \alpha \mapsto \frac{1}{2} \|G S(\alpha) - y_d\|_{L^2(\Omega)}^2$$

and the feasible set

$$K := \left\{ \alpha \in \mathbb{R}^n \mid 0 \leq \alpha \text{ and } \sum_{i=1}^n \alpha_i = 1 \right\}.$$

Now, the upper-level problem (IOP) is equivalent to the minimization of  $\psi$  over  $K$ . We remark that the results in this section also hold true for arbitrary compact and convex  $K \subset [0, 1]^n$ . We collect some results that follow directly from the results in Section 2.

#### Corollary 3.1.

1. The reduced objective function  $\psi$  is globally Lipschitz continuous.
2. For every  $\alpha \in [0, 2]^n$  and a feasible direction  $\beta \in \mathcal{T}_{[0, 2]^n}(\alpha)$  the reduced objective function  $\psi$  is directionally differentiable. The directional derivative is given by

$$\psi'(\alpha; \beta) = (G S(\alpha) - y_d, G S'(\alpha; \beta))_{L^2(\Omega)}.$$

3. The upper-level problem (IOP) has a solution.
4. Let  $(\bar{y}, \bar{u}, \bar{\alpha})$  be a local solution of (IOP). Then, the primal optimality condition

$$\psi'(\bar{\alpha}; \beta) \geq 0 \quad \forall \beta \in \mathcal{T}_K(\bar{\alpha}) \tag{3.1}$$

holds, where  $\mathcal{T}_K(\bar{\alpha})$  is the tangent cone to  $K$  at  $\bar{\alpha}$ .

In case that the lower-level problem is unconstrained, i.e.,  $U_{\text{ad}} = L^2(\Omega)$ , the directional derivatives  $S'(\alpha, \cdot)$  and  $\psi'(\alpha, \cdot)$  are even linear. This is caused by the fact that the critical cone in (2.14) is the entire space  $L^2(\Omega)$ . Now, we utilize this linearity in order to transform (3.1) into an optimality condition involving multipliers. To this end, we introduce the objective of the lower-level problem in dependence of all parameters of the upper-level problem

$$F(\alpha, y, u) := \sum_{i=0}^n \alpha_i f_i(y, u) = \sum_{i=1}^n \alpha_i f_i(y, u) + \frac{\alpha_0}{2} \|u\|_{L^2(\Omega)}^2. \tag{3.2}$$

**Theorem 3.2.** Consider the unconstrained case  $U_{\text{ad}} = L^2(\Omega)$  and let  $(\bar{\alpha}, \bar{y}, \bar{u})$  be a local solution of (IOP). Then, there are functions  $(\bar{\mu}, \bar{\nu}, \bar{\rho}) \in H_0^1(\Omega) \times L^2(\Omega) \times H_0^1(\Omega)$  such



that

$$\begin{pmatrix} D_{yy}F(\cdot) & D_{yu}F(\cdot) & -\Delta \\ D_{uy}F(\cdot) & D_{uu}F(\cdot) & -I \\ -\Delta & -I & 0 \end{pmatrix} \begin{pmatrix} \bar{\mu} \\ \bar{\nu} \\ \bar{\rho} \end{pmatrix} = \begin{pmatrix} \bar{y} - y_d \\ 0 \\ 0 \end{pmatrix} \quad (3.3a)$$

$$-D_{\alpha y}F(\cdot)\bar{\mu} - D_{\alpha u}F(\cdot)\bar{\nu} \in -\mathcal{N}_K(\bar{\alpha}). \quad (3.3b)$$

Here,  $(\cdot)$  is an abbreviation for the argument  $(\bar{\alpha}, \bar{y}, \bar{u})$  and  $\mathcal{N}_K(\bar{\alpha})$  is the normal cone of  $K$  at  $\bar{\alpha}$  (in the sense of convex analysis).

*Proof.* For given  $\beta \in \mathcal{T}_K(\bar{\alpha})$ , we set  $\delta u := S'(\bar{\alpha}; \beta)$ . Using that the critical cone in (2.14) becomes  $L^2(\Omega)$ , a short calculation shows that  $\delta u$  together with some functions  $\delta y, \delta p \in H_0^1(\Omega)$  solves

$$\begin{pmatrix} D_{yy}F(\cdot) & D_{yu}F(\cdot) & -\Delta \\ D_{uy}F(\cdot) & D_{uu}F(\cdot) & -I \\ -\Delta & -I & 0 \end{pmatrix} \begin{pmatrix} \delta y \\ \delta u \\ \delta p \end{pmatrix} = \begin{pmatrix} -D_{y\alpha}F(\cdot)\beta \\ -D_{u\alpha}F(\cdot)\beta \\ 0 \end{pmatrix}.$$

Now, Corollary 3.1 implies

$$0 \leq \psi'(\bar{\alpha}; \beta) = (\bar{y} - y_d, \delta y)_{L^2(\Omega)}.$$

Using the convexity of the function  $F$  w.r.t.  $y$  and  $u$ , and the coercivity w.r.t.  $u$ , it can be shown that the operator on the left-hand side of (3.3a) is continuously invertible. By defining  $(\bar{\mu}, \bar{\nu}, \bar{\rho})$  via (3.3a), we find

$$0 \leq (\bar{y} - y_d, \delta y)_{L^2(\Omega)} = (-D_{\alpha y}F(\cdot)\bar{\mu} - D_{\alpha u}F(\cdot)\bar{\nu}, \beta)_{\mathbb{R}^n}.$$

Now, (3.3b) follows from the definition of the normal cone.

We remark that the obtained stationarity system is the system of strong stationarity, see Definition 3.3.

### 3.1 FORMAL DERIVATION OF STATIONARITY SYSTEMS

In this subsection we want to describe some stationarity systems for (IOP) that arise from a formal calculation. These will be similar to stationarity conditions for MPCCs in finite dimensions. For stationarity concepts in infinite dimensions, we refer to Mehlig, 2017, Section 3.1.

Due to Corollary 2.8, we rewrite the bilevel optimization problem (IOP) into the equiv-

alent infinite-dimensional MPCC

$$\min_{\alpha, y, u, p, \lambda} \quad \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 \quad (3.4a)$$

$$\text{s.t.} \quad -\Delta p - D_y F(\alpha, y, u) = 0, \quad (3.4b)$$

$$p + D_u F(\alpha, y, u) + \lambda = 0, \quad (3.4c)$$

$$-u - \Delta y = 0, \quad (3.4d)$$

$$(u, \lambda) \in \text{gph } \mathcal{N}_{U_{\text{ad}}}, \quad (3.4e)$$

$$\alpha \in K. \quad (3.4f)$$

Here,

$$\text{gph } \mathcal{N}_{U_{\text{ad}}} = \{(u, \lambda) \in L^2(\Omega)^2 \mid u \in U_{\text{ad}}, \lambda \in \mathcal{N}_{U_{\text{ad}}}(u)\}$$

is the graph of the normal cone mapping to the admissible set  $U_{\text{ad}}$ . We recall that the condition (3.4e) is equivalent to the pointwise a.e. conditions

$$0 \leq u_b - u \perp \lambda^+ \geq 0, \quad 0 \leq u - u_a \perp \lambda^- \geq 0. \quad (3.5)$$

Here,  $\lambda^+$ ,  $\lambda^-$  are the positive and the negative part of  $\lambda$ , i.e.,  $\lambda^+ = \max(\lambda, 0)$  and  $\lambda^- = \max(-\lambda, 0)$ .

We note that due to Corollary 2.8, (3.4b)–(3.4e) are a characterization of a solution  $(y, u)$  of the lower-level problem.

The Lagrange function of (3.4) is given by

$$\begin{aligned} \mathcal{L}(\alpha, y, u, \lambda, p; \mu, \hat{\nu}, \nu, \rho, z, \gamma) &= \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + (u, z)_{L^2(\Omega)} + (\nu, \lambda)_{L^2(\Omega)} + (\alpha, \gamma)_{\mathbb{R}^n} \\ &\quad + \langle -\Delta p - D_y F(\alpha, y, u), \mu \rangle_{L^1(\Omega) \times L^\infty(\Omega)} \\ &\quad + (p + D_u F(\alpha, y, u) + \lambda, \hat{\nu})_{L^2(\Omega)} + (u + \Delta y, \rho)_{L^2(\Omega)}. \end{aligned}$$

Setting the partial derivatives  $D_\lambda \mathcal{L}$ ,  $D_y \mathcal{L}$ ,  $D_u \mathcal{L}$ ,  $D_p \mathcal{L}$ ,  $D_\alpha \mathcal{L}$  to zero yields

$$\begin{aligned} \nu + \hat{\nu} &= 0, \\ y - y_d - D_{yy} F(\alpha, y, u) \mu + D_{yu} F(\alpha, y, u) \hat{\nu} + \Delta \rho &= 0, \\ z - D_{uy} F(\alpha, y, u) \mu + D_{uu} F(\alpha, y, u) \hat{\nu} + \rho &= 0, \\ -\Delta \mu + \hat{\nu} &= 0, \\ -D_{\alpha y} F(\alpha, y, u) \mu + D_{\alpha u} F(\alpha, y, u) \hat{\nu} + \gamma &= 0. \end{aligned}$$

As usual, we also have the condition  $\gamma \in \mathcal{N}_K(\alpha)$  for the multiplier  $\gamma$  of the constraint  $\alpha \in K$ . Similarly, we get a condition for the multipliers  $(z, \nu)$ . However, since  $\text{gph } \mathcal{N}_{U_{\text{ad}}}$  is not convex, this condition has the form

$$(z, \nu) \in \mathcal{N}_{\text{gph } \mathcal{N}_{U_{\text{ad}}}}^\#(u, \lambda) \quad (3.6)$$

for an unspecified normal cone  $\mathcal{N}^\#$  to the nonconvex set  $\text{gph } \mathcal{N}_{U_{\text{ad}}}$ . Let us give some examples for these normal cones to  $\text{gph } \mathcal{N}_{U_{\text{ad}}}$ . Since this nonconvex set consists of pointwise

constraints in  $L^2(\Omega)^2$ , we can apply the results of Mehlitz, Wachsmuth, 2017, Section 5 to characterize the various normal cones to  $\text{gph } \mathcal{N}_{U_{\text{ad}}}$ . Using the Clarke normal cone instead of  $\mathcal{N}^\sharp$  yields the conditions

$$\begin{aligned} z &= 0 \text{ a.e. on } \{u_a < u < u_b\}, \\ \nu &= 0 \text{ a.e. on } \{\lambda \neq 0\}, \end{aligned}$$

that correspond to weak stationarity. Moreover, in this situation, the limiting normal cone to  $\text{gph } \mathcal{N}_{U_{\text{ad}}}$  coincides with the Clarke normal cone. Using the Fréchet normal cone instead of  $\mathcal{N}^\sharp$  results in the additional sign conditions

$$\begin{aligned} \nu &\geq 0, \quad z \leq 0 \text{ a.e. on } \{\lambda = 0\} \cap \{u = u_a\}, \\ \nu &\leq 0, \quad z \geq 0 \text{ a.e. on } \{\lambda = 0\} \cap \{u = u_b\}. \end{aligned}$$

These sign conditions will appear in the system of strong stationarity. Note that these stationarity conditions are pointwise versions of the stationarity concepts with the same name in the theory of MPCCs in finite dimensions.

We will summarize the above formal stationarity conditions in the following definition. As a slight simplification, we will avoid the variables  $\hat{\nu}$  and  $\gamma$  by simple substitution and we will add the intermediate stationarity concepts C- and M-stationarity between strong and weak stationarity that correspond to their finite-dimensional analogues.

**Definition 3.3.** A point  $(\bar{\alpha}, \bar{y}, \bar{u}) \in \mathbb{R}^n \times H_0^1(\Omega) \times L^2(\Omega)$  is called *weakly stationary* if there exist functions  $\bar{p}, \bar{\rho}, \bar{\mu} \in H_0^1(\Omega)$  and  $\bar{\nu}, \bar{\lambda}, \bar{z} \in L^2(\Omega)$  such that

$$-\Delta \bar{p} - D_y F(\cdot) = 0, \quad (3.7a)$$

$$\bar{p} + D_u F(\cdot) + \bar{\lambda} = 0, \quad (3.7b)$$

$$-\Delta \bar{y} = \bar{u}, \quad (3.7c)$$

$$\bar{u} \in U_{\text{ad}}, \quad \bar{\lambda} \in \mathcal{N}_{U_{\text{ad}}}(\bar{u}), \quad (3.7d)$$

$$\bar{\alpha} \in K, \quad (3.7e)$$

$$\begin{pmatrix} D_{yy}F(\cdot) & D_{yu}F(\cdot) & -\Delta \\ D_{uy}F(\cdot) & D_{uu}F(\cdot) & -I \\ -\Delta & -I & 0 \end{pmatrix} \begin{pmatrix} \bar{\mu} \\ \bar{\nu} \\ \bar{\rho} \end{pmatrix} = \begin{pmatrix} \bar{y} - y_d \\ \bar{z} \\ 0 \end{pmatrix}, \quad (3.7f)$$

$$D_{\alpha y}F(\cdot)\bar{\mu} + D_{\alpha u}F(\cdot)\bar{\nu} \in \mathcal{N}_K(\bar{\alpha}), \quad (3.7g)$$

$$\bar{\nu} = 0 \text{ a.e. on } \{\bar{\lambda} \neq 0\}, \quad (3.7h)$$

$$\bar{z} = 0 \text{ a.e. on } \{u_a < \bar{u} < u_b\}. \quad (3.7i)$$

Again, we use  $(\cdot)$  as an abbreviation for  $(\bar{\alpha}, \bar{y}, \bar{u})$ . We call the point  $(\bar{\alpha}, \bar{y}, \bar{u})$  *C-stationary*, if it is weakly stationary and the multipliers  $\bar{\nu}, \bar{z}$  satisfy

$$\bar{\nu} \bar{z} \leq 0 \text{ a.e. on } \Omega. \quad (3.8)$$

For  $M$ -stationarity of  $(\bar{\alpha}, \bar{y}, \bar{u})$  we require

$$\begin{aligned} \bar{\nu} \bar{z} &= 0 \vee (\bar{\nu} \geq 0 \wedge \bar{z} \leq 0) \text{ a.e. on } \{\bar{\lambda} = 0\} \cap \{\bar{u} = u_a\} \\ \bar{\nu} \bar{z} &= 0 \vee (\bar{\nu} \leq 0 \wedge \bar{z} \geq 0) \text{ a.e. on } \{\bar{\lambda} = 0\} \cap \{\bar{u} = u_b\} \end{aligned} \quad (3.9)$$

in addition to the conditions for weak stationarity. Finally, we call the point  $(\bar{\alpha}, \bar{y}, \bar{u})$  *strongly stationary* if it is weakly stationary and the multipliers  $\bar{\nu}, \bar{z}$  satisfy the conditions

$$\begin{aligned} \bar{\nu} &\geq 0, \bar{z} \leq 0 \text{ a.e. on } \{\bar{\lambda} = 0\} \cap \{\bar{u} = u_a\} \\ \bar{\nu} &\leq 0, \bar{z} \geq 0 \text{ a.e. on } \{\bar{\lambda} = 0\} \cap \{\bar{u} = u_b\}. \end{aligned} \quad (3.10)$$

Note that all stationarity concepts coincide in the case that the biactive sets  $\{\bar{\lambda} = 0\} \cap \{\bar{u} = u_a\}$  and  $\{\bar{\lambda} = 0\} \cap \{\bar{u} = u_b\}$  have measure zero. This is in particular the case in the setting of [Theorem 3.2](#). More general, one can prove strong stationarity similar to [Theorem 3.2](#) when the biactive set has measure zero, because the directional derivative  $\psi'(\alpha; \beta)$  is linear in  $\beta$  in this case. We note that strong stationarity in this case is also a trivial consequence of [Theorem 4.8](#).

We observe that every feasible point  $(\bar{y}, \bar{u})$  of (IOP) where the inactive set has measure zero (i.e.,  $\bar{u} = u_a$  or  $\bar{u} = u_b$  a.e. in  $\Omega$ ) is an M-stationary point. This can be seen by setting  $\bar{\mu} = 0, \bar{\nu} = 0, \bar{\rho} = G(\bar{y} - y_d), \bar{z} = -\bar{\rho}$  in [Definition 3.3](#).

### 3.2 COUNTEREXAMPLE TO STRONG STATIONARITY

We give a counterexample showing that a local minimizer does not need to fulfill the system of strong stationarity.

**Example 3.4.** Let  $n = 2, \alpha_0 = \frac{1}{10}, \Omega = (-1, 1) \subset \mathbb{R}^1, f_1(y, u) := \|y + 1\|_{L^2(\Omega)}^2, f_2(y, u) = \|y - 1\|_{L^2(\Omega)}^2, u_a = 0, u_b = \infty,$  and  $y_d = \chi_{(-1,0)} - \chi_{(0,1)}$ . Then, the point  $\bar{\alpha} = (\frac{1}{2}, \frac{1}{2})$  is a global minimizer of the bilevel optimal control problem, but it is not a strongly stationary point.

*Proof.* First, we show that  $\bar{\alpha}$  is indeed a solution of the upper-level problem. Note that for  $\bar{\alpha}$  the solution of the lower-level problem is given by  $(\bar{y}, \bar{u}) = (0, 0)$ . For arbitrary  $\alpha \in K$  it follows from the symmetry and convexity of the functions  $f_i$  that solutions  $(y, u)$  of the lower-level problem are even functions, i.e., they satisfy  $u(x) = u(-x), y(x) = y(-x)$  a.e. in  $\Omega$ . Then, the objective function of the upper-level problem satisfies

$$\frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 = \frac{1}{2} \int_0^1 |y - 1|^2 + |y + 1|^2 dx = \int_0^1 y^2 + 1 dx.$$

Hence,  $\bar{y} = 0$  is the global minimizer of this function. In particular,  $\bar{\alpha}$  is a global minimizer of the bilevel optimal control problem.

For the disproof of strong stationarity, let  $\bar{\rho}, \bar{\nu}, \bar{\mu}, \bar{z}$  be given such that the system of strong stationarity is satisfied. From (3.7a), (3.7b) with  $\bar{y} = 0, \bar{u} = 0$  it follows that  $\bar{\lambda} = 0$  and therefore the biactive set is equal to  $\Omega = (-1, 1)$ . Next we consider condition (3.7g). We have  $\mathcal{N}_K(\bar{\alpha}) = \text{lin}\{(1, 1)\}$  and therefore

$$((-2, \bar{\mu})_{L^2(\Omega)}, (2, \bar{\mu})_{L^2(\Omega)}) \in \text{lin}\{(1, 1)\}.$$

It follows that  $(1, \bar{\mu})_{L^2(\Omega)} = 0$ . Due to strong stationarity, we have  $\bar{\nu} \geq 0, \bar{z} \leq 0$ . By maximum principle it follows that  $\bar{\mu} \geq 0$  and therefore  $(1, \bar{\mu})_{L^2(\Omega)} = 0$  implies  $\bar{\mu} = 0$  and  $\bar{\nu} = 0$ . Now it follows from (3.7f) that  $-\Delta \bar{\rho} = -y_d$  and  $\bar{\rho} = -\bar{z}$ . Thus we can directly calculate that  $\bar{\rho}(x) = \frac{1}{2}x(1 - |x|)$  on  $\Omega$ . Since  $\bar{z} \leq 0$  by assumption, this is a contradiction.

However, the point  $(\bar{\alpha}, \bar{y}, \bar{u})$  in this example is M-stationary. This follows from the remark after Definition 3.3 because  $\bar{u}$  is active a.e., i.e.  $\{\bar{u} = u_a\} = \Omega$ .

## 4 C-STATIONARITY OF LOCAL SOLUTIONS

In this section, we are going to prove C-stationarity of a local minimizer in the general case  $U_{\text{ad}} \neq L^2(\Omega)$ . Our strategy is to apply the optimality condition from Theorem 3.2 to a penalized problem. Consequently, an optimality condition is obtained by passing to the limit.

First, we are going to state a penalized version of the lower-level problem  $(P(\alpha))$ . W.l.o.g., we may assume  $u_a, u_b \in L^\infty(\Omega)$ , see Theorem 2.7. We will use the penalty function  $P(u) := \int_\Omega \pi(u) dx$ , where

$$\pi(s) = \begin{cases} 0 & s \leq 0, \\ 2s^3 - s^4 & 0 < s < 1, \\ 2s - 1 & s \geq 1. \end{cases}$$

Note that  $\pi$  is twice continuously differentiable and its second derivative is bounded. As a penalty functional we will use  $\Pi(u) := P(u_a - u) + P(u - u_b)$ . It is easy to see that  $u \in U_{\text{ad}} \Leftrightarrow \Pi(u) = 0$  and  $\Pi'(u) = 0, \Pi''(u) = 0$  for  $u \in U_{\text{ad}}$ . For a penalty parameter  $k > 0$ , we arrive at the penalized lower-level problem

$$\begin{aligned} \min_{y, u} \quad & \sum_{i=1}^n \alpha_i f_i(y, u) + \frac{\alpha_0}{2} \|u\|_{L^2(\Omega)}^2 + k \Pi(u) \\ \text{s.t.} \quad & -\Delta y = u \quad \text{in } H^{-1}(\Omega) \\ & y \in H_0^1(\Omega) \end{aligned} \tag{P(k, \alpha)}$$

Note that our Assumption 2.1 is satisfied for fixed  $k$  if we would use an additional function  $f_{n+1}(y, u) = k\Pi(u)$ . Therefore our previous results can be applied for  $(P(k, \alpha))$  using  $U_{\text{ad}} = L^2(\Omega)$  and  $\tilde{K} = K \times \{1\} \subset \mathbb{R}^{n+1}$  as feasible sets.

We denote the solution operator  $\alpha \mapsto u$  of this penalized lower-level problem by  $S_k(\alpha)$ . Note that the solution operator  $\alpha \mapsto u$  of the unpenalized lower-level problem with control constraints is still denoted by  $S(\alpha)$ .

**Lemma 4.1.** Assume that  $\alpha^k \rightarrow \tilde{\alpha} \in K$ . Let  $u_k = S_k(\alpha)$  denote the associated optimal control and let  $y_k = G u_k$  denote the optimal state. Then, the sequences  $\{u_k\}$  and  $\{y_k\}$  are uniformly bounded in  $L^\infty(\Omega)$  and  $u_k \rightarrow S(\tilde{\alpha})$  in  $L^2(\Omega)$ .

*Proof.* First, we apply [Theorem 2.7](#) to obtain the boundedness of  $u_k$  in  $L^\infty(\Omega)$ . Note that this bound is uniform in  $k$  since the penalty parameter does not influence the quantities, on which  $C$  from [\(2.8\)](#) depends, cf. [\(2.2\)](#).

W.l.o.g. we may assume  $u_k \rightharpoonup \tilde{u}$  in  $L^2(\Omega)$ . This is justified since we will see that the limit satisfies  $\tilde{u} = S(\tilde{\alpha})$  and, thus, is unique. The weak convergence implies  $y_k := G u_k \rightarrow \tilde{y} = G \tilde{u}$  in  $H_0^1(\Omega)$ . We use again the function  $F$  from [\(3.2\)](#). This yields

$$F(\tilde{\alpha}, \tilde{y}, \tilde{u}) \leq \liminf_{k \rightarrow \infty} \left[ F(\tilde{\alpha}, \tilde{y}, \tilde{u}) + k \Pi(\tilde{u}) \right] \quad (4.1a)$$

$$\leq \liminf_{k \rightarrow \infty} \left[ F(\alpha^k, y_k, u_k) + k \Pi(u_k) \right] \quad (4.1b)$$

$$\leq \limsup_{k \rightarrow \infty} \left[ F(\alpha^k, y_k, u_k) + k \Pi(u_k) \right] \quad (4.1c)$$

$$\leq \limsup_{k \rightarrow \infty} \left[ F(\alpha^k, G S(\alpha^k), S(\alpha^k)) \right] \quad (4.1d)$$

$$= F(\tilde{\alpha}, G S(\tilde{\alpha}), S(\tilde{\alpha})) < \infty. \quad (4.1e)$$

Inequality [\(4.1a\)](#) follows from  $\Pi(\tilde{u}) \geq 0$ . The next inequality [\(4.1b\)](#) follows from the weak lower semicontinuity of  $(\alpha, y, u) \mapsto F(\alpha, y, u)$  and by a distinction of the cases  $\Pi(\tilde{u}) = 0$  or  $\Pi(\tilde{u}) > 0$ . Since  $(y_k, u_k)$  solves  $(P(k, \alpha))$  and  $\Pi(S(\alpha^k)) = 0$ , [\(4.1d\)](#) follows. Finally, [\(4.1e\)](#) follows from the continuity of  $F$  and  $S$ , see [Theorem 2.12](#).

Altogether, the chain of inequalities [\(4.1\)](#) implies  $\Pi(\tilde{u}) = 0$ , i.e.,  $\tilde{u} \in U_{\text{ad}}$ . Since the original lower-level problem  $(P(\alpha))$  has a unique solution,  $\tilde{u} = S(\tilde{\alpha})$  follows. Hence, [\(4.1\)](#) holds with equality. In particular,

$$F(\alpha^k, y_k, u_k) + k \Pi(u_k) \rightarrow F(\tilde{\alpha}, \tilde{y}, \tilde{u}).$$

Using again the weak lower semicontinuity of  $F$ , we find  $k \Pi(u_k) \rightarrow 0$ . Finally,

$$\sum_{i=1}^n \alpha_i^k f_i(y_k, u_k) + \frac{\alpha_0}{2} \|u_k\|_{L^2(\Omega)}^2 \rightarrow \sum_{i=1}^n \tilde{\alpha}_i f_i(\tilde{y}, \tilde{u}) + \frac{\alpha_0}{2} \|\tilde{u}\|_{L^2(\Omega)}^2$$

implies the convergence  $\|u_k\|_{L^2(\Omega)} \rightarrow \|\tilde{u}\|_{L^2(\Omega)}$ . The claim follows.

From now on, we consider a fixed local minimizer  $\bar{\alpha}$  of [\(IOP\)](#). If  $\varepsilon > 0$  is the optimality

radius of the local minimizer  $\bar{\alpha}$  (i.e.  $\psi(\alpha) \geq \psi(\bar{\alpha})$  if  $|\alpha - \bar{\alpha}| < \varepsilon$ ) then we define

$$K_0 := K \cap \overline{B_\varepsilon(\bar{\alpha})} \subset \mathbb{R}^n$$

where  $\overline{B_\varepsilon(\bar{\alpha})} \subset \mathbb{R}^n$  is the closed ball with radius  $\varepsilon$  and centered at  $\bar{\alpha}$ . It follows that  $\bar{\alpha}$  is a global minimizer if we restrict the feasible set to  $K_0$ . We consider the modified upper-level problem

$$\begin{aligned} \min_{\alpha, y, u} \quad & \frac{1}{2} \|y - y_d\|_{L^2(\Omega)} + \frac{1}{2} |\alpha - \bar{\alpha}|_2^2 \\ \text{s.t.} \quad & \alpha \in K_0 \\ & (y, u) \text{ solves } (P(\alpha)). \end{aligned} \tag{4.2}$$

It is clear that  $\bar{\alpha}$  is a unique global minimizer of (4.2).

Now we combine (4.2) and  $(P(k, \alpha))$  into the penalized bilevel problem, which yields

$$\begin{aligned} \min_{\alpha, y, u} \quad & \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{1}{2} |\alpha - \bar{\alpha}|_2^2 \\ \text{s.t.} \quad & \alpha \in K_0 \\ & (y, u) \text{ solves } (P(k, \alpha)). \end{aligned} \tag{Q(k)}$$

This optimization problem will be the basis for deriving optimality conditions for (IOP).

**Lemma 4.2.** For each  $k$  let  $\alpha^k$  be a local minimizer of  $(Q(k))$ . Then we have  $\alpha^k \rightarrow \bar{\alpha}$ .

*Proof.* We denote the first term of the objective function of the upper-level problem with

$$\psi_k(\alpha) := \frac{1}{2} \|GS_k(\alpha) - y_d\|_{L^2(\Omega)}.$$

Because  $\alpha^k$  is bounded, there exists a convergent subsequence. We denote its limit by  $\tilde{\alpha}$  and w.l.o.g. we can say that  $\alpha^k \rightarrow \tilde{\alpha}$ . According to Lemma 4.1 we have  $S_k(\alpha^k) \rightarrow S(\tilde{\alpha})$ . Therefore the convergence  $\psi_k(\alpha^k) \rightarrow \psi(\tilde{\alpha})$  follows.

Another consequence of Lemma 4.1 is that  $S_k(\bar{\alpha}) \rightarrow S(\bar{\alpha})$  and therefore  $\psi_k(\bar{\alpha}) \rightarrow \psi(\bar{\alpha})$ . We have

$$\psi(\tilde{\alpha}) + \frac{1}{2} |\tilde{\alpha} - \bar{\alpha}|^2 = \lim_{k \rightarrow \infty} (\psi_k(\alpha^k) + \frac{1}{2} |\alpha^k - \bar{\alpha}|^2) \leq \lim_{k \rightarrow \infty} \psi_k(\bar{\alpha}) = \psi(\bar{\alpha}) \leq \psi(\tilde{\alpha}).$$

This implies  $\bar{\alpha} = \tilde{\alpha}$ . Because every subsequence of  $\{\alpha_k\}_{k \in \mathbb{N}}$  has a subsequence that converges to  $\bar{\alpha}$ , it follows that  $\alpha^k \rightarrow \bar{\alpha}$ .

From now on we consider a fixed sequence  $(\alpha^k, y_k, u_k)$  of solutions of  $(Q(k))$ . Our strategy is to apply Theorem 3.2 to obtain necessary optimality conditions for each  $k \in \mathbb{N}$  and then to take the limit.

We recall that [Lemma 4.1](#) yields  $u_k = S_k(\alpha^k) \rightarrow S(\bar{\alpha}) =: \bar{u}$  in  $L^2(\Omega)$ . It follows that  $y_k := Gu_k$  converges to  $\bar{y} := G\bar{u}$  in  $H_0^1(\Omega)$ . The solution  $(y_k, u_k)$  of  $(P(k, \alpha^k))$  satisfies the optimality condition

$$-\Delta p_k - D_y F(\alpha^k, y_k, u_k) = 0 \quad (4.3a)$$

$$p_k + D_u F(\alpha^k, y_k, u_k) + k\Pi'(u_k) = 0, \quad (4.3b)$$

which follow from [Theorem 2.4](#). For the limit we obtain the following lemma.

**Lemma 4.3.** The term  $k\Pi'(u_k)$  converges in  $L^2(\Omega)$  and  $p_k$  in  $H_0^1(\Omega)$ . The limits  $\bar{\lambda} := \lim_{k \rightarrow \infty} k\Pi'(u_k)$  and  $\bar{p} := \lim_{k \rightarrow \infty} p_k$  satisfy the conditions

$$-\Delta \bar{p} - D_y F(\bar{\alpha}, \bar{y}, \bar{u}) = 0,$$

$$\bar{p} + D_u F(\bar{\alpha}, \bar{y}, \bar{u}) + \bar{\lambda} = 0,$$

$$0 \leq u_b - \bar{u} \perp \bar{\lambda}^+ \geq 0,$$

$$0 \leq \bar{u} - u_a \perp \bar{\lambda}^- \geq 0.$$

*Proof.* Due to [\(2.2b\)](#) the functions  $(y, u) \mapsto D_y F(\alpha, y, u)$ ,  $(y, u) \mapsto D_u F(\alpha, y, u)$  are pointwise a.e. Lipschitz continuous for every  $\alpha \in K$  with Lipschitz constant  $nC_M$  if  $M$  is a uniform upper bound on  $y, u$ . Therefore we have  $D_{(y,u)} F(\alpha^k, y_k, u_k) \rightarrow D_{(y,u)} F(\bar{\alpha}, G\bar{y}, \bar{u})$  in  $L^2(\Omega)^2$ . As a consequence we first obtain from [\(4.3a\)](#) that  $p_k \rightarrow \bar{p}$  in  $H_0^1(\Omega)$  and then from [\(4.3b\)](#)  $k\Pi'(u_k)$  converges in  $L^2(\Omega)$  and the limits  $\bar{p}, \bar{\lambda}$  are given such that they satisfy the equations in the claim.

In order to show the pointwise complementarity conditions, we consider a subsequence such that both  $u_k$  and  $k\Pi'(u_k)$  converge pointwise. Due to the definition of  $\Pi$ , it can be seen that

$$(u_b - u_k)(k\Pi'(u_k))^+ = k(u_b - u_k)(\pi'(u_k - u_b)) \leq 0 \quad \text{a.e. in } \Omega$$

because  $\pi'(u_k - u_b) = 0$  a.e. on  $\{u_k \leq u_b\}$ . Taking the limit, this implies  $(u_b - \bar{u})\bar{\lambda}^+ \leq 0$  a.e. in  $\Omega$ . Similarly, we can show that  $(\bar{u} - u_a)\bar{\lambda}^- \leq 0$  a.e. in  $\Omega$ . Finally, using  $u_a \leq \bar{u} \leq u_b$  and  $\bar{\lambda}^\pm \geq 0$ , the result follows.

Now we apply [Theorem 3.2](#) to  $(Q(k))$  (while taking the additional term in the objective function into account). Thus there exist  $(\mu_k, \nu_k, \rho_k)$  such that

$$\begin{pmatrix} D_{yy} F(\alpha^k, y_k, u_k) & D_{yu} F(\alpha^k, y_k, u_k) & -\Delta \\ D_{uy} F(\alpha^k, y_k, u_k) & D_{uu} F(\alpha^k, y_k, u_k) + k\Pi''(u_k) & -I \\ -\Delta & -I & 0 \end{pmatrix} \begin{pmatrix} \mu_k \\ \nu_k \\ \rho_k \end{pmatrix} = \begin{pmatrix} y_k - y_d \\ 0 \\ 0 \end{pmatrix} \quad (4.4a)$$

$$\bar{\alpha} - \alpha^k + D_{\alpha y} F(\alpha^k, y_k, u_k)\mu_k + D_{\alpha u} F(\alpha^k, y_k, u_k)\nu_k \in \mathcal{N}_{K_0}(\alpha^k). \quad (4.4b)$$

Since the parameter corresponding to  $k\Pi(\cdot)$  is fixed, its stationarity condition, which arises from [\(3.7g\)](#), contains no information and has been removed.



The goal for the next lemmas is to show that a subsequence of  $(\mu_k, \nu_k, \rho_k)$  converges (weakly) in suitable spaces and that the limit point satisfies the system of C-stationarity (3.7), (3.8). First, we address the convergence of these dual variables and check that the limit point satisfies (3.7f), (3.7g).

**Lemma 4.4.** There exist  $\bar{\mu}, \bar{\rho} \in H_0^1(\Omega), \bar{\nu}, \bar{z} \in L^2(\Omega)$  such that

$$\mu_k \rightarrow \bar{\mu} \quad \text{in } H_0^1(\Omega), \quad (4.5a)$$

$$\nu_k \rightharpoonup \bar{\nu} \quad \text{in } L^2(\Omega), \quad (4.5b)$$

$$\rho_k \rightarrow \bar{\rho} \quad \text{in } H_0^1(\Omega), \quad (4.5c)$$

$$-k\Pi''(u_k)\nu_k \rightharpoonup \bar{z} \quad \text{in } L^2(\Omega) \quad (4.5d)$$

along a subsequence. The limits  $\bar{\mu}, \bar{\nu}, \bar{\rho}, \bar{z}$  satisfy

$$D_{yy}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\mu} + D_{yu}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\nu} - \Delta\bar{\rho} = \bar{y} - y_d, \quad (4.5e)$$

$$D_{uy}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\mu} + D_{uu}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\nu} - \bar{\rho} = \bar{z}, \quad (4.5f)$$

$$-\Delta\bar{\mu} - \bar{\nu} = 0, \quad (4.5g)$$

$$D_{\alpha y}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\mu} + D_{\alpha u}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\nu} \in \mathcal{N}_K(\bar{\alpha}). \quad (4.5h)$$

*Proof.* Using some substitutions, (4.4a) can be rewritten as

$$(\tilde{A}(\alpha^k, y_k, u_k) + k\Pi''(u_k))\nu_k = G(y_k - y_d) \quad (4.6)$$

where  $\tilde{A}(\alpha^k, y_k, u_k) : L^2(\Omega) \rightarrow L^2(\Omega)$  is a suitable bounded linear operator with coercivity constant  $\alpha_0 > 0$ . In particular, this constant does not depend on  $k$ . Thus, we can apply the Lax-Milgram theorem to obtain a solution  $\nu_k \in L^2(\Omega)$  of (4.6) and this solution satisfies  $\|\nu_k\|_{L^2(\Omega)} \leq \frac{1}{\alpha_0} \|G(y_k - y_d)\|_{L^2(\Omega)}$ . Hence  $\nu_k$  converges weakly along a subsequence. We denote the weak limit by  $\bar{\nu}$  and w.l.o.g. we can say that  $\nu_k \rightharpoonup \bar{\nu}$ . It follows that  $\mu_k$  converges strongly to  $\bar{\mu} := G\bar{\nu}$  in  $H_0^1(\Omega)$ .

The convergences  $D_{\alpha y}F(\alpha^k, y_k, u_k) \rightarrow D_{\alpha y}F(\bar{\alpha}, \bar{y}, \bar{u})$ ,  $D_{\alpha u}F(\alpha^k, y_k, u_k) \rightarrow D_{\alpha u}F(\bar{\alpha}, \bar{y}, \bar{u})$  in  $L^2(\Omega)$  can be shown similar to the convergence of  $D_{(y,u)}F(\alpha, y_k, u_k)$  in the proof of Lemma 4.3. By passing to the limit with (4.4b), (4.5h) follows. Indeed, since  $\bar{\alpha}$  is in the interior of  $\bar{B}_\varepsilon(\bar{\alpha})$  we have  $\mathcal{N}_{K_0}(\bar{\alpha}) = \mathcal{N}_K(\bar{\alpha})$ .

Now, consider a subsequence of  $u_k$  (without renaming) that converges pointwise almost everywhere. By continuity of  $D_{yu}\phi_i(x, \cdot, \cdot)$  it follows that  $D_{yu}\phi_i(\cdot, y_k, u_k) \rightarrow D_{yu}\phi_i(\cdot, \bar{y}, \bar{u})$  a.e. and by Lebesgue's dominated convergence we have

$$D_{yu}\phi_i(\cdot, y_k, u_k)w \rightarrow D_{yu}\phi_i(\cdot, \bar{y}, \bar{u})w \quad \text{in } L^2(\Omega) \quad (4.7)$$

for all  $w \in L^2(\Omega), i \in \{1, \dots, n\}$ . It follows that

$$\begin{aligned} (w, D_{yu}F(\alpha^k, y_k, u_k)\nu_k)_{L^2(\Omega)} &= (\nu_k, D_{yu}F(\alpha^k, y_k, u_k)w)_{L^2(\Omega)} \\ &\rightarrow (\bar{\nu}, D_{yu}F(\bar{\alpha}, \bar{y}, \bar{u})w)_{L^2(\Omega)} \\ &= (w, D_{yu}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\nu})_{L^2(\Omega)}. \end{aligned}$$

One can proceed in the same way for the term  $D_{uu}F(\alpha^k, y_k, u_k)\nu_k$  which results in the weak  $L^2(\Omega)$ -convergences

$$\begin{aligned} D_{yu}F(\alpha^k, y_k, u_k)\nu_k &\rightharpoonup D_{yu}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\nu}, \\ D_{uu}F(\alpha^k, y_k, u_k)\nu_k &\rightharpoonup D_{uu}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\nu}. \end{aligned} \quad (4.8)$$

Using the triangle inequality and the uniform boundedness of  $D_{yy}F(\alpha^k, y_k, u_k)$  and  $D_{uy}F(\alpha^k, y_k, u_k)$ , (4.7) even implies the strong  $L^2(\Omega)$  convergence

$$D_{uy}F(\alpha^k, y_k, u_k)\mu_k \rightarrow D_{uy}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\mu}. \quad (4.9)$$

Likewise, the  $L^2(\Omega)$  convergence

$$D_{yy}F(\alpha^k, y_k, u_k)\mu_k \rightarrow D_{yy}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\mu} \quad (4.10)$$

can be shown. From (4.4a) it follows that  $-\Delta\rho_k$  converges weakly in  $L^2(\Omega)$  and thus  $\rho_k$  converges strongly in  $H_0^1(\Omega)$  and the limit  $\bar{\rho}$  satisfies (4.5e). Now using (4.8),(4.9),(4.10) and other previous results, the weak convergence (4.5d) and the satisfaction of (4.5f) follow.

The next goal is to show the complementarities necessary for the system of weak stationarity. Our plan is to use Egorov's theorem to a pointwise convergent subsequence of  $u_k$  to obtain uniform convergence on the set  $\Omega_\varepsilon$  where  $\varepsilon > 0$  is arbitrary and  $\text{meas}(\Omega \setminus \Omega_\varepsilon) < \varepsilon$ . We will use this for the next two lemmas in order to show the conditions necessary for weak stationarity.

**Lemma 4.5.** The condition

$$\bar{\nu} = 0 \text{ a.e. on } \{\bar{\lambda} \neq 0\}$$

holds.

*Proof.* First, we consider a subsequence such that  $\nu_k \rightharpoonup \bar{\nu}$  holds. From Lemmas 4.1 and 4.3 we know  $u_k \rightarrow \bar{u}$  and  $k\Pi'(u_k) \rightarrow \bar{\lambda}$  in  $L^2(\Omega)$ . Let  $\varepsilon > 0$  be given. Then, by Egorov's theorem there exists  $\Omega_\varepsilon \subset \Omega$  with  $\text{meas}(\Omega \setminus \Omega_\varepsilon) < \varepsilon$  such that  $u_k \rightarrow \bar{u}$  and  $k\Pi'(u_k) \rightarrow \bar{\lambda}$  in  $L^\infty(\Omega_\varepsilon)$  along a further subsequence of  $\{u_k\}_{k \in \mathbb{N}}$ .

Consider the set  $\Omega_{\varepsilon,+} := \Omega \cap \{\bar{\lambda}^+ > \varepsilon\}$ . We have  $P'(u_k - u_b) = \Pi'(u_k)^+ > 0$  on  $\Omega_{\varepsilon,+}$  for large  $k$ . Using the definition of our penalty functional  $P$ , we have  $P'(u_k - u_b) \leq 6(u_k - u_b)^2$ . Since  $kP'(u_k - u_b) \rightarrow \bar{\lambda}^+$  uniformly on  $\Omega_\varepsilon$ , this implies

$$6(u_k - u_b)^2 \geq P'(u_k - u_b) \geq \frac{1}{2k}\bar{\lambda}^+ \geq \frac{\varepsilon}{2k} \quad \text{on } \Omega_{\varepsilon,+}$$

for large  $k \in \mathbb{N}$ . Again, using the definition of  $P(u_k - u_b)$  it follows that  $kP''(u_k - u_b) \geq k(u_k - u_b)$  for large  $k$ . Then we have

$$k\Pi''(u_k) \geq kP''(u_k - u_b) \geq k(u_k - u_b) \geq k\sqrt{\frac{\varepsilon}{12k}} \quad \text{on } \Omega_{\varepsilon,+}.$$

In particular,  $k\Pi''(u_k) \rightarrow \infty$  pointwise on  $\Omega_{\varepsilon,+}$ . Because  $k\Pi''(u_k)\nu_k$  is bounded in  $L^2(\Omega)$  according to [Lemma 4.4](#), it follows that  $\nu_k \rightarrow 0$  pointwise on  $\Omega_{\varepsilon,+}$ . Using  $\nu_k \rightharpoonup \bar{\nu}$ , it follows that  $\bar{\nu} = 0$  a.e. on  $\Omega_{\varepsilon,+}$ .

Similarly, it can be shown that  $\bar{\nu} = 0$  a.e. on  $\Omega_{\varepsilon,-} := \Omega_\varepsilon \cap \{\bar{\lambda}^- > \varepsilon\}$ . Finally, because

$$\text{meas}(\{\bar{\lambda} \neq 0\} \setminus \bigcup_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} (\Omega_{\varepsilon,+} \cup \Omega_{\varepsilon,-})) = 0,$$

the claim follows.

**Lemma 4.6.** The limits  $\bar{z}, \bar{u}$  satisfy the condition

$$\bar{z} = 0 \text{ a.e. on } \{u_a < \bar{u} < u_b\}.$$

*Proof.* First, we consider a subsequence such that  $-k\Pi''(u_k)\nu_k \rightharpoonup \bar{z}$  holds. From [Lemma 4.1](#) we know  $u_k \rightarrow \bar{u}$ . Let  $\varepsilon > 0$  be given. Then, by Egorov's theorem there exists  $\Omega_\varepsilon \subset \Omega$  with  $\text{meas}(\Omega \setminus \Omega_\varepsilon) < \varepsilon$  such that  $u_k \rightarrow \bar{u}$  uniformly on  $\Omega_\varepsilon$  along a subsequence of  $\{u_k\}_{k \in \mathbb{N}}$ .

Then we define  $\Omega_{\varepsilon,0} := \Omega_\varepsilon \cap \{u_a + \varepsilon < \bar{u} < u_b - \varepsilon\}$ . Because  $u_k$  converges uniformly, we have  $u_a < u_k < u_b$  on  $\Omega_{\varepsilon,0}$  for large  $k$ . Therefore  $k\Pi''(u_k)\nu_k = 0$  a.e. on  $\Omega_{\varepsilon,0}$ . Due to the weak convergence  $-k\Pi''(u_k)\nu_k \rightharpoonup \bar{z}$  and the pointwise a.e. convergence on  $\Omega_{\varepsilon,0}$ , we know that  $\bar{z} = 0$  a.e. on  $\Omega_{\varepsilon,0}$ . Finally, because

$$\{u_a < \bar{u} < u_b\} = \bigcup_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \Omega_{\varepsilon,0}$$

up to a set of measure zero, the claim follows.

The previous results show that any local minimizer of [\(IOP\)](#) are weakly stationary. It remains to upgrade this to C-stationarity.

**Lemma 4.7.** We have  $\bar{\nu}\bar{z} \leq 0$  a.e. in  $\Omega$ .

*Proof.* In this proof, we consider a subsequence such that the convergences in [Lemma 4.4](#), [\(4.7\)](#), [\(4.8\)](#), and [\(4.9\)](#) hold. Let  $E \subset \Omega$  be an arbitrary measurable set. If we test [\(4.4a\)](#) with  $(0, \nu_k \chi_E, 0)$  and consider that  $\chi_E \Pi''(u_k) \geq 0$  we obtain

$$(\nu_k \chi_E, D_{uy}F(\alpha^k, y_k, u_k)\mu_k + D_{uu}F(\alpha^k, y_k, u_k)\nu_k - \rho_k)_{L^2(\Omega)} \leq 0. \quad (4.11)$$

Because  $\nu_k \chi_E \rightharpoonup \bar{\nu} \chi_E$  in  $L^2(\Omega)$ , [\(4.9\)](#) yields

$$(\nu_k \chi_E, D_{uy}F(\alpha^k, y_k, u_k)\mu_k)_{L^2(\Omega)} \rightarrow (\bar{\nu} \chi_E, D_{uy}F(\bar{\alpha}, \bar{y}, \bar{u})\bar{\mu})_{L^2(\Omega)}.$$

We define the abbreviation  $w_k := \chi_E D_{uu}F(\alpha^k, y_k, u_k) \in L^\infty(\Omega)$ . Because  $w_k$  is nonnegative, we have

$$(\bar{\nu}, w_k \bar{\nu})_{L^2(\Omega)} \leq 2(\bar{\nu} - \nu_k, w_k \bar{\nu})_{L^2(\Omega)} + (\nu_k, w_k \nu_k)_{L^2(\Omega)}.$$

On the other hand, by using arguments similar to those leading to (4.7), we have

$$(\bar{\nu}, w_k \bar{\nu})_{L^2(\Omega)} \rightarrow (\bar{\nu}, \chi_E D_{uu}F(\bar{\alpha}, \bar{y}, \bar{u}) \bar{\nu})_{L^2(\Omega)}.$$

Combining this with the previous inequality, (4.8) and the weak convergence  $\nu_k \rightharpoonup \bar{\nu}$  yields

$$(\bar{\nu}, \chi_E D_{uu}F(\bar{\alpha}, \bar{y}, \bar{u}) \bar{\nu})_{L^2(\Omega)} \leq \liminf_{k \rightarrow \infty} (\nu_k, \chi_E D_{uu}F(\alpha^k, y_k, u_k) \nu_k)_{L^2(\Omega)}.$$

Together with (4.5c) and (4.11), this gives

$$\begin{aligned} (\bar{\nu}, \chi_E D_{uu}F(\bar{\alpha}, \bar{y}, \bar{u}) \bar{\nu})_{L^2(\Omega)} &\leq \liminf_{k \rightarrow \infty} (\nu_k \chi_E, \rho_k - D_{uy}F(\alpha^k, y_k, u_k) \mu_k)_{L^2(\Omega)} \\ &= (\bar{\nu} \chi_E, \bar{\rho} - D_{uy}F(\bar{\alpha}, \bar{y}, \bar{u}) \bar{\mu})_{L^2(\Omega)} \end{aligned}$$

Now, using the equation (4.5f) for  $\bar{z}$  yields

$$0 \geq (\bar{\nu} \chi_E, D_{uu}F(\bar{\alpha}, \bar{y}, \bar{u}) \bar{\nu} - \bar{\rho} + D_{uy}F(\bar{\alpha}, \bar{y}, \bar{u}) \bar{\mu})_{L^2(\Omega)} = (\bar{\nu} \chi_E, \bar{z})_{L^2(\Omega)}.$$

Because  $E \subset \Omega$  was arbitrary, the claim follows.

Now we have shown that for each local minimizer  $(\bar{\alpha}, \bar{y}, \bar{u})$  of (IOP), the sequence  $(\mu_k, \nu_k, \rho_k, -k\Pi''(u_k)\nu_k)$  of multipliers has a (weakly) convergent subsequence and the accumulation point satisfies the system of C-stationarity from Definition 3.3. Because this is our main result and for the sake of completeness, we state this as a theorem.

**Theorem 4.8.** Every local minimizer  $(\bar{\alpha}, \bar{y}, \bar{u})$  of (IOP) is a C-stationary point.

Finally, we recall that Example 3.4 shows that local minimizers might not be strongly stationary. However, it is not clear whether all local minimizers are M-stationary.

## ACKNOWLEDGMENTS

This work is supported by DFG grant WA 3636/4-1 within the **Priority Program SPP 1962** (Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization).

## REFERENCES

- Albrecht, Sebastian; Leibold, Marion; Ulbrich, Michael (2012). “A bilevel optimization approach to obtain optimal cost functions for human arm movements”. *Numerical Algebra, Control and Optimization* 2.1, pp. 105–127. ISSN: 2155-3289. DOI: [10.3934/naco.2012.2.105](https://doi.org/10.3934/naco.2012.2.105).
- Albrecht, Sebastian; Ulbrich, Michael (2017). “Mathematical programs with complementarity constraints in the context of inverse optimal control for locomotion”. *Optimization Methods & Software* 32.4, pp. 670–698. ISSN: 1055-6788. DOI: [10.1080/10556788.2016.1225212](https://doi.org/10.1080/10556788.2016.1225212).
- Christof, C.; Wachsmuth, G. (2017). *Differential Sensitivity Analysis of Variational Inequalities with Locally Lipschitz Continuous Solution Operators*. Preprint. arXiv: [1711.02720](https://arxiv.org/abs/1711.02720).
- Dempe, Stephan (2002). *Foundations of bilevel programming*. Dordrecht: Kluwer Academic Publishers.
- Hatz, Kathrin; Schlöder, Johannes P.; Bock, Hans Georg (2012). “Estimating parameters in optimal control problems”. *SIAM Journal on Scientific Computing* 34.3, A1707–A1728. ISSN: 1064-8275. DOI: [10.1137/110823390](https://doi.org/10.1137/110823390).
- Malanowski, K. (2002). “Sensitivity Analysis for Parametric Optimal Control of Semilinear Parabolic Equations”. *Journal of Convex Analysis* 9.2, pp. 543–561.
- 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.; Wachsmuth, G. (2016). “Weak and strong stationarity in generalized bilevel programming and bilevel optimal control”. *Optimization* 65.5, pp. 907–935. DOI: [10.1080/02331934.2015.1122007](https://doi.org/10.1080/02331934.2015.1122007).
- Mehlitz, P.; Wachsmuth, G. (2017). “The weak sequential closure of decomposable sets in Lebesgue spaces and its application to variational geometry”. *Set-Valued and Variational Analysis*. To appear. DOI: [10.1007/s11228-017-0464-1](https://doi.org/10.1007/s11228-017-0464-1).
- Mombaur, Katja; Truong, Anh; Laumond, Jean-Paul (2010). “From human to humanoid locomotion—an inverse optimal control approach”. *Autonomous Robots* 28.3, pp. 369–383. ISSN: 1573-7527. DOI: [10.1007/s10514-009-9170-7](https://doi.org/10.1007/s10514-009-9170-7).
- Stampacchia, Guido (1960–1961). “Équations elliptiques à données discontinues”. *fre. Séminaire Schwartz* 5, pp. 1–16.
- Tröltzsch, F. (2000). “Lipschitz Stability of Solutions of Linear-Quadratic Parabolic Control Problems with Respect to Perturbations”. *Dynamics of Continuous, Discrete and Impulsive Systems Series A Mathematical Analysis* 7.2, pp. 289–306.
- Tröltzsch, F. (2010). *Optimal Control of Partial Differential Equations*. Vol. 112. Graduate Studies in Mathematics. Providence: American Mathematical Society. DOI: [10.1090/gsm/112](https://doi.org/10.1090/gsm/112).
- Unger, A. (1997). “Hinreichende Optimalitätsbedingungen zweiter Ordnung und Konvergenz des SQP-Verfahrens für semilineare elliptische Randsteuerprobleme”. PhD thesis. Technische Universität Chemnitz, Germany.

Wachsmuth, G. (2015). “Mathematical Programs with Complementarity Constraints in Banach Spaces”. *Journal of Optimization Theory and Applications* 166.2, pp. 480–507. ISSN: 0022-3239. DOI: [10.1007/s10957-014-0695-3](https://doi.org/10.1007/s10957-014-0695-3).