Several Approaches for the Derivation of Stationarity Conditions for Elliptic MPECs with Upper-Level Control Constraints

M. Hintermüller*, B. S. Mordukhovich† and T. Surowiec‡

Abstract

The derivation of multiplier-based optimality conditions for elliptic mathematical programs with equilibrium constraints (MPEC) is essential for the characterization of solutions and development of numerical methods. Though much can be said for broad classes of elliptic MPECs in both polyhedric and non-polyhedric settings, the calculation becomes significantly more complicated when additional constraints are imposed on the control. In this paper we develop three derivation methods for constrained MPEC problems: via concepts from variational analysis, via penalization of the control constraints, and via penalization of the lower-level problem with the subsequent regularization of the resulting nonsmoothness. The developed methods and obtained results are then compared and contrasted.

1 Introduction

In a previous work [9], the first and third authors applied and further developed certain techniques from convex and nonsmooth analysis to derive first-order optimality conditions for a class of bilevel optimization problems known as mathematical programs with equilibrium constraints, or simply MPECs, in function spaces. Such models are known to arise in many application areas such as mathematical elasticity, finance, economics, etc. Nevertheless, the techniques were only applicable to a certain class of MPECs in which the so-called upper-level variables or controls are not subject to any constraints. In fact, the literature on the derivation of explicit (i.e., multiplier-based) necessary optimality conditions for MPECs in function spaces with upper-level constraints is rather scarce; though there are some results available in [17, 18] and [8].

We thus aim to present several techniques for the derivation of multiplier-based first-order optimality conditions. Throughout the text we compare and contrast their applicability to the development of numerical methods based on the amount of information they require from the user as well as their theoretical strength, in terms of their selectivity.

In the literature on optimization problems governed by partial differential equations, regularization/penalization techniques employed for the derivation of necessary optimality conditions are relatively widespread. Conversely, techniques from set-valued and variational analysis provide powerful tools for the direct derivation of multiplier-based optimality conditions. Currently it is unclear as to how these techniques compare from both the analytical perspective, e.g., the selectivity of the derived conditions and the generality of their applicability, as well as in terms of numerics, e.g., the development of mesh-independent solvers.

*Department of Mathematics, Humboldt University of Berlin, Germany and Department of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, Austria (hint@math.hu-berlin.de).

†Department of Mathematics, Wayne State University, USA (boris@math.wayne.edu).

‡Department of Mathematics, Humboldt University of Berlin, Germany (surowiec@math.hu-berlin.de).

In this paper we mainly concern ourselves with the following class of MPECs:

$$\begin{aligned} &\min \tfrac{1}{2}||y-y_d||^2_{L^2(\Omega)} + \tfrac{\alpha}{2}||u||^2_{L^2(\Omega)} \text{ over } (u,y) \in L^2(\Omega)\times H^1_0(\Omega) \\ &\text{s.t.} \quad a \le u \le b \text{ almost everywhere (a.e.) in } \Omega, \\ &\qquad Ay + N_M(y) \ni Bu. \end{aligned} \tag{1}$$

Here $\alpha > 0$, $\Omega \subset \mathbb{R}^n$ with $1 \le n \le 3$ is open and bounded, and there exists $\beta \in \mathbb{R}$ such that $b - a \ge \beta > 0$ *a.e.* Ω, where $a, b \in L^\infty(\Omega)$. The notation $L^2(\Omega)$ is used to represent the standard Lebesgue space of square integrable functions, while $H^1_0(\Omega)$ stands for the Sobolev space of $L^2(\Omega)$-functions y with $y|_{\partial\Omega} = 0$ whose distributional derivatives ∇y belong to $L^2(\Omega)$. We use the symbol $H^{-1}(\Omega)$ to represent the dual of $H^1_0(\Omega)$ throughout the entirety of this paper. The bounded linear operator $A \in \mathcal{L}(H^1_0(\Omega), H^{-1}(\Omega))$ is assumed to be coercive, i.e., we assume that there exists a constant $\xi > 0$ such that

$$\langle Ay, y\rangle_{H^{-1},H^1_0} \ge \xi \|y\|^2_{H^1_0} \quad \text{for all} \quad y \in H^1_0(\Omega)$$

whereas, unless otherwise stated, $B \in \mathcal{L}(L^2(\Omega), H^{-1}(\Omega))$. Finally, we define the closed and convex subset $M \subset H^1_0(\Omega)$ by

$$M := \left\{ y \in H^1_0(\Omega) \mid y \ge \psi \ a.e.\, \Omega \right\}, \tag{2}$$

where $\psi \in H^1(\Omega)$ with $\psi|_{\partial\Omega} \le 0$. The operator $N_M(y)$ for $y \in M$ signifies the classical normal cone of convex analysis defined by

$$N_M(y) := \left\{ y^* \in H^{-1}(\Omega) \,\middle|\, \langle y^*, y - y'\rangle_{H^{-1},H^1_0} \le 0, \ \forall y' \in M \right\}.$$

Accordingly, we could rewrite the generalized equation in (1) as the variational inequality

$$\langle Ay - u, y - y'\rangle_{H^{-1},H^1_0} \ge 0, \ \forall y' \in M.$$

The remaining notational assumptions are fairly standard, however, for completeness we provide them here for quick reference. We use $\langle\cdot,\cdot\rangle_{X^*,X}$ to represent the duality pairing between a topological vector space X and its dual X^* and $(\cdot,\cdot)_X$ for the inner product on X when X is a Hilbert space. The arrows $\to_X$ and $\rightharpoonup_X$ are used to represent, respectively, strong and weak convergence of sequences in topology on X. All the subscripts are omitted when it is clear in context.

Recall that the contingent cone (also known as the Bouligand-Severi tangent cone) to a closed set $C \subset X$ of a Banach space X at a point $x \in C$ is defined by

$$T_C(x) := \left\{ h \in X \,\middle|\, \exists t_k \to 0^+, \exists h_k \to_X h : x + t_k h_k \in C, \forall k \right\}.$$

In the event that the set C is convex and the space X is reflexive, the aforementioned normal cone of convex analysis can be defined as the polar cone to $T_C(x)$, i.e.,

$$N_C(x) := [T_C(x)]^-_X := \{ x^* \in X^* \mid \langle x^*, h\rangle_{X^*,X} \le 0, \ \forall h \in T_C(x) \}.$$

As they play a central role in our paper, we define in what follows various stationarity concepts for MPECs in the current context that are studied in the subsequent sections. We mainly base our notation and definitions on [7] and [8]. In keeping with the terminology of the finite-dimensional literature, we occasionally refer to the variational inequality in (1) as the *lower-level problem* and the remaining constraints, i.e., those on the control, as being in the *upper-level* part of the MPEC.

Definition 1.1 (**C- and S-Stationarity**). *A point* $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ *feasible to the MPEC* (1) *is called a* C-STATIONARY POINT *of the MPEC if there exist multipliers* $\bar{s} \in L^2(\Omega)$, $\bar{v} \in H^{-1}(\Omega)$, $\bar{p} \in H_0^1(\Omega)$, *and* $\bar{r} \in H^{-1}(\Omega)$ *for which*

$$0 = \alpha\bar{u} + B^*\bar{p} + \bar{s}, \tag{3}$$
$$0 = \bar{y} - y_d - A^*\bar{p} + \bar{r}, \tag{4}$$
$$0 = A\bar{y} - B\bar{u} + \bar{v}, \tag{5}$$

where the multipliers satisfy the following conditions:

$$0 \le \bar{s},\ a.e.\, \mathcal{A}_a(\bar{u}) \quad \bar{s} = 0,\ a.e.\, \mathcal{J}(\bar{u}), \quad \bar{s} \ge 0\ a.e.\, \mathcal{A}_b(\bar{u}), \tag{6}$$
$$0 \ge \langle \bar{v}, \varphi \rangle_{H^{-1}, H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi \ge 0\ a.e.\, \mathcal{A}(\bar{y}), \tag{7}$$
$$0 = \langle \bar{v}, \varphi \rangle_{H^{-1}, H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi = 0\ a.e.\, \mathcal{A}(\bar{y}), \tag{8}$$
$$0 = \langle B\bar{u} - A\bar{y}, \bar{p} \rangle_{H^{-1}, H_0^1}, \tag{9}$$
$$0 = \langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi = 0\ a.e.\, \mathcal{A}(\bar{y}), \tag{10}$$
$$0 \ge \langle \bar{r}, \bar{p} \rangle_{H^{-1}, H_0^1}. \tag{11}$$

Here, we use the notation

$$\mathcal{A}(\bar{y}) := \{ x \in \Omega \mid y(x) = \psi(x) \} \ \text{ and } \ \mathcal{I}(\bar{y}) := \Omega \setminus \mathcal{A}(\bar{y})$$

to represent the active and inactive sets for the lower-level problem, respectively, and

$$\begin{aligned} \mathcal{A}_a(\bar{u}) &:= \{ x \in \Omega \,|\, u(x) = a(x) \}, \\ \mathcal{A}_b(\bar{u}) &:= \{ x \in \Omega \,|\, u(x) = b(x) \}, \\ \mathcal{J}(\bar{u}) &:= \Omega \setminus (\mathcal{A}_a(\bar{u}) \cup \mathcal{A}_b(\bar{u})) \end{aligned}$$

for the lower active, upper active, and inactive sets for the upper-level constraints, respectively. If in addition to the above conditions we have $\bar{v} \in L^2(\Omega)$ *and*

$$0 \le \bar{p}\ a.e.\, \mathcal{B}, \tag{12}$$
$$0 \le \langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi \ge 0,\ a.e.\, \mathcal{B} \text{ and } \varphi = 0,\ a.e.\, \mathcal{A}(\bar{y}) \setminus \mathcal{B}, \tag{13}$$

then $(\bar{u}, \bar{y})$ *is said to be a* S(TRONG)-STATIONARY POINT, *where the notation*

$$\mathcal{B} := \{ x \in \mathcal{A}(\bar{y}) \mid \bar{v}(x) = 0 \},$$

is used to denote the so-called bi-active set.

We note that (13) could also be defined when $v \in H^{-1}(\Omega)$. In this case, one has

$$0 \le \langle \bar{r}, \varphi \rangle,\ \forall \varphi \in H_0^1(\Omega) : \langle \bar{u} - A\bar{y}, \varphi \rangle = 0 \text{ and } \varphi \ge 0,\ a.e.\, \mathcal{A}(\bar{y}).$$

An additional stationarity concept, introduced in [7] and [8] and unique to function space settings, also appears later in the paper. For convenience we define it here.

Definition 1.2 (ε-**Almost C-Stationarity**). *A point* $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ *feasible to MPEC* (1) *is called an* ε-ALMOST C-STATIONARY POINT *if there exist multipliers* $\bar{s} \in L^2(\Omega)$, $\bar{v} \in H^{-1}(\Omega)$, $\bar{p} \in H_0^1(\Omega)$, *and* $\bar{r} \in H^{-1}(\Omega)$ *for which the relationships* (3)–(5), (6)–(9), *and* (11) *are satisfied and, instead of* (10), *the following condition holds: for every* $\varepsilon > 0$ *there exists a subset* $E_\varepsilon \subset \mathcal{I}(\bar{y})$ *with* $meas(\mathcal{I}(\bar{y}) \setminus E_\varepsilon) \le \varepsilon$ *such that*

$$0 = \langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi = 0,\ \Omega \setminus E_\varepsilon.$$

The terms C-stationarity and S-stationarity are originally attributed to Scheel and Scholtes [19], where the "C" reflects the fact that the notions from Clarke's nonsmooth calculus were used in the derivation process. In the sense that only the product of the multipliers $\bar{r}$ and $\bar{p}$ has a sign, C-stationarity conditions are not "true" KKT conditions for the MPEC (1). Nevertheless, it has been argued in a finite-dimensional context in the recent papers by Jongen, Rückmann and Shikhman (see, e.g., [11] and the references therein) that C-stationarity is the fundamental stationarity concept needed for the global study of critical points in mathematical programs with complementarity constraints (MPCCs), an important subclass of MPECs.

There are, however, other concepts in the finite-dimensional literature, namely, *B(ouligand)-stationarity*, a primal optimality concept similar to the result found in Theorem 2.1 below, as well as *M(ordukhovich)-stationarity* and *W(eak)-stationarity*. In a function space context, Outrata, Jarušek and Starà in [17] and [18] successfully applied elements of the limiting variational calculus by Mordukhovich to problems similar to ours. Unfortunately, these results are only applicable in the case of control constraints when $\Omega \subset \mathbb{R}$ and the controls u belong to $H^{-1}(\Omega)$.

Following Outrata et al., a point $(\bar{u}, \bar{y})$ is *M-stationary* if conditions (3)–(11) hold and, instead of (12) and (13), we have

$$\langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1} = 0, \forall \varphi \in H_0^1(\Omega) : \varphi < 0 \ a.e. \, \mathcal{B}, \tag{14}$$

$$\langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1} \leq 0, \forall \varphi \in H_0^1(\Omega) : \varphi > 0 \ a.e. \, \mathcal{B}. \tag{15}$$

Since (14) and (15) imply further restrictions on $\bar{r}$, we see that M-stationarity is indeed a more stringent concept than C-stationarity albeit weaker than S-stationarity. The "M" in M-stationarity refers to the fact that multiplier conditions arose in the context of finite-dimensional MPECs as part of optimality conditions through the usage of Mordukhovich's generalized differential constructions and calculus. Due to the low regularity of $\bar{r}$, we see in the subsequent sections that deriving conditions (14) and (15) in our context seems to be problematic in general.

The rest of the paper is structured as follows. In Section 2 we recall some well-known results concerning the existence of solutions to the MPECs under consideration and regularity properties of solution maps to variational inequalities. Following this, we derive primal first-order optimality conditions similar to the B-stationarity conditions mentioned above. In Section 3 we define certain notions from the limiting variational calculus and then apply these concepts to our class of MPECs. We demonstrate that the standard data assumptions made above suffice for the fulfillment of the qualification conditions needed for the application of the limiting variational constructions. Afterwards we characterize the so-called Mordukhovich limiting coderivative (see Section 3 for definition) of the solution mapping to the variational inequality. These results lead us to the derivation of new limiting stationarity conditions surprisingly weaker than C-stationarity. Section 4 is devoted to a hybrid derivation method utilizing the results from [9] for MPECs without upper-level constraints by penalizing the control constraints with a smooth penalty function. In Section 5 we recall a penalization-regularization method extended to elliptic MPECs by Hintermüller and Kopacka in [8] and establish its important consequences. The concluding Section 6 compares both the quality of the results as well as the similarities of the methods and their usefulness for the development of numerical techniques.

2 Preliminaries and B-Stationarity

Throughout the paper we denote by S the solution mapping from $H^{-1}(\Omega)$ into $H_0^1(\Omega)$ defined by

$$S(w) := \left\{ y \in H_0^1(\Omega) \mid Ay + N_M(y) \ni w \right\} \tag{16}$$

and referred to as the *solution mapping* associated with the lower-level problem in our original MPEC (1). This mapping can be easily shown to be single-valued and Lipschitz continuous by utilizing the coercivity of A and the variational form of the generalized equation in (16); see, e.g., [12] or [3] as well as [9]. Moreover, it is well-known that S is in fact (Hadamard) directionally differentiable at every $w \in \operatorname{dom} S$, i.e., the limits

$$S'(w,\cdot) := \lim_{\substack{t\to 0^+ \\ h' \to_{H^{-1}} h}} \frac{S(w+th') - S(w)}{t}$$

exist for all $h \in H^{-1}(\Omega)$. The graph of S' is directly characterized by

$$\operatorname{gph} S'(w,\cdot) = \left\{ (h,d) \in H^{-1}(\Omega) \times H_0^1(\Omega) \,\middle|\, Ad + N_{\mathcal{K}(y,v)}(d) \ni h \right\}. \tag{17}$$

Here, we let $(w,y) \in \operatorname{gph} S$ and use $v \in N_M(y)$ such that $w - Ay = v$ define

$$\mathcal{K}(y,v) := T_M(y) \cap \{v\}^\perp, \tag{18}$$

i.e., the classical *critical cone* from optimization theory. This differentiability result is essentially due to Mignot [13], but it was rederived for a broader class of problems in [9]. Furthermore, since the operator B is linear and bounded from $L^2(\Omega)$ into $H^{-1}(\Omega)$, we know that

$$d = (S \circ B)'(u;h) = S'(Bu;Bh) \Longleftrightarrow Ad + N_{\mathcal{K}(y,v)}(d) \ni Bh.$$

The reader is referred to [4, Chapter 2.2] for more details on these concepts.

Concerning existence of solutions, we mention that the MPEC (1) admits a solution in this setting, see e.g., [14]. This follows from the Lipschitz continuity of $S \circ B$, the weak lower semicontinuity of the objective functional, and the closedness and convexity of the set of admissible controls U_{ad} defined by

$$U_{ad} := \left\{ u \in L^2(\Omega) \,|\, a \le u \le b \text{ a.e. } \Omega \right\},$$

see e.g., [14] or [3] for more details and discussions.

We are now ready to establish our first result. The choice of terminology below is based on the similarity to the corresponding finite-dimensional concept. It can be observed from the proof of this, as well as many of the subsequent results, that it is certainly possible to work with more general objective functionals J than the tracking-type functional chosen for our setting. Nevertheless, we decide to use the tracking-type functional as it often appears in applications and since it helps to better illustrate essential features of our methods and results.

Theorem 2.1 (**B-Stationarity of an Optimal Solution**). *Let $(\bar{u},\bar{y})$ be a locally optimal solution to the original MPEC (1). Then the following optimality condition holds*

$$\alpha(\bar{u},h)_{L^2} + (\bar{y} - y_d, d)_{L^2} \ge 0, \forall (h,d) \in \left[T_{U_{ad}}(\bar{u}) \times H_0^1(\Omega) \right] \cap \operatorname{gph} S'(B\bar{u}; B\cdot). \tag{19}$$

Equivalently, if $(\bar{u},\bar{y})$ is a locally optimal solution to the MPEC (1), then the origin in $L^2(\Omega)\times H_0^1(\Omega)$ is a solution to the following MPEC

$$\begin{aligned} &\min \ \alpha(\bar{u},h)_{L^2} + (\bar{y} - y_d, d)_{L^2} \ \textit{over} \ (h,d) \in L^2(\Omega) \times H_0^1(\Omega) \\ &s.t. \quad h \in T_{U_{ad}}(\bar{u}), \qquad Ad + N_{\mathcal{K}(\bar{y},\bar{v})}(d) \ni Bh. \end{aligned} \tag{20}$$

Proof. Given the properties of the solution mapping S to the variational inequality described above, we can reformulate the MPEC (1) as the following nonsmooth optimization problem:

$$\begin{array}{l} \min V(u) := \frac{1}{2}||S(Bu) - y_d||_{L^2}^2 + \frac{\alpha}{2}||u||_{L^2}^2 \text{ over } u \in L^2(\Omega) \\ s.t. \quad u \in U_{ad}. \end{array} \tag{21}$$

Clearly the mapping $V : L^2(\Omega) \to \mathbb{R}$ is directionally differentiable and Lipschitz continuous. Next we modify the nonsmooth problem (21) one step further to

$$\min V(u) + I_{U_{ad}}(u) \text{ over } u \in L^2(\Omega). \tag{22}$$

Given an arbitrary locally optimal solution $\bar{u}$ to problem (22), observe that the corresponding pair $(\bar{u}, \bar{y})$ is a locally optimal solution to the original MPEC (1), and vice versa. Moreover, it can be argued (see e.g. [2] Chapter 6.1.3) that the following condition must hold

$$\liminf_{\substack{t \to 0^+ \\ h' \to_{L^2} h}} \frac{V(\bar{u} + th') - V(\bar{u}) + I_{U_{ad}}(\bar{u} + th') - I_{U_{ad}}(\bar{u})}{t} \geq 0, \ \forall h \in L^2(\Omega). \tag{23}$$

To proceed, we first note that if $h \in L^2(\Omega)$ but $h \notin T_{U_{ad}}(\bar{u})$, then there either exist no sequences $t_k \to 0^+$ or $h_k \to_{L^2} h$ such that $\bar{u} + t_k h_k \in U_{ad}$. Thus, for such h the limit inferior in (23) is equal to $+\infty$. Suppose now that $h \in T_{U_{ad}}(\bar{u})$. Then by definition there exist sequences $t_k \to 0^+$ and $h_k \to_{L^2} h$ such that $\bar{u} + t_k h_k \in U_{ad}$. For such sequences, the difference quotients in (23) reduce to

$$\frac{V(\bar{u} + t_k h_k) - V(\bar{u})}{t_k}.$$

Then by using the directional differentiability and the fact that V is Lipschitz continuous (and therefore $V'(\bar{u}; \cdot)$ as well), we further reduce the difference quotients to

$$\frac{V(\bar{u} + t_k h_k) - V(\bar{u})}{t_k} = \frac{V(\bar{u}) + t_k V'(\bar{u}; h_k) + o(t_k) - V(\bar{u})}{t_k} = V'(\bar{u}; h_k) + \frac{o(t_k)}{t_k},$$

which implies in turn that

$$V'(\bar{u}; h) \geq 0, \forall h \in T_{U_{ad}}(\bar{u}).$$

The final step of the proof requires us to compute the derivative $V'(\bar{u}, h)$. By definition, we need to calculate the following limit:

$$\lim_{t \to 0^+} \frac{\frac{1}{2}||S(B(\bar{u} + th)) - y_d||_{L^2}^2 + \frac{\alpha}{2}||\bar{u} + th||_{L^2}^2 - \frac{1}{2}||S(B\bar{u}) - y_d||_{L^2}^2 - \frac{\alpha}{2}||\bar{u}||_{L^2}^2}{t}.$$

We observe first that

$$\frac{\frac{\alpha}{2}||\bar{u} + th||_{L^2}^2 - \frac{\alpha}{2}||\bar{u}||_{L^2}^2}{t} = \alpha(\bar{u}, h)_{L^2} + \frac{\alpha t}{2}||h||_{L^2}^2.$$

Similarly, we reduce the remaining terms (using the directional differentiability of S) to

$$\frac{\frac{1}{2}||S(B\bar{u} + th)) - y_d||_{L^2}^2 - \frac{1}{2}||S(B\bar{u}) - y_d||_{L^2}^2}{t} =$$

$$(S(B\bar{u}) - y_d, S'(B\bar{u}; Bh) + \frac{o(t)}{t})_{L^2} + t||S'(B\bar{u}; Bh) + \frac{o(t)}{t}||_{L^2}^2.$$

Then by adding the reduced terms and passing to the limit, we obtain the equality

$$\lim_{t\to 0^+} \frac{\frac{1}{2}||S(B(\bar{u}+th))-y_d||^2_{L^2}+\frac{\alpha}{2}||\bar{u}+th||^2_{L^2}-\frac{1}{2}||S(B\bar{u})-y_d||^2_{L^2}-\frac{\alpha}{2}||\bar{u}||^2_{L^2}}{t} =$$
$$\alpha(\bar{u},h)_{L^2}+(S(B\bar{u})-y_d,S'(B\bar{u};Bh))_{L^2},$$

which completes the proof of the theorem via substitution. □

In comparison with the dual stationarity concepts (e.g., S- or C-stationarity), B-stationarity translates more directly into function space settings provided, of course, that the needed regularity properties of the solution map S are available. In fact, since $\mathcal{K}(\bar{y},\bar{v})$ is a closed convex cone in $H^1_0(\Omega)$, we can equivalently rewrite MPEC (20) as

$$\min\ \alpha(\bar{u},h)_{L^2}+(\bar{y}-y_d,d)_{L^2} \text{ over } (h,d)\in L^2(\Omega)\times H^1_0(\Omega)$$

$$s.t.\quad h\in T_{U_{ad}}(\bar{u}),$$

$$Bh-Ad\in[\mathcal{K}(\bar{y},\bar{v})]^-,\qquad d\in\mathcal{K}(\bar{y},\bar{v}),\qquad \langle Bh-Ad,d\rangle_{H^{-1},H^1_0}=0.$$

Continuing, we note that Theorem 2.1 also shows that if $\bar{u}\in U_{ad}$ such that $T_{U_{ad}}(\bar{u})=L^2(\Omega)$, then the so-called strong stationarity conditions as seen in [9] can be rederived without major difficulties provided that the operator B satisfies certain requirements; in particular, if B is the identity of $L^2(\Omega)$. By directly adapting the proof of Lemma 6.34 in [4], we obtain the following description:

$$T_{U_{ad}}=\left\{h\in L^2(\Omega)\,\middle|\,\begin{array}{l} h\geq 0,\ \text{a.e. on}\ \{x\in\Omega\,|u(x)=a(x)\}\\ h\leq 0,\ \text{a.e. on}\ \{x\in\Omega\,|u(x)=b(x)\}\end{array}\right\}.$$

Therefore, it can be argued that if the Lebesgue measure of the active set $\mathcal{A}_a(\bar{u})\cup\mathcal{A}_b(\bar{u})$ equals zero, then $T_{U_{ad}}(\bar{u})=L^2(\Omega)$. Thus, even though U_{ad} has an empty interior in $L^2(\Omega)$, there exist admissible points such that the tangent cone is equal to the entire space.

Under a fairly restrictive assumption, it is easy to derive the following corollary from Theorem 2.1, which yields a dual form of the B-stationarity conditions.

Corollary 2.2 (**Dual Form of B-Stationarity**). *Let $(\bar{u},\bar{y})$ be a locally optimal solution to the MPEC (1), where $\bar{u}\in U_{ad}$ such that $S'(B\bar{u};B\cdot)=:\Sigma_{\bar{u}}(\cdot)$ is a bounded linear operator from $L^2(\Omega)$ into $H^1_0(\Omega)$. Then the following optimality condition holds*

$$\alpha(\bar{u},h)_{L^2}+(B^*\Sigma^*_{\bar{u}}(\bar{y}-y_d),h)_{L^2}\geq 0,\forall h\in T_{U_{ad}}(\bar{u}),$$

which in dual form is equivalent to the inclusion

$$0\in\alpha\bar{u}+B^*\Sigma^*_{\bar{u}}(\bar{y}-y_d)+N_{U_{ad}}(\bar{u})$$

or equivalently the variational inequality

$$(\alpha\bar{u}+B^*\Sigma^*_{\bar{u}}(\bar{y}-y_d),\bar{u}-u')_{L^2}\geq 0,\ \forall u'\in U_{ad}.$$

In order to obtain workable KKT-type optimality conditions in the case where $S'(B\bar{u};B\cdot)$ is not a bounded linear operator, we would need to calculate the following polar cone:

$$\left[\big(T_{U_{ad}}(\bar{u})\times H^1_0(\Omega)\big)\cap \operatorname{gph} S'(B\bar{u};B\cdot)\right]^-_{L^2\times H^1_0}.$$

Unfortunately, it appears to be a difficult, if not impossible, task. Thus the need for a different set of more constructive tools for the derivation of dual conditions (in both finite and infinite dimensions) is evident. In the next section we proceed in this direction by using advanced tools of variational analysis and generalized differentiation.

3 Dual Optimality Conditions via Limiting Variational Calculus

In order to develop our technique, we first recall several definitions and concepts from variational analysis and generalized differentiation. Our main source is the two-volume monograph [15, 16] Throughout the following section, unless otherwise noted, all spaces we will be assumed to be Hilbert spaces. Nevertheless, we stress that these objects along with the accompanying results can be defined/proved in much more general settings than Hilbert spaces, see e.g., [15, 16].

Definition 3.1 (**The Regular/Fréchet Normal Cone**). *Let $C \subset X$ and assume that X and its dual X^* possess compatible topologies. Then the multifunction (set-valued mapping) $\widehat{N}_C : X \rightrightarrows X^*$ defined by*

$$\widehat{N}_C(x) := \left\{ x^* \in X^* \;\middle|\; \limsup_{\substack{x' \to_X x \\ x' \in C}} \frac{\langle x^*, x' - x \rangle_{X^*,X}}{||x' - x||_X} \leq 0 \right\}, \quad x \in C, \tag{24}$$

and $\widehat{N}_C(x) := \emptyset$ for $x \notin C$ is called the REGULAR/FRÉCHET NORMAL CONE *to C.*

Unfortunately, it is know that $\widehat{N}_C$, which is convex for each $x \in C$, does not admit a satisfactory calculus. This restricts the scope of applications of (24)—in particular, to deriving workable, multiplier-based optimality conditions for the class of MPECs under consideration. The situation changes significantly when we apply an appropriate limiting procedure to the mapping $\widehat{N}_C(\cdot)$.

Definition 3.2 (**The Limiting/Mordukhovich Normal Cone**). *Let $C \subset X$ and assume that X and its dual X^* possess compatible topologies. The multifunction $N_C : X \rightrightarrows X^*$ defined by*

$$N_C(x) := \left\{ x^* \in X^* \;\middle|\; \exists x_k \to_X \bar{x},\, \exists x_k^* \rightharpoonup_{X^*} x^* : \, x_k^* \in \widehat{N}_C(x_k),\ \forall k \in \mathbb{N} \right\} \tag{25}$$

is called the LIMITING/MORDUKHOVICH NORMAL CONE *to C.*

If the set C is convex, both cones (24) and (25) agree with the normal cone of convex analysis. However, for general sets C the limiting normal cone (25) and the corresponding coderivative and subdifferential constructions admit a full set of calculus rules often needed for various applications.

Next we define the notions of *coderivatives* for set-valued (in particular, single-valued) mappings generated by the corresponding normal cones (24) and (25).

Definition 3.3 (**Coderivatives**). *Let $\Phi : X \rightrightarrows Y$ be a set-valued mapping between (paired) spaces X and Y, and let $(x, y) \in \operatorname{gph} \Phi$. The* REGULAR/FRÉCHET CODERIVATIVE *of Φ at (x, y) is the multifunction $\widehat{D}^*\Phi(x, y)\colon Y^* \rightrightarrows X^*$ defined by*

$$h^* \in \widehat{D}^*\Phi(x, y)(d^*) \Longleftrightarrow (h^*, -d^*) \in \widehat{N}_{\operatorname{gph} \Phi}(x, y). \tag{26}$$

The LIMITING/MORDUKHOVICH CODERIVATIVE *$D^*\Phi(x, y)$ of Φ at $(x, y) \in \operatorname{gph} \Phi$ is similarly defined by*

$$h^* \in D^*\Phi(x, y)(d^*) \Longleftrightarrow (h^*, -d^*) \in N_{\operatorname{gph} \Phi}(x, y). \tag{27}$$

We observe from (24)–(27) that the limiting coderivative (27) admits the following representation:

$$h^* \in D^*\Phi(x, y)(d^*) \Longleftrightarrow \left\{ \begin{array}{l} x_k \to_X x \\ y_k \to_Y y \\ d_k^* \rightharpoonup_{Y^*} d^* \\ h_k^* \rightharpoonup_{X^*} h^* \end{array} \right\} : h_k^* \in \widehat{D}^*\Phi(x_k, y_k)(d_k^*). \tag{28}$$

Note that if the condition "$d_k^* \to_{Y^*} d^*$" is replaced by the weaker condition "$d_k^* \rightharpoonup_{Y^*} \bar{d}^*$" in (28), then the corresponding construction $D_M^*\Phi(x,y)$ is known as the *mixed coderivative* of Φ at $(x,y) \in \operatorname{gph}\Phi$.

In the case of $\Phi\colon X \to Y$ strictly differentiable at x (in particular, C^1 around this point) with the derivative $\nabla\Phi(x)$, all the three coderivatives above reduce to the adjoint derivative operator

$$\widehat{D}^*\Phi(x)(y^*) = D^*\Phi(x)(y^*) = D_M^*\Phi(x)(y^*) = \big\{\nabla\Phi(x)^*y^*\big\}, \quad y^* \in Y^*,$$

where $y = \Phi(x)$ is omitted due to single-valuedness. In general these coderivative mappings are positively homogeneous in y^* with *full calculi* for $D^*\Phi$ and $D_M^*\Phi$ and a rather restrictive one for $\widehat{D}^*\Phi$. For mappings between infinite-dimensional spaces the aforementioned calculus rules require appropriate "normal compactness" conditions, which are automatic in finite dimensions. The weakest ones among such conditions are given in the next definition.

Definition 3.4 (Sequential Normal Compactness). *Let $\Phi : X \rightrightarrows Y$ be a set-valued mapping between (paired) spaces X and Y, and let $(x,y) \in \operatorname{gph}\Phi$. We say that $\Phi : X \rightrightarrows Y$ is* SEQUENTIALLY NORMALLY COMPACT (SNC) *at $(x,y) \in \operatorname{gph}\Phi$ if for any collection of sequences $\{x_k\} \subset X$, $\{y_k\} \subset Y$, $\{x_k^*\} \subset X^*$, and $\{y_k^*\} \subset Y^*$ satisfying*

$$\left.\begin{array}{l} x_k \to_X \bar{x} \\ y_k \to_Y \bar{y} \\ x_k^* \rightharpoonup_{X^*} 0 \\ y_k^* \rightharpoonup_{Y^*} 0 \end{array}\right\} \text{ with } y_k^* \in \widehat{D}^*\Phi(x_k,y_k)(x_k^*)$$

it follows that $||x_k^*||_{X^*} \to 0$ and $||y_k^*||_{Y^*} \to 0$. If the requirement that $y_k^* \rightharpoonup_{X^*} 0$ above is replaced by $||y_k^*||_{Y^*} \to 0$, then Φ is said to be PARTIALLY SEQUENTIALLY COMPACT (PSNC) *at (x,y).*

Observe that, besides finite-dimensional settings, the SNC and PSNC properties automatically hold under appropriate Lipschitzian conditions imposed on set-valued and single-valued mappings and are preserved under various compositions; see [15] for such SNC and PSNC calculi. In particular, Φ is SNC at (x,y) if its graph is *compactly epi-Lipschitzian* around this point in the sense of Borwein-Strojwás; see [15, Theorem 1.26]. The weaker PSNC property always holds for locally Lipschitzian and Lipschitz-like mappings from the next definition. Unfortunately, in the context of Lebesgue spaces $L^p(\Omega)$ and Sobolev spaces $W^{m,p}(\Omega)$ with $1 \le p < +\infty$, the aforementioned compactly epi-Lipschitzian property cannot hold in general for the important class of subsets defined by pointwise constraints.

Definition 3.5 (The Aubin Property). *Let $\Phi : X \rightrightarrows Y$ be a set-valued mapping between (paired) spaces X and Y, and let $(x,y) \in \operatorname{gph}\Phi$. We say that Φ has the* AUBIN PROPERTY *or is* LIPSCHITZ-LIKE/PSEUDO-LIPSCHITZ *at (x,y) if there are neighborhoods $\mathcal{U}$ of x and $\mathcal{V}$ of y together with a constant $L > 0$ such that*

$$\|y - y'\|_Y \le L\|u - u'\|_X, \ \forall (x,y),(x',y') \in [\mathcal{U} \times \mathcal{V}] \cap \operatorname{gph}\Phi. \tag{29}$$

It immediately follows from (29) that for single-valued mappings $\Phi\colon X \to Y$, the Aubin property reduces to the classical local Lipschitz continuity. Moreover, the coderivative criterion from [15, Theorem 4.0] asserts that a closed-graph mapping $\Phi : X \rightrightarrows Y$ has the Aubin property around $(x,y) \in \operatorname{gph}\Phi$ if and only if it is PSNC at this point and the injectivity condition "$D_M^*\Phi(x,y)(0) = \{0\}$" holds.

Now we apply the facts above together with coderivative calculus and related results from [15, 16] to derive necessary optimality conditions for the original MPEC (1), which we restate in compact form for convenience as follows:

$$\begin{array}{l} \min \frac{1}{2}||y-y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)} \text{ over } (u,y) \in L^2(\Omega) \times H^1_0(\Omega) \\ s.t. \quad u \in U_{ad}, \quad y = S(Bu). \end{array} \tag{30}$$

Our first result provides a necessary optimality condition for the MPEC (1) in terms of the limiting coderivative of S and the convex normal cone to the control set U_{ad}.

Proposition 3.6 (**Limiting Optimality Conditions for the MPEC**). *Let $(\bar{u}, \bar{y})$ be a locally optimal solution to the MPEC (1). Then we have*

$$0 \in \alpha\bar{u} + B^*D^*S(B\bar{u}, y)(\bar{y} - y_d) + N_{U_{ad}}(\bar{u}). \tag{31}$$

Proof. As follows from our discussions in Section 2, the solution map S is single-valued and Lipschitz continuous under the standing assumptions. Since the operator B is linear and bounded, the composition $S \circ B$ is Lipschitz continuous as well. Thus, by the previously mentioned coderivative criterion from [15, Theorem 4.10], we get the equivalent description of local Lipschitz continuity:

- $S \circ B$ is partially sequentially normally compact (PSNC) at $(\bar{u}, \bar{y})$, and
- $D^*_M(S \circ B)(\bar{u}, \bar{y})(0) = \{0\}$ (injectivity via the mixed coderivative).

Applying now the necessary optimality conditions for abstract MPECs established in [16, Theorems 5.33 and 5.34] to the case of our MPEC (1), considered in form (30), and taking into account that the cost functional therein is smooth, we conclude that the PSNC and qualification assumptions required by [16, Theorems 5.33 and 5.34] are satisfied. It follows that

$$0 \in \alpha\bar{u} + D^*(S \circ B)(\bar{u}, \bar{y})(\bar{y} - y_d) + N_{U_{ad}}(\bar{u}).$$

Finally, it follows from the calculus result of [15, Corollary 3.16] that

$$D^*(S \circ B)(\bar{u}, \bar{y})(\bar{y} - y_d) \subset B^*D^*S(B\bar{u}, \bar{y})(\bar{y} - y_d),$$

and therefore, the asserted optimality condition (31) holds. □

Remark 3.7 (**Regularity of the Optimal Control**). We mention that if $U_{ad} = L^2(\Omega)$, then $N_{U_{ad}}(\bar{u}) = \{0\}$. Moreover, if B acts as the identity on $L^2(\Omega)$, then $B^*y^* \in H^1_0(\Omega)$ for all $y^* \in D^*S(B\bar{u}, y)(\bar{y} - y_d)$. Thus it follows from Proposition 3.6 that the optimal solution $\bar{u}$ enjoys an increased regularity in this case. Observe that the above arguments can be easily extended to more general situations.

The remaining part of this section is dedicated to the explicit characterization of the coderivative in the necessary optimality condition of Proposition 3.6. Developing this derivation technique, we arrive at multiplier-based optimality conditions for the original MPEC.

We start by first observing the following description of the coderivative in (31) in light of [15, Corollary 2.36]:

$$\begin{aligned} D^*S(B\bar{u}, \bar{y})(\bar{y} - y_d) = \big\{ \bar{p}^* \in H^1_0(\Omega) \,| \\ \exists y^*_k \to_{H^{-1}} B\bar{u}, \exists y_k \to_{H^1_0} \bar{y}, \exists q^*_k \rightharpoonup_{H^{-1}} \bar{y} - y_d, \exists p^*_k \rightharpoonup_{H^1_0} \bar{p}^* : p^*_k \in \widehat{D}^*S(y^*_k, y_k)(q^*_k), \forall k \Big\}. \end{aligned}$$

By simply referring to the definition of the regular coderivative (26), we know that the previous equation can be understood as

$$D^*S(B\bar{u},\bar{y})(\bar{y}-y_d) = \big\{\bar{p}^* \in H_0^1(\Omega) \,|\, \exists y_k^* \to_{H^{-1}} B\bar{u}, \exists y_k \to_{H_0^1} \bar{y}, \exists q_k^* \rightharpoonup_{H^{-1}} \bar{y}-y_d, \exists p_k^* \rightharpoonup_{H_0^1} \bar{p}^* : (p_k^*,-q_k^*) \in \widehat{N}_{\mathrm{gph}\, S}(y_k^*,y_k), \forall k \big\}.$$

Using [15, Theorem 1.10], we approximate the limiting coderivative of S by replacing $\widehat{N}_{\mathrm{gph}\, S}(y_k^*,y_k)$ with the larger polar contingent cone $[T_{\mathrm{gph}\, S}(y_k^*,y_k)]^-$. Note that the contingent cone to the graph of S coincides with the graph of the so-called contingent derivative of S; see [2]. In the current setting with S being single-valued, Lipschitz continuous and Hadamard directionally differentiable, the contingent derivative coincides with the Hadamard directional derivative. It was shown in the proof of [9, Theorem 4.6] that

$$(p_k^*,-q_k^*) \in [T_{\mathrm{gph}\, S}(y_k^*,y_k)]^- \Longleftrightarrow p_k^* \in \mathcal{K}(y_k,v_k),\ A^*p_k^* - q_k^* \in [\mathcal{K}(y_k,v_k)]^-,$$

where $v_k \in N_M(y_k)$ such that $v_k = y_k^* - Ay_k$, and where $\mathcal{K}(\bar{y},\bar{v})$ is the critical cone (18). This leads to the following characterization of the coderivative.

Proposition 3.8 (**Characterizing the Coderivative via the Critical Cone**). *Let $v_k := y_k^* - Ay_k$. Then elements of the limiting coderivative of the solution map* (16) *are described by*

$$\bar{p} \in D^*S(B\bar{u},\bar{y})(\bar{y}-y_d) \Longrightarrow \left.\begin{array}{l} \exists y_k^* \to_{H^{-1}} B\bar{u} \\ \exists y_k \to_{H_0^1} \bar{y} \\ \exists q_k \rightharpoonup_{H^{-1}} \bar{y}-y_d \\ \exists p_k \rightharpoonup_{H_0^1} \bar{p} \end{array}\right\} : p_k \in \mathcal{K}(y_k,v_k) \text{ and } A^*p_k - q_k \in [\mathcal{K}(y_k,v_k)]^-.$$

In order to provide our preliminary set of optimality conditions. We will need a few auxiliary results. In the following we make use of a constraint qualification due to Jeyakumar and Wolkowicz (see Lemma 2.2 b in [10]).

Lemma 3.9 (**Closedness of Conic Sums**). *Let C, D be two closed and convex cones in a Hilbert space X. Assume that the angle between C and $-D$ is positive, i.e.,*

$$\inf\{(c,d)_X \,|\, 1 = ||c||_X = ||d||_X,\, c \in C,\, d \in D\} < 1.$$

Then the conic sum $C + D$ is closed in X.

Using Lemma 3.9, we can easily draw the following conclusion.

Corollary 3.10 (**Conic Summation with Straight Lines**). *Suppose that X is a Hilbert space and C is a (nonempty) closed convex cone. Then for any $c \in C$, the sum $C + \mathbb{R}c$ is closed in X.*

Proof. Let $b \in \mathbb{R}c$ such that $b = \frac{-1}{||c||_X}c$. Since C is a cone and $c \in C$, it follows that $a = \frac{1}{||c||_X}c \in C$. But then $||a||_X = ||b||_X = 1$ and $(a,b)_X = -1 < 1$. Therefore,

$$\inf\{(c,d)_X \,|\, ||c||_X = ||d||_X = 1,\, c \in C,\, d \in D\} \leq -1 < 1.$$

It follows that the angle between C and $\mathbb{R}c$ is positive. Thus, by Lemma 3.9, the assertion holds. □

Returning now to a setting closer to our MPEC, we employ the above lemmas to derive a useful representation result.

Corollary 3.11 (**The Polar Critical Cone**). *Let M be a closed and convex subset of a Hilbert space Y, and let $v \in N_M(y)$. Then the following hold:*

(i) $N_M(y) + \mathbb{R}v$ *is closed in* Y^*,

(ii) $[\mathcal{K}(y,v)]^- = N_M(y) + \mathbb{R}v$.

Proof. Since Y^* is a Hilbert space and the normal cone $N_M(y) \subset Y^*$ is closed and convex, (i) follows from Corollary 3.10. For the proof of (ii), we observe that both cones $T_M(y)$ and $\{v\}^\perp$ are closed and convex in Y. Thus it follows from convex analysis and (i) that

$$[\mathcal{K}(y,v)]^- = \Big[T_M(y) \cap \{v\}^\perp\Big]^- = \text{cl } \{N_M(y) + \mathbb{R}v\}_{Y^*} = N_M(y) + \mathbb{R}v.$$

□

Corollary 3.11 helps us to explicitly characterize the sequences from the polar critical cones $[\mathcal{K}(y_k, v_k)]^-$ in Proposition 3.8 that weakly converge in $H^{-1}(\Omega)$. Before establishing this result, we point out a simple fact concerning the convergence of normal cone mappings to closed convex sets. Suppose that X is a Banach space and $C \subset X$ is a closed convex subset. Let $x_k \in C$ be such that $x_k \to x$ in X and let $z_k \in N_C(x_k)$ be such that $z_k \rightharpoonup z$ in X^*. Then by the definition of the normal cone we have $\langle z_k, x_k - x' \rangle \leq 0$ for all $x' \in C$. For an arbitrary element $x' \in C$, it follows that $\langle z_k, x_k - x' \rangle \to \langle z, x - x' \rangle$ and thus $z \in N_C(x)$. We make use of this property in the proof of the following proposition.

Proposition 3.12 (**Limits of the Polar Critical Cones**). *Let $M := \big\{y \in H_0^1(\Omega) \mid y \geq \psi \text{ a.e. } \Omega\big\}$, where $\psi \in H^1(\Omega)$ with $\psi|_{\partial\Omega} \leq 0$, and where $\Omega \subset \mathbb{R}^n$ with $1 \leq n \leq 3$ is open and bounded. Let $r_k \in [\mathcal{K}(y_k, v_k)]^-$ from the polar critical cones be such that*

$$r_k \rightharpoonup_{H^{-1}} \bar{r}, \quad y_k \to_{H_0^1} \bar{y}, \quad v_k \to_{H^{-1}} \bar{v}$$

with $v_k \in N_M(y_k)$, where the convex set M is defined in (2). *Then we have the condition*

$$0 = \langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi = 0 \text{ a.e. } \mathcal{A}(\bar{y}) \tag{32}$$

and either

$$0 \leq \langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi \geq 0 \text{ a.e. } \mathcal{A}(\bar{y}), \tag{33}$$

or

$$0 \leq \langle \bar{r}, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi \leq 0 \text{ a.e. } \mathcal{A}(\bar{y}). \tag{34}$$

Proof. Corollary 3.11 yields $r_k \in N_M(y_k) + \mathbb{R}v_k$ for all k. Then arguing similarly to the proof of [4, Theorem 6.57], for each k we find elements $w_k \in N_M(y_k)$ and $\alpha_k \in \mathbb{R}$ such that $r_k = w_k + \alpha_k v_k$ and the conditions

$$\begin{aligned} 0 &\geq \langle w_k, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi \geq 0 \text{ a.e. } \mathcal{A}(y_k), \\ 0 &= \langle w_k, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi = 0 \text{ a.e. } \mathcal{A}(y_k) \end{aligned}$$

are satisfied. The same conditions also hold for v_k and $\bar{v}$ (with y_k replaced by $\bar{y}$) since $v_k \in N_M(y_k)$ and $\bar{v} \in N_M(\bar{y})$.

Consider now that $v_k = (r_k - w_k)/\alpha_k$ and assume first, without loss of generality, that $\alpha_k > 0$ for all k; otherwise, take a subsequence. By the definition of the normal cone, we infer

$$\begin{aligned}
0 &\geq \langle (r_k - w_k)/\alpha_k, y_k - y' \rangle_{H^{-1},H_0^1},\ \forall y' \in M \\
&\Longrightarrow \langle r_k, y_k - y' \rangle_{H^{-1},H_0^1} \leq \langle w_k, y_k - y' \rangle_{H^{-1},H_0^1},\ \forall y' \in M \\
&\Longrightarrow \langle r_k, y_k - y' \rangle_{H^{-1},H_0^1} \leq 0,\ \forall y' \in M.
\end{aligned}$$

As argued in the paragraph preceding the statement of the proposition, we get $\bar{r} \in N_M(\bar{y})$, which implies (32) and (33) by the structure of the normal cone to the convex set M.

Suppose now that $\alpha_k < 0$ taking a subsequence if necessary. Arguing as above, we get from the definition of the normal cone that

$$\begin{aligned}
0 &\geq \langle (r_k - w_k)/\alpha_k, y_k - y' \rangle_{H^{-1},H_0^1},\ \forall y' \in M \\
&\Longrightarrow \langle r_k, y_k - y' \rangle_{H^{-1},H_0^1} \geq \langle w_k, y_k - y' \rangle_{H^{-1},H_0^1},\ \forall y' \in M \\
&\Longrightarrow \langle -r_k, y_k - y' \rangle_{H^{-1},H_0^1} \leq \langle w_k, y' - y_k \rangle_{H^{-1},H_0^1},\ \forall y' \in M.
\end{aligned}$$

Since $y_k = \psi\ a.e.\,\mathcal{A}(y_k)$ and the inclusion $y' \in M$ implies that $y' \geq \psi\ a.e.\,\Omega$, it follows that $y' - y_k = y' - \psi \geq 0\ a.e.\,\mathcal{A}(y_k)$. Then using the characterization of w_k, we have that

$$\langle -r_k, y_k - y' \rangle_{H^{-1},H_0^1} \leq 0, \forall y' \in M.$$

Arguments similar to those above yield (32) and (34). This completes the proof of the proposition. □

Now we are ready to establish the main result of this section, which provides necessary optimality conditions for the MPEC setting under consideration in terms of the initial data of problem (1). Note that these conditions are close to M-stationarity while somewhat different from it; see the corresponding discussion in Section 1.

Theorem 3.13 (Limiting Stationarity Conditions). *Let $(\bar{u}, \bar{y})$ be a locally optimal solution to MPEC (1). Then there exist $\bar{p} \in H_0^1(\Omega)$, $\bar{r} \in H^{-1}(\Omega)$, $\bar{v} \in N_M(\bar{y})$, and $\bar{s} \in N_{U_{ad}}(\bar{u})$ along with sequences $\{p_k\} \subset H_0^1(\Omega)$ and $\{r_k\} \subset H^{-1}(\Omega)$ such that*

$$p_k \rightharpoonup_{H_0^1(\Omega)} \bar{p}, \quad r_k \rightharpoonup_{H^{-1}(\Omega)} \bar{r}, \tag{35}$$

and the following limiting conditions

$$0 = \alpha\bar{u} + B^*\bar{p} + \bar{s}, \tag{36}$$

$$0 = \bar{y} - y_d - A^*\bar{p} + \bar{r}, \tag{37}$$

$$0 = A\bar{y} - B\bar{u} + \bar{v}, \tag{38}$$

$$0 = \langle \bar{v}, \bar{p} \rangle_{H^{-1},H_0^1}, \tag{39}$$

$$0 \geq \limsup_{k\to\infty} \langle r_k, p_k \rangle_{H^{-1},H_0^1}, \tag{40}$$

$$0 = \langle \bar{r}, \varphi \rangle_{H^{-1},H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi = 0\ a.e.\,\mathcal{A}(\bar{y}) \tag{41}$$

are satisfied together with one of the alternative conditions: either

$$0 \leq \langle \bar{r}, \varphi \rangle_{H^{-1},H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi \geq 0\ a.e.\,\mathcal{A}(\bar{y}) \tag{42}$$

or

$$0 \leq \langle \bar{r}, \varphi \rangle_{H^{-1},H_0^1},\ \forall \varphi \in H_0^1(\Omega) : \varphi \leq 0\ a.e.\,\mathcal{A}(\bar{y}). \tag{43}$$

Proof. Equations (36) and (37) follow directly from Proposition 3.6, whereas equation (38) is due to the feasibility of an optimal solution.

According to Proposition 3.8, we have $p_k \in T_M(y_k) \cap \{y_k^* - Ay_k\}^\perp$. Hence, $\langle p_k, y_k^* - Ay_k\rangle_{H_0^1,H^{-1}} = 0$ for all k. Since $p_k \rightharpoonup \bar{p}$ in H_0^1 and $y_k^* - Ay_k \to B\bar{u} - A\bar{y}$ in $H^{-1}(\Omega)$, it follows that $\langle \bar{p}, B\bar{u} - A\bar{y}\rangle_{H_0^1,H^{-1}} = 0$. Thus, we get (39). Further, the existence of the sequence p_k claimed above follows from Proposition 3.8, whereas r_k is defined by

$$r_k := A^* p_k - q_k \in [\mathcal{K}(y_k, v_k)]^- ,$$

with $q_k \rightharpoonup \bar{y} - y_d$ in $H^{-1}(\Omega)$. However, despite the fact that $\langle r_k, p_k\rangle \le 0$ for all k, we are only provided with weak convergence of the sequences. Therefore, no statement can be made about the product $\langle \bar{r}, \bar{p}\rangle$. This leads to equation (40).

Finally, we see that the relations in (41)-(43) follow directly from Proposition 3.12. □

Remark 3.14 (**Discussions on the Limiting Stationarity Conditions**)**.** The following observations on the limiting stationarity conditions obtained above are in order:

(i) The reader has most likely noted that we did not use the inclusion $p_k^* \in T_M(y_k)$ in the proof of Theorem 3.13 to further characterize $\bar{p}$. We elaborate on this here. For each k the inclusion $p_k \in T_M(y_k)$ implies the existence of sequences $t_l^k \to 0^+$ and $d_l^k \to p_k$ with $y_k + t_l^k d_l^k \in M$. Then for each k there exists a natural number l_k such that

$$||p_k - d_l^k||_{H_0^1} \le 1/k, \ \forall l \ge l_k.$$

Now let $\{q_k\}$ be a sequence of real numbers such that $q_k \to 0^+$ strictly monotonically. For each k select another natural number L_k with $t_l^k \le q_k$ whenever $l \ge L_k$. Using these sequences and constants, we define two new sequences by letting $n(k) := \max\{l_k, L_k\}$ for each k such that

$$\delta_k := d_{n(k)}^k \ \text{ and } \ \tau_k := t_{n(k)}^k, \ \forall k.$$

It follows that $\|p_k - \delta_k\|_{H_0^1} \to 0$ as $k \to \infty$, which yields $\delta_k \rightharpoonup \bar{p}$ in $H_0^1(\Omega)$, and $\tau_k \to 0^+$. In addition we have $y_k + \tau_k\delta_k \in M$ for all k. This implies, in particular, that $\bar{p}$ is an element of the weak Painlevé-Kuratowski outer limit of the family of sets $\{t^{-1}(M - y)\}_{t>0}$, which amounts to the so-called (weak) *paratingent cone*; see, e.g., [2].

It is well known that the paratingent cone is too large to provide any meaningful local linearization of the set M. For example, let $M := [0, 1] \subset \mathbb{R}$. Here, $T_M(0) = \mathbb{R}_+$, which is locally a reasonable approximation of the set. Conversely, for a sequence $\varepsilon_k \to 0^+$, $T_M(\varepsilon_k) = \mathbb{R}$ for all k. Therefore, the paratingent cone at 0 becomes the entire space. Due to such a property even for the simplest of convex sets in finite dimensions, we do not attempt to characterize this cone in $H_0^1(\Omega)$ without more knowledge of the involved structures.

(ii) By consulting the literature (see, e.g., [4, Chapter 6]), we can easily calculate the involved convex normal cones. It follows then that $\bar{v} \in N_M(\bar{y})$ and $\bar{s} \in N_{U_{ad}}(\bar{u})$ imply the conditions

$$\begin{aligned}
&0 \le \bar{s} \ a.e. \, \mathcal{A}_l(\bar{u}) \quad \bar{s} = 0, \ a.e. \, \mathcal{I}(\bar{u}), \quad \bar{s} \ge 0 \ a.e. \, \mathcal{A}_u(\bar{u}),\\
&0 \ge \langle \bar{v}, \varphi\rangle_{H^{-1},H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi \ge 0 \ a.e. \, \mathcal{A}(\bar{y}),\\
&0 = \langle \bar{v}, \varphi\rangle_{H^{-1},H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi = 0 \ a.e. \, \mathcal{A}(\bar{y}).
\end{aligned}$$

Upon comparison with conditions (3)–(11), we see that those in (36)–(41) satisfy almost all the conditions for being C-stationary. In addition, (42) and (43) provide more information than is found in the C-stationarity conditions.

(iii) Unfortunately, nothing can be said about the sign of the product $\langle \bar{r}, \bar{p}\rangle_{H^{-1},H_0^1}$ unless it is known that $\bar{p}$ has a constant sign on the entire active set $\mathcal{A}(\bar{y})$; a fact, which would almost be

tantamount to obtaining strong stationary conditions. Indeed, the very definition of the coderivative provides the existence of the sequences p_k and q_k in Theorem 3.13, which are in turn used to define r_k. These sequences only need to converge weakly in their respective spaces despite the fact that q_k converges to an element with higher regularity—a clear disadvantage. In the next two sections we are not provided with the existence of these sequences; rather we must derive them. The advantage then becomes evident as we can show that $q_k \to \bar{y} - y_d$ in $L^2(\Omega)$, not merely weakly in $H^{-1}(\Omega)$.

4 Stationarity Conditions via Penalization of the Control Constraints

By using a combination of penalization techniques along with methods from variational analysis, we are assured to have strong stationarity conditions for each "penalized" MPEC. Moreover, we can show that the global solutions of the penalized problems converge to a global solution of the original MPEC (1). This information is used to obtain sequences $(\bar{p}_k, \bar{q}_k, \bar{r}_k) \in H_0^1(\Omega) \times L^2(\Omega) \times H^{-1}(\Omega)$ similar to those found in Proposition 3.6. The limit points of these sequences are then shown to satisfy a system of relations stronger than those found in Theorem 3.13.

We begin this section by simplifying the model class through the removal of the constraint $u \geq a$ *a.e.* Ω. It should be clear that the same arguments remain valid so that their application to bilateral control constraints can also be considered. Our new model problem becomes

$$\begin{array}{l} \min \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)} \text{ over } (u, y) \in L^2(\Omega) \times H_0^1(\Omega), \\ s.t. \quad u \leq b \text{ a.e. } \Omega, \\ \qquad Ay + N_M(y) \ni Bu. \end{array} \tag{44}$$

Thus from now on we denote by U_{ad} the set

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

Moreover, we assume further that the linear operator B is the identity on $L^2(\Omega)$ and henceforth cease to explicitly use it in the results below. All the other data assumptions for (1) remain the same, unless otherwise stated.

Continuing with the reduced model class (44), we now penalize the constraint on u with an L^2-penalty function derived from the Moreau-Yosida regularization of the indicator function of U_{ad}. This gives rise to the following class of MPECs:

$$\begin{array}{l} \min J_\gamma(u, y) := \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)} + \frac{\gamma}{2}||(u - b)_+||^2_{L^2} \text{ over } (u, y) \in L^2(\Omega) \times H_0^1(\Omega), \\ s.t. \quad y = S(u) \end{array} \tag{45}$$

with $\gamma > 0$, where $(\cdot)_+ := \max(0, \cdot)$ pointwise almost everywhere.

First we justify the required well-posedness of the penalization procedure.

Proposition 4.1 (Well-Posedness of the Penalization). *Let $\gamma_n \to \infty$ as $n \to \infty$. Then for each $n \in \mathbb{N}$ the MPEC problem (45) with $\gamma := \gamma_n$ has a (globally) optimal solution $(\bar{u}_n, \bar{y}_n)$. Moreover, if $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ is optimal to (44), then there exists a subsequence of $\{(\bar{u}_n, \bar{y}_n)\}$, denoted still by n, such that $(\bar{u}_n, \bar{y}_n) \to (\bar{u}, \bar{y})$ in the weak-strong topology on $L^2(\Omega) \times H_0^1(\Omega)$.*

Proof. Although the following arguments are rather standard, we present them for completeness. Upon recognizing that the penalty functional $||(\cdot - b)_+||^2_{L^2} : L^2(\Omega) \to \mathbb{R}$ is weakly lower semicontinuous and bounded from below, we can apply a classical argument (see, e.g., [14]) to show that

MPEC (45) has an optimal solution $(\bar{u}_n, \bar{y}_n)$ for each $\gamma_n > 0$. It follows from the definition that

$$\frac{1}{2}||\bar{y}_n - y_d||^2_{L^2} + \frac{\alpha}{2}||\bar{u}_n||^2_{L^2} + \frac{\gamma_n}{2}||(\bar{u}_n - b)_+||^2_{L^2} \leq \frac{1}{2}||y - y_d||^2_{L^2} + \frac{\alpha}{2}||u||^2_{L^2} + \frac{\gamma_n}{2}||(u-b)_+||^2_{L^2}, \quad \forall (u,y) \in L^2(\Omega) \times H^1_0(\Omega) : Ay + N_M(y) \ni u.$$

Then letting $(\bar{u}, \bar{y})$ be a globally optimal solution to (44), we obtain the inequality

$$\frac{1}{2}||\bar{y}_n - y_d||^2_{L^2} + \frac{\alpha}{2}||\bar{u}_n||^2_{L^2} + \frac{\gamma_n}{2}||(\bar{u}_n - b)_+||^2_{L^2} \leq \frac{1}{2}||\bar{y} - y_d||^2_{L^2} + \frac{\alpha}{2}||\bar{u}||^2_{L^2}, \tag{46}$$

from which the following conclusions are deduced:

(i) $\{u_n\}$ is bounded in $L^2(\Omega)$;

(ii) $\frac{\gamma_n}{2}||(\bar{u}_n - b)_+||^2_{L^2} \to 0$ as $n \to \infty$.

Hence there exists a control $u^* \in L^2(\Omega)$ and a subsequence $\{u_{n_l}\}$ such that $u_{n_l} \rightharpoonup u^*$ in $L^2(\Omega)$. Using the Lipschitz continuity of y as a function of u from $H^{-1}(\Omega)$ into $H^1_0(\Omega)$, we have

$$||\bar{y}_{n_l} - y^*||_{H^1_0} \leq C||\bar{u}_{n_l} - u^*||_{H^{-1}},$$

where $\bar{y}_{n_l}, y^*$ are solutions to the variational inequality associated with $u_{n_l}, u^* \in L^2(\Omega)$, respectively, and where $C > 0$. Since $L^2(\Omega)$ is compactly embedded in $H^{-1}(\Omega)$, there exists a subsequence $\{u_{n_{l_k}}\}$ with $u_{n_{l_k}} \to u^*$ in $H^{-1}(\Omega)$. This implies that $y_{n_{l_k}} \to y^*$ in $H^1_0(\Omega)$. Furthermore, since

$$\langle Ay_{n_{l_k}} - u_{n_{l_k}}, y_{n_{l_k}} - y' \rangle_{H^{-1}, H^1_0} \geq 0, \ \forall y' \in M,$$

passing to the limit as $k \to \infty$ yields

$$\langle Ay^* - u^*, y^* - y' \rangle_{H^{-1}, H^1_0} \geq 0, \ \forall y' \in M,$$

and thus $y^* = S(Bu^*)$. Then it is easy to check that (u^*, y^*) is in fact a feasible point of the original MPEC (44). Indeed, since the functional $F(\cdot) := ||(\cdot - b)_+||^2_{L^2} : L^2(\Omega) \to \mathbb{R}$ is weakly lower semicontinuous, it follows that

$$0 = \lim_{n\to\infty} F(u_n) = \liminf_{n\to\infty} F(u_n) \geq F(u^*) \Rightarrow ||(u^* - b)_+||^2_{L^2} = 0,$$

and hence, $u^* \leq b$ *a.e.* Ω. Taking now the limit inferior in (46) ensures that $(u^*, y^*) = (\bar{u}, \bar{y})$. □

By applying the same arguments as in the proof of Corollary 3.6, we check that any locally optimal solution $(\bar{u}, \bar{y})$ to (45) satisfies the necessary optimality condition

$$0 \in \nabla_u J_\gamma(\bar{u}, \bar{y}) + B^* D^* S(\bar{u}, \bar{y})\big(\nabla_y J_\gamma(\bar{u}, \bar{y})\big). \tag{47}$$

Since $B = B^*$ is the identity on $L^2(\Omega)$ and $\bar{p} \in D^*S(\bar{u}, \bar{y})(\nabla_y J_\gamma(\bar{u}, \bar{y}))$ is an element of $H^1_0(\Omega)$, we can argue that $\nabla_u J_\gamma(\bar{u}, \bar{y}) \in H^1_0(\Omega)$. This leads us to the following proposition.

Proposition 4.2 (Increased Regularity at a Solution). *If $(\bar{u}_\gamma, \bar{y}_\gamma)$ is a locally optimal solution of (45), then*

$$\alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+ \in H^1_0(\Omega).$$

Proof. Since $\nabla_u J_\gamma(\bar{u}_\gamma, \bar{y}_\gamma) = \alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+$, the result follows from the argument directly preceeding the statement of this proposition. □

Based on the results in [9], we now derive primal and dual optimality conditions for MPECs of type (45).

Theorem 4.3 (**S-Stationarity for Penalized MPECs**). *Let $(\bar{u}_\gamma, \bar{y}_\gamma)$ be a local optimal solution to MPEC (45). Then we have*

$$(\alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+, h)_{L^2} + (\bar{y}_\gamma - y_d, d)_{L^2} \geq 0, \forall (h, d) \in \operatorname{gph} S'(\bar{u}_\gamma, \cdot). \tag{48}$$

Moreover, there exist $\bar{p}_\gamma \in H_0^1(\Omega)$, $\bar{r}_\gamma \in H^{-1}(\Omega)$, and $\bar{v}_\gamma \in H^{-1}(\Omega)$ such that

$$0 = \alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+ + \bar{p}_\gamma, \tag{49}$$
$$0 = \bar{y}_\gamma - y_d - A^*\bar{p}_\gamma + \bar{r}_\gamma, \tag{50}$$
$$0 = A\bar{y}_\gamma - \bar{u}_\gamma + \bar{v}_\gamma \tag{51}$$

with the primal-dual triple $(\bar{p}_\gamma, \bar{r}_\gamma, \bar{v}_\gamma)$ satisfying the inclusions

$$\bar{p}_\gamma \in \mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma), \quad \bar{r}_\gamma \in [\mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma)]^-, \quad \bar{v}_\gamma \in N_M(\bar{y}_\gamma). \tag{52}$$

Proof. As the penalty functional is Fréchet differentiable from $L^2(\Omega)$ into $\mathbb{R}$ for each $\gamma > 0$, the primal optimality condition (48) can be derived by using the same argument that was applied in order to prove Theorem 2.1.

By the data assumptions, $h \in L^2(\Omega)$. Therefore, we can rewrite (48) as

$$(\alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+, h)_{L^2} + (\bar{y}_\gamma - y_d, d)_{L^2} \geq 0, \forall (h, d) \in \operatorname{gph} S'(\bar{u}_\gamma, \cdot).$$

Using the characterization (17) of $\operatorname{gph} S'(\bar{u}_\gamma, \cdot)$ and the result from Proposition 4.2, we may write

$$\langle \alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+, Ad + w\rangle_{H_0^1, H^{-1}} + \langle \bar{y}_\gamma - y_d, d\rangle_{H^{-1}, H_0^1} \geq 0, \forall (d, w) \in \operatorname{gph} N_{\mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma)}.$$

This is equivalent to defining $\bar{p}_\gamma \in H_0^1(\Omega)$ such that

$$\langle -A^*\bar{p}_\gamma + \bar{y}_\gamma - y_d, d\rangle_{H^{-1}, H_0^1} + \langle \bar{p}_\gamma, w\rangle_{H_0^1, H^{-1}} \geq 0, \forall (d, w) \in \operatorname{gph} N_{\mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma)}.$$

where

$$0 = \alpha\bar{u}_\gamma + \gamma(\bar{u}_\gamma - b)_+ + \bar{p}_\gamma.$$

Then since $[\operatorname{gph} N_{\mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma)}]^- = [\mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma)]^- \times \mathcal{K}(\bar{y}, \bar{v})$ in the $H^{-1}(\Omega) \times H_0^1(\Omega)$-topology (see, e.g., the proof of Theorem 4.6 in [9]), we obtain the relation

$$0 = \bar{y}_\gamma - y_d - A^*\bar{p}_\gamma + \bar{r}_\gamma$$

where

$$\bar{p}_\gamma \in \mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma) \text{ and } \bar{r}_\gamma \in [\mathcal{K}(\bar{y}_\gamma, \bar{v}_\gamma)]^-.$$

From which we obtain the assertion; (51) follows from feasibility. □

Given the well-known characterizations of the cones involved in the dual conditions of Theorem 4.3, see e.g. Lemmas 2.2 and 3.2 in [18], we have the additional sign conditions:

$$\bar{p}_\gamma \geq 0, \ a.e. \, \mathcal{A}(\bar{y}_\gamma), \tag{53}$$

$$0 = \langle \bar{u}_\gamma - A\bar{y}_\gamma, \bar{p}_\gamma \rangle_{H^{-1}, H_0^1}, \tag{54}$$

$$0 = \langle \bar{r}_\gamma, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi = 0, \ a.e. \, \mathcal{A}(\bar{y}_\gamma), \tag{55}$$

$$0 \geq \langle \bar{r}_\gamma, \varphi \rangle_{H^{-1}, H_0^1}, \ \forall \varphi \in H_0^1(\Omega) : \varphi \geq 0, \ a.e. \, \mathcal{A}(\bar{y}_\gamma) \text{ and } \langle \bar{u}_\gamma - A\bar{y}_\gamma, \varphi \rangle_{H^{-1}, H_0^1} = 0. \tag{56}$$

Thus, the optimality conditions (49)-(56) amount to S-stationarity conditions for the penalized MPEC (45). This result is not surprising, as the results in [9] provide S-stationarity conditions for much more general settings than considered here, provided the objective functional is Fréchet differentiable and there are no upper-level constraints. We have nevertheless decided to provide the derivation above in order to partially demonstrate the technique.

Though it was not needed for the proofs of the preceeding results, we can extended the result of Proposition 4.1. This is needed for the derivation of a limiting stationarity system.

Proposition 4.4 (**Strong Convergence of Penalized Solutions**). *Let $\gamma_n \to \infty$. Then for each $n \in \mathbb{N}$ the penalized MPEC* (45) *admits an optimal solution $(\bar{u}_n, \bar{y}_n)$. Moreover, if $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ is an optimal solution to* (44)*, then there is a subsequence of $\{(\bar{u}_n, \bar{y}_n)\}$, still indexed by n, which converges to $(\bar{u}, \bar{y})$ in the strong-strong topology on $L^2(\Omega) \times H_0^1(\Omega)$.*

Proof. Returning to the proof of Proposition 4.1, we can deduce from (46) that

$$||\bar{u}_n||_{L^2}^2 - ||\bar{u}||_{L^2}^2 \leq \frac{1}{\alpha}\Big(||\bar{y}_n - y_d||_{L^2}^2 - ||\bar{y} - y_d||_{L^2}^2\Big).$$

Using then the strong convergence of $\bar{y}_n \to \bar{y}$ in $H_0^1(\Omega)$ gives us

$$0 = ||\bar{u}||_{L^2}^2 - ||\bar{u}||_{L^2}^2 \leq \liminf_{n\to\infty} ||\bar{u}_n||_{L^2}^2 - ||\bar{u}||_{L^2}^2 \leq \liminf_{n\to\infty} \frac{1}{\alpha}\Big(||\bar{y}_n - y_d||_{L^2}^2 - ||\bar{y} - y_d||_{L^2}^2\Big) = 0$$

as well as

$$\limsup_{n\to\infty} ||\bar{u}_n||_{L^2}^2 - ||\bar{u}||_{L^2}^2 \leq \limsup_{n\to\infty} \frac{1}{\alpha}\Big(||\bar{y}_n - y_d||_{L^2}^2 - ||\bar{y} - y_d||_{L^2}^2\Big) = 0.$$

Thus $\bar{u}_n \rightharpoonup \bar{u}$ in $L^2(\Omega)$ and $||u_n||_{L^2} \to ||u||_{L^2}$. Since $L^2(\Omega)$ is a Hilbert space, the latter implies the strong convergence $\bar{u}_n \to \bar{u}$ in $L^2(\Omega)$. □

Next, we derive some auxiliary results needed for the main result of this section. Recall the following two notions of variational convergence:

Definition 4.5 (**Mosco Epi-Convergence and Graph Convergence**). *For $n \geq 1$, let $\phi_n, \phi : X \to \overline{\mathbb{R}}$ be proper convex lower semicontinuous functions and X a reflexive Banach space. One says that ϕ_n* EPI-CONVERGES IN THE SENSE OF MOSCO *to ϕ, denoted by $\phi_n \xrightarrow{\text{M}-\text{epi}} \phi$, provided the following two conditions hold for all $x \in X$:*

1. *$\forall x_n \rightharpoonup_X x$, $\phi(x) \leq \liminf_{n\to\infty} \phi_n(x_n)$,*
2. *$\exists x_n \to_X x$ such that $\phi(x) \geq \limsup_{n\to\infty} \phi_n(x_n)$.*

For $n \geq 1$, let A_n and A be maximal monotone operators from X into X^*. The sequence A_n is said to GRAPH CONVERGE *to A, denoted by $A_n \xrightarrow{\text{G}} A$, if the following property holds:*

For every $(x, y) \in \operatorname{gph} A$, there exists a sequence $(x_n, y_n) \in \operatorname{gph} A_n$ such that $x_n \to x$ strongly in X and $y_n \to y$ strongly in X^*.

We refer the reader to the monograph by Attouch [1] for more on these and related topics. After defining graph convergence, Attouch points out in Proposition 3.59 in [1] that for a sequence of maximal monotone operators $A_n \xrightarrow{\mathrm{G}} A$, the following holds:

For every sequence $(x_n, y_n) \in \operatorname{gph} A_n$ such that $x_n \to x$ strongly in X and $y_n \rightharpoonup y$ weakly in X^*, $(x, y) \in \operatorname{gph} A$ (and vice versa, by exchanging strong and weak).

This result shows that the convergence properties of sequences of normal cone mappings to convex sets discussed in Section 3 extends to the much broader class of maximal monotone operators.

We now apply these notions and results on variational convergence to our problem.

Lemma 4.6 (**Moreau-Yosida Approximations of Unilateral Pointwise Constraints**). *Let $\gamma_n \to \infty$, and let $b \in L^2(\Omega)$. Define the Moreau-Yosida regularization $F_n : L^2(\Omega) \to \overline{\mathbb{R}}$ by*

$$F_n(u) := \frac{\gamma_n}{2} ||(u - b)_+||^2_{L^2}, \ \forall u \in L^2(\Omega).$$

Then $F_n \xrightarrow{\mathrm{M-epi}} I_{U_{ad}}$, where $I_{U_{ad}}$ stands for the indicator function of the set U_{ad} given by

$$U_{ad} := \left\{ u \in L^2(\Omega) \mid u \leq b \ a.e. \, \Omega \right\}.$$

Proof. We begin by assuming that $u \notin U_{ad}$. For any $u_n \rightharpoonup u$ in $L^2(\Omega)$, we can use the weak lower semicontinuity of $||(\cdot - b)_+||^2_{L^2}$ in order to deduce the existence of some $\varepsilon > 0$ such that

$$\liminf_{n \to \infty} ||(u_n - b)_+||^2_{L^2} \geq \varepsilon > 0.$$

It follows that $\liminf_{n\to\infty} F_n(u_n) = +\infty$. Conversely, suppose that $u \in U_{ad}$, then since the trivial sequence $u_n = u$ converges weakly to u in $L^2(\Omega)$, we have found a sequence such that $\liminf_{n\to\infty} F_n(u_n) = 0$. Therefore, it holds for all $u \in L^2(\Omega)$ that

$$\forall u_n \to_{L^2(\Omega)} u, \ I_{U_{ad}}(u) \leq \liminf_{n \to \infty} F_n(u_n).$$

The remaining argument requires us to demonstrate the existence of a strongly convergent sequence such that the limit superior condition in Definition 4.5 holds for all $u \in L^2(\Omega)$. Of course, if $u \notin U_{ad}$, then $I_{U_{ad}}(u) = +\infty$. Thus for any sequence u_n strongly converging to u in $L^2(\Omega)$, it follows that

$$+\infty = I_{U_{ad}}(u) \geq \limsup_{n \to \infty} F_n(u_n).$$

Finally, if $u \in U_{ad}$, then by taking the trivial sequence $u_n = u$, we see that $F_n(u_n) = 0$ for all n. Hence,

$$0 = I_{U_{ad}}(u) \geq \limsup_{n \to \infty} F_n(u_n),$$

as was to be shown. □

Combining Lemma 4.6 with [1, Theorem 3.66], we arrive at the following result.

Proposition 4.7 (**Convergence of Approximations**). *Let $\gamma_n \to \infty$, $b \in L^2(\Omega)$, and $u_n \to u$ in $L^2(\Omega)$. If $w_n \to w$ in $L^2(\Omega)$ for $w_n := \gamma_n(u_n - b)_+$, then we have $w \in N_{U_{ad}}(u)$.*

Proof. The aforementioned theorem by Attouch states that the Mosco epi-convergence for a sequence of proper, convex, and lower semicontinuous functions is equivalent to the graph convergence of their subdifferentials (plus a normalizing condition). Using F_n and F from Lemma 4.6, we see that

$$\partial F_n(u) = \gamma_n(u-b)_+ \text{ and } \partial F(u) = \partial I_{U_{ad}} = N_{U_{ad}}(u), \quad u \in L^2(\Omega).$$

Then by [1, Proposition 3.59] (see above), the assertion holds. □

We are now ready to derive the main result of this section.

Theorem 4.8 (Improved Limiting Stationarity Conditions for the Constrained MPEC). *Let $\gamma_n \to \infty$, and let $(\bar{u}, \bar{y}) \in L^2(\Omega) \times H_0^1(\Omega)$ be an optimal solution to (44). Then there exist sequences*

$$\bar{u}_n \to_{L^2} \bar{u}, \quad \bar{y}_n \to_{H_0^1} \bar{y}, \quad \bar{p}_n \rightharpoonup_{H_0^1} \bar{p}, \quad \bar{r}_n \rightharpoonup_{H^{-1}} \bar{r}, \tag{57}$$

where $(\bar{u}_n, \bar{y}_n) \in L^2(\Omega) \times H_0^1(\Omega)$ solves the penalized MPEC (45) for each $n \in \mathbb{N}$, with $\gamma := \gamma_n$ and

$$(\bar{u}_n, \bar{y}_n, \bar{p}_n, \bar{r}_n) \in L^2(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega) \times H^{-1}(\Omega)$$

satisfying the strong stationarity system (49)–(56). Moreover, the quadruple $(\bar{u}, \bar{y}, \bar{p}, \bar{r})$ satisfies the limiting stationarity conditions (36)–(43) with (40) replaced by

$$0 \geq \langle \bar{r}, \bar{p} \rangle_{H^{-1}, H_0^1}.$$

Proof. Let $(\bar{u}_n, \bar{y}_n)$ be a globally optimal solution of (45) with $\gamma := \gamma_n$. The existence of $\bar{u}_n$ and $\bar{y}_n$ follows from Proposition 4.4. Since each pair is an optimal solution, we have from Theorem 4.3 the existence of $(\bar{p}_n, \bar{r}_n)$ such that the conditions (49)-(56) hold.

Using now the properties of $\bar{p}_n$ and $\bar{r}_n$, we have from (50), after multiplying with $\bar{p}_n$, the following equation

$$\langle A^* \bar{p}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1} = (\bar{y}_n - y_d, \bar{p}_n)_{L^2} + \langle \bar{r}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1}.$$

Using the coercivity of A and the fact that $\langle \bar{r}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1} \leq 0$, we know there exists a $\xi > 0$ such that

$$\xi ||\bar{p}_n||^2_{H_0^1} \leq (\bar{y}_n - y_d, \bar{p}_n)_{L^2} \leq ||\bar{y}_n - y_d||_{L^2} ||\bar{p}_n||_{L^2}.$$

Then by dividing through by $||\bar{p}_n||_{L^2}$ and using the fact that $H_0^1(\Omega)$ is continuously embedded into $L^2(\Omega)$, we derive the existence of some $\kappa > 0$ such that

$$||\bar{p}_n||_{H_0^1} \leq \kappa ||\bar{y}_n - y_d||_{L^2}.$$

It follows that $\{\bar{p}_n\}$ is bounded in $H_0^1(\Omega)$. Therefore, there exists $\bar{p} \in H_0^1(\Omega)$ and a subsequence $\{\bar{p}_{n_l}\}$ such that $\bar{p}_{n_l} \rightharpoonup_{H_0^1} \bar{p}$. Moreover, we can use this sequence along with (50) to conclude the existence of a sequence $\{\bar{r}_n\}$ in $H^{-1}(\Omega)$ which converges weakly in $H^{-1}(\Omega)$ to some $\bar{r} \in H^{-1}(\Omega)$. Thus, the sequences $(\bar{u}_n, \bar{y}_n, \bar{p}_n, \bar{r}_n)$ satisfy the same requirements as those arising from the definition of the limiting coderivative in Proposition 3.8.

Next, since for all n

$$-\bar{p}_n - \alpha \bar{u}_n = \gamma_n(\bar{u}_n - b)_+,$$

where $p_n \rightharpoonup p$ in $H_0^1(\Omega)$, therefore strongly in $L^2(\Omega)$, and $u_n \to u$ in $L^2(\Omega)$, we can apply Proposition 4.7 in order to deduce the limiting condition

$$0 \in p + \alpha u + N_{U_{ad}}(u).$$

Hence, $(\bar{u}, \bar{y}, \bar{p}, \bar{r})$ fulfills the relations (36)-(42) via the same results which were used to prove Theorem 3.13.

Finally, since $\langle \bar{r}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1} \leq 0$ for all $n \geq 1$, with $\bar{r}_n := A^* \bar{p}_n + y_d - \bar{y}_n$, we obtain

$$\begin{aligned} 0 \geq \langle A^* \bar{p}_n + y_d - \bar{y}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1} = \langle A^* \bar{p}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1} + (y_d - \bar{y}_n, \bar{p}_n)_{L^2} \geq \\ \liminf_{n \to \infty} \langle A^* \bar{p}_n, \bar{p}_n \rangle_{H^{-1}, H_0^1} + (y_d - \bar{y}_n, \bar{p}_n)_{L^2} \geq \\ \langle A^* \bar{p}, \bar{p} \rangle_{H^{-1}, H_0^1} + (y_d - \bar{y}, \bar{p})_{L^2} = \langle \bar{r}, \bar{p} \rangle_{H^{-1}, H_0^1}. \end{aligned}$$

□

Remark 4.9 (**Obtaining Strong Stationarity**). Since $p_k \geq 0$, *a.e.* $\mathcal{A}(y_k)$ and $p_k \to \bar{p}$ in $L^2(\Omega)$, it would be easy to obtain a full sign condition for $\bar{p}$ on $\mathcal{A}(\bar{y})$ provided the active sets $\mathcal{A}(y_k)$ were to converge in the sense of characteristic functions to $\mathcal{A}(\bar{y})$. This idea was presented in [7] in a similar setting, but without upper-level constraints on the control. Nevertheless, this does not guarantee that $\bar{r}$ has the needed sign condition for strong stationarity as described in Definition 1.1. In finite dimensions or if $\bar{r}$ were more regular, e.g., in $L^2(\Omega)$, one could then argue that $\bar{r} = 0$ pointwise almost everywhere on $\mathcal{I}(\bar{y})$. Moreover, the inequality $\bar{r}\bar{p} \leq 0$ along with the non-negativity of $\bar{p}$ on the active set, would imply that $\bar{r} \leq 0$. Hence, in certain regular cases, it would be possible to derive strong stationarity conditions from our limiting stationarity conditions.

Finally, we note that Theorem 4.8 is interesting ultimately for the fact that by using a simple smooth penalization of the control constraints, we are able to apply the technique developed in [9] in order to guarantee that each solution of the penalized problems satisfies strong stationarity conditions. In addition, when passing to the limit, we obtain a better system than the stationarity conditions found in Theorem 3.13. Thus, we arrive at a stationarity system better than C-stationarity but weaker than strong stationarity.

5 Stationarity Conditions for Constrained MPECs via Regularization-Penalization Techniques

In this section we explore yet another approximation approach to study the class of our constrained elliptic MPECs. Such a penalization-approximation technique has been recently applied to MPECs by Hintermüller and Kopacka in [8] while it has been widely employed before in various frameworks of single-level optimal control and related problems governed by partial differential equations; see, e.g., the books by Barbu [3, Chapter 3.2] and Mordukhovich [16, Chapter 7.4] with the bibliographies therein. Note also that the concept of penalizing the nonsmoothness/multivaluedness via a sequence of parameter-dependent differentiable functions goes back to earlier developments presented in [12] and [6]. Our notation and terminology are based on [8].

For simplicity we consider in this section the class of MPECs (44) described at the beginning of Section 4 with imposing two additional assumptions. The first one changes the variational inequality used to define (1) by adding an $L^2(\Omega)$-function f to the right-hand side. Thus, in what follows S denotes the solution map of the variational inequality (lower-level problem)

$$Ay + N_M(y) \ni u + f \tag{58}$$

in the modified MPEC. As noted in [8], under the assumption that the obstacle $\psi \in H^2(\Omega)$ with $\psi|_{\partial\Omega} \leq 0$, we can transform the lower-level problem such that $y := y - \psi$ and $f := f + A\psi$. For this reason, we henceforth use the constraint $y \geq 0$, *a.e.* Ω in place of $y \geq \psi$, *a.e.* Ω. In order to

easily state the results from [8], we make this assumption here. Second, we assume that A is the second-order differential operator associated with the bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ defined by

$$a(v,w) = \sum_{i,j=1}^{l} \int_\Omega a_{ij} \frac{\partial v}{\partial x_j} \frac{\partial w}{\partial x_i} dx + \sum_{i=1}^{l} \int_\Omega b_i \frac{\partial v}{\partial x_i} w dx + \int_\Omega cvw dx, \ \forall v, w \in H_0^1(\Omega), \tag{59}$$

where a_{ij}, b_i, c are in $L^\infty(\Omega)$. Assume, in addition, that $a_{ij} \in C^{0,1}(\bar{\Omega})$, i.e., Lipschitz continuous on the closure of Ω, $c \geq 0$, and that $a(\cdot,\cdot)$ is both bounded and coercive.

The new data assumptions given above, we now proceed with the description of the technique. Suppose that $\pi : H_0^1(\Omega) \to H^{-1}(\Omega)$ is Lipschitz continuous and monotone with the condition $\ker(\pi) = M$. Then the variational inequality (58) can be approximated by a quasi-linear second-order partial differential equation written here in the form

$$a(y,\varphi) + \frac{1}{\beta} \langle \pi(y), \varphi \rangle_{H^{-1}, H_0^1} = (u,\varphi)_{L^2} + (f,\varphi)_{L_2}, \ \forall \varphi \in H_0^1(\Omega),$$

where $\beta > 0$ is a penalty parameter. The assumptions imposed on the penalty operator π ensure that the above partial differential equation (PDE) has a unique solution $y_\beta(u)$. Moreover, it can be shown that $y_\beta(u) \to y(u)$ in $H_0^1(\Omega)$ as $\beta \to 0^+$, where $y(u)$ solves the original variational inequality (58); see e.g., [6, 8]. Note that in [8] the mapping π was defined by using the maximum operator

$$\pi(v) := -\max(0, -v), \ \forall v \in H_0^1(\Omega).$$

Since the pointwise maximum $\max(0,\cdot)$ is nondifferentiable, certain regularized (i.e., smoothed) operators dependent on some parameter $\varepsilon > 0$ were considered in [8]. These smoothed operators, which we denote now by $\max_\varepsilon(0,\cdot)$, act almost identically to the $\max(0,\cdot)$ operators with the only difference that the "kink" at zero is smoothed out on a neighborhood depending on ε. One such example is given explicitly by

$$\max_\varepsilon(0,r) := \begin{cases} r - \dfrac{\varepsilon}{2} & \text{if } \ r \geq \varepsilon, \\ \dfrac{r^2}{2\varepsilon} & \text{if } \ r \in (0,\varepsilon), \\ 0 & \text{if } \ r \leq 0. \end{cases}$$

Under relatively weak assumptions it is shown in [8, Theorem 2.3] that solutions y_β to the regularized penalized problems

$$Ay_\beta - \frac{1}{\beta} \max_\varepsilon(0, -y_\beta) = u_\beta + f,$$

with $u_\beta, u \in L^2(\Omega)$ and $u_\beta \to u$ in $H^{-1}(\Omega)$, converge strongly in $H_0^1(\Omega)$ as $\beta \to 0^+$ to the solution $y(u)$ of the original variational inequality (58). By using the penalized regularized variational inequality, i.e., the quasi-linear partial differential equation, we define the following smoothed penalized problem that approximates MPEC (44) under consideration:

$$\begin{array}{ll} \min \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)} \text{ over } (u,y) \in L^2(\Omega) \times H_0^1(\Omega) \\ s.t. \quad a \leq u \leq b \ a.e. \, \Omega, \\ \qquad Ay - \frac{1}{\beta} \max_\varepsilon(0, -y) = u + f. \end{array} \tag{60}$$

Since (60) is no longer an MPEC, more classical methods for the derivation of optimality conditions can be applied. The process is roughly as follows: the regularization of the non-smoothness can be used to show that the solution mapping S of the PDE is Fréchet differentiable for each $\varepsilon > 0$.

After rewriting the problem in terms of the control u, one can then characterize the solutions via a variational inequality, which after introducing the proper slack variables, leads to the following result.

Theorem 5.1 (**Necessary Optimality Conditions for the Penalized-Regularized Problems**). *Let $\beta, \varepsilon > 0$ and $(y, u) \in H_0^1(\Omega) \times L^2(\Omega)$ be an optimal solution to* (60). *Then there exists an adjoint state $p \in H_0^1(\Omega)$ such that*

$$y + A^* p + \frac{1}{\beta} \max_\varepsilon'(0, -y)p = y_d, \tag{61}$$

$$Ay - \frac{1}{\beta} \max_\varepsilon(0, -y) = u + f, \tag{62}$$

$$u \in U_{ad},\ (\alpha u - p, v - u)_{L^2} \geq 0,\ \forall v \in U_{ad}. \tag{63}$$

By defining sequences of *stationary points*, rather than local or global minimizers in the primal variables, satisfying (61)–(63) along a sequence of positive numbers $\beta \to 0^+$ and a bounded sequence $\{\varepsilon(\beta)\}$ with $\varepsilon(\beta)/\beta \to 0$ as $\beta \to 0^+$, it is shown in [8, Theorem 3.4] that there exists a quintuple

$$(\tilde{u}, \tilde{y}, \tilde{v}, \tilde{p}, \tilde{r}) \in L^2(\Omega) \times H_0^1(\Omega) \times L^2(\Omega) \times H_0^1(\Omega) \times H^{-1}(\Omega)$$

and a subsequence of the stationary points, which we again denote by β, such that

$$\begin{aligned}
u_\beta &\to \tilde{u}, \quad \text{in } L^2(\Omega), \\
y_\beta &\to \tilde{y}, \quad \text{in } H_0^1(\Omega), \\
\frac{1}{\beta} \max_{\varepsilon(\beta)}(0, -y_\beta) &\to \tilde{v}, \quad \text{in } H^{-1}(\Omega), \\
p_\beta &\rightharpoonup \tilde{p}, \quad \text{in } H_0^1(\Omega) \\
\frac{1}{\beta} \max_{\varepsilon(\beta)}'(0, -y_\beta)p_\beta &\rightharpoonup \tilde{r}, \quad \text{in } H^{-1}(\Omega),
\end{aligned}$$

where $(\tilde{u}, \tilde{y})$ is ε-almost C-stationary for the original MPEC; see Definitions 1.1 and 1.2.

In fact, the multiplier $\tilde{r}$ can be shown to be in $(L^\infty(\Omega))^*$, which is more regular in the sense that each $\tilde{r}$ is then a finitely additive, finite signed-measure, see e.g., [5, Theorem IV.8.16]. For simplicity we confine ourselves to the case when the coefficient functions b_j in the bilinear form (59) equal to zero.

Proposition 5.2 (**Increased Regularity of the Multiplier $\tilde{r}$**). *The limiting multiplier $\tilde{r}$ in Theorem* 5.1 *is an element of $H^{-1}(\Omega) \cap (L^\infty(\Omega))^*$.*

Proof. In principle, this follows from Theorem 3.3 [3], however in the interest of completeness, we include a short proof here. We begin by letting $\operatorname{sign}(\cdot)$ represent the pointwise sign function and suppose that $\sigma(\cdot)$ is a monotonic smoothing of $\operatorname{sign}(\cdot)$, which has the property

$$\sigma(x) < 0,\ \text{if } x < 0,\ \sigma(0) = 0,\ \sigma(x) > 0,\ \text{if } x > 0.$$

For an arbitrarily fixed number $\beta > 0$, multiply equation (61) above by $\sigma(p_\beta)$ and obtain the equality

$$\langle A^* p_\beta, \sigma(p_\beta) \rangle_{H^{-1}, H_0^1} + \Big(\frac{1}{\beta} \max_{\varepsilon(\beta)}'(0, -y_\beta)p_\beta, \sigma(p_\beta) \Big)_{L^2} = \big(y_d - y_\beta, \sigma(p_\beta) \big)_{L^2}.$$

To see that the first term of the latter equation is always nonnegative, we refer back to the definition of the bilinear form $a(\cdot,\cdot)$ in (59). This gives

$$\langle A^* p_\beta, \sigma(p_\beta)\rangle_{H^{-1},H_0^1} = \int_\Omega \sigma'(p_\beta) \sum_{i,j} a_{ij} |\nabla p_\beta|^2 \, dx + \int_\Omega c p_\beta \sigma(p_\beta) \, dx. \tag{64}$$

The assumptions imposed above ensure that $cp_\beta\sigma(p_\beta) \geq 0$ almost everywhere on Ω. Furthermore, observe that the first term on the right-hand side of equation (64) is positive since the derivative of σ is either zero or positive and since the operator A is coercive. It follows from the convergence results of Theorem 5.1 that there exists a constant $\kappa > 0$ such that

$$0 \leq \Big(\frac{1}{\beta}\max\nolimits'_{\varepsilon(\beta)}(0, -y_\beta)p_\beta, \sigma(p_\beta)\Big)_{L^2} \leq \big(y_d - y_\beta, \sigma(p_\beta)\big)_{L^2} \leq \kappa,$$

Given the positivity of the integrand $\frac{1}{\beta}\max'_{\varepsilon(\beta)}(0, -y_\beta)p_\beta\sigma(p_\beta)$, it follows that

$$\int_\Omega |\frac{1}{\beta}\max\nolimits'_{\varepsilon(\beta)}(0, -y_\beta)p_\beta\sigma(p_\beta)|dx \leq \kappa, \forall \beta > 0.$$

Then by letting $\sigma \to \text{sign}(\cdot)$, we can argue that $\frac{1}{\beta}\max'_{\varepsilon(\beta)}(0, -y_\beta)p_\beta$ is bounded in $L^1(\Omega)$. In which case, we deduce the existence of a subsequence, still denoted by β, and an element $r^* \in (L^\infty(\Omega))^*$, such that $\frac{1}{\beta}\max'_{\varepsilon(\beta)}(0, -y_\beta)p_\beta \overset{*}{\rightharpoonup} r^*$ in $(L^\infty(\Omega))^*$. It follows that $r^* = \tilde{r} \in (L^\infty(\Omega))^*$. □

6 Conclusions and Comparisons

In this paper, we considered four possibilities for the derivation of dual optimality/stationarity conditions of the MPEC (1), namely, via

1. Dualization of a primal optimality condition derived using directional derivatives,
2. Limiting variational calculus,
3. Penalization of the upper-level control constraints,
4. Penalization and regularization of the non-smooth and multivalued term.

From a theoretical standpoint, we saw that the dualization of B-stationarity conditions seems extremely difficult to directly characterize in order to obtain a multiplier-based, dual optimality condition.

Then exploiting our knowledge of the directional derivative of the solution map to the variational inequality in the MPEC under consideration, we were able to work with the limiting variational objects developed and popularized by the second author. This led us to a new system of limiting optimality conditions similar to M-stationary conditions. Due to a lack of regularity of certain sequences, we saw that the limiting conditions obtained are different from C-stationarity. On one hand, nothing can be said about the product of the adjoint states $\bar{p}$ and $\bar{r}$, the multipliers associated with the lower-level multiplier $\bar{v}$, unless $\bar{p}$ has a constant sign on the active set $\mathcal{A}(\bar{y})$. On the other hand, the multiplier $\bar{r}$ associated with the complementarity constraint has an increased level of information in comparison with C-stationarity.

Further, by employing a smooth Moreau-Yosida-type penalization of the upper-level constraints, we reduced the original constrained MPEC setting to that in which the results from [9] became directly applicable. These results guarantee that every solution to the penalized problem is strongly

stationary. We demonstrated in this way that a sequence of optimal solutions to the penalized MPECs satisfies a limiting stationarity system stronger than the one obtained using the limiting variational calculus. The relations of this system are more selective than those of C-stationarity but weaker than strong stationarity.

Lastly, we applied a more classical penalization-regularization technique examined recently in [8] in some MPEC settings to derive ε-almost C-stationary conditions for the class of constrained MPECs under consideration in this paper.

In terms of the usefulness of the results for the development of numerical methods, the penalization/regularization technique has the clear advantage. Indeed, in this setting the practitioner is required to solve a sequence KKT systems arising from smooth non-linear programs (NLPs). Moreover, the limit of subsequences of solutions to the NLPs is guaranteed under weak assumptions to satisfy a type of stationarity conditions weaker than C-stationarity, yet stronger than so-called weak stationarity [19].

The development of a numerical method from the derivation technique described in Section 4 is somewhat more difficult. In contrast to the previously discussed method, in which one speaks of the convergence of stationary points, this method requires knowledge of optimal solutions for each of the penalized MPECs. However, provided with this information, one is guaranteed that each member of the sequence is strongly stationary and that the limit of subsequences of these solutions will satisfy the limiting stationarity system. The development of numerical methods realizing strong stationary points is one possible future direction.

Finally, though they provide us with a significant amount of insight in terms of the limits of solutions satisfying strong stationarity conditions, the limiting calculus appears to impose certain restrictions on the ability to construct numerical methods in function spaces. This relates mainly to the fact that the existence of the sequences in Proposition 3.8 is guaranteed by the *definition* of the limiting coderivative while the sequences, along with their characteristics, in Theorem 4.8 had to be derived. Moreover, the relationship in (40) is clearly difficult to handle. If such a method were available, then the minimal requirements placed on the operator B would allow the practitioner to consider examples in which the control perturbation of the variational inequality is not distributed on the entire domain Ω.

Acknowledgements

The first and third authors would like to acknowledge the financial support from the DFG Research Center MATHEON Project C28, the SPP 1253 "Optimization with Partial Differential Equations", and the START Project Y 305 "Interfaces and Free Boundaries" funded by the Austrian Ministry of Science and Education and administered by the Austrian Science Fund FWF. The second author would especially like to thank the USA National Science Foundation for their support under grant DMS-1007132, the Australian Research Council under grant DP-12092508, and the Portuguese Foundation of Science and Technologies under grant MAT/11109.

References

[1] Attouch, H. *Variational Convergence for Functions and Operators*. Pitman Advanced Publishing Program, Boston, London, Melbourne, 1984.

[2] Aubin, J.-P., and Frankowska, H. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.

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

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

[5] Dunford, N., and Schwartz, J. T. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988.

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

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

[8] Hintermüller, M., and Kopacka, I. A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs. *Computational Optimization and Applications* (2009). DOI 10.1007/s10589-009-9307-9.

[9] Hintermüller, M., and Surowiec, T. First order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis. DFG Research Center MATHEON Preprint, Available at: http://www.matheon.de/research/list_preprints.asp, July 2010.

[10] Jeyakumar, V., and Wolkowicz, H. Generalizations of slater's constraint qualification for infinite convex programs. *Mathematical Programming 57* (1992), 85–101.

[11] Jongen, H. T., Rückmann, J.-J., and Shikhman, V. MPCC: critical point theory. *SIAM J. Optim. 20*, 1 (2009), 473–484.

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

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

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

[15] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory.* Springer-Verlag, Berlin, 2006.

[16] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. Vol. 2: Applications.* Springer-Verlag, Berlin, 2006.

[17] Outrata, J. V., and Jařusek, J. On sharp necessary optimality conditions in control of contact problems with strings. *Nonlinear Analysis 67*, 4 (2007), 1117–1128.

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

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