SIAM J. OPTIM.
Vol. 21, No. 4, pp. 1561–1593

FIRST-ORDER OPTIMALITY CONDITIONS FOR ELLIPTIC MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS VIA VARIATIONAL ANALYSIS*

M. HINTERMÜLLER† AND T. SUROWIEC‡

Abstract. Mathematical programs in which the constraint set is partially defined by the solutions of an elliptic variational inequality, so-called "elliptic MPECs," are formulated in reflexive Banach spaces. With the goal of deriving explicit first-order optimality conditions amenable to the development of numerical procedures, variational analytic concepts are both applied and further developed. The paper is split into two main parts. The first part concerns the derivation of conditions in which the (lower-level) state constraints are assumed to be polyhedric sets. This part is then completed by two examples, the latter of which involves pointwise bilateral bounds on the gradient of the state. The second part focuses on an important nonpolyhedric example, namely, when the lower-level state constraints are presented by pointwise bounds on the Euclidean norm of the gradient of the state. A formula for the second-order (Mosco) epiderivative of the indicator function for this convex set is derived. This result is then used to demonstrate the (Hadamard) directional differentiability of the solution mapping of the variational inequality, which then leads to the derivation of explicit strong stationarity conditions for this problem.

Key words. elliptic MPEC, strong stationarity, generalized differentiation, pointwise gradient constraints, epiconvergence, epiderivative

AMS subject classifications. 49J53, 49K20, 65K10, 90C33

DOI. 10.1137/100802396

1. Introduction.

The mathematical modeling of real-world phenomena often leads to infinite dimensional, i.e., function space, problem formulations containing variational inequalities. For example, certain problems in elasticity [27], elastoplasticity [17, 26], and mathematical finance [1] all lead to models in which a variational inequality arises. In addition, minimization problems involving certain classes of nonsmooth functionals result in variational inequalities via Fenchel–Legendre dualization and associated Euler–Lagrange conditions [14]. Due to their practical relevance, many research efforts have been devoted to the study of variational inequalities and their numerical solution since their conception; see, e.g., [15, 16, 28] and the references therein.

Frequently one is interested in controlling the solution of a variational inequality in order to achieve a desired state or to minimize a target quantity. On an abstract level this leads to minimization problems of the type

$$\begin{aligned} &\text{minimize} && J(u,y) \quad \text{over } (u,y) \in \mathcal{U} \times Y \\ &\text{subject to (s.t.)} && y \in S(u), \end{aligned} \tag{1}$$

*Received by the editors July 16, 2010; accepted for publication (in revised form) September 6, 2011; published electronically December 15, 2011. This work was supported by the DFG Research Center MATHEON Project C28, the SPP 1253 "Optimization with Partial Differential Equations," and the START Project Y 305-N18 "Interfaces and Free Boundaries" funded by the Austrian Ministry of Science and Education and administered by the Austrian Science Fund FWF.

http://www.siam.org/journals/siopt/21-4/80239.html

†Department of Mathematics, Humboldt University of Berlin, Berlin, Germany, and Department of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, 8010 Graz, Austria (hint@math.hu-berlin.de).

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

where (u, y) denotes the associated control-state pair with respective control space $\mathcal{U}$ and state space Y, J is a (usually Fréchet differentiable) objective function from $\mathcal{U} \times Y \to \mathbb{R}$, and $S : \mathcal{U} \to Y$ (or $S : Y^* \to Y$) represents the solution operator of the underlying variational inequality. Problems of type (1) are sometimes called mathematical programs with equilibrium constraints (MPECs), as the variational inequality often represents an equilibrium condition, e.g., first-order optimality conditions of a convex optimization problem. Although the literature on finite dimensional MPECs has reached a certain level of sophistication, as evidenced by the monographs [32, 36, 38] and the many references therein, far less is known about MPECs in function spaces. With respect to the latter, we mention [6, 37] and the selected papers [7, 8, 9, 21, 25, 34]. We would also like to mention that parameter identification problems for variational inequalities lead to problems of type (1); see, e.g., [19, 20] and the references therein.

From a mathematical optimization point of view, the difficulties associated with (1) result from a lack of constraint regularity, which in turn prevents the application of well-known results for mathematical programs in Banach space; see, e.g., [47]. Moreover, upon reducing (1) to a problem in u by considering $y = y(u) = S(u)$, following the so-called implicit programming approach, the problem typically becomes a nonsmooth and nonconvex problem, which is then hard to tackle analytically as well as numerically. In particular, the explicit representation of first-order optimality conditions suitable for numerical realization remains an issue.

In a recent work motivated by similar results in finite dimensions as found in [44], an attempt at systematizing stationarity conditions for function-space–based problems of the type (1) was undertaken in [21]. Remarkably, versions of weak, C, and strong stationarity were derived that paralleled the concepts in finite dimensions, and whereas many approaches applied in the past relied on penalization techniques, the method of [21] utilizes a relaxation approach yielding stronger stationarity conditions than those resulting from penalty techniques. Penalization or relaxation techniques have the advantage that they readily facilitate the application of the well-established theory on mathematical programs in Banach space for the existence of Lagrange multipliers. Moreover, these techniques may be turned into algorithmic frameworks by closely following the derivation of first-order optimality systems for the MPEC. Using these techniques, as was demonstrated in [21, 22], makes the problems amenable to the application of fast solvers such as semismooth Newton and multigrid methods.

Despite the appeal of penalization and regularization methods, variational analysis (see, e.g., [5, 43, 35]) provides a different set of analytical tools able to directly derive sharp, i.e., strong, stationarity conditions without needing to pass to the limit with respect to certain parameters that arise in the relaxation/penalty approaches. Oftentimes one need only verify that the data of the lower-level problem satisfy certain constraint qualifications in order to derive first-order optimality conditions, thereby providing a means for avoiding relaxation/penalization techniques and limit processes in a problem-dependent fashion.

In this respect, the aim of this work is twofold: (i) We utilize and extend tools from variational analysis in order to derive an abstract first-order optimality system (in the sense of strong stationarity) for a rather broad class of control problems of elliptic variational inequalities. (ii) We treat systems involving (pointwise) gradient constraints of the type

$$M := \left\{ y \in H_0^1(\Omega) \,|\, |\nabla y| \leq \psi \text{ almost everywhere (a.e.) on } \Omega \right\}.$$

Gradient constraints in function-space MPECs have yet to be treated in the literature, despite having been mentioned as an important (open) problem class in [34]. For control of obstacle-type problems (i.e., with pointwise unilateral constraints on the state, rather than its gradient) it turns out that we recover the strong stationarity conditions derived earlier by Mignot and Puel in [34], who applied a mix of penalization techniques and so-called "conical derivatives."

The rest of the paper is organized as follows. In section 2, we introduce notation and basic concepts. The remaining sections subdivide our investigation into a polyhedric setting (sections 4–5) and an important nonpolyhedric example arising in the theory of elastoplasticity utilizing gradient constraints (section 6). Section 4 is devoted to studying differentiability properties of the control-to-state mapping, i.e., the solution operator of the underlying variational inequality. Our results extend those obtained by Mignot in his fundamental paper [33] on the conical derivative of the solution operator associated with the obstacle problem. In section 3, these results are applied to derive strong stationarity conditions for the MPEC, and in section 5 two case studies are performed yielding first the well-known stationarity result of Mignot and Puel [34] (here in the sense of a validation of our technique) and then explicit strong stationarity conditions in the presence of pointwise constraints on the gradient of the state. The extension to the nonpolyhedric case is more delicate as the variational arguments in the polyhedric case do not immediately apply. Hence, in section 6, we establish a new result for the second-order Mosco epiderivative of the indicator function of the considered closed convex set. This result enables us to derive strong stationarity in this challenging nonpolyhedric case.

2. Notation and basic concepts from variational analysis. Throughout the text we make significant use of certain objects that are more or less standard in the literature. New or lesser known concepts are introduced throughout the text so that they may be better understood in context.

Assumptions and notation. Throughout the entirety of this paper, we will consider only real Banach spaces, and we make the additional assumption that the topologies of some Banach space X along with its topological dual X^* are compatible. If X is, in addition, reflexive, then the strong topologies on both spaces are considered; otherwise we assume X^* is equipped with the weak*-topology, so that the dual of X^* is isometric to X. For more on this subject, the reader is referred to any standard reference on functional analysis, e.g., [46]. We denote the dual pairing between X and X^* by $\langle \cdot, \cdot \rangle_{X^*,X}$ and denote strong convergence in any space, e.g., X, via the symbol "$\to_X$" and weak convergence by "$\rightharpoonup_X$". The embedding of a space X into Y is denoted $X \hookrightarrow Y$. If X is an inner product space, then the inner product will be denoted by $(\cdot, \cdot)_X$ and the norm defining the topology on X is denoted by $|| \cdot ||_X$. We denote the closure in the topology on X by $\mathrm{cl}\{\cdot\}_X$. In all cases, we leave off the subscript "X" if it is clear in context. Finally, if $x, y \in \mathbb{R}^l$, then $x \cdot y$ represents their scalar product, and for any subset $A \subseteq \mathbb{R}^l$, we use "a.e. A" to represent "almost everywhere on A." We use the standard notation $\mathbb{B}_\varepsilon(0)$ for the closed ε ball in some space X and "$o(t)$" to be a function such that $o(t)/t \to 0$ as $t \to 0^+$.

A few important function spaces. At some points in this paper, we provide examples in which certain function spaces are present. We always assume that the subset $\Omega \subseteq \mathbb{R}^l$ is a bounded open subset with Lipschitz boundary $\partial\Omega$ and let $l \geq 1$. We denote the standard Lebesgue space of square integrable functions/vector fields by $L^2(\Omega)^l$, leaving off the "l" subscript if $l = 1$, and we denote the space of all infinitely differentiable functions whose (compact) support is contained in Ω by

$C_0^\infty(\Omega)$. We then define the Sobolev space $H_0^1(\Omega)$ as the completion of the space $C_0^\infty(\Omega)$ with respect to the norm $||x||_{H_0^1(\Omega)} := ||\nabla x||_{L^2(\Omega)^l}$. The usage of this norm for defining $H_0^1(\Omega)$ follows from the boundedness of Ω. Here, the gradient of x is understood in a weak sense. Finally, we denote the *pointwise* ∞-norm and 2-norm of the gradients of H_0^1-functions for any $x \in \Omega$ by $|\nabla y(x)|_\infty := \max_{1\le i\le l} |\left(\nabla y(x)\right)_i|$ and $|\nabla y(x)|_2 := (\sum_{i=1}^l (\nabla y(x))_i^2)^{1/2}$, respectively. When it is clear in context, we leave off the arguments "x". For more on Sobolev spaces we refer the reader to [2].

Variational analytic concepts. Throughout this subsection we assume that X is some arbitrary Banach space paired with its dual X^* and let $C \subseteq X$ be a nonempty closed convex set. In addition, we define the indicator function of C by $I_C(x) := 0$ if $y \in C$ and set $I_C(x) := \infty$ otherwise.

The *tangent cone* to C at $x \in C$ is defined by

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

In the event that C is only closed but not convex, we refer to this cone as the *contingent cone*. As we will see in a moment, the tangent cone can be derived via calculating the polar cone to another variational object. This, however, is not true when C is not convex.

Given another arbitrary Banach space Y, we refer to any mapping F from X into the set of subsets of Y as a *multifunction* or *set-valued mapping*. We use the notation $F : X \rightrightarrows Y$ to denote that F is a multifunction. The *graph* of a multifunction is defined by $\operatorname{gph} F := \{(x,y) \in X \times Y \,| y \in F(x)\}$. Clearly, $\operatorname{gph} F \subset X \times Y$.

Though multifunctions are very different from single-valued mappings, we can still define (generalized) derivatives using contingent cones. Accordingly, we define the *contingent derivative* of F at a point $(x,y) \in \operatorname{gph} F$ to be the mapping $DF[(x,y)] : X \rightrightarrows Y$ whose graph equals $T_{\operatorname{gph} F}(x,y)$, i.e.,

$$w \in DF[(x,y)](u) \Leftrightarrow (u,w) \in T_{\operatorname{gph} F}(x,y) \Leftrightarrow \left\{ \begin{array}{l} \exists t_k \to 0^+, \exists u_k \to_X u, \exists w_k \to_Y w : \\ y + t_k w_k \in F(x + t_k u_k). \end{array} \right.$$

Note that the third implication is precisely the definition of the Painlevé–Kuratowski upper limit of the difference quotient $t^{-1}(\operatorname{gph} F - (x,y))$. For more on these and related concepts, see, e.g., [5].

We can define an even stronger concept of generalized derivative, known as *proto-derivatives*, in a similar way. We define the *proto-derivative* of F at a point $(x,y) \in \operatorname{gph} F$ to be the mapping $PF[(x,y)] : X \rightrightarrows Y$ whose graph satisfies the following condition:

$$\operatorname{gph} PF[(x,y)] = \operatorname{Lim\,sup}_{t\to 0^+} \frac{\operatorname{gph} F - (x,y)}{t} = \operatorname{Lim\,inf}_{t\to 0^+} \frac{\operatorname{gph} F - (x,y)}{t},$$

where "Lim sup" and "Lim inf" are the Painlevé–Kuratowski upper and lower limits, respectively. Therefore, if $PF[(x,y)]$ exists, then $PF[(x,y)] = DF[(x,y)]$, but not necessarily vice versa. For more on this topic, see [31, 42].

By definition, $(w,d) \in \operatorname{Lim\,inf}_{t\to 0^+} t^{-1}(\operatorname{gph} F - (x,y))$ implies that for all sequences $t_k \to 0^+$, $w_k \to_X w$, $d_k \to_Y d$, $(w_k, d_k) \in t_k^{-1}(\operatorname{gph} F - (x,y))$. Therefore, we can argue that if F is single-valued, then

$$PF[(x,y)](w) = \lim_{\substack{t\to 0^+ \\ w' \to_X w}} t^{-1}(F(x + tw') - y).$$

If $PF[(x,y)](w)$ exists for all $w \in X$, then it coincides with the classical *Hadamard directional derivative*. We will generally denote the directional derivative by $F'(x)(w)$. The Hadamard directional derivative has also appeared in the literature under the names *B-derivative* by Robinson in [40] as well as *semiderivative* by Levy and Rockafellar in [31]. Many of the fundamental results on Hadamard directional derivatives are compiled in [10].

Another important object in our study is the so-called *normal cone*. The normal cone to a closed convex set $C \subseteq X$ at some point $x \in C$ is defined by

$$N_C(x) := \{x^* \in X^* \,|\langle x^*, x' - x\rangle_{X^*,X} \leq 0 \;\forall x' \in C\}\,.$$

In addition, $N_C(x)$ can also be (equivalently) defined by

$$N_C(x) := \{x^* \in X^* \,|\langle x^*, d\rangle_{X^*,X} \leq 0 \;\forall d \in T_C(x)\} = [T_C(x)]^-\,,$$

i.e., as the negative polar/dual cone of the tangent cone. Given that X and X^* are paired spaces, it also holds that $[N_C(x)]^- = [[T_C(x)]^-]^- = T_C(x)$, as T_C itself is a closed convex cone.

When C is merely closed and not convex, we can define another type of normal cone known as the *Fréchet normal cone*

$$\widehat{N}_C(x) := \left\{x^* \in X^* \left| \limsup_{x' \to x} \frac{\langle x^*, x' - x\rangle}{||x' - x||_X} \leq 0,\; x' \in C \right.\right\}.$$

Note that the lim sup in the previous definition must hold for *all* sequences $x' \to_X x$ such that $x' \in C$. This often makes the Fréchet normal cone extremely difficult to calculate explicitly. Nevertheless, Theorem 1.10 in [35] provides the following upper approximation:

$$\widehat{N}_C(y) \subset [T_C(y)]^-\,,$$

where equality holds if X is reflexive and one considers weak limits in the definition of T_C or X is finite dimensional.

Finally, we define the *subdifferential* of a convex lower-semicontinuous function $f : X \to \bar{\mathbb{R}}$ at $x \in \operatorname{dom} f$ by

$$\partial f(x) := \{x^* \in X^* \,|\langle x^*, x' - x\rangle + f(x) \leq f(x') \;\forall x' \in X\}\,.$$

Note that for any nonempty closed convex set $C \subseteq X$, $\partial I_C(x) = N_C(x)$.

3. Existence of a solution and a primal optimality condition. In this short section, we provide a result in which the existence of a solution of a certain class of MPECs is proven along with an important primal optimality condition. Though the result is not entirely surprising, it motivates our interest in the generalized differentiability concepts used in the following sections.

THEOREM 3.1. *Let Y and $\mathcal{U}$ be Hilbert spaces such that the embeddings $Y \hookrightarrow \mathcal{U} \hookrightarrow Y^*$ represent a Gelfand triple. Suppose that $S : Y^* \to Y$ is the Lipschitz continuous solution mapping defined by*

$$S(u) := \{y \in Y \,|u \in Ay + N_M(y)\}\,, \quad u \in Y^*.$$

Here, A is a bounded linear elliptic operator between Y and Y^*, i.e., $A \in \mathcal{L}(Y, Y^*)$, and there exists a $\xi \in \mathbb{R}_+ \setminus \{0\}$ such that

$$\langle Ay, y\rangle_{Y^*,Y} \geq \xi ||y||_Y^2 \;\forall y \in Y,$$

and M is a nonempty closed convex subset of Y. If

1. *$J(\cdot, y)$ is weakly lower-semicontinuous for all $y \in Y$;*
2. *$J(u, \cdot)$ is lower-semicontinuous for all $u \in \mathcal{U}$;*
3. *J is bounded from below;*
4. *$J(\cdot, y)$ is coercive for all $y \in Y$;*
5. *the embedding $\mathcal{U} \hookrightarrow Y^*$ is compact;*

then there exists a solution $(\bar{u}, \bar{y})$ to the following MPEC:

$$\begin{cases} \min J(u, y) \\ s.t.\ u \in \mathcal{U},\ \ y \in S(u). \end{cases} \tag{2}$$

In addition, if $J : \mathcal{U} \times Y \to \mathbb{R}$ is Fréchet differentiable and S is Hadamard directionally differentiable from Y^* to Y, which in particular implies that S admits a first-order expansion of the type

$$\forall w \in Y^*,\ S(\bar{u} + tw) = S(\bar{u}) + tS'(\bar{u})(w) + o(t), \quad t \geq 0,$$

then the following optimality condition holds:

$$\langle \nabla_u J(\bar{u}, \bar{y}), w\rangle_{Y,Y^*} + \langle \nabla_y J(\bar{u}, \bar{y}), d\rangle_{Y^*,Y} \leq 0\ \forall (w, d) \in T_{\mathrm{gph}\, S}(\bar{u}, \bar{y}), \tag{3}$$

where the contingent cone is defined using the $Y^* \times Y$-topology.

Proof. Under the current assumptions, one can show using classical arguments that S is Lipschitz continuous from $Y^* \to Y$ and that (2) possesses a solution. We have chosen to place these proofs in the appendix as they are fairly standard. We refer the reader to [10, Chapter 2.2.1] and [45] for results on the Hadamard directional derivative. Note that $S'(\bar{u})(\cdot)$ is Lipschitz continuous on Y^*, and therefore on any subset of Y^*. Moreover, the Lipschitz continuity of S implies that $S'(\bar{u})(\cdot)$ coincides with the classical Gâteaux directional derivative.

Upon a direct generalization of the proofs of Theorem 2.1 and Corollary 2.1 in [23] (this can be done by replacing $H^1_0(\Omega)$ with Y, $L^2(\Omega)$ with $\mathcal{U}$, and $H^{-1}(\Omega)$ with Y^*), one immediately comes to the statement

$$\nabla_u J(\bar{u}, \bar{y}) \in Y, \tag{4}$$

where $(\bar{u}, \bar{y})$ is a locally optimal solution of (2).

Given that any (locally) optimal solution of (2) fulfills the inclusion

$$0 \in \hat{\partial}(J(\bar{u}, \bar{y}) + I_{[\mathcal{U}\times Y]\cap \mathrm{gph}\, S}(\bar{u}, \bar{y})),$$

where we take the convergences used to define the Fréchet subdifferential in the strong topology on $\mathcal{U} \times Y$, we can apply Proposition 5.1 in [36], which states that

$$(-\nabla_u J(\bar{u}, \bar{y}), -\nabla_y J(\bar{u}, \bar{y})) \in \widehat{N}_{[\mathcal{U}\times Y]\cap \mathrm{gph}\, S}(\bar{u}, \bar{y}).$$

Using now the inclusion

$$\widehat{N}_{[\mathcal{U}\times Y]\cap \mathrm{gph}\, S}(\bar{u}, \bar{y}) \subset \left[T_{[\mathcal{U}\times Y]\cap \mathrm{gph}\, S}(\bar{u}, \bar{y})\right]^-$$

from Theorem 1.10 in [35], we arrive at the condition

$$(\nabla_u J(\bar{u}, \bar{y}), w)_{\mathcal{U}} + \langle \nabla_y J(\bar{u}, \bar{y}), d\rangle_{Y^*,Y} \leq 0\ \forall (w, d) \in T_{[\mathcal{U}\times Y]\cap \mathrm{gph}\, S}(\bar{u}, \bar{y}).$$

Ones concludes directly from the definition of the contingent cone that the previous statement is equivalent to the following:

$$(\nabla_u J(\bar{u},\bar{y}), w)_{\mathcal{U}} + \langle \nabla_y J(\bar{u},\bar{y}), d\rangle_{Y^*,Y} \le 0 \ \forall (w,d) \in T_{\operatorname{gph} S|_{\mathcal{U}}}(\bar{u},\bar{y}),$$

where we note that $T_{\operatorname{gph} S|_{\mathcal{U}}}(\bar{u},\bar{y})$ is defined using the $\mathcal{U}\times Y$-topology. Given (4), we can extend this inequality so that

$$\langle \nabla_u J(\bar{u},\bar{y}), w\rangle_{Y,Y^*} + \langle \nabla_y J(\bar{u},\bar{y}), d\rangle_{Y^*,Y} \le 0 \ \forall (w,d) \in \operatorname{cl}\{T_{\operatorname{gph} S|_{\mathcal{U}}}(\bar{u},\bar{y})\}_{Y^*\times Y}.$$

We now demonstrate (3). Note that the inclusion

$$\operatorname{cl}\{T_{\operatorname{gph} S|_{\mathcal{U}}}(\bar{u},\bar{y})\}_{Y^*\times Y} \subseteq T_{\operatorname{gph} S}(\bar{u},\bar{y}),$$

with $T_{\operatorname{gph} S}(\bar{u},\bar{y})$ defined in the $Y^*\times Y$-topology, is trivial and follows directly from the definition.

Let $(w,d) \in T_{\operatorname{gph} S}(\bar{u},\bar{y})$, where the contingent cone is defined in the $Y^*\times Y$-topology. Then by definition, there exist sequences $t_k \to 0^+$, $w_k \to w$ in Y^*, and $d_k \to d$ in Y such that

$$\bar{y} + t_k d_k = S(\bar{u} + t_k w_k) \ \forall k.$$

As $\mathcal{U}$ is dense in Y^*, there exists for each k a sequence $\{w_l^k\} \subset \mathcal{U}$ such that $w_l^k \to w_k$ in Y^*. In addition, we define $\{w_n^{k,l}\} \subset \mathcal{U}$ such that $w_n^{k,l} \to w_l^k$ in $\mathcal{U}$ (and, in particular, in Y^*). We now define sequences d_l^k and $d_n^{k,l}$ such that

$$d_l^k := S'(\bar{u})(w_l^k) \text{ and } d_n^{k,l}(\tau) := \frac{S(\bar{u}+\tau w_n^{k,l}) - \bar{y}}{\tau} \ \forall \tau > 0.$$

As $S'(\bar{u})(\cdot)$ is (Lipschitz) continuous on Y^* and $w_n^{k,l} \to w_l^k$ in $\mathcal{U}$, $d_n^{k,l} \to d_l^k$ in Y. Since the previous argument holds for all $\tau > 0$, there exist sequences $\tau_k \to 0^+$, $w_n^{k,l} \to w_l^k$ in $\mathcal{U}$ and $d_n^{l,k} \to d_l^k$ in Y such that

$$\bar{y} + \tau_k d_n^{l,k} = S(\bar{u} + \tau w_n^{k,l}).$$

Therefore $(w_l^k, d_l^k) \in T_{\operatorname{gph} S|_{\mathcal{U}}}(\bar{u},\bar{y})$. Finally, we note that there exists $L > 0$ such that

$$||d_l^k - d||_Y = ||S'(\bar{u})(w_l^k) - S'(\bar{u})(w)||_Y \le L||w_l^k - w||_{Y^*} \le L||w_l^k - w_k||_{Y^*} + L||w_k - w||_{Y^*}$$

Hence, $(w,d) \in \operatorname{cl}\{T_{\operatorname{gph} S|_{\mathcal{U}}}(\bar{u},\bar{y})\}_{Y^*\times Y}$, as was to be shown. ☐

We note that in certain classical texts, e.g., Barbu's monograph [6], $S : \mathcal{U} \to Y$ is the solution mapping of the generalized equation

$$Bu \in Ay + N_M(y),$$

where B is a compact bounded linear operator from $\mathcal{U}$ into Y^*. See, for example, section 3.1 in [6]. By requiring the embedding of $\mathcal{U}$ into Y^* to be compact, we are essentially making the same requirement, albeit less general than Barbu.

Remark 3.2 (strength of the assumptions). One example of an objective functional that satisfies the needed requirements is

$$J(u,y) = \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)},$$

where $\alpha > 0$ and $y_d \in L^2(\Omega)$ with $Y = H_0^1(\Omega)$, $Y^* = H^{-1}(\Omega)$, and $\mathcal{U} = L^2(\Omega)$. It is well known that $H_0^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ forms a Gelfand triple. Moreover, it is also known that the embedding $L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ is compact.

4. Generalized differentiation of S: Polyhedric case.

In this section, we demonstrate the application of some known theoretical results and obtain a general formula for the proto-/contingent derivative of the solution of the variational inequality in the case where the set of (lower-level) state constraints is polyhedric.

Throughout this section, we define $S : Y^* \to Y$ to be the solution mapping of the variational inequality

$$u \in Ay + N_M(y),$$

where A is a coercive bounded linear operator from Y into Y^*, i.e., $A \in \mathcal{L}(Y, Y^*)$, and there exists a $\xi \in \mathbb{R}_+ \setminus \{0\}$ such that

$$\langle Ay, y\rangle_{Y^*,Y} \geq \xi ||y||_Y^2 \ \forall y \in Y.$$

In addition, we define M to be a closed convex subset of the reflexive Banach space Y and $u \in Y^*$. It is well known (see, e.g., Chapter 3 in [28]) that S is in fact a single-valued and locally Lipschitz function of u. We have included a concise proof of this in the appendix.

In order to provide a formula for the proto-derivative of S, we need a way of characterizing the proto-derivative of N_M. We begin with the case of so-called *polyhedric* sets M.

Definition 4.1 (polyhedric sets). *A closed convex set C of a Banach space X is called polyhedric if for all $x \in C$*

$$T_C(x) \cap \{v\}^{\perp} = \mathrm{cl}\left\{R_C(y) \cap \{v\}^{\perp}\right\}_X,$$

where $R_C(y)$ represents the so-called radial cone and is defined by

$$R_C(y) := \{h \in X \,|\exists \tau^* > 0 : \forall \tau \in [0, \tau^*],\, y + \tau h \in C\}$$

and $v \in N_C(y)$.

Note that, in general, the tangent cone to M contains the radial cone, in fact, it can also be defined as the closure of R_M. We will later provide two (nontrivial) examples containing polyhedric sets, but first we state the following important result due to Levy [29].

Theorem 4.2 (see [29, Theorem 3.1]). *Let M be a polyhedric subset of some reflexive Banach space Y and $v \in N_M(y)$. Then for any $(y, v) \in \mathrm{gph}\, N_M$, the normal cone mapping is proto-differentiable and the following are equivalent:*

1. $w \in PN_M[(y, v)](d)$.
2. $w \in N_{\mathcal{K}(y,v)}(d)$.
3. $(d, w) \in \mathcal{K}(y, v) \times [\mathcal{K}(y, v)]^- : \langle w, d\rangle = 0$.

Here, $\mathcal{K}(y, v) := T_M(y) \cap \{v\}^{\perp}$, i.e., the critical cone.

We would also like to bring to the reader's attention the fact that Theorem 4.2 can be easily derived by using equation (2.13) in Example 2.10 together with Theorem 3.9 in [13].

By appealing to a special calculus rule for proto-derivatives of certain classes of multifunctions, we prove the next important corollary.

Corollary 4.3 (the proto-derivative of S). *Let M be a polyhedric subset of a reflexive Banach space Y and S be as above. If $(u, y) \in \mathrm{gph}\, S$, then S is proto-differentiable and the following are equivalent:*

1. $d \in PS[(u, y)](w)$.

2. $w \in Ad + N_{\mathcal{K}(y,v)}(d)$.

3. $w - Ad \in [\mathcal{K}(y,v)]^-,\ d \in \mathcal{K}(y,v),\ \langle w - Ad, d\rangle = 0$,

or equivalently

$$T_{\mathrm{gph}\, S}(u,y) = \left\{(w,d) \in Y^* \times Y \,\middle|\, w - Ad \in [\mathcal{K}(y,v)]^-,\ d \in \mathcal{K}(y,v),\ \langle w - Ad, d\rangle = 0\right\}.$$

Here, $v := u - Ay$.

Proof. Define

$$S^{-1}(y) := \{u \in Y^* \,|u \in Ay + N_M(y)\}.$$

It can be easily derived from the definition of the proto-derivative (see, e.g., Proposition 2.3 in [31]) that

$$d \in PS[(u,y)](w) \Leftrightarrow w \in PS^{-1}[(y,u)](d).$$

Then since A is Fréchet differentiable from Y to Y^*, we can refer to Proposition 3.4 in [31], which states that

$$w \in PS^{-1}[(y,u)](d) = Ad + PN_M[(y, u - Ay)](d).$$

The rest follows from Theorem 4.2. □

Remark 4.4 (proto-derivatives vs. conical derivatives). In his seminal 1976 paper [33], Mignot introduces a type of one-sided directional derivative for continuous mappings between Banach spaces called the *conical derivative*. Perhaps the most stunning result concerning these derivatives is found in Theorem 3.3 of [33], where solutions to a specific class of variational inequalities are shown to admit a conical derivative for every perturbation parameter (similar to w in Corollary 4.3). In turn, the conical derivative is the solution of a variational inequality of the type found in Corollary 4.3. However, Mignot does so only for a restricted class of function spaces and choice of M. In this sense, Corollary 4.3 extends Mignot's result to reflexive Banach spaces for all state constraints M, provided the sets are polyhedric. Indeed, 2. in Corollary 4.3 can be viewed as the necessary and sufficient optimality conditions to the optimization problem

$$\min_{d\in Y} \left\{\frac{1}{2}\langle Ad, d\rangle - \langle w, d\rangle + I_{\mathcal{K}(y,v)}(d)\right\}.$$

Since this objective function is strictly convex, coercive, and lower-semicontinuous, the optimization problem always has a unique solution; see, e.g., Theorem 3.3.4 in [4]; this holds regardless of the choice of $w \in Y^*$. Moreover, it is easy to determine the Lipschitz continuity of these solutions as functions of w using the coercivity of the operator A. Hence, for any fixed $(u,y) \in \mathrm{gph}\, S$, i.e., a solution y of the original variational inequality for a given u, we see that at every point $w \in Y^*$, this solution admits a proto-derivative $PS[(u,y)](w) = d$, where d is the unique solution of the variational inequality:

$$\text{Find } d \in \mathcal{K}(y,v) : \langle Ad, d' - d\rangle \geq \langle w, d' - d\rangle\ \forall d' \in \mathcal{K}(y,v).$$

In Corollary 6.10, we extend this result further, beyond the realm of polyhedricity, to an important class of MPECs arising in the study of elastoplasticity.

Based on the discussions in Remark 4.4 and at the end of section 2, we have demonstrated the following result. Note that this result is more or less due to Haraux (cf. [18]); we have simply used a different technique.

PROPOSITION 4.5 (Hadamard directional differentiability of S). *Let M be a polyhedric subset of a reflexive Banach space Y and S be as above. If $(u, y) \in \operatorname{gph} S$, then S is Hadamard directionally differentiable at $\bar{u}$.*

We now present our first main result providing dual optimality conditions for (2).

THEOREM 4.6 (strong stationarity conditions I). *Let $(\bar{u}, \bar{y})$ be a locally optimal solution of (2); then there exist multipliers $p^* \in \mathcal{K}(\bar{y}, \bar{v})$, $r^* \in [\mathcal{K}(\bar{y}, \bar{v})]^-$, and $\bar{v} \in N_M(\bar{y})$ such that*

$$0 = \nabla_u J(\bar{u}, \bar{y}) + p^*, \tag{5}$$

$$0 = \nabla_y J(\bar{u}, \bar{y}) + r^* - A^* p^*, \tag{6}$$

$$0 = A\bar{y} - \bar{u} + \bar{v}. \tag{7}$$

Proof. The result follows from (3) upon the calculation of $[T_{\operatorname{gph} S}(\bar{u}, \bar{y})]^-$. By definition of the polar cone we see that

$$[T_{\operatorname{gph} S}(u, y)]^- = \{(p^*, q^*) \in Y \times Y^* \,|\langle w, p^* \rangle + \langle q^*, d \rangle \leq 0 \ \forall (w, d) \in T_{\operatorname{gph} S}(u, y)\}.$$

Using Corollary 4.3, we can proceed in a manner similar to the proof of Lemma 3.1 in [39],

$$\begin{aligned}
[T_{\operatorname{gph} S}(u, y)]^- &= \{(p^*, q^*) \in Y \times Y^* \,|\langle w, p^* \rangle + \langle q^*, d \rangle \leq 0 \ \forall (w, d) \in T_{\operatorname{gph} S}(u, y)\} \\
&= \left\{(p^*, q^*) \in Y \times Y^* \,\middle|\langle Ad + h, p^* \rangle + \langle q^*, d \rangle \leq 0 \ \forall (d, h) \in \operatorname{gph} N_{\mathcal{K}(y,v)}\right\} \\
&= \left\{(p^*, q^*) \in Y \times Y^* \,\middle|\langle A^* p^* + q^*, d \rangle + \langle h, p^* \rangle \leq 0 \ \forall (d, h) \in \operatorname{gph} N_{\mathcal{K}(y,v)}\right\}.
\end{aligned}$$

Recall from Theorem 4.2 that

$$\operatorname{gph} N_{K(y,v)} = \left\{(d, h) \in Y \times Y^* \,\middle|\, d \in \mathcal{K}(y, v),\ h \in [\mathcal{K}(y, v)]^-,\ \langle h, d \rangle = 0\right\}.$$

Then by ignoring the complementarity relation $\langle h, d \rangle = 0$, we observe that

$$[T_{\operatorname{gph} S}(u, y)]^- \supset \left\{(p^*, q^*) \in Y \times Y^* \,\middle|\, A^* p^* + q^* \in [\mathcal{K}(y, v)]^-,\ p^* \in \mathcal{K}(y, v)\right\}.$$

Conversely, since $\mathcal{K}(y, v) \times \{0\}$ and $\{0\} \times [\mathcal{K}(y, v)]^-$ are subsets of $\operatorname{gph} N_{\mathcal{K}(y,v)}$ we obtain the reverse inclusion, i.e.,

$$[T_{\operatorname{gph} S}(u, y)]^- \subset \left\{(p^*, q^*) \in Y \times Y^* \,\middle|\, A^* p^* + q^* \in [\mathcal{K}(y, v)]^-,\ p^* \in \mathcal{K}(y, v)\right\},$$

as was to be shown. ∎

We would also like to note that a similar technique was used to derive optimality conditions for the class of MPECs in which $Y = H_0^1(\Omega)$, $\mathcal{U} = Y^* = H^{-1}(\Omega)$, and $M := \left\{y \in H_0^1(\Omega) \,|y \geq 0 \text{ a.e.}\, \Omega\right\}$ in [39] and when $\mathcal{U} = L^2(\Omega)$ in [23].

5. Two examples with polyhedric state constraints M.

In this section, we provide examples in which the conditions (5) and (6) from Theorem 4.6 are made explicit. We begin with a classical example for illustration, following which we derive new conditions for an important example from the study of elastoplasticity (cf. [26, 27]).

5.1. Optimal control of the obstacle problem. In [33], Mignot demonstrates the polyhedricity of constraint sets of the type

$$\left\{ y \in H_0^1(\Omega) \,|\, \varphi \le y \le \psi \text{ a.e.}\, \Omega \right\},$$

where $\psi, \varphi \in H^1(\Omega)$ are appropriately chosen. Furthermore, many other obstacle-problem–type constraints sets, i.e., box-constraints, in L^p-spaces are shown to be polyhedric in Chapter 6 of [10]. Using the optimality conditions derived in the previous section, along with the well-known characterizations for both the associated critical cone and its dual, we quickly rederive the well-known conditions of Mignot and Puel [34]. As their conditions are considered to be the best possible for the optimal control of the obstacle problem, we consider this brief derivation as a type of validation for the optimality of our conditions.

In the following, we have that

- $Y := H_0^1(\Omega)$;
- $\mathcal{U} := L^2(\Omega)$;
- $M := \left\{ y \in H_0^1(\Omega) \,|\, y \ge 0 \text{ a.e.}\, \Omega \right\}$;
- $J : L^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, Fréchet differentiable;
- $\mathcal{A}(y) := \{ x \in \Omega \,| y(x) = 0 \}$.

For the definitions and assumptions on these spaces, we refer the reader to section 2.

By Theorem 4.6, if $(\bar{u}, \bar{y})$ is a locally optimal solution of the associated elliptic MPEC, then there exist $(p^*, r^*) \in H_0^1(\Omega) \times H^{-1}(\Omega)$ such that

$$\begin{aligned} 0 &= \nabla_u J(\bar{u}, \bar{y}) + p^*, \\ 0 &= \nabla_y J(\bar{u}, \bar{y}) + r^* - A^* p^*, \end{aligned}$$

where $p^* \in T_M(\bar{y}) \cap \{\bar{u} - A\bar{y}\}^\perp$ and $r^* \in \left[T_M(\bar{y}) \cap \{\bar{u} - A\bar{y}\}^\perp \right]^-$. Then by referring to Lemma 3.2 in [39] (along with the discussion following Lemma 2.2), it holds that

$$\begin{aligned} p^* &\ge 0 \text{ a.e.}\, \mathcal{A}(\bar{y}), \\ \langle \bar{u} - A\bar{y}, p^* \rangle &= 0, \\ \langle r^*, \varphi \rangle &= 0 \;\forall \varphi \in H_0^1(\Omega) : \varphi = 0 \text{ a.e.}\, \mathcal{A}(\bar{y}), \\ \langle r^*, \varphi \rangle &\le 0 \;\forall \varphi \in M : \langle \bar{u} - A\bar{y}, \varphi \rangle = 0. \end{aligned}$$

By carefully comparing these conditions to Theorem 3.3 in [34], we see that our conditions yield those of Mignot and Puel. Note that in [34], $J(u, y) := \frac{1}{2}||y - z_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u||^2_{L^2(\Omega)}$, with $\alpha > 0$. As expected in this setting, we see that the regularity of the optimal control $\bar{u}$ is better than $L^2(\Omega)$, in fact, $\bar{u} \in H_0^1(\Omega)$. In the recent paper [21], efforts were successfully made to calculate and/or obtain strong stationary points of this type numerically.

5.2. Pointwise constraints on the gradient of the state using the ∞-norm. Many important problems in the study of elastoplasticity require the pointwise bounding of the gradient of the displacement, i.e., the stress on an isotropic body at each point in the presence of a given force. The optimal control problem then results in an elliptic MPEC in which the gradients of the state y are pointwise bounded (almost everywhere). In the following example, we consider a setting in which the gradients of the state are pointwise bounded using the ∞-norm on vectors in $\mathbb{R}^l$. In section 8, we consider the 2-norm instead, which does not allow a simple reformulation to a bilateral setting.

In the following, we have that

- $Y := H_0^1(\Omega)$;
- $\mathcal{U} := L^2(\Omega)$;
- $M := \left\{ y \in H_0^1(\Omega) \,|\, |\nabla y|_\infty \le \psi \text{ a.e.} \,\Omega \right\}$, $\psi \in L^\infty(\Omega)$, and $\exists \underline{\psi} \in \mathbb{R}_+ \setminus \{0\}$: $\psi \ge \underline{\psi} > 0$ a.e. Ω;
- $J : L^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, Fréchet differentiable.

We use ∇y to represent the full gradient and $(\nabla y)_i$ for its components. This simple rule will be applied throughout this example for all vectors and their components. It is easy to see that M can be equivalently defined by

$$M = \left\{ y \in H_0^1(\Omega) \,|\, -\psi \le (\nabla y)_i \le \psi \text{ a.e.} \,\Omega, 1 \le i \le l \right\}.$$

In addition to these basic assumptions, we reduce the space of the gradient used in the state constraints, i.e., more specifically,

- $\nabla : H_0^1(\Omega) \to G(\Omega)$, where $G(\Omega) := \nabla(H_0^1(\Omega))$, i.e., the image space of the gradient.

In shrinking the image space of the gradient, we obtain a surjective bounded linear operator. This leads to the new formulation of the constraint set M:

$$M = \left\{ y \in H_0^1(\Omega) \,|\, \nabla y \in B_\psi \right\}, \tag{8}$$

where

$$B_\psi := \{ \mathbf{z} \in G(\Omega) \,|\, -\psi \le z_i \le \psi \text{ a.e.} \,\Omega,\ 1 \le i \le l \}.$$

The image space is not merely chosen for its convenience. Indeed, we have the well-known Helmholtz decomposition (see, e.g., Proposition 1 (Chapter IX, section 1) in [12]) that $L^2(\Omega)^l$ can be written as the orthogonal direct sum

$$L^2(\Omega)^l = G(\Omega) \oplus H(\operatorname{div} 0, \Omega),$$

where

$$H(\operatorname{div} 0, \Omega) := \left\{ \mathbf{z} \in L^2(\Omega)^l \,|\, \operatorname{div} \mathbf{z} = 0 \right\}.$$

This then allows us to use the $L^2(\Omega)^l$-norm on $G(\Omega)$. Not only is $G(\Omega)$ closed with respect to the $L^2(\Omega)^l$-norm, but it is also a Hilbert space with inner product $(\cdot,\cdot)_{L^2(\Omega)^l}$. Using these facts, it is easy to see the B_ψ is closed and convex in $G(\Omega)$. We now show, using essentially the same argument as in the proof of Proposition 6.33 in [10], that B_ψ is polyhedric in $G(\Omega)$. In order to continue, we will need the following definitions.

Definition 5.1 (the active and inactive sets). *For some $y \in M$, we define the upper active set for the ith component of ∇y, $\mathcal{A}_i^+(\nabla y) \subseteq \Omega$, such that*

$$\mathcal{A}_i^+(\nabla y) := \{ x \in \Omega \,|\, (\nabla y(x))_i = \psi(x) \},$$

and the lower active set for the ith component of ∇y, $\mathcal{A}_i^-(\nabla y) \subseteq \Omega$, such that

$$\mathcal{A}_i^-(\nabla y) := \{ x \in \Omega \,|\, (\nabla y(x))_i = -\psi(x) \}.$$

The ith inactive set, $\mathcal{I}_i(\nabla y)$, is therefore defined by

$$\mathcal{I}_i(\nabla y) := \Omega \setminus \left(\mathcal{A}_i^+(\nabla y) \cup \mathcal{A}_i^-(\nabla y) \right).$$

Note that we can analogously define the active and inactive sets for any element of B_ψ.

PROPOSITION 5.2 (the tangent and normal cones to B_ψ). *Let the set B_ψ be defined as above and $\mathbf{z} \in B_\psi$. Then*

$$T_{B_\psi}(\mathbf{z}) = \left\{\mathbf{h} \in G(\Omega) \,\middle|\, h_i \le 0 \; a.e. \, \mathcal{A}_i^+(\mathbf{z}),\; h_i \ge 0 \; a.e. \, \mathcal{A}_i^-(\mathbf{z}), 1 \le i \le l \right\}.$$

Moreover, the properties of $G(\Omega)$ allow the elements of the normal cone $N_{B_\psi}(\mathbf{z})$ to be identified with elements of $G(\Omega)$ so that

$$\begin{aligned} &N_{B_\psi}(\mathbf{z}) \\ &= \left\{\boldsymbol{\lambda} \in G(\Omega) \,\middle|\, \lambda_i \ge 0 \; a.e. \, \mathcal{A}_i^+(\mathbf{z}),\; \lambda_i \le 0 \; a.e. \, \mathcal{A}_i^-(\mathbf{z}),\; \lambda_i = 0 \; a.e. \, \mathcal{I}_i(\mathbf{z}),\; 1 \le i \le l \right\}. \end{aligned}$$

Therefore, the critical cone to B_ψ at $\mathbf{z}$ for some $\boldsymbol{\lambda} \in N_{B_\psi}(\mathbf{z})$ is characterized as

$$\mathcal{K}(\mathbf{z}, \boldsymbol{\lambda}) = T_{B_\psi}(\mathbf{z}) \cap \{\boldsymbol{\lambda}\}^\perp = \left\{ \mathbf{h} \in G(\Omega) \,\middle|\, \begin{array}{ll} h_i \le 0 & a.e. \, \mathcal{A}_i^+(\mathbf{z}) : \lambda_i = 0, \\ h_i = 0 & a.e. \, \mathcal{A}_i^+(\mathbf{z}) : \lambda_i > 0, \\ h_i \ge 0 & a.e. \, \mathcal{A}_i^-(\mathbf{z}) : \lambda_i = 0, \\ h_i = 0 & a.e. \, \mathcal{A}_i^-(\mathbf{z}) : \lambda_i < 0, 1 \le i \le l \end{array} \right\},$$

where

$$\{\boldsymbol{\lambda}\}^\perp := \left\{\mathbf{h} \in G(\Omega) \,\middle|\, (\boldsymbol{\lambda}, \mathbf{h})_{L^2(\Omega)^l} = 0 \right\}.$$

Here, conditions of the type "$a.e. \, \mathcal{A}_i^+(\mathbf{z}) : \lambda_i = 0$" are to be understood: "almost everywhere on the set $\left\{x \in \mathcal{A}_i^+(\mathbf{z}) \,|\, \lambda_i(x) = 0 \right\}$."

Proof. Let $\mathbf{h} \in T_{B_\psi}(\mathbf{z})$. Then by definition, there exist sequences $t_k \to 0^+$ and $\mathbf{h}_k \to \mathbf{h}$ in $G(\Omega)$ such that

$$\mathbf{z} + t_k \mathbf{h}_k \in B_\psi \;\forall k.$$

Then for any i such that $1 \le i \le l$, it holds that

$$-\psi \le z_i + t_k h_k^i \le \psi \text{ a.e.}\, \Omega \;\forall k.$$

Thus,

$$0 \le h_k^i \le \frac{\psi - z_i}{t_k} \text{ a.e.}\, \mathcal{A}_i^-(\mathbf{z}) \;\forall k$$

and

$$\frac{-\psi - z_i}{t_k} \le h_k^i \le 0 \text{ a.e.}\, \mathcal{A}_i^+(\mathbf{z}) \;\forall k.$$

Here, we use h_k^i to represent the ith component of the $\mathbf{h}_k$. Hence,

$$T_{B_\psi}(\mathbf{z}) \subseteq \left\{\mathbf{h} \in G(\Omega) \,\middle|\, h_i \le 0 \text{ a.e.}\, \mathcal{A}_i^+(\mathbf{z}),\; h_i \ge 0 \text{ a.e.}\, \mathcal{A}_i^-(\mathbf{z}), 1 \le i \le l \right\}. \tag{9}$$

Now let $\mathbf{h} \in G(\Omega)$ satisfy the right-hand side of (9) and define the following class of vector fields indexed by $\tau > 0$:

$$r_\tau := \frac{\Pi_{B_\psi}(\mathbf{z} + \tau \mathbf{h}) - \mathbf{z}}{\tau}.$$

Here, Π_{B_ψ} represents the metric projection onto B_ψ. Since $\mathbf{z} + \tau r_\tau = \Pi_{B_\psi}(\mathbf{z} + \tau \mathbf{h})$, $r_\tau \in R_{B_\psi}(\mathbf{z})$ for all $\tau > 0$.

Consider now that for almost every $x \in \Omega$, $r_\tau(x) \to \mathbf{h}(x)$. Indeed, pointwise, we can always find $\tau > 0$ small enough such that $\mathbf{z}(x) + \tau\mathbf{h}(x) \in B_\psi(x)$, where $B_\psi(x) := \left\{ q \in \mathbb{R}^l \,\middle|\, -\psi(x) \le q_i \le \psi(x) \right\}$. Moreover, since $G(\Omega)$ is a Hilbert space and $B_\psi(x)$ is closed and convex for almost every $x \in \Omega$, the metric projection is single-valued and Lipschitz continuous with modulus 1 (nonexpansive). Therefore, it holds that $|r_\tau(x)| \le |h(x)|$ for almost every $x \in \Omega$. Then given $G(\Omega)$ is a closed subspace of $L^2(\Omega)^l$, we can apply Lebesgue's dominating convergence theorem, which yields $r_\tau \to \mathbf{h}$ in $G(\Omega)$. As the set of all r_τ is contained in $R_{B_\psi}(\mathbf{z})$ and $\mathrm{cl}\left\{R_{B_\psi}(\mathbf{z})\right\}_{G(\Omega)} = T_{B_\psi}(\mathbf{z})$, (9) holds as an equality.

We now move on to the derivation of the normal cone. By definition

$$N_{B_\psi}(\mathbf{z}) = \left[T_{B_\psi}(\mathbf{z})\right]^- = \left\{ \boldsymbol{\lambda} \in G(\Omega)^* \,\middle|\, \langle \boldsymbol{\lambda}, \mathbf{h}\rangle_{G(\Omega)^*,G(\Omega)} \le 0 \,\forall h \in T_{B_\psi}(\mathbf{z}) \right\}.$$

By virtue of the Riesz representation theorem, there exists a unique $\tilde{\boldsymbol{\lambda}} \in G(\Omega)$ for each $\boldsymbol{\lambda} \in G(\Omega)^*$ such that

$$\langle \boldsymbol{\lambda}, \mathbf{h}\rangle_{G(\Omega)^*,G(\Omega)} = \left(\tilde{\boldsymbol{\lambda}}, \mathbf{h}\right)_{L^2(\Omega)^l}.$$

Hence, we identify all $\boldsymbol{\lambda} \in N_{B_\psi}(\mathbf{z})$ with their $G(\Omega)$-counterparts, so that

$$\langle \boldsymbol{\lambda}, \mathbf{h}\rangle_{G(\Omega)^*,G(\Omega)} \le 0 \,\forall h \in T_{B_\psi}(\mathbf{z}) \Leftrightarrow \left(\tilde{\boldsymbol{\lambda}}, \mathbf{h}\right)_{L^2(\Omega)^l} \le 0,\ \forall h \in T_{B_\psi}(\mathbf{z}).$$

Defining $\mathcal{A}^+(\mathbf{z}) := \mathcal{A}_1^+(\mathbf{z}) \times \cdots \times \mathcal{A}_l^+(\mathbf{z})$, $\mathcal{A}^-(\mathbf{z}) := \mathcal{A}_1^-(\mathbf{z}) \times \cdots \times \mathcal{A}_l^-(\mathbf{z})$, and $\mathcal{I}(\mathbf{z}) := \mathcal{I}_1(\mathbf{z}) \times \cdots \times \mathcal{I}_l(\mathbf{z})$, the polarity inequality becomes

$$\left(\tilde{\boldsymbol{\lambda}}, \mathbf{h}\right)_{L^2(\Omega)^l} = \int_{\mathcal{A}^+(\mathbf{z})} \tilde{\boldsymbol{\lambda}} \cdot \mathbf{h} dx + \int_{\mathcal{A}^-(\mathbf{z})} \tilde{\boldsymbol{\lambda}} \cdot \mathbf{h} dx + \int_{\mathcal{I}(\mathbf{z})} \tilde{\boldsymbol{\lambda}} \cdot \mathbf{h} dx \le 0 \ \forall h \in T_{B_\psi}(\mathbf{z}).$$

Referring to the above, we see that if $\mathbf{h} \in G(\Omega)$ such that $h_i = 0$ a.e. on $\mathcal{A}_i^+(\mathbf{z}) \cup \mathcal{A}_i^-(\mathbf{z})$ and free on $\mathcal{I}_i(\mathbf{z})$ for all $i = 1, \ldots, l$, then $\mathbf{h} \in T_{B_\psi}(\mathbf{z})$. Therefore, $\tilde{\boldsymbol{\lambda}}$ must equal zero a.e. on $\mathcal{I}(\mathbf{z})$. Then since the components of $\mathbf{h}$ are a.e. nonpositive on $\mathcal{A}^+(\mathbf{z})$ and a.e. nonnegative on $\mathcal{A}^-(\mathbf{z})$ for all $\mathbf{h} \in T_{B_\psi}(\mathbf{z})$, $\tilde{\boldsymbol{\lambda}}$ must always have the opposite signs (a.e.) on these sets. By identifying the $\tilde{\boldsymbol{\lambda}}$ with the $\boldsymbol{\lambda}$, the asserted formula for the normal cone holds.

Given the formulae for the tangent and normal cones, the characterization for the critical cone follows trivially. ∎

COROLLARY 5.3 (polyhedricity of B_ψ). *The set B_ψ as defined above is polyhedric in $G(\Omega)$; i.e., for any $\boldsymbol{\lambda} \in N_{B_\psi}(\mathbf{z})$, it holds that*

$$T_{B_\psi}(\mathbf{z}) \cap \{\boldsymbol{\lambda}\}^\perp = \mathrm{cl}\left\{R_{B_\psi}(\mathbf{z}) \cap \{\boldsymbol{\lambda}\}^\perp\right\}_{G(\Omega)}.$$

Proof. The argument follows analogously to the derivation of the tangent cone and mirrors the proof of Proposition 6.33 in [10]. ∎

This leads to our next result.

PROPOSITION 5.4 (the tangent and normal cones to M). *Let $y \in M$, where M is defined as in* (8). *Then*

$$T_M(y) = \left\{ d \in H_0^1(\Omega) \,\middle|\, \nabla d \in T_{B_\psi}(\nabla y) \right\}$$

and

$$N_M(y) = \left\{ -\operatorname{div} \boldsymbol{\lambda} \in H^{-1}(\Omega) \,\middle|\, \boldsymbol{\lambda} \in G(\Omega) : \boldsymbol{\lambda} \in N_{B_\psi}(\nabla y) \right\}.$$

Here, the associated critical cone to M at $y \in M$ for some $v \in N_M(y)$ is characterized by

$$\mathcal{K}(y,v) = T_M(y) \cap \{v\}^\perp = \left\{ d \in H_0^1(\Omega) \,\middle|\, \begin{array}{ll} \nabla d_i \le 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0, \\ \nabla d_i = 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i > 0, \\ \nabla d_i \ge 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0, \\ \nabla d_i = 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i < 0, 1 \le i \le l \end{array} \right\}.$$

Here, $\boldsymbol{\lambda} \in N_{B_\psi}(\nabla y)$ such that $v = -\operatorname{div} \boldsymbol{\lambda}$.

Proof. Due to the assumption on the range space of ∇, the classical generalized Slater condition, i.e.,

$$0 \in \operatorname{int} \left\{ \nabla(H_0^1(\Omega)) - B_\psi \right\},$$

automatically holds (understood in the strong topology on $G(\Omega)$). Indeed, since $\nabla(H_0^1(\Omega)) = G(\Omega)$ and $B_\psi \subset G(\Omega)$, there always exists an $\varepsilon > 0$ such that $\mathbb{B}_\varepsilon(0) \subset \nabla(H_0^1(\Omega)) - B_\psi$. Thus, the assertions hold for $T_M(y)$ and $N_M(y)$ (cf., e.g., [5, Chapter 4.2]). Note that the adjoint of ∇ is $-\operatorname{div}$, understood in a weak sense.

Due to Proposition 5.2, for each $v \in N_M(y)$, there exists a $\boldsymbol{\lambda} \in N_{B_\psi}(\nabla y)$ such that $v = -\operatorname{div} \boldsymbol{\lambda}$. Therefore, taking any arbitrary $d \in \mathcal{K}(y,v)$ requires

$$\langle -\operatorname{div} \boldsymbol{\lambda}, d \rangle_{H^{-1}, H_0^1} = (\boldsymbol{\lambda}, \nabla d)_{L^2(\Omega)^l} = 0.$$

Continuing the previous relation further yields

$$\begin{aligned} (\boldsymbol{\lambda}, \nabla d)_{L^2(\Omega)^l} &= \int_{\mathcal{A}^+(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d dx + \int_{\mathcal{A}^-(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d dx + \int_{\mathcal{I}(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d dx \\ &= \int_{\mathcal{A}^+(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d dx + \int_{\mathcal{A}^-(\nabla y)} \boldsymbol{\lambda} \cdot \nabla d dx = 0. \end{aligned}$$

The rest follows from the fact that ∇d_i and λ_i always have opposite signs. The reverse inclusion is trivial. ∎

Given the explicit formula for the critical cone associated with M provided by the previous proposition, we now demonstrate that M is in fact polyhedric.

PROPOSITION 5.5 (polyhedricity of M). *Given M as above, $y \in M$, and $v \in N_M(y)$, it holds that*

$$\mathcal{K}(y,v) = T_M(y) \cap \{v\}^\perp = \operatorname{cl} \left\{ R_M(y) \cap \{v\}^\perp \right\}_{H_0^1(\Omega)},$$

i.e., M is polyhedric.

Proof. Since ∇ is onto and B_ψ is polyhedric (Proposition 5.2), Proposition 3.54 in [10] implies

$$\operatorname{cl} \left\{ d \in \mathcal{K}(y,v) \,\middle|\, \nabla d \in R_{B_\psi}(\nabla y) \right\}_{H_0^1(\Omega)} = \mathcal{K}(y,v).$$

Given

$$T_M(y) \cap \{v\}^\perp \supseteq \operatorname{cl} \left\{ R_M(y) \cap \{v\}^\perp \right\}_{H_0^1(\Omega)},$$

it suffices to show that

$$\left\{d \in \mathcal{K}(y,v) \,\middle|\, \nabla d \in R_{B_\psi}(\nabla y)\right\} \subset R_M(y) \cap \{v\}^\perp .$$

By definition $\nabla d \in R_{B_\psi}(\nabla y)$ implies the existence of a $\tau > 0$ such that $\nabla y + \tau \nabla d \in B_\psi$. Hence, $d \in R_M(y)$. Then since $d \in \mathcal{K}(y,v)$ implies $\langle v, d\rangle_{H^{-1},H_0^1} = 0$, it holds that $d \in R_M(y) \cap \{v\}^\perp$. ∎

Given the previous results, we need one last component in order to provide the explicit stationarity conditions for the elliptic MPEC.

PROPOSITION 5.6 (the polar cone $[\mathcal{K}(y,v)]^-$). *Given M as above, $y \in M$, and $v \in N_M(y)$, it holds that*

$$[\mathcal{K}(y,v)]^- \\ = \left\{ -\mathrm{div}\,\boldsymbol{\mu} \in H^{-1}(\Omega) \,\middle|\, \boldsymbol{\mu} \in G(\Omega) : \begin{array}{ll} \mu_i \geq 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 \\ \mu_i \leq 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0,\ 1 \leq i \leq l \\ \mu_i = 0 & a.e.\, \mathcal{I}_i(\nabla y) \end{array} \right\}.$$

Proof. By definition,

$$[\mathcal{K}(y,v)]^- = \left\{ d^* \in H^{-1}(\Omega) \,\middle|\, \langle d^*, d\rangle_{H^{-1},H_0^1} \leq 0\, \forall d \in \mathcal{K}(y,v) \right\}.$$

Let $\boldsymbol{\mu}$ satisfy the requirements for the right-hand side of the asserted result. Then by Proposition 5.4, for any $d \in \mathcal{K}(y,v)$, we have

$$\langle -\mathrm{div}\,\boldsymbol{\mu}, d\rangle_{H^{-1},H_0^1} = (\boldsymbol{\mu}, \nabla d)_{L^2} \leq 0.$$

Therefore, the inclusion "$\supseteq$" holds. For the reverse direction, define

$$L := \left\{ \boldsymbol{\mu} \in G(\Omega) \,\middle|\, \begin{array}{ll} \mu_i \geq 0 & \text{a.e. } \mathcal{A}_i^+(\nabla y) : \lambda_i = 0, \\ \mu_i \leq 0 & \text{a.e. } \mathcal{A}_i^-(\nabla y) : \lambda_i = 0,\ 1 \leq i \leq l \\ \mu_i = 0 & \text{a.e. } \mathcal{I}_i(\nabla y) \end{array} \right\}.$$

It is easy to show that L is a closed convex cone in $G(\Omega)$. The image $-\mathrm{div}\,(L)$ is clearly convex; we refer the reader to Lemma 6.14 to see that $-\mathrm{div}\,(L)$ is closed as well. Assume now that there exists some $d^* \in [\mathcal{K}(y,v)]^-$ such that $d^* \notin -\mathrm{div}\,(L)$. Then there must exist some $\delta \in H_0^1(\Omega)$ strongly separating d^* from $-\mathrm{div}\,(L)$; see, e.g., TVS II.38, Prop. 4 of [11], i.e.,

$$\langle d^*, \delta\rangle_{H^{-1},H_0^1} > 0, \quad \langle -\mathrm{div}\,\boldsymbol{\mu}, \delta\rangle_{H^{-1},H_0^1} \leq 0,\ \forall \boldsymbol{\mu} \in L.$$

Then δ cannot be in $\mathcal{K}(y,v)$. However, for an arbitrary $\boldsymbol{\mu} \in L$, it holds that

$$0 \geq \langle -\mathrm{div}\,\boldsymbol{\mu}, \delta\rangle_{H^{-1},H_0^1} = (\boldsymbol{\mu}, \nabla\boldsymbol{\delta})_{L^2(\Omega)^l}.$$

Since the previous relation must hold for all $\boldsymbol{\mu} \in L$, we deduce that $\delta \in \mathcal{K}(y,v)$, a contradiction. The assertion follows. ∎

Now that we have all the necessary characterizations, we can provide explicit strong stationarity conditions for the elliptic MPEC via Theorem 4.6.

PROPOSITION 5.7 (explicit strong stationarity conditions). *Under the given data assumptions of this subsection, let $(\bar{u}, \bar{y})$ be a (locally) optimal solution to the corresponding MPEC. Then there exist multipliers $p \in H_0^1(\Omega)$, $\boldsymbol{\lambda} \in G(\Omega)$, and $\boldsymbol{\mu} \in G(\Omega)$*

such that

$$0 = \nabla_u J(\bar{u}, \bar{y}) + p, \tag{10}$$

$$0 = \nabla_y J(\bar{u}, \bar{y}) - \operatorname{div} \boldsymbol{\mu} - A^* p, \tag{11}$$

$$0 = A\bar{y} - \bar{u} - \operatorname{div} \boldsymbol{\lambda}, \tag{12}$$

(13)

where for all $i = 1, \ldots, l$

$$\left\{ \begin{array}{ll|ll|ll} \nabla p_i \le 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 & \mu_i \ge 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i = 0 & \lambda_i \ge 0 & a.e.\, \mathcal{A}_i^+(\nabla y) \\ \nabla p_i = 0 & a.e.\, \mathcal{A}_i^+(\nabla y) : \lambda_i > 0 & \mu_i \le 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 & \lambda_i \le 0 & a.e.\, \mathcal{A}_i^-(\nabla y) \\ \nabla p_i \ge 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i = 0 & \mu_i = 0 & a.e.\, \mathcal{I}_i(\nabla y) & \lambda_i = 0 & a.e.\, \mathcal{I}_i(\nabla y) \\ \nabla p_i = 0 & a.e.\, \mathcal{A}_i^-(\nabla y) : \lambda_i < 0 & & & & \end{array} \right\}$$

and

$$\begin{aligned} \mathcal{A}_i^+(\nabla y) &= \{ x \in \Omega \,|\, (\nabla y(x))_i = \psi(x) \}, \\ \mathcal{A}_i^-(\nabla y) &= \{ x \in \Omega \,|\, (\nabla y(x))_i = -\psi(x) \}, \\ \mathcal{I}_i(\nabla y) &= \Omega \setminus \mathcal{A}_i^+(\nabla y) \cup \mathcal{A}_i^-(\nabla y). \end{aligned}$$

Proof. The result follows from Theorem 4.6 via Propositions 5.4, 5.5, and 5.6. ∎

As in the simple obstacle case, we see that if the tracking functional

$$J(u, y) = \frac{1}{2} ||y - y_d||_{L^2}^2 + \frac{\alpha}{2} ||u||_{L^2(\Omega)}^2$$

is chosen with $\alpha > 0$, then the optimal control $\bar{u} \in H_0^1(\Omega)$.

6. Optimality conditions in the absence of polyhedricity: Gradient constraints using the 2-norm.

In the nonpolyhedric setting, the ability to directly obtain an explicit formula for the generalized derivative of the normal cone is much more difficult. For this reason, we restrict ourselves to a class of nonpolyhedric convex sets described by nonsmooth convex functions, namely, we consider

$$M := \left\{ y \in H_0^1(\Omega) \,|\, |\nabla y|_2 \le \psi \text{ a.e.}\, \Omega \right\},$$

where, as before, $\nabla : H_0^1(\Omega) \to G(\Omega)$ with $G(\Omega) := \nabla(H_0^1(\Omega)) \subset L^2(\Omega)^l$, a closed (Hilbert) subspace of $L^2(\Omega)^l$. Once again, we have the following data assumptions with M as above:

- $Y := H_0^1(\Omega)$;
- $\mathcal{U} := L^2(\Omega)$;
- $J : L^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, Fréchet differentiable;
- $\psi \in L^\infty(\Omega)$ has a lower bound, $\underline{\psi} \in \mathbb{R}_+ \setminus \{0\}$, i.e., $\psi \ge \underline{\psi} > 0$.
- $\nabla : H_0^1(\Omega) \to G(\Omega)$, where $G(\Omega) := \nabla(H_0^1(\Omega))$, i.e., the image space of the gradient.

As it significantly simplifies the computations, we reformulate M as

$$M = \left\{ y \in H_0^1(\Omega) \,|\nabla y \in K_\psi \right\}, \tag{14}$$

where

$$K_\psi := \left\{ \mathbf{z} \in G(\Omega) \,\middle|\, |\mathbf{z}|_2^2 \le \psi^2 \text{ a.e.}\, \Omega \right\}.$$

Since ∇, as defined here, is a surjective bounded linear operator, we need only calculate T_{K_ψ} and N_{K_ψ} in order to derive formulae for T_M and N_M.

By using a calculus rule related to a certain type of second-order directional derivative of lower-semicontinuous functions, we are able to derive new formulae for the contingent derivatives of the normal cone and solution mappings. In order to state our result, we first need to introduce a few important concepts gathered from the literature.

DEFINITION 6.1 (Mosco epiconvergence). *Let $\{\varphi_t\}$ be a family of functions from a Banach space X into the extended reals $\bar{\mathbb{R}}$ parameterized by $t > 0$ and $\varphi : X \to \bar{\mathbb{R}}$. Then the family φ_t is said to Mosco epiconverge to φ as $t \to 0^+$ if for all sequences $t_n \to 0^+$ and every $x \in X$, the following two conditions hold:*

$$\forall x_n \rightharpoonup x, \quad \varphi(x) \leq \liminf_n \varphi_{t_n}(x_n), \tag{15}$$

$$\exists x_n \to x, \quad \varphi(x) \geq \limsup_n \varphi_{t_n}(x_n). \tag{16}$$

For more on this and related types of variational convergence, we refer the reader to [3]. In the next definition, we will use second-order differential quotients associated with some proper convex lower-semicontinuous function $f : X \to \bar{\mathbb{R}}$, where X is again some arbitrary Banach space. We assume f is finite at $x \in X$, $x^* \in X^*$, and $h \in X$ arbitrary. The so-called *second-order difference quotient* associated with f is then defined by

$$(\Delta_t^2 f)_{x,x^*}(h) := \frac{f(x+th) - f(x) - t\langle x^*, x\rangle}{\frac{1}{2}t^2}.$$

DEFINITION 6.2 (second-order Mosco epiderivatives). *Let $f : X \to \bar{\mathbb{R}}$ be a proper convex lower-semicontinuous function, let X be a Banach space, and let $x^* \in X^*$. If the family of associated second-order difference quotients $(\Delta_t^2 f)_{x,x^*}$ Mosco epiconverges to some function φ as $t \to 0^+$ with $\varphi(0) \neq -\infty$, then f is said to be twice Mosco epidifferentiable at x relative to x^*. Here, φ represents the second-order Mosco epiderivative of f at x relative to x^*, which we denote by f''_{x,x^*}.*

Second-order Mosco epiderivatives were introduced by Rockafellar for extended real-valued functionals from $\mathbb{R}^n$ in [41], and there is a compendium of results for finite dimensional objects in [43]. Some important references for the infinite dimensional setting include, but are by no means limited to, [13, 24, 29].

We can now state the following important calculus rule, which forms the basis for our interest in second-order Mosco epiderivatives.

THEOREM 6.3 (Do [13, Theorem 3.9]). *Let $f : X \to \bar{\mathbb{R}}$ be a proper convex lower-semicontinuous function, let X be a reflexive Banach space, and let $f(x)$ be finite. Then the following statements are equivalent:*

- *f is twice Mosco epidifferentiable at $x \in X$ relative to $x^* \in X^*$.*
- *The subdifferential ∂f is proto-differentiable at $(x, x^*) \in \operatorname{gph} \partial f$.*

In addition, it holds that

$$\partial \left(\frac{1}{2} f''_{x,x^*}\right)(h) = P\partial f[(x, x^*)](h), \quad h \in X. \tag{17}$$

Upon referring to the results in Chapter 13 of [43] as well as Theorem 7 in [30], one notes the presence of a second-order term in nonpolyhedral settings. As may be expected by the reader, the derivation of such results requires an explicit structure of the feasible set. Additionally, one also needs a constraint qualification in order to guarantee that the normals $v \in N_M(y)$ have a specific structure.

Before presenting our main result concerning the second-order epiderivative of the indicator function associated with M, we explicitly calculate the necessary cones for our setting.

DEFINITION 6.4. *For $y \in M$, where M is defined as in* (14)*, we define the active set by*

$$\mathcal{A}(y) := \left\{ x \in \Omega \,\middle|\, |\nabla y(x)|_2^2 \leq \psi^2(x) \right\}.$$

Accordingly, we define the inactive set by

$$\mathcal{I}(y) := \Omega \setminus \mathcal{A}(y).$$

PROPOSITION 6.5 (the tangent and normal cones to K_ψ). *Let K_ψ be defined as above and assume $\mathbf{z} \in K_\psi$. Then*

$$T_{K_\psi}(\mathbf{z}) = \{ \mathbf{h} \in G(\Omega) \,|\, \mathbf{z} \cdot \mathbf{h} \leq 0 \ a.e.\ \mathcal{A}(\mathbf{z}) \}$$

and

$$N_{K_\psi}(\mathbf{z}) = \left\{ 2\lambda \mathbf{z} \in G(\Omega) \,\middle|\, \lambda \in L^2(\Omega) : \lambda \geq 0 \ a.e.\ \mathcal{A}(\mathbf{z}), \lambda = 0 \ a.e.\ \mathcal{I}(\mathbf{z}) \right\}.$$

Proof. Let $\mathbf{h} \in T_{K_\psi}(\mathbf{z})$. Then there exist sequences $t_k \to 0^+$ and $\mathbf{h}_k \to \mathbf{h}$ in $G(\Omega)$ such that $|\mathbf{z} + t_k \mathbf{h}_k|_2^2 \leq \psi^2$ a.e. Ω. By rearranging terms and passing to the limit, we observe that $\mathbf{z} \cdot \mathbf{h} \leq 0$ a.e. $\mathcal{A}(\mathbf{z})$. Thus, the inclusion "$\subseteq$" holds. For the reverse direction, we again use an argument based on Lebesgue's dominating convergence theorem. Indeed, by defining the family of $G(\Omega)$-vector fields

$$\mathbf{p}_\tau := \frac{\Pi_{K_\psi}(\mathbf{z} + \tau \mathbf{h}) - \mathbf{z}}{\tau},$$

it holds that for all $\tau > 0$, $\mathbf{p}_\tau \in R_{K_\psi}(\mathbf{z})$. Moreover, for almost every x, $\mathbf{p}_\tau(x) \to \mathbf{h}(x)$ as $\tau \to 0$ and since $G(\Omega)$ is a Hilbert space and K_ψ is closed and convex, the metric projection is nonexpansive, i.e., $|\mathbf{p}_\tau(x)| \leq |\mathbf{h}(x)|$. As $G(\Omega)$ is a closed subspace of $L^2(\Omega)^l$ in the $L^2(\Omega)^l$-norm, it holds that $\mathbf{p}_\tau \to \mathbf{h}$ in $G(\Omega)$. Therefore, the reverse inclusion holds.

We now characterize the normal cone. By definition,

$$N_{K_\psi}(\mathbf{z}) = \left\{ \mu^* \in G(\Omega)^* \,\middle|\, \langle \mu^*, \mathbf{h} \rangle_{G^*,G} \leq 0 \ \forall \mathbf{h} \in T_{K_\psi}(\mathbf{z}) \right\}.$$

Nevertheless, since $G(\Omega)$ is a Hilbert space with $L^2(\Omega)^l$ inner product, we can associate with every $\mu^* \in N_{K_\psi}(\mathbf{z})$ its $G(\Omega)$-counterpart μ. Therefore, we view the normal cone as follows:

$$N_{K_\psi}(\mathbf{z}) = \left\{ \mu \in G(\Omega) \,\middle|\, (\mu, \mathbf{h})_{L^2(\Omega)^l} \leq 0 \ \forall \mathbf{h} \in T_{K_\psi}(\mathbf{z}) \right\}.$$

Let $\mathbf{z} \in K_\psi$. By definition, $\mathbf{z} \in G(\Omega)$. In addition, let $\lambda \in L^2(\Omega)$ such that $\lambda \geq 0$ a.e. $\mathcal{A}(\mathbf{z})$ and $\lambda = 0$ a.e. $\mathcal{I}(\mathbf{z})$. Since $\mathbf{z} \in K_\psi$, it holds by definition that $|\mathbf{z}|_2^2 \leq \psi^2$ a.e. Ω. Therefore, given $\psi \in L^\infty(\Omega)$, it follows that $\mathbf{z} \in L^\infty(\Omega)^l$, in which case we have $\lambda \mathbf{z} \in L^2(\Omega)^l$. Moreover, we note that since ψ has a strictly positive lower bound, $|\mathbf{z}| > 0$ a.e. $\mathcal{A}(\mathbf{z})$.

Without loss of generality, suppose $\lambda \mathbf{z} \neq 0 \in L^2(\Omega)^l$. This implies in particular that the Lebesgue measure of the strongly active set $\mathcal{A}(\mathbf{z})_+$ is strictly positive. Indeed, by assuming otherwise, we see that $\lambda = 0$ a.e. Ω; thus $\lambda \mathbf{z} = 0$ a.e. Ω (componentwise). Clearly, $|\lambda \mathbf{z}|_2^2 = 0$ a.e. Ω as well. Thus, $\lambda \mathbf{z} = 0 \in L^2(\Omega)^l$.

Define the set

$$N^0(\mathbf{z}) := \left\{ 2\lambda \mathbf{z} \in G(\Omega) \,\middle|\, \lambda \in L^2(\Omega) : \lambda \geq 0 \text{ a.e. } \mathcal{A}(\mathbf{z}), \lambda = 0 \text{ a.e. } \mathcal{I}(\mathbf{z}) \right\}.$$

By the previous argument and convexity of K_ψ in $G(\Omega)$, it holds that $N_{K_\psi}(\mathbf{z}) \supseteq N^0(\mathbf{z})$. To see that $N^0(\mathbf{z})$ is convex in $G(\Omega)$, let $\alpha \in (0,1)$ and $\mathbf{v}_1, \mathbf{v}_2 \in N^0(\mathbf{z})$. Then since

$$\alpha \mathbf{v}_1(x) + (1-\alpha)\mathbf{v}_2(x) = 2(\alpha\lambda_1 + (1-\alpha)\lambda_2)\mathbf{z}$$

and the set

$$\left\{ \lambda \in L^2(\Omega) \,|\, \lambda \geq 0 \text{ a.e. } \mathcal{A}(\mathbf{z}), \lambda = 0 \text{ a.e. } \mathcal{I}(\mathbf{z}) \right\}$$

is convex, it holds that $\alpha \mathbf{v}_1 + (1-\alpha)\mathbf{v}_2 \in N^0(\mathbf{z})$.

Next, we demonstrate that $N^0(\mathbf{z})$ is closed in $G(\Omega)$. Let $\mathbf{v}_n \in N^0(\mathbf{z})$ such that $\mathbf{v}_n \to \mathbf{v}$ in $L^2(\Omega)^l$. By definition, there exist $\lambda_n \in L^2(\Omega)$, where $\lambda_n \geq 0$ a.e. on $\mathcal{A}(\mathbf{z})$ and $\lambda_n = 0$ a.e. on $\mathcal{I}(\mathbf{z})$ such that $\mathbf{v}_n = 2\lambda_n \mathbf{z}$.

Given $\mathbf{v}_n \to \mathbf{v}$ in $L^2(\Omega)^l$, there exists a positive constant C such that $||\mathbf{v}_n||^2_{L^2(\Omega)^l} \leq C$. Hence,

$$C \geq ||\mathbf{v}_n||^2_{L^2(\Omega)^l} = 4\int_\Omega \lambda_n^2 |\mathbf{z}|^2 dx = 4\int_{\mathcal{A}(\mathbf{z})} \lambda_n^2 \psi^2 dx \geq 4\underline{\psi}^2 \int_{\mathcal{A}(\mathbf{z})} \lambda_n^2 dx = 4\underline{\psi}^2 ||\lambda_n||^2_{L^2(\Omega)}.$$

Therefore, there exists $\hat{\lambda} \in L^2(\Omega)$ and a subsequence $\{n_k\}_{k=1}^\infty \subset \{n\}_{n=1}^\infty$ such that $\lambda_{n_k} \rightharpoonup \hat{\lambda}$ in $L^2(\Omega)$. Moreover, since the set

$$U = \left\{ \lambda \in L^2(\Omega) \,|\, \lambda \geq 0 \text{ a.e. } \mathcal{A}(\mathbf{z}), \lambda = 0 \text{ a.e. } \mathcal{I}(\mathbf{z}) \right\}$$

is closed and convex in $L^2(\Omega)$, and thus, weakly closed in $L^2(\Omega)$, $\hat{\lambda} \in U$ as well.

Continuing, we use the fact that $G(\Omega)$ is a Hilbert space with the $L^2(\Omega)^l$ inner product to rewrite $\mathbf{v}$. Indeed, by letting $\alpha, \beta \in \mathbb{R} \setminus \{0\}$, we can write $\mathbf{v} = 2\alpha\hat{\lambda}\mathbf{z} + \beta\mathbf{w}$, where $\mathbf{w} \in \{\hat{\lambda}\mathbf{z}\}^\perp$. Consider then that

$$\begin{aligned} ||\mathbf{v}_n - \mathbf{v}||^2_{L^2(\Omega)^l} &= ||\mathbf{v}_n||^2_{L^2(\Omega)} - 2(\mathbf{v}_n, \mathbf{v})_{L^2(\Omega)^l} + ||\mathbf{v}||^2_{L^2(\Omega)^l} \\ &= ||2\lambda_n \mathbf{z}||^2_{L^2(\Omega)} - 2(2\lambda_n \mathbf{z}, \alpha\hat{\lambda}\mathbf{z} + \beta\mathbf{w})_{L^2(\Omega)^l} + ||\alpha\hat{\lambda}\mathbf{z} + \beta\mathbf{w}||^2_{L^2(\Omega)^l} \\ &= ||2\lambda_n \mathbf{z} - \alpha\hat{\lambda}\mathbf{z}||^2_{L^2(\Omega)^l} - 2\beta(2\lambda_n \mathbf{z}, \mathbf{w})_{L^2(\Omega)^l} + ||\beta\mathbf{w}||^2_{L^2(\Omega)^l}. \end{aligned}$$

Then since $\lambda_{n_k} \rightharpoonup \hat{\lambda}$ in $L^2(\Omega)$ and $\mathbf{w}$ is orthogonal to $2\hat{\lambda}\mathbf{z}$,

$$-2\beta(2\lambda_{n_k}\mathbf{z}, \mathbf{w})_{L^2(\Omega)^l} = -2\beta\int_{\mathcal{A}(\mathbf{z})} 2\lambda_{n_k}\mathbf{z} \cdot \mathbf{w} dx \to -2\beta\int_{\mathcal{A}(\mathbf{z})} 2\hat{\lambda}\mathbf{z} \cdot \mathbf{w} dx = 0.$$

However, this implies that $||2\lambda_{n_k}\mathbf{z} - \alpha\hat{\lambda}\mathbf{z}||^2_{L^2(\Omega)^l} + ||\beta\mathbf{w}||^2_{L^2(\Omega)^l} \to 0$, which can only hold when $||\beta\mathbf{w}||^2_{L^2(\Omega)^l} = 0$ and $||\lambda_{n_k}\mathbf{z} - \alpha\hat{\lambda}\mathbf{z}||^2_{L^2(\Omega)^l} \to 0$. Finally, we note that

$$||2\lambda_{n_k}\mathbf{z} - \alpha\hat{\lambda}\mathbf{z}||^2_{L^2(\Omega)^l} = \int_{\mathcal{A}(\mathbf{z})} \left|(2\lambda_{n_k} - \alpha\hat{\lambda})\mathbf{z}\right|_2^2 dx \geq \underline{\psi}^2 \int_{\mathcal{A}(\mathbf{z})} \left|(2\lambda_{n_k} - \alpha\hat{\lambda})\right|_2^2 dx.$$

Hence, $2\lambda_n \to \alpha\hat{\lambda}$ in $L^2(\Omega)$, and $\mathbf{v}$ has the form $2\tilde{\lambda}\mathbf{z}$, where $\tilde{\lambda} \in U$ such that $\tilde{\lambda} = \alpha\hat{\lambda}$. Therefore, $N^0(\mathbf{z})$ is closed.

Finally, suppose there exists $\mathbf{v}^* \in N_{K_\psi}(\mathbf{z})$ such that $\mathbf{v}^* \notin N^0(\mathbf{z})$. Then there exists a $\mathbf{u}^* \in G(\Omega)^*$ strongly separating $\mathbf{v}^*$ from $N^0(\mathbf{z})$ (see, e.g., TVS II.38, Prop. 4 of [11]). That is,

$$\langle \mathbf{u}^*, \mathbf{v}^* \rangle_{G^*,G} > 0, \quad \langle \mathbf{u}^*, \mathbf{v} \rangle_{G^*,G} \le 0 \; \forall \mathbf{v} \in N^0(\mathbf{z}).$$

Using again the fact that $G(\Omega)$ is a Hilbert space, we immediately identify $\mathbf{u}^*$ with its counterpart $\mathbf{u} \in G(\Omega)$ so that the previous relations become

$$\exists \mathbf{u} \in G(\Omega) : (\mathbf{u}, \mathbf{v}^*)_{L^2(\Omega)^l} > 0, \quad (\mathbf{u}, \mathbf{v})_{L^2(\Omega)^l} \le 0 \; \forall \mathbf{v} \in N^0(\mathbf{z}).$$

The first inequality, being strict, implies that $\mathbf{u} \notin T_{K_\psi}(\mathbf{z})$. However,

$$(\mathbf{u}, \mathbf{v})_{L^2(\Omega)^l} \le 0 \; \forall \mathbf{v} \in N^0(\mathbf{z}) \Leftrightarrow \int_{\mathcal{A}(\mathbf{z})} 2\lambda \mathbf{z} \cdot \mathbf{u} dx \le 0 \quad \forall \lambda \in U.$$

But then $\mathbf{z} \cdot \mathbf{u} \le 0$ a.e. $\mathcal{A}(\mathbf{z})$, a contradiction. Therefore, $N_{K_\psi}(\mathbf{z}) \setminus N^0(\mathbf{z}) = \emptyset$, as was to be shown. ∎

Using the surjectivity of $\nabla : H_0^1(\Omega) \to G(\Omega)$, we immediately obtain our next result.

PROPOSITION 6.6 (the tangent and normal cones to M). *Let $y \in M$, where M is defined as in* (14). *Then*

$$T_M(y) = \left\{ d \in H_0^1(\Omega) \,\middle|\, \nabla y \cdot \nabla d \in T_{K_\psi}(\nabla y) \right\}$$

and

$$\begin{aligned} &N_M(y) \\ &\quad = \left\{ -2\mathrm{div}\,(\lambda \nabla y) \in H^{-1}(\Omega) \,\middle|\, \lambda \in L^2(\Omega) : \lambda \ge 0 \text{ a.e. } \mathcal{A}(\nabla y), \lambda = 0 \text{ a.e. } \mathcal{I}(\nabla y) \right\}. \end{aligned}$$

In addition, the critical cone to M at y for any normal $v \in N_M(y)$ becomes

$$\begin{aligned} \mathcal{K}(y, v) &= T_M(y) \cap \{v\}^\perp \\ &= \left\{ d \in H_0^1(\Omega) \,\middle|\, \nabla y \cdot \nabla d \le 0 \text{ a.e. } \mathcal{A}^0(\nabla y), \nabla y \cdot \nabla d = 0 \text{ a.e. } \mathcal{A}^+(\nabla y) \right\}, \end{aligned}$$

where

$$\mathcal{A}^0(\nabla y) := \{ x \in \mathcal{A}(\nabla y) \,|\, \lambda(x) = 0 \}, \quad \mathcal{A}^+(\nabla y) := \{ x \in \mathcal{A}(\nabla y) \,|\, \lambda(x) > 0 \},$$

i.e., the weakly (biactive) and the strongly active sets, respectively.

Proof. The same argument as used in the proof of Proposition 5.4 applies to the derivation of the normal and tangent cones. For the critical cone, let $d \in T_M(y) \cap \{v\}^\perp$. Then

$$\nabla y \cdot \nabla d \le 0 \text{ a.e. } \mathcal{A}(\nabla y) \wedge \langle v, d \rangle_{H^{-1}, H_0^1} = 0.$$

Using the characterization of the normals v, it holds for all d in the critical cone that

$$\begin{aligned} \langle v, d \rangle_{H^{-1}, H_0^1} &= \langle -2\mathrm{div}\, \lambda \nabla y, d \rangle_{H^{-1}, H_0^1} = (2\lambda \nabla y, \nabla d)_{L^2(\Omega)^l} \\ &= 2 \int_{\mathcal{A}(\nabla y)} \lambda \nabla y \cdot \nabla d dx = 2 \int_{\mathcal{A}^+(\nabla y)} \lambda \nabla y \cdot \nabla d dx = 0. \end{aligned}$$

Hence, the inclusion "$\subseteq$" holds, whereas the reverse direction is trivial and follows via a direct verification. Note that $y \in M$ implies that $\nabla y \in L^\infty(\Omega)^l$ (cf. the proof of Proposition 6.5). Therefore, the function $\lambda \nabla y \cdot \nabla d$ is integrable. $\square$

Given Proposition 6.6, we see that any normal $v \in N_M(y)$ has the structure $-2\mathrm{div}\,(\lambda \nabla y)$, where λ is a type of Lagrange multiplier. We are now ready for our main result.

THEOREM 6.7 (the second-order Mosco epiderivative of I_M). *Let $y \in M$, where M is defined as above. For $v \in \partial I_M(y) = N_M(y)$, if there exists $\lambda \in L^\infty(\Omega)$ with $\lambda \geq 0$ a.e. on $\mathcal{A}(y)$ and $\lambda = 0$ a.e. on $\mathcal{I}(y)$ such that $v = -2\mathrm{div}\,(\lambda\nabla y) \in H^{-1}(\Omega)$, then the indicator function $I_M : Y \to \bar{\mathbb{R}}$ is twice Mosco epidifferentiable at y relative to v and the second-order Mosco epiderivative is characterized as follows:*

$$(I_M^{''})_{y,v}(d) = \begin{cases} Q(d;\lambda), & d \in T_M(y) \cap \{v\}^\perp, \\ \infty & \textit{otherwise}. \end{cases} \tag{18}$$

Here, $Q(\cdot;\lambda) : H_0^1(\Omega) \to \mathbb{R}$ is the convex continuous functional defined by

$$Q(d;\lambda) = 2\int_{\mathcal{A}(y)} \lambda |\nabla d|_2^2 dx = 2\langle -\mathrm{div}\,(\lambda \nabla d), d\rangle.$$

In other words,

$$(I_M^{''})_{y,v}(d) = Q(d;\lambda) + I_{\mathcal{K}(y,v)}(d),$$

where

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

Proof. The fact that v has such a form follows directly from Proposition 6.6. Moreover, it is easy to see that the indicator function of a nonempty closed convex set is proper, convex, and lower-semicontinuous, and for any $y \in M$, $I_M(y) = 0$, i.e., I_M is finite.

Begin by letting $t_n \to 0^+$ be an arbitrary sequence of scalars converging to zero from above and let $d_n \rightharpoonup_Y d$ for an arbitrary $d \in Y$. By the definition of the indicator function, the lower limit

$$\liminf_n \frac{I_M(y + t_n d_n) - I_M(y) - t_n\langle v, d_n\rangle}{t_n^2/2} \tag{19}$$

will be equal to infinity unless there exists a subsequence of d_n or some large $N_0 \in \mathbb{N}$ such that $y + t_n d_n \in M$ for all $n \geq N_0$. Therefore, suppose such a sequence exists. Given any $v' \in N_M(y)$, the closure and convexity of M imply that $\langle v', y' - y\rangle \leq 0$ for all $y' \in M$, in which case it follows from the assumption that $\langle v', d_n\rangle \leq 0$. Since $d_n \rightharpoonup_Y d$, we observe that $\langle v', d\rangle \leq 0$ for all $v' \in N_M(y)$. Then from the convexity of M, we deduce that $d \in T_M(y)$. In addition, we see that the second-order difference quotients are all nonnegative and in fact reduce to

$$\frac{-2\langle v, d_n\rangle}{t_n} \geq 0.$$

Hence, if $\langle v, d_n\rangle$ does not converge to zero, $t_n \to 0^+$ implies that the lower limit tends to infinity, which leads to the following inequality:

$$\liminf_n \frac{I_M(y + t_n d_n) - I_M(y) - t_n\langle v, d_n\rangle}{t_n^2/2} \geq \begin{cases} 0, & d \in T_M(y) \cap \{v\}^\perp, \\ \infty & \text{otherwise}. \end{cases} \tag{20}$$

Since t_n was arbitrarily chosen, (20) holds for all sequences $t_n \to 0^+$ and $d_n \rightharpoonup d$. Though the right-hand side of (20) may seem like a good candidate for the second-order Mosco epiderivative, we can use our knowledge of v along with the given structure to obtain a better lower estimate. In what follows we continue with the assumption that $t_n \to 0^+$ and $d_n \rightharpoonup_Y d$ with $y + t_n d_n \in M$ and $d \in \mathcal{K}(\bar{y}, \bar{v})$.

By letting $v = -2\mathrm{div}\,(\lambda \nabla y)$ as in Proposition 6.6, the second-order difference quotients can be equivalently written as follows:

$$\frac{I_M(y + t_n d_n) - I_M(y) - t_n \langle -2\mathrm{div}\,(\lambda\nabla y), d_n\rangle}{\frac{1}{2}t_n^2},$$

which under the assumptions of the theorem reduces to the right-hand side of the following relation:

$$\begin{aligned} \frac{I_M(y + t_n d_n) - I_M(y) - t_n \langle -2\mathrm{div}\,(\lambda\nabla y), d_n\rangle}{\frac{1}{2}t_n^2} &= \frac{-2\langle -2\mathrm{div}\,(\lambda\nabla y), d_n\rangle}{t_n} \\ &= \frac{-2(2\lambda\nabla y, \nabla d_n)_{L^2}}{t_n}. \end{aligned} \tag{21}$$

For $g := |\cdot|_2^2$, we let

$$\Delta_{t_n} g(y)(d_n) = \frac{g(y + t_n d_n) - g(y)}{t_n},$$

denote the first-order difference quotient of g. Then by adding "zero" to the reduced quotient in (21), we further transform the second difference quotients to

$$\frac{-2(\lambda, \Delta_{t_n} g(y)(d_n))_{L^2}}{t_n} + \frac{(\lambda, \Delta_{t_n} g(y)(d_n))_{L^2} - (2\lambda\nabla y, \nabla d_n)_{L^2}}{\frac{1}{2}t_n}.$$

Consider now that the second summand can be written

$$\begin{aligned} &\frac{2}{t_n}\int_{\mathcal{A}(y)} \lambda(\Delta_{t_n} g(y)(d_n) - 2\nabla y \cdot \nabla d_n)dx \\ &\quad = \frac{2}{t_n}\int_{\mathcal{A}(y)} \lambda\left(\frac{|\nabla y|_2^2 + 2t_n \nabla y \cdot \nabla d_n + t_n^2 |\nabla d_n|_2^2 - |\nabla y|_2^2}{t_n} - 2\nabla y \cdot \nabla d_n\right)dx \\ &\qquad\qquad = 2\int_{\mathcal{A}(y)} \lambda |\nabla d_n|_2^2 dx = Q(d_n; \lambda). \end{aligned}$$

Since $Q(\cdot;\lambda)$ is convex and continuous from Y to $\mathbb{R}$, it is also weakly lower-semicontinuous. Therefore, for any $d_n \rightharpoonup_Y d$, $\liminf_n Q(d_n;\lambda) \geq Q(d;\lambda)$. Conversely, we have from the assumed feasibility that $|\nabla y + t_n \nabla d_n|_2^2 \leq \psi^2$ a.e. on Ω. Thus,

$$g(y + t_n d_n) - g(y) = 2t_n \nabla y \cdot \nabla d_n + t_n^2 |\nabla d_n|_2^2 \leq 0$$

a.e. on the active set $\mathcal{A}(y)$. It follows then that

$$\frac{-2\langle \lambda, \Delta_{t_n} g(y)(d_n)\rangle}{t_n} = \frac{-2}{t_n}\int_{\mathcal{A}(y)} \lambda \Delta_{t_n} g(y)(d_n)dx \geq 0.$$

Using these new observations we have (starting from (19)) that

$$\begin{aligned}\liminf_n \frac{I_M(y+t_n d_n) - I_M(y) - t_n\langle v, d_n\rangle}{t_n^2/2} \\ \geq \liminf_n Q(d_n;\lambda) + \liminf_n \frac{-2\langle \lambda, \Delta_{t_n} g(y)(d_n)\rangle}{t_n} \\ \geq \begin{cases} Q(d;\lambda), & d \in \mathcal{K}(\bar{y},\bar{v}), \\ +\infty & \text{otherwise.}\end{cases}\end{aligned} \tag{22}$$

Recalling Definition 6.2, we see that in order to complete the proof, we need to show for all $t_n \to 0^+$ that there exists a strongly converging sequence $d_n \to_Y d$ such that

$$\limsup_n \frac{I_M(y+t_n d_n) - I_M(y) - t_n\langle v, d_n\rangle}{t_n^2/2} \leq \begin{cases} Q(d;\lambda), & d \in \mathcal{K}(\bar{y},\bar{v}), \\ +\infty & \text{otherwise.}\end{cases}$$

In what follows, let $t_n \to 0^+$ be arbitrary. Clearly, if $d \notin T_M(y) \cap \{v\}^\perp$, then the inequality will always hold. Therefore, we need only construct sequences for $d \in T_M(y) \cap \{v\}^\perp$.

Since both $T_M(y)$ and $\{v\}^\perp$ are strongly closed in Y, their intersection is as well. Therefore, for all $d \in T_M(y) \cap \{v\}^\perp$ there exists a strongly convergent sequence $\delta_n \to d$ with $\delta_n \in T_M(y) \cap \{v\}^\perp$ for all n such that $\langle v, \delta_n\rangle = 0$ and, by the definition of $T_M(y)$, sequences $\tau_k^n \to 0^+$ and $\delta_k^n \to \delta_n$ such that $y + \tau_k^n \delta_k^n \in M$ for all k and each n. From the convexity of M, we infer that for any $t \in [0, \tau_k^n]$ we have

$$y + t\delta_k^n = \left(1 - \frac{t}{\tau_k^n}\right) y + \frac{t}{\tau_k^n}(y + \tau_k^n \delta_k^n) \in M \ \forall k, \forall n. \tag{23}$$

Moreover, we deduce that $\langle v, \delta_k^n\rangle \to 0$ as $k \to \infty$ and similar to before, we see that $\langle v, \delta_k^n\rangle \leq 0$. Hence, for some $\varepsilon > 0$, there exists $K_n \in \mathbb{N}$ for each n such that

1. $|2\langle v, \delta_k^n\rangle| \leq t_{n_{min}}^{1+\varepsilon}$ for all $k \geq K_n$, where $n_{min} := \operatorname{argmin}_n \{t_1, \ldots, t_n\}$.
2. $||\delta_n - \delta_k^n|| \leq t_n$ for all $k \geq K_n$.
3. $\tau_k^n \leq f_n$ for all $k \geq K_n$, with f_n arbitrary such that $f_n \to 0^+$ monotonely.

We now build our strongly converging sequence d_n. Begin by fixing $m_1 \in \mathbb{N}$ and define

$$\varepsilon_{m_1} := \min_{1 \leq i \leq m_1} \tau_{K_i}^i.$$

Given $t_n \to 0^+$, there exists an $N(\varepsilon_{m_1}) \in \mathbb{N}$ such that for all $n \geq N(\varepsilon_{m_1})$

$$t_n \leq \varepsilon_{m_1}.$$

By the definition of ε_{m_1}, it also holds that

$$t_n \leq \tau_{K_{m_1}}^{m_1}.$$

For all $n < N(\varepsilon_{m_1})$, set $d_n := \delta_{K_1}^1$. Now define j_1 to be the smallest index such that

$$t_{N(\varepsilon_{m_1})} > \tau_{K_{m_1+j_1}}^{m_1+j_1},$$

and define $l_1 \geq 1$ to be the first index such that

$$t_{N(\varepsilon_{m_1})+l_1} \leq \tau_{K_{m_1+j_1}}^{m_1+j_1}.$$

By the third assumption on K_n, j_1 exists since $\tau^n_{K_n} \to 0^+$, with n, and l_1 exists since $t_n \to 0^+$. Then using these indices, we set $d_n = \delta^{m_1}_{K_{m_1}}$ for all $n = N(\varepsilon_{m_1}), \ldots, N(\varepsilon_{m_1}) + l_1 - 1$.

Given for all $n = N(\varepsilon_{m_1}) + i$ with $i = 0, \ldots, l_1 - 1$, $t_n \in [0, \tau^{m_1}_{K_{m_1}}]$ and $d_n = \delta^{m_1}_{K_{m_1}}$, it holds that $y + t_n d_n \in M$ (cf. (23)). Thus, $I_M(y + t_n d_n) = 0$.

At this point we define $m_2 := m_1 + j_1$ and repeat the process described above with m_2 in place of m_1. Clearly, $\varepsilon_{m_2} \le \varepsilon_{m_1}$ and $N(\varepsilon_{m_2}) \ge N(\varepsilon_{m_1})$. If $N(\varepsilon_{m_1}) + l_1 - 1 < N(\varepsilon_{m_2})$, then set $d_n = \delta^{m_1}_{K_{m_1}}$ for all n such that $N(\varepsilon_{m_1}) + l_1 - 1 \le n < N(\varepsilon_{m_2})$. As $l_1 \ge 1$, $t_n \le \varepsilon_{m_1}$ so that $t \in [0, \tau^{m_1}_{K_{m_1}}]$ and $y + t_n d_n \in M$ still holds. Also note that the case "$N(\varepsilon_{m_2}) < N(\varepsilon_{m_1}) + l_1 - 1$" cannot happen. Indeed, this would require, for all $n = N(\varepsilon_{m_2}), \ldots, N(\varepsilon_{m_1}) + l_1 - 1$, that $\tau^{m_2}_{K_{m_2}} < t_n \le \varepsilon_{m_2}$, but $\varepsilon_{m_2} \le \tau^{m_2}_{K_{m_2}}$ by definition, a contradiction. The rest continues as before.

To see that the process continues indefinitely, let $p \ge 1$ and consider that the convergence of $\tau^n_{K_n} \to 0^+$ ensures the existence of a $j_p \in \mathbb{N}$ such that $t_{N(\varepsilon_{m_p})} > \tau^{m_p + j_p}_{K_{m_p + j_p}}$. With j_p fixed, we now look to increase the number n larger than $N(\varepsilon_{m_p})$. We do this by checking if $t_{N(\varepsilon_{m_p})+i} > \tau^{m_p+j_p}_{K_{m_p+j_p}}$ for $i = 1, 2, \ldots.$ As j_p is fixed, so is $\tau^{m_p+j_p}_{K_{m_p+j_p}}$. Therefore, the convergence of $t_n \to 0^+$ implies the existence of some l_p such that $t_{N(\varepsilon_{m_p})+l_p} \le \tau^{m_p+j_p}_{K_{m_p+j_p}}$, by definition, we define m_{p+1} and continue as before. This ensures that the process is perpetual.

Summarizing, we describe the construction via the following diagram:

$$d_n = \underbrace{\delta^1_{K_1}, \ldots, \delta^1_{K_1}}_{1 \le n < N(\varepsilon_{m_1})} \mid \underbrace{\delta^{m_1}_{K_{m_1}}, \ldots, \delta^{m_1}_{K_{m_1}}}_{N(\varepsilon_{m_1}) \le n < N(\varepsilon_{m_2})} \mid \underbrace{\delta^{m_2}_{K_{m_2}}, \ldots, \delta^{m_2}_{K_{m_2}}}_{N(\varepsilon_{m_2}) \le n < N(\varepsilon_{m_3})} \mid \ldots,$$

that is, d_n is a subsequence of $\delta^n_{K_n}$. Furthermore, we have for all $\mu > 0$ and large n that

$$||d - \delta^n_{K_n}|| \le ||d - \delta_n|| + ||\delta_n - \delta^n_{K_n}|| \le \mu + t_n \le 2\mu,$$

so that $\delta^n_{K_n} \to d$ as $n \to \infty$, and thus d_n as well.

Due to the fact that $y + t_n d_n \in M$, recall that the second-order difference quotients can be transformed into the sum

$$Q(d_n; \lambda) + \frac{-2\langle \lambda, \Delta_{t_n} g(y)(d_n) \rangle}{t_n}.$$

Clearly, $|\nabla y + t_n \nabla d_n|_2^2 \ge |\nabla y|_2^2 + 2t_n \nabla y \cdot \nabla d_n$ a.e. on Ω. Therefore,

$$\begin{aligned}\frac{-2\langle \lambda, \Delta_{t_n} g(y)(d_n) \rangle}{t_n} &= \frac{-2}{t_n} \int_{\mathcal{A}(y)} \lambda \left(\frac{|\nabla y + t_n \nabla d_n|_2^2 - |\nabla y|_2^2}{t_n} \right) dx \\ &\le \frac{-2}{t_n} \int_{\mathcal{A}(y)} 2\lambda \nabla y \cdot \nabla d_n dx = \frac{-2\langle v, d_n \rangle}{t_n} \le t_n^{\varepsilon}.\end{aligned}$$

It follows that

$$\begin{aligned}&\limsup_{n\to\infty} \left[\frac{-2\langle v, d_n \rangle}{t_n} + Q(d_n; \lambda) \right] \\ &\qquad \le \limsup_{n\to\infty} \frac{-2\langle v, d_n \rangle}{t_n} + \limsup_{n\to\infty} Q(d_n; \lambda) = \limsup_{n\to\infty} Q(d_n; \lambda) = Q(d; \lambda).\end{aligned}$$

Here, the last equality follows from the continuity of $Q(\cdot,\lambda)$, which is demonstrated via the following implication:

$$d_n \to d \text{ in } H_0^1(\Omega) \Rightarrow |\nabla d_n|_2^2 \to |\nabla d|_2^2 \text{ in } L^1(\Omega).$$

Lastly, since $0 \in T_M(y) \cap \{v\}^\perp$ the function

$$\varphi(d) := \begin{cases} Q(d;\lambda), & d \in T_M(y) \cap \{v\}^\perp, \\ +\infty & \text{otherwise,} \end{cases}$$

equals zero at zero. Hence, φ is the Mosco epilimit of the second-order difference quotients and therefore amounts to the second-order Mosco epiderivative of I_M, as was to be shown. ∎

Remark 6.8. The reader should note the reliance of the proof on the fact that $\lambda \in L^\infty(\Omega)$. This is a crucial point which allows $Q(d;\lambda)$ to be well defined and provides the validity of the convergence arguments for the associated integrals. This represents a substantial difference from arguments carried out in finite dimensions, where no such consideration on λ is necessary. We refer the reader to the end of this paper for an example in which $\lambda \in L^\infty(\Omega)$.

We immediately obtain the following corollary.

COROLLARY 6.9 (the proto-derivative of N_M). *Let $y \in M$, where M is defined as above and $v \in N_M(y)$. If there exists $\lambda \in L^\infty(\Omega)$ with $\lambda \geq 0$ a.e. $\mathcal{A}(y)$ and $\lambda = 0$ a.e. $\mathcal{I}(y)$ such that $v = -2\mathrm{div}\,(\lambda\nabla y)$, then the normal cone operator is proto-differentiable and the following are equivalent:*

1. $w \in PN_M[(y,v)](d)$.
2. $w \in -2\mathrm{div}\,(\lambda\nabla d) + N_{\mathcal{K}(y,v)}(d)$.
3. $(w + 2\mathrm{div}\,(\lambda\nabla d), d) \in \mathcal{K}(y,v) \times [\mathcal{K}(y,v)]^- : \langle w + 2\mathrm{div}\,(\lambda\nabla d), d\rangle = 0$.

Here, $\mathcal{K}(y,v) := T_M(y) \cap \{v\}^\perp$, i.e., the critical cone.

Proof. The result follows from (17) via the formula (18) in Theorem 6.7. Indeed, due to the differentiability of the quadratic form as a function of d and the convexity of the critical cone, $\partial(Q(d;\lambda) + I_{\mathcal{K}(y,v)}(d)) = \nabla Q(d;\lambda) + N_{\mathcal{K}(y,v)}(d)$. Furthermore, since $\mathcal{K}(y,v)$ is a cone, we have from Lemma 4.2.5 in [5] that

$$u \in N_{\mathcal{K}(y,v)}(d) \Longleftrightarrow (d,u) \in \mathcal{K}(y,v) \times [\mathcal{K}(y,v)]^- : \langle u, d\rangle = 0.$$

The assertion then follows. ∎

The next result follows similarly to the polyhedric case.

COROLLARY 6.10 (the proto-derivative of S). *Let $(u,y) \in \mathrm{gph}\,S$, where $u = Ay + v$ with $v \in N_M(y)$. If there exists $\lambda \in L^\infty(\Omega)$ with $\lambda \geq 0$ a.e. $\mathcal{A}(y)$ and $\lambda = 0$ a.e. $\mathcal{I}(y)$ such that $v = -2\mathrm{div}\,(\lambda\nabla y)$, then S is proto-differentiable and the following are equivalent:*

1. $d \in PS[(u,y)](w)$.
2. $w \in Ad - 2\mathrm{div}\,(\lambda\nabla d) + N_{\mathcal{K}(y,v)}(d)$.
3. $w - Ad + 2\mathrm{div}\,(\lambda\nabla d) \in [\mathcal{K}(y,v)]^-$, $d \in \mathcal{K}(y,v)$, $\langle w - Ad + 2\mathrm{div}\,(\lambda\nabla d), d\rangle = 0$,

or equivalently,

$$T_{\mathrm{gph}\,S}(u,y) = \Big\{(w,d) \in Y^* \times Y \Big| \\ w - Ad + 2\mathrm{div}\,(\lambda\nabla d) \in \big[\mathcal{K}(y,v)\big]^-, d \in \mathcal{K}(y,v), \langle w - Ad + 2\mathrm{div}\,(\lambda\nabla d), d\rangle = 0\Big\}.$$

Here, $v := u - Ay$.

Proof. The proof is analogous to that of Theorem 4.6. The rest follows from Corollary 6.9. ◻

It remains to argue that $Ad - 2\mathrm{div}\,(\lambda \nabla d)$ is a coercive bounded linear operator in order to demonstrate that we have characterized not only the proto-derivative of S, but also the Hadamard directional derivative. Indeed, by letting $\tilde{A}\cdot := A\cdot - 2\mathrm{div}\,(\lambda\nabla\cdot)$, the proto-derivative of S in direction $w \in H^{-1}(\Omega)$ is characterized as the solution(s) to the following variational inequality:

$$\text{Find } d \in H_0^1(\Omega) : \langle \tilde{A}d - w, d' - d \rangle_{H^{-1},H_0^1} \geq 0 \ \forall d' \in \mathcal{K}(\bar{y}, \bar{v}).$$

As argued in Remark 4.4, if $\tilde{A}$ is coercive, bounded, and linear, then classical arguments can be applied to demonstrate that for any $w \in H^{-1}(\Omega)$, d is the unique solution of the variational inequality and $d(\cdot)$ is Lipschitz continuous on all of $H^{-1}(\Omega)$. Consequently, $PS[(\bar{u}, \bar{y})]$ would be single-valued and Lipschitz and therefore, coincide with the Hadamard directional derivative of S (cf. section 2).

PROPOSITION 6.11 (Hadamard directional differentiability of S). *Let $(u, y) \in \mathrm{gph}\, S$, where $u = Ay + v$ with $v \in N_M(y)$. If there exists $\lambda \in L^\infty(\Omega)$ with $\lambda \geq 0$ a.e. $\mathcal{A}(y)$ and $\lambda = 0$ a.e. $\mathcal{I}(y)$ such that $v = -2\mathrm{div}\,(\lambda\nabla y)$, then S is Hadamard directionally differentiable.*

Proof. As argued above, it suffices to demonstrate that $A\cdot - 2\mathrm{div}\,(\lambda\nabla\cdot)$ is bounded, linear, and coercive from $H_0^1(\Omega)$ into $H^{-1}(\Omega)$.

Since linearity is obvious, we focus on proving the boundedness and coercivity of this operator. Let $d \in H_0^1(\Omega)$ be arbitrary and consider the following argument:

$$\begin{aligned}
&||Ad - 2\mathrm{div}\,(\lambda\nabla d)||_{H^{-1}} \\
&\leq ||A||_{\mathcal{L}}||d||_{H_0^1} + 2 \sup_{\substack{\varphi \in H_0^1(\Omega) \\ ||\varphi||_{H_0^1} = 1}} \left| \int_{\mathcal{A}(y)} \lambda \nabla d \cdot \nabla\varphi dx \right| \leq ||A||_{\mathcal{L}}||d||_{H_0^1} + 2||\lambda||_{L^\infty}||d||^2_{H_0^1},
\end{aligned}$$

whence we obtain the boundedness. To see that coercivity holds, let $d \in H_0^1(\Omega)$ be arbitrarily fixed and consider

$$\langle Ad - 2\mathrm{div}\,(\lambda\nabla d), d \rangle_{H^{-1},H_0^1} \geq \xi ||d||^2_{H_0^1} + 2\int_\Omega \lambda |\nabla d|^2 dx \geq \xi ||d||^2_{H_0^1}.$$

It follows that $A\cdot - 2\mathrm{div}\,(\lambda\nabla\cdot)$ is linear, bounded, and coercive, as was to be shown. ◻

Corollary 6.10 once again demonstrates the power of the variational analytic method, as we now have a formula characterizing the Hadamard directional derivative of the solution mapping S of an important class of linear elliptic variational inequalities in function space in which nonpolyhedric constraints on the state are involved. As mentioned in Remark 4.4, we have now extended Mignot's classical result beyond the realm of polyhedricity.

As in the polyhedric case, we will also require the following cone in order to derive explicit stationarity conditions.

PROPOSITION 6.12 (the polar cone $[\mathcal{K}(y, v)]^-$). *Let $y \in M$, where M is defined as in* (14). *Then*

$$\begin{aligned}
&[\mathcal{K}(y, v)]^- \\
&= \left\{ -2\mathrm{div}\,(\mu\nabla y) \in H^{-1}(\Omega) \,\middle|\, \mu \in L^2(\Omega) : \mu \geq 0 \text{ a.e. } \mathcal{A}^0(\nabla y),\ \mu = 0 \text{ a.e. } \mathcal{I}(\nabla y) \right\}.
\end{aligned}$$

Proof. We begin by demonstrating the inclusion "$\supseteq$" for the assertion and denote the right-hand side of the equation by K^0. Let $w \in K^0$ and consider an arbitrary $d \in \mathcal{K}(y, v)$. Then

$$\begin{aligned} \langle w, d\rangle_{H^{-1},H_0^1} &= \langle -2\mathrm{div}\,(\mu\nabla y), d\rangle_{H^{-1},H_0^1} = (2\mu\nabla y, \nabla d)_{L^2(\Omega)^l} \\ &= \int_{\mathcal{A}^0(\nabla y)} 2\mu\nabla y \cdot \nabla d dx + \int_{\mathcal{A}^+(\nabla y)} 2\mu\nabla y \cdot \nabla d dx. \end{aligned}$$

Continuing, we recall the characterization of d provided in Proposition 6.6, which provides us with

$$\int_{\mathcal{A}^0(\nabla y)} 2\mu\nabla y \cdot \nabla d dx + \int_{\mathcal{A}^+(\nabla y)} 2\mu\nabla y \cdot \nabla d dx = \int_{\mathcal{A}^0(\nabla y)} 2\mu\nabla y \cdot \nabla d dx \leq 0.$$

Hence, the inclusion holds.

We now use an analogous argument as in the proof of Proposition 6.5 to demonstrate equality. We first need to argue that K^0 is closed and convex. Since the argument for convexity is identical to the one in Proposition 6.5 we need only demonstrate closedness of K^0. First note that K^0 is the image of the set

$$L^0 := \left\{2\mu\nabla y \in G(\Omega) \,\middle|\, \mu \in L^2(\Omega) : \mu \geq 0 \text{ a.e.}\, \mathcal{A}^0(\nabla y)\ \mu = 0 \text{ a.e.}\, \mathcal{I}(\nabla y) \right\}$$

under the negative divergence operator. The image set $-\mathrm{div}\,(L^0)$ is clearly convex; we refer the reader to Lemma 6.14 to see that $-\mathrm{div}\,(L^0)$ is closed provided L^0 is closed. Let $\mathbf{w}_n \in L^0$ such that $\mathbf{w}_n \to \mathbf{w}$ in $L^2(\Omega)^l$. Then by the closedness of $G(\Omega)$, $\mathbf{w} \in G(\Omega)$. By definition, $\mathbf{w}_n \in L^0$ implies that there exist $\mu_n \in L^2(\Omega)$ such that $\mathbf{w}_n = \mu_n\nabla y$. The rest follows analogously to the closure argument for $N^0(\mathbf{z})$ found in the proof of Proposition 6.5. Hence, L^0 and, as argued above, K^0 are closed in their respective spaces.

Assume now there exists $w^* \in [\mathcal{K}(y, v)]^-$ such that $w^* \notin K^0$. Then there exists a $\delta \in H_0^1(\Omega)$ strongly separately the two sets, i.e.,

$$\langle w^*, \delta\rangle_{H^{-1},H_0^1} > 0,\ \langle w, \delta\rangle_{H^{-1},H_0^1} \leq 0\ \forall w \in K^0.$$

Therefore, $\delta \notin \mathcal{K}(y, v)$. Conversely, using the definition of K^0 and the characterization of $\mathcal{K}(y, v)$ provided by Proposition 6.6, we obtain $\delta \in \mathcal{K}(y, v)$, a contradiction. Hence, the equality holds. ∎

We can now derive the optimality conditions for the associated MPEC.

PROPOSITION 6.13 (explicit strong stationarity conditions). *Under the given data assumptions, let $(\bar{u}, \bar{y})$ be a (locally) optimal solution to the corresponding MPEC. If there exist $\lambda \in L^\infty(\Omega)$ such that $\bar{u} = A\bar{y} - \mathrm{div}\,(\lambda\nabla\bar{y})$, then there exist multipliers $p \in H_0^1(\Omega)$ and $\mu \in L^2(\Omega)$ such that*

$$0 = \nabla_u J(\bar{u}, \bar{y}) + p, \tag{24}$$

$$0 = \nabla_y J(\bar{u}, \bar{y}) - A^*p + 2\mathrm{div}\,(\lambda\nabla p) - 2\mathrm{div}\,(\mu\nabla\bar{y}), \tag{25}$$

$$0 = A\bar{y} - \bar{u} - 2\mathrm{div}\,\lambda\nabla\bar{y}, \tag{26}$$

where

$$\left\{ \begin{array}{ll|ll|ll} \nabla\bar{y} \cdot \nabla p \leq 0 & a.e.\, \mathcal{A}^0(\nabla\bar{y}) & \mu \geq 0 & a.e.\, \mathcal{A}^0(\nabla\bar{y}) & \lambda \geq 0 & a.e.\, \mathcal{A}(\nabla\bar{y}) \\ \nabla\bar{y} \cdot \nabla p = 0 & a.e.\, \mathcal{A}^+(\nabla\bar{y}) & \mu = 0 & a.e.\, \mathcal{I}(\nabla\bar{y}) & \lambda = 0 & a.e.\, \mathcal{I}(\nabla\bar{y}) \end{array} \right\}.$$

Proof. The proof follows analogously to that which was used for Theorem 4.6. We therefore need to characterize $[T_{\mathrm{gph}\,S}(\bar{u},\bar{y})]^{-}$.

By definition

$$[T_{\mathrm{gph}\,S}(u,y)]^{-}$$
$$= \left\{(p^*,q^*) \in H_0^1(\Omega) \times H^{-1}(\Omega) \,|\, \langle p^*, w\rangle + \langle q^*, d\rangle \le 0 \;\forall (w,d) \in T_{\mathrm{gph}\,S}(u,y)\right\}.$$

Then by using the characterization from Corollary 6.10, we have the equivalent relation

$$[T_{\mathrm{gph}\,S}(u,y)]^{-} = \big\{(p^*,q^*) \in H_0^1(\Omega) \times H^{-1}(\Omega) \,|$$
$$\langle p^*, Ad - 2\mathrm{div}\,(\lambda\nabla d) + r\rangle + \langle q^*, d\rangle \le 0 \;\forall (d,r) \in \mathrm{gph}\, N_{\mathcal{K}(y,v)}\big\}.$$

Rearranging terms, it follows that

$$[T_{\mathrm{gph}\,S}(u,y)]^{-} = \big\{(p^*,q^*) \in H_0^1(\Omega) \times H^{-1}(\Omega) \,|$$
$$\langle A^*p^* - 2\mathrm{div}\,(\lambda\nabla p) + q^*, d\rangle + \langle p^*, r\rangle_{Y^*,Y} \le 0 \;\forall (d,r) \in \mathrm{gph}\, N_{\mathcal{K}(y,v)}\big\}.$$

This is equivalent to

$$(A^*p^* - 2\mathrm{div}\,(\lambda\nabla p) + q^*, p^*) \in \left[\mathrm{gph}\, N_{\mathcal{K}(y,v)}\right]^{-} = [\mathcal{K}(y,v)]^{-} \times \mathcal{K}(y,v).$$

The assertion then follows from Propositions 6.6 and 6.12. ∎

Note that $\lambda \in L^2(\Omega)$ always exists for a solution $(\bar{u},\bar{y})$ such that $\bar{u} = A\bar{y} - \mathrm{div}\,(\lambda\nabla\bar{y})$. Moreover, as λ is in essence the multiplier associated with the inequality used in the description of K_ψ, which is incidentally to be understood in the range space $L^1(\Omega)$, it does not seem unreasonable to expect $\lambda \in L^\infty(\Omega)$. In the final part of this section, we provide a typical example of an MPEC with gradient constraints for which we demonstrate that $\lambda \in L^\infty(\Omega)$ at a solution. Using M as defined throughout this section, consider the following MPEC:

$$\begin{array}{ll} \min & \frac{1}{2}||y-y_d||^2_{L^2(\Omega)} + \frac{\alpha}{2}||u-c||^2_{L^2(\Omega)} \\ \text{s.t.} & u \in L^2(\Omega),\, y \in H_0^1(\Omega), \\ & u \in -\Delta y + N_M(y). \end{array}$$

Here, Δ is the Laplacian, $\psi \equiv 1$, $\Omega = \mathbb{B}_R(0) \subset \mathbb{R}^l$ is the open ball of radius R. (R sufficiently large), c is a positive constant, and y_d is a radially symmetric function (in $H_0^1(\Omega)$) defined by

$$y_d(x_1,x_2) = \begin{cases} R - \frac{c}{4}|x|_2^2 - \frac{1}{c}, & 0 \le |x|_2 < \frac{2}{c}, \\ R - |x|_2, & \frac{2}{c} \le |x|_2 \le R. \end{cases}$$

Clearly, (c, y_d) is a global minimizer of the objective functional. Moreover, we have

$$|\nabla y_d(x)|_2^2 = \begin{cases} \frac{c^2}{4}|x|_2^2, & 0 \le |x|_2 < \frac{2}{c}, \\ 1, & \frac{2}{c} \le |x|_2 \le R. \end{cases}$$

It follows that $y_d \in M$ with

$$\mathcal{I}(y_d) = \left\{0 \le |x|_2 < \frac{2}{c}\right\} \quad \text{and} \quad \mathcal{A}(y_d) = \left\{\frac{2}{c} \le |x|_2 \le R\right\}.$$

For any $v \in N_M(y_d)$, we know by Proposition 6.6 that there exists $\lambda \in L^2(\Omega)$ such that $v = -2\mathrm{div}\,(\lambda \nabla y)$. Thus, given the feasibility of y_d, it remains to show that $\lambda \in L^\infty(\Omega)$ such that the complementarity conditions hold and that y_d is a solution to the following partial differential equation:

$$(c, \phi)_{L^2} = (\nabla y, \nabla \phi)_{L^2} + 2(\lambda \nabla y, \nabla \phi)_{L^2} \;\forall \phi \in H_0^1(\Omega). \tag{27}$$

We begin with the ansatz for the multiplier

$$\lambda = 0 \text{ a.e.} \,\mathcal{I}(y_d) \quad \text{ and } \quad \lambda = \frac{1}{2}\left(\frac{c}{2}|x|_2 - 1\right) \text{ a.e.} \,\mathcal{A}(y_d).$$

It is easy to check that λ and $|\nabla y_d|_2^2 - 1$ are complementary, and, given that the free boundary is defined by all (x_1, x_2) such that $|x|_2 = 2/c$, we see that λ is in fact continuous. Consider then that for a fixed arbitrary $\phi \in H_0^1(\Omega)$,

$$\begin{aligned}
&(\nabla y_d, \nabla \phi)_{L^2} + 2(\lambda \nabla y, \nabla \phi)_{L^2}\\
&\quad = \int_{\mathcal{A}(y)} \frac{-x}{|x|_2} \cdot \nabla \phi dx - \frac{c}{2}\int_{\mathcal{I}(y)} x \cdot \nabla \phi dx - 2\int_{\mathcal{A}(y)} \lambda \frac{x}{|x|_2} \cdot \nabla \phi dx\\
&\quad = -\int_{\mathcal{A}(y)} (1+2\lambda)\frac{x}{|x|_2} \cdot \nabla \phi dx - \frac{c}{2}\int_{\mathcal{I}(y)} x \cdot \nabla \phi dx\\
&\quad = \int_{\mathcal{A}(y)} \nabla \cdot \left((1+2\lambda)\frac{x}{|x|_2}\right)\phi dx + \int_{\partial \mathcal{I}(y)} \phi ds - \frac{c}{2}\int_{\mathcal{I}(y)} x \cdot \nabla \phi dx\\
&\quad = \int_{\mathcal{A}(y)} \nabla \cdot \left((1+2\lambda)\frac{x}{|x|_2}\right)\phi dx + \int_{\partial \mathcal{I}(y)} \phi ds + c\int_{\mathcal{I}(y)} \phi dx - \int_{\partial \mathcal{I}(y)} \phi ds.
\end{aligned}$$

By substituting the value of λ on the active set, the previous equation yields the desired result. Hence,

$$(\nabla y_d, \nabla \phi) + 2(\lambda \nabla y_d, \nabla \phi) = (c, \phi).$$

It follows that (c, y_d) is a feasible point for the MPEC and therefore, as argued above, the global minimizer.

We note that it is certainly possible to create further examples, and we conjecture that a multiplier $\lambda \in L^\infty(\Omega)$ will exist in the presence of a smooth enough boundary and control.

Appendix. Proof of Theorem 3.1.

Proof. We first derive that S is a Lipschitz continuous function with respect to $u \in Y^*$. By Theorem 3.3.4 in [4], there exists a unique y solving the variational inequality for each $u \in Y^*$. Let (u_1, y_1) and (u_2, y_2) be two arbitrary control-state pairs $(S(u_i) = y_i)$. Then by the convexity of M, the generalized equation used to describe S can be formulated in variational form for each pair as follows:

$$\langle Ay_1 - u_1, y' - y_1\rangle_{Y^*,Y} \geq 0 \;\forall y' \in M, \tag{28}$$

$$\langle Ay_2 - u_2, y'' - y_2\rangle_{Y^*,Y} \geq 0 \;\forall y'' \in M. \tag{29}$$

Substituting $y' = y_2$ and $y'' = y_1$ into (28) and (29), respectively, and recognizing that

$$\langle Ay_2 - u_2, y'' - y_2\rangle = \langle u_2 - Ay_2, y_2 - y''\rangle_{Y^*.Y},$$

we add the two inequalities together and obtain

$$\langle Ay_1 - u_1 + u_2 - Ay_2, y_2 - y_1\rangle_{Y^*.Y} \geq 0.$$

By the coercivity of A, there exists a $\xi \in \mathbb{R}_+ \setminus \{0\}$ such that

$$\begin{aligned}&\xi||y_2 - y_1||_Y^2\\ &\leq \langle A(y_2 - y_1), y_2 - y_1\rangle_{Y^*.Y} \leq \langle u_2 - u_1, y_2 - y_1\rangle_{Y^*,Y} \leq ||u_2 - u_1||_{Y^*}||y_2 - y_1||_Y.\end{aligned}$$

It follows that there exists $L = 1/\xi$ such that

$$||S(u_2) - S(u_1)||_Y \leq L||u_2 - u_1||_{Y^*} \ \forall u_1, u_2 \in Y^*. \tag{30}$$

Therefore, we can rewrite (2) as follows:

$$\min\left\{\tilde{J}(u)\,|u \in \mathcal{U}\right\},$$

where $\tilde{J}(u) := J(u, S(u))$. Due to the continuity of S, $\tilde{J}$ remains coercive and bounded from below by some K. Therefore, the level sets of $\tilde{J}$, i.e., the sets defined by

$$\mathrm{lev}_\gamma \tilde{J} := \left\{u \in \mathcal{U} \,\middle|\, \tilde{J}(u) \leq \gamma\right\}, \gamma \in \mathbb{R},$$

are bounded in $\mathcal{U}$ for all $\gamma \in \mathbb{R}$ (cf. Proposition 3.2.8 of [4]). Now let u_k be an infimizing sequence of $\tilde{J}$, that is,

$$\lim_k \tilde{J}(u_k) = \inf_{u\in\mathcal{U}} \tilde{J}(u).$$

Clearly, there exists some $\gamma_0 \in \mathbb{R}_+\setminus\{0\}$ such that $u_k \in \mathrm{lev}_{\gamma_0}\tilde{J}$ for all k large. Since $\mathcal{U}$ is a Hilbert space and therefore reflexive, we can select a weakly converging subsequence of $\{u_k\}$, denoted by $\{u_{k_l}\}$, such that $u_{k_l} \rightharpoonup \bar{u}$. Moreover, the compactness of the embedding of $\mathcal{U}$ into Y^*, which itself is a Banach space, implies that there exists a further subsequence $u_{k_{l_n}} \to_{Y^*} \bar{u}$. Therefore, $y_{k_{l_n}} = S(u_{k_{l_n}}) \to S(\bar{u}) = \bar{y}$, via (30), so that the assumptions on J imply

$$-\infty < K \leq \inf_{u\in\mathcal{U}} \tilde{J}(u) \leq \tilde{J}(\bar{u}) \leq \liminf_n \tilde{J}(u_{k_{l_n}}) = \lim_k \tilde{J}(u_k) = \inf_{u\in\mathcal{U}} \tilde{J}(u) \leq \gamma_0,$$

as was to be shown. ∎

The following technical lemma guarantees the closure of the image of certain $G(\Omega)$-sets under the $-\mathrm{div}$ operator.

LEMMA 6.14. *If $C \subset G(\Omega)$ is a nonempty, closed, convex cone, then $-\mathrm{div}\,(C)$ is closed in the strong topology on $H^{-1}(\Omega)$.*

Proof. Let $z^* \in \mathrm{cl}\,\{-\mathrm{div}\,(C)\}$. Then there exists a sequence $\{x_n\} \subset C$ such that $-\mathrm{div}\,(x_n) \to z^*$ in $H^{-1}(\Omega)$. To see that $\{x_n\}$ is bounded in $G(\Omega)$, we note the existence of a positive constant $\kappa > 0$ independent of n such that $||-\mathrm{div}\,(x_n)||_{H^{-1}} \leq \kappa$ and, for each n, a function $\varphi_n \in H_0^1(\Omega)$ such that $x_n = \nabla\varphi_n$. Without loss of generality, we assume that $x_n \neq 0$ for all n; otherwise, we could take a subsequence. It follows that

$$\kappa \geq ||-\mathrm{div}\,(x_n)||_{H^{-1}} = \sup_{\substack{\varphi \in H_0^1(\Omega) \\ \varphi \neq 0}} \frac{|\langle -\mathrm{div}\,(x_n), \varphi \rangle_{H^{-1},H_0^1}|}{||\varphi||_{H_0^1}} \geq \frac{|(x_n, \nabla \varphi_n)_{L^2}|}{||\nabla \varphi_n||_{L^2}} = ||x_n||_{L^2},$$

from which we deduce the boundedness of $\{x_n\}$ in $G(\Omega)$.

Since $G(\Omega)$ is a reflexive Banach space, there exists a subsequence $\{n_l\} \subset \{n\}$ and a function $x^* \in G(\Omega)$ such that $x_{n_l} \rightharpoonup x^*$ in $G(\Omega)$. Since $x_{n_l} \in C$ and C is closed and convex, and therefore, weakly closed, $x^* \in C$. Moreover, $-\mathrm{div}\,(x_{n_l}) \rightharpoonup -\mathrm{div}\,(x^*)$ in $H^{-1}(\Omega)$. But then for any $z^* \in \mathrm{cl}\,\{-\mathrm{div}\,(C)\}$, there exists some $x^* \in C$ such that $z^* = -\mathrm{div}\,(x^*)$, from which the assertion follows. ∎

Acknowledgments. We would very much like to thank the reviewers for their important critical remarks, which certainly helped improve this paper. Finally, we thank Antoine Laurain for the helpful discussions leading to the construction of the example following Proposition 6.13.

REFERENCES

[1] Y. Achdou and O. Pironneau, *Computational Methods for Option Pricing*, Frontiers Appl. Math. 30, SIAM, Philadelphia, 2005.

[2] R. A. Adams and J. J.-F. Fournier, *Sobolev Spaces*, 2nd ed., Elsevier, Amsterdam, 2008.

[3] H. Attouch, *Variational Convergence for Functions and Operators*, Pitman (Advanced Publishing Program), Boston, 1984.

[4] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces*, MPS/SIAM Ser. Optim. 6, SIAM, MPS, Philadelphia, 2006.

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

[6] V. Barbu, *Optimal Control of Variational Inequalities*, Res. Notes in Math. 100, Pitman (Advanced Publishing Program), Boston, 1984.

[7] M. Bergounioux, *Optimal control of an obstacle problem*, Appl. Math. Optim., 36 (1997), pp. 147–172.

[8] M. Bergounioux, *Optimal control of problems governed by abstract elliptic variational inequalities with state constraints*, SIAM J. Control Optim., 36 (1998), pp. 273–289.

[9] M. Bergounioux and S. Lenhart, *Optimal control of bilateral obstacle problems*, SIAM J. Control Optim., 43 (2004), pp. 240–255.

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

[11] N. Bourbaki, *Topological Vector Spaces: Chapters* 1–5, Springer-Verlag, Berlin, 1987.

[12] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Vol. 3, Springer-Verlag, Berlin, 1990.

[13] C. Do, *Generalized second-order derivatives of convex functions in reflexive Banach spaces*, Trans. Amer. Math. Soc., 334 (1992), pp. 281–301.

[14] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Classics Appl. Math. 28, SIAM, Philadelphia, 1999.

[15] A. Friedman, *Variational Principles and Free-Boundary Problems*, John Wiley & Sons, Inc., New York, 1982.

[16] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer Ser. Comput. Phys., Springer-Verlag, New York, 1984.

[17] W. Han and B. D. Reddy, *Plasticity*, Interdiscip. Appl. Math. 9, Springer-Verlag, New York, 1999.

[18] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities.*, J. Math. Soc. Japan, 29 (1977), pp. 615–631.

[19] M. Hintermüller, *Inverse coefficient problems for variational inequalities: optimality conditions and numerical realization*, M2AN Math. Model. Numer. Anal., 35 (2001), pp. 129–152.

[20] M. Hintermüller, *An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities*, Inverse Problems, 24 (2008), 034017.

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

[22] M. Hintermüller and I. Kopacka, *A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs*, Comput. Optim. Appl., 50 (2011), pp. 111–145.

[23] M. Hintermüller and J. Outrata, *A Note on Optimality Conditions in Control of Elliptic Variational Inequalities*. Preprint IFB-Report 47 (09/2011); Institute for Mathematik and Scientific Computing, Karl-Franzens-Universität Graz, Graz, Austria, 2011.

[24] A. Ioffe, *Variational analysis of a composite function: A formula for the lower second order epi-derivative*, J. Math. Anal. Appl., 160 (1991), pp. 379–405.

[25] K. Ito and K. Kunisch, *Optimal control of elliptic variational inequalities*, Appl. Math. Optim., 41 (2000), pp. 343–364.

[26] L. Katchanov, *Élements de la Théorie de la Plasticité*, MIR, Moscow, 1975.

[27] N. Kikuchi and J. T. Oden, *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*, SIAM Stud. Appl. Math. 8, SIAM, Philadelphia, 1988.

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

[29] A. B. Levy, *Sensitivity of solutions to variational inequalities on Banach spaces*, SIAM J. Control Optim., 38 (1999), pp. 50–60.

[30] A. B. Levy, *Second-order epi-derivatives of composite functionals*, Ann. Oper. Res., 101 (2001), pp. 267–281.

[31] A. B. Levy and R. T. Rockafellar, *Sensitivity analysis of solutions to generalized equations*, Trans. Amer. Math. Soc., 345 (1994), pp. 661–671.

[32] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.

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

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

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

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

[37] P. Neittaanmaki, J. Sprekels, and D. Tiba, *Optimization of Elliptic Systems*, Springer Monogr. Math., Springer, New York, 2006.

[38] J. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Nonconvex Optim. Appl. 28, Kluwer Academic Publishers, Dordrecht, 1998.

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

[40] S. M. Robinson, *An implicit function theorem for a class of nonsmooth functions*, Math. Oper. Res., 16, pp. 292–309.

[41] R. T. Rockafellar, *First- and second-order epi-differentiability in nonlinear programming*, Trans. Amer. Math. Soc., 307 (1988), pp. 75–108.

[42] R. T. Rockafellar, *Proto-differentiability of set-valued mappings and its applications in optimization*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 6 (1989), pp. 449–482.

[43] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[44] H. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.

[45] A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl., 66 (1990), pp. 477–487.

[46] K. Yosida, *Functional Analysis*, 6th ed., Spinger-Verlag, Berlin, New York, 1980.

[47] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.