Mathematical Programming manuscript No.
(will be inserted by the editor)

On M-stationarity conditions in MPECs and the associated qualification conditions

Lukáš Adam · **René Henrion** · **Jiří Outrata**

Received: date / Accepted: date

Abstract Depending on whether a mathematical program with equilibrium constraints (MPEC) is considered in its original or its enhanced (via KKT conditions) form, the assumed qualification conditions as well as the derived necessary optimality conditions may differ significantly. In this paper, we study this issue when imposing one of the weakest possible qualification conditions, namely the calmness of the perturbation mapping associated with the respective generalized equations in both forms of the MPEC. It is well known that the calmness property allows one to derive the so-called M-stationarity conditions. The restrictiveness of assumptions and the strength of conclusions in the two forms of the MPEC is also strongly related to the qualification conditions on the "lower level". For instance, even under the Linear Independence Constraint Qualification (LICQ) for a lower level feasible set described by $\mathscr{C}^1$ functions, the calmness properties of the original and the enhanced perturbation mapping are drastically different. When passing to $\mathscr{C}^{1,1}$ data, this difference still remains true under the weaker Mangasarian-Fromovitz Constraint Qualification, whereas under LICQ both the calmness assumption and the derived optimality conditions are fully equivalent for the original and the enhanced form of the MPEC. After clarifying these relations, we provide a compilation of practically relevant consequences of our analysis in the derivation of necessary optimality conditions. The obtained results are finally applied to MPECs with structured equilibria.

This paper is dedicated to Prof. R. T. Rockafellar on the occasion of his 80th birthday.

This work was partially supported by the Grant Agency of the Czech Republic under grant 15-00735S (L. Adam) and in part by the Australian Research Council under grant DP-160100854 (R. Henrion, J. Outrata).

The final publication is available at Springer via http://dx.doi.org/10.1007/s10107-017-1146-3

L. Adam, J. Outrata
ÚTIA, Czech Academy of Sciences
Pod Vodárenskou věží 4, 18208 Prague, Czech Republic
E-mail: adam@utia.cas.cz, outrata@utia.cas.cz

R. Henrion
Weierstrass Institute
Mohrenstrasse 39, 10117 Berlin, Germany
E-mail: henrion@wias-berlin.de

Keywords mathematical programs with equilibrium constraints · optimality conditions · constraint qualification · calmness · perturbation mapping

Mathematics Subject Classification (2000) 65K10 · 90C30 · 90C31 · 90C46

1 Introduction

Starting with [22], efficient necessary optimality conditions for various types of *mathematical programs with equilibrium constraints* (MPECs) have been developed on the basis of the generalized differential calculus of Mordukhovich, e.g. [13,15,16,21]. Following [19], we speak about M-stationarity conditions. Let us consider an MPEC of the form

$$\begin{aligned} &\underset{x,y}{\text{minimize}}\ \varphi(x,y) \\ &\text{subject to}\ \ 0 \in F(x,y) + \hat{N}_\Gamma(y), \\ &\qquad\qquad\ \ x \in \omega, \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the *control* variable, $y \in \mathbb{R}^m$ is the *state* variable, $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is the objective, $\omega \subset \mathbb{R}^n$ is a closed set of admissible controls, $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is a continuously differentiable mapping, and the constraint set $\Gamma \subset \mathbb{R}^m$ is given by inequalities

$$\Gamma = \{y \in \mathbb{R}^m \mid q_i(y) \le 0,\ i = 1,\dots,s\} \tag{2}$$

with a continuously differentiable mapping $q = (q_1,\dots,q_s)^\top : \mathbb{R}^m \to \mathbb{R}^s$. Further, $\hat{N}$ refers to the *regular (Fréchet) normal* cone (see Definition 1).

Let $(\bar{x},\bar{y})$ be a (local) solution of (1). When Γ satisfies the *Mangasarian-Fromovitz Constraint Qualification* (MFCQ) at $\bar{y}$ (see Definition 4), one has the representation

$$\hat{N}_\Gamma(y) = N_\Gamma(y) = (\nabla q(y))^\top N_{\mathbb{R}^s_-}(q(y))$$

on a neighborhood of $\bar{y}$ so that the following equivalence holds true for the *generalized equation* in (1):

$$0 \in F(x,y) + N_\Gamma(y) \Leftrightarrow \exists \lambda : 0 \in H(x,y,\lambda) + N_{\mathbb{R}^m \times \mathbb{R}^s_+}(y,\lambda), \tag{3}$$

provided y is close to $\bar{y}$ and $H(x,y,\lambda) := (F(x,y) + (\nabla q(y))^\top \lambda, -q(y))$. This relation suggests also to consider the *enhanced* MPEC

$$\begin{aligned} &\underset{x,y,\lambda}{\text{minimize}}\ \varphi(x,y) \\ &\text{subject to}\ \ 0 \in H(x,y,\lambda) + N_{\mathbb{R}^m \times \mathbb{R}^s_+}(y,\lambda), \\ &\qquad\qquad\ \ x \in \omega \end{aligned} \tag{4}$$

in variables (x,y,λ). The generalized equation in (4) has a substantially simpler constraint set than the generalized equation in (1). As the price for it, one has to do with an additional variable λ. Let us introduce the multifunction $\Lambda : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^s$ by

$$\Lambda(x,y) := \left\{ \lambda \in \mathbb{R}^s \,\middle|\, 0 = F(x,y) + (\nabla q(y))^\top \lambda,\ q(y) \in N_{\mathbb{R}^s_+}(\lambda) \right\} \tag{5}$$

so that $\Lambda(x,y)$ is the set of *Lagrange multipliers* associated with a pair (x,y), feasible with respect to the generalized equation from (1). It is easy to see that under MFCQ we have that $\Lambda(\bar{x},\bar{y}) \neq \emptyset$ and $(\bar{x},\bar{y})$ is a local solution to problem (1) if and only if $(\bar{x},\bar{y},\lambda)$ is a local solution to (4) for all $\lambda \in \Lambda(\bar{x},\bar{y})$. Likewise, it is known that for a local solution $(\bar{x},\bar{y},\bar{\lambda})$ of (4) the pair $(\bar{x},\bar{y})$ need not be a local solution of (1), see [2] in the context of bilevel programming. It follows that numerical methods computing M-stationary points of (4) may terminate at points which are not M-stationary with respect to the original (1). A complete analysis of this issue requires, however, to compare also the *qualification conditions* imposed in the course of derivation of the M-stationarity conditions for (1) and (4), respectively. As in [15,22] we will make use of the so-called calmness qualification conditions [10] which ensure a certain Lipschitzian behavior of the canonically perturbed constraint maps in (1) and (4), cf. Definition 3 and formula (7). It turns out that, very often, the calmness qualification condition related to (1) is satisfied, whereas the qualification condition of (4) may be not fulfilled for some or even for any multipliers λ. The main aim of this paper is thus a thorough analysis of both these qualification conditions and their mutual relationship. Not surprisingly, in the achieved results an important role is played by the *constraint qualifications* (CQs) which Γ fulfills at $\bar{y}$. The choice between M-stationarity conditions of (1) and (4) depends, however, also on some other circumstances. First, it is the question of workable criteria for the considered calmness qualification conditions which are typically somewhat simpler in the case of (4). Further, one has to take into account also the possibility to express M-stationarity conditions of (1) in terms of problem data because otherwise the results do not have a practical value.

In the paper, all these aspects will be considered. To state our aims rigorously, one needs some basic notions from variational analysis. They are introduced at the beginning of Section 2.1. Section 2.2 is then devoted to a proper problem setting. We define here the perturbation mappings M and $\tilde{M}$ associated with problems (1) and (4). In Section 2.3 we present several auxiliary results needed in the sequel. Since calmness of M and $\tilde{M}$ allows us to derive necessary optimality conditions, Section 3 deals with the relations between calmness of M and $\tilde{M}$ under various CQs imposed on Γ. Another important issue is to find workable criteria (in terms of problem data) ensuring the calmness of M and $\tilde{M}$. This will be considered in Section 4. One finds there in Theorem 8 also a compilation of the main results of the paper. In Section 5 we illustrate the application of our results to a structured family of MPECs or bilevel problems.

Our notation is standard. For $f:\mathbb{R}\to\mathbb{R}$ by f' we mean its derivative. For a vector $x\in\mathbb{R}^n$ and a set $C\subset\mathbb{R}^n$, by $\|x\|$ we mean the (Euclidean) norm of x and by $d(x,C)$ the distance of x from C. By $o(h)$ we understand any function such that $\lim_{h\searrow 0}\frac{o(h)}{\|h\|}=0$. Finally, by $\#S$ we mean the cardinality of a set S.

2 Problem setting and preliminaries

Throughout the whole paper we consider equilibria governed by the *generalized equation* from (1), where Γ is given in (2). With minor modifications, however, the whole theory applies also to the case when Γ is given by inequalities and *equalities*. For the sake of brevity we assume (without any loss of generality) that, at the considered point $\bar{y}$, all inequality constraints are active, i.e,

$$q_i(\bar{y}) = 0,\ i = 1,\dots,s.$$

2.1 Background from variational analysis

Definition 1 For a closed set $A \subset \mathbb{R}^n$ and $\bar{x} \in A$ we define the *Fréchet* and *limiting (Mordukhovich) normal cone* to A at $\bar{x}$ by

$$\hat{N}_A(\bar{x}) = \{x^* \mid \langle x^*, x-\bar{x}\rangle \le o(\|x-\bar{x}\|) \text{ for all } x \in A\}$$
$$N_A(\bar{x}) = \operatorname*{Limsup}_{x\to\bar{x}} \hat{N}_A(x) := \left\{x^* \,\middle|\, \exists (x_k, x_k^*) : x_k^* \in \hat{N}_A(x_k),\ x_k \to \bar{x},\ x_k^* \to x^* \right\}.$$

If A happens to be convex, both normal cones coincide and are equal to the normal cone in the sense of convex analysis

$$\hat{N}_A(\bar{x}) = N_A(\bar{x}) = \{x^* \mid \langle x^*, x-\bar{x}\rangle \le 0 \text{ for all } x \in A\}.$$

It follows from [18, Exercise 10.26(d)] that under the MFCQ at $\bar{y}$ we have $\hat{N}_\Gamma(y) = N_\Gamma(y)$ for all y from a neighborhood of $\bar{y}$ and therefore one can replace the regular normal cone in (1) by the limiting one, having a better calculus.

Definition 2 For a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and for any $\bar{y} \in M(\bar{x})$ we define the *(limiting) coderivative* $D^*M(\bar{x},\bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ at this point as

$$D^*M(\bar{x},\bar{y})(y^*) = \left\{x^* \,\middle|\, (x^*, -y^*) \in N_{\operatorname{gph} M}(\bar{x},\bar{y}) \right\},$$

where $\operatorname{gph} M$ stands for the graph of M.

Definition 3 We say that a multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ has the *Aubin property* around $(\bar{x},\bar{y}) \in \operatorname{gph} M$ if there exist a nonnegative modulus L and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that for all $x, x' \in U$ and all $y \in M(x) \cap V$ we have

$$d(y, M(x')) \le L\|x - x'\|.$$

Similarly, we say that M is *calm* at $(\bar{x},\bar{y}) \in \operatorname{gph} M$ if there exist a nonnegative modulus L and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that for all $x \in U$ and $y \in M(x) \cap V$ we have

$$d(y, M(\bar{x})) \le L\|x - \bar{x}\|. \tag{6}$$

Note that the calmness may be significantly weaker than the Aubin property. For example any polyhedral mapping (mapping whose graph is a finite union of convex polyhedra) satisfies the calmness property at any point of its graph but may fail to have the Aubin property at the same time.

In our analysis we make use of some basic CQs from nonlinear programming. For the reader's convenience, we recall them in the next definition, where $I(y)$ denotes the set of active constraints, i.e.,

$$I(y) := \{i \in \{1,\dots,S\} \mid q_i(y) = 0\}.$$

Definition 4 Consider a set Γ defined by inequalities (2) and some point $\bar{y} \in \Gamma$. We say that Γ satisfies LICQ *(linear independence constraint qualification)* at $\bar{y}$ if the gradients corresponding to all active constraints are linearly independent, hence

$$\sum_{i \in I(\bar{y})} \mu_i \nabla q_i(\bar{y}) = 0 \implies \mu_i = 0 \text{ for all } i \in I(\bar{y}).$$

Similarly, we say that Γ satisfies MFCQ *(Mangasarian-Fromovitz constraint qualification)* at $\bar{y}$ if the gradients corresponding to all active constraints are positively linearly independent, hence

$$\sum_{i\in I(\bar{y})} \mu_i \nabla q_i(\bar{y}) = 0,\ \mu_i \geq 0 \implies \mu_i = 0 \text{ for all } i \in I(\bar{y}).$$

We have used here the dual formulation of MFCQ. Finally, Γ satisfies CRCQ *(constant rank constraint qualification)* at $\bar{y}$ if there is a neighborhood U of $\bar{y}$ such that for all subsets I of active indices $I(\bar{y})$ we have that $\operatorname{rank}\{\nabla q_i(y) \,|\, i \in I\}$ is a constant value for all $y \in U$.

Note that both, MFCQ and CRCQ are strictly weaker conditions than LICQ (even when imposed jointly) and that neither of the two implies the other.

2.2 Problem setting

The notions defined above enable us to state the investigated problem rigorously. The perturbation mappings associated with MPECs (1) and (4) attain the form

$$\begin{aligned} M(z) &:= \{(x,y) \,|\, x \in \omega, z \in F(x,y) + N_\Gamma(y)\}, \\ \tilde{M}(z_1,z_2) &:= \left\{(x,y,\lambda) \,\middle|\, x \in \omega, (z_1,z_2) \in H(x,y,\lambda) + N_{\mathbb{R}^m\times\mathbb{R}^s_+}(y,\lambda)\right\} \\ &= \left\{(x,y,\lambda) \,\middle|\, x \in \omega,\ z_1 = F(x,y) + (\nabla q(y))^\top \lambda,\ z_2 \in -q(y) + N_{\mathbb{R}^s_+}(\lambda)\right\}, \end{aligned} \tag{7}$$

respectively. The M-stationarity conditions for (1) can be formulated as follows.

Theorem 1 ([22], Theorem 3.2) *Let $(\bar{x},\bar{y})$ be a local solution to (1). If M is calm at $(0,\bar{x},\bar{y})$, then there exists an MPEC multiplier $v \in \mathbb{R}^m$ such that*

$$\begin{aligned} 0 &\in \nabla_x \varphi(\bar{x},\bar{y}) + [\nabla_x F(\bar{x},\bar{y})]^\top v + N_\omega(\bar{x}), \\ 0 &\in \nabla_y \varphi(\bar{x},\bar{y}) + [\nabla_y F(\bar{x},\bar{y})]^\top v + D^* N_\Gamma(\bar{y}, -F(\bar{x},\bar{y}))(v). \end{aligned} \tag{8}$$

Since MPEC (4) has exactly the same structure as MPEC (1), the respective M-stationarity condition can be derived in the same way upon putting

$$x := x,\ y := (y,\lambda),\ F := H,\ \Gamma := \mathbb{R}^m \times \mathbb{R}^s_+.$$

Instead of keeping a co-derivative expression $D^* N_{\mathbb{R}^m\times\mathbb{R}^s_+}$ similar to $D^* N_\Gamma$ in (8), one can make this fully explicit now by relying on well-known formulae (e.g., [14]). We obtain the following twin theorem to Theorem 1:

Theorem 2 *Let $(\bar{x},\bar{y},\bar{\lambda})$ be a local solution to (4) and assume that $q \in \mathscr{C}^2$. If $\tilde{M}$ is calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$, then there exist some multipliers $v \in \mathbb{R}^m$ and $w \in \mathbb{R}^s$ such that*

$$\begin{aligned} &0 = \nabla_x \varphi(\bar{x},\bar{y}) + [\nabla_x F(\bar{x},\bar{y})]^\top v + N_\omega(\bar{x}), \\ &0 = \nabla_y \varphi(\bar{x},\bar{y}) + [\nabla_y F(\bar{x},\bar{y})]^\top v + \sum_{i=1}^{s} \bar{\lambda}_i \nabla^2 q_i(\bar{y}) v - [\nabla q(\bar{y})]^\top w, \\ &0 = \nabla q_i(\bar{y}) v && \forall i : \bar{\lambda}_i > 0, \\ &0 = w_i && \forall i : q_i(\bar{y}) < 0, \\ &0 \geq w_i, 0 \leq \nabla q_i(\bar{y}) v \quad \textit{or} \quad 0 = w_i \quad \textit{or} \quad 0 = \nabla q_i(\bar{y}) v && \forall i : \bar{\lambda}_i = q_i(\bar{y}) = 0. \end{aligned} \tag{9}$$

Theorem 2 can be interpreted as a variant of Theorem 1 in a different disguise addressing the same topic of MPEC (1) with differing assumptions and differing stationarity conditions. By taking into account the relationships between local solutions to (1) and (4) mentioned above, the combination of both theorems immediately leads to the following result.

Corollary 1 *Let* $(\bar{x},\bar{y})$ *be a local solution to* (1) *and assume that MFCQ is satisfied at* $\bar{y}$*. Then there exist multipliers* v *and* w *such that* (9) *holds true for those* $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$ *for which* $\tilde{M}$ *is calm at* $\left(0,0,\bar{x},\bar{y},\bar{\lambda}\right)$.

We observe first that Theorem 1 requires the computation of a coderivative while Theorem 2 provides fully explicit stationarity conditions. Precise formulae for this coderivative in terms of the problem data are available provided that Γ is polyhedral ([9, Theorem 3.2]), under LICQ at $\bar{y}$ ([7, Theorem 3.1]) or under a relaxation of MFCQ combined with the so-called 2-regularity ([5, Theorem 3]). An upper estimate has been derived in [7, Theorem 3.3] and further worked out in the Section 3.2 (Corollary 3). Moreover, Corollary 1 enables us to circumvent the difficulties associated with the coderivative in (8) and to benefit from the explicit stationary conditions (9). This gain in convenience is bought by the need to check a calmness condition for $\tilde{M}$ which may be more restrictive than the calmness condition for M imposed in Theorem 1.

2.3 Auxiliary results

At several places of the paper we will make use of the following statement from [12] which ensures the calmness of the intersection of two independently perturbed multifunctions.

Theorem 3 ([12], Theorem 3.6) *Consider the following multifunctions* $S_1 : \mathbb{R}^{n_1} \rightrightarrows \mathbb{R}^m$ *and* $S_2 : \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^m$ *and a point* $\bar{u} \in S_1(0) \cap S_2(0)$*. Then* $\Sigma(z_1,z_2) := S_1(z_1) \cap S_2(z_2)$ *is calm at* $(0,0,\bar{u})$ *provided the following conditions are satisfied:*

1. S_1 *is calm at* $(0,\bar{u})$*;*
2. S_2 *is calm at* $(0,\bar{u})$*;*
3. S_1^{-1} *has the Aubin property at* $(\bar{u},0)$*;*
4. $S_1 \cap S_2(0)$ *is calm at* $(0,\bar{u})$.

In the next two lemmas we present a convenient way of verifying the assumptions of Theorem 3 and then we apply it to a special structure arising later in the manuscript. Note that the following lemma is a compilation of well-known results:

Lemma 1 *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$ *be a differentiable function. Then* f^{-1} *is calm at* $(f(\bar{x}),\bar{x})$ *if at least one of the following conditions holds:*

1. f *is piecewise linear;*
2. $\nabla f(\bar{x})$ *has full row rank;*
3. $\nabla f(\bar{x})$ *has full column rank.*

Proof The first case is the classical result of Robinson [17, Proposition 1]. The second one implies the Aubin property of f^{-1} at $(f(\bar{x}),\bar{x})$ and the third one the isolated calmness property of f^{-1} at $(f(\bar{x}),\bar{x})$ by [3, Corollary 3I.11]. Since both these properties imply calmness, the proof is complete. □

Lemma 2 *Consider a multifunction* $\phi : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^p \times \mathbb{R}^t$ *with the separable structure*

$$\phi(u,v) = \phi_1(u) \times \phi_2(v),$$

and assume that $(\bar{w},\bar{z}) \in \phi_1(\bar{u}) \times \phi_2(\bar{v})$*, where* ϕ_1 *is calm at* $(\bar{u},\bar{w})$ *and* ϕ_2 *is calm at* $(\bar{v},\bar{z})$*. Then* ϕ *is calm at* $((\bar{u},\bar{v}),(\bar{w},\bar{z}))$*.*

Proof Let us equip the Cartesian product $\mathbb{R}^p \times \mathbb{R}^t$ with the sum norm. Then one has for all $w \in \phi_1(u)$ and $z \in \phi_1(v)$ that

$$d((w,z),\phi(\bar{u},\bar{v})) = d(w,\phi_1(\bar{u})) + d(z,\phi_2(\bar{v})) \le L_1\|u-\bar{u}\| + L_2\|v-\bar{v}\| \tag{10}$$

whenever (u,v) and (w,z) are sufficiently close to $(\bar{u},\bar{v})$ and $(\bar{w},\bar{z})$, respectively. In (10), L_1 and L_2 signify the calmness moduli of ϕ_1 and ϕ_2 at $(\bar{u},\bar{w})$ and $(\bar{v},\bar{z})$, respectively. We immediately conclude that ϕ is calm at the respective point. □

Lemma 3 *Consider* $u = (u_1,u_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = \mathbb{R}^n$*, continuously differentiable mappings* $H_1 : \mathbb{R}^n \to \mathbb{R}^m$*,* $H_2 : \mathbb{R}^n \to \mathbb{R}^{n_2}$*, closed sets* $\Delta \subset \mathbb{R}^n$*,* $\Omega \subset \mathbb{R}^{n_2}$ *and the following multifunctions*

$$\begin{aligned} S_1(z_1) &:= \{u \,|\, H_1(u) - z_1 = 0\}, \\ S_2(z_2) &:= \{u \in \Delta \,|\, H_2(u) - z_2 \in N_\Omega(u_2)\}. \end{aligned} \tag{11}$$

Consider further a point $\bar{u} \in S_1(0) \cap S_2(0)$ *with the following properties:* S_1 *is calm at* $(0,\bar{u})$*,* S_2 *is calm at* $(0,\bar{u})$ *and the following qualification condition holds:*

$$(\nabla H_1(\bar{u}))^\top a \in \begin{pmatrix} 0 & \nabla_{u_1} H_2(\bar{u})^\top \\ I & \nabla_{u_2} H_2(\bar{u})^\top \end{pmatrix} N_{\mathrm{gph}\, N_\Omega}(\bar{u}_2, H_2(\bar{u})) + N_\Delta(\bar{u}) \implies a = 0. \tag{12}$$

Then $\Sigma(z_1,z_2) := S_1(z_1) \cap S_2(z_2)$ *is calm at* $(0,0,\bar{u})$*.*

Proof Imitating the proof of [20, Proposition 5.2], it can be shown that Σ is calm at $(0,0,\bar{u})$ if and only if $S_1 \cap \tilde{S}_2$ is calm at $(0,0,0,\bar{u})$ with

$$\tilde{S}_2(z_2,z_3) := \left\{ u \in \Delta \,\middle|\, \begin{pmatrix} u_2 - z_3 \\ H_2(u) - z_2 \end{pmatrix} \in \mathrm{gph}\, N_\Omega \right\}.$$

We will now apply Theorem 3 to S_1 and $\tilde{S}_2$. Due to [20, Proposition 5.2] the calmness of $\tilde{S}_2$ at $(0,0,\bar{u})$ is equivalent to the calmness of S_2 at $(0,\bar{u})$, which is satisfied by our assumptions. The multifunction $S_1^{-1} = H_1$ is single-valued and locally Lipschitz continuous, and thus satisfies the Aubin property everywhere. Calmness of S_1 at $(0,\bar{u})$ is satisfied due to the assumptions.

To show that $G(z) := S_1(z) \cap \tilde{S}_2(0,0)$ is calm at $(0,\bar{u})$, we claim that (12) implies even the Aubin property of G around $(0,\bar{u})$, which by virtue of the Mordukhovich criterion [18, Theorem 9.40] is equivalent to the implication

$$\begin{pmatrix} a \\ 0 \end{pmatrix} \in N_{\mathrm{gph}\, G}(0,\bar{u}) \implies a = 0.$$

By [18, Theorem 6.14] this is implied by

$$(\nabla H_1(\bar{u}))^\top a \in N_{\tilde{S}_2(0,0)}(\bar{u}) \implies a = 0. \tag{13}$$

Since $\tilde{S}_2$ is calm at $(0,0,\bar{u})$, we may use [6, Theorem 4.1] to deduce

$$N_{\tilde{S}_2(0,0)}(\bar{u}) \subset \begin{pmatrix} 0 & I \\ \nabla_{u_1} H_2(\bar{u}) & \nabla_{u_2} H_2(\bar{u}) \end{pmatrix}^\top N_{\mathrm{gph}\, N_\Omega}(\bar{u}_2, H_2(\bar{u})) + N_\Delta(\bar{u}). \tag{14}$$

However, due to (14), it is clear that (12) implies (13) and hence G has the Aubin property around $(0,\bar{u})$, which means that Σ is indeed calm at $(0,0,\bar{u})$. □

3 Relations of calmness properties of M and $\tilde{M}$

This section is devoted to a study of the general relationship between the calmness properties of M and $\tilde{M}$ defined in (7). Since we do not make use of any result from second-order variational analysis in this section, we require functions q_i to be of class $\mathscr{C}^1$.

3.1 Calmness under MFCQ and $\mathscr{C}^1$ inequalities

Before proving our first result concerning the relation between the calmness properties of M and $\tilde{M}$, we state the following two propositions. For the first one, we omit its standard proof.

Proposition 1 *Fix any $(\bar{x},\bar{y}) \in M(0)$ and assume that MFCQ holds at $\bar{y} \in \Gamma$ (described by $\mathscr{C}^1$ inequalities). Then there exist a constant L and a neighborhood $\mathscr{U}$ of $(0,0,\bar{x},\bar{y})$ such that $\|\lambda\| \le L$ for all $(z_1,z_2,x,y) \in \mathscr{U}$ and $(x,y,\lambda) \in \tilde{M}(z_1,z_2)$.*

Proposition 2 *Let MFCQ hold at $\bar{y} \in \Gamma$ (described by $\mathscr{C}^1$ inequalities). Then the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for all $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$ implies the calmness of M at $(0,\bar{x},\bar{y})$.*

Proof Assume by contradiction that M is not calm at $(0,\bar{x},\bar{y})$, which means that there exist sequences $x_k \to \bar{x}$, $y_k \to \bar{y}$ and $p_k \to 0$ with $x_k \in \omega$ such that

$$p_k \in F(x_k,y_k) + N_\Gamma(y_k), \tag{15}$$

$$d((x_k,y_k),M(0)) > k\|p_k\|. \tag{16}$$

Since for k sufficiently large MFCQ holds for Γ at y_k, it follows from (15) the existence of λ_k with

$$p_k = F(x_k,y_k) + (\nabla q(y_k))^T \lambda_k, \quad q(y_k) \in N_{\mathbb{R}^s_+}(\lambda_k). \tag{17}$$

In particular, $(x_k,y_k,\lambda_k) \in \tilde{M}(p_k,0)$. From Proposition 1 we obtain that the sequence $\{\lambda_k\}$ is bounded and thus we may assume, by taking a subsequence if necessary, that $\{\lambda_k\}$ converges to some $\bar{\lambda}$. Then, passing to the limit in (17) and taking into account the closedness of the graph of the normal cone mapping, we derive that

$$0 = F(\bar{x},\bar{y}) + (\nabla q(\bar{y}))^T \bar{\lambda}, \quad q(\bar{y}) \in N_{\mathbb{R}^s_+}(\bar{\lambda}).$$

In other words, $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$ (see (5)). Since $M(0)$ is the canonical projection of $\tilde{M}(0,0)$ onto the space of the first two variables, one obtains from (16) and $(x_k,y_k,\lambda_k) \in \tilde{M}(p_k,0)$ that

$$d((x_k,y_k,\lambda_k),\tilde{M}(0,0)) \ge d((x_k,y_k),M(0)) > k\|p_k\|$$

and hence $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for some $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$ which provides a contradiction.
□

The reverse implication of Proposition 2 cannot be expected to hold true even when strengthening MFCQ to LICQ as shown in the following example:

Example 1 Consider the function $q : \mathbb{R} \to \mathbb{R}$ defined as

$$q(y) = \begin{cases} y + y^{3/2} & \text{if } y \ge 0 \\ y - |y|^{3/2} & \text{if } y < 0. \end{cases}$$

Further define $F(x,y) = -1$, $\omega = \mathbb{R}$ and fix the reference point $(\bar{x},\bar{y},\bar{\lambda}) = (0,0,1)$. Since $q'(0) = 1$, LICQ is satisfied around $\bar{y}$. Moreover, it is clear that $\Gamma = (-\infty,0]$ and that q' is continuous at 0 but it is not Lipschitz continuous there. For all p close to 0 it holds true that

$$M(p) = \{(x,y) \,|\, p+1 \in N_\Gamma(y)\} = \mathbb{R}\times\{0\}$$

and thus M is calm at $(0,\bar{x},\bar{y})$. Since $\bar{\lambda} = 1$, we may find a neighborhood $U(\bar{x},\bar{y},\bar{\lambda})$ of the reference point such that

$$\tilde{M}(z_1,z_2)\cap U(\bar{x},\bar{y},\bar{\lambda}) = \{(x,y,\lambda) \,|\, z_1+1 = q'(y)\lambda,\, q(y) = -z_2\}$$

and thus, due to Lemma 2, the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ is equivalent to the calmness of $\hat{M}$ at $(0,0,\bar{y},\bar{\lambda})$ with

$$\hat{M}(z_1,z_2) := \{(y,\lambda) \,|\, z_1+1 = q'(y)\lambda,\, q(y) = -z_2\}.$$

Since q is continuously differentiable and $q'(0)\neq 0$, the inverse function theorem implies that there exists a continuously differentiable function h such that on some neighborhood of 0, relation $-q(y) = z_2$ is equivalent to $h(z_2) = y$. Further we have $h'(z_2) = -\frac{1}{q'(h(z_2))}$, which directly implies

$$\hat{M}(z_1,z_2) = \{(y,\lambda) \,|\, \lambda = -h'(z_2)(z_1+1),\, y = h(z_2).\}.$$

This means that $\hat{M}$ is single-valued and to show that $\hat{M}$ is not calm at $(0,0,\bar{y},\bar{\lambda})$ it is sufficient to show that $p\mapsto h'(p)$ is not calm at 0. Since h' is continuous, we do not have to consider a neighborhood in the range from the definition of calmness. It is easy to see that

$$\frac{|h'(p)-h'(0)|}{|p-0|} = \frac{1}{|q'(h(p))q'(h(0))|}\frac{|q'(h(0))-q'(h(p))|}{|p-0|} \geq \frac{|q'(h(0))-q'(h(p))|}{2|h(p)-h(0)|} \overset{p\to 0}{\to} \infty$$

because q' is not Lipschitz at 0. In the inequality we have used the estimate

$$\frac{1}{|q'(h(p))q'(h(0))|}\frac{|h(p)-h(0)|}{|p-0|} \geq \frac{1}{2},$$

for all p sufficiently close to zero as $q'(0) = 1$ and $h'(0) = -\frac{1}{q'(0)} = -1$ and both q and h are continuously differentiable at 0. But the previous inequality implies directly from (6) that h' is not calm at 0. Thus, we have managed to find an example, in which LICQ holds, M is calm at $(0,\bar{x},\bar{y})$ but $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. △

Note that in this example q was of class $\mathscr{C}^1$ only. This raises the question of whether the reverse direction of Proposition 2 could be established under smoother data. The answer is still negative if one assumes just MFCQ as in Proposition 2. This is shown in the following example.

Example 2 Consider the following data for (1) and (2)

$$q(y_1,y_2) := \begin{pmatrix} y_1^2 - y_2 \\ -y_2 \end{pmatrix}, \quad F(x,y_1,y_2) := \begin{pmatrix} x \\ 1 \end{pmatrix}, \quad (\bar{x},\bar{y}_1,\bar{y}_2) := (0,0,0)$$

and $\omega = \mathbb{R}$. Note that MFCQ is satisfied for Γ at $\bar{y}$ but LICQ is not. Some elementary calculus shows that, locally around $(0,0)$, we have

$$M(p_1,p_2) = \left\{ (x,y_1,y_2) \,\middle|\, y_1 = \frac{p_1 - x}{2(1-p_2)},\, y_2 = \frac{(p_1-x)^2}{4(1-p_2)^2} \right\}.$$

Since we can write $M(p_1,p_2)=\{(x,y_1,y_2)|\ G(p_1,p_2,x,y_1,y_2)=0\}$ for a certain smooth mapping G with $\nabla_{x,y_1,y_2}G(0,0,0,0,0)$ having full row rank, we obtain that M has the Aubin property at $(0,0,0,0,0)$ due to [13, Corollary 4.42] and, hence, is calm there.

It can be easily computed that $\Lambda(\bar{x},\bar{y})=\{\lambda\geq 0|\ \lambda_1+\lambda_2=1\}$. For $k\in\mathbb{N}$ we define

$$(z_{k1},z_{k2},z_{k3},z_{k4},x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2}):=(0,0,-k^{-2},0,0,k^{-1},0,0,1)$$

and observe that

$$(x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2})\in\tilde{M}(z_{k1},z_{k2},z_{k3},z_{k4}).$$

Now, let $(\tilde{x},\tilde{y}_1,\tilde{y}_2,\tilde{\lambda}_1,\tilde{\lambda}_2)\in\tilde{M}(0,0,0,0)$ be arbitrarily given, where $(\tilde{\lambda}_1,\tilde{\lambda}_2)$ is close to $(0,1)$. By construction of the example, one has that $\tilde{x}=\tilde{y}_1=\tilde{y}_2=0$. Consequently, one arrives at

$$\begin{aligned} d((x_k,y_{k1},y_{k2},\lambda_{k1},\lambda_{k2}),\tilde{M}(0,0,0,0)) &= \|(0,-k^{-1},0,0,1)-(0,0,0,0,1)\| \\ &= k^{-1}=k\|(z_{k1},z_{k2},z_{k3},z_{k4})\|, \end{aligned}$$

which implies that $\tilde{M}$ is not calm at $(0,0,0,0,\bar{x},\bar{y}_1,\bar{y}_2,\bar{\lambda}_1,\bar{\lambda}_2)$ with $\bar{\lambda}=(0,1)$. △

It is even possible to strengthen the previous counterexample in the following sense: In the Appendix, we construct a set Γ described by $\mathscr{C}^2$ inequalities satisfying MFCQ at given $\bar{y}$ and a function F such that M is calm at $(0,\bar{x},\bar{y})$ while $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\lambda)$ for **any** $\lambda\in\Lambda(\bar{x},\bar{y})$.

Examples 1 and 2 have shown that a reversion of Proposition 2 is not possible under $\mathscr{C}^1$ data even under LICQ and for smooth data under MFCQ. This raises the question about achieving the desired reversion by combining smooth data with LICQ. This time the answer is affirmative as will be shown in Section 3.3 (actually, $\mathscr{C}^{1,1}$ data will be sufficient). Before addressing this issue, we insert a calmness result for the perturbed complementarity constraints which on the one hand is a basic prerequisite for all following sections but on the other hand also of some independent interest (for instance with respect to a calculus rule for coderivatives, see Corollary 3 below).

3.2 Calmness of perturbed complementarity constraints

In this section we investigate the calmness of the multifunction $T:\mathbb{R}^s\rightrightarrows\mathbb{R}^m\times\mathbb{R}^s$ defined by

$$T(p):=\left\{(y,\lambda)\,|\,q(y)-p\in N_{\mathbb{R}^s_+}(\lambda)\right\}. \tag{18}$$

which represents a perturbation of the complementarity constraints. First, we provide an equivalent characterization of the calmness of T in terms of the calmness systems of perturbed inequality/equality subsystems of the given constraint $q(y)\leq 0$ defining the set Γ. The latter is much more explicit and easier to check than calmness of T itself. To this aim, we introduce for each arbitrary index set $I\subset\{1,\dots,s\}$ the multifunctions $T_I,\hat{T}_I:\mathbb{R}^s\rightrightarrows\mathbb{R}^m$ by

$$\begin{aligned} T_I(p) &:= \{y|\,q_i(y)=p_i\ \ (i\in I)\,,\,q_i(y)\leq 0\ \ (i\notin I)\}\,,\\ \hat{T}_I(p) &:= \{y|\,q_i(y)=p_i\ \ (i\in I)\,,\,q_i(y)\leq p_i\ \ (i\notin I)\}\,. \end{aligned} \tag{19}$$

Lemma 4 *Let $\bar{y}\in q^{-1}(0)$ be arbitrary. Then we have the following statements:*

1. $\hat{T}_I$ is calm at $(0,\bar{y})$ for every $I\subset\{1,\dots,s\}$ $\Longrightarrow$ T_I is calm at $(0,\bar{y})$ for every $I\subset\{1,\dots,s\}$ $\Longrightarrow$ T is calm at all $\left(0,\bar{y},\bar{\lambda}\right)\in\operatorname{gph}T$.

2. *T is calm at some $\left(0,\bar{y},\bar{\lambda}\right) \in \operatorname{gph} T \implies \hat{T}_I$ is calm at $(0,\bar{y})$ for $I := \{i | \bar{\lambda}_i > 0\} \implies T_I$ is calm at $(0,\bar{y})$ for $I := \{i | \bar{\lambda}_i > 0\}$.*

Proof The first implication of 1. and the second implication of 2. are immediate consequences of the fact that calmness of the richer perturbed mapping $\hat{T}_I$ implies that of T_I. The second implication of 1. has been shown in [7, Proposition 3.1]. It remains to show the first implication of 2. To do so, assume that T is calm at $(0,\bar{y},\bar{\lambda})$ and that $\hat{T}_I$ fails to be calm at $(0,\bar{y})$ for the I from the lemma statement. Then there exists a sequence $(p_k,y_k) \to (0,\bar{y})$ such that for all k

$$q_i(y_k) = (p_k)_i \ (i \in I), \qquad q_i(y_k) \le (p_k)_i \ (i \notin I) \tag{20}$$

and

$$d(y_k, \hat{T}_I(0)) > k\|p_k\|. \tag{21}$$

Necessarily we have $p_k \neq 0$ because otherwise both sides of the inequality are zeros.

We claim now that, for k large enough,

$$d((y_k,\bar{\lambda}),T(0)) = d((y_k,\bar{\lambda}),T(0) \cap \{(y,\lambda) | \lambda_i > 0 \ (i \in I)\}). \tag{22}$$

Indeed, if this relation did not hold, then there would exist some $(\tilde{y}_k,\tilde{\lambda}_k) \in T(0)$ such that

$$\|(y_k,\bar{\lambda}) - (\tilde{y}_k,\tilde{\lambda}_k)\| = d((y_k,\bar{\lambda}),T(0)) < d((y_k,\bar{\lambda}),T(0) \cap \{(y,\lambda) | \lambda_i > 0 \ (i \in I)\}),$$

which implies that $(\tilde{\lambda}_k)_j = 0$ for some $j \in I$. On the other hand, $\bar{\lambda}_j > 0$ by assumption. Consequently, due to $(y_k,\bar{\lambda}) \to (\bar{y},\bar{\lambda}) \in T(0)$, we end up at the contradiction

$$0 < \bar{\lambda}_j = |\bar{\lambda}_j - (\tilde{\lambda}_k)_j| \le \|(y_k,\bar{\lambda}) - (\tilde{y}_k,\tilde{\lambda}_k)\| = d((y_k,\bar{\lambda}),T(0)) \to d((\bar{y},\bar{\lambda}),T(0)) = 0.$$

Consequently, there exists a minimizing sequence to the distance function on (22), thus some $(\tilde{y}_k,\tilde{\lambda}_k) \in T(0)$ such that $(\tilde{\lambda}_k)_i > 0$ for all $i \in I$ and

$$d((y_k,\bar{\lambda}),T(0)) \ge \|(y_k,\bar{\lambda}) - (\tilde{y}_k,\tilde{\lambda}_k)\| - \|p_k\|. \tag{23}$$

Since $q(\tilde{y}_k) \in N_{\mathbb{R}^s_+}(\tilde{\lambda}_k)$, it follows that $q_i(\tilde{y}_k) = 0$ for all $i \in I$ and $q_i(\tilde{y}_k) \le 0$ for all $i \notin I$. In other words, $\tilde{y}_k \in \hat{T}_I(0)$. Now, (21) implies that $\|y_k - \tilde{y}_k\| > k\|p_k\|$. Combining this with (23) yields that

$$d((y_k,\bar{\lambda}),T(0)) > k\|p_k\| - \|p_k\|.$$

Now, (20) along with $\bar{\lambda}_i = 0$ for $i \notin I$ implies that $(y_k,\bar{\lambda}) \in T(p_k)$. Altogether, we have shown that

$$(y_k,\bar{\lambda}) \in T(p_k), \quad (p_k,y_k,\bar{\lambda}) \to (0,\bar{y},\bar{\lambda}), \quad d((y_k,\bar{\lambda}),T(0)) > (k-1)\|p_k\|,$$

which violates the calmness of T at $(0,\bar{y},\bar{\lambda})$. This finishes the proof. □

The lemma above may be used in order to check the calmness of T by means of that of certain inequality/equality subsystems. It turns out, however, that this check is not even necessary, whenever our set Γ satisfies CRCQ.

Corollary 2 *Let $\bar{y} \in q^{-1}(0)$ be arbitrary. If Γ satisfies CRCQ at $\bar{y}$, then T is calm at all $(0,\bar{y},\bar{\lambda}) \in \operatorname{gph} T$.*

Proof Fix an arbitrary index set $I \subset \{1,\dots,s\}$ and consider the system

$$q_i(y) = 0 \quad (i \in I), \quad q_i(y) \le 0 \quad (i \notin I). \tag{24}$$

By our assumption $\bar{y} \in q^{-1}(0)$, all constraints are active at $\bar{y}$ both in the inequality system (2) describing the set Γ and in the mixed system (24). Consequently, the assumed CRCQ for (2) at $\bar{y}$ implies CRCQ for (24) at $\bar{y}$. Referring to [11, Proposition 2.5], we conclude that the multifunction T_I is calm at $(0,\bar{y})$. Since $I \subset \{1,\dots,s\}$ was arbitrary, Lemma 4 yields the calmness of T at all $(0,\bar{y},\bar{\lambda}) \in \operatorname{gph} T$. □

Although deriving calmness of T via CRCQ is very convenient, it may happen that CRCQ is violated, yet calmness can still be checked on the basis of Lemma 4. This is the case in the following example:

Example 3 Let $\bar{y} := (0,0)$ and

$$g_1(y_1,y_2) := -y_1; \quad g_2(y_1,y_2) := -y_2; \quad g_3(y_1,y_2) := \begin{cases} -y_2 & (y_1 \ge 0) \\ y_1^2 - y_2 & (y_1 \le 0) \end{cases}.$$

Then, the g_i are continuously differentiable and Γ satisfies MFCQ but violates CRCQ at $\bar{y}$. On the other hand, elementary computations, which we omit here, show that all multifunctions T_I introduced in (19) are calm at $(0,\bar{y})$ for all $I \subset \{1,2,3\}$. Hence, the multifunction T in (18) is calm at all $(0,\bar{y},\bar{\lambda}) \in \operatorname{gph} T$ thanks to Lemma 4. △

Finally, we mention that in [7,14] the authors computed an upper estimate of the coderivative $D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))$ under MFCQ at $\bar{y}$ and under the assumption that T is calm at $(0,\bar{y},\lambda)$ for all $\lambda \in \Lambda(\bar{x},\bar{y})$. By combining [7, Theorem 3.3] and Corollary 2, one arrives directly at the next statement.

Corollary 3 *Assume that $q \in \mathscr{C}^2$ and both MFCQ as well as CRCQ are fulfilled at $\bar{y}$. Then one has with for all $v^* \in \mathbb{R}^m$ the estimate*

$$D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))(v^*) \subset \bigcup_{\lambda \in \Lambda(\bar{x},\bar{y})} \left\{ \left(\sum_{i=1}^s \lambda_i \nabla^2 q_i(\bar{y}) \right) v^* + (\nabla q(\bar{y}))^\top D^*N_{\mathbb{R}^s_-}(q(\bar{y}),\lambda)(\nabla q(\bar{y})v^*) \right\}.$$

3.3 LICQ and $\mathscr{C}^{1,1}$ inequalities or MFCQ and linear inequalities

We now address again the issue discussed at the end of Section 3.1 on the reversion of Proposition 2 when strengthening MFCQ and the smoothness of q. For the main theorem, we will define two auxiliary multifunctions which will be of use when partitioning $\tilde{M}$

$$\begin{aligned} S_1(z_1) &:= \{(x,y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s | F(x,y) + (\nabla q(y))^\top \lambda - z_1 = 0\} \\ S_2(z_2) &:= \left\{ (x,y,\lambda) \in \omega \times \mathbb{R}^m \times \mathbb{R}^s \,\middle|\, \begin{pmatrix} \lambda \\ q(y) - z_2 \end{pmatrix} \in \operatorname{gph} N_{\mathbb{R}^s_+} \right\}. \end{aligned} \tag{25}$$

Theorem 4 *Let q be of class $\mathscr{C}^{1,1}$. Fix an arbitrary $(\bar{x},\bar{y}) \in M(0)$ and assume that LICQ is satisfied at $\bar{y} \in \Gamma$. Then the calmness of M at $(0,\bar{x},\bar{y})$ is equivalent to the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ for the unique (by LICQ) $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$.*

Proof Recall first that, without loss of generality, we may assume $q(\bar{y}) = 0$. One theorem implication follows directly from Proposition 2. Hence, it suffices to show that the calmness of M at $(0,\bar{x},\bar{y})$ implies the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ at the unique (by LICQ) $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$. We will show that there are constants $\kappa \geq 0$ and $\varepsilon_1 > 0$ such that for all $(z_1,z_2,x',y',\lambda') \in \operatorname{gph}\tilde{M} \cap \mathbb{B}_{\varepsilon_1}(0,0,\bar{x},\bar{y},\bar{\lambda})$ we have

$$d((x',y',\lambda'),\tilde{M}(0,0)) \leq \kappa \|(z_1,z_2)\|. \tag{26}$$

We observe first that S_2 defined in 25 is calm at $\left(0,\bar{x},\bar{y},\bar{\lambda}\right)$. Indeed, as LICQ implies CRCQ, Corollary 2 ensures the calmness of the multifunction T defined in (18) at $(0,\bar{y},\bar{\lambda})$. Now, the calmness of S_2 is evident from Lemma 2.

Without loss of generality, we will work with the maximum norm throughout this proof. First we collect all information that is at our disposal in the following relations, where $\varepsilon, L > 0$ are certain positive constants which may be assumed to have common values in all of them:

$$\|F(x_1,y_1) - F(x_2,y_2)\| \leq L\|(x_1,y_1) - (x_2,y_2)\| \quad \forall (x_1,y_1),(x_2,y_2) \in \mathbb{B}_\varepsilon((\bar{x},\bar{y})), \tag{27a}$$
$$\|F(x,y)\| \leq L \quad \forall (x,y) \in \mathbb{B}_\varepsilon((\bar{x},\bar{y})), \tag{27b}$$
$$\|q(y_1) - q(y_2)\| \leq L\|y_1 - y_2\| \quad \forall y_1,y_2 \in \mathbb{B}_\varepsilon(\bar{y}), \tag{27c}$$
$$\|\nabla q(y_1) - \nabla q(y_2)\| \leq L\|y_1 - y_2\| \quad \forall y_1,y_2 \in \mathbb{B}_\varepsilon(\bar{y}), \tag{27d}$$
$$\|\nabla q(y)\| \leq L \quad \forall y \in \mathbb{B}_\varepsilon(\bar{y}), \tag{27e}$$
$$d((x,y),M(0)) \leq L\|z\| \quad \forall (z,x,y) \in \mathbb{B}_\varepsilon(0,\bar{x},\bar{y}) : (x,y) \in M(z), \tag{27f}$$
$$d((x,y,\lambda),S_2(0)) \leq L\|z\| \quad \forall (z,x,y,\lambda) \in \mathbb{B}_\varepsilon(0,\bar{x},\bar{y},\bar{\lambda}) : (x,y,\lambda) \in S_2(z), \tag{27g}$$
$$\|\lambda\| \leq L \quad \forall \lambda\ \forall (z_1,z_2,x,y) \in \mathbb{B}_\varepsilon(0,0,\bar{x},\bar{y}) : (x,y,\lambda) \in \tilde{M}(z_1,z_2). \tag{27h}$$

Here, (27a)-(27e) follow from the differentiability assumptions we have made, (27f) corresponds to the assumed calmness of M at $(0,\bar{x},\bar{y})$. Inequality (27g) means the calmness of S_2 at $\left(0,\bar{x},\bar{y},\bar{\lambda}\right)$ observed above. Finally, formula (27h) is a consequence of Proposition 1.

In order to verify the asserted calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$, define

$$\varepsilon_1 := \min\left\{\frac{\varepsilon}{2}, \frac{\varepsilon}{2L}, \frac{\varepsilon}{1+2L^2+L^3}, \frac{\varepsilon}{1+2L+2L^3+L^4}\right\} \tag{28}$$

and consider an arbitrary triple $(x',y',\lambda') \in \tilde{M}(z_1,z_2)$ with $(z_1,z_2,x',y',\lambda') \in \mathbb{B}_{\varepsilon_1}\left(0,0,\bar{x},\bar{y},\bar{\lambda}\right)$. Since $\tilde{M}(z_1,z_2) = S_1(z_1) \cap S_2(z_2)$ and $S_2(0)$ is a closed set, we may use (27g) to obtain the existence of some $(\tilde{x},\tilde{y},\tilde{\lambda}) \in S_2(0)$ such that

$$\max\left\{\|x'-\tilde{x}\|, \|y'-\tilde{y}\|, \|\lambda'-\tilde{\lambda}\|\right\} \leq L\|z_2\|. \tag{29}$$

By definition of S_2, relation $(\tilde{x},\tilde{y},\tilde{\lambda}) \in S_2(0)$ implies that $q(\tilde{y}) \in N_{\mathbb{R}^s_+}(\tilde{\lambda})$, which further means that $(\tilde{x},\tilde{y},\tilde{\lambda}) \in \tilde{M}(a,0)$ and $(\tilde{x},\tilde{y}) \in M(a)$ with

$$a := F(\tilde{x},\tilde{y}) + [\nabla q(\tilde{y})]^\top \tilde{\lambda}. \tag{30}$$

Moreover, since $(x',y',\lambda') \in S_1(z_1)$, we obtain

$$\begin{aligned}
\|a\| &= \|F(\tilde{x},\tilde{y}) + [\nabla q(\tilde{y})]^\top \tilde{\lambda} + z_1 - F(x',y') - [\nabla q(y')]^\top \lambda'\| \\
&\leq \|z_1\| + \|F(\tilde{x},\tilde{y}) - F(x',y')\| + \|[\nabla q(\tilde{y})]^\top \tilde{\lambda} - \left[\nabla q(y')\right]^\top \lambda'\| \\
&\leq \|z_1\| + \left\|F(\tilde{x},\tilde{y}) - F(x',y')\right\| + \left\|\lambda'\right\| \left\|\nabla q(\tilde{y}) - \nabla q(y')\right\| + \|\lambda'-\tilde{\lambda}\| \, \|\nabla q(\tilde{y})\|.
\end{aligned} \tag{31}$$

Next, the relation $(x',y',\lambda') \in \mathbb{B}_{\varepsilon_1}(\bar{x},\bar{y},\bar{\lambda})$ and (28, first case) imply that

$$(x',y',\lambda') \in \mathbb{B}_{\varepsilon/2}(\bar{x},\bar{y},\bar{\lambda}).$$

Combining (29) with (28, second case) and recalling that $z_2 \in \mathbb{B}_{\varepsilon_1}(0)$ yields

$$(\tilde{x},\tilde{y},\tilde{\lambda}) \in \mathbb{B}_{L\|z_2\|}\left(x',y',\lambda'\right) \subset \mathbb{B}_{\varepsilon/2}\left(x',y',\lambda'\right) \subset \mathbb{B}_{\varepsilon}(\bar{x},\bar{y},\bar{\lambda}). \tag{32}$$

Now, relations (27a), (27d), (27e), (27h), and (28, third case) together with (29) allow us to continue our estimation from (31) and to obtain

$$\|a\| \le \|z_1\| + L^2\|z_2\| + L^3\|z_2\| + L^2\|z_2\| \le \left(1+2L^2+L^3\right)\|(z_1,z_2)\| \le \varepsilon. \tag{33}$$

Therefore, we are now allowed to apply (27f) and make use of the fact that $(\tilde{x},\tilde{y}) \in M(a)$ implies the existence of some $(x^*,y^*) \in M(0)$ such that

$$\max\{\|x^*-\tilde{x}\|,\|y^*-\tilde{y}\|\} \le L\|a\|. \tag{34}$$

Note that (34) along with (33) implies

$$\max\{\|x^*-\tilde{x}\|,\|y^*-\tilde{y}\|\} \le L\left(1+2L^2+L^3\right)\|(z_1,z_2)\|. \tag{35a}$$

Further due to (35a) with (29) we can deduce

$$\max\{\left\|x^*-x'\right\|,\left\|y^*-y'\right\|\} \le L\left(2+2L^2+L^3\right)\|(z_1,z_2)\| \tag{35b}$$

and finally (35b) together with (28, fourth case) and the initial assumption $(z_1,z_2,x',y') \in \mathbb{B}_{\varepsilon_1}(0,0,\bar{x},\bar{y})$ leads to

$$\max\{\|x^*-\bar{x}\|,\|y^*-\bar{y}\|\} \le \left(1+2L+2L^3+L^4\right)\varepsilon_1 \le \varepsilon. \tag{35c}$$

Since LICQ is satisfied at $\bar{y}$, then due to assumption $q(\bar{y})=0$ we have that $\nabla q(\bar{y})$ is surjective and we may assume ε to be small enough to guarantee that the surjectivity pertains for all $\nabla q(y)$ and for all $y \in \mathbb{B}_{\varepsilon}(\bar{y})$. This allows us to define the mapping

$$V(y) := [\nabla q(y)\nabla q(y)^\top]^{-1}\nabla q(y) \quad \forall y \in \mathbb{B}_{\varepsilon}(\bar{y}).$$

With V being continuous on $\mathbb{B}_{\varepsilon}(\bar{y})$, we may assume that $\|V(y)\| \le L'$ for some L' and all $y \in \mathbb{B}_{\varepsilon}(\bar{y})$. Moreover, $y^* \in \mathbb{B}_{\varepsilon}(\bar{y})$ entails that $\nabla q(y^*)$ is surjective and, hence, LICQ is satisfied at y^*. For this reason, the relation $(x^*,y^*) \in M(0)$ implies the existence of a unique multiplier λ^* such that $(x^*,y^*,\lambda^*) \in \tilde{M}(0,0)$. By definition of V and $\tilde{M}$, we have that

$$\lambda^* = -V(y^*)F(x^*,y^*); \quad \tilde{\lambda} = V(y^*)\nabla q(y^*)^\top\tilde{\lambda}.$$

Hence,

$$\|\lambda^*-\tilde{\lambda}\| \le L'\|\nabla q(y^*)^\top\tilde{\lambda} + F(x^*,y^*)\|. \tag{36}$$

To estimate the right-hand side of (36), we realize first that (32) and (35c) allow us to employ the relations (27). We use (30), (33), (27h) coupled with $(\tilde{x},\tilde{y},\tilde{\lambda}) \in \tilde{M}(a,0)$, (27d), (27a) and (35a) to obtain some constant $c>0$ such that

$$\begin{aligned}\|\nabla q(y^*)^\top\tilde{\lambda} + F(x^*,y^*)\| &= \|a + (\nabla q(y^*) - \nabla q(\tilde{y}))^\top\tilde{\lambda} + F(x^*,y^*) - F(\tilde{x},\tilde{y})\| \\ &\le \|a\| + \|\tilde{\lambda}\|\|\nabla q(y^*) - \nabla q(\tilde{y})\| + \|F(x^*,y^*) - F(\tilde{x},\tilde{y})\| \\ &\le c\|(z_1,z_2)\|.\end{aligned} \tag{37}$$

Then, estimates (29), (36) and (37) yield

$$\|\lambda^* - \lambda'\| \le \|\lambda^* - \tilde{\lambda}\| + \|\tilde{\lambda} - \lambda'\| \le L'c\|(z_1, z_2)\| + L\|z_2\|.$$

Adding this to (35b), we arrive at existence of some κ such that

$$\left\|(x', y', \lambda') - (x^*, y^*, \lambda^*)\right\| \le \kappa \|(z_1, z_2)\| \tag{38}$$

Since $(x^*, y^*, \lambda^*) \in \tilde{M}(0,0)$, we have shown (26). This finishes the proof. □

We next provide a second instance under which the desired equivalence of calmness for M and $\tilde{M}$ can be guaranteed.

Theorem 5 *Let Γ be a polyhedral set, i.e., $q(y) = Ay - b$ for some matrix A of order (s, m) and some $b \in \mathbb{R}^s$. Assume that Γ has nonempty interior, that $A\bar{y} = b$ and that the rows a_i of A satisfy*

$$\operatorname{rank}\{a_i\}_{i \in I} = \min\{m, \#I\} \quad \forall I \subseteq \{1, \ldots, s\}. \tag{39}$$

Then, the calmness of M at $(0, \bar{x}, \bar{y})$ is equivalent to the calmness of $\tilde{M}$ at $(0, 0, \bar{x}, \bar{y}, \bar{\lambda})$ for all $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$.

Proof Observe first that our assumption on Γ having nonempty interior is equivalent with Γ satisfying MFCQ at all its points. By Proposition 2 it is sufficient to prove that the calmness of M at $(0, \bar{x}, \bar{y})$ implies the calmness of $\tilde{M}$ at $(0, 0, \bar{x}, \bar{y}, \bar{\lambda})$ for any $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$. We fix an arbitrary such $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$. If $s \le m$, then (39) implies the surjectivity of A so that LICQ is satisfied at $\bar{y}$. Hence, the assertion follows from Theorem 4. Therefore, we may assume the opposite case ($s > m$), in which (39) implies the injectivity of A. We are going to prove the assertion of this theorem by means of Theorem 3 applied to the multifunctions S_1, S_2 defined in (25). We will check next, all hypotheses of that Theorem.

Introducing the function $f(x, y, \lambda) := F(x, y) + A^\top \lambda$, we observe that $f = S_1^{-1}$. Since f is single-valued and continuously differentiable, it follows that S_1^{-1} trivially fulfills the Aubin property. Furthermore, the Jacobian

$$\nabla f(\bar{x}, \bar{y}, \bar{\lambda}) = \left(\nabla_x F(\bar{x}, \bar{y}) \middle| \nabla_y F(\bar{x}, \bar{y}) \middle| A^\top\right)$$

is surjective by injectivity of A. Hence, S_1 is calm at $(0, \bar{x}, \bar{y}, \bar{\lambda})$ as a consequence of 2. in Lemma 1. Since CRCQ is satisfied for Γ by linearity of the describing inequalities, S_2 is calm at $(0, \bar{x}, \bar{y}, \bar{\lambda})$ due to Corollary 2 with the same argument already used in the proof of Theorem 4 (see below (26)).

It remains to verify 4. in Theorem 3, i.e., the calmness of $S_1 \cap S_2(0)$ at $(0, \bar{x}, \bar{y}, \bar{\lambda})$. To do so, let $\varepsilon, L > 0$ refer to the definition of the supposed calmness of M at $(0, \bar{x}, \bar{y})$. Select an arbitrary $(z, x, y, \lambda) \in \mathbb{B}_\varepsilon(0, \bar{x}, \bar{y}, \bar{\lambda})$ such that $(x, y, \lambda) \in S_1(z) \cap S_2(0)$. We conclude that $\lambda \ge 0$ and $(x, y) \in M(z)$. Thus, by calmness of M at $(0, \bar{x}, \bar{y})$, there exists some $(x^*, y^*) \in M(0)$ such that

$$\|(x^*, y^*) - (x, y)\| \le L\|z\|. \tag{40}$$

Note that $(x^*, y^*) \in M(0)$ entails that $y^* \in \Gamma$. Since Γ is defined by linear inequalities, it follows that

$$\Lambda(x^*, y^*) = \{\mu \,|\, A^\top \mu = -F(x^*, y^*),\ Ay^* - b \in N_{\mathbb{R}^s_+}(\mu)\} \ne \emptyset$$

We claim that $\Lambda(x^*,y^*) = P$, where

$$P := \{\mu |\, A^\top \mu = -F(x^*,y^*), \mu \geq 0\}.$$

Clearly, $\Lambda(x^*,y^*) \subseteq P$. The reverse inclusion is evident if $y^* = \bar{y}$ due to $A\bar{y} = b$. If $y^* \neq \bar{y}$, then define the set of active rows a_i of A at y^* as

$$I := \{i |\, \langle a_i, y^*\rangle = b_i\}.$$

If $\#I \geq m$, then rank$\{a_i | i \in I\} = m$ by (39) and the linear equality system $\langle a_i, y\rangle = b_i (i \in I)$ has the unique solution $\bar{y}$ by our assumption $A\bar{y} = b$. Since y^* also solves this system, we necessarily have $y^* = \bar{y}$, which is a contradiction. Thus, $\#I < m$. Select an arbitrary $\lambda' \in \Lambda(x^*,y^*) \neq \emptyset$ and $\mu \in P$. We will show that necessarily $\lambda' = \mu$ finally implying the desired equality $\Lambda(x^*,y^*) = P$. By definition we have

$$A^\top(\lambda' - \mu) = 0. \tag{41}$$

Multiplying this relation by y^* and using $\lambda'_i = 0$, $\mu_i \geq 0$ and $\langle a_i, y^*\rangle < b_i$ for $i \notin I$, we arrive at

$$\begin{aligned} 0 = (Ay^*)^\top(\lambda' - \mu) &= \sum_{i\in I}(\lambda'_i - \mu_i)b_i + \sum_{i\notin I}(\lambda'_i - \mu_i)\langle a_i, y^*\rangle \\ &\geq \sum_{i\in I}(\lambda'_i - \mu_i)b_i + \sum_{i\notin I}(\lambda'_i - \mu_i)b_i = b^\top(\lambda' - \mu) = (A\bar{y})^\top(\lambda' - \mu) = 0, \end{aligned}$$

where the last equality follows from (41). This means that we can replace the inequality by an equality and as a part of it we get the relation

$$\sum_{i\notin I}\mu_i\langle a_i, y^*\rangle = \sum_{i\notin I}\mu_i b_i$$

which yields $\mu_i = 0$ for all $i \notin I$. But then (41) reduces to

$$\sum_{i\in I}(\lambda'_i - \mu_i)a_i = 0. \tag{42}$$

Since $\#I < m$, the $\{a_i | i \in I\}$ are linearly independent thanks to (39) and thus (42) yields that $\mu_i = \lambda'_i$ for $i \in I$. Combining this with $\mu_i = \lambda'_i = 0$ for $i \notin I$ we conclude that $\lambda' = \mu$, as was to be shown.

Now, Hoffman's Lemma guarantees the existence of some constant c (only depending on A) such that

$$d(\mu, \Lambda(x^*,y^*)) = d(\mu, P) \leq c\|A^\top\mu + F(x^*,y^*)\| \quad \forall \mu \geq 0.$$

In particular, this applies to our multiplier $\lambda \geq 0$ selected above:

$$d(\lambda, \Lambda(x^*,y^*)) \leq c\|A^\top\lambda + F(x^*,y^*)\| = c\|z - F(x,y) + F(x^*,y^*)\|.$$

Here, we exploited that $(x,y,\lambda) \in S_1(z)$. Consequently, there exists some $\lambda^* \in \Lambda(x^*,y^*)$ such that

$$\|\lambda - \lambda^*\| \leq c\|z - F(x,y) + F(x^*,y^*)\| \leq c\|z\| + cL'\|(x,y) - (x^*,y^*)\|,$$

where L' denotes a local Lipschitz constant of F around $(\bar{x},\bar{y})$. Along with (40), it results in

$$\|(x^*,y^*,\lambda^*) - (x,y,\lambda)\| \leq \tilde{L}\|z\|$$

for some constant $\tilde{L}$. Since $(x^*,y^*) \in M(0)$ and $\lambda^* \in \Lambda(x^*,y^*)$ amount to $(x^*,y^*,\lambda^*) \in S_1(0) \cap S_2(0)$, we have shown that

$$d((x,y,\lambda),S_1(0)\cap S_2(0)) \le \tilde{L}\|z\|,$$

which is the asserted calmness of $S_1 \cap S_2(0)$ at $(0,\bar{x},\bar{y},\bar{\lambda})$. Thus, we have finally verified all assumptions of Theorem 3 and may conclude the desired calmness of the mapping $\tilde{M}(z_1,z_2) = S_1(z_1)\cap S_2(z_2)$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. □

Observe that the previous Theorem does not relate to a fully linear generalized equation in (1) which would automatically guarantee the desired calmness of $\tilde{M}$ thanks to Robinson's Theorem on upper Lipschitz continuity of polyhedral multifunctions. Rather, we allow that the mapping F is nonlinear but, in such a case, the calmness of M needs to be satisfied in addition. As an example for a polyhedral set Γ violating LICQ at 0 but satisfying the assumptions of Theorem 5, one may take the set defined by the inequality $y_3 \ge \max\{|y_1|,|y_2|\}$ (resolved as a linear system).

4 Main results

In the first part of this section we address the question how the calmness property of M and $\tilde{M}$ can be ensured by suitable point-based conditions. Concerning the calmness of M, we present here only a standard result in which one enforces in fact even the (substantially more restrictive) Aubin property. In [18] and [13], exclusively this type of qualification conditions is used. We are aware about the possibility to employ to this purpose some less restrictive calmness criteria from, e.g., [4,10].

Theorem 6 *Assume that the implication*

$$\left.\begin{array}{r}(\nabla_x F(\bar{x},\bar{y}))^\top a \in -N_\omega(\bar{x})\\ -(\nabla_y F(\bar{x},\bar{y}))^\top a \in D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))(a)\end{array}\right\} \implies a=0 \tag{43}$$

is fulfilled. Then M has the Aubin property around $(0,\bar{x},\bar{y})$ *and hence it is also calm at this point.*

Proof The assertion follows immediately from the Mordukhovich criterion [18, Theorem 9.40] and the standard first-order calculus. □

For the verification of the calmness of $\tilde{M}$, however, we present here a new condition based on Lemma 4. To this aim, we define the Lagrangian as

$$\mathscr{L}(x,y,\lambda) := F(x,y) + (\nabla q(y))^\top \lambda. \tag{44}$$

Theorem 7 *Assume that* $(\bar{x},\bar{y},\bar{\lambda}) \in \tilde{M}(0,0)$*, that* $q \in \mathscr{C}^2$ *and that the implication*

$$\left.\begin{array}{ll}(\nabla_x F(\bar{x},\bar{y}))^\top a \in -N_\omega(\bar{x}) & \\ (\nabla_y \mathscr{L}(\bar{x},\bar{y},\bar{\lambda}))^\top a + (\nabla q(\bar{y}))^\top c = 0 & \\ 0 = \nabla q_i(\bar{y})a & \forall i: \bar{\lambda}_i > 0\\ 0 = c_i & \forall i: q_i(\bar{y}) < 0\\ 0 \le c_i, 0 \le \nabla q_i(\bar{y})a \quad \text{or} \quad 0 = c_i \quad \text{or} \quad 0 = \nabla q_i(\bar{y})a & \forall i: \bar{\lambda}_i = q_i(\bar{y}) = 0.\end{array}\right\}$$

$$\Longrightarrow a = 0. \tag{45}$$

holds true. Assume, moreover, that the multifunctions $T_I : \mathbb{R}^s \to \mathbb{R}^m$ defined in (19) *are calm at $(0,\bar{y})$ for all $I \subset \{1,\ldots,s\}$ (which holds automatically true under CRCQ by Corollary 2). Then $\tilde{M}$ is calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$.*

Proof Taking into account that $\tilde{M}(z_1,z_2) = S_1(z_1) \cap S_2(z_2)$ with S_1 and S_2 defined in (25), to obtain the calmness of $\tilde{M}$ at $(0,0,\bar{x},\bar{y},\bar{\lambda})$ it suffices to verify the assumptions of Lemma 3 for the following data: $u_1 = (x,y)$, $u_2 = \lambda$, $H_1(u) = \mathscr{L}(x,y,\lambda)$, $H_2(u) = q(y)$, $\Delta = \omega \times \mathbb{R}^m \times \mathbb{R}^s$ and $\Omega = \mathbb{R}^s_+$. It is not difficult to show that condition (12) takes the form (45) and so it remains to show that S_1 and S_2 are calm at $(0,\bar{x},\bar{y},\bar{\lambda})$.

In order to verify that S_1 has this property, we will apply Lemma 1 according to which it is sufficient to show that $\nabla\mathscr{L}(\bar{x},\bar{y},\bar{\lambda})$ has full row rank. Hence consider any a such that $\nabla\mathscr{L}(\bar{x},\bar{y},\bar{\lambda})^\top a = 0$. But then $(a,0)$ satisfies the relations on the left-hand side of (45) and thus $a = 0$, implying that S_1 is indeed calm at $(0,\bar{x},\bar{y},\bar{\lambda})$. On the other hand, Lemma 4 yields the calmness of T defined in (18) at $(0,\bar{y},\bar{\lambda})$ and, hence, S_2 is calm at $(0,\bar{x},\bar{y},\bar{\lambda})$ by Lemma 2. □

Note that if ω is a convex set, then N_ω is the standard normal cone in the sense of convex analysis. Moreover, if $\omega = \mathbb{R}^n$, then $N_\omega(\bar{x}) = \{0\}$ and the inclusion reduces to an equality. In the MPEC literature, one finds under various names (GMFCQ, NNAMCQ) a qualification condition similar to (45) with the difference that $a = c = 0$ is required instead of only $a = 0$. Clearly, under LICQ at $\bar{y}$, both these conditions coincide. However, if we impose only MFCQ and CRCQ at $\bar{y}$, (45) is strictly better (less restrictive) than GMFCQ.

In the remainder of this section we will state the main result of the paper. It comprises in a concise form the information which we have gained in the course of our analysis about the relationship between Theorems 1 and 2. It leads to several useful conclusions in deriving workable M-stationarity conditions for MPEC (1).

Theorem 8 *Let $(\bar{x},\bar{y})$ be a local solution to* (1) *and assume that $q \in \mathscr{C}^2$ and that MFCQ holds at $\bar{y} \in \Gamma$.*

1. *If CRCQ holds at $\bar{y}$, then for those $\lambda \in \Lambda(\bar{x},\bar{y})$ satisfying the qualification condition* (45)*, there exist v and w fulfilling the stationarity conditions* (9)*.*
2. *If CRCQ holds at $\bar{y}$ and M is calm at $(0,\bar{x},\bar{y})$, then there exist $\lambda \in \Lambda(\bar{x},\bar{y})$, v and w fulfilling the stationarity conditions* (9)*.*
3. *If Γ is a polyhedral set with nonempty interior satisfying* (39) *and M is calm at $(0,\bar{x},\bar{y})$, then for all $\lambda \in \Lambda(\bar{x},\bar{y})$ there exist v and w fulfilling the stationarity conditions* (9)*.*
4. *If even LICQ holds at $\bar{y} \in \Gamma$, then Theorems 1 and 2 are completely equivalent in their assumptions and their results.*

Before proving this Theorem, we include some comments on the statements 1-3. The big progress of statement 1 over Theorems 1 and 2 or Corollary 1 is that under MFCQ and CRCQ it completely frees us from the necessity of checking any calmness condition or computing the complicated coderivative $D^*N_\Gamma(\bar{y},-F(\bar{x},\bar{y}))$. It just relies on checking the explicit qualification condition (45) and provides explicit stationarity conditions (9). For instance, in order to exclude $(\bar{x},\bar{y})$ from being a local solution to (1), it will be sufficient to find some $\lambda \in \Lambda(\bar{x},\bar{y})$ satisfying (45) and violating (9) for all v and w. Unfortunately, it is not excluded that the set of $\lambda \in \Lambda(\bar{x},\bar{y})$ satisfying (45) is empty so that statement 1 cannot be applied. But even then, one might be successful in checking the calmness of M and thus in applying statement 2. Excluding $(\bar{x},\bar{y})$ from being a local solution to (1) would then amount to verifying that (9) is violated for all $\lambda \in \Lambda(\bar{x},\bar{y})$ and all v and w. Statement 3 provides an instance under which we do not have to care about specific $\lambda \in \Lambda(\bar{x},\bar{y})$. This facilitates the task of excluding $(\bar{x},\bar{y})$ from being a local solution to (1) in the sense that we just have to find some $\lambda \in \Lambda(\bar{x},\bar{y})$ such that (9) is violated for any v and w.

Proof (of Theorem 8) First recall that under MFCQ at $\bar{y}$, $(\bar{x},\bar{y},\lambda)$ is a local solution of MPEC (4) for all $\lambda \in \Lambda(\bar{x},\bar{y})$. Concerning statement 1, observe that under CRCQ at $\bar{y}$ we have that $\tilde{M}$ is calm at all points $(0,0,\bar{x},\bar{y},\lambda)$ with $\lambda \in \Lambda(\bar{x},\bar{y})$ satisfying (45) by virtue of Theorem 7. Statement 1 thus follows from Theorem 2. Statement 2 is a direct consequence of Theorem 1 and Corollary 3, where one needs just to express the coderivative $D^*N_{\mathbb{R}^s_-}(q(\bar{y}),\lambda)$ in Corollary 3 in terms of $q(\bar{y})$ and λ. To prove statement 3, it suffices to combine Theorem 2 with Theorem 5. Finally, in statement 4, the equivalence of the calmness assumptions in Theorems 1 and 2 follows from Theorem 4. On the other hand, the equivalence of the obtained stationarity conditions in both theorems relies on a well-known formula for making explicit the coderivative D^*N_Γ in case that Γ is described by smooth inequalities satisfying LICQ (see, e.g., [7, Theorem 3.1]). □

5 MPECs with structured equilibria

Some of the tools and/or results from the preceding part of the paper can be utilized in deriving stationarity conditions for MPECs with equilibria governed by generalized equations having a special structure. In Section 5.1 we illustrate this fact by such an equilibrium with a polyhedral constraint set. In Section 5.2 we then apply these results to a class of bilevel programming problems arising in electricity spot market modelling.

5.1 Structured equilibria with polyhedral constraint sets

Let us consider a generalized equation of the considered type where

$$F(x,y) = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix}, \; q(y) = Ay - b \tag{46}$$

with $F_1 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{m_1}$, $F_2 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{m_2}$, $A = (A_1, A_2)$ and $y = (y_1, y_2) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$. Even though there is no structural difference between F_1 and F_2 yet, we will impose different assumptions on them later in the text. Structure (46) with $F_2(x,y) \equiv F_2(y)$ arises typically in a hierarchical bilevel multileader game where one looks for a Nash equilibrium on the upper level. In this case we obtain a finite number of MPECs in which the equilibria on the lower level are governed by generalized equation having the special structure (46), see e.g. [8].

It is appropriate to define the mappings S_1, S_2, employed in Section 3, in a different way here, namely:

$$\begin{aligned} S_1(z_1) &:= \left\{ (x,y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \,\middle|\, z_1 = F_1(x,y) + A_1^\top \lambda \right\}, \\ S_2(z_2,z_3) &:= \left\{ (x,y,\lambda) \in \omega \times \mathbb{R}^m \times \mathbb{R}^s \,\middle|\, z_2 = F_2(x,y) + A_2^\top \lambda,\; q(y) - z_3 \in N_{\mathbb{R}^s_+}(\lambda) \right\}. \end{aligned} \tag{47}$$

We will derive two results with differing assumptions and results.

Theorem 9 *In the setting of* (46) *fix some* $(\bar{x},\bar{y}) \in M(0)$ *and* $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$. *Assume that* $\omega = \mathbb{R}^n$, $F_2(x,y) \equiv F_2(y)$ *is affine linear and that* $\nabla_x F_1(\bar{x},\bar{y})$ *is surjective. Then* $\tilde{M}$ *is calm at* $(0,0,0,\bar{x},\bar{y},\lambda)$ *for all* $\lambda \in \Lambda(\bar{x},\bar{y})$. *If in addition* Γ *has nonempty interior, then* M *is calm at* $(0,\bar{x},\bar{y})$.

Proof Clearly $\tilde{M}(z_1,z_2,z_3) = S_1(z_1) \cap S_2(z_2,z_3)$. We will apply Lemma 3. By Lemma 1 and the assumed surjectivity of $\nabla_x F_1(\bar{x},\bar{y})$ we obtain that S_1 is calm at $(0,\bar{x},\bar{y},\bar{\lambda})$. As S_2 has polyhedral graph, it is calm at every point of its graph and it remains to verify condition (12), which takes the form

$$\left.\begin{aligned} (\nabla_x F_1(\bar{x},\bar{y}))^\top a &= 0 \\ (\nabla_y F_1(\bar{x},\bar{y}))^\top a + (\nabla_y F_2(\bar{y}))^\top d + A^\top c &= 0 \\ -A_1 a - A_2 d &\in D^* N_{\mathbb{R}^s_+}(\bar{\lambda}, A\bar{y} - b)(-c) \end{aligned}\right\} \implies a = 0.$$

However, we easily conclude that this condition is fulfilled by virtue of the surjectivity of $\nabla_x F_1(\bar{x},\bar{y})$. The last statement follows directly from Proposition 2 and the equivalence of nonempty interior and MFCQ for polyhedral sets. ☐

Under the assumption of Theorem 9 we may thus take advantage of the sharp M-stationarity conditions (8) where, thanks to the affine linearity of q, $D^* N_\Gamma$ can be computed on the basis of an explicit formula (see [9, Prop. 3.2]). In the next result we relax the assumptions of this theorem. Note that Theorem 9 immediately follows from Theorem 10.

Theorem 10 *In the setting of* (46) *fix some* $(\bar{x},\bar{y}) \in M(0)$ *and* $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$. *Assume first that the function* $G(x,y,\lambda) := F_1(x,y) + A_1^\top \lambda$ *satisfies the assumptions of Lemma 1 and that the following system is satisfied*

$$\left.\begin{aligned} (\nabla_x F_1(\bar{x},\bar{y}))^\top a + (\nabla_x F_2(\bar{x},\bar{y}))^\top d &\in -N_\omega(\bar{x}) \\ (\nabla_y F_1(\bar{x},\bar{y}))^\top a + (\nabla_y F_2(\bar{x},\bar{y}))^\top d + A^\top c &= 0 \\ -A_1 a - A_2 d &\in D^* N_{\mathbb{R}^s_+}(\bar{\lambda}, A\bar{y} - b)(-c) \end{aligned}\right\} \implies a = 0. \quad (48)$$

Moreover, assume that at least one of the three following assumptions is satisfied:

1. F_2 *is affine linear;*
2. $\omega = \mathbb{R}^n$, *condition* (39) *is satisfied and* $\nabla_x F_2(\bar{x},\bar{y})$ *has full row rank;*
3. Γ *has nonempty interior, condition* (39) *is satisfied and for all* $c \in \operatorname{Ker} \nabla_x F_2(\bar{x},\bar{y})^\top \setminus \{0\}$ *we have*

$$c^\top \nabla_{y_2} F_2(\bar{x},\bar{y}) c > 0. \quad (49)$$

Then $\tilde{M}$ *is calm at* $(0,0,0,\bar{x},\bar{y},\bar{\lambda})$.

Proof Again we will employ Lemma 3 with the same representation of $\tilde{M}$ in terms of S_1 and S_2 as in Theorem 9. Since (12) takes the form of (48), it remains to verify the calmness of S_2 at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. It is easy to see that this property holds under assumption 1.

Concerning assumption 2. and 3., we define

$$\hat{S}_2(z_1,z_2) := \left\{ (x,y,v) \in \omega \times \mathbb{R}^m \times \mathbb{R}^s \;\middle|\; \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} v \\ F_2(x,y) \end{pmatrix} + N_\Gamma(y) \right\}$$

and show that $\hat{S}_2$ possesses the Aubin property around $(0,0,\bar{x},\bar{y},-A_1\bar{\lambda}) = (0,0,\bar{x},\bar{y},F_1(\bar{x},\bar{y}))$. By Theorem 6, this is equivalent with the following implication

$$\left.\begin{array}{c}(\nabla_x F_2(\bar{x},\bar{y}))^\top c \in N_\omega(\bar{x}) \\ \begin{pmatrix}(\nabla_y F_2(\bar{x},\bar{y}))^\top c \\ 0 \\ c\end{pmatrix} \in N_{\mathrm{gph}\, N_\Gamma}(\bar{y},-F_1(\bar{x},\bar{y}),-F_2(\bar{x},\bar{y}))\end{array}\right\} \implies c = 0. \tag{50}$$

This implication is satisfied under assumption 2. If assumption 3. holds true and if c satisfies the left-hand side of (50), then the polyhedrality of Γ and [9, Proposition 3.2] tells us that

$$0 \geq c^\top \nabla_y F_2(\bar{x},\bar{y}) \begin{pmatrix}0\\c\end{pmatrix} = c^\top (\nabla_{y_1} F_2(\bar{x},\bar{y}), \nabla_{y_2} F_2(\bar{x},\bar{y})) \begin{pmatrix}0\\c\end{pmatrix} = c^\top \nabla_{y_2} F_2(\bar{x},\bar{y}) c.$$

From (49) follows that $c = 0$, and thus in both cases 2. and 3. we have the Aubin property of $\hat{S}_2$ at $(0,0,\bar{x},\bar{y},-A_1\bar{\lambda})$, which implies calmness at the same point.

Since q is affine linear and (39) holds, we may apply Theorem 5 with $M = \hat{S}_2$ and $\tilde{M} = \tilde{S}$ defined by

$$\tilde{S}_2(z_1,z_2,z_3) := \left\{ (x,y,\lambda,v) \,\middle|\, x \in \omega,\ \begin{pmatrix}z_1\\z_2\end{pmatrix} = \begin{pmatrix}v\\F_2(x,y)\end{pmatrix} + \begin{pmatrix}A_1^\top\\A_2^\top\end{pmatrix}\lambda,\ q(y) - z_3 \in N_{\mathbb{R}^s_+}(\lambda) \right\}$$

to obtain that $\tilde{S}$ is calm at $(0,0,0,\bar{x},\bar{y},\bar{\lambda},-A_1^\top\bar{\lambda})$. But since

$$\tilde{S}_2(z_1,z_2,z_3) = \left\{ (x,y,\lambda,v) \,\middle|\, (x,y,\lambda) \in S_2(z_2,z_3),\ v = z_1 - A_1^\top \lambda \right\},$$

the calmness of $\tilde{S}_2$ at $(0,0,\bar{x},\bar{y},\bar{\lambda},-A_1\bar{\lambda})$ implies the calmness of S_2 at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Thus, we have verified all assumptions of Lemma 3 and thus $\tilde{M} = S_1 \cap S_2$ is indeed calm at $(0,0,0,\bar{x},\bar{y},\bar{\lambda})$. □

5.2 Application to a class of bilevel programming problems

As an application of the results from the previous section we introduce a special class of bilevel programming problems automatically satisfying the calmness conditions required for deriving necessary optimality conditions according to Theorem 1. Consider an MPEC

$$\begin{aligned} \underset{x,y}{\text{minimize}}\ & \varphi(x,y) \\ \text{subject to}\ & y \in \operatorname{argmin}_{y^*}\{f(x,y^*) \mid y^* \in \Gamma\}, \\ & x \in \omega \end{aligned} \tag{51}$$

with

$$f(x,y) := \langle x_1, By_1\rangle + f_1(x_2,y_1) + f_2(y_2).$$

Here, $x = (x_1,x_2)$, $y = (y_1,y_2)$, Γ is a polyhedral set described by the linear inequality system $\Gamma := \{y \mid Ay \leq b\}$ with nonempty interior and $A = (A_1,A_2)$, φ is a continuously differentiable function, f_1 is twice continuously differentiable and convex in the second variable, f_2 is twice continuously differentiable and ω is a closed set. Moreover, we assume that $(A_1^\top, B^\top)$ has full row rank and that at least one of the following conditions is satisfied:

1. f_2 is convex quadratic;

2. f_2 is strongly convex and condition (39) is satisfied.

Due to the convexity of the lower level, we may equivalently recast it into

$$0 \in \begin{pmatrix} F_1(x,y) \\ F_2(y) \end{pmatrix} + N_\Gamma(y) := \begin{pmatrix} B^\top x_1 + \nabla_{y_1} f_1(x_2,y_1) \\ \nabla_{y_2} f_2(y_2) \end{pmatrix} + N_\Gamma(y).$$

Then we have the following optimality conditions of the MPEC above.

Theorem 11 *Let $(\bar{x},\bar{y})$ be a solution to* (51). *Apart from the assumptions above, we assume that implication*

$$\begin{pmatrix} Ba \\ \nabla^2_{x_2y_1} f_1(\bar{x}_2,\bar{y}_1)^\top a \end{pmatrix} \in N_\omega(\bar{x}) \implies a = 0, \tag{52}$$

holds true. Then there exist multipliers $u^* = (u_1^*, u_2^*)$ and $v^* = (v_1^*, v_2^*)$ such that

$$\begin{aligned} 0 &\in \begin{pmatrix} \nabla_{x_1} \varphi(\bar{x},\bar{y}) + Bv_1^* \\ \nabla_{x_2} \varphi(\bar{x},\bar{y}) + \nabla^2_{x_2y_1} f_1(\bar{x}_2,\bar{y}_1)^\top v_1^* \end{pmatrix} + N_\omega(\bar{x}), \\ 0 &= \nabla_{y_1} \varphi(\bar{x},\bar{y}) + \nabla^2_{y_1y_1} f_1(\bar{x}_2,\bar{y}_1) v_1^* + u_1^*, \\ 0 &= \nabla_{y_2} \varphi(\bar{x},\bar{y}) + \nabla^2_{y_2y_2} f_2(\bar{y}_2) v_1^* + u_2^*, \\ u^* &\in D^* N_\Gamma(\bar{y}, -F(\bar{x},\bar{y}))(v_1^*, v_2^*). \end{aligned}$$

Proof We want to employ Theorem 10. Since $(A_1^\top, B^\top)$ has full row rank due to the assumptions, the Jacobian of $G(x,y,\lambda) := B^\top x_1 + \nabla_{y_1} f_1(x_2,y_1) + A_1^\top \lambda$ has full row rank and thus satisfies the assumptions of Lemma 1. Moreover, (52) implies (48). If f_2 is convex quadratic, then F_2 is affine linear. On the other hand, if f_2 is strongly convex, then $\nabla^2_{y_2y_2} F_2(\bar{y}_2)$ is positive definite, which implies (49). Thus, we have verified all assumptions of Theorem 10 and this theorem implies the calmness of $\tilde{M}$ at $(0,0,0,\bar{x},\bar{y},\bar{\lambda})$ for all $\bar{\lambda} \in \Lambda(\bar{x},\bar{y})$. As Γ has nonempty interior, we may apply Proposition 2 to obtain that M is calm $(0,0,\bar{x},\bar{y})$. The rest then follows from Theorem 1. □

For a specific application, we mention the electricity spot market problem which may be modelled via the *Equilibrium Problems with Equilibrium Constraints* (EPECs), see [1, 8]. In this model, we have N power producers. Producer i provides the so-called bidding curve $c_i(q_i)$, which determines the unit price for which he is willing to sell quantity q_i. After all producers submit their bids, the ISO (independent system operator) decides how much electricity each producer may create. We assume that the bidding curves are quadratic, i.e.,

$$c_i(g_i) = \alpha_i g_i + \beta_i g_i^2$$

for some parameters $\alpha_i, \beta_i \geq 0$. The true production cost for each producer is assumed to be equal to

$$C_i(g_i) = \gamma_i g_i + \delta_i g_i^2$$

for known parameters $\gamma_i, \delta_i \geq 0$. In the pay-as-clear model, each producer maximizes the difference between the clearing price and the costs

$$c_i'(g_i) g_i - C_i(g_i) = (\alpha_i - \gamma_i) g_i + (2\beta_i - \delta_i) g_i^2.$$

The ISO wants to minimize the total cost which has to be payed to producers provided that the demand is satisfied. This leads to the following bilevel problem

$$\begin{aligned} &\underset{\alpha_i,\beta_i}{\text{maximize}}\ (\alpha_i-\gamma_i)g_i+(2\beta_i-\delta_i)g_i^2 \\ &\text{subject to}\ \ (g,t)\in\operatorname{argmin}_{(\tilde g,\tilde t)}\left\{\textstyle\sum_{j=1}^N \alpha_j\tilde g_j+\beta_j\tilde g_j^2 |\ (\tilde g,\tilde t)\in\Gamma\right\}, \\ &\qquad\qquad \alpha_i\ge 0,\ \beta_i\ge 0 \end{aligned} \tag{53}$$

for variables $(\alpha_i,\beta_i)\in\mathbb{R}^2$, where the constraint set

$$\Gamma:=\{(g,t)|\ g+Bt\ge d,\ g\ge 0\}$$

ensures that the demand d is satisfied at all nodes. Here, g is the produced amount at all nodes, B is the incidence matrix of the network, and thus Bt describes the amount of electricity transmitted between nodes. Naturally, the produced amount q has to be nonnegative.

We arrive at the following result. Note that no constraint qualification is needed and that the assumption on $\bar\alpha_i$ and $\bar\beta_i$ is reasonable because $\bar\alpha_i=\bar\beta_i=0$ means that the producer is willing to provide electricity for free.

Theorem 12 *Let $(\bar\alpha_i,\bar\beta_i)$ be a local solution to* (51) *and let (g,t) be the corresponding solution of its lower level. Assume that $\bar\alpha_i>0$ or that $\bar\beta_i g_i\ne 0$. Then there exist multipliers v^* and w^* such that*

$$\begin{aligned} &0\in -g_i+v_i^*+N_{[0,\infty)}(\bar\alpha_i), \\ &0\in -2g_i^2+2g_iv_i^*+N_{[0,\infty)}(\bar\beta_i), \\ &0\in \begin{pmatrix} e^i\cdot(\gamma-\bar\alpha)+2e^i\cdot(\delta-2\bar\beta)\cdot g+2\beta\cdot v^*s \\ 0\end{pmatrix}+D^*N_\Gamma(g,t,-F(\bar\alpha_i,\bar\beta_i,g,t))(v^*,w^*), \end{aligned}$$

where e^i is vector of zeros with one on position i and $\beta\cdot v$ denotes the Hadamard (componentwise) product of two vectors.

Proof We apply Theorem 11 to the MPEC with structure (46), where

$$\begin{aligned} &x_1=\bar\alpha_i,\ x_2=\bar\beta_i,\ y_1=g_i,\ y_2=(g_{-i},t),\ B=1,\ \omega=\mathbb{R}^2_+, \\ &\varphi(x,y)=(\gamma_i-\alpha_i)g_i+(\delta_i-2\beta_i)g_i^2,\ f_1(x_2,y_1)=\beta_i g_i^2,\ f_2(y_2)=\sum_{j\ne i}(\alpha_j g_j+\beta_j g_j^2). \end{aligned}$$

Here g_{-i} denotes vector g without component i and φ was multiplied by -1 to switch from a maximization to a minimization problem. Condition (52) reads

$$\begin{pmatrix} a \\ 2ag_i\end{pmatrix}\in N_\omega(\bar\alpha_i,\bar\beta_i)\implies a=0,$$

which is satisfied due to the imposed assumptions. Theorem 11 then implies the result. □

Acknowledgements The authors would like to thank two anonymous referees for their critical comments which led to a substantial improvement of the paper. We are particularly indebted to Prof. Helmut Gfrerer for pointing us to an incorrect result in the first version of the paper.

References

1. Aussel, D., Correa, R., Marechal, M.: Electricity spot market with transmission losses. Journal of Industrial and Management Optimization **9**, 275–290 (2013)
2. Dempe, S., Dutta, J.: Is bilevel programming a special case of a mathematical program with complementarity constraints? Math. Program. A **131**(1-2), 37–48 (2012)
3. Dontchev, A., Rockafellar, R.T.: Implicit Functions and Solution Mappings. Springer (2009)
4. Gfrerer, H.: First order and second order characterizations of metric subregularity and calmness of constraint set mappings. SIAM J. Optim. **21**(4), 1439–1474 (2011)
5. Gfrerer, H., Outrata, J.: On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications. Optimization (accepted)
6. Henrion, R., Jourani, A., Outrata, J.: On the calmness of a class of multifunctions. SIAM J. Optim. **13**(2), 603–618 (2002)
7. Henrion, R., Outrata, J., Surowiec, T.: On the co-derivative of normal cone mappings to inequality systems. Nonlinear Anal. Theory, Methods Appl. **71**(3-4), 1213–1226 (2009)
8. Henrion, R., Outrata, J., Surowiec, T.: Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market. ESAIM Control. Optim. Calc. Var. **18**(2), 295–317 (2012)
9. Henrion, R., Römisch, W.: On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. Appl. Math. **52**(6), 473–494 (2007)
10. Ioffe, A.D., Outrata, J.: On metric and calmness qualification conditions in subdifferential calculus. Set-Valued Anal. **16**(2-3), 199–227 (2008)
11. Janin, R.: Directional Derivative of the Marginal Function in Nonlinear Programming. Math. Programming Study **21**, 110–126 (1984)
12. Klatte, D., Kummer, B.: Constrained minima and Lipschitzian penalties in metric spaces. SIAM J. Optim. **13**(2), 619–633 (2002)
13. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I. Springer (2006)
14. Mordukhovich, B.S., Outrata, J.: Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. SIAM J. Optim. **18**(2), 389–412 (2007)
15. Outrata, J.: A generalized mathematical program with equilibrium constraints. SIAM J. Control Optim. **38**(5), 1623–1638 (2000)
16. Outrata, J.V.: Optimality conditions for a class of mathematical programs with equilibrium constraints. Mathematics of Operations Research **24**(3), 627–644 (1999)
17. Robinson, S.M.: Some continuity properties of polyhedral multifunctions. In: H. König, B. Korte, K. Ritter (eds.) Mathematical Programming at Oberwolfach, *Mathematical Programming Studies*, vol. 14, pp. 206–214. Springer Berlin Heidelberg (1981)
18. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer (1998)
19. Scholtes, S., Stöhr, M.: How stringent is the linear independence assumption for mathematical programs with complementarity constraints? Mathematics of Operations Research **26**(4), 851–863 (2001)
20. Surowiec, T.: Explicit stationarity conditions and solution characterization for equilibrium problems with equilibrium constraints. PhD thesis, Humboldt University Berlin (2010)
21. Ye, J.J.: Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. SIAM J. Optim. **10**(4), 943–962 (2000)
22. Ye, J.J., Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints. Math. Oper. Res. **22**(4), 977–997 (1997)

A A strong counterexample to the reversion of Proposition 2 under MFCQ and $\mathscr{C}^2$ data for Γ

In Example 2 we have shown that under MFCQ and smooth inequalities describing the set Γ, the mapping M may be calm, whereas the enhanced mapping $\tilde{M}$ fails to be calm for some multiplier. In the following stronger counterexample we construct a set Γ described by $\mathscr{C}^2$ inequalities satisfying MFCQ at given $\bar{y}$ and a function F such that M is calm at $(0,\bar{x},\bar{y})$ while $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\lambda)$ for **any** $\lambda \in \Lambda(\bar{x},\bar{y})$.

Define first $\varphi_1, \varphi_2 : [-1,1] \to \mathbb{R}$ and $q_1, q_2 : [-1,1] \times \mathbb{R} \to \mathbb{R}$ as

$$\varphi_1(t) := \begin{cases} (-1)^k \left(t - \frac{1}{k}\right)^3 \left(t - \frac{1}{k+1}\right)^3 & \text{for } t \in \left[\frac{1}{k+1}, \frac{1}{k}\right],\ k \in \mathbb{N} \\ 0 & \text{for } t \le 0, \end{cases}$$

$$\varphi_2(t) := \begin{cases} (-1)^k \left(t - \frac{1}{k}\right)^5 \left(t - \frac{1}{k+1}\right)^5 & \text{for } t \in \left[\frac{1}{k+1}, \frac{1}{k}\right],\ k \in \mathbb{N} \\ 0 & \text{for } t \le 0, \end{cases}$$

$$q_1(y) := \varphi_1(y_1) - y_2,$$
$$q_2(y) := \varphi_2(y_1) - y_2,$$

put $\omega = \mathbb{R}$ and as the reference point take $(\bar{x}, \bar{y}_1, \bar{y}_2) = (0,0,0)$. These functions are depicted in Figure 1. Note first that MFCQ is indeed satisfied for Γ and that φ_1 and φ_2 are twice continuously differentiable.

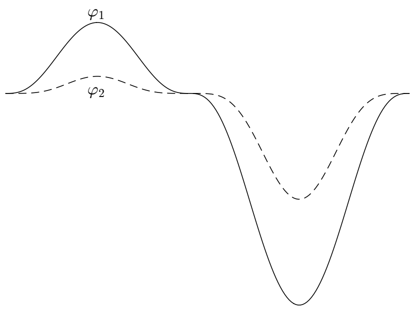

Fig. 1 Segments of graphs φ_1 and $2.3 \cdot 10^9 \varphi_2$. The constant in front of φ_2 is used for graphical purposes.

Define further

$$\phi(t) := \max\{\varphi_1(t), \varphi_2(t)\}.$$

Because $\phi'(\frac{1}{k}) = \phi''(\frac{1}{k}) = 0$ for all $k \in \mathbb{N}$, the twice continuous differentiability of ϕ is obvious apart from 0. At 0 we compute

$$\lim_{t \to 0} t^{-1} |\phi(t) - \phi(0)| = \lim_{t \to 0} t^{-1} |\varphi_1(t)| = 0,$$

which implies that $|\phi'(0)| = 0$. Similarly we obtain $\phi''(0) = 0$ and that ϕ is twice continuously differentiable. Finally, we define $F(x,y) := (-\phi'(y_1),\ 1)$. By construction of ϕ, we obtain that F is continuously differentiable. Since $\Gamma = \operatorname{epi} \phi$ we have that

$$M(0) = \left\{ (x,y) \,\middle|\, \begin{pmatrix} \phi'(y_1) \\ -1 \end{pmatrix} \in N_\Gamma(y) \right\} = \mathbb{R} \times \operatorname{gph} \phi.$$

As $M(p) \subset M(0)$ for all p small enough, we obtain that M is calm at $(0, \bar{x}, \bar{y})$.

It is easy to see that $\Lambda(\bar{x}, \bar{y}) = \{\lambda \ge 0 | \lambda_1 + \lambda_2 = 1\}$. We will show now that $\tilde{M}$ is not calm at $(0, 0, \bar{x}, \bar{y}, \lambda)$ for any $\lambda \in \Lambda(\bar{x}, \bar{y})$. Define

$$\Omega_1 := \{t \in [0,1] | \varphi_1(t) = \varphi_2(t)\},$$
$$\Omega_2 := \{t \in [0,1] | \varphi_1(t) \ne \varphi_2(t),\ \varphi_1'(t) = \varphi_2'(t)\},$$
$$\Omega_3 := [0,1] \setminus (\Omega_1 \cup \Omega_2)$$

and note that for all $t \in \Omega_2 \cup \Omega_3$ small enough it holds that $|\varphi_2(t)| < |\varphi_1(t)|$ and for all $t \in \Omega_3$ small enough we have $|\varphi_2'(t)| < |\varphi_1'(t)|$.

We will show first that $\hat{T}_{\{1\}}$ defined in (19) is not calm at $(0, \bar{y})$. From the definition we see that

$$\hat{T}_{\{1\}}(p) = \{y | \varphi_1(y_1) = y_2 + p_1, \varphi_2(y_1) \le y_2 + p_2\}.$$

and thus

$$\hat{T}_{\{1\}}(0)=\{y|\,\varphi_1(y_1)=y_2,\varphi_2(y_1)\le y_2\}=\{(y_1,\varphi_1(y_1))|\,\varphi_1(y_1)\ge 0\}.$$

Now pick any sequence $y_{k1}>0$, $y_{k1}\to 0$ such that $y_{k1}\in\Omega_2$ and $\varphi_1(y_{k1})<0$ and define $p_{k1}:=0$, $y_{k2}:=\varphi_1(y_{k1})$ and $p_{k2}:=\varphi_2(y_{k1})-y_{k2}$. Then $y_k\in\hat{T}_{\{1\}}(p_k)$. Moreover, as φ_1 and φ_2 have the same signs

$$0<\|p_k\|=p_{k2}=\varphi_2(y_{k1})-y_{k2}=\varphi_2(y_{k1})-\varphi_1(y_{k1})\le|\varphi_1(y_{k1})|.$$

Consider now a point $\tilde{y}_{k1}\in\Omega_1$ at which $d(y_{k1},\Omega_1)$ is realized. Since $\Omega_1\subset\hat{T}_{\{1\}}(0)$ and φ_1 is zero on Ω_1, we obtain

$$\frac{|d(y_k,\hat{T}_{\{1\}}(0))|}{|p_k|}\ge\frac{|d(y_{k1},\Omega_1)|}{|\varphi_1(y_{k1})|}=\frac{|y_{k1}-\tilde{y}_{k1}|}{|\varphi_1(y_{k1})-\varphi_1(\tilde{y}_{k1})|}=\frac{1}{\varphi_1'(\xi_k)},$$

where in the last equality we have used the mean value theorem to find some ξ_k which lies in the line segment connecting y_{k1} and $\tilde{y}_{k1}$. Since φ_1 is twice continuously differentiable with $\varphi_1'(0)=0$, we have proved that $\hat{T}_{\{1\}}$ is not calm at $(0,\bar{y})$. For $\hat{T}_{\{2\}}$ we proceed with a similar construction. In this case we have

$$\hat{T}_{\{2\}}(0)=\{y|\,\varphi_1(y_1)\le y_2,\varphi_2(y_1)=y_2\}=\{(y_1,\varphi_2(y_1))|\,\varphi_1(y_1)\le 0\}$$

and for the contradicting sequence we choose some $y_{k1}>0$, $y_{k1}\to 0$ such that $y_{k1}\in\Omega_2$ and $\varphi_1(y_{k1})>0$ and define again $p_{k1}:=0$, $y_{k2}:=\varphi_1(y_{k1})$ and $p_{k2}:=\varphi_2(y_{k1})-y_{k2}$ and perform the estimates as in the previous case. Since for $\hat{T}_{\{1,2\}}$ we have

$$\hat{T}_{\{1,2\}}(0)=\{y|\,\varphi_1(y_1)=y_2,\varphi_2(y_1)=y_2\}=\{(y_1,\varphi_1(y_1))|\,\varphi_1(y_1)=0\},$$

either of the previous contradicting sequences can be chosen.

Fix now any $\bar{\lambda}\in\Lambda(\bar{x},\bar{y})$ and consider the corresponding index set $I=\{i|\,\bar{\lambda}_i>0\}$. In the previous several paragraphs we have shown that $\hat{T}_I$ is not calm at $(0,\bar{y})$ and found a sequence $(\tilde{p}_k,\tilde{y}_k)$ violating the calmness property. By virtue of Lemma 4 we obtain that T is not calm at $(0,\bar{y},\bar{\lambda})$. Moreover, from the proof of this lemma we see that the sequence (p_k,y_k,λ_k), which violates the calmness of T at $(0,\bar{y},\bar{\lambda})$, can be taken in such a way that $p_k=\tilde{p}_k$, $y_k=\tilde{y}_k$ and $\lambda_k=\bar{\lambda}$ with $(\tilde{y}_k,\bar{\lambda})\in T(\tilde{p}_k)$ and

$$d((\tilde{y}_k,\bar{\lambda}),T(0))>(k-1)\|\tilde{p}_k\|. \tag{54}$$

Furthermore, in all the previous cases we have chosen $\tilde{y}_k$ in such a way that $\tilde{y}_{k1}\in\Omega_2$.

We will show that $\tilde{M}$ is not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Consider sequence

$$(0,0,\tilde{p}_{k1},\tilde{p}_{k2},\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2)\to(0,0,0,0,\bar{x},0,0,\bar{\lambda}_1,\bar{\lambda}_2) \tag{55}$$

and show first that $(\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2)\in\tilde{M}(0,0,\tilde{p}_{k1},\tilde{p}_{k2})$, which amounts to showing

$$\begin{aligned}\begin{pmatrix}0\\0\end{pmatrix}&=\begin{pmatrix}-\phi'(\tilde{y}_{k1})\\1\end{pmatrix}+\begin{pmatrix}\varphi_1'(\tilde{y}_{k1}) & \varphi_2'(\tilde{y}_{k1})\\-1 & -1\end{pmatrix}\begin{pmatrix}\bar{\lambda}_1\\\bar{\lambda}_2\end{pmatrix},\\ q(\tilde{y}_k)-\tilde{p}_k&\in N_{\mathbb{R}^2_+}(\bar{\lambda}).\end{aligned}$$

We know that $(\tilde{y}_k,\bar{\lambda})\in T(\tilde{p}_k)$ and hence the inclusion is satisfied. Moreover, as $\tilde{y}_{k1}\in\Omega_2$ by construction of this sequence and as $\bar{\lambda}_1+\bar{\lambda}_2=1$, we indeed obtain

$$(\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2)\in\tilde{M}(0,0,\tilde{p}_{k1},\tilde{p}_{k2}). \tag{56}$$

From the respective definitions of $\tilde{M}$ and T, we infer that $\tilde{M}(0,0,0,0)\subset\mathbb{R}^n\times T(0,0)$ and consequently due to (54) we obtain

$$d((\bar{x},\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2),\tilde{M}(0,0,0,0))\ge d((\tilde{y}_{k1},\tilde{y}_{k2},\bar{\lambda}_1,\bar{\lambda}_2),T(0,0))>(k-1)\|\tilde{p}_k\|.$$

This together with (55) and (56) implies that $\tilde{M}$ is indeed not calm at $(0,0,\bar{x},\bar{y},\bar{\lambda})$. Since $\bar{\lambda}$ was chosen arbitrarily from $\Lambda(\bar{x},\bar{y})$, the construction has been completed.