Math. Program., Ser. B (2012) 136:111–131
DOI 10.1007/s10107-012-0553-8

FULL LENGTH PAPER

On regular coderivatives in parametric equilibria with non-unique multipliers

R. Henrion · J. V. Outrata · T. Surowiec

Received: 12 January 2011 / Accepted: 4 October 2011 / Published online: 26 May 2012

Abstract This paper deals with the computation of regular coderivatives of solution maps associated with a frequently arising class of generalized equations (GEs). The constraint sets are given by (not necessarily convex) inequalities, and we do not assume linear independence of gradients to active constraints. The achieved results enable us to state several versions of sharp necessary optimality conditions in optimization problems with equilibria governed by such GEs. The advantages are illustrated by means of examples.

This work was supported by the DFG Research Center Matheon "Mathematics for key technologies" in Berlin, by the Grant Agency of the Czech Academy of Sciences (Grant No. IAA 100750802) and by the Australian Research Council (Project DP 110102011).

R. Henrion
Weierstrass Institute for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117 Berlin, Germany
e-mail: henrion@wias-berlin.de

J. V. Outrata (✉)
Institute of Information Theory and Automation, Pod vodárenskou vezí 4,
18208 Praha 8, Czech Republic
e-mail: outrata@utia.cas.cz

J. V. Outrata
Centre for Informatics and Applied Optimization, University of Ballarat,
P.O. Box 663, Ballarat, VIC 3353, Australia

T. Surowiec
Humboldt University Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: surowiec@math.hu-berlin.de

Keywords Parameterized generalized equation · Regular and limiting coderivative · Constant rank CQ · Mathematical program with equilibrium constraint

Mathematics Subject Classification 49J53 · 90C31 · 90C46

1 Introduction

The main object of this study is local analysis of solutions to the parameter-dependent generalized equation (GE)

$$0 \in F(x, y) + \widehat{N}_{\Gamma}(y), \tag{1}$$

where $x \in \mathbb{R}^n$ is the *parameter*, $y \in \mathbb{R}^m$ is the *decision variable*, $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is a continuously differentiable mapping and $\widehat{N}_{\Gamma}(y)$ stands for the *regular normal cone* to a closed, not necessarily convex set $\Gamma \subseteq \mathbb{R}^m$ at y, cf. Definition 2.1. If Γ is convex, then (1) amounts to a standard variational inequality of the first kind. In the form of GE (1), one can model a large variety of parameter-dependent equilibrium problems ranging from parametric optimization over Nash equilibria, second-order cone programs up to discretized contact problems from continuum mechanics.

We denote throughout the text by S the *solution map* to (1) defined by

$$S(x) := \{y | 0 \in F(x, y) + \widehat{N}_{\Gamma}(y)\}, \tag{2}$$

which assigns the corresponding set of equilibria to each value of the parameter x. Our main aim in this paper is to analyze the local behavior of S around a certain *reference point* $(\bar{x}, \bar{y})$ from the graph of S. Such analysis is useful both in the so-called post-optimal analysis, where one examines the dependence of an equilibrium on (a part of) problem data, as well as in mathematical programs with equilibrium constraints (MPECs), where (1) arises as a part of the constraint system. It may be useful also in numerical analysis, where the parameter could express, e.g., the accuracy with which the Karush–Kuhn–Tucker (KKT) condition is used as the stopping criterion in a numerical method.

Local analysis of solution maps belongs to the core of problems of modern variational analysis and has been studied in connection with similar models in a large amount of papers and many monographs since 1968 [4]. From the recently published advanced monographs, we highlight the monographs [3,10,19]. Further results can be found in [9,15] and [1], where the motivation comes from the mentioned MPECs. In this paper, our main tool is the notion of *regular coderivative* (Definition 2.3). This is the main difference with respect to the papers [8,11] in which similar models are analyzed with the help of *limiting coderivatives* (Definition 2.4). We place a special amount of emphasis on the case in which Γ is given via inequalities and the formula for $N_{\Gamma}(\cdot)$ involves non-unique multipliers.

The paper is structured as follows. The preliminary Sect. 2 contains definitions and some basic auxiliary results used throughout the following sections. Section 3 contains our main results. Under the Mangasarian-Fromovitz constraint qualification

(MFCQ), constant rank qualification condition (CRCQ) and the *Strong Second-Order Sufficient Condition*, we compute the desired regular coderivative of S either exactly or we provide a tight upper estimate of it in dependence on the properties of F and Γ. These results enable us, among other things, to obtain new sharp optimality conditions for a class of MPECs.

In Sect. 4, the Strong Second-Order Sufficient Condition is omitted, but one has to require the surjectivity of $\nabla_x F(\bar{x}, \bar{y})$. The resulting formula could again be applied to a class of MPECs where, however, no other conditions on x and y can be imposed. Both sections depend heavily on a consequence of [16, Theorem 2] which is our main reference and will be stated in the next section.

Our notation is basically standard. Gr Φ is the graph of a multifunction Φ, $f'(x; h)$ denotes the directional derivative of a single-valued mapping f at x in direction h and K^0 stands for the negative polar to a cone K. For a finite set I, $\#I$ denotes its cardinality.

2 Definitions and preliminaries

In order to proceed with the statement of some necessary preliminary results we will need to define a few standard notions from variational analysis. Let $C \subseteq \mathbb{R}^m$ be a nonempty closed subset and $x \in C$. Then we define the *contingent* or *Bouligand cone* by

$$T_C(x) := \left\{d \in \mathbb{R}^m \,|\exists t_k \downarrow 0, d_k \to d : x + t_k d_k \in C, \forall k\right\}.$$

Analogously to similar constructions in classical convex analysis, we use the contingent cone $T_C(x)$ to define the so-called *regular* or *Fréchet normal cone*.

Definition 2.1 For $C \subseteq \mathbb{R}^m$ be a nonempty closed subset and $x \in C$, the *regular* or *Fréchet normal cone* is defined by

$$\widehat{N}_C(x) := T_C(x)^0,$$

i.e., the regular normal cone to C at x is defined as the polar cone of the contingent cone $T_C(x)$.

Note that one can also define the regular normal cone by

$$\widehat{N}_C(x) := \left\{ v \in \mathbb{R}^m \,\middle|\, \limsup_{\substack{x' \to x \\ x' \in C}} \frac{\langle v, x' - x\rangle}{||x' - x||} \leq 0 \right\}.$$

The regular normal cone is then used to introduce another important normal cone.

Definition 2.2 Let $C \subseteq \mathbb{R}^m$ be a nonempty closed subset and $\bar{x} \in C$, then the so-called *limiting* or *Mordukhovich normal cone* is defined by

$$N_C(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \to \bar{x} \\ x \in C}} \widehat{N}_C(x).$$

where "Lim sup" represents the Painlevé-Kuratowski upper-limit, cf. [10, formula (1.1)].

We mention that both these cones coincide with the normal cone from convex analysis when C is convex.

The two notions of normal cones are often used to define two types of generalized (co)derivatives of special interest in this paper.

Definition 2.3 Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be an arbitrary multifunction and suppose $(x, y) \in \mathrm{Gr}\,\Phi$. Then we define the *regular coderivative* of Φ at (x, y) in direction y^* by

$$\widehat{D}^*\Phi(x, y)(y^*) := \left\{x^* \in \mathbb{R}^n \,\middle|\, (x^*, -y^*) \in \widehat{N}_{\mathrm{Gr}\,\Phi}(x, y)\right\}.$$

Similar to the above, we have a limiting concept for coderivatives as well.

Definition 2.4 Let $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be an arbitrary multifunction and suppose $(x, y) \in \mathrm{Gr}\,\Phi$. The *limiting coderivative* of Φ can then be defined by

$$D^*\Phi(\bar{x}, \bar{y})(\bar{y}^*) := \mathop{\mathrm{Lim\,sup}}_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in \mathrm{Gr}\,\Phi \\ y^*\to\bar{y}^*}} \widehat{D}^*\Phi(x, y)(y^*)$$

or equivalently by using the limiting normal cone to $\mathrm{Gr}\,\Phi$:

$$D^*\Phi(\bar{x}, \bar{y})(\bar{y}^*) := \left\{\bar{x}^* \in \mathbb{R}^n \,\middle|\, (\bar{x}^*, -\bar{y}^*) \in N_{\mathrm{Gr}\,\Phi}(\bar{x}, \bar{y})\right\}.$$

If Φ is single-valued, we write simply $\widehat{D}^*\Phi(x)(y^*)$, $D^*\Phi(x)(y^*)$. We refer the reader to the monographs [19] and [10] for more on these concepts.

The following theorem recalls a basic transformation formula for coderivatives [10, Theorem 1.127]:

Theorem 2.1 *Let $C = g^{-1}(\Xi)$, where $g : \mathbb{R}^k \to \mathbb{R}^l$ is twice continuously differentiable and $\Xi \subseteq \mathbb{R}^l$ is some closed subset. Consider points $\bar{z} \in C$ and $\bar{v} \in N_C(\bar{z})$. If the Jacobian $\nabla g(\bar{z})$ is surjective, then*

$$D^*N_C(\bar{z}, \bar{v})(v^*) = \left(\sum_{i=1}^{l} \bar{\mu}_i \nabla^2 g_i(\bar{z})\right) v^* + (\nabla g\,(\bar{z}))^T\, D^*N_\Xi\,(g(\bar{z}), \bar{\mu})\left(\nabla g\,(\bar{z})\, v^*\right). \quad (3)$$

Here, $\bar{\mu}$ is the unique solution of the equation $\nabla^T g(\bar{z})\mu = \bar{v}$.

In the special case, when $\Xi := \{0\}_t \times \mathbb{R}^{l-t}_-$ $(t \le l)$ serves to model a finite set of smooth equality and inequality constraints, (3) reduces to

$$D^*N_C(\bar{z}, \bar{v})(v^*) = \left(\sum_{i=1}^{l} \bar{\mu}_i \nabla^2 g_i(\bar{z})\right) v^* + (\nabla g\,(\bar{z}))^T\,(A_1 \times \cdots \times A_l), \quad (4)$$

where the sets $A_i \subseteq \mathbb{R}$ are given by

$$A_i = D^* N_{\{0\}}(0, \bar{\mu}_i)\left(\nabla g_i(\bar{z})\, v^*\right) = \begin{cases} \mathbb{R} & \text{if } \nabla g_i(\bar{z})\, v^* = 0 \\ \emptyset & \text{else} \end{cases} \quad (i = 1, \dots t),$$

$$\begin{aligned} A_i &= D^* N_{\mathbb{R}_-}(g_i(\bar{z}), \bar{\mu}_i)\left(\nabla g_i(\bar{z})\, v^*\right) \\ &= \begin{cases} \{0\} & \text{if } g_i(\bar{z}) < 0 \\ \mathbb{R} & \text{if } g_i(\bar{z}) = 0,\ \bar{\mu}_i > 0 \text{ and } \nabla g_i(\bar{z})\, v^* = 0 \\ \emptyset & \text{if } g_i(\bar{z}) = 0,\ \bar{\mu}_i > 0 \text{ and } \nabla g_i(\bar{z})\, v^* \neq 0 \\ \{0\} & \text{if } g_i(\bar{z}) = \bar{\mu}_i = 0 \text{ and } \nabla g_i(\bar{z})\, v^* < 0 \\ \mathbb{R}_+ & \text{if } g_i(\bar{z}) = \bar{\mu}_i = 0 \text{ and } \nabla g_i(\bar{z})\, v^* \geq 0. \end{cases} \quad (i = t + 1, \dots l) \end{aligned}$$

In this special structure for Ξ, we can replace the surjectivity condition required in Theorem 2.1 by the weaker Linear Independence Constraint Qualification (LICQ) at $\bar{z}$. In this condition, one requires that the gradients $\nabla g_i(\bar{z})$ are linearly independent for all i such that $g_i(\bar{z}) = 0$.

Throughout the whole paper, we will specify the set Γ in (1) as

$$\Gamma = \{y \in \mathbb{R}^m | q(y) \in \mathbb{R}^s_-\}, \tag{5}$$

where the mapping $q : \mathbb{R}^m \to \mathbb{R}^s$ is twice continuously differentiable. Let $\bar{y} \in \Gamma$. We say that Γ fulfills the MFCQ at $\bar{y}$, provided

$$\left.\begin{aligned} (\nabla q(\bar{y}))^T \lambda = 0 \\ \lambda \geq 0 \\ \langle q(\bar{y}), \lambda \rangle = 0 \end{aligned}\right\} \Rightarrow \lambda = 0.$$

It is well known (cf., e.g., [19, Exercise 10.26]) that under MFCQ there is a neighborhood $\mathcal{N}$ of $\bar{y}$ such that

$$N_\Gamma(y) = (\nabla q(y))^T N_{\mathbb{R}^s_-}(q(y))$$

for all $y \in \mathcal{N}$. Since $N_\Gamma(y) \supset \widehat{N}_\Gamma(y)$ by the definition,

$$\widehat{N}_\Gamma(y) \supset (\nabla q(y))^T \widehat{N}_{\mathbb{R}^s_-}(q(y))$$

by virtue of [19, Theorem 6.14] and $\widehat{N}_{\mathbb{R}^s_-}(q(y)) = N_{\mathbb{R}^s_-}(q(y))$, it follows that

$$\begin{aligned} \widehat{N}_\Gamma(y) &= (\nabla q(y))^T N_{\mathbb{R}^s_-}(q(y)) \\ &= \left\{(\nabla q(y))^T \lambda | q(y) \leq 0,\ \lambda \geq 0,\ \langle q(y), \lambda \rangle = 0\right\} \end{aligned}$$

for all $y \in \mathcal{N}$. This enables us to replace the GE (1) for our purposes by the GE

$$0 \in F(x, y) + (\nabla q(y))^T N_{\mathbb{R}^s_-}(q(y)), \tag{6}$$

or by the KKT system

$$0 = \mathcal{L}(x, y, \lambda),\ \ q(y) \leq 0,\ \ \lambda \geq 0,\ \ \langle q(y), \lambda \rangle = 0, \tag{7}$$

where

$$\mathcal{L}(x, y, \lambda) := F(x, y) + (\nabla q(y))^T \lambda$$

is the *Lagrangian*, associated with the GE (1).

The *critical cone* to Γ at y with respect to a vector $v \in \mathbb{R}^m$ is given by

$$\mathcal{C}(y, v) := T_\Gamma(y) \cap (v)^\perp. \tag{8}$$

Assume that the MFCQ is fulfilled at $\bar{y}$, y is sufficiently close to $\bar{y}$, and let (x, y, λ) be a feasible triple with respect to (7). In the usual way we associate with (y, λ) the index sets

$$\begin{aligned} I(y) &:= \{i \in \{1, 2, \ldots, s\} | q_i(y) = 0\}, \\ I_+(y, \lambda) &:= \{i \in I(y) | \lambda_i > 0\}, \\ I_0(y, \lambda) &:= I(y) \backslash I_+(y, \lambda) \end{aligned}$$

of *active, strongly active* and *weakly active* inequalities, respectively. Then it can easily be shown that the critical cone to Γ at y with respect to $F(x, y)$, denoted henceforth by $K(x, y)$, amounts to

$$\begin{aligned} K(x, y) &= \mathcal{C}(y, F(x, y)) = T_\Gamma(y) \cap \{F(x, y)\}^\perp \\ &= \{v \in \mathbb{R}^m | \nabla q(y) v \in T_{\mathbb{R}^s_-}(q(y)) \cap \{\lambda\}^\perp\} \\ &= \{v | \langle \nabla q_i(y), v \rangle = 0 \text{ for } i \in I_+(y, \lambda), \langle \nabla q_i(y), v \rangle \leq 0 \text{ for } i \in I_0(y, \lambda)\}. \end{aligned} \tag{9}$$

By the definition, this cone is independent of a concrete choice of the multiplier λ but it does depend on x.

In Theorem 2.2 below we will employ also another constraint qualification. We say that the constant rank qualification condition (CRCQ) holds at $\bar{y}$, provided there is a neighborhood $\mathcal{M}$ of $\bar{y}$ such that for any subsets I of $I(\bar{y})$, the family of gradients $\{\nabla q_i(y) | i \in I\}$ has the same rank for all $y \in \mathcal{M}$.

Note that (MFCQ) and (CRCQ) are independent in the sense that they do not imply each other.

Theorem 2.2 *Consider the GE* (1) *with* Γ *given by* (5) *around a reference point* $(\bar{x}, \bar{y}) \in \operatorname{Gr} S$. *Assume that*

(i) Γ *fulfills MFCQ and CRCQ at* $\bar{y}$;
(ii) *For each* λ *satisfying* (7) *with* $(x, y) = (\bar{x}, \bar{y})$, *and each* $v \neq 0$ *such that* $\langle \nabla q_i(\bar{y}), v \rangle = 0$ *if* $i \in I_+(\bar{y}, \lambda)$,

$$\langle v, \nabla_y \mathcal{L}(\bar{x}, \bar{y}, \lambda) v \rangle > 0.$$

(iii) *either all functions q_i are convex or $F(x, y)$ amounts to the partial gradient with respect to the second variable of a function $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.*

Then the following statements hold.

(1) *S admits a single-valued Lipschitz localization at $(\bar{x}, \bar{y})$, i.e., there are neighborhoods $\mathcal{U}$ of $\bar{x}$, $\mathcal{V}$ of $\bar{y}$ and a Lipschitz function $\sigma : \mathcal{U} \to \mathcal{V}$ such that*

$$\sigma(\bar{x}) = \bar{y} \ \text{ and } \ S(x) \cap \mathcal{V} = \{\sigma(x)\} \ \text{ for } \ x \in \mathcal{U}.$$

(2) *For each $x \in \mathcal{U}$, $d \in \mathbb{R}^n$, σ is directionally differentiable at x in the direction d, and $\sigma'(x; d) = v$, the unique solution of the GE*

$$0 \in \nabla_x F(x, y)d + \nabla_y \mathcal{L}(x, y, \mu)v + N_{K(x,y)}(v), \tag{10}$$

where $y = \sigma(x)$ and μ is an associated multiplier in the sense of (7).

Proof Consider first the case when the functions q_i are convex. Statement (1) then follows from [9, Theorem 4.2.16], because (ii) implies the so-called strong coherent orientation condition (used in that result). Statement (2) follows from [9, Theorem 4.2.25], because q does not depend on x and (ii) implies the uniqueness of solutions to (10) for all $d \in \mathbb{R}^n$. In the nonconvex case the statement goes back to [16, Theorem 2 and Corollary 4]. □

Remark 2.1 According to the assumption (iii) one can distinguish between two cases. If the functions q_i are convex, then (1) amounts to the standard parameter-dependent VI of the first kind. Otherwise, (1) is a stationarity condition of the parametric nonlinear program (NLP)

$$\begin{array}{l} \text{minimize } \ \varphi(x, y) \\ \text{subject to} \\ y \in \Gamma, \end{array} \tag{11}$$

and for $x \in \mathcal{U}$, $\sigma(x)$ is an isolated minimizer of (11).

Condition (ii) is the already mentioned strong second-order sufficient condition used extensively in the literature in connection with parametric problems.

3 When S is directionally differentiable

We start with the following simple statement:

Proposition 3.1 *Consider a mapping $H : \mathbb{R}^p \to \mathbb{R}^q$ which is Lipschitz and directionally differentiable on a neighborhood $\mathcal{U}$ of $\bar{u}$. Then,*

$$\hat{D}^* H(u)(y^*) = \hat{D}^* \mathcal{H}_u(0)(y^*) \quad \forall y^* \in \mathbb{R}^q \ \forall u \in \mathcal{U},$$

where $\mathcal{H}_u(\cdot) = H'(u; \cdot)$.

Proof It suffices to observe that, under the imposed assumptions, for $\tilde{u} \in \mathcal{U}$ and $\tilde{v} = H(\tilde{u})$ one has $T_{\mathrm{Gr}\,H}(\tilde{u}, \tilde{v}) = \mathrm{Gr}\,\mathcal{H}_{\tilde{u}}$. Consequently,

$$\begin{aligned}\hat{N}_{\mathrm{Gr}\,H}(\tilde{u}, \tilde{v}) &= (T_{\mathrm{Gr}\,H}(\tilde{u}, \tilde{v}))^0 \\ &= \left\{(u^*, v^*) \mid \langle u^*, d\rangle + \langle v^*, h\rangle \le 0 \quad \forall (d, h) \in T_{\mathrm{Gr}H}(\tilde{u}, \tilde{v})\right\} \\ &= \left\{(u^*, v^*) \mid \langle u^*, d\rangle + \langle v^*, \mathcal{H}_{\tilde{u}}(d)\rangle \le 0 \quad \forall d\right\}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\hat{N}_{\mathrm{Gr}\,\mathcal{H}_{\tilde{u}}}(0, 0) &= \left\{(u^*, v^*) \mid \langle u^*, d\rangle + \langle v^*, h\rangle \le o(\|(d, h)\|) \quad \forall (d, h) \in \mathrm{Gr}\,\mathcal{H}_{\tilde{u}}\right\} \\ &= \left\{(u^*, v^*) \mid \langle u^*, d\rangle + \langle v^*, \mathcal{H}_{\tilde{u}}(d)\rangle \le \tilde{o}(\|d\|) \quad \forall d\right\}\end{aligned}$$

by virtue of the Lipschitz continuity of $\mathcal{H}_{\tilde{u}}$. It follows, due to the positive homogeneity of $\mathcal{H}_{\tilde{u}}$, that

$$\begin{aligned}\widehat{N}_{\mathrm{Gr}\mathcal{H}_{\tilde{u}}}(0, 0) &= \left\{(u^*, v^*) \,\middle|\, \limsup_{d \to 0, d \neq 0} \left[\left\langle u^*, \frac{d}{\|d\|}\right\rangle + \left\langle v^*, \mathcal{H}_{\tilde{u}}\left(\frac{d}{\|d\|}\right)\right\rangle\right] \le 0\right\} \\ &= \left\{(u^*, v^*) \,\middle|\, \limsup_{\|z\|=1} \left[\langle u^*, z\rangle + \langle v^*, \mathcal{H}_{\tilde{u}}(z)\rangle\right] \le 0\right\} \\ &= \left\{(u^*, v^*) \mid \langle u^*, z\rangle + \langle v^*, \mathcal{H}_{\tilde{u}}(z)\rangle \le 0 \quad \forall z\right\}.\end{aligned}$$

This proves the assertion. □

Corollary 3.1 *For all $\bar{y}^* \in \mathbb{R}^q$*

$$D^* H(\bar{u})(\bar{y}^*) = \operatorname*{Limsup}_{u \to \bar{u}, y^* \to \bar{y}^*} \widehat{D}^* \mathcal{H}_u(0)(y^*).$$

Thus in some situations, the statement above enables us to replace a difficult single-valued mapping by a simpler one when computing coderivatives. In this section we exploit this idea for the solution mapping S on the basis of Theorem 2.2.

Theorem 3.1 *Under the assumptions of Theorem 2.2 let $\bar{\lambda}$ be an arbitrary multiplier satisfying conditions (7) with $(x, y) = (\bar{x}, \bar{y})$. Then for all $y^* \in \mathbb{R}^m$ one has that*

$$\begin{aligned}&\hat{D}^* S(\bar{x}, \bar{y})(y^*) \\ &\quad \supseteq \left\{(\nabla_x F(\bar{x}, \bar{y}))^T b \mid 0 \in y^* + \left(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\right)^T b + K^0(\bar{x}, \bar{y}), \ -b \in K(\bar{x}, \bar{y})\right\},\end{aligned} \tag{12}$$

where

$$\begin{aligned}&K(\bar{x}, \bar{y}) \\ &\quad = \left\{v \in \mathbb{R}^m \mid \langle \nabla q_i(\bar{y}), v\rangle = 0 \, \textit{for } i \in I_+(\bar{y}, \bar{\lambda}), \ \langle \nabla q_i(\bar{y}), v\rangle \le 0 \, \textit{for } i \in I_0(\bar{y}, \bar{\lambda})\right\}.\end{aligned}$$

If in addition the inclusion

$$\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}) K(\bar{x}, \bar{y}) \subseteq \operatorname{Im} \nabla_x F(\bar{x}, \bar{y}) \tag{13}$$

is satisfied, then (12) *holds as an equality.*

Proof Referring back to the single-valued Lipschitz localization σ of S in statement (1) of Theorem 2.2, we observe that $\hat{N}_{\mathrm{Gr}\, S} = \hat{N}_{\mathrm{Gr}\, \sigma}$ locally around $(\bar{x}, \bar{y})$. By Proposition 3.1, $\hat{N}_{\mathrm{Gr}\, \sigma}(\bar{x}, \bar{y}) = \hat{N}_{\mathrm{Gr}\, \sigma'(\bar{x};\cdot)}(0, 0)$ and so it remains to compute $\hat{N}_{\mathrm{Gr}\, \sigma'(\bar{x};\cdot)}(0, 0)$. According to (10), we have that

$$\mathrm{Gr}\, \sigma'(\bar{x}; \cdot) = G^{-1} \mathrm{Gr}\, N_{K(\bar{x}, \bar{y})}, \tag{14}$$

where G is the linear mapping defined by

$$G(d, v) := \begin{bmatrix} 0 & I \\ -\nabla_x F(\bar{x}, \bar{y}) & -\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}) \end{bmatrix} \begin{pmatrix} d \\ v \end{pmatrix}. \tag{15}$$

Moreover, since $K(\bar{x}, \bar{y})$ is a polyhedral cone, it holds that (see [2, Proof of Th. 2])

$$\hat{N}_{\mathrm{Gr}\, N_{K(\bar{x}, \bar{y})}}(0, 0) = K^0(\bar{x}, \bar{y}) \times K(\bar{x}, \bar{y}).$$

We now invoke [19, Theorem 6.14], according to which

$$\begin{aligned} \hat{N}_{\mathrm{Gr}\, \sigma'(\bar{x};\cdot)}(0, 0) &\supseteq [\nabla G\,(0, 0)]^T\, \hat{N}_{\mathrm{Gr}\, N_{K(\bar{x}, \bar{y})}}(0, 0) \\ &= \begin{bmatrix} 0 & -(\nabla_x F(\bar{x}, \bar{y}))^T \\ I & -\left(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\right)^T \end{bmatrix} \left(K^0(\bar{x}, \bar{y}) \times K(\bar{x}, \bar{y})\right) \\ &= \{(-(\nabla_x F(\bar{x}, \bar{y}))^T\, b, s) | b \in K(\bar{x}, \bar{y}), \\ &\quad s \in -\left(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\right)^T b + K^0(\bar{x}, \bar{y})\}. \end{aligned}$$

This shows inclusion (12).

To verify the reverse inclusion under the additional condition (13), let $x^* \in \hat{D}^* S(\bar{x}, \bar{y})(y^*)$ be arbitrary. By definition,

$$\langle x^*, d\rangle - \langle y^*, v\rangle \le 0 \quad \forall\, (d, v) \in T_{\mathrm{Gr}\, S}(\bar{x}, \bar{y}). \tag{16}$$

Since $T_{\mathrm{Gr}\, S} = T_{\mathrm{Gr}\, \sigma}$ locally around $(\bar{x}, \bar{y})$ and $T_{\mathrm{Gr}\, \sigma}(\bar{x}, \bar{y}) = \mathrm{Gr}\, \sigma'(\bar{x}; \cdot)$, as already stated in the beginning of the proof of Proposition 3.1, (14) provides that $T_{\mathrm{Gr}\, S}(\bar{x}, \bar{y}) = G^{-1} \mathrm{Gr}\, N_{K(\bar{x}, \bar{y})}$. Recalling that

$$\mathrm{Gr}\, N_{K(\bar{x}, \bar{y})} = \{(r, s) \in K(\bar{x}, \bar{y}) \times K^0(\bar{x}, \bar{y}) |\, \langle r, s\rangle = 0\}$$

because $K(\bar{x},\bar{y})$ is a convex cone, we derive that

$$T_{\mathrm{Gr}\,S}(\bar{x},\bar{y}) = \left\{ (d,v) \left| \begin{array}{l} v \in K(\bar{x},\bar{y}) \\ -\nabla_x F(\bar{x},\bar{y})d - \nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})v \in K^0(\bar{x},\bar{y}) \\ \left\langle v, \nabla_x F(\bar{x},\bar{y})d + \nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})v \right\rangle = 0 \end{array} \right. \right\}. \tag{17}$$

Clearly, $0 \in K(\bar{x},\bar{y})$, hence (16) and (17) entail that

$$\left\langle x^*, d \right\rangle \le 0 \quad \forall d : -\nabla_x F(\bar{x},\bar{y})d \in K^0(\bar{x},\bar{y}).$$

As a consequence,

$$x^* \in \left[-(\nabla_x F(\bar{x},\bar{y}))^{-1} K^0(\bar{x},\bar{y}) \right]^0 = -\left(\nabla_x F(\bar{x},\bar{y})\right)^T K(\bar{x},\bar{y}).$$

Therefore,

$$x^* = -\left(\nabla_x F(\bar{x},\bar{y})\right)^T \bar{u} \tag{18}$$

for some $\bar{u} \in K(\bar{x},\bar{y})$. This allows us to write (16) as

$$\left\langle -\left(\nabla_x F(\bar{x},\bar{y})\right)^T \bar{u}, d \right\rangle - \left\langle y^*, v \right\rangle \le 0 \quad \forall (d,v) \in T_{\mathrm{Gr}\,S}(\bar{x},\bar{y}). \tag{19}$$

Now, fixing an arbitrary $v \in K(\bar{x},\bar{y})$, condition (13) yields the existence of some d_v such that $-\nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})v = \nabla_x F(\bar{x},\bar{y})d_v$. It follows from (17) that $(d_v, v) \in T_{\mathrm{Gr}\,S}(\bar{x},\bar{y})$, whence (19) yields that

$$\begin{aligned} 0 &\ge \left\langle -\left(\nabla_x F(\bar{x},\bar{y})\right)^T \bar{u}, d_v \right\rangle - \left\langle y^*, v \right\rangle = \langle -\bar{u}, \nabla_x F(\bar{x},\bar{y})d_v \rangle - \left\langle y^*, v \right\rangle \\ &= \left\langle \bar{u}, \nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})v \right\rangle - \left\langle y^*, v \right\rangle = \left\langle \left(\nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})\right)^T \bar{u} - y^*, v \right\rangle. \end{aligned}$$

As $v \in K(\bar{x},\bar{y})$ was arbitrary, we get that $\left(\nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})\right)^T \bar{u} - y^* \in K^0(\bar{x},\bar{y})$. Along with (18) this last relation yields that x^* belongs to the set on the right-hand side of inclusion (12). □

Corollary 3.2 *Consider the setting of Theorem 2.2 and let $\bar{\lambda}$ be an arbitrary multiplier satisfying conditions (7) with $(x,y) = (\bar{x},\bar{y})$. If $\nabla_x F(\bar{x},\bar{y})$ is surjective, then one has for all $y^* \in \mathbb{R}^m$ that*

$$\begin{aligned} &\hat{D}^* S(\bar{x},\bar{y})(y^*) \\ &\quad = \left\{ \left(\nabla_x F(\bar{x},\bar{y})\right)^T b \middle| 0 \in y^* + \left(\nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})\right)^T b + K^0(\bar{x},\bar{y}),\ -b \in K(\bar{x},\bar{y}) \right\}. \end{aligned} \tag{20}$$

We emphasize that condition (13) allows us via Theorem 3.1 to relax the surjectivity of $\nabla_x F(\bar{x},\bar{y})$ as required in Corollary 3.2. This is illustrated in the following example:

Example 3.1 Consider the solution mapping S of the following parametric convex optimization problem

$$\min \left\{ y_1^2 + y_2^2 + xy_1 | y_2 \geq 1 \right\}.$$

Clearly, S can be equivalently described by the GE (1), where

$$F(x, y_1, y_2) = \begin{pmatrix} 2y_1 + x \\ 2y_2 \end{pmatrix}$$

and Γ is of the form (5) with $q(y_1, y_2) = 1 - y_2$. Let us consider the reference point

$$(\bar{x}, \bar{y}_1, \bar{y}_2) = (0, 0, 1) \in \operatorname{Gr} S.$$

Note that all assumptions of Theorem 2.2 are fulfilled. One easily checks that $K(\bar{x}, \bar{y}_1, \bar{y}_2) = \mathbb{R} \times \{0\}$. Moreover,

$$\nabla_y F(\bar{x}, \bar{y}_1, \bar{y}_2) = 2I, \quad \nabla_x F(\bar{x}, \bar{y}_1, \bar{y}_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

It follows that condition (13) is satisfied no matter what the choice of $\bar{\lambda}$ is:

$$\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}) K(\bar{x}, \bar{y}) = \nabla_y F(\bar{x}, \bar{y}_1, \bar{y}_2)\, K(\bar{x}, \bar{y}) = \mathbb{R} \times \{0\} = \operatorname{Im} \nabla_x F(\bar{x}, \bar{y}).$$

On the other hand, $\nabla_x F(\bar{x}, \bar{y}_1, \bar{y}_2)$ fails to be surjective.

Now, from the last statement of Theorem 3.1, it follows that

$$\hat{D}^* S(\bar{x}, \bar{y})(y^*) = \left\{ -\frac{y_1^*}{2} \right\}.$$

If condition (13) is violated, we have to confine ourselves to the following upper estimate of $\hat{D}^* S(\bar{x}, \bar{y})$:

Theorem 3.2 *Consider the setting of Theorem 2.2 with $\bar{\lambda}$ being an arbitrary multiplier satisfying (7) for $(x, y) = (\bar{x}, \bar{y})$. Then, for all $y^* \in \mathbb{R}^m$,*

$$\begin{aligned} &\hat{D}^* S(\bar{x}, \bar{y})(y^*) \\ &\quad \subseteq \left\{ (\nabla_x F(\bar{x}, \bar{y}))^T b | 0 \in y^* + \left(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\right)^T b + D^* N_{K(\bar{x}, \bar{y})}(0, 0)(b) \right\}. \end{aligned} \tag{21}$$

Proof Using the same arguments as in the beginning of the proof of Theorem 3.1 we have that

$$\hat{N}_{\operatorname{Gr} S}(\bar{x}, \bar{y}) = \hat{N}_{\operatorname{Gr} \sigma'(\bar{x}; \cdot)}(0, 0).$$

In order to estimate the right-hand side of the last equation, we apply a calculus rule from [6, Th. 4.1] yielding

$$\hat{N}_{\mathrm{Gr}\,\sigma'(\bar{x};\cdot)}(0,0) \subseteq \begin{bmatrix} 0 & -(\nabla_x F(\bar{x},\bar{y}))^T \\ I & -\left(\nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})\right)^T \end{bmatrix} N_{\mathrm{Gr}\,N_{K(\bar{x},\bar{y})}}(0,0)$$

and we are done. Note that the mentioned calculus rule requires the so-called calmness property (see, e.g., [19] p.399) of the multifunction

$$M(p) := \left\{(d,v)\in\mathbb{R}^n\times\mathbb{R}^m \,\middle|\, \begin{pmatrix} v \\ -\nabla_x F(\bar{x},\bar{y})d - \nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{\lambda})v \end{pmatrix} - p \in \mathrm{Gr}\,N_{K(\bar{x},\bar{y})}\right\}.$$

This property is satisfied because $N_{K(\bar{x},\bar{y})}$ is a polyhedral mapping (cf. [17]). □

Due to the polyhedrality of $K(\bar{x},\bar{y})$ an exact computation of $D^*N_{K(\bar{x},\bar{y})}(0,0)$ is possible even if the inequalities defining $K(\bar{x},\bar{y})$ are linearly dependent. We return to this question at the end of this section.

For some applications it is convenient to extend the preceding result to the mapping $\tilde{S}:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined by

$$\tilde{S}(x) := \begin{cases} S(x) & \text{if } x\in\omega \\ \emptyset & \text{otherwise,} \end{cases} \tag{22}$$

where $\omega\subseteq\mathbb{R}^n$ is a nonempty closed set. Clearly, Gr $\tilde{S}$ =Gr $S\cap(\omega\times\mathbb{R}^m)$.

Theorem 3.3 *Consider the setting of Theorem 3.2 and the mapping $\tilde{S}$ given by (22). Then, for all $y^*\in\mathbb{R}^m$,*

$$\hat{D}^*\tilde{S}(\bar{x},\bar{y})(y^*) \subseteq \left\{(\nabla_x F(\bar{x},\bar{y}))^T b + N_\omega(\bar{x}) \,|\, 0\in y^* + \left(\nabla_y\mathcal{L}(\bar{x},\bar{y},\bar{\lambda})\right)^T b + D^*N_{K(\bar{x},\bar{y})}(0,0)(b)\right\}. \tag{23}$$

Proof Let y^* be arbitrary and $x^*\in\hat{D}^*\tilde{S}(\bar{x},\bar{y})(-y^*)$. By virtue of [19, Th. 6.11], there is a smooth function $\varphi:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ such that φ attains its global minimum over Gr $\tilde{S}$ at $(\bar{x},\bar{y})$ and $\nabla\varphi(\bar{x},\bar{y}) = -(x^*,y^*)$. By virtue of the standard optimality condition one has thus

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} \in \widehat{N}_{\mathrm{Gr}\,\tilde{S}}(\bar{x},\bar{y}).$$

This is equivalent to stating that

$$\langle -x^*, d\rangle + \langle -y^*, v\rangle \geq 0 \text{ for all } (d,v)\in T_{\mathrm{Gr}\,\tilde{S}}(\bar{x},\bar{y}),$$

i.e., (0, 0) is a solution to the optimization problem

$$\min \left\{ \langle -x^*, d\rangle + \langle -y^*, v\rangle \mid (d, v) \in T_{\mathrm{Gr}\,\tilde{S}}(\bar{x}, \bar{y}) \right\}. \tag{24}$$

From the properties of S stated in Theorem 2.2 one can easily infer that

$$T_{\mathrm{Gr}\,\tilde{S}}(\bar{x}, \bar{y}) = \{(d, v) \in T_{\mathrm{Gr}\,S}(\bar{x}, \bar{y}) \mid d \in T_\omega(\bar{x})\},$$

cf., e.g., [13]. Further recall that $T_{\mathrm{Gr}S}(\bar{x}, \bar{y}) = T_{\mathrm{Gr}\sigma}(\bar{x}, \bar{y}) = \mathrm{Gr}\sigma'(\bar{x}; \cdot)$ so that, on the basis of Theorem 2.2, problem (24) amounts to the (linearized) MPEC

$$\begin{aligned} &\text{minimize } \langle -x^*, d\rangle + \langle -y^*, v\rangle \\ &\text{subject to} \\ &\qquad 0 \in \nabla_x F(\bar{x}, \bar{y})d + \nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})v + N_{K(\bar{x},\bar{y})}(v), \;\; d \in T_\omega(\bar{x}). \end{aligned} \tag{25}$$

It remains to write down the standard M-stationarity conditions for (25) at (0, 0). To this end, observe that $K(\bar{x}, \bar{y})$ is a polyhedral cone with vertex at 0 and the critical cone to $K(\bar{x}, \bar{y})$ at 0 with respect to 0 is again $K(\bar{x}, \bar{y})$. Since $\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})$ is strictly copositive with respect to $K(\bar{x}, \bar{y}) - K(\bar{x}, \bar{y})$ by virtue of assumption (ii) in Theorem 2.2, it follows e.g. from [15, Theorem 5.4] that the GE in (25) is strongly regular at (0, 0) [18]. One can thus invoke [12, Proposition 3.2] ensuring that the standard qualification condition

$$\left.\begin{array}{l} (\nabla_x F(\bar{x}, \bar{y}))^T c \in -N_{T_\omega(x)}(0) \\ 0 \in (\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T c + D^* N_{K(\bar{x},\bar{y})}(0, 0)(c) \end{array}\right\} \Rightarrow c = 0$$

is fulfilled. Thus, by [12, Theorem 3.1], there is an MPEC multiplier $b \in \mathbb{R}^m$ such that

$$\begin{aligned} x^* &\in (\nabla_x F(\bar{x}, \bar{y}))^T b + N_{T_\omega(\bar{x})}(0) \\ y^* &\in \left(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\right)^T b + D^* N_{K(\bar{x},\bar{y})}(0, 0)(b). \end{aligned}$$

Since $N_{T_\omega(\bar{x})}(0) \subseteq N_\omega(\bar{x})$ (see [19, Prop. 6.27]), the statement has been proved. □

Based on the previous result, we are now able to state the following new necessary optimality conditions for the MPEC

$$\min \left\{ f(x, y) | 0 \in F(x, y) + \hat{N}_\Gamma(y), \;\; x \in \omega \right\}, \tag{26}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable, $\omega \subseteq \mathbb{R}^n$ is closed and the GE in (26) satisfies all assumptions of Theorem 2.2.

Theorem 3.4 *Consider the MPEC (26) under the assumptions listed above and assume that $(\bar{x}, \bar{y})$ is its (local) solution. Then there exists an MPEC multiplier $\bar{b} \in \mathbb{R}^m$, such that*

$$0 \in \nabla_x f(\bar{x}, \bar{y}) + (\nabla_x F(\bar{x}, \bar{y}))^T \bar{b} + N_\omega(\bar{x}) \tag{27}$$

$$0 \in \nabla_y f(\bar{x}, \bar{y}) + \left(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\right)^T \bar{b} + D^* N_{K(\bar{x},\bar{y})}(0, 0)(\bar{b}), \tag{28}$$

where $\bar{\lambda}$ *is an arbitrary multiplier satisfying the KKT conditions* (7) *with* $(x, y) = (\bar{x}, \bar{y})$.

Proof The assertion follows immediately from the standard optimality condition

$$0 \in \nabla f(\bar{x}, \bar{y}) + \hat{N}_{\mathrm{Gr}\,\tilde{S}}(\bar{x}, \bar{y})$$

by invoking the inclusion (23). □

In the MPEC theory there are several stationarity concepts among which the M- and S-stationarity play a prominent role. When equilibria governed by (1) are considered, M-stationarity conditions involve the limiting coderivative of the multi-valued part $(D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y})))$ whereas in the S-stationary conditions the respective regular coderivative arises. Both conditions differ substantially also in the imposed qualification conditions which are typically stronger in the case of S-stationarity. We refer the reader to [10,20] and [5] for a deeper discussion on this subject.

In the M-stationarity conditions for MPEC (26) (which hold under weaker assumptions compared with Theorem 3.4), the inclusion (28) in the statement of Theorem 3.4 is replaced by

$$0 \in \nabla_y f(\bar{x}, \bar{y}) + \left(\nabla_y F(\bar{x}, \bar{y})\right)^T \bar{d} + D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(\bar{d}), \tag{29}$$

where $\bar{d} \in \mathbb{R}^m$ is a counterpart of the MPEC multiplier $\bar{b}$. The following proposition shows that (29) and (28) are equivalent in case that Γ satisfies LICQ. In the absence of LICQ, which is not required in Theorem 3.4, the coderivative $D^* N_{K(\bar{x},\bar{y})}(0, 0)$ can be still evaluated exactly in (28), whereas the coderivative $D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))$ in (29) may be hard to evaluate at least exactly. An instance of this situation will be illustrated in Example 3.2 following this proposition.

Proposition 3.2 *Consider the setting of Theorem* 3.4 *and assume that the constraint system defining* Γ *fulfills LICQ at* $\bar{y}$. *Then, for any* $b \in \mathbb{R}^m$ *one has*

$$D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(b) = \sum_{i=1}^{s} \bar{\lambda}_i \nabla^2 q_i(\bar{y}) b + D^* N_{K(\bar{x},\bar{y})}(0, 0)(b),$$

where $\bar{\lambda}$ *is the unique solution of the equation* $(\nabla q(\bar{y}))^T \lambda = -F(\bar{x}, \bar{y})$.

Proof Note first that the local structure of Γ and, hence, $D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))$ do not depend on components q_i which are inactive at $\bar{y}$. Similarly, $K(\bar{x}, \bar{y})$ and, hence, $D^* N_{K(\bar{x},\bar{y})}(0, 0)$ are only defined via the components q_i which are active at $\bar{y}$ [see (9)]. Therefore, we may assume without loss of generality that $q(\bar{y}) = 0$. Due to LICQ, one has by (4) (with $l = s$, $t = 0$ and $g = q$),

$$D^* N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))(b) = \sum_{i=1}^{s} \bar{\lambda}_i \nabla^2 q_i(\bar{y}) b + (\nabla q(\bar{y}))^T (A_1 \times \cdots \times A_s), \tag{30}$$

where, for $i = 1, \ldots, s$,

$$A_i = D^* N_{\mathbb{R}_-}\left(0, \bar{\lambda}_i\right)(\nabla q_i(\bar{y})b) = \begin{cases} \mathbb{R} & \text{if } \bar{\lambda}_i > 0 \text{ and } \nabla q_i(\bar{y})\, b = 0 \\ \emptyset & \text{if } \bar{\lambda}_i > 0 \text{ and } \nabla q_i(\bar{y})\, b \neq 0 \\ \{0\} & \text{if } \bar{\lambda}_i = 0 \text{ and } \nabla q_i(\bar{y})\, b < 0 \\ \mathbb{R}_+ & \text{if } \bar{\lambda}_i = 0 \text{ and } \nabla q_i(\bar{y})\, b \geq 0. \end{cases}$$

On the other hand, since Γ fulfills LICQ at $\bar{y}$, the same holds true for $K(\bar{x}, \bar{y})$ in its description (9). Hence, we may apply the same formula (4) with $l = s$, $t = \#I_+(\bar{y}, \bar{\lambda})$ and $g(\cdot) = \nabla q(\bar{y})(\cdot)$. Observing, that $\nabla g \equiv \nabla q(\bar{y})$, it follows that $\nabla^2 g \equiv 0$. Moreover, $g(0) = 0$, and the uniquely defined (by LICQ) multiplier $\bar{\mu}$ satisfying

$$\nabla^T g(0)\mu = \nabla^T q(\bar{y})\mu = 0$$

is $\bar{\mu} = 0$ due to LICQ. Therefore,

$$D^* N_{K(\bar{x},\bar{y})}(0,0)(b) = (\nabla q(\bar{y}))^T (B_1 \times \cdots \times B_s), \tag{31}$$

where, for $i = 1, \ldots, \#I_+(\bar{y}, \bar{\lambda})$ (so that $\bar{\lambda}_i > 0$),

$$B_i = D^* N_{\{0\}}(0,0)(\nabla q_i(\bar{y})b) = \begin{cases} \mathbb{R} & \text{if } \nabla q_i(\bar{y})\, b = 0 \\ \emptyset & \text{else} \end{cases},$$

and, for $i = \#I_+(\bar{y}, \bar{\lambda}) + 1, \ldots, s$ (so that $\bar{\lambda}_i = 0$),

$$B_i = D^* N_{\mathbb{R}_-}(0,0)(\nabla q_i(\bar{y})b) = \begin{cases} \{0\} & \text{if } \nabla q_i(\bar{y})\, b < 0 \\ \mathbb{R}_+ & \text{if } \nabla q_i(\bar{y})\, b \geq 0 \end{cases}.$$

By comparison, it follows that $A_i = B_i$ for all i. Now, combining (30) and (31) yields the assertion. □

The following example illustrates some advantages of the conditions in Theorem 3.4.

Example 3.2 Consider the MPEC

$$\begin{aligned} &\text{minimize} \quad -x + 4y_3 \\ &\text{subject to} \\ &\qquad 0 \in y - \begin{pmatrix} 1 \\ 0 \\ x \end{pmatrix} + N_\Gamma(y), \;\; x \in \mathbb{R}, \;\; y \in \mathbb{R}^3, \end{aligned} \tag{32}$$

where

$$\Gamma = \left\{ y \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} y_1^2 + y_1 - y_3 \\ y_1^2 - y_1 - y_3 \\ y_2^2 + y_2 - y_3 \\ y_2^2 - y_2 - y_3 \end{pmatrix} \in \mathbb{R}^4_- \right\}, \tag{33}$$

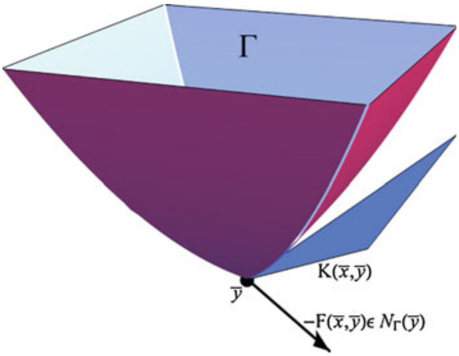

Fig. 1 Illustration of the feasible set Γ defined by (33) and of the critical cone $K(\bar{x}, \bar{y})$

(see Fig. 1). It can easily be seen that $(\bar{x}, \bar{y}) = (-1, 0, 0, 0)$ is a solution of (32). This problem satisfies all assumptions imposed in Theorem 3.4. In particular, MFCQ and CRCQ are fulfilled for Γ. Since, however, LICQ is violated, a straightforward application of the standard M-stationarity condition (29) is not possible because the coderivative $D^*N_\Gamma(\bar{y}, -F(\bar{x}, \bar{y}))$ in (29) is hard to evaluate exactly. Note that also condition (13) is violated so that Theorem 3.1 cannot be applied. Clearly, the KKT conditions (7) are fulfilled with $\bar{\lambda} = (1, 0, 0, 0)$,

$$\nabla_x F(\bar{x}, \bar{y}) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad \nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and

$$K(\bar{x}, \bar{y}) = \{v \in \mathbb{R}^3 | v_1 = v_3, \;\; v_2 \le v_3, \;\; -v_2 \le v_3\}. \tag{34}$$

Since the equalities and inequalities in (34) fulfill the LICQ at (0, 0, 0), we can invoke (4) and deduce that for all $b \in \mathbb{R}^3$

$$D^*N_{K(\bar{x},\bar{y})}(0, 0)(b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ -1 & -1 & -1 \end{pmatrix} (A_1 \times A_2 \times A_3),$$

where

$$A_1 = \begin{cases} \mathbb{R} & \text{if } b_1 = b_3 \\ \emptyset & \text{else} \end{cases} \quad A_2 = \begin{cases} \mathbb{R}_+ & \text{if } b_2 \ge b_3 \\ \{0\} & \text{else} \end{cases} \quad A_3 = \begin{cases} \mathbb{R}_+ & \text{if } \; b_2 + b_3 \le 0 \\ \{0\} & \text{else.} \end{cases}.$$

Conditions (27) take the form

$$b_3 = -1, \;\; 0 \in \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 - d_3 \\ -d_1 - d_2 - d_3 \end{pmatrix}$$

with $d_i \in A_i$ $(i = 1, 2, 3)$. These conditions are fulfilled, e.g., with

$$\bar{b} = (-1, 0, -1)^T, \quad \bar{d} = (3, 0, 0)^T.$$

Alternatively, one might be tempted to apply [12, Theorem 3.1] to the associated "enhanced" MPEC

$$\begin{aligned} &\text{minimize} \ \ -x + 4y_3 \\ &\text{subject to} \\ &\qquad 0 \in y - \begin{pmatrix} 1 \\ 0 \\ x \end{pmatrix} + (\nabla q(y))^T \lambda, \ \ (\lambda, q(y)) \in \operatorname{Gr} N_{\mathbb{R}^4_+}, \end{aligned} \tag{35}$$

at the "enhanced" solution $(\bar{x}, \bar{y}, \bar{\lambda})$ [where q is specified in (33)]. It turns out, however, that the qualification condition of this statement is violated. Another option would be to apply the upper estimate of $D^*S(\bar{x}, \bar{y})$ along the lines of [11, Theorem 3.1]. This, however, would require the verification of a calmness condition whose validity is not evident in this example.

We note that in the previous example one could easily calculate the coderivative $D^*N_{K(\bar{x},\bar{y})}(0, 0)$ because the small dimension of $m = 3$ led to LICQ being satisfied for the critical cone $K(\bar{x}, \bar{y})$ at 0 even if LICQ was violated for Γ. This can no longer be expected in larger dimension. In that case, LICQ may be violated for $K(\bar{x}, \bar{y})$ too, thus prohibiting the application of transformation formula (4). Nonetheless, one may fall back on a transformation formula for coderivatives working in the polyhedral case without any constraint qualification [7, Proposition 3.2]. Specifying this formula to the case of the polyhedral critical cone $K(\bar{x}, \bar{y})$, one would arrive at the following formula:

$$D^*N_{K(\bar{x},\bar{y})}(0, 0)(b) = \left\{ d \,\middle|\, (d, -b) \in \bigcup_{I_+(\bar{y},\bar{\lambda}) \subseteq I_1 \subseteq I_2 \subseteq I(\bar{y}), C_{I_2} \neq \emptyset} A_{I_1,I_2} \times A^0_{I_1,I_2} \right\},$$

where

$$\begin{aligned} A_{I_1,I_2} &:= \left\{ \sum_{i \in I_2} \mu_i \nabla^T q_i(\bar{y}) \,\middle|\, \mu_i \geq 0 \quad (i \in I_2 \backslash I_1) \right\} \\ C_{I_2} &:= \{v | \nabla q_i(\bar{y})v = 0 \quad (i \in I_2), \quad \nabla q_i(\bar{y})v < 0 \quad (i \in I(\bar{y}) \backslash I_2)\}. \end{aligned}$$

4 When S is not directionally differentiable

Consider again Γ given by (5) and assume that all functions q_i are convex. Our first aim in this section is the computation of $\hat{N}_{\operatorname{Gr} N_\Gamma}$ at a pair $(\bar{y}, \bar{v}) \in \operatorname{Gr} N_\Gamma$. Let $\mathcal{K}$ denote the critical cone to Γ at $\bar{y}$ with respect to $\bar{v}$, i.e.,

$$\mathcal{K} = T_\Gamma(\bar{y}) \cap \{\bar{v}\}^\perp.$$

Proposition 4.1 *Let $\bar{v} \in N_\Gamma(\bar{y})$ and assume that both MFCQ and CRCQ hold at $\bar{y}$. Further, let λ be an arbitrary multiplier satisfying conditions*

$$\bar{v} = \sum_{i=1}^{s} \lambda_i \nabla q_i(\bar{y}), \lambda \geq 0, \langle q(\bar{y}), \lambda \rangle = 0. \tag{36}$$

Then

$$\hat{N}_{\mathrm{Gr}\, N_\Gamma}(\bar{y}, \bar{v}) = \left\{ \left(y^*, v^*\right) \in \mathbb{R}^m \times \mathbb{R}^m | y^* + \left(\sum_{i=1}^{s} \lambda_i \nabla^2 q_i\,(\bar{y}) \right) v^* \in \mathcal{K}^0,\;\; v^* \in \mathcal{K} \right\}. \tag{37}$$

Proof Put $\bar{x} = \bar{y} + \bar{v}$ and consider the graph of the projection mapping on Γ denoted by P_Γ, at $(\bar{x}, \bar{y})$. A straightforward application of Corollary 3.2 to the GE

$$x \in y + N_\Gamma(y)$$

yields the formula

$$\hat{N}_{\mathrm{Gr}\, P_\Gamma}(\bar{x}, \bar{y}) = \left\{ \left(x^*, y^*\right) \in \mathbb{R}^m \times \mathbb{R}^m | x^* \in \mathcal{K},\;\; y^* + A^T x^* \in \mathcal{K}^0 \right\},$$

where $A = I + \sum_{i=1}^{s} \lambda_i \nabla^2 q_i\,(\bar{y})$. Next we observe that

$$(y, v) \in \mathrm{Gr}\, N_\Gamma \Longleftrightarrow (y + v, y) \in \mathrm{Gr}\, P_\Gamma.$$

This allows us to invoke [10, Theorem 1.66], according to which

$$\begin{aligned} \hat{N}_{\mathrm{Gr}\, N_\Gamma}(\bar{y}, \bar{v}) &= \begin{bmatrix} I & I \\ I & 0 \end{bmatrix} \hat{N}_{\mathrm{Gr}\, P_\Gamma}(\bar{y} + \bar{v}, \bar{y}) \\ &= \left\{ \left(y^*, v^*\right) \in \mathbb{R}^m \times \mathbb{R}^m | y^* - v^* + A^T v^* \in \mathcal{K}^0,\;\; v^* \in \mathcal{K} \right\}, \end{aligned}$$

which leads directly to (37). The statement has been established. □

By invoking [10, Theorem 1.66] once more, we can now compute the desired regular coderivative of the solution mapping to (1) without assuming condition (ii) of Theorem 2.2.

Theorem 4.1 *Consider the GE (1) with Γ given by (5), where all functions q_i are convex. Let $(\bar{x}, \bar{y}) \in \mathrm{Gr}\, S$ and assume that Γ fulfills MFCQ and CRCQ at $\bar{y}$. Furthermore, suppose that $\nabla_x F(\bar{x}, \bar{y})$ is surjective. Then, formula (20) holds true.*

Proof Clearly, by the surjectivity of $\nabla_x F(\bar{x}, \bar{y})$, one has for all $y^* \in \mathbb{R}^m$

$$\hat{D}^* S(\bar{x}, \bar{y})(y^*) = \Big\{ (\nabla_x F(\bar{x}, \bar{y}))^T b \,\Big|\, 0 = y^* + a + \big(\nabla_y F(\bar{x}, \bar{y})\big)^T b,\ (a, -b) \in \hat{N}_{\mathrm{Gr}\, N_\Gamma}(\bar{y}, -F(\bar{x}, \bar{y})) \Big\}.$$

The rest follows from (37) by putting $\bar{v} = -F(\bar{x}, \bar{y})$ and from the fact that $\mathcal{K} = K(\bar{x}, \bar{y})$, given by (9). □

Let us point out the difference in the assumptions of Theorems 3.1 and 4.1. While in the latter the Strong Second-Order Sufficient Condition is not imposed, one has to pay the price of requiring the convexity of all functions q_i independently of F. Moreover, the surjectivity of $\nabla_x F(\bar{x}, \bar{y})$ cannot be replaced by the weaker condition (13) as imposed in Theorem 3.1. When (1) amounts to KKT conditions of a parametric NLP, then, roughly speaking, Theorem 3.1 requires a kind of strong convexity of the objective, whereas Theorem 4.1 needs a convex constraint set.

On the basis of Theorem 4.1 we immediately obtain necessary optimality conditions for the MPEC

$$\min\{f(x, y) \mid 0 \in F(x, y) + N_\Gamma(y)\}, \tag{38}$$

where one does not have any non-equilibrium constraints.

Theorem 4.2 *Let $(\bar{x}, \bar{y})$ be a (local) solution of the MPEC (38), where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable near $(\bar{x}, \bar{y})$ and the GE fulfills all assumptions made in Theorem 4.1. Then, there is an MPEC multiplier $\bar{b} \in -K(\bar{x}, \bar{y})$ such that*

$$0 = \nabla_x f(\bar{x}, \bar{y}) + (\nabla_x F(\bar{x}, \bar{y}))^T \bar{b} \tag{39}$$

$$0 \in \nabla_y f(\bar{x}, \bar{y}) + \big(\nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda})\big)^T \bar{b} + K^0(\bar{x}, \bar{y}), \tag{40}$$

where $\bar{\lambda}$ is an arbitrary multiplier satisfying conditions (7) with $(x, y) = (\bar{x}, \bar{y})$.

Proof The statement follows immediately from the standard optimality condition

$$0 \in \nabla f(\bar{x}, \bar{y}) + \hat{N}_{\mathrm{Gr}\, S}(\bar{x}, \bar{y})$$

by virtue of Theorem 4.1. □

These conditions correspond to the notion of S-stationarity as introduced in [20]. Their nature is illustrated by the next example.

Example 4.1 Consider the bilevel program

$$\begin{aligned} &\text{minimize} \quad y_3 - 0.5\tfrac{x_3}{x_1} \\ &\text{subject to} \\ &\qquad\qquad 0 \in x + N_\Gamma(y),\ \ x, y \in \mathbb{R}^3, \end{aligned} \tag{41}$$

where Γ is given in Example 3.2. Observe that the GE in (41) represents the necessary and sufficient optimality conditions of the NLP

$$\min\{\langle x, y\rangle \,|\, y \in \Gamma\}.$$

It can be easily verified that $\bar{x} = (-1, 0, 1)$, $\bar{y} = 0$ is a local solution to (41). Note that the results of Sect. 3 are not applicable because assumption (ii) of Theorem 2.2 is violated. On the other hand, all assumptions of Theorem 4.2 are fulfilled. Clearly, the equality

$$0 = \bar{x} + \sum_{i=1}^{4} \lambda_i \nabla q_i\,(\bar{y})$$

holds, for instance, with $\bar{\lambda}_1 = 1, \bar{\lambda}_2 = \bar{\lambda}_3 = \bar{\lambda}_4 = 0$. Conditions (39), (40) attain the form

$$0 = \begin{pmatrix} 0.5 \\ 0 \\ 0.5 \end{pmatrix} + \bar{b}; \quad 0 \in \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bar{b} + K^0(\bar{x}, \bar{y}). \tag{42}$$

Since $\bar{b} := (-0.5, 0, -0.5)^T$ belongs to $-K(\bar{x}, \bar{y})$ and the second condition in (42) is fulfilled as well, the pair $(\bar{x}, \bar{y})$ satisfies the optimality condition of Theorem 4.2.

5 Conclusions

We have considered a of GEs with non-polyhedral constraint sets given by inequalities under MFCQ, CRCQ and various additional conditions. In this setting we have provided an exact formula for the regular coderivative of the respective solution map and, under weakened assumptions, a tight upper estimate of it. These results have been utilized in various types of MPECs and lead to various new sharp necessary optimality conditions. Some of them (e.g., Theorem 4.2) are related to the notion of S-stationarity and extend thus the available results for equilibria governed by complementarity problems. The conditions of Theorem 3.4, however, are closer to the notion of M-stationarity and eliminate possible problems with qualification conditions arising in the standard approach (cf., e.g., [12,21]) in the absence of LICQ. Finally, we would like to stress the fact that many of the previous results concerning formulae for the coderivative of the solution mapping to GEs of the type considered here are formulated as inclusions in which the larger set is expressed as the union over *all* (Lagrange) multipliers, which in the absence of LICQ can result in a poor estimate of the true expression. The advantage of the results here is then evident, as only one multiplier is required.

Acknowledgments The second author would like to express his gratitude to S. Dempe, D. Ralph and S. Scholtes who have motivated his interest in MPECs with non-unique lower-level multipliers. All authors would like to express their gratitude to both reviewers for their important suggestions and comments.

References

1. Dempe, S.: Foundations of Bilevel Programming. Kluwer, Dordrecht (2002)
2. Dontchev, A.L., Rockafellar, R.T.: Characterization of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. **7**, 1087–1105 (1996)
3. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings. Springer, Heidelberg (2009)
4. Fiacco, A.V., McCormick, G.P.: Nonlinear Programming: Sequential Unconstrained Minimization Techniques. Wiley, New York (1968)
5. Flegel, M., Kanzow, C., Outrata, J.V.: Optimality conditions for disjunctive programs with application to mathematical program with equilibrium constraints. Set-Valued Anal. **15**, 139–162 (2007)
6. Henrion, R., Jourani, A., Outrata, J.V.: On the calmness of a class of multifunctions. SIAM J. Optim. **13**, 603–618 (2002)
7. Henrion, R., Römisch, W.: On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. Appl. Math. **52**, 473–494 (2007)
8. Levy, A.B., Mordukhovich, B.S.: Coderivatives in parametric optimization. Math. Program. **99**, 311–327 (2004)
9. Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
10. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory, vol. 2: Applications. Springer, Berlin (2006)
11. Mordukhovich, B.S., Outrata, J.V.: Coderivative analysis of quasivariational inequalities with applications to stability and optimization. SIAM J. Optim. **18**, 389–412 (2007)
12. Outrata, J.V.: A generalized mathematical program with equilibrium constraints. SIAM J. Control Optim. **38**, 1623–1638 (2000)
13. Outrata, J.V., Červinka, M.: On the implicit programming approach in a class of mathematical programs with equilibrium constraints. Control Cybernet. **38**, 1557–1574 (2009)
14. Outrata, J.V., Jarušek, J., Stará, J.: On optimality conditions in control of elliptic varriational inequalities. Set Valued Var. Anal. **19**, 23–42 (2011)
15. Outrata, J.V., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results. Kluwer, Dordrecht (1998)
16. Ralph, D., Dempe, S.: Directional derivatives of the solution of a parametric nonlinear program. Math. Program. **70**, 159–172 (1995)
17. Robinson, S.M.: Some continuity properties of polyhedral multifunctions. Math. Program. Stud. **14**, 206–214 (1976)
18. Robinson, S.M.: Strongly regular generalized equations. Math. Oper. Res. **5**, 43–62 (1980)
19. Rockafellar, R.T., Wets, J.-B.: Variational Analysis. Springer, Berlin (1998)
20. Scheel, H., Scholtes, S.: Mathematical Programs with Complementarity Constraints: Stationarity, Optimality and Sensitivity. Math. Oper. Res. **25**, 1–22 (2000)
21. Ye, J.J.: Constraint qualifications and necessary optimality conditions for optimality conditions for optimization problems with variational inequality constraints. SIAM J. Optim. **10**, 943–962 (2000)