AN ABADIE-TYPE CONSTRAINT QUALIFICATION FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

Michael L. Flegel and Christian Kanzow

University of Würzburg
Institute of Applied Mathematics and Statistics
Am Hubland
97074 Würzburg
Germany

e-mail: flegel@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de

November 12, 2002

Abstract. Mathematical programs with equilibrium constraints (MPECs) are nonlinear programs which do not satisfy any of the common constraint qualifications. In order to obtain first order optimality conditions, constraint qualifications tailored to MPECs have been developed and researched in the past. In this paper we introduce a new Abadie-type constraint qualification for MPECs. We investigate necessary conditions for this new CQ, discuss its relationship to several existing MPEC constraint qualifications and introduce a new Slater-type constraint qualification. Finally, we prove a new stationarity concept to be a necessary optimality condition under our new Abadie-type CQ.

Key Words. Mathematical programs with equilibrium constraints, Abadie constraint qualification, Slater constraint qualification, optimality conditions.

1 Introduction

Consider the constrained optimization problem

$$
\begin{array}{ll}
\min & f(z) \\
\text{s.t.} & g(z) \leq 0, \quad h(z) = 0, \\
& G(z) \geq 0, \quad H(z) \geq 0, \quad G(z)^T H(z) = 0.
\end{array} \tag{1}
$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^p$, $G : \mathbb{R}^n \to \mathbb{R}^l$, and $H : \mathbb{R}^n \to \mathbb{R}^l$ are continuously differentiable functions. Problems of this kind are usually called *mathematical programs with equilibrium constraints*, MPEC for short, or sometimes *mathematical programs with complementarity constraints*. See, e.g., the two monographs [6, 10] for more information.

It is well-known (see, e.g., [2, 14]) and easily verified that the MPEC (1) does not satisfy most of the common constraint qualifications known from standard nonlinear programming at any feasible point. (One exception to this is the Guignard constraint qualification, see [4] for details.) Consequently, the usual Karush-Kuhn-Tucker conditions associated with the program (1) can, in general, not be viewed as first order optimality conditions for (1).

It has therefore been the subject of intensive research [8, 9, 13] during the last few years to find suitable MPEC constraint qualifications under which a local minimizer of the problem (1) satisfies some first order optimality conditions.

In this paper we introduce a new Abadie-type constraint qualification for MPECs. We examine the relationship between this constraint qualification and existing ones, and show it to be weaker than any of the existing ones. We also introduce a new Slater-type constraint qualification for MPECs and show that it also implies our Abadie-type condition. Furthermore, we introduce a new optimality condition which holds under our Abadie-type CQ.

The organization of this paper is as follows: Section 2 reviews some existing constraint qualifications together with some stationarity concepts related to the problem (1). Section 3 contains an Abadie-type approach to first order optimality conditions for (1). In Section 4 we proceed to compare our Abadie-type constraint qualification to some constraint qualifications used in [6]. It turns out that, in a special situation, our Abadie-type constraint qualification is equivalent to one used in [6], although its derivation is different. We conclude this paper with a summary of our results in Section 5.

The notation used in this paper is rather standard: $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, we simply write (x, y) for the $(n+m)$-dimensional column vector $(x^T, y^T)^T$. Given $x \in \mathbb{R}^n$ and a subset $\delta \subseteq \{1, \ldots, n\}$, we denote by x_δ the subvector in $\mathbb{R}^{|\delta|}$ consisting of all components x_i with $i \in \delta$. Finally, inequalities of vectors are defined componentwise.

2 Constraint Qualifications and Stationarity Concepts

This section reviews some existing constraint qualifications for the MPEC (1) as well as some existing first order optimality conditions. Additional (new) constraint qualifications and stationary concepts will be introduced in Section 3.

Before we begin, we need to introduce some notation. Given a feasible vector z^* of the MPEC (1), we define the following sets of indices:

$$\alpha := \alpha(z^*) := \{i \mid G_i(z^*) = 0,\ H_i(z^*) > 0\}, \tag{2a}$$
$$\beta := \beta(z^*) := \{i \mid G_i(z^*) = 0,\ H_i(z^*) = 0\}, \tag{2b}$$
$$\gamma := \gamma(z^*) := \{i \mid G_i(z^*) > 0,\ H_i(z^*) = 0\}. \tag{2c}$$

The set β is known as the *degenerate* set. If it is empty, the vector z^* is said to fulfill *strict complementarity.* As we shall see, it will become convenient to split β into its partitions, which are defined as follows:

$$\mathcal{P}(\beta) := \{(\beta_1, \beta_2) \mid \beta_1 \cup \beta_2 = \beta, \beta_1 \cap \beta_2 = \emptyset\}. \tag{3}$$

To define altered constraint qualifications, we introduce the following program, dependent on z^*, and called the *tightened nonlinear program TNLP* $:=$ *TNLP*(z^*):

$$\begin{array}{lll} \min & f(z) & \\ \text{s.t.} & g(z) \leq 0, & h(z) = 0, \\ & G_{\alpha\cup\beta}(z) = 0, & G_\gamma(z) \geq 0, \\ & H_\alpha(z) \geq 0, & H_{\gamma\cup\beta}(z) = 0. \end{array} \tag{4}$$

The above nonlinear program is called *tightened* since the feasible region is a subset of the feasible region of the MPEC (1). This implies that if z^* is a local minimizer of the MPEC (1), then it is also a local minimizer of the corresponding tightened nonlinear program TNLP(z^*).

The TNLP (4) can now be used to define suitable MPEC variants of the standard linear independence, Mangasarian-Fromovitz- and strict Mangasarian-Fromovitz constraint qualifications (LICQ, MFCQ, and SMFCQ for short).

Definition 2.1 *The MPEC (1) is said to satisfy the* MPEC-LICQ (MPEC-MFCQ, MPEC-SMFCQ) *in a feasible vector z^* if the corresponding TNLP(z^*) satisfies the LICQ (MFCQ, SMFCQ) in that vector z^*.*

Since we will need them in Section 3, we explicitly write down the conditions for MPEC-MFCQ: The gradient vectors

$$\begin{array}{ll} \nabla h_i(z^*) & \forall i = 1, \dots, p, \\ \nabla G_i(z^*) & \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*) & \forall i \in \gamma \cup \beta \end{array} \tag{5a}$$

are linearly independent, and there exists a vector $d \in \mathbb{R}^n$ such that

$$\begin{aligned} \nabla h_i(z^*)^T d &= 0 \qquad && \forall i = 1, \dots, p, \\ \nabla G_i(z^*)^T d &= 0 && \forall i \in \alpha \cup \beta, \\ \nabla H_i(z^*)^T d &= 0 && \forall i \in \gamma \cup \beta, \\ \nabla g_i(z^*)^T d &< 0 && \forall i \in \mathcal{I}_g. \end{aligned} \tag{5b}$$

MPEC-LICQ and MPEC-SMFCQ can be expanded similarly.

As mentioned earlier, classic KKT conditions are not appropriate in the context of MPECs. We therefore introduce two stationarity conditions used in [11, 13].

A feasible point z of the MPEC (1) is called *weakly stationary* [13] if there exists a Lagrange multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that the following conditions hold:

$$\begin{aligned} &0 = \nabla f(z) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z) + \lambda_i^H H_i(z)\right], \\ &\lambda_\alpha^G \ \text{ free}, \qquad \lambda_\beta^G \ \text{ free}, \qquad \lambda_\gamma^G = 0, \\ &\lambda_\gamma^H \ \text{ free}, \qquad \lambda_\beta^H \ \text{ free}, \qquad \lambda_\alpha^H = 0, \\ &g(z) \le 0, \qquad \lambda^g \ge 0, \qquad (\lambda^g)^T g(z) = 0. \end{aligned} \tag{6}$$

A feasible point z of the MPEC (1) is called *strongly stationary* [13] or *primal-dual stationary* [11] if there exists a Lagrange multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that the following conditions hold:

$$\begin{aligned} &0 = \nabla f(z) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(z) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(z) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(z) + \lambda_i^H H_i(z)\right], \\ &\lambda_\alpha^G \ \text{ free}, \qquad \lambda_\beta^G \ge 0, \qquad \lambda_\gamma^G = 0, \\ &\lambda_\gamma^H \ \text{ free}, \qquad \lambda_\beta^H \ge 0, \qquad \lambda_\alpha^H = 0, \\ &g(z) \le 0, \qquad \lambda^g \ge 0, \qquad (\lambda^g)^T g(z) = 0. \end{aligned} \tag{7}$$

Note that the difference between the two stationarity conditions is the sign-restriction imposed on λ_β^G and λ_β^H in the case of strong stationarity. It is easily verified that strong stationarity coincides with the KKT conditions of the MPEC (1) (see, e.g., [4]). Furthermore, in the nondegenerate case, i.e. if $\beta = \emptyset$, strong stationarity is identical to weak stationarity.

Other stationary conditions are derived and examined elsewhere, among which are C-stationarity [13] and M-stationarity [8, 9]. Both lie between the weak and strong stationarity conditions (6) and (7), respectively. Hence they all coincide in the nondegenerate case, whereas, in general, differences occur in the properties of the multipliers λ_β^G and λ_β^H.

Since we will need it repeatedly in the following sections, we now define another program derived from the MPEC (1). Given a partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$, let $\mathrm{NLP}_*(\beta_1, \beta_2)$ denote the following nonlinear program:

$$
\begin{array}{rll}
\min & f(z) & \\
\text{s.t.} & g(z) \le 0, & h(z) = 0, \\
& G_{\alpha\cup\beta_1}(z) = 0, & H_{\alpha\cup\beta_1}(z) \ge 0, \\
& G_{\gamma\cup\beta_2}(z) \ge 0, & H_{\gamma\cup\beta_2}(z) = 0.
\end{array} \tag{8}
$$

Note that the program $\mathrm{NLP}_*(\beta_1, \beta_2)$ depends on the vector z^*.

Also, a local minimizer z^* of the MPEC (1) is a local minimizer of the $\mathrm{NLP}_*(\beta_1, \beta_2)$ since z^* is feasible for the latter program and its feasible region is a subset of the feasible region of the MPEC (1).

3 Abadie-type Approach to Optimality Conditions

We divide this section into three parts: Section 3.1 contains the definition of our Abadie-type constraint qualification (MPEC-Abadie CQ) as well as some discussion on it. Several sufficient conditions for the MPEC-Abadie CQ to hold are given in Section 3.2. Finally, Section 3.3 provides a stationarity concept which holds under the MPEC-Abadie CQ.

3.1 The MPEC-Abadie Constraint Qualification

In Definition 2.1 we introduced MPEC variants of some common constraint qualifications. The question arises whether a suitable variant of the Abadie constraint qualification can be found. Since it is among the weaker constraint qualification of standard nonlinear programming, this question seems to be of some importance.

Some background is needed before we can delve into this question, however. Consider therefore the MPEC (1) and let $\mathcal{Z}$ denote its feasible region. Then the tangent cone of (1) in a feasible point z^* is defined by

$$
\mathcal{T}(z^*) := \big\{ d \in \mathbb{R}^n \;\big|\; \exists \{z^k\} \subset \mathcal{Z}, \exists t_k \searrow 0 \,:\, z^k \to z^* \text{ and } \frac{z^k - z^*}{t_k} \to d \big\}. \tag{9}
$$

Note that the tangent cone is closed, but not convex, in general.

If we set $\theta(z) := G(z)^T H(z)$, then the standard linearized tangent cone of the MPEC (1) in a feasible point z^* is given by

$$
\begin{aligned}
\mathcal{T}^{lin}(z^*) := \{ d \in \mathbb{R}^n \mid \; & \nabla g_i(z^*)^T d \le 0, && \forall i \in \mathcal{I}_g, \\
& \nabla h_i(z^*)^T d = 0, && \forall i = 1, \dots, p, \\
& \nabla G_i(z^*)^T d \ge 0, && \forall i \in \alpha \cup \beta, \\
& \nabla H_i(z^*)^T d \ge 0, && \forall i \in \gamma \cup \beta, \\
& \nabla \theta(z^*)^T d = 0 \, \}. &&
\end{aligned} \tag{10}
$$

An easy calculation shows that the following characterization of the linearized tangent cone is equivalent to the one in (10):

$$
\begin{aligned}
\mathcal{T}^{lin}(z^*) = \{d \in \mathbb{R}^n \mid\ & \nabla g_i(z^*)^T d \le 0, && \forall i \in \mathcal{I}_g, \\
& \nabla h_i(z^*)^T d = 0, && \forall i = 1, \dots, p, \\
& \nabla G_i(z^*)^T d = 0, && \forall i \in \alpha, \\
& \nabla H_i(z^*)^T d = 0, && \forall i \in \gamma, \\
& \nabla G_i(z^*)^T d \ge 0, && \forall i \in \beta, \\
& \nabla H_i(z^*)^T d \ge 0, && \forall i \in \beta\,\}.
\end{aligned} \tag{11}
$$

It is well-known that the inclusion

$$\mathcal{T}(z^*) \subseteq \mathcal{T}^{lin}(z^*). \tag{12}$$

holds, and the standard Abadie constraint qualification (abbreviated ACQ) for nonlinear programs requires equality in (12):

$$\mathcal{T}(z^*) = \mathcal{T}^{lin}(z^*). \tag{13}$$

While this condition is likely to be satisfied for standard nonlinear programming, it is not appropriate for the MPEC (1). As simple examples show, the tangent cone $\mathcal{T}(z^*)$ is, in general, not convex, whereas the linearized cone $\mathcal{T}^{lin}(z^*)$ is polyhedral and hence, in particular, convex. This circumstance suggests the definition of the following cone:

$$
\begin{aligned}
\mathcal{T}^{lin}_{\text{MPEC}}(z^*) := \{d \in \mathbb{R}^n \mid\ & \nabla g_i(z^*)^T d \le 0, \forall i \in \mathcal{I}_g, \\
& \nabla h_i(z^*)^T d = 0, \forall i = 1, \dots, p, \\
& \nabla G_i(z^*)^T d = 0, \forall i \in \alpha, \\
& \nabla H_i(z^*)^T d = 0, \forall i \in \gamma, \\
& \nabla G_i(z^*)^T d \ge 0, \forall i \in \beta, \\
& \nabla H_i(z^*)^T d \ge 0, \forall i \in \beta, \\
& (\nabla G_i(z^*)^T d) \cdot (\nabla H_i(z^*)^T d) = 0, \quad \forall i \in \beta\,\}.
\end{aligned} \tag{14}
$$

This set has appeared in [13, 11] before, but was not investigated further in either paper. Here, it will play an important role.

Obviously, we have

$$\mathcal{T}^{lin}_{\text{MPEC}}(z^*) \subseteq \mathcal{T}^{lin}(z^*), \tag{15}$$

which can be seen by comparing $\mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ with the representation (11) of $\mathcal{T}^{lin}(z^*)$. However, the relationship between $\mathcal{T}(z^*)$ and $\mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ is not that apparent. To shed some light on this, consider the nonlinear program $\text{NLP}_*(\beta_1, \beta_2)$ associated with an arbitrary partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$. Let $\mathcal{T}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*)$ be the tangent cone of $\mathrm{NLP}_*(\beta_1, \beta_2)$, and let $\mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*)$ be the standard linearized tangent cone of $\mathrm{NLP}_*(\beta_1, \beta_2)$, i.e.,

$$
\begin{aligned}
\mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*) = \{d \in \mathbb{R}^n \mid \; & \nabla g_i(z^*)^T d \leq 0, && \forall i \in \mathcal{I}_g, \\
& \nabla h_i(z^*)^T d = 0, && \forall i = 1, \dots, p, \\
& \nabla G_i(z^*)^T d = 0, && \forall i \in \alpha \cup \beta_1, \\
& \nabla H_i(z^*)^T d = 0, && \forall i \in \gamma \cup \beta_2, \\
& \nabla G_i(z^*)^T d \geq 0, && \forall i \in \beta_2, \\
& \nabla H_i(z^*)^T d \geq 0, && \forall i \in \beta_1 \, \}.
\end{aligned}
\tag{16}
$$

Then the following result holds.

Lemma 3.1 *Let z^* be a feasible vector of the MPEC (1). Then the following statements hold:*

(a) $\mathcal{T}(z^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}_{NLP_*(\beta_1,\beta_2)}(z^*)$,

(b) $\mathcal{T}^{lin}_{MPEC}(z^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}^{lin}_{NLP_*(\beta_1,\beta_2)}(z^*)$.

Proof. Both equalities are easily verified. Also, the one in (b) has been stated previously in [11]. □

Lemma 3.1 demonstrates that the tangent set $\mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*)$ is non-convex in general, which gives hope that it may be more likely to be equal to the tangent cone $\mathcal{T}(z^*)$. This of course gives rise to the question whether the inclusion (12) from nonlinear program transfers to the tangent set $\mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*)$. This is indeed so, as we prove in the following lemma.

Lemma 3.2 *The inclusion*

$$
\mathcal{T}(z^*) \subseteq \mathcal{T}^{lin}_{MPEC}(z^*). \tag{17}
$$

holds for all feasible z^*.

Proof. It is known from nonlinear programming that

$$
\mathcal{T}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*) \subseteq \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*)
$$

(see (12)). It follows immediately that

$$
\bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*) \subseteq \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*). \tag{18}
$$

Together with the two equalities of Lemma 3.1, this yields

$$
\mathcal{T}(z^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*) \subseteq \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*) = \mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*),
$$

which proves the result. □

In view of Lemma 3.2, the inclusion (15) is supplemented to yield the following:

$$\mathcal{T}(z^*) \subseteq \mathcal{T}^{lin}_{\text{MPEC}}(z^*) \subseteq \mathcal{T}^{lin}(z^*). \tag{19}$$

As mentioned before, $\mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ is, in general, a nonconvex cone, and it therefore seems reasonable to require that the equality $\mathcal{T}(z^*) = \mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ holds. This is precisely how we define our MPEC variant of the Abadie constraint qualification.

Definition 3.3 *The MPEC (1) is said to satisfy* MPEC-ACQ *in a feasible vector z^* if*

$$\mathcal{T}(z^*) = \mathcal{T}^{lin}_{MPEC}(z^*)$$

holds.

As in the case of standard nonlinear programming, it is not difficult to find examples where the MPEC-ACQ is not satisfied. In fact, such an example may be taken from nonlinear programming and expanded with unproblematic equilibrium constraints ($G(\cdot)$ and $H(\cdot)$) to serve as an example in our scenario.

Far more interesting is, of course, when we might expect MPEC-ACQ to hold.

3.2 Sufficient Conditions for MPEC-ACQ

The following corollary to Lemma 3.1 gives a first insight into this question.

Corollary 3.4 *If, for every partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$, the Abadie constraint qualification holds for $NLP_*(\beta_1, \beta_2)$, i.e.*

$$\mathcal{T}_{NLP_*(\beta_1,\beta_2)}(z^*) = \mathcal{T}^{lin}_{NLP_*(\beta_1,\beta_2)}(z^*) \qquad \forall(\beta_1, \beta_2) \in \mathcal{P}(\beta),$$

then

$$\mathcal{T}(z^*) = \mathcal{T}^{lin}_{MPEC}(z^*),$$

i.e. MPEC-ACQ holds.

Note that the condition that Abadie CQ hold for all $\text{NLP}_*(\beta_1, \beta_2)$ has appeared in [11] and was used in conjunction with other assumptions to prove a result for a necessary optimality condition.

Before we can prove our next result, which will clarify the relationship between MPEC-MFCQ and MPEC-ACQ, we need the following lemma.

Lemma 3.5 *If a feasible point z^* of the MPEC (1) satisfies MPEC-MFCQ, then classic MFCQ is satisfied in z^* by the corresponding nonlinear program $NLP_*(\beta_1, \beta_2)$ (8) for any partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$.*

Proof. The set of linear equations

$$\begin{pmatrix} \nabla h_i(z^*)^T & (i = 1, \ldots, p) \\ \nabla G_i(z^*)^T & (i \in \alpha \cup \beta_1) \\ \nabla H_i(z^*)^T & (i \in \gamma \cup \beta_2) \\ \nabla G_i(z^*)^T & (i \in \beta_2) \\ \nabla H_i(z^*)^T & (i \in \beta_1) \end{pmatrix} \hat{d} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}}\right\} p + |\alpha \cup \beta_1| + |\gamma \cup \beta_2| \\ \left.\vphantom{\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix}}\right\} |\beta_2| + |\beta_1| \end{matrix}$$

has a solution $\hat{d}$ because the coefficient matrix has full rank (MPEC-MFCQ holds in z^*, see (5a)).

We now choose $d \in \mathbb{R}^n$ to satisfy the conditions (5b) of MPEC-MFCQ and set

$$d(\delta) := d + \delta \hat{d}.$$

It is easy to see that $d(\delta)$ satisfies the following conditions for all $\delta > 0$:

$$\begin{aligned} \nabla h_i(z^*)^T d(\delta) &= 0, & \forall i &= 1, \ldots, p, \\ \nabla G_i(z^*)^T d(\delta) &= 0, & \forall i &\in \alpha \cup \beta_1, \\ \nabla H_i(z^*)^T d(\delta) &= 0, & \forall i &\in \gamma \cup \beta_2, \\ \nabla G_i(z^*)^T d(\delta) &> 0, & \forall i &\in \beta_2, \\ \nabla H_i(z^*)^T d(\delta) &> 0, & \forall i &\in \beta_1. \end{aligned} \tag{20}$$

Since the inequality in (5b) is strict, we have

$$\nabla g_i(z^*)^T d(\delta) < 0, \qquad \forall i \in \mathcal{I}_g \tag{21}$$

for all $\delta > 0$ sufficiently small.

We have therefore shown the existence of a vector satisfying the conditions (20) and (21), which is part of the MFCQ for $\mathrm{NLP}_*(\beta_1, \beta_2)$. Additionally, the gradient vectors

$$\begin{aligned} &\nabla h_i(z^*), & \forall i &= 1, \ldots, p, \\ &\nabla G_i(z^*), & \forall i &\in \alpha \cup \beta_1, \\ &\nabla H_i(z^*), & \forall i &\in \gamma \cup \beta_2 \end{aligned}$$

are linearly independent since they are a subset of the linearly independent gradient vectors (5a) in MPEC-MFCQ. This completes the conditions for MFCQ for $\mathrm{NLP}_*(\beta_1, \beta_2)$, and hence the proof. □

We are now able to prove the following theorem which shows that, analogous to standard nonlinear programming, MPEC-MFCQ implies MPEC-ACQ.

Theorem 3.6 *If a feasible point z^* satisfies MPEC-MFCQ, it also satisfies MPEC-ACQ.*

Proof. Since, by Lemma 3.5, MPEC-MFCQ implies MFCQ for every $\text{NLP}_*(\beta_1, \beta_2)$ with $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$, which in turn implies that Abadie CQ holds for every such $\text{NLP}_*(\beta_1, \beta_2)$ (see, e.g., [1]), we have, by Corollary 3.4, that MPEC-ACQ holds. □

The following example (taken from [11]) demonstrates that the reverse of Theorem 3.6 does not hold:

$$\begin{array}{ll} \min & f(z) := z_1 - z_2 \\ \text{s.t.} & g(z) := z_2 \leq 0, \\ & G(z) := z_1 \geq 0, \\ & H(z) := z_1 + z_2 \geq 0, \\ & G(z)^T H(z) = z_1(z_1 + z_2) = 0. \end{array}$$

The origin $z^* = (0, 0)$ is the unique minimizer. It is easily verified that

$$\mathcal{T}(z^*) = \mathcal{T}^{lin}_{\text{MPEC}}(z^*) = \{(d_1, d_2) \mid d_2 \leq 0, d_1 + d_2 = 0\},$$

showing that MPEC-ACQ holds in z^*, whereas MPEC-MFCQ does not hold in z^*.

The fact that $\mathcal{T}(z^*) = \mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ holds in the above example also follows from the following result.

Theorem 3.7 *Let z^* be a feasible point of the MPEC (1) and assume that the functions $g(\cdot)$, $h(\cdot)$, $G(\cdot)$, and $H(\cdot)$ are affine linear. Then MPEC-ACQ holds in z^*.*

Proof. By assumption, the functions making up the constraints of each $\text{NLP}_*(\beta_1, \beta_2)$ with $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$ are affine-linear. Hence Abadie CQ holds for all $\text{NLP}_*(\beta_1, \beta_2)$ (see, e.g., [1, Lemma 5.1.4]). By Corollary 3.4, MPEC-ACQ holds. □

Note that Theorem 3.7 demonstrates MPEC-ACQ to be a fairly weak constraint qualification.

In classical mathematical program the special case of a convex program is often considered. A constraint qualification closely linked with convex programs are the Slater and weak Slater constraint qualifications (referred to as SCQ and WSCQ respectively in the following), see, e.g., [7].

Since SCQ requires equality constraints to be linear, the MPEC (1) will never satisfy SCQ in any feasible point due to its complementarity term. In order to define a suitable variant of SCQ for MPECs, we must first formulate the subset of MPECs we want to consider.

Consider therefore the following program:

$$\begin{array}{lll} \min & f(z) & \\ \text{s.t.} & g(z) \leq 0, \quad h(z) = 0, & \\ & G(z) \geq 0, \quad H(z) \geq 0, \quad G(z)^T H(z) = 0, & \end{array} \tag{22a}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable and the component functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are convex. Furthermore, $h : \mathbb{R}^n \to \mathbb{R}^p$, $G : \mathbb{R}^n \to \mathbb{R}^l$, and $H : \mathbb{R}^n \to \mathbb{R}^l$ are affine linear, i.e. they take the following format:

$$\begin{aligned} h_i(z) &:= v_i^T z + \eta_i, && i = 1, \ldots, p, \\ G_i(z) &:= w_i^T z + \chi_i, && i = 1, \ldots, l, \\ H_i(z) &:= x_i^T z + \xi_i, && i = 1, \ldots, l. \end{aligned} \tag{22b}$$

In analogy to standard nonlinear programming, we call a program of this type *MPEC-convex*. Note, however, that (22) is *not* convex.

We shall now define an appropriate modification of the Slater constraint qualification.

Definition 3.8 *The program (22) is said to satisfy* weak MPEC-SCQ *or* MPEC-WSCQ *in a feasible vector z^* if there exists a vector $\hat{z}$ such that*

$$\begin{aligned} g_i(\hat{z}) &< 0, && \forall i \in \mathcal{I}_g, \\ h_i(\hat{z}) &= 0, && \forall i = 1, \ldots, p, \\ G_i(\hat{z}) &= 0, && \forall i \in \alpha \cup \beta, \\ H_i(\hat{z}) &= 0, && \forall i \in \gamma \cup \beta. \end{aligned} \tag{23}$$

It is said to satisfy MPEC-SCQ *if there exists a vector $\hat{z}$ such that*

$$\begin{aligned} g_i(\hat{z}) &< 0, && \forall i = 1, \ldots, m, \\ h_i(\hat{z}) &= 0, && \forall i = 1, \ldots, p, \\ G_i(\hat{z}) &= 0, && \forall i = 1, \ldots, l, \\ H_i(\hat{z}) &= 0, && \forall i = 1, \ldots, l. \end{aligned} \tag{24}$$

Note that, just as in standard nonlinear programming, the weak MPEC-Slater CQ is characterized by the fact that it depends on the vector z^* (through the sets $\mathcal{I}_g$, α, β, and γ). The appeal of the MPEC-Slater CQ is that it is stated independently of z^*. However, MPEC-SCQ is potentially much more restrictive than MPEC-WSCQ, since we require equality for $G_\gamma(\hat{z}) = 0$ and $H_\alpha(\hat{z}) = 0$. Note that MPEC-SCQ implies MPEC-WSCQ, analogeous to standard nonlinear programming.

We now show the weak MPEC-Slater CQ to imply MPEC-ACQ. Unfortunately it is not possible to reduce the proof of the following theorem to results from nonlinear programming and Coroallary 3.4 as was the case for Theorems 3.6 and 3.7. This is because neither MPEC-WSCQ, nor MPEC-SCQ implies that the Slater constraint qualification holds for any $\mathrm{NLP}_*(\beta_1, \beta_2)$. We therefore have to fall back on a more elementary proof.

Theorem 3.9 *Let z^* be a feasible point of the program (22). If it satisfies MPEC-WSCQ, it also satisfies MPEC-ACQ.*

Proof. By virtue of Lemma 3.2 we know that $\mathcal{T}(z^*) \subseteq \mathcal{T}^{lin}_{\text{MPEC}}(z^*)$. Therefore, all that remains to be shown is

$$\mathcal{T}^{lin}_{\text{MPEC}}(z^*) \subseteq \mathcal{T}(z^*). \tag{25}$$

We take the same path here as was taken in [5]. To this end, we define the following cone:

$$\begin{aligned}
\mathcal{T}^{strict}_{\text{MPEC}}(z^*) := \{d \in \mathbb{R}^n \mid & \nabla g_i(z^*)^T d < 0, && \forall i \in \mathcal{I}_g, \\
& \nabla h_i(z^*)^T d = 0, && \forall i = 1, \dots, p, \\
& \nabla G_i(z^*)^T d = 0, && \forall i \in \alpha, \\
& \nabla H_i(z^*)^T d = 0, && \forall i \in \gamma, \\
& \nabla G_i(z^*)^T d \geq 0, && \forall i \in \beta, \\
& \nabla H_i(z^*)^T d \geq 0, && \forall i \in \beta, \\
& (\nabla G_i(z^*)^T d) \cdot (\nabla H_i(z^*)^T d) = 0, && \forall i \in \beta \}.
\end{aligned} \tag{26}$$

Note that the difference between $\mathcal{T}^{strict}_{\text{MPEC}}(z^*)$ and $\mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ (see (14)) lies in the strict inequality for $g(\cdot)$.

Now, to prove (25), we show the following two inclusions:

$$\mathcal{T}^{lin}_{\text{MPEC}}(z^*) \subseteq \text{cl}(\mathcal{T}^{strict}_{\text{MPEC}}(z^*)) \subseteq \mathcal{T}(z^*). \tag{27}$$

To prove the first inclusion of (27), we take a vecor $\hat{z}$ satisfying the Slater conditions (23) and set

$$\hat{d} := \hat{z} - z^*.$$

Now the following holds for all $i \in \mathcal{I}_g$ because $g(\cdot)$ is convex by assumption:

$$\nabla g_i(z^*)^T \hat{d} \leq \underbrace{g_i(\hat{z})}_{<0} - \underbrace{g_i(z^*)}_{=0} < 0, \qquad \forall i \in \mathcal{I}_g. \tag{28}$$

Similarly, the following hold for the linear functions $h(\cdot)$, $G(\cdot)$, and $H(\cdot)$:

$$\begin{aligned}
\nabla h_i(z^*)^T \hat{d} &= v_i^T \hat{z} - v_i^T z^* = \underbrace{h_i(\hat{z})}_{=0} - \underbrace{h_i(z^*)}_{=0} = 0, && \forall i = 1, \dots, p, \\
\nabla G_i(z^*)^T \hat{d} &= w_i^T \hat{z} - w_i^T z^* = \underbrace{G_i(\hat{z})}_{=0} - \underbrace{G_i(z^*)}_{=0} = 0, && \forall i \in \alpha \cup \beta, \\
\nabla H_i(z^*)^T \hat{d} &= x_i^T \hat{z} - x_i^T z^* = \underbrace{H_i(\hat{z})}_{=0} - \underbrace{H_i(z^*)}_{=0} = 0, && \forall i \in \gamma \cup \beta.
\end{aligned} \tag{29}$$

We now use the vector $\hat{d}$, which we know has the properties (28) and (29), to define the following function:

$$d(\delta) := d + \delta \hat{d},$$

where $d \in \mathcal{T}^{lin}_{\text{MPEC}}(z^*)$ is chosen arbitrarily.

We want to show that $d(\delta) \in \mathcal{T}^{strict}_{\text{MPEC}}(z^*)$ for all $\delta > 0$. To this end, let $\delta > 0$ be fixed for the time being. Then the following holds:

$$\nabla g_i(z^*)^T d(\delta) = \underbrace{\nabla g_i(z^*)^T d}_{\leq 0} + \delta \underbrace{\nabla g_i(z^*)^T \hat{d}}_{<0} < 0, \qquad \forall i \in \mathcal{I}_g,$$

$$\nabla h_i(z^*)^T d(\delta) = \nabla h_i(z^*)^T d + \delta \nabla h_i(z^*)^T \hat{d} = 0, \qquad \forall i = 1, \ldots, p,$$

$$\nabla G_i(z^*)^T d(\delta) = \begin{cases} \underbrace{\nabla G_i(z^*)^T d}_{=0} + \delta \underbrace{\nabla G_i(z^*)^T \hat{d}}_{=0} = 0, & \forall i \in \alpha, \\ \underbrace{\nabla G_i(z^*)^T d}_{\geq 0} + \delta \underbrace{\nabla G_i(z^*)^T \hat{d}}_{=0} \geq 0, & \forall i \in \beta, \end{cases}$$

$$\nabla H_i(z^*)^T d(\delta) = \begin{cases} \underbrace{\nabla H_i(z^*)^T d}_{=0} + \delta \underbrace{\nabla H_i(z^*)^T \hat{d}}_{=0} = 0, & \forall i \in \gamma, \\ \underbrace{\nabla H_i(z^*)^T d}_{\geq 0} + \delta \underbrace{\nabla H_i(z^*)^T \hat{d}}_{=0} \geq 0, & \forall i \in \beta. \end{cases}$$

Furthermore, for all $i \in \beta$ it holds that

$$\begin{aligned} (\nabla G_i(z^*)^T d(\delta)) \cdot (\nabla H_i(z^*)^T d(\delta)) = & \underbrace{(\nabla G_i(z^*)^T d) \cdot (\nabla H_i(z^*)^T d)}_{=0} \\ & + \delta (\nabla G_i(z^*)^T d) \cdot \underbrace{(\nabla H_i(z^*)^T \hat{d})}_{=0} \\ & + \delta \underbrace{(\nabla G_i(z^*)^T \hat{d})}_{=0} \cdot (\nabla H_i(z^*)^T d) \\ & + \delta^2 \underbrace{(\nabla G_i(z^*)^T \hat{d})}_{=0} \cdot \underbrace{(\nabla H_i(z^*)^T \hat{d})}_{=0} \\ = & \ 0. \end{aligned}$$

Comparing the properties of $d(\delta)$ to $\mathcal{T}^{strict}_{\text{MPEC}}(z^*)$ (see (26)), we see that $d(\delta) \in \mathcal{T}^{strict}_{\text{MPEC}}(z^*)$ for all $\delta > 0$. Since $\text{cl}(\mathcal{T}^{strict}_{\text{MPEC}}(z^*))$ is closed by definition, it holds that $d = \lim_{\delta \searrow 0} d(\delta) \in \text{cl}(\mathcal{T}^{strict}_{\text{MPEC}}(z^*))$. Hence, the inclusion

$$\mathcal{T}^{lin}_{\text{MPEC}}(z^*) \subseteq \text{cl}(\mathcal{T}^{strict}_{\text{MPEC}}(z^*)) \tag{30}$$

holds.

To prove the second inclusion of (27), let $d \in \mathcal{T}^{strict}_{\text{MPEC}}(z^*)$ and $\{t_k\}$ be a sequence with $t_k \searrow 0$. Setting

$$z^k := z^* + t_k d,$$

we have

$$z^k \to z^* \quad \text{and} \quad \frac{z^k - z^*}{t_k} \to d.$$

If we can prove that $\{z^k\} \subset \mathcal{Z}$, it would follow that $d \in \mathcal{T}(z^*)$ (see (9)) and we would have shown $\mathcal{T}_{\text{MPEC}}^{strict}(z^*) \subseteq \mathcal{T}(z^*)$, and, since $\mathcal{T}(z^*)$ is closed, also that $\text{cl}(\mathcal{T}_{\text{MPEC}}^{strict}(z^*)) \subseteq \mathcal{T}(z^*)$. Let us therefore check whether $z^k \in \mathcal{Z}$. By the mean value theorem, it holds that there exists a vector ζ on the connecting line between z^k and z^* such that the following holds:

$$\begin{aligned} g_i(z^k) &= g_i(z^*) + \nabla g_i(\zeta)^T(z^k - z^*) \\ &= g_i(z^*) + t_k \nabla g_i(\zeta)^T d \\ &= \left\{ \begin{array}{ll} \underbrace{g_i(z^*)}_{=0} + t_k \underbrace{\nabla g_i(\zeta)^T d}_{<0,\, \forall k > k_0}, & \forall i \in \mathcal{I}_g \\ \underbrace{g_i(z^*)}_{<0} + t_k \nabla g_i(\zeta)^T d, & \forall i \notin \mathcal{I}_g \end{array} \right\} \le 0, \qquad \forall k > k_1, \end{aligned}$$

where $k_0 \ge 0$ and $k_1 \ge k_0$ are sufficiently large integers. Note that the convexity of $g(\cdot)$ does not enter here.

The linear functions are handled similarly:

$$h_i(z^k) = v_i^T z^k + \eta_i = \underbrace{v_i^T z^* + \eta_i}_{=0} + t_k \underbrace{v_i^T d}_{=0} = 0, \qquad \forall i = 1, \dots, p,$$

$$G_i(z^k) = w_i^T z^k + \chi_i = \left\{ \begin{array}{ll} \underbrace{w_i^T z^* + \chi_i}_{=0} + t_k \underbrace{w_i^T d}_{=0}, & \forall i \in \alpha \\ \underbrace{w_i^T z^* + \chi_i}_{=0} + t_k \underbrace{w_i^T d}_{\ge 0}, & \forall i \in \beta \\ \underbrace{w_i^T z^* + \chi_i}_{>0} + t_k w_i^T d, & \forall i \in \gamma \end{array} \right\} \ge 0, \qquad \forall k > k_2,$$

$$H_i(z^k) = x_i^T z^k + \xi_i = \left\{ \begin{array}{ll} \underbrace{x_i^T z^* + \xi_i}_{>0} + t_k x_i^T d, & \forall i \in \alpha \\ \underbrace{x_i^T z^* + \xi_i}_{=0} + t_k \underbrace{x_i^T d}_{\ge 0}, & \forall i \in \beta \\ \underbrace{x_i^T z^* + \xi_i}_{=0} + t_k \underbrace{x_i^T d}_{=0}, & \forall i \in \gamma \end{array} \right\} \ge 0, \qquad \forall k > k_3,$$

with $k_2 \ge 0$ and $k_3 \ge 0$ sufficiently large integers.

Finally, taking into consideration the definition of $\mathcal{T}_{\text{MPEC}}^{strict}(z^*)$ (see (26)), the following holds for the product $G_i(z^k) \cdot H_i(z^k)$:

$$G_i(z^k) \cdot H_i(z^k) = \left\{ \begin{array}{ll} (\underbrace{w_i^T z^* + \chi_i}_{=0} + t_k \underbrace{w_i^T d}_{=0})(x_i^T z^* + \xi_i + t_k x_i^T d) & \forall i \in \alpha \\ (\underbrace{w_i^T z^* + \chi_i}_{=0} + t_k w_i^T d)(\underbrace{x_i^T z^* + \xi_i}_{=0} + t_k x_i^T d) = t_k^2 \underbrace{(w_i^T d)(x_i^T d)}_{=0} & \forall i \in \beta \\ (w_i^T z^* + \chi_i + t_k w_i^T d)(\underbrace{x_i^T z^* + \xi_i}_{=0} + t_k \underbrace{x_i^T d}_{=0}) & \forall i \in \gamma \end{array} \right\} = 0.$$

The above results demonstrate that $z^k \in \mathcal{Z}$ for all $k \geq \max\{k_0, k_1, k_2, k_3\}$ and hence $d \in \mathcal{T}(z^*)$. As was already mentioned, $\mathcal{T}(z^*)$ is closed, yielding that $\mathrm{cl}(\mathcal{T}^{strict}_{\mathrm{MPEC}}(z^*)) \subseteq \mathcal{T}(z^*)$.

Together with (30), we therefore have $\mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*) \subseteq \mathcal{T}(z^*)$ and hence MPEC-ACQ is satisfied, concluding this proof. □

The following corollary follows immediately from Theorem 3.9 and Definition 3.8.

Corollary 3.10 *Let z^* be a feasible point of the program (22). If it satisfies MPEC-SCQ, it also satisfies MPEC-ACQ.*

3.3 First Order Optimality Conditions

The next question is, of course, which first order optimality conditions can be obtained using MPEC-ACQ. We cannot expect strong stationarity since MPEC-ACQ is implied by MPEC-MFCQ which, in turn, is not sufficient for strong stationarity (see [13, 3]). However, our next result shows that MPEC-ACQ does imply a weaker stationarity concept.

Theorem 3.11 *Let $z^* \in \mathbb{R}^n$ be a local minimizer of the MPEC (1). If MPEC-ACQ holds in z^*, then there exists a Lagrange multiplier λ^* such that (z^*, λ^*) satisfies the following stationarity conditions:*

$$
\begin{aligned}
&0 = \nabla f(z^*) + \sum_{i=1}^{m} (\lambda_i^g)^* \nabla g_i(z^*) + \sum_{i=1}^{p} (\lambda_i^h)^* \nabla h_i(z^*) - \sum_{i=1}^{l} \left[(\lambda_i^G)^* \nabla G_i(z^*) + (\lambda_i^H)^* H_i(z^*) \right], \\
&\begin{matrix} (\lambda_\alpha^G)^* \quad \text{free}, \\ (\lambda_\gamma^H)^* \quad \text{free}, \end{matrix} \qquad (\lambda_i^G)^* \geq 0 \vee (\lambda_i^H)^* \geq 0 \quad \forall i \in \beta \qquad \begin{matrix} (\lambda_\gamma^G)^* = 0, \\ (\lambda_\alpha^H)^* = 0, \end{matrix} \\
&g(z^*) \leq 0, \qquad (\lambda^g)^* \geq 0, \qquad g(z^*)^T (\lambda^g)^* = 0.
\end{aligned}
\tag{31}
$$

In particular, (z^*, λ^*) is weakly stationary.

Proof. It is known from nonlinear programming (cf., e.g., [5]) that if z^* is a local minimizer of the MPEC (1), then it is B-stationary, i.e. the following holds:

$$
\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}(z^*). \tag{32}
$$

Since MPEC-ACQ holds, (32) is equivalent to the following:

$$
\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*) = \bigcup_{(\beta_1, \beta_2) \in \mathcal{P}(\beta)} \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1, \beta_2)}(z^*), \tag{33}
$$

cf. Lemma 3.1 (b).

Choosing a partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$ arbitrarily, it follows from (33) that

$$\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(z^*). \tag{34}$$

We now follow standard arguments from nonlinear programming. The condition (34) can also be written as

$$-\nabla f(z^*)^T d \leq 0 \qquad \forall d \in \mathbb{R}^n \text{ with } Ad \leq 0,$$

where the matrix $A \in \mathbb{R}^{(|\mathcal{I}_g|+2p+2|\alpha|+2|\gamma|+3|\beta|)\times n}$ is given by

$$A := \begin{pmatrix} \nabla g_i(z^*)^T & (i \in \mathcal{I}_g) \\ \nabla h_i(z^*)^T & (i = 1, \dots, p) \\ -\nabla h_i(z^*)^T & (i = 1, \dots, p) \\ \nabla G_i(z^*)^T & (i \in \alpha \cup \beta_1) \\ -\nabla G_i(z^*)^T & (i \in \alpha \cup \beta_1) \\ \nabla H_i(z^*)^T & (i \in \gamma \cup \beta_2) \\ -\nabla H_i(z^*)^T & (i \in \gamma \cup \beta_2) \\ -\nabla G_i(z^*)^T & (i \in \beta_2) \\ -\nabla H_i(z^*)^T & (i \in \beta_1) \end{pmatrix}.$$

Farkas' theorem of the alternative (cf., e.g., [7, Theorem 2.4.6]) yields that

$$A^T y = -\nabla f(z^*), \qquad y \geq 0$$

has a solution. Now let us denote the components of y by $\lambda^g_{\mathcal{I}_g}$, λ^{h+}, λ^{h-} ($\lambda^{h+}, \lambda^{h-} \in \mathbb{R}^p$), $\lambda^{G+}_{\alpha\cup\beta_1}$, $\lambda^{G-}_{\alpha\cup\beta_1}$, $\lambda^{H+}_{\gamma\cup\beta_2}$, $\lambda^{H-}_{\gamma\cup\beta_2}$, $\lambda^G_{\beta_2}$, and $\lambda^H_{\beta_1}$, in that order. We then set $\lambda^h := \lambda^{h+} - \lambda^{h-}$, $\lambda^G_{\alpha\cup\beta_1} := \lambda^{G+}_{\alpha\cup\beta_1} - \lambda^{G-}_{\alpha\cup\beta_1}$, and $\lambda^H_{\gamma\cup\beta_2} := \lambda^{H+}_{\gamma\cup\beta_2} - \lambda^{H-}_{\gamma\cup\beta_2}$. Additionally, we set $\lambda^g_i := 0$ ($i \notin \mathcal{I}_g$), $\lambda^G_\gamma := 0$, and $\lambda^H_\alpha := 0$. The resulting vector $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ satisfies the conditions for A-stationarity. Setting $\lambda^* := \lambda$ completes the proof. □

Note that the proof of Theorem 3.11 holds for an arbitrary partition (β_1, β_2) of the index set β. Hence we can choose, a priori, such a partition and obtain corresponding Lagrange multipliers $(\lambda^G)^*$ and $(\lambda^H)^*$ such that $(\lambda^G_i)^* \geq 0$ for all $i \in \beta_1$ and $(\lambda^H_i)^* \geq 0$ for all $i \in \beta_2$.

Motivated by Theorem 3.11, we call a weakly stationary point z^* of the MPEC (1) *A-stationary* if there exists a corresponding Lagrange multiplier λ^* such that

$$(\lambda^G_i)^* \geq 0 \quad \text{or} \quad (\lambda^H_i)^* \geq 0 \quad \forall i \in \beta,$$

i.e., z^* is A-stationary if and only if (31) holds for some multiplier λ^*. Here, the letter 'A' may stand for 'alternative' since, for each $i \in \beta$, we have the alternative that either $(\lambda^G_i)^* \geq 0$ or $(\lambda^H_i)^* \geq 0$ (or both) hold. However, the letter 'A' may also be interpreted as an abbreviation for 'Abadie' since this stationarity concept holds under MPEC-Abadie CQ.

The condition (33), i.e.

$$\nabla f(z^*)^T d \geq 0 \qquad \forall d \in \mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*), \tag{35}$$

has appeared before in [13] under the name "B-stationarity". This is somewhat misleading, since more commonly, (32) is called B-stationarity (cf., e.g., [11]). We shall therefore call (35) *MPEC-linearized B-stationarity* to avoid confusion. The name is motivated by the fact that (35) involves the MPEC-linearized cone $\mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*)$.

Note that the proof of Theorem 3.11 yields that MPEC-linearized B-stationarity implies A-stationarity.

Also note that under both MPEC-LICQ and MPEC-SMFCQ, A-stationarity implies strong stationarity. This is easily verified if one considers that both these constraint qualifications imply the uniqueness of the Lagrange multiplier and the fact that the partition (β_1, β_2) for A-stationarity should hold can be chosen a priori. See [3] for more details, as well as a different derivation of A-stationarity.

4 Comparison between the basic CQ [6] and MPEC-ACQ

The question was raised by Danny Ralph [12] whether MPEC-ACQ may be identical to the *basic CQ* used in [6]. In this section we try to shed some light on this. It may be noted in advance that in the absense of upper-level constraints and if the lower level constraints take on a certain form, then the basic CQ and MPEC-ACQ do indeed coincide. To see this, however, some effort is required. The remainder of this section is dedicated to that question.

Let us therefore consider an MPEC in the notation used in [6]:

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & (x,y) \in Z \subset \mathbb{R}^{n+m}, \\ & y \in \mathcal{S} := \mathrm{SOL}(F(x,\cdot), C(x)) \end{array} \tag{36}$$

with

$$C(x) := \{y \in \mathbb{R}^m \mid g(x,y) \leq 0\}. \tag{37}$$

Here $y \in \mathrm{SOL}(F(x,\cdot), C(x))$ denotes a solution of the variational inequality

$$\begin{array}{l} y \in C(x), \\ (v-y)^T F(x,y) \geq 0, \qquad \forall v \in C(x). \end{array} \tag{38}$$

Since we are interested in mathematical programs with nonlinear complementarity problems as constraints rather than variational inequalities, we consider the case where $g(x,y) :=$

$-y$. Then the program (36) reduces to the following program:

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & (x,y) \in Z, \\ & y \geq 0, \\ & (v-y)^T F(x,y) \geq 0, \qquad \forall v \geq 0, \end{array}$$

which obviously is equivalent to the following program:

$$\begin{array}{ll} \min & f(x,y) \\ \text{s.t.} & (x,y) \in Z, \\ & y \geq 0, \quad F(x,y) \geq 0, \quad y^T F(x,y) = 0. \end{array} \tag{39}$$

Now, the *basic CQ* from [6] is said to hold in (x^*, y^*) if there exists a nonempty set M' of Lagrange multipliers of the variational inequality (38) such that

$$\mathcal{T}((x^*,y^*);\mathcal{F}) = \mathcal{T}((x^*,y^*);Z) \cap \left(\bigcup_{\lambda \in M'} \mathrm{Gr}(\mathcal{LS}_{(x^*,y^*,\lambda)}) \right), \tag{40}$$

where $\mathcal{F}$ denotes the feasible region of (39), $\mathrm{Gr}(\cdot)$ denotes the graph of a multifunction and

$$\mathcal{LS}_{(x^*,y^*,\lambda)}(dx) = \mathrm{SOL}(\nabla_x L(x^*,y^*,\lambda)dx, \nabla_y L(x^*,y^*,\lambda), \mathcal{K}(x^*,y^*,\lambda;dx)) \tag{41}$$

is the solution set of the affine variational inequality

$$\begin{array}{l} dy \in \mathcal{K}(x^*,y^*,\lambda;dx), \\ (v-dy)^T(\nabla_x L(x^*,y^*,\lambda)dx + \nabla_y L(x^*,y^*,\lambda)dy) \geq 0, \quad \forall v \in \mathcal{K}(x^*,y^*,\lambda;dx), \end{array} \tag{42}$$

where

$$L(x,y,\lambda) := F(x,y) + \sum_{i=1}^{l} \lambda_i \nabla_y g_i(x,y) = F(x,y) - \lambda \tag{43}$$

is the Lagrangian,

$$\mathcal{K}(x^*,y^*,\lambda;dx) := \{dy \in \mathbb{R}^m \mid (dx,dy) \in \mathcal{K}(x^*,y^*,\lambda)\} \tag{44}$$

is the *directional critical set*, and

$$\begin{aligned} \mathcal{K}(x^*,y^*,\lambda) := \big\{ &(dx,dy) \in \mathbb{R}^{n+m} \mid \\ &\nabla_x g_i(x^*,y^*)^T dx + \nabla_y g_i(x^*,y^*)^T dy \leq 0, \quad \forall i \in \{i \mid g_i(x^*,y^*) = 0 \wedge \lambda_i = 0\}, \\ &\nabla_x g_i(x^*,y^*)^T dx + \nabla_y g_i(x^*,y^*)^T dy = 0, \quad \forall i \in \{i \mid g_i(x^*,y^*) = 0 \wedge \lambda_i > 0\} \big\} \\ = \big\{ &(dx,dy) \in \mathbb{R}^{n+m} \mid \\ &e_i^T dy \geq 0, \quad \forall i \in \{i \mid y_i^* = 0 \wedge \lambda_i = 0\}, \\ &e_i^T dy = 0, \quad \forall i \in \{i \mid \lambda_i > 0\} \big\} \end{aligned} \tag{45}$$

is the *lifted critical cone* ($e_i \in \mathbb{R}^m$ is the i-th unit vector).

If, in (40), M' is assumed to be the whole set of Lagrange multipliers, the *full CQ* is said to hold.

The Lagrange multipliers λ of the variational inequality (38) must satisfy the following conditions:

$$\begin{aligned} &L(x, y, \lambda) = F(x, y) - \lambda = 0, \\ &\lambda \geq 0, \quad y \geq 0, \quad \lambda^T y = 0. \end{aligned} \tag{46}$$

Hence $\lambda^* := F(x^*, y^*)$ is the only Lagrange multiplier and $M' := \{\lambda^*\}$ reduces to a singleton, so that the basic and full CQ coincide. Note also that, because of the complementarity conditions in (46), $\lambda_i > 0$ implies $g_i(x^*, y^*) = -y_i^* = 0$ in (45).

Let us again consider the basic CQ (40). Note that it is somewhat different from our MPEC-ACQ since we also linearize the upper level constraints. We void this difference in our subsequent discussion by setting $Z := \mathbb{R}^{n+m}$. Combining this with the fact that M' reduces to a singleton for our program (39) results in the basic (or full) CQ reducing to

$$\mathcal{T}((x^*, y^*); \mathcal{F}) = \mathrm{Gr}(\mathcal{LS}_{(x^*, y^*, \lambda^*)}). \tag{47}$$

Now, by virtue of [6, (1.3.14)], $dy \in \mathcal{LS}_{(x^*, y^*, \lambda^*)}(dx)$ holds if and only if there exists a multiplier μ such that

$$\begin{aligned} \nabla_x L(x^*, y^*, \lambda^*)dx + \nabla_y L(x^*, y^*, \lambda^*)dy - A^T\mu &= 0, \\ Ady - b &\geq 0, \\ \mu &\geq 0, \\ \mu^T(Ady - b) &= 0, \end{aligned} \tag{48}$$

with

$$A := \begin{pmatrix} e_i^T & i \in \{i \mid y_i^* = 0 \wedge \lambda_i^* = 0\} \\ -e_i^T & i \in \{i \mid \lambda_i^* > 0\} \\ e_i^T & \end{pmatrix}, \qquad b := 0.$$

Combining $z := (x, y)$ and $d = (dx, dy)$, and remembering that $\lambda^* = F(x^*, y^*)$, the set $\mathrm{Gr}(\mathcal{LS}_{(x^*, y^*, \lambda^*)})$ can be written as follows:

$$\begin{aligned} \mathrm{Gr}(\mathcal{LS}_{(z^*, \lambda^*)}) = \{d \in \mathbb{R}^{n+m} \mid \exists \mu \quad : \quad & \nabla F(z^*)^T d - A^T\mu = 0, && \text{(49a)} \\ & \mu \geq 0, \quad \mu^T A dy = 0, && \text{(49b)} \\ & e_i^T dy \geq 0, \quad i \in \beta && \text{(49c)} \\ & e_i^T dy = 0, \quad i \in \gamma\}, && \text{(49d)} \end{aligned}$$

where β and γ (and α, see below) are defined as in (2), for the program (39).

The conditions (49c) and (49d) reduce to $d_{n+i} \geq 0$ for $i \in \beta$ and $d_{n+i} = 0$ for $i \in \gamma$ respectively. From (49a) it follows that $\nabla F_i(z^*)^T d = 0$ for $i \in \alpha$, and $\nabla F_i(z^*)^T d = \mu_{j(i)} \geq 0$ for $i \in \beta$, where the index j depends on i.

Let us now consider (49b):

$$\begin{aligned}
\mu^T A dy &= \mu^T \begin{pmatrix} d_{n+i} & i \in \beta \\ \overbrace{-d_{n+i}}^{=0} & \\ \underbrace{d_{n+i}}_{=0} & i \in \gamma \end{pmatrix} \\
&= \sum_{i \in \beta} \mu_{j(i)} d_{n+i} \\
&= \sum_{i \in \beta} \underbrace{(\nabla F_i(z^*)^T d)}_{\geq 0} \underbrace{d_{n+i}}_{\geq 0} = 0.
\end{aligned}$$

From this, it follows that

$$(\nabla F_i(z^*)^T d) \cdot d_{n+i} = 0, \qquad \forall i \in \beta.$$

Collecting everything we have gathered about $\mathrm{Gr}(\mathcal{LS}_{(z^*,\lambda^*)})$, we arrive at the following result:

$$\begin{aligned}
\mathrm{Gr}(\mathcal{LS}_{(z^*,\lambda^*)}) = \{d \in \mathbb{R}^{n+m} \mid\ & \nabla F_i(z^*)^T d = 0, && \forall i \in \alpha, \\
& d_{n+i} = 0, && \forall i \in \gamma, \\
& \nabla F_i(z^*)^T d \geq 0, && \forall i \in \beta, \\
& d_{n+i} \geq 0, && \forall i \in \beta, \\
& (\nabla F_i(z^*)^T d) \cdot d_{n+i} = 0, && \forall i \in \beta\},
\end{aligned}$$

which is equal to $\mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*)$ of the MPEC (39). Hence the basic CQ [6] and MPEC-ACQ coincide in this case. Note however, that we considered a program in the format (39) and in the absense of upper-level constraints.

5 Summary

If we collect the constraint qualifications we have covered in this paper and put them in relation to each other, we get the following chain of implications for any feasible point z^*:

$$\begin{array}{ccccccc}
 & & & & \text{affine MPEC} & \searrow & \\
\text{MPEC-LICQ} & \Longrightarrow & \text{MPEC-SMFCQ} & \Longrightarrow & \text{MPEC-MFCQ} & \Longrightarrow & \text{MPEC-ACQ.} \\
 & & \text{MPEC-SCQ} & \Longrightarrow & \text{MPEC-WSCQ} & \nearrow &
\end{array}$$

The implications MPEC-LICQ $\Rightarrow$ MPEC-SMFCQ $\Rightarrow$ MPEC-MFCQ transfer directly from nonlinear programming since the constraint qualifications involved are defined via TNLP, which is a nonlinear program (see Definition 2.1). The other implications were proved in Section 3 (see Theorems 3.6, 3.7, and 3.9 and Definition 3.8).

Of perhaps even greater interest is the relationship between the stationarity concepts mentioned in this paper. To this end, let us recall the various stationarity concepts: A feasible vector z^* of the MPEC (1) is called

- *B-stationary* if $\nabla f(z^*)^T d \geq 0 \quad \forall d \in \mathcal{T}(z^*)$;
- *MPEC-linearized B-stationary* if $\nabla f(z^*)^T d \geq 0 \quad \forall d \in \mathcal{T}^{lin}_{\mathrm{MPEC}}(z^*)$;
- *linearized B-stationary* if $\nabla f(z^*)^T d \geq 0 \quad \forall d \in \mathcal{T}^{lin}(z^*)$.

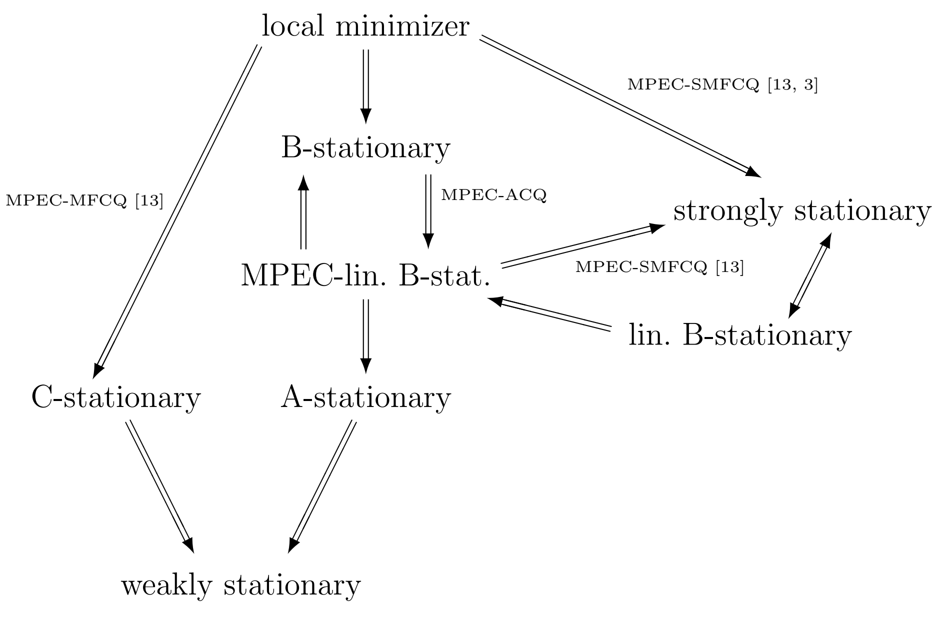
The notion of linearized B-stationarity is new and included here only for the sake of completeness. In view of the inclusions (19), we have

$$\text{lin. B-stationarity} \implies \text{MPEC-lin. B-stationarity} \implies \text{B-stationarity.}$$

We further recall that a weakly stationary point z^* (see (6)) with corresponding Lagrange multiplier λ^* is called

- *A-stationary* if $(\lambda_i^G)^* \geq 0$ or $(\lambda_i^H)^* \geq 0$ for all $i \in \beta$;
- *C-stationary* if $(\lambda_i^G)^*(\lambda_i^H)^* \geq 0$ for all $i \in \beta$ (see [13]);
- *strongly stationary* if $(\lambda_i^G)^* \geq 0$ and $(\lambda_i^H)^* \geq 0$ for all $i \in \beta$.

The following graph now attempts to clarify the relationship between the various stationarity concepts.



The implications

$$\text{local minimizer} \overset{\text{MPEC-MFCQ}}{\Longrightarrow} \text{C-stationary},$$

$$\text{MPEC-linearized B-stationary} \overset{\text{MPEC-SMFCQ}}{\Longrightarrow} \text{strongly stationary},$$

and

$$\text{local minimizer} \overset{\text{MPEC-SMFCQ}}{\Longrightarrow} \text{strongly stationary}$$

have been shown in [13] and do not follow from our analysis, whereas the equivalence

$$\text{strongly stationary} \iff \text{linearized B-stationary}$$

follows by the following arguments: For the forward implication, multiply the first line of (7) by d with $d \in \mathcal{T}^{lin}(z^*)$:

$$\begin{aligned}
\nabla f(z^*)^T d \;\; = \;\; & -\sum_{i \in \mathcal{I}_g} \underbrace{\lambda_i^g}_{\geq 0} \underbrace{\nabla g_i(z^*)^T d}_{\leq 0} - \sum_{i=1}^{p} \lambda_i^h \underbrace{\nabla h_i(z^*)^T d}_{=0} \\
& + \sum_{i \in \alpha} \lambda_i^G \underbrace{\nabla G_i(z^*)^T d}_{=0} + \sum_{i \in \gamma} \lambda_i^H \underbrace{\nabla H_i(z^*)^T d}_{=0} \\
& + \sum_{i \in \beta} \Big[\underbrace{\lambda_i^G}_{\geq 0} \underbrace{\nabla G_i(z^*)^T d}_{\geq 0} + \underbrace{\lambda_i^H}_{\geq 0} \underbrace{\nabla H_i(z^*)^T d}_{\geq 0} \Big] \;\geq\; 0.
\end{aligned}$$

For the reverse implication, we express linearized B-stationarity as

$$-\nabla f(z^*)^T d \leq 0 \qquad \forall d \text{ with } Ad \leq 0$$

with an appropriare matrix A. Farkas' theorem of the alternative (cf., e.g., [7, Theorem 2.4.6]) yields the existence of a $y \geq 0$, such that

$$A^T y = -\nabla f(z^*).$$

From this, it is immediately apparent that multipliers satisfying strong stationarity (7) exist. (Compare also to the proof of Theorem 3.11, where the same technique was used.)

All other implications follow from our results in Section 3.

The implication

$$\text{strongly stationary} \Longrightarrow \text{C-stationary}$$

was not included in the graph for esthetic reasons, although it does, of course, hold.

References

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, second ed., 1993.

[2] Y. Chen and M. Florian, *The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions*, Optimization, 32 (1995), pp. 193–209.

[3] M. L. Flegel and C. Kanzow, *A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints.* Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint, October 2002.

[4] ——, *On the Guignard constraint qualification for mathematical programs with equilibrium constraints.* Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint 248, October 2002.

[5] C. Geiger and C. Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer-Verlag, Berlin · Heidelberg, 2002.

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

[7] O. L. Mangasarian, *Nonlinear Programming*, Classics in Applied Mathematics, SIAM, Philadelphia, Classics ed., 1994. Identical to the 1969 book published by McGraw-Hill.

[8] J. V. Outrata, *Optimality conditions for a class of mathematical programs with equilibrium constraints*, Mathematics of Operations Research, 24 (1999), pp. 627–644.

[9] ——, *A generalized mathematical program with equilibrium constraints*, SIAM Jounral of Control and Optimization, 38 (2000), pp. 1623–1638.

[10] J. V. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[11] J.-S. Pang and M. Fukushima, *Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints*, Computational Optimization and Applications, 13 (1999), pp. 111–136.

[12] D. Ralph, *Personal communication.* Third International Conference on Complementarity Problems, Cambridge, England, July 2002.

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

[14] J. J. Ye and D. L. Zhu, *Optimality conditions for bilevel programming problems*, Optimization, 33 (1995), pp. 9–27.