OPTIMALITY CONDITIONS FOR DISJUNCTIVE PROGRAMS WITH APPLICATION TO MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

Michael L. Flegel[1], Christian Kanzow[2], and Jiří V. Outrata[3]

Preprint 256 October 2004

[1,2] 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

[3] Academy of Sciences
Institute of Information Theory and Automation
18208 Prague
Czech Republic

e-mail: outrata@utia.cas.cz

October 29, 2004

Abstract. We consider optimization problems with a disjunctive structure of the feasible set. Using Guignard-type constraint qualifications for these optimization problems and exploiting some results for the limiting normal cone by Mordukhovich, we derive different optimality conditions. Furthermore, we specialize these results to mathematical programs with equilibrium constraints. In particular, we show that a new constraint qualification, weaker than any other constraint qualification used in the literature, is enough in order to show that a local minimum results in a so-called M-stationary point. Additional assumptions are also discussed which guarantee that such an M-stationary point is in fact a strongly stationary point.

Key Words. disjunctive programs, mathematical programs with equilibrium constraints, M-stationarity, strong stationarity, Guignard constraint qualification.

1 Introduction

In the past decade applied mathematicians have been paying increasing interest to optimization problems where, among the constraints, so-called equilibrium constraints occur. This equilibrium is described mostly by a (lower-level) optimization problem, a variational inequality or a complementarity problem. Following [10], all these problems are currently termed *mathematical programs with equilibrium constraints*, or *MPECs*. Most of them can be written down in form of (smooth) nonlinear programs which, however, do not satisfy most of the standard constraint qualifications (CQs). This has lead to several weakened stationarity notions that have been introduced in connection with optimality conditions and numerical approaches.

Among these new stationarity notions an important role is played by a concept associated with the generalized differential calculus of Mordukhovich, cf. [26, 13, 24, 14]. Following [19], we will call it *M-stationarity*. The advantage of this concept consists, above all, in the fact that it requires only very weak constraint qualifications, cf. [4, 25]. Simultaneously, starting with [10], another, stronger, stationarity notion has been investigated ([23, 15]), referred to by the moniker *strong stationarity* in [19]. Strong stationarity is the natural candidate for a stationarity concept for MPECs, because it is equivalent to the standard KKT conditions if the MPEC can be expressed as a standard nonlinear program. As mentioned above, however, this stationarity concept is too restrictive in the context of MPECs.

In this paper we will investigate the above mentioned stationarity notions by means of a general disjunctive program which embeds most MPECs considered in the literature so far. In particular, we will show that local minimizers are M-stationary under an appropriate variant of the Guignard CQ [7], known to be the weakest CQ used in classical nonlinear programming. Moreover, we will derive a new condition which, together with this variant of Guignard CQ, ensures the strong stationarity of local minimizers.

Finally, these results will then be translated to the language of MPECs where the constraints are described by a complementarity problem. In particular, we will show that M- and strong stationarity indeed amount to the well-known conditions common in the literature of these programs. Additionally, a new constraint qualification is introduced to such MPECs by transferring the variant of Guignard CQ for our disjunctive programs to an MPEC setting. This yields the weakest result known to date, M-stationarity as a first order condition under this very weak CQ.

The organization of the paper is as follows. Section 2 is devoted to some preliminary results including important definitions and some initial observations. The main results for the disjunctive program are collected in Section 3. Finally, in Section 4, we apply the results of Section 3 to a standard MPEC, where the equilibrium is modelled by a nonlinear complementarity problem. We then close with some final remarks in Section 5.

Our notation is standard. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$, its nonnegative and nonpositive orthant by $\mathbb{R}^n_+$ and $\mathbb{R}^n_-$, respectively. If unclear from context, the size of a zero vector is denoted by an appropriate subscript: 0_n is the zero vector in $\mathbb{R}^n$. Given a vector $x \in \mathbb{R}^n$, its components are denoted by x_i. Furthermore, given an index set $\omega \subseteq \{1, \dots, n\}$, we denote by $x_\omega \in \mathbb{R}^{|\omega|}$ the vector consisting of those components x_i corresponding to the indices in ω. The same terminology applies to functions. Given a function $F : \mathbb{R}^n \to \mathbb{R}^m$, we denote its Jacobian by $\nabla F(z) \in \mathbb{R}^{m \times n}$. If $m = 1$, the gradient $\nabla F(z) \in \mathbb{R}^n$ is considered a column vector. We call a set valued function a multifunction. Given such a multifunction $\Phi : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$, its graph is defined by $\operatorname{gph} \Phi := \{(u, v) \in \mathbb{R}^{p+q} \mid v \in \Phi(u)\}$. Finally, given an arbitrary set $K \subseteq \mathbb{R}^n$, it's polar is defined by $K^\circ := \{v \in \mathbb{R}^n \mid v^T w \leq 0 \, \forall w \in K\}$.

2 Preliminaries

For the reader's convenience we start with the definitions of several basic notions in variational analysis which will be extensively used throughout the paper.

Let $A \subseteq \mathbb{R}^n$ be an arbitrary closed set and $u \in A$. The nonempty cone

$$\begin{aligned} T_A(u) &:= \limsup_{\tau \searrow 0} \frac{A - u}{\tau} \\ &:= \{d \in \mathbb{R}^n \mid \exists \{u^k\} \subset A, \exists \{\tau_k\} \searrow 0 : u^k \to u, \frac{u^k - u}{\tau_k} \to d\} \end{aligned}$$

is called the *contingent* (also *Bouligand* or *tangent*) *cone* to A at u. Consider the family of closed sets A_i, $i = 1, \dots, k$, and a point $u \in \bigcap_{i=1}^k A_i$. Then it follows directly from the above definition that

$$T_{\bigcup_{i=1}^k A_i}(u) = \bigcup_{i=1}^k T_{A_i}(u). \tag{1}$$

Furthermore, we use the contingent cone to define the *Fréchet normal cone*

$$\hat{N}_A(u) := (T_A(u))^\circ \tag{2}$$

to A at u. Note that the Fréchet normal cone is sometimes referred to as the regular normal cone. This is most notably the case in [18]. Again, if we consider the family of closed sets A_i, $i = 1, \dots, k$, and a point $x \in \bigcap_{i=1}^k A_i$, it follows from (1) and the properties of the polar cone that

$$\hat{N}_{\bigcup_{i=1}^k A_i}(u) = \big(T_{\bigcup_{i=1}^k A_i}(u)\big)^\circ = \bigcap_{i=1}^k (T_{A_i}(u))^\circ = \bigcap_{i=1}^k \hat{N}_{A_i}(u) \tag{3}$$

(see [1, Theorem 3.1.9]).

The cornerstone of the generalized differential calculus of Mordukhovich is the *limiting normal cone*, also called the *Mordukhovich normal cone*, defined by

$$\begin{aligned} N_A(u) &:= \limsup_{u' \xrightarrow{A} u} \hat{N}_A(u') \\ &:= \{\lim_{k \to \infty} w^k \mid \exists \{u^k\} \subset A : \lim_{k \to \infty} u^k = u, w^k \in \hat{N}_A(u^k)\}. \end{aligned}$$

In contrast to the Fréchet normal cone, $N_A(u)$ can be nonconvex, but admits a widely developed calculus, cf. [11, 18]. In this calculus, an important role is played by stability of multifunctions reflecting the local structure of A. This will become important when we introduce constraint qualifications further down. Note that if $N_A(u) = \hat{N}_A(u)$, we say that the set A is *(normally) regular* at u.

Consider now the general mathematical program

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & F(z) \in \Lambda, \end{array} \tag{4}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $F : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions and $\Lambda \subseteq \mathbb{R}^m$ is a nonempty closed (possibly nonconvex) set. It is clear that the constraint $F(z) \in \Lambda$ can also incorporate geometric constraints of the form $z \in \Omega$, where $\Omega \subseteq \mathbb{R}^n$ is an arbitrary closed set.

Utilizing the normal cones introduced above, we are now able to define two stationarity concepts which will play a central role in this paper.

Definition 2.1 *Let $\hat{z}$ be feasible for the program (4).*

(a) We say that $\hat{z}$ is M-stationary *if there exists a KKT vector $\hat{\lambda} \in N_\Lambda(F(\hat{z}))$ such that*

$$0 = \nabla f(\hat{z}) + (\nabla F(\hat{z}))^T \hat{\lambda}. \tag{5}$$

(b) We say that $\hat{z}$ is strongly stationary *if there exists a KKT vector $\hat{\lambda} \in \hat{N}_\Lambda(F(\hat{z}))$ such that*

$$0 = \nabla f(\hat{z}) + (\nabla F(\hat{z}))^T \hat{\lambda}. \tag{6}$$

Note that M- and strong stationarity may be expressed as

$$0 \in \nabla f(\hat{z}) + (\nabla F(\hat{z}))^T N_\Lambda(F(\hat{z}))$$

and

$$0 \in \nabla f(\hat{z}) + (\nabla F(\hat{z}))^T \hat{N}_\Lambda(F(\hat{z})),$$

respectively.

The name M-stationarity is motivated by the fact that it involves the Mordukhovich normal cone. It was coined in the context of MPECs by Scholtes [20]. Similarly, we use the name *strong stationarity* in accordance with [19], where it was also used in the context of MPECs. In the MPEC setting, strong stationarity is sometimes also referred to as primal-dual stationarity [15].

As already mentioned, M-stationarity is a weaker stationarity concept than strong stationarity. However, the limiting normal cone (though nonconvex in general) has an extensive calculus, whereas the fuzzy calculus, developed for the Fréchet normal cone is not useful for our purposes.

Having introduced stationarity concepts for the program (4), we now turn our attention to constraint qualifications. To this end, we need the concept of calmness of multifunctions and the closely related idea of upper Lipschitz continuity.

Definition 2.2 *Let* $\Phi : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ *be a multifunction with a closed graph and* $(u, v) \in \operatorname{gph} \Phi$. *We say that* Φ *is* calm at (u, v) *provided there exist neighborhoods* $\mathcal{U}$ *of* u, $\mathcal{V}$ *of* v, *and a modulus* $L \geq 0$ *such that*

$$\Phi(u') \cap \mathcal{V} \subset \Phi(u) + L\|u' - u\|\, \mathbb{B} \quad \text{for all } u' \in \mathcal{U}. \tag{7}$$

If (7) holds true with $\mathcal{V} = \mathbb{R}^q$, Φ *is said to be* locally upper Lipschitz continuous *at* u, *cf. [17].*

Note that calmness is sometimes also referred to as *pseudo upper Lipschitz continuity* (see, e.g., [26]) and that in [18] both calmness and local upper Lipschitz continuity are referred to as calmness.

To obtain a constraint qualification for the program (4), we define a multifunction $M : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ associated with the constraint system of (4) in the following fashion:

$$M(p) := \{z \in \mathbb{R}^n \mid F(z) + p \in \Lambda\}. \tag{8}$$

We can now use this multifunction to define a constraint qualification under which M-stationarity is a first order optimality condition. This is made precise in the following theorem, a proof of which may be found, e.g., in [14, Theorem 2.4].

Theorem 2.3 *Let* $\hat{z}$ *be a local minimizer of (4). If the multifunction* M *(see (8)) is calm at* $(0, \hat{z})$, $\hat{z}$ *is M-stationary.*

Note that if we drop the continuous differentiability of f and F in favor of Lipschitz continuity near $\hat{z}$ of these functions, the result of Theorem 2.3 remains true if we state M-stationarity (5) in terms of the subgradient of f and the coderivative of an appropriate function, see [26] for details.

A constraint qualification better known in nonlinear programming is the Mangasarian-Fromovitz constraint qualification (MFCQ). We define a generalization of this CQ.

Definition 2.4 *Let* $\hat{z}$ *be feasible for the program (4). We say that the* generalized Mangasarian-Fromovitz constraint qualification (GMFCQ) *holds at* $\hat{z}$ *if the implication*

$$\left.\begin{array}{l} \nabla F(\hat{z})^T \lambda = 0 \\ \lambda \in N_\Lambda(F(\hat{z})) \end{array}\right\} \Longrightarrow \lambda = 0 \tag{9}$$

holds.

If $\Lambda = \mathbb{R}^m_-$, the condition (9) reduces to the classical Mangasarian-Fromovitz constraint qualification, justifying the name GMFCQ.

We now show that M-stationarity is a first order condition under GMFCQ.

Corollary 2.5 *Let* $\hat{z}$ *be a local minimizer of the program (4) at which GMFCQ holds. Then* $\hat{z}$ *is M-stationary.*

Proof. Due to Theorem 2.3, it suffices to show that GMFCQ implies the calmness of M at $(0, \hat{z})$. From [12, Corollary 4.4], we readily infer that GMFCQ implies the so-called Aubin property of M around $(0, \hat{z})$. Since this property is stronger than the required calmness (see, e.g., [18, Definition 9.36]), the result follows. □

For more information on the Aubin property, we refer the reader to the book by Rockafellar and Wets [18].

Having discussed both calmness and GMFCQ as constraint qualifications, we now introduce generalizations of the Abadie and Guignard constraint qualifications known from standard nonlinear programming.

Definition 2.6 *Let $\hat{z}$ be feasible for the program (4).*

(a) We say that the generalized Abadie constraint qualification (GACQ) *holds at $\hat{z}$ if*

$$T_{F^{-1}(\Lambda)}(\hat{z}) = L(\hat{z}), \tag{10}$$

where

$$L(\hat{z}) := \{h \in \mathbb{R}^n \mid \nabla F(\hat{z})\, h \in T_\Lambda(F(\hat{z}))\} \tag{11}$$

is called the linearized cone *at $\hat{z}$.*

(b) We say that the generalized *(or* dual*)* Guignard constraint qualification (GGCQ) *holds at $\hat{z}$ if*

$$\hat{N}_{F^{-1}(\Lambda)}(\hat{z}) = (L(\hat{z}))^\circ. \tag{12}$$

To justify the names of the above constraint qualifications, we apply them to a standard nonlinear program that takes the form

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0, \quad h(z) = 0, \end{array} \tag{13}$$

with continuously differentiable $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, and $h : \mathbb{R}^n \to \mathbb{R}^p$. To formulate this program in the fashion of (4), we set

$$F(z) := (g(z), h(z)) \qquad \text{and} \qquad \Lambda := \mathbb{R}^m_- \times 0_p.$$

It is then easily verified that

$$\begin{aligned} L(\hat{z}) = \{d \in \mathbb{R}^n \mid \ & \nabla g_i(\hat{z})^T d \le 0, \qquad \forall i \in \mathcal{I}_g, \\ & \nabla h_i(\hat{z})^T d = 0, \qquad \forall i = 1, \ldots, p\,\}, \end{aligned}$$

where $\mathcal{I}_g := \{i \mid g_i(\hat{z}) = 0\}$ is the set of active inequality constraints.

Clearly, in this case GACQ and GGCQ reduce to the standard definitions of Abadie and Guignard CQ, respectively, as they are most commonly stated, see, e.g., [1]. However, Guignard [7] originally stated her CQ in the primal form. Her definition is equivalent to the above dual characterization of Guignard CQ if Λ is convex. In the nonconvex case the relationship is not as straightforward. Since Guignard CQ is commonly defined using the dual formulation (see, e.g., [16, 1, 6, 21]), however, we feel justified in calling (12) the generalized Guignard CQ.

Note that since $\hat{N}_{F^{-1}(\Lambda)}(\hat{z}) = T_{F^{-1}(\Lambda)}(\hat{z})^\circ$ by definition (see (2)), GACQ at $\hat{z}$ obviously implies GGCQ at $\hat{z}$. Furthermore, as proved in [9, Proposition 1], calmness of M at $(0, \hat{z})$ implies GACQ at $\hat{z}$. Together with the discussion preceding Corollary 2.5, this yields the following chain of implications:

$$\text{GMFCQ at } \hat{z} \Longrightarrow \text{calmness of } M \text{ at } (0, \hat{z}) \Longrightarrow \text{GACQ at } \hat{z} \Longrightarrow \text{GGCQ at } \hat{z}. \tag{14}$$

In general, none of the implications can be reversed. Specifically, the following example demonstrates that GGCQ does not, in general, imply GACQ.

Example 2.7 Consider the program

$$\begin{array}{ll} \text{minimize} & z_1^2 + z_2^2 \\ \text{subject to} & \begin{bmatrix} z_1^2 \\ z_2^2 \end{bmatrix} \in \Lambda := \Lambda_1 \cup \Lambda_2, \end{array}$$

with $\Lambda_1 := \mathbb{R}_+ \times \{0\}$ and $\Lambda_2 := \{0\} \times \mathbb{R}_+$. Obviously, the origin is the unique minimizer. It is then easily verified that the contingent cone takes the form

$$T_{F^{-1}(\Lambda)}(0) = \{h \in \mathbb{R}^2 \mid h_1 h_2 = 0\},$$

while the linearized cone takes the form

$$L(0) = \{h \in \mathbb{R}^2 \mid (0,0)h \in T_\Lambda(0)\} = \mathbb{R}^2.$$

Obviously, $T_{F^{-1}(\Lambda)}(0) \neq L(0)$, so GACQ is not satisfied. However, it is easily verified that

$$T_{F^{-1}(\Lambda)}(0)^\circ = \{0\} = L(0)^\circ,$$

demonstrating that GGCQ does hold.

An important question surrounding GACQ and GGCQ is their connection with optimality conditions. It is known that if Λ is convex, strong stationarity is a first order optimality condition under GGCQ and hence under GACQ (see the arguments following Definition 2.6). In the following sections we will investigate how this can be extended to nonconvex Λ.

Another question is when we might expect strong stationarity to be a first order optimality condition. Since f is continuously differentiable, we have the well-known optimality condition

$$0 \in \nabla f(\hat{z}) + \hat{N}_{F^{-1}(\Lambda)}(\hat{z}), \tag{15}$$

cf., e.g., [18, Theorem 6.12]. Unfortunately, we only have the inclusion

$$\hat{N}_{F^{-1}(\Lambda)}(\hat{z}) \supset (\nabla F(\hat{z}))^T \ \hat{N}_\Lambda (F(\hat{z})) \tag{16}$$

which turns out to be an equality provided GMFCQ is satisfied at $\hat{z}$ and Λ is (normally) regular at $F(\hat{z})$ (see [18, Theorem 6.14]). However, this condition never holds for equilibrium constraints. This question is important, because we need to have equality in (16) in order to explicitly determine the cone $\hat{N}_{F^{-1}(\Lambda)}(\hat{z})$, and with it to be able to verify whether strong stationarity holds.

3 The Disjunctive Program

In this section we will investigate GACQ and GGCQ in context with M- and strong stationarity under the structural assumption that

$$\Lambda = \bigcup_{i=1}^{k} \Lambda_i, \tag{17}$$

where all $\Lambda_i, i = 1, \ldots, k$, are convex polyhedra. This assumption is satisfied by a large class of equilibrium constraints, see Section 4 for an example of this.

The sets Λ_i are called *components* and with each $z \in \Lambda$ we associate the index set $\mathbb{I}(z)$ of *active components* defined by

$$\mathbb{I}(z) := \{i \in \{1, \ldots, k\} \mid F(z) \in \Lambda_i\}.$$

For the remainder of this section, we will confine ourselves to the program

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & F(z) \in \Lambda = \bigcup_{i=1}^{k} \Lambda_i, \end{array} \tag{18}$$

where the constraints are given by (17).

For such a program, we are indeed able to show that M-stationarity is a necessary first order condition under GGCQ and hence GACQ.

Theorem 3.1 *Let $\hat{z}$ be a local minimizer of the program (18). Then, if GGCQ is satisfied at $\hat{z}$, $\hat{z}$ is M-stationary.*

Proof. As mentioned in the previous section, relation (15) holds true. Thus, due to GGCQ at $\hat{z}$, one has

$$0 \in \nabla f(\hat{z}) + (L(\hat{z}))^{\circ}. \tag{19}$$

By definition of the polar cone, this means that $h = 0$ is the (unique) solution of the program

$$\begin{array}{ll} \text{minimize} & \nabla f(\hat{z})^T h \\ \text{subject to} & \nabla F(\hat{z})\, h \in T_{\Lambda}(F(\hat{z})). \end{array} \tag{20}$$

By virtue of (1),

$$T_{\Lambda}(F(\hat{z})) = \bigcup_{i \in \mathbb{I}(\hat{z})} T_{\Lambda_i}(F(\hat{z})).$$

Consider now the multifunction $\tilde{M} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, defined by

$$\tilde{M}(p) := \{h \in \mathbb{R}^n \mid \nabla F(\hat{z})\, h + p \in T_\Lambda(F(\hat{z}))\}\,.$$

Clearly, the constraint in (20) can equivalently be written as $h \in \tilde{M}(0)$. Moreover, since the contingent cones $T_{\Lambda_i}(F(\hat{z}))$, $i \in \mathbb{I}(\hat{z})$, are convex polyhedra,

$$\begin{aligned} \operatorname{gph} \tilde{M} &= \big\{(p,h) \in \mathbb{R}^m \times \mathbb{R}^n \,|\, \nabla F(\hat{z})\, h + p \in \bigcup_{i \in \mathbb{I}(\hat{z})} T_{\Lambda_i}(F(\hat{z}))\big\} \\ &= \bigcup_{i \in \mathbb{I}(\hat{z})} \{(p,h) \in \mathbb{R}^m \times \mathbb{R}^n \,|\, \nabla F(\hat{z})\, h + p \in T_{\Lambda_i}(F(\hat{z}))\} \end{aligned}$$

is a union of finitely many convex polyhedra (the second equality is trivially verified). From [17, Proposition 1] it follows that $\tilde{M}$ is locally upper Lipschitz at 0 and, a fortiori, calm at $(0, \hat{z})$. Thus, by [8, Theorem 4.1] we have

$$N_{\tilde{M}(0)}(0) \subseteq (\nabla F(\hat{z}))^T N_{T_\Lambda(F(\hat{z}))}(0). \tag{21}$$

It remains to invoke [18, Example 6.47] according to which, due to (1), there exists a neighborhood $\mathcal{O}$ of $F(\hat{z})$ in $\mathbb{R}^m$ such that

$$\Lambda \cap \mathcal{O} = [F(\hat{z}) + T_\Lambda(F(\hat{z}))] \cap \mathcal{O}.$$

Consequently, by the definition of the limiting normal cone,

$$N_{T_\Lambda(F(\hat{z}))}(0) = N_{F(\hat{z})+T_\Lambda(F(\hat{z}))}(F(\hat{z})) = N_\Lambda(F(\hat{z})). \tag{22}$$

Finally, remembering that $\hat{h} = 0$ is a minimizer of (20), applying the optimality condition (15) to the program (20) yields (note that h is the variable)

$$\begin{aligned} 0 &\overset{(15)}{\in} \nabla f(\hat{z}) + \hat{N}_{\tilde{M}(0)}(0) \\ &\subseteq \nabla f(\hat{z}) + N_{\tilde{M}(0)}(0) \\ &\overset{(21)}{\subseteq} \nabla f(\hat{z}) + (\nabla F(\hat{z}))^T N_{T_\Lambda(F(\hat{z}))}(0) \\ &\overset{(22)}{=} \nabla f(\hat{z}) + (\nabla F(\hat{z}))^T N_\Lambda(F(\hat{z})). \end{aligned}$$

This is the condition for M-stationarity (5), completing the proof. □

Note that the continuous differentiability of f enters crucially in constructing the program (20). The statement of Theorem 3.1 cannot be proved using the above technique if we replace continuously differentiable data with locally Lipschitz data.

We now turn our attention to strong stationarity of $\hat{z}$ if it is a minimizer of a program (18). Let us consider the following academic example, taken from [23].

Example 3.2 Consider the program

$$\begin{array}{ll} \text{minimize} & -z_2 \\ \text{subject to} & \begin{bmatrix} z_1 \\ z_2 \\ z_1 - z_2 \end{bmatrix} \in \Lambda_1 \cup \Lambda_2, \end{array}$$

with $\Lambda_1 = \mathbb{R}_+ \times \{0\} \times \{0\}$ and $\Lambda_2 = \{0\} \times \mathbb{R}_+ \times \{0\}$. It is easy to verify that $\hat{z} = 0$ is the unique minimizer of this program and that GMFCQ is satisfied at the origin. Hence, $\hat{z}$ is M-stationary and the vector

$$\hat{\lambda} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in N_\Lambda(0) \tag{23}$$

belongs to the set of KKT vectors. On the other hand,

$$\hat{N}_\Lambda(0) = \hat{N}_{\Lambda_1}(0) \cap \hat{N}_{\Lambda_2}(0) = \left\{\lambda \in \mathbb{R}^3 \,|\, \lambda_1 \leq 0,\ \lambda_2 \leq 0\right\}$$

(see (3)) and we easily verify that there exists no $\tilde{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3) \in \hat{N}_\Lambda(0)$ such that

$$0 \in \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ \tilde{\lambda}_3 \end{bmatrix}.$$

This example shows (by virtue of (14)) that none of the CQs mentioned in Section 2 can ensure strong stationarity of local minima in the considered program. We therefore need an additional condition which we introduce in the following definition.

Definition 3.3 *Let $\hat{z}$ be feasible for the program (18). We say that the* intersection property *holds at $\hat{z}$ if*

$$\bigcap_{i \in \mathbb{I}(\hat{z})} (\nabla F(\hat{z}))^T \hat{N}_{\Lambda_i}(F(\hat{z})) = (\nabla F(\hat{z}))^T \bigcap_{i \in \mathbb{I}(\hat{z})} \hat{N}_{\Lambda_i}(F(\hat{z})). \tag{I}$$

In order to use the intersection property to acquire a stronger optimality condition, we first need the following result, concerning the polar of a polyhedral cone, the proof of which may be found in [1, Theorem 3.2.2].

Lemma 3.4 *Let the cones*

$$\begin{aligned} \mathcal{K}_1 := \{d \in \mathbb{R}^n \mid\ & a_i^T d \geq 0, \qquad \forall i = 1, \dots, k, \\ & b_j^T d = 0, \qquad \forall j = 1, \dots, l\} \end{aligned} \tag{24}$$

and

$$\begin{aligned} \mathcal{K}_2 = \{v \in \mathbb{R}^n \mid\ v = & \sum_{i=1}^k \alpha_i a_i + \sum_{j=1}^l \beta_j b_j, \\ & \alpha_i \leq 0, \quad \forall i = 1, \dots, k\} \end{aligned} \tag{25}$$

be given. Then $\mathcal{K}_1 = \mathcal{K}_2^\circ$ and $\mathcal{K}_1^\circ = \mathcal{K}_2$.

We are now able to state a stronger optimality condition than in Theorem 3.1.

Theorem 3.5 *Let $\hat{z}$ be a local minimizer of the program (18). If GGCQ and the intersection property (I) are satisfied at $\hat{z}$, then $\hat{z}$ is strongly stationary.*

Proof. As in the proof of Theorem 3.1 we start with the relation

$$0 \in \nabla f(\hat{z}) + (L(\hat{z}))^\circ \tag{26}$$

which holds by virtue of the local optimality of $\hat{z}$ (see (15)) and GGCQ. From (1) we infer that

$$\begin{aligned}(L(\hat{z}))^\circ &= \Big(\bigcup_{i \in \mathbb{I}(\hat{z})} \{h \in \mathbb{R}^n \mid \nabla F(\hat{z})\, h \in T_{\Lambda_i}(F(\hat{z}))\} \Big)^\circ \\ &= \bigcap_{i \in \mathbb{I}(\hat{z})} \big(\{h \in \mathbb{R}^n \mid \nabla F(\hat{z})\, h \in T_{\Lambda_i}(F(\hat{z}))\}\big)^\circ.\end{aligned} \tag{27}$$

Due to the polyhedrality of $T_{\Lambda_i}(F(\hat{z}))$, there exist $a_i \in \mathbb{R}^m$, $i = 1, \dots, q$ for an appropriate q, such that

$$T_{\Lambda_i}(F(\hat{z})) = \{d \in \mathbb{R}^m \mid a_i^T d \geq 0, \quad \forall i = 1, \dots, q\}.$$

Hence we can write

$$\begin{aligned} &\{h \in \mathbb{R}^n \mid \nabla F(\hat{z})\, h \in T_{\Lambda_i}(F(\hat{z}))\} \\ &= \{h \in \mathbb{R}^n \mid a_i^T(\nabla F(\hat{z}) h) \geq 0, \quad \forall i = 1, \dots, q\} \\ &= \{h \in \mathbb{R}^n \mid (\nabla F(\hat{z})^T a_i)^T h \geq 0, \quad \forall i = 1, \dots, q\}.\end{aligned}$$

We now apply Lemma 3.4 to obtain the polar of this set:

$$\begin{aligned} &\big(\{h \in \mathbb{R}^n \mid \nabla F(\hat{z})\, h \in T_{\Lambda_i}(F(\hat{z}))\}\big)^\circ \\ &= \{v \in \mathbb{R}^n \mid v = \sum_{i=1}^q \alpha_i \nabla F(\hat{z})^T a_i, \quad \alpha \leq 0\} \\ &= \{v \in \mathbb{R}^n \mid v = \nabla F(\hat{z})^T \sum_{i=1}^q \alpha_i a_i, \quad \alpha \leq 0\} \\ &= \nabla F(\hat{z})^T T_{\Lambda_i}(F(\hat{z}))^\circ = \nabla F(\hat{z})^T \hat{N}_{\Lambda_i}(F(\hat{z})),\end{aligned}$$

where the penultimate equality is verified again by applying Lemma 3.4. Substituting this into (27), we get

$$(L(\hat{z}))^\circ = \bigcap_{i \in \mathbb{I}(\hat{z})} (\nabla F(\hat{z}))^T \hat{N}_{\Lambda_i}(F(\hat{z})) = (\nabla F(\hat{z}))^T \bigcap_{i \in \mathbb{I}(\hat{z})} \hat{N}_{\Lambda_i}(F(\hat{z})), \tag{28}$$

where the last equality is exactly condition (I). Since

$$\hat{N}_\Lambda(F(\hat{z})) = \bigcap_{i \in \mathbb{I}(\hat{z})} \hat{N}_{\Lambda_i}(F(\hat{z}))$$

(see (3)), we obtain the statement of the theorem by substituting (28) in (26). □

We easily observe that it is just condition (I) which is missing in Example 3.2. Indeed,

$$(\nabla F(\hat{z}))^T \hat{N}_{\Lambda_1}(F(\hat{z})) \cap (\nabla F(\hat{z}))^T \hat{N}_{\Lambda_2}(F(\hat{z})) = \mathbb{R}^2,$$

whereas

$$(\nabla F(\hat{z}))^T \left(\hat{N}_{\Lambda_1}(F(\hat{z})) \cap \hat{N}_{\Lambda_2}(F(\hat{z}))\right) = \left\{ \begin{bmatrix} \lambda^1 + \lambda^3 \\ \lambda^2 - \lambda^3 \end{bmatrix} \middle| \lambda_1, \lambda_2 \le 0 \right\} \ne \mathbb{R}^2, \tag{29}$$

i.e. the KKT vector $\hat{\lambda}$ from (23) violates the constraints in (29).

One readily infers that the surjectivity of $\nabla F(\hat{z})$ insures (I). This is, however, a substantial strengthening even of GMFCQ. In the following section we will discuss two conditions that imply the intersection property (I) in the context of MPECs.

4 Application to MPECs

In this section we consider a special case of the program (18), *mathematical programs with equilibrium* or *complementarity constraints* (*MPECs* for short). They take the form

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0, \quad h(z) = 0, \\ & G(z) \ge 0, \quad H(z) \ge 0, \quad G(z)^T H(z) = 0, \end{array} \tag{30}$$

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.

It is possible to write the program (30) in the form of (18) in several ways. One way we already discussed in the context of program (13). We will go into this, more obvious, choice later. First, however, we will introduce a particular reformulation which will render useful statements when we apply the results from Section 3.

Let us therefore consider a local minimizer $\hat{z}$ of (30). From the complementarity term in (30) it is clear that either $G_i(\hat{z})$, or $H_i(\hat{z})$, or both must be zero. To distinguish among these cases, we divide the indices of G and H into three sets:

$$\alpha := \alpha(\hat{z}) := \{i \mid G_i(\hat{z}) = 0,\ H_i(\hat{z}) > 0\}, \tag{31a}$$

$$\beta := \beta(\hat{z}) := \{i \mid G_i(\hat{z}) = 0,\ H_i(\hat{z}) = 0\}, \tag{31b}$$

$$\gamma := \gamma(\hat{z}) := \{i \mid G_i(\hat{z}) > 0,\ H_i(\hat{z}) = 0\}. \tag{31c}$$

The set β is called the *degenerate* or, a term which has come into use more recently, the *biactive set*. Note that the sets (31) may also be defined for arbitrary feasible points of (30).

Next we define the function F in (18) utilizing these sets:

$$F(z) := \big(g(z), h(z), G_\alpha(z), H_\alpha(z), G_\beta(z), H_\beta(z), G_\gamma(z), H_\gamma(z)\big). \tag{32}$$

Finally, we take a pair (β_1, β_2) from $\mathcal{P}(\beta)$, the set of all partitions of beta ($\mathcal{P}(\beta) := \{(\beta_1, \beta_2) \mid \beta_1 \cup \beta_2 = \beta, \beta_1 \cap \beta_2 = \emptyset\}$). We then proceed to define the sets

$$\Lambda_{\beta_1,\beta_2} := \mathbb{R}^m_- \times 0_p \times 0_{|\alpha|} \times \mathbb{R}^{|\alpha|}_+ \times \Delta_{\beta_1,\beta_2} \times \Delta_{\beta_2,\beta_1} \times \mathbb{R}^{|\gamma|}_+ \times 0_{|\gamma|} \tag{33}$$

with

$$(\Delta_{\mu,\nu})_j := \begin{cases} 0 & : \ j \in \mu \\ \mathbb{R}_+ & : \ j \in \nu. \end{cases} \tag{34}$$

Obviously, the sets $\Lambda_{\beta_1,\beta_2}$ are convex polyhedra.

It is now easy to see that (30) is locally equivalent to

$$\begin{array}{ll} \min & f(z) \\ \text{s.t.} & F(z) \in \Lambda \end{array} \tag{35}$$

with

$$\Lambda := \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \Lambda_{\beta_1,\beta_2}. \tag{36}$$

Note that the definition of this program depends on the solution $\hat{z}$ of the program (30). Therefore, local equivalence refers to the fact that the feasible regions of the programs (30) and (35) are equal for a whole neighborhood of the local minimizer $\hat{z}$ that was used to define (35). Further note that every component $\Lambda_{\beta_1,\beta_2}$ is active at $\hat{z}$. Central to the definition of both GACQ and GGCQ is the linearized cone $L(\hat{z})$ (see (11)). To transfer the conditions for GACQ and GGCQ to the MPEC setting, we investigate this linearized tangent cone $L(\hat{z})$ in the following lemma.

Lemma 4.1 *Let $\hat{z}$ be a local minimizer of the MPEC (30) and consider the locally equivalent program (35). Then the linearized cone may be expressed as follows:*

$$\begin{aligned} L(\hat{z}) = \{d \in \mathbb{R}^n \mid & \nabla g_i(\hat{z})^T d \le 0 && \forall i \in \mathcal{I}_g, \\ & \nabla h_i(\hat{z})^T d = 0 && \forall i = 1, \dots, p, \\ & \nabla G_i(\hat{z})^T d = 0 && \forall i \in \alpha, \\ & \nabla H_i(\hat{z})^T d = 0 && \forall i \in \gamma, \\ & \nabla G_i(\hat{z})^T d \ge 0 && \forall i \in \beta, \\ & \nabla H_i(\hat{z})^T d \ge 0 && \forall i \in \beta, \\ & (\nabla G_i(\hat{z})^T d) \cdot (\nabla H_i(\hat{z})^T d) = 0, && \forall i \in \beta \,\}. \end{aligned} \tag{37}$$

Proof. In order to prove this result, we need to determine the contingent cones of the components $\Lambda_{\beta_1,\beta_2}$ at the point $F(\hat{z})$ (it is important to note here that the index sets α, β, and γ depend on the point $\hat{z}$ at which we evaluate F, and that some components of $F(\hat{z})$ are in the interior of $\Lambda_{\beta_1,\beta_2}$):

$$T_{\Lambda_{\beta_1,\beta_2}}(F(\hat{z})) = (\Delta_{\overline{\mathcal{I}}_g,\mathcal{I}_g})^\circ \times 0_p \times 0_{|\alpha|} \times \mathbb{R}^{|\alpha|} \times \Delta_{\beta_1,\beta_2} \times \Delta_{\beta_2,\beta_1} \times \mathbb{R}^{|\gamma|} \times 0_{|\gamma|}, \tag{38}$$

where $\mathcal{I}_g := \{i \mid g_i(\hat{z}) = 0\}$ is the set of active inequality constraints and $\overline{\mathcal{I}}_g := \{1, \dots, m\} \backslash \mathcal{I}_g$ is its complement. Note that

$$(\Delta_{\overline{\mathcal{I}}_g, \mathcal{I}_g})_j^\circ = \begin{cases} \mathbb{R} & : j \notin \mathcal{I}_g \\ \mathbb{R}_- & : j \in \mathcal{I}_g. \end{cases}$$

Since each $\Lambda_{\beta_1,\beta_2}$ is the Cartesian product of (normally) regular sets, we can apply [18, Proposition 6.41] in addition to the relation (1) to obtain an explicit representation of $T_\Lambda(F(\hat{z}))$ and with it the representation (37) of the linearized cone $L(\hat{z})$. □

In the MPEC literature, the linearized cone $L(\hat{z})$ is commonly called the *MPEC-linearized tangent cone* and is denoted by $\mathcal{T}^{lin}_{\text{MPEC}}(\hat{z})$ (see, e.g., [5, 3, 2, 25]). In the following, we will use this terminology to avoid confusion with the linearized tangent cone $\mathcal{T}^{lin}(\hat{z})$, which we acquire if the reformulate the program (30) in the fashion to be described in the following.

By setting

$$F(z) := \big(g(z), h(z), G(z), H(z), G(z)^T H(z)\big)$$

and

$$\Lambda := \mathbb{R}^m_- \times 0_p \times \mathbb{R}^l_+ \times \mathbb{R}^l_+ \times 0_1,$$

we obviously obtain a trivial equivalent formulation of (30) in the form of (18). It is then easily verified that the corresponding linearized cone $L(\hat{z})$ can be expressed as

$$\begin{aligned} \mathcal{T}^{lin}(\hat{z}) = \{d \in \mathbb{R}^n \mid \; & \nabla g_i(\hat{z})^T d \le 0 && \forall i \in \mathcal{I}_g, \\ & \nabla h_i(\hat{z})^T d = 0 && \forall i = 1, \dots, p, \\ & \nabla G_i(\hat{z})^T d = 0 && \forall i \in \alpha, \\ & \nabla H_i(\hat{z})^T d = 0 && \forall i \in \gamma, \\ & \nabla G_i(\hat{z})^T d \ge 0 && \forall i \in \beta, \\ & \nabla H_i(\hat{z})^T d \ge 0 && \forall i \in \beta\,\}. \end{aligned} \tag{39}$$

See [5] for more discussion on the linearized tangent cone $\mathcal{T}^{lin}(\hat{z})$.

We can now use the different formulations of the MPEC (30) with their different linearized cones to define variants of the GACQ and GGCQ constraint qualifications. We call the condition

$$T_{F^{-1}(\Lambda)}(\hat{z}) = \mathcal{T}^{lin}_{\text{MPEC}}(\hat{z}) \tag{40}$$

MPEC-ACQ, as has been adapted by the MPEC community, while we call

$$T_{F^{-1}(\Lambda)}(\hat{z}) = \mathcal{T}^{lin}(\hat{z}) \tag{41}$$

simply Abadie CQ (or ACQ), because it is obviously the classical definition of Abadie CQ.

Similarly, we call the condition

$$T_{F^{-1}(\Lambda)}(\hat{z})^\circ = \mathcal{T}^{lin}_{\text{MPEC}}(\hat{z})^\circ \tag{42}$$

MPEC-GCQ, and call

$$T_{F^{-1}(\Lambda)}(\hat{z})^\circ = \mathcal{T}^{lin}(\hat{z})^\circ \tag{43}$$

Guignard CQ, it also being the classical definition of Guignard CQ.

We will now proceed to apply Theorem 3.1 to an MPEC of the form (30). In order to do this, we exploit the following proposition which we will use to reformulate the complementarity constraints in (30).

Proposition 4.2 *Let the set*

$$\mathcal{C} := \{(a, b) \in \mathbb{R}^{2|\beta|} \mid a \geq 0, b \geq 0, a^T b = 0\} \tag{44}$$

be given. Then the Fréchet and limiting normal cones at the origin are given by

$$\hat{N}_{\mathcal{C}}(0) = \mathbb{R}^{2|\beta|}_-, \tag{45}$$

$$N_{\mathcal{C}}(0) = \mathbb{R}^{2|\beta|}_- \cup \mathcal{C}, \tag{46}$$

respectively.

Proof. We can map $\mathcal{C}$ isomorphically to the Cartesian product

$$\tilde{\mathcal{C}} := \mathcal{C}_1 \times \ldots \times \mathcal{C}_{|\beta|}$$

with

$$\mathcal{C}_i := \{(a_i, b_i) \in \mathbb{R}^2 \mid a_i \geq 0, b_i \geq 0, a_i b_i = 0\}$$

by simply rearranging the components of $\mathcal{C}$ in an appropriate fashion.

It follows directly from (2) that

$$\hat{N}_{\mathcal{C}_i}(0) = \mathbb{R}^2_-. \tag{47}$$

Since the sets $\mathcal{C}_i$ are obviously closed, we can apply [18, Proposition 6.41] to obtain

$$\hat{N}_{\tilde{\mathcal{C}}}(0) = \hat{N}_{\mathcal{C}_1}(0) \times \ldots \times \hat{N}_{\mathcal{C}_{|\beta|}}(0).$$

Together with (47), this yields that

$$\hat{N}_{\tilde{\mathcal{C}}}(0) = \mathbb{R}^{2|\beta|}_-.$$

Applying the appropriate inverse mapping to acquire $\hat{N}_{\mathcal{C}}(0)$ from $\hat{N}_{\tilde{\mathcal{C}}}(0)$, we obtain (45).

The proof for the limiting normal cone $N_{\mathcal{C}}(0)$ is identical to the method used above, except that

$$N_{\mathcal{C}_i}(0) = \mathbb{R}^2_- \cup \mathcal{C}_i,$$

as can be gleaned from the proof of [13, Lemma 2.2]. Using this, we immediately acquire (46) employing the same arguments as above. □

We are now able to state the conditions for M-stationarity for MPECs in a much more tangible fashion than in the general case (see (5)).

Theorem 4.3 *Let $\hat{z}$ be a local minimizer of the MPEC (30) at which MPEC-GCQ holds. Then there exist KKT vectors λ^g, λ^h, λ^G, and λ^H such that*

$$0 = \nabla f(\hat{z}) + \sum_{i=1}^{m} \lambda_i^g \nabla g_i(\hat{z}) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\hat{z}) - \sum_{i=1}^{l} \left[\lambda_i^G \nabla G_i(\hat{z}) + \lambda_i^H \nabla H_i(\hat{z})\right],$$

$$\begin{array}{lll} \lambda_\alpha^G \ \text{free}, & & \lambda_\gamma^G = 0, \\ & (\lambda_i^G > 0 \wedge \lambda_i^H > 0) \ \vee \ \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta & \\ \lambda_\gamma^H \ \text{free}, & & \lambda_\alpha^H = 0, \end{array} \tag{48}$$

$$g(\hat{z}) \leq 0, \qquad \lambda^g \geq 0, \qquad g(\hat{z})^T \lambda^g = 0.$$

In this case we call $\hat{z}$ M-stationary.

Proof. We apply Theorem 3.1 to an MPEC using the formulation (35), (36).

All that remains is to show that M-stationarity takes on the form (48). To this end, we need to determine $N_\Lambda(F(\hat{z}))$, see (5). Referring to (33) and (36), it is easy to see that Λ can be written as

$$\Lambda = \mathbb{R}_-^m \times 0_p \times 0_{|\alpha|} \times \mathbb{R}_+^{|\alpha|} \times \mathcal{C} \times \mathbb{R}_+^{|\gamma|} \times 0_{|\gamma|} \tag{49}$$

where $\mathcal{C}$ is defined in (44).

Observing that Λ is the Cartesian product of closed sets (see (49)), we can apply [18, Proposition 6.41] to obtain

$$\begin{aligned} N_\Lambda(F(\hat{z})) &= N_{\mathbb{R}_-^m \times 0_p \times 0_{|\alpha|} \times \mathbb{R}_+^{|\alpha|} \times \mathcal{C} \times \mathbb{R}_+^{|\gamma|} \times 0_{|\gamma|}}(F(\hat{z})) \\ &= N_{\mathbb{R}_-^m}(g(\hat{z})) \times N_{0_p}(h(\hat{z})) \times N_{0_{|\alpha|}}(G_\alpha(\hat{z})) \times N_{\mathbb{R}_+^{|\alpha|}}(H_\alpha(\hat{z})) \times \\ &\qquad N_{\mathcal{C}}((G_\beta(\hat{z}), H_\beta(\hat{z}))) \times N_{\mathbb{R}_+^{|\gamma|}}(G_\gamma(\hat{z})) \times N_{0_{|\gamma|}}(H_\gamma(\hat{z})). \end{aligned} \tag{50}$$

Note that the limiting normal cones to orthants and isolated points are equal to the standard convex normal cone since these sets are convex (see [18, Theorem 6.9]). Additionally, we apply Proposition 4.2 to get $N_{\mathcal{C}}((G_\beta(\hat{z}), H_\beta(\hat{z}))) = N_{\mathcal{C}}((0,0)) = \mathbb{R}_-^{2|\beta|} \cup \mathcal{C}$. Substitituting the resulting limiting normal cone $N_\Lambda(F(\hat{z}))$ into (5), we obtain the conditions (48) for M-stationarity. Observe the signs used in the Lagrangian in (48). □

Note that even though results similar to Theorem 4.3 have appeared before (cf. [25, 4]) these results assume that the stronger MPEC-ACQ holds at the local minimizer $\hat{z}$. As demonstrated by Example 2.7, MPEC-GCQ is indeed a weaker assumption than MPEC-ACQ, making Theorem 4.3 a stronger result. Note that Example 2.7 is an MPEC of type (35), (36).

We now turn our focus to strong stationarity. In the previous section, the intersection property (I) was needed in addition to GGCQ in order for strong stationarity to be a necessary first order condition. We now state which form (I) assumes in an MPEC setting.

To this end, we first introduce an auxiliary program associated with the MPEC (30) at an arbitrary feasible point $\hat{z}$: Given a partition $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$, let $\mathrm{NLP}_*(\beta_1, \beta_2)$ denote the following nonlinear program:

$$\begin{array}{lll} \text{minimize} & f(z) & \\ \text{subject to} & g(z) \le 0, & h(z) = 0, \\ & G_{\alpha\cup\beta_1}(z) = 0, & H_{\alpha\cup\beta_1}(z) \ge 0, \\ & G_{\gamma\cup\beta_2}(z) \ge 0, & H_{\gamma\cup\beta_2}(z) = 0. \end{array} \tag{51}$$

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

Furthermore, the linearized tangent cone of the $\mathrm{NLP}_*(\beta_1, \beta_2)$ (51) is given by

$$\begin{aligned} \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(\hat{z}) = \{d \in \mathbb{R}^n \mid & \nabla g_i(\hat{z})^T d \le 0, && \forall i \in \mathcal{I}_g, \\ & \nabla h_i(\hat{z})^T d = 0, && \forall i = 1, \ldots, p, \\ & \nabla G_i(\hat{z})^T d = 0, && \forall i \in \alpha \cup \beta_1, \\ & \nabla H_i(\hat{z})^T d = 0, && \forall i \in \gamma \cup \beta_2, \\ & \nabla G_i(\hat{z})^T d \ge 0, && \forall i \in \beta_2, \\ & \nabla H_i(\hat{z})^T d \ge 0, && \forall i \in \beta_1 \}. \end{aligned} \tag{52}$$

It is easily verified that

$$\mathcal{T}^{lin}_{\mathrm{MPEC}}(\hat{z}) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(\hat{z}) \tag{53}$$

(compare (37) with (52)).

We are now able to characterize the intersection property (I) for MPECs.

Lemma 4.4 *The intersection property (I) amounts to*

$$\mathcal{T}^{lin}_{\mathrm{MPEC}}(\hat{z})^\circ = \mathcal{T}^{lin}(\hat{z})^\circ \tag{$\mathrm{I}^{\mathrm{MPEC}}$}$$

for programs of type (30).

Proof. First, let us consider the Fréchet normal cone to the set $\Lambda_{\beta_1,\beta_2}$ at the point $F(\hat{z})$. This is simply the polar of $T_{\Lambda_{\beta_1,\beta_2}}(F(\hat{z}))$ (see (38)):

$$\hat{N}_{\Lambda_{\beta_1,\beta_2}}(F(\hat{z})) = \Delta_{\overline{\mathcal{I}}_g,\mathcal{I}_g} \times \mathbb{R}^p \times \mathbb{R}^{|\alpha|} \times 0_{|\alpha|} \times (\Delta_{\beta_1,\beta_2})^\circ \times (\Delta_{\beta_2,\beta_1})^\circ \times 0_{|\gamma|} \times \mathbb{R}^{|\gamma|}. \tag{54}$$

We now insert this expression on the left hand side of (I) (note that all $(\beta_1, \beta_2) \in \mathcal{P}(\beta)$ are active):

$$\bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} (\nabla F(\hat{z}))^T \hat{N}_{\Lambda_{\beta_1,\beta_2}}(F(\hat{z}))$$

$$= \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \{\sum_{i\in\mathcal{I}_g} \mu_i^g \nabla g_i(\hat{z}) + \sum_{i=1}^p \mu_i^h \nabla h_i(\hat{z}) + \sum_{i\in\alpha\cup\beta} \mu_i^G \nabla G_i(\hat{z}) + \sum_{i\in\gamma\cup\beta} \mu_i^H \nabla H_i(\hat{z}) \mid \mu_{\mathcal{I}_g}^g \geq 0, \mu_{\beta_2}^G \leq 0, \mu_{\beta_1}^H \leq 0\}$$

$$= \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}^{lin}_{\mathrm{NLP}_*(\beta_1,\beta_2)}(\hat{z})^\circ = \mathcal{T}^{lin}_{\mathrm{MPEC}}(\hat{z})^\circ, \tag{55}$$

where the penultimate equality is acquired by applying Lemma 3.4 to (52), and the final equality is verified by taking the polar of (53) and applying [1, Theorem 3.1.9].

We now turn our attention to the right hand side of (I), again inserting (54):

$$(\nabla F(\hat{z}))^T \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \hat{N}_{\Lambda_{\beta_1,\beta_2}}(F(\hat{z}))$$

$$= (\nabla F(\hat{z}))^T \big\{\mu \mid \mu \in \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \big(\Delta_{\overline{\mathcal{I}}_g,\mathcal{I}_g} \times \mathbb{R}^p \times \mathbb{R}^{|\alpha|} \times 0_{|\alpha|} \times \Delta^\circ_{\beta_1,\beta_2} \times \Delta^\circ_{\beta_2,\beta_1} \times 0_{|\gamma|} \times \mathbb{R}^{|\gamma|}\big)\big\}$$

$$= (\nabla F(\hat{z}))^T \big\{\mu \mid \mu \in \Delta_{\overline{\mathcal{I}}_g,\mathcal{I}_g} \times \mathbb{R}^p \times \mathbb{R}^{|\alpha|} \times 0_{|\alpha|} \times \mathbb{R}_-^{|\beta|} \times \mathbb{R}_-^{|\beta|} \times 0_{|\gamma|} \times \mathbb{R}^{|\gamma|}\big\}$$

$$= \{\sum_{i\in\mathcal{I}_g} \mu_i^g \nabla g_i(\hat{z}) + \sum_{i=1}^p \mu_i^h \nabla h_i(\hat{z}) + \sum_{i\in\alpha\cup\beta} \mu_i^G \nabla G_i(\hat{z}) + \sum_{i\in\gamma\cup\beta} \mu_i^H \nabla H_i(\hat{z}) \mid \mu_{\mathcal{I}_g}^g \geq 0, \mu_\beta^G \leq 0, \mu_\beta^H \leq 0\}$$

$$= \mathcal{T}^{lin}(\hat{z})^\circ. \tag{56}$$

Again, the last equality is verified by applying Lemma 3.4 to (39).

Finally, equations (55) and (56) represent the left and right hand sides of (I), respectively, showing that (I) does indeed reduce to ($\mathrm{I}^{\mathrm{MPEC}}$) in an MPEC setting. □

We now use this result to apply Theorem 3.5 to an MPEC.

Theorem 4.5 *Let $\hat{z}$ be a local minimizer of the MPEC (30) at which MPEC-GCQ as well as (I^{MPEC}) holds. Then there exist KKT vectors λ^g, λ^h, λ^G, and λ^H such that*

$$0 = \nabla f(\hat{z}) + \sum_{i=1}^m \lambda_i^g \nabla g_i(\hat{z}) + \sum_{i=1}^p \lambda_i^h \nabla h_i(\hat{z}) - \sum_{i=1}^l \big[\lambda_i^G \nabla G_i(\hat{z}) + \lambda_i^H \nabla H_i(\hat{z})\big],$$

$$\begin{aligned} &\lambda_\alpha^G \quad \text{free}, \qquad \lambda_\beta^G \geq 0, \qquad \lambda_\gamma^G = 0, \\ &\lambda_\gamma^H \quad \text{free}, \qquad \lambda_\beta^H \geq 0, \qquad \lambda_\alpha^H = 0, \end{aligned} \tag{57}$$

$$g(\hat{z}) \leq 0, \qquad \lambda^g \geq 0, \qquad g(\hat{z})^T \lambda^g = 0.$$

In this case we call $\hat{z}$ strongly stationary*.*

Proof. We apply Theorem 3.5 using the formulation ($\mathrm{I}^{\mathrm{MPEC}}$) of the intersection property (I). We then proceed identical to the proof of Theorem 4.3, substituting the Fréchet normal cone $\hat{N}$ for the limiting normal cone N where appropriate.

The only difference is that we obtain $\hat{N}_{\mathcal{C}}((G_\beta(\hat{z}), H_\beta(\hat{z}))) = \hat{N}_{\mathcal{C}}(0,0) = \mathbb{R}_-^{2|\beta|}$ when we apply Proposition 4.2.

Substituting the resulting Fréchet normal $\hat{N}_\Lambda(F(\hat{z}))$ into (6) yields the conditions (57) for strong stationarity, completing the proof. □

The result of Theorem 4.5 may be acquired using a different approach, which we will investigate with the aid of the following lemma, the proof of which follows immediately from the definition of Guignard CQ, MPEC-GCQ and the intersection property ($\mathrm{I}^{\mathrm{MPEC}}$).

Lemma 4.6 *Let $\hat{z}$ be a feasible point of the MPEC (30). Then standard Guignard CQ holds in $\hat{z}$ if and only if MPEC-GCQ and (I^{MPEC}) hold in $\hat{z}$.*

Consider the program (13) and the corresponding set Λ. Clearly, $N_\Lambda(u) = \hat{N}_\Lambda(u)$ for any u, and so the two cases from Definition 2.1 coincide. We then speak only about the existence of a KKT vector. It is well-known that in such a case the existence of a KKT vector is a necessary first order optimality condition under standard Guignard CQ (see, e.g., [1]). Furthermore, it can be easily shown that the concept of strong stationarity is equivalent to the existence of a KKT vector for an MPEC if we conside it as a nonlinear program of the form (13) (see, e.g., [2]).

Together with this result, the statement of Theorem 4.5 follows immediately from Lemma 4.6.

Conversely, consider the following theorem, originally due to [6].

Theorem 4.7 *Let a feasible point $\hat{z}$ of the MPEC (30) be given. Further suppose that for every continuously differentiable objective function f which assumes a local minimizer at $\hat{z}$ under the constraints of the MPEC (30), there exists a KKT vector λ such that the conditions (57) for strong stationarity hold. Then Guignard CQ holds at $\hat{z}$.*

Proof. Recalling that the classical KKT conditions at $\hat{z}$ of the MPEC (30) are equivalent $\hat{z}$ being strongly stationary (see, e.g., [2]), this result immediately follows from [1, Theorem 6.3.2]. □

Together with the preceeding discussion, it follows that Theorem 4.7 and Lemma 4.6 together yield that strong stationarity is, in a sense, equivalent to the assumption pair MPEC-GCQ and ($\mathrm{I}^{\mathrm{MPEC}}$).

This shows ($\mathrm{I}^{\mathrm{MPEC}}$) to be the minimum we need to assume in addition to MPEC-GCQ in order to have strong stationarity be a necessary first order condition. Unfortunately, ($\mathrm{I}^{\mathrm{MPEC}}$) is not easy to verify. We therefore dedicate the remainder of this section to finding sufficient conditions for ($\mathrm{I}^{\mathrm{MPEC}}$).

The spadework for this has been done by Pang and Fukushima [15], albeit with a slightly different purpose in mind. We will therefore extensively fall back on their results in the following discussion.

To this end, we must first introduce the concept of nonsingularity, as used in [15, 22].

Definition 4.8 *Given the linear system*

$$Ax \leq b, \qquad Cx = d, \tag{58}$$

an inequality $a_i x \leq b_i$ *is said to be* nonsingular *if there exists a feasible solution of the system (58) which satisfies this inequality strictly. Here* a_i *denotes the* i*-th row of the matrix* A.

We will now apply nonsingularity to the linearized tangent cone $\mathcal{T}^{lin}(\hat{z})$ (see (39)). To this end we introduce two new sets: Let β^G denote the subset of β consisting of all indices $i \in \beta$ such that the inequality $\nabla G_i(\hat{z})^T d \geq 0$ is nonsingular in the system defining $\mathcal{T}^{lin}(\hat{z})$. Similarly, we denote by β^H the nonsingular set pertaining to the inequalities $\nabla H_i(\hat{z})^T d \geq 0$. Note that β^G and β^H depend on $\hat{z}$.

Using the sets β^G and β^H renders the following representation of $\mathcal{T}^{lin}(\hat{z})$ (cf. (39)):

$$\begin{aligned}\mathcal{T}^{lin}(\hat{z}) = \{d \in \mathbb{R}^n \mid\ & \nabla g_i(\hat{z})^T d \leq 0 && \forall i \in \mathcal{I}_g,\\ & \nabla h_i(\hat{z})^T d = 0 && \forall i = 1, \ldots, p,\\ & \nabla G_i(\hat{z})^T d = 0 && \forall i \in \alpha \cup \beta \backslash \beta^G,\\ & \nabla H_i(\hat{z})^T d = 0 && \forall i \in \gamma \cup \beta \backslash \beta^H,\\ & \nabla G_i(\hat{z})^T d \geq 0 && \forall i \in \beta^G,\\ & \nabla H_i(\hat{z})^T d \geq 0 && \forall i \in \beta^H\ \}.\end{aligned} \tag{59}$$

We will now also use the sets β^G and β^H to define the following assumption (A). Note that (A) is equivalent to [15, (A2)] by Lemma 1 of the same reference.

(A) Given the feasible point $\hat{z}$, there exists a partition $(\beta_1^{GH}, \beta_2^{GH}) \in \mathcal{P}(\beta^G \cap \beta^H)$ such that

$$\sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(\hat{z}) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\hat{z}) - \sum_{i \in \alpha \cup \beta} \lambda_i^G \nabla G_i(\hat{z}) - \sum_{i \in \gamma \cup \beta} \lambda_i^H \nabla H_i(\hat{z}) = 0$$
$$\Longrightarrow \begin{cases} \lambda_i^G = 0, & \forall i \in \beta_1^{GH} \\ \lambda_i^H = 0, & \forall i \in \beta_2^{GH}. \end{cases}$$

We are now able to prove that assumption (A) implies the intersection property ($\mathrm{I}^{\mathrm{MPEC}}$).

Lemma 4.9 *If a feasible point* $\hat{z}$ *of the MPEC (30) satisfies assumption (A), it also satisfies the intersection property (*I^{MPEC}*).*

Proof. By looking at the representation (39) and (37) of $\mathcal{T}^{lin}(\hat{z})$ and $\mathcal{T}^{lin}_{\mathrm{MPEC}}(\hat{z})$ respectively, it follows immediately that

$$\mathcal{T}^{lin}_{\mathrm{MPEC}}(\hat{z}) \subseteq \mathcal{T}^{lin}(\hat{z}),$$

and hence

$$\mathcal{T}^{lin}(\hat{z})^\circ \subseteq \mathcal{T}^{lin}_{\text{MPEC}}(\hat{z})^\circ.$$

Therefore, all that remains to be shown is that

$$\mathcal{T}^{lin}_{\text{MPEC}}(\hat{z})^\circ \subseteq \mathcal{T}^{lin}(\hat{z})^\circ. \tag{60}$$

Also recall that

$$\mathcal{T}^{lin}_{\text{MPEC}}(\hat{z})^\circ = \bigcap_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \mathcal{T}^{lin}_{\text{NLP}_*(\beta_1,\beta_2)}(\hat{z})^\circ \tag{61}$$

(see (53)). Now to prove (60), we take an arbitrary $v \in \mathcal{T}^{lin}_{\text{MPEC}}(\hat{z})^\circ$. By virtue of (61), we have

$$v \in \mathcal{T}^{lin}_{\text{NLP}_*(\beta_1,\beta_2)}(\hat{z})^\circ \qquad \forall(\beta_1,\beta_2) \in \mathcal{P}(\beta). \tag{62}$$

Consider the specific partition of β given by

$$\hat{\beta}_1 := \beta^H \backslash \beta_2^{GH}, \quad \hat{\beta}_2 := \beta \backslash \hat{\beta}_1. \tag{63}$$

Here $(\beta_1^{GH}, \beta_2^{GH}) \in \mathcal{P}(\beta^G \cap \beta^H)$ is a partition of $\beta^G \cap \beta^H$ that satisfies assumption (A). Note that $\beta^G \backslash \beta_1^{GH} \subseteq \hat{\beta}_2$.

Since $v \in \mathcal{T}^{lin}_{\text{NLP}_*(\hat{\beta}_1,\hat{\beta}_2)}(\hat{z})^\circ$ as well as $v \in \mathcal{T}^{lin}_{\text{NLP}_*(\hat{\beta}_2,\hat{\beta}_1)}(\hat{z})^\circ$, we can apply Lemma 3.4 to both of these cones, yielding the existence of vectors $u = (u^g, u^h, u^G, u^H)$ and $w = (w^g, w^h, w^G, w^H)$ with

$$\begin{aligned}
u_i^g &\geq 0 \quad \forall i \in \mathcal{I}_g, & w_i^g &\geq 0 \quad \forall i \in \mathcal{I}_g,\\
u_i^G &\geq 0 \quad \forall i \in \hat{\beta}_2, & w_i^G &\geq 0 \quad \forall i \in \hat{\beta}_1,\\
u_i^H &\geq 0 \quad \forall i \in \hat{\beta}_1, & w_i^H &\geq 0 \quad \forall i \in \hat{\beta}_2
\end{aligned}$$

such that

$$\begin{aligned}
v &= -\sum_{i\in\mathcal{I}_g} u_i^g \nabla g_i(\hat{z}) - \sum_{i=1}^p u_i^h \nabla h_i(\hat{z}) + \sum_{i\in\alpha\cup\beta} u_i^G \nabla G_i(\hat{z}) + \sum_{i\in\gamma\cup\beta} u_i^H \nabla H_i(\hat{z})\\
&= -\sum_{i\in\mathcal{I}_g} w_i^g \nabla g_i(\hat{z}) - \sum_{i=1}^p w_i^h \nabla h_i(\hat{z}) + \sum_{i\in\alpha\cup\beta} w_i^G \nabla G_i(\hat{z}) + \sum_{i\in\gamma\cup\beta} w_i^H \nabla H_i(\hat{z}).
\end{aligned} \tag{64}$$

The choice of the sets $\hat{\beta}_1$ and $\hat{\beta}_2$ guarantee, in particular, that

$$\begin{aligned}
u_i^G &\geq 0 \quad \forall i \in \beta^G \backslash \beta_1^{GH}, & w_i^G &\geq 0 \quad \forall i \in \beta_1^{GH},\\
u_i^H &\geq 0 \quad \forall i \in \beta^H \backslash \beta_2^{GH}, & w_i^H &\geq 0 \quad \forall i \in \beta_2^{GH}.
\end{aligned} \tag{65}$$

Taking the difference of the two representations of v in (64) and applying assumption (A) yields that

$$\begin{aligned}
u_i^G - w_i^G &= 0 \qquad \forall i \in \beta_1^{GH} \quad \text{and}\\
u_i^H - w_i^H &= 0 \qquad \forall i \in \beta_2^{GH}.
\end{aligned}$$

Together with (65), this yields that

$$u_i^G \geq 0 \quad \forall i \in \beta^G \quad \text{and} \quad u_i^H \geq 0 \quad \forall i \in \beta^H.$$

Finally, applying Lemma 3.4 to the representation (59) of $\mathcal{T}^{lin}(\hat{z})$ yields that $v \in \mathcal{T}^{lin}(\hat{z})^\circ$. This concludes the proof. □

The purpose of introducing assumption (A) was to offer a more tangible and more easily verifiable property than the intersection property (I^{MPEC}). Determining the sets β^G and β^H requires checking whether an inequality is nonsingular in the sense of Definition 4.8. To avoid having to do this, we introduce the following concept, stronger than assumption (A).

Definition 4.10 *Let a feasible point $\hat{z}$ of the MPEC (30) be given. The* partial MPEC-LICQ *is said to hold at $\hat{z}$ if the implication*

$$\sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(\hat{z}) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(\hat{z}) - \sum_{i \in \alpha \cup \beta} \lambda_i^G \nabla G_i(\hat{z}) - \sum_{i \in \gamma \cup \beta} \lambda_i^H \nabla H_i(\hat{z}) = 0$$
$$\Longrightarrow \begin{cases} \lambda_\beta^G = 0 \\ \lambda_\beta^H = 0 \end{cases}$$

holds.

Note that partial MPEC-LICQ obviously implies assumption (A) since $\beta_1^{GH} \subseteq \beta$ and $\beta_2^{GH} \subseteq \beta$. This, together with Lemmas 4.6 and 4.9, and the discussion following Lemma 4.6 gives the following corollary.

Corollary 4.11 *Let $\hat{z}$ be a local minimizer of the MPEC (30) satisfying MPEC-GCQ. Furthermore, let $\hat{z}$ satisfy any one of the following conditions*

(a) intersection property (I^{MPEC});

(b) assumption (A);

(c) partial MPEC-LICQ.

Then there exist a KKT vector $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ satisfying the relations (57), i.e. $\hat{z}$ is strongly stationary.

A few notes on Corollary 4.11 are in order. If we replace MPEC-GCQ with the stronger MPEC-ACQ, statements (b) and (c) have already been shown in [2]. If MPEC-GCQ is replaced by the still stronger assumption that all the associated nonlinear programs $\text{NLP}_*(\beta_1, \beta_2)$ satisfy the standard Abadie CQ, statement (b) has been shown in [15].

5 Final Remarks

In this paper, we have introduced new constraint qualifications and new stationarity concepts for a class of difficult optimization problems with disjunctive constraints. In particular, we have shown that specializations of these concepts result in new conditions for a local minimizer of an MPEC to be an M-stationary point and, under additional conditions, to be a strongly stationary point. We believe, however, that our general results can also be specialized to other optimization problems, and that they would give new insights into these problems as well. We leave this as a future research topic.

Acknowledgments

The third author would like to express his gratitude to the Grant Agency of the Academy of Sciences of the Czech Republic for its support with Grant IAA1075402.

References

[1] M. S. Bazaraa and C. M. Shetty, *Foundations of Optimization*, vol. 122 of Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, Heidelberg, New York, 1976.

[2] M. L. Flegel and C. Kanzow, *On the Guignard constraint qualification for mathematical programs with equilibrium constraints.* Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint 248, October 2002.

[3] ——, *A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints*, Optimization, 52 (2003), pp. 277–286.

[4] ——, *A direct proof for M-stationarity under MPEC-ACQ for mathematical programs with equilibrium constraints.* Institute of Applied Mathematics and Statistics, University of Würzburg, Preprint, May 2004.

[5] ——, *Abadie-type constraint qualification for mathematical programs with equilibrium constraints*, Journal of Optimization Theory and Applications, 124 (2005), pp. ???–???

[6] F. J. Gould and J. W. Tolle, *A necessary and sufficient qualification for constrained optimization*, SIAM Journal on Applied Mathematics, 20 (1971), pp. 164–172.

[7] M. Guignard, *Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space*, SIAM Journal on Control, 7 (1969), pp. 232–241.

[8] R. Henrion, A. Jourani, and J. Outrata, *On the calmness of a class of multifunctions*, SIAM Journal on Optimization, 13 (2002), pp. 603–618.

[9] R. Henrion and J. V. Outrata, *Calmness of constraint systems with applications*, Tech. Report 929, Weierstrass-Institut für Angewandte Analysis und Stochastik, 2004.

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

[11] B. S. Mordukhovich, *Generalized differential calculus for nonsmooth and set-valued mappings*, Journal of Mathematical Analysis and Applications, (1994), pp. 250–288.

[12] ——, *Lipschitzian stability of constraint systems and generalized equations*, Nonlinear Analysis, Theory, Methods & Applications, 22 (1994), pp. 173–206.

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

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

[15] 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.

[16] D. W. Peterson, *A review of constraint qualifications in finite-dimensional spaces*, SIAM Review, 15 (1973), pp. 639–654.

[17] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Mathematical Programming Study, 14 (1981), pp. 206–214.

[18] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317 of A Series of Comprehensive Studies in Mathematics, Springer, Berlin, Heidelberg, 1998.

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

[20] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal on Optimization, 11 (2001), pp. 918–936.

[21] P. Spellucci, *Numerische Verfahren der nichtlinearen Optimierung*, Internationale Schriftenreihe zur Numerischen Mathematik: Lehrbuch, Birkhäuser, Basel · Boston · Berlin, 1993.

[22] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions I*, vol. 163 of Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Springer-Verlag, Berlin · Heidelberg · New York, 1970.

[23] J. J. Ye, *Optimality conditions for optimization problems with complementarity constraints*, SIAM Journal on Optimization, 9 (1999), pp. 374–387.

[24] ——, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM Journal on Optimization, 10 (2000), pp. 943–962.

[25] ——, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, Journal on Mathematical Analysis and Applications, (2005), pp. ??–??

[26] J. J. Ye and X. Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Mathematics of Operations Research, 22 (1997), pp. 977–997.