STATIONARY CONDITIONS FOR MATHEMATICAL PROGRAMS WITH VANISHING CONSTRAINTS USING WEAK CONSTRAINT QUALIFICATIONS[1]

Tim Hoheisel and Christian Kanzow

Preprint 274 Dezember 2006

University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany

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

December 7, 2006

[1]This research was partially supported by the DFG (Deutsche Forschungsgemeinschaft) under grant KA1296/15-1.

Abstract. We consider a class of optimization problems that is called a *mathematical program with vanishing constraints* (MPVC for short). This class has some similarities to *mathematical programs with equilibrium constraints* (MPECs, for short), and typically violates standard constraint qualifications, hence the well-known Karush-Kuhn-Tucker conditions do not provide necessary optimality criteria. In order to obtain reasonable first order conditions under very weak assumptions, we introduce several MPVC-tailored constraint qualifications, discuss their relation, and prove an optimality condition which may be viewed as the counterpart of what is called M-stationarity in the MPEC-field.

Key Words: Mathematical programs with vanishing constraints, Mathematical programs with equilibrium constraints, Optimality conditions, Constraint qualifications, Limiting normal cone

Mathematics Subject Classification: 90C30, 90C33

1 Introduction

Consider the optimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0 \quad \forall i = 1, \dots, m, \\ & h_j(x) = 0 \quad \forall j = 1, \dots, p, \\ & H_i(x) \ge 0 \quad \forall i = 1, \dots, l, \\ & G_i(x) H_i(x) \le 0 \quad \forall i = 1, \dots, l \end{array} \tag{1}$$

with continuously differentiable functions $f, g_i, h_j, G_i, H_i : \mathbb{R}^n \to \mathbb{R}$. Following [2], we call (1) a *mathematical program with vanishing constraints*, MPVC for short. It serves as a model for many problems from structural and topology optimization, see [2] for more details. For example, vanishing constraints occur in truss topology design problems if a bar is not realized in the optimal structure so that constraints (like minimum thickness) disappear at the solution. Loosely speaking, this is reflected in the program (1) by the fact that the implicit constraint $G_i(x) \le 0$ vanishes whenever the corresponding inequality $H_i(x) \ge 0$ is active, cf. [2].

According to [2], the MPVC can, in principle, be reformulated as a *mathematical program with equilibrium constraints*, MPEC for short. Such an MPEC is an optimization problem of the form

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0 \quad \forall i = 1, \dots, m, \\ & h_j(x) = 0 \quad \forall j = 1, \dots, p, \\ & G_i(x) \ge 0, \quad H_i(x) \ge 0, \quad G_i(x) H_i(x) = 0 \quad \forall i = 1, \dots, l, \end{array}$$

see, for example, the two books [20, 27] for a general treatment and many applications of MPECs, the more recent works [10, 11, 12, 25, 26, 28, 34, 36, 37] for some more refined theoretical results, or [4, 7, 8, 13, 14, 18, 19, 30, 35] for a number of suitable methods. Therefore, it would be possible to apply the whole MPEC machinery to an MPVC. However, the reformulation of an MPVC as an MPEC given in [2] has some disadvantages. In particular, it increases the dimension and, more importantly, it involves a nonuniqueness so that isolated solutions of the MPVC are, in general, not locally unique solutions of the corresponding MPEC. Furthermore, it seems that the MPVC, though being a difficult nonconvex optimization problem, is somewhat simpler than an MPEC.

This motivates to consider the MPVC itself. So far, the literature on MPVCs is rather limited. From an application (engineering) point of view, it was considered in [1]. The first formal theoretical treatment can be found in [2]. In particular, the paper [2] shows that the MPVC typically does not satisfy standard constraint qualifications like the linear independence or Mangasarian-Fromovitz constraint qualifications. Hence standard optimization methods are likely to fail at MPVCs. The subsequent paper [17] investigates the Abadie and Guignard constraint qualifications in the context of MPVCs. It shows that also the Abadie constraint qualification is too strong an assumption for MPVCs, while the

Guignard constraint qualification holds in many situations, and some sufficient conditions are presented in [17].

While the Guignard constraint qualification implies that the usual KKT conditions are necessary optimality criteria for an MPVC, it has at least two major disadvantages from a practical point of view: First, it is difficult to see whether a given MPVC satisfies the Guignard constraint qualification. It would be nice, for example, if one could say that a certain condition holds in the case where all mappings g_i, h_j, G_i, H_i are linear, since this can be checked a priori. Second, the Guignard constraint qualification is certainly not enough in order to prove nice global or local convergence results for suitable algorithms. These algorithms typically require some LICQ- or MFCQ-type conditions, see, for example, the forthcoming paper [3].

The aim of this paper is therefore to introduce some MPVC-tailored constraint qualification, a corresponding optimality result which holds under very weak conditions, as well as several sufficient conditions for the different constraint qualifications.

To this end, we first recall in Section 2 a number of preliminary results. In Section 3, we use an MPVC-variant of the Guignard constraint qualification in order to establish a first order condition which is only slightly weaker than the usual KKT conditions. Section 4 gives some relatively simple sufficient conditions for our MPVC-tailored Guignard constraint qualification to hold. In particular, this includes the case where all functions g_i, h_j, G_i, H_i are linear. MPVC-versions of some other standard constraint qualifications are introduced and discussed in Section 5. We then close with some final remarks in Section 6.

Notation: $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_+ := [0, +\infty)$ is the set of nonnegative real numbers, and $\mathbb{R}_- := (-\infty, 0]$ are the nonpositive numbers. Given a(n index) set I, we write $\mathcal{P}(I)$ for the set of all partitions of I into two disjoint subsets of I, i.e. $(\beta_1, \beta_2) \in \mathcal{P}(I)$ if and only if $\beta_1 \cup \beta_2 = I$ and $\beta_1 \cap \beta_2 = \emptyset$. The closure of a set $X \subseteq \mathbb{R}^n$ is denoted by $cl(X)$. Furthermore, we write $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ for a multifunction or set-valued map, i.e., $\Phi(x)$ is a subset of $\mathbb{R}^n$. Its graph is defined as $\text{gph}\Phi := \{(x, y) \mid y \in \Phi(x)\}$. Following [31], $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is called a polyhedral multifunction if its graph is the union of finitely many polyhedral convex sets.

2 Preliminaries

In this section we recall some basic definitions from optimization, introduce several index sets and state some preliminary results that will be used in our subsequent analysis. We begin with the definition of the dual and polar cone.

Definition 2.1 *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a nonempty set. Then*

(a) $\mathcal{C}^* := \{v \in \mathbb{R}^n \,|\, v^T d \geq 0 \ \ \forall d \in \mathcal{C}\}$ *is the* dual cone *of* $\mathcal{C}$.

(b) $\mathcal{C}^\circ := \{v \in \mathbb{R}^n \,|\, v^T d \leq 0 \ \ \forall d \in \mathcal{C}\}$ *is the* polar cone *of* $\mathcal{C}$.

Note that $v \in \mathcal{C}^*$ if and only if $-v \in \mathcal{C}^\circ$, hence $\mathcal{C}^\circ$ is the negative of C^*.

Next consider a general optimization problem of the form

$$\begin{array}{ll} \min & \tilde{f}(x) \\ \text{s.t.} & \tilde{g}_i(x) \leq 0 \quad \forall i = 1, \ldots, \tilde{m}, \\ & \tilde{h}_j(x) = 0 \quad \forall j = 1, \ldots, \tilde{p}, \end{array} \tag{2}$$

where all functions $\tilde{f}, \tilde{g}_i, \tilde{h}_j : \mathbb{R}^{\tilde{n}} \to \mathbb{R}$ are assumed to be continuously differentiable. Let $\tilde{X}$ denote the feasible set of this optimization problem. Then the *tangent cone* at a feasible point $\tilde{x} \in \tilde{X}$ is defined by

$$\mathcal{T}(\tilde{x}) := \left\{ d \in \mathbb{R}^{\tilde{n}} \;\middle|\; \exists \{\tilde{x}^k\} \subseteq \tilde{X}, \{t_k\} \downarrow 0 : \tilde{x}^k \to \tilde{x} \text{ and } \frac{\tilde{x}^k - \tilde{x}}{t_k} \to d \right\}.$$

Furthermore, the *linearized cone* at $\tilde{x} \in \tilde{X}$ is defined by

$$\begin{aligned} \mathcal{L}(\tilde{x}) = \big\{ d \in \mathbb{R}^{\tilde{n}} \mid \nabla \tilde{g}_i(\tilde{x})^T d \leq 0 \quad & (i : \tilde{g}_i(\tilde{x}) = 0), \\ \nabla \tilde{h}_j(\tilde{x})^T d = 0 \quad & (j = 1, \ldots, \tilde{p}) \big\}. \end{aligned}$$

The following constraint qualifications are standard in optimization, see, e.g., [6, 29].

Definition 2.2 *Let $\tilde{x} \in \tilde{X}$ be a feasible point of the program* (2)*. Then*

(a) the linear independence constraint qualification *(LICQ for short) holds at $\tilde{x}$ if the gradients $\nabla \tilde{h}_j(\tilde{x})$ $(j = 1, \ldots, \tilde{p})$, $\nabla \tilde{g}_i(\tilde{x})$ $(i : \tilde{g}_i(\tilde{x}) = 0)$ are linearly independent.*

(b) the Mangasarian-Fromovitz constraint qualification *(MFCQ for short) holds at $\tilde{x}$ if the gradients $\nabla \tilde{h}_j(\tilde{x})$ $(j = 1, \ldots, \tilde{p})$ are linearly independent, and there exists a vector $\tilde{d}$ such that $\nabla \tilde{h}_j(\tilde{x})^T \tilde{d} = 0$ $(j = 1, \ldots, \tilde{p})$ and $\nabla \tilde{g}_i(\tilde{x})^T \tilde{d} < 0$ $(i : \tilde{g}_i(\tilde{x}) = 0)$.*

(c) the Abadie constraint qualification *(ACQ for short) holds at $\tilde{x}$ if $\mathcal{L}(\tilde{x}) = \mathcal{T}(\tilde{x})$.*

(d) the Guignard constraint qualification *(GCQ for short) holds at $\tilde{x}$ if $\mathcal{L}(\tilde{x})^* = \mathcal{T}(\tilde{x})^*$.*

The following implications are known to hold:

$$\text{LICQ} \Longrightarrow \text{MFCQ} \Longrightarrow \text{ACQ} \Longrightarrow \text{GCQ},$$

whereas the converse directions do not hold in general. If $\tilde{x}$ denotes a local minimum of (2) such that GCQ (or any of the other stronger constraint qualifications) is satisfied at $\tilde{x}$, then it is known that there exist certain Lagrange multipliers such that the usual KKT conditions hold. In fact, GCQ is known to be the weakest constraint qualification which guarantees that the KKT conditions are necessary optimality conditions, in a sense discussed in [15, 6].

Let us come back to our MPVC from (1). It was already noted in [2] that both LICQ and MFCQ are usually violated at an arbitary feasible point. ACQ and GCQ were then discussed in more detail in the subsequent work [17]. In order to get a better understanding of these results, let X denote the feasible set of (1), and let $x^* \in X$ be an arbitrary feasible point. Then define the index sets

$$\begin{array}{rcl} I_g & := & \{i \mid g_i(x^*) = 0\}, \\ I_+ & := & \{i \mid H_i(x^*) > 0\}, \\ I_0 & := & \{i \mid H_i(x^*) = 0\}. \end{array} \tag{3}$$

Furthermore, we divide the index set I_+ into the following subsets:

$$\begin{array}{rcl} I_{+0} & := & \{i \mid H_i(x^*) > 0, G_i(x^*) = 0\}, \\ I_{+-} & := & \{i \mid H_i(x^*) > 0, G_i(x^*) < 0\}. \end{array} \tag{4}$$

Similarly, we partition the set I_0 in the following way:

$$\begin{array}{rcl} I_{0+} & := & \{i \mid H_i(x^*) = 0, G_i(x^*) > 0\}, \\ I_{00} & := & \{i \mid H_i(x^*) = 0, G_i(x^*) = 0\}, \\ I_{0-} & := & \{i \mid H_i(x^*) = 0, G_i(x^*) < 0\}. \end{array} \tag{5}$$

Note that the first subscript indicates the sign of $H_i(x^*)$, whereas the second subscript stands for the sign of $G_i(x^*)$. Using these index sets, we can state the following representation of the linearized cone at a feasible point of our MPVC. Its elementary proof can be found in [2, Lemma 4].

Lemma 2.3 *Let $x^* \in X$ be a feasible point for (1). Then the linearized cone at x^* is given by*

$$\begin{array}{rll} \mathcal{L}(x^*) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x^*)^T d \leq 0 & (i \in I_g), \\ & \nabla h_j(x^*)^T d = 0 & (j = 1, \ldots, p), \\ & \nabla H_i(x^*)^T d = 0 & (i \in I_{0+}), \\ & \nabla H_i(x^*)^T d \geq 0 & (i \in I_{00} \cup I_{0-}), \\ & \nabla G_i(x^*)^T d \leq 0 & (i \in I_{+0})\}. \end{array} \tag{6}$$

It is also possible to get an explicit representation of the tangent cone itself. To this end, let $x^* \in X$ once again be feasible for the program (1), and let $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$ be an arbitrary partition of the index set I_{00}. Then let $NLP_*(\beta_1, \beta_2)$ denote the nonlinear program

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leq 0 & \forall i = 1, \ldots, m, \\ & h_j(x) = 0 & \forall j = 1, \ldots, p, \\ & H_i(x) = 0 & \forall i \in I_{0+}, \\ & H_i(x) \geq 0 & \forall i \in I_{0-}, \\ & G_i(x) \leq 0 & \forall i \in I_{+0}, \\ & H_i(x) \geq 0 & \forall i \in \beta_1, \\ & G_i(x) \leq 0 & \forall i \in \beta_1, \\ & H_i(x) = 0 & \forall i \in \beta_2, \\ & H_i(x) \geq 0 & \forall i \in I_+, \\ & G_i(x) \leq 0 & \forall i \in I_{+-} \cup I_{0-}. \end{array} \tag{7}$$

The tangent cone of $NLP_*(\beta_1,\beta_2)$ is denoted by $\mathcal{T}_{NLP_*(\beta_1,\beta_2)}(x^*)$, whereas $\mathcal{L}_{NLP_*(\beta_1,\beta_2)}(x^*)$ is the corresponding linearized cone. This linearized cone is given by

$$
\begin{aligned}
\mathcal{L}_{NLP_*(\beta_1,\beta_2)}(x^*) = \big\{d \in \mathbb{R}^n \mid\ & \nabla g_i(x^*)^T d \le 0 \quad (i \in I_g),\\
& \nabla h_j(x^*)^T d = 0 \quad (j = 1,\dots,p),\\
& \nabla H_i(x^*)^T d = 0 \quad (i \in I_{0+}),\\
& \nabla H_i(x^*)^T d \ge 0 \quad (i \in I_{0-}),\\
& \nabla G_i(x^*)^T d \le 0 \quad (i \in I_{+0}),\\
& \nabla H_i(x^*)^T d \ge 0 \quad (i \in \beta_1),\\
& \nabla G_i(x^*)^T d \le 0 \quad (i \in \beta_1),\\
& \nabla H_i(x^*)^T d = 0 \quad (i \in \beta_2)\big\}.
\end{aligned} \tag{8}
$$

Following [17], we also define the *MPVC-linearized cone*

$$
\begin{aligned}
\mathcal{L}_{MPVC}(x^*) := \big\{d \in \mathbb{R}^n \mid\ & \nabla g_i(x^*)^T d \le 0 && (i \in I_g),\\
& \nabla h_j(x^*)^T d = 0 && (j = 1,\dots,p),\\
& \nabla H_i(x^*)^T d = 0 && (i \in I_{0+}),\\
& \nabla H_i(x^*)^T d \ge 0 && (i \in I_{00} \cup I_{0-}),\\
& \nabla G_i(x^*)^T d \le 0 && (i \in I_{+0}),\\
& (\nabla H_i(x^*)^T d)(\nabla G_i(x^*)^T d) \le 0 && (i \in I_{00})\big\}.
\end{aligned} \tag{9}
$$

Note that $\mathcal{L}_{MPVC}(x^*)$ is, in general, a nonconvex cone, and that the only difference between $\mathcal{L}_{MPVC}(x^*)$ and the linearized cone $\mathcal{L}(x^*)$ is that we have an additional quadratic term in the last line of (9), cf. Lemma 2.3.

Using these definitions and cones, the following result was shown in [17, Lemma 2.4]. (Similar results for MPECs may be found in [20, 28, 10].)

Lemma 2.4 *Let x^* be feasible for (1). Then the following statements hold:*

(a) $\mathcal{T}(x^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1,\beta_2)}(x^*)$.

(b) $\mathcal{L}_{MPVC}(x^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \mathcal{L}_{NLP_*(\beta_1,\beta_2)}(x^*)$.

Lemma 2.4 shows that the tangent cone $\mathcal{T}(x^*)$ is usually the union of finitely many cones and, therefore, not convex in general. Since the linearized cone $\mathcal{L}(x^*)$ is polyhedral and, therefore, always closed and convex, this shows that ACQ usually does not hold for MPVCs. On the other hand, the discussion in [17] indicates that GCQ has a good chance to hold, and several sufficient conditions for GCQ to be satisfied are given in [17]. Using GCQ, we get the following result from [2, Theorem 1].

Theorem 2.5 *Let x^* be a local minimum of (1) such that GCQ holds at x^*. Then there exist Lagrange multipliers $\lambda_i \in \mathbb{R}\,(i = 1,\dots,m), \mu_j \in \mathbb{R}\,(j = 1,\dots,p), \eta_i^H, \eta_i^G \in \mathbb{R}\,(i = 1,\dots,l)$ such that*

$$
\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{j=1}^p \mu_j \nabla h_j(x^*) - \sum_{i=1}^l \eta_i^H \nabla H_i(x^*) + \sum_{i=1}^l \eta_i^G \nabla G_i(x^*) = 0 \tag{10}
$$

and

$$\begin{aligned}
&\lambda_i \geq 0, \;\; g_i(x^*) \leq 0, \;\; \lambda_i g_i(x^*) = 0 \quad \forall i = 1, \ldots, m, \\
&\eta_i^H = 0 \, (i \in I_+), \;\; \eta_i^H \geq 0 \, (i \in I_{00} \cup I_{0-}), \;\; \eta_i^H \textit{ free } (i \in I_{0+}), \\
&\eta_i^G = 0 \, (i \in I_0 \cup I_{+-}), \;\; \eta_i^G \geq 0 \, (i \in I_{+0}).
\end{aligned} \tag{11}$$

Note that (10) and (11) are the usual KKT conditions of our MPVC, cf. their derivation in [2].

Motivated by the fact that most standard constraint qualifications are violated and taking into account that GCQ is not enough in order to prove convergence of suitable algorithms or sensitivity results for MPVCs, we now introduce several MPVC-tailored variants of LICQ, MFCQ etc. To this end, let $x^* \in X$ be once again a feasible point of MPVC. Then consider the nonlinear program

$$\begin{array}{ll}
\min & f(x) \\
\text{s.t.} & g_i(x) \leq 0 \quad \forall i = 1, \ldots, m, \\
& h_j(x) = 0 \quad \forall j = 1, \ldots, p, \\
& H_i(x) = 0 \quad \forall i \in I_{0+} \cup I_{00}, \\
& H_i(x) \geq 0 \quad \forall i \in I_{0-} \cup I_+, \\
& G_i(x) \leq 0 \quad \forall i = 1, \ldots, l
\end{array} \tag{12}$$

that we call the *tightened nonlinear program*, $TNLP(x^*)$ for short, since its feasible set is obviously contained in X. (Another tightened nonlinear program in the context of MPECs was also used in [34] in order to define MPEC-tailored constraint qualifications.)

Definition 2.6 *The MPVC* (1) *satisfies MPVC-LICQ (MPVC-MFCQ) at a feasible point x^*, if $TNLP(x^*)$ satisfies LICQ (MFCQ) at x^*.*

Note that the above definition of MPVC-LICQ coincides with the definition given in [17], and it follows immediately that MPVC-LICQ implies MPVC-MFCQ, since standard LICQ always implies standard MFCQ.

As we will use MPVC-MFCQ in the subsequent analysis, we write it down explicitly using Definition 2.2: MPVC-MFCQ holds at a feasible point x^* of (1) if and only if the gradients

$$\nabla h_j(x^*) \;\, (j = 1, \ldots, p) \quad \text{and} \quad \nabla H_i(x^*) \;\, (i \in I_{0+} \cup I_{00}) \tag{13}$$

are linearly independent, and there exists a vector d such that

$$\begin{array}{ll}
\nabla g_i(x^*)^T d < 0 & \forall i \in I_g, \\
\nabla H_i(x^*)^T d > 0 & \forall i \in I_{0-}, \\
\nabla G_i(x^*)^T d < 0 & \forall i \in I_{+0} \cup I_{00}, \\
\nabla h_j(x^*)^T d = 0 & \forall j = 1, \ldots, p, \\
\nabla H_i(x^*)^T d = 0 & \forall i \in I_{0+} \cup I_{00}.
\end{array} \tag{14}$$

In order to define the MPVC-counterparts of ACQ and GCQ, we also recall the following result from [17, Corollary 2.5].

Lemma 2.7 *Given a feasible point $x^* \in X$ of* (1)*, the inclusions $\mathcal{T}(x^*) \subseteq \mathcal{L}_{MPVC}(x^*) \subseteq \mathcal{L}(x^*)$ hold.*

While the usual ACQ requires that $\mathcal{T}(x^*) = \mathcal{L}(x^*)$ which, in the context of MPVCs, was noted to be too strong due to the usual nonconvexity of the tangent cone $\mathcal{T}(x^*)$, Lemma 2.7 motivates to replace this equality by the weaker assumption $\mathcal{T}(x^*) = \mathcal{L}_{MPVC}(x^*)$, especially since the MPVC-linearized cone $\mathcal{L}_{MPVC}(x^*)$ is, in general, also nonconvex by definition. This gives the following MPVC-counterparts of ACQ and GCQ.

Definition 2.8 *Let $x^* \in X$ be a feasible point of* (1)*. Then*

(a) MPVC-ACQ *holds at x^* if $\mathcal{T}(x^*) = \mathcal{L}_{MPVC}(x^*)$.*

(b) MPVC-GCQ *holds at x^* if $\mathcal{T}(x^*)^* = \mathcal{L}_{MPVC}(x^*)^*$.*

MPVC-ACQ was introduced earlier in [17, Definition 2.6], see also [9, 12] for similar definitions in the context of MPECs and disjunctive programs. Note that MPVC-ACQ and MPVC-GCQ are not defined via the tightened nonlinear program $TNLP(x^*)$ and, in fact, are usually different from standard ACQ and standard GCQ of this tightened program.

As one might expect, the following implications hold:

$$\text{MPVC-LICQ} \Longrightarrow \text{MPVC-MFCQ} \Longrightarrow \text{MPVC-ACQ} \Longrightarrow \text{MPVC-GCQ}. \tag{15}$$

The first and third implications are direct consequences of the corresponding definitions, whereas the second implication will be shown in Theorem 4.4 below.

Using Lemma 2.7, it follows immediately from Definition 2.8 that the standard GCQ (standard ACQ) implies MPVC-GCQ (MPVC-ACQ). The converse is not true in general. This is illustrated by the following counterexample where MPVC-ACQ (and therefore also MPVC-GCQ) holds, whereas GCQ is violated and, thus, ACQ is not satisfied either.

Example 2.9 Consider the optimization problem

$$\begin{array}{ll} \min & f(x) := x_1^2 + x_2^2 \\ \text{s.t.} & H_1(x) := x_2 \geq 0, \\ & G_1(x)H_1(x) := (x_2 - x_1^3)x_2 \leq 0. \end{array} \tag{16}$$

The unique solution of (16) is $x^* := (0,0)^T$. A simple calculation (invoking Lemma 2.4, for example) shows that the tangent cone at x^* is given by $\mathcal{T}(x^*) = \{d \in \mathbb{R}^2 \mid d_2 = 0\}$. Hence, its dual cone is $\mathcal{T}(x^*)^* = \{v \in \mathbb{R}^2 \mid v_1 = 0\}$. Furthermore, the MPVC-linearized cone at x^* is given by $\mathcal{L}_{MPVC}(x^*) = \{d \in \mathbb{R}^2 \mid d_2 = 0\}$, cf. (9). Hence $\mathcal{T}(x^*) = \mathcal{L}_{MPVC}(x^*)$ and thus, MPVC-ACQ holds. In turn, the linearized cone at x^* is given by $\mathcal{L}(x^*) = \{d \in \mathbb{R}^2 \mid d_2 \geq 0\}$ and its dual is $\mathcal{L}(x^*)^* = \{v \in \mathbb{R}^2 \mid v_1 = 0,\ v_2 \geq 0\}$. Hence, we have $\mathcal{L}(x^*)^* \subsetneq \mathcal{T}(x^*)^*$ and thus, GCQ is violated. $\Diamond$

The next example shows that MPVC-GCQ has a chance to be satisfied even if MPVC-ACQ is not and thus, MPVC-GCQ happens to be a strictly weaker constraint qualification than MPVC-ACQ, cf. (15).

Example 2.10 Consider the optimization problem

$$\begin{array}{ll} \min & f(x) := x_1^2 + x_2^2 \\ \text{s.t.} & g_1(x) := -x_2 \leq 0, \\ & H_1(x) := x_2 - x_1^3 \geq 0, \\ & G_1(x)H_1(x) := x_1^3(x_2 - x_1^3) \leq 0. \end{array} \tag{17}$$

Its unique solution is $x^* := (0,0)^T$. One can easily see by geometric arguments or by Lemma 2.4 that $\mathcal{T}(x^*) = \{d \in \mathbb{R}^2 \mid d_2 \geq 0,\ d_1 d_2 \leq 0\}$. One can also compute that $\mathcal{L}_{MPVC}(x^*) = \{d \in \mathbb{R}^2 \mid d_2 \geq 0\}$. Thus, MPVC-ACQ is obviously violated, whereas MPVC-GCQ holds, since we have $\mathcal{T}(x^*)^* = \{v \in \mathbb{R}^2 \mid v_1 = 0,\ v_2 \geq 0\} = \mathcal{L}_{MPVC}(x^*)^*$. $\Diamond$

3 Optimality Conditions under MPVC-GCQ

In this section, we want to present optimality conditions under the MPVC-GCQ assumption. Since this means that GCQ does not necessarily hold, and because GCQ is the weakest constraint qualification such that the standard KKT conditions are necessary first order conditions, it follows that the optimality conditions to be derived in this section must be weaker than those from Theorem 2.5. However, we will see that we do not lose much if we replace GCQ by the MPVC-GCQ condition.

Our technique of proof is motivated by the corresponding analysis carried out in [11] for MPECs, and is based on the so-called limiting normal cone.

Definition 3.1 *Let $C \subseteq \mathbb{R}^n$ be a nonempty, closed set, and let $a \in C$. Then*

(a) the Fréchet normal cone *to C at a is defined by $\hat{N}(a, C) := (\mathcal{T}_C(a))^\circ$, i.e., the Fréchet normal cone is the polar of the tangent cone.*

(b) the limiting normal cone *to C at a is defined by*

$$N(a, C) := \big\{ \lim_{k\to\infty} w^k \mid \exists \{a^k\} \subseteq C : a^k \to a,\ w^k \in \hat{N}(a^k, C) \big\}. \tag{18}$$

The Fréchet normal cone is sometimes also called the *regular normal cone*, most notably in [33], whereas the limiting normal cone comes with a number of different names, including *normal cone*, *basic normal cone*, and *Mordukhovich normal cone* due to the many contributions of Mordukhovich in this area, see, in particular, [22, 23] for an extensive treatment and many applications of this cone. In case of a convex set C, both the Fréchet normal cone and the limiting normal cone coincide with the standard normal cone from convex analysis, cf. [32].

For the remainder, we put

$$q := |I_{00}|.$$

The following result calculates both the Fréchet and the limiting normal cone of a particular set that will play an essential role in the analysis of MPVCs.

Lemma 3.2 *Let the set*

$$C := \{(\nu,\rho) \in \mathbb{R}^q \times \mathbb{R}^q \mid \rho_i \geq 0, \rho_i\nu_i \leq 0 \ \forall i = 1,\dots,q\}$$

be given. Then the following statements hold:

(a) $\hat{N}\big((0,0),C\big) = \big\{(u,v) \mid u = 0, v \leq 0\big\}$.

(b) $N\big((0,0),C\big) = \big\{(u,v) \mid u_i \geq 0, u_iv_i = 0 \ \forall i = 1,\dots,q\big\}$.

Proof. Reordering the elements of the set C in a suitable way, we see that C can be expressed as a Cartesian product $C_1 \times \cdots \times C_q$ with closed sets $C_i := \{(\nu_i,\rho_i) \in \mathbb{R}^2 \mid \rho_i \geq 0, \rho_i\nu_i \leq 0\}$. Invoking [33, Proposition 6.41], it follows that we simply have to calculate the Fréchet and the limiting normal cones of the set $M := \big\{(\nu,\rho) \in \mathbb{R}^2 \mid \rho \geq 0, \rho\nu \leq 0\big\}$ at $(0,0) \in \mathbb{R}^2$.

(a) Because of the above remark, it suffices to show that $\hat{N}((0,0),M) = \{0\} \times \mathbb{R}_-$. It is easy to see, however, that $\mathcal{T}_M((0,0)) = M$ holds. Thus, the Fréchet normal cone is given by $\hat{N}((0,0),M) = M^\circ = \{(c,d) \in \mathbb{R}^2 \mid c = 0, \ d \leq 0\} = \{0\} \times \mathbb{R}_-$, which proves assertion (a).

(b) It suffices to show that $N((0,0),M) = \big\{(r,s) \in \mathbb{R}^2 \mid r \geq 0, rs = 0\big\}$ holds.

$'\subseteq'$: In view of the definition of the limiting normal cone in (18), we first need to figure out how the Fréchet normal cone of M at an arbitrary point $(\nu,\rho) \in M$ looks like. To this end, we consider five cases:

1) $\nu < 0, \rho > 0$: This implies $\mathcal{T}_M(\nu,\rho) = \mathbb{R}^2$. Hence $\hat{N}\big((\nu,\rho),M\big) = \{0\} \times \{0\} =: A_1$.

2) $\nu = 0, \rho > 0$: This implies $\mathcal{T}_M(\nu,\rho) = \mathbb{R}_- \times \mathbb{R}$. Hence $\hat{N}\big((\nu,\rho),M\big) = \mathbb{R}_+ \times \{0\} =: A_2$.

3) $\nu < 0, \rho = 0$: This implies $\mathcal{T}_M(\nu,\rho) = \mathbb{R} \times \mathbb{R}_+$. Hence $\hat{N}\big((\nu,\rho),M\big) = \{0\} \times \mathbb{R}_- =: A_3$.

4) $\nu > 0, \rho = 0$: This implies $\mathcal{T}_M(\nu,\rho) = \mathbb{R} \times \{0\}$. Hence $\hat{N}\big((\nu,\rho),M\big) = \{0\} \times \mathbb{R} =: A_4$.

5) $\nu = \rho = 0$: This implies $\mathcal{T}_M(\nu,\rho) = M$. Hence $\hat{N}\big((\nu,\rho),M\big) = \{0\} \times \mathbb{R}_- = A_3$.

Now let $w \in N\big((0,0),M\big)$. Then there is a sequence $\{w^k\} \to w$ such that $w^k \in \hat{N}\big((\nu_k,\rho_k),M\big)$ for all $k \in \mathbb{N}$ and some sequence $\{(\nu_k,\rho_k)\} \subseteq M$ converging to $(0,0)$. Then it follows from the above five cases that all w^k belong to the set $A_1 \cup A_2 \cup A_3 \cup A_4 = A_2 \cup A_4 = \mathbb{R}_+ \times \{0\} \cup \{0\} \times \mathbb{R} = \{(r,s) \in \mathbb{R}^2 \mid r \geq 0, rs = 0\}$. Since this set is closed, the limiting element w also belongs to this set. This gives the desired inclusion.

$'\supseteq'$: Let $(a,b) \in \big\{(r,s) \in \mathbb{R}^2 \mid r \geq 0, rs = 0\big\}$. First, we consider the case $a > 0$ (hence $b = 0$). In order to prove $(a,b) \in N\big((0,0),M\big)$, we define the sequence $\{(u_k,v_k)\} \subseteq M$ by putting $u_k := 0$ and selecting v_k such that we have $v_k \downarrow 0$. Then we are in the above second case for all $k \in \mathbb{N}$. Consequently, we have $(a_k,b_k) := (a,0) \in \hat{N}\big((u_k,v_k),M\big)$ for all

$k \in \mathbb{N}$ which proves the desired inclusion. Next, consider the case $a = 0$ (and b arbitrary). Then let $\{(u_k, v_k)\} \subseteq M$ be any sequence with $u_k \downarrow 0$ and $v_k = 0$ for all $k \in \mathbb{N}$. Then the above fourth case shows that $\hat{N}\big((u_k, v_k), M\big) = \{0\} \times \mathbb{R}$. Defining $(a_k, b_k) := (0, b)$ for all $k \in \mathbb{R}$, it therefore follows that $(a_k, b_k) \in \hat{N}\big((u_k, v_k), M\big)$ for all $k \in \mathbb{N}$, and this gives the desired inclusion also in this case. $\square$

Now let $\mathcal{D}_1$ and $\mathcal{D}_2$ denote the following sets:

$$\mathcal{D}_1 := \Big\{(d, \nu, \rho) \in \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \,\Big|\, \begin{array}{ll} \nabla g_i(x^*)^T d \leq 0 & (i \in I_g), \\ \nabla h_j(x^*)^T d = 0 & (j = 1, \ldots, p), \\ \nabla H_i(x^*)^T d = 0 & (i \in I_{0+}), \\ \nabla H_i(x^*)^T d \geq 0 & (i \in I_{0-}), \\ \nabla G_i(x^*)^T d \leq 0 & (i \in I_{+0}), \\ \nabla G_i(x^*)^T d - \nu_i = 0 & (i \in I_{00}), \\ \nabla H_i(x^*)^T d - \rho_i = 0 & (i \in I_{00}) \Big\}. \end{array} \tag{19}$$

and

$$\mathcal{D}_2 := \Big\{(d, \nu, \rho) \in \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^q \,\Big|\, \rho_i \geq 0,\ \nu_i \rho_i \leq 0\ \forall i = 1, \ldots, q\Big\}. \tag{20}$$

These two sets will be crucial for the proof of our upcoming main result.

Lemma 3.3 *Let the multifunction $\Phi : \mathbb{R}^{n+2q} \rightrightarrows \mathbb{R}^{n+2q}$ be given by*

$$\Phi(v) := \Big\{w \in \mathcal{D}_1 \mid v + w \in \mathcal{D}_2\Big\}. \tag{21}$$

Then Φ is a polyhedral multifunction.

Proof. Since the graph of Φ may be expressed as

$$\mathrm{gph}\Phi = \Big\{(d^v, \nu^v, \rho^v, d^w, \nu^w, \rho^w) \mid \begin{array}{ll} \nabla g_i(x^*)^T d^w \leq 0 & (i \in I_g), \\ \nabla h_j(x^*)^T d^w = 0 & (j = 1, \ldots, p), \\ \nabla H_i(x^*)^T d^w = 0 & (i \in I_{0+}), \\ \nabla H_i(x^*)^T d^w \geq 0 & (i \in I_{0-}), \\ \nabla G_i(x^*)^T d^w \leq 0 & (i \in I_{+0}), \\ \nabla G_i(x^*)^T d^w - \nu_i^w = 0 & (i \in I_{00}), \\ \nabla H_i(x^*)^T d^w - \rho_i^w = 0 & (i \in I_{00}), \\ \rho^v + \rho^w \geq 0, & \\ (\rho_i^v + \rho_i^w)(\nu_i^v + \nu_i^w) \leq 0 & (i = 1, \ldots, q) \Big\} \end{array}$$

$$= \bigcup_{(\alpha_1,\alpha_2)\in\mathcal{P}(\{1,\ldots,q\})} \Big\{(d^v,\nu^v,\rho^v,d^w,\nu^w,\rho^w) \mid \begin{array}{ll} \nabla g_i(x^*)^T d^w \leq 0 & (i \in I_g), \\ \nabla h_j(x^*)^T d^w = 0 & (j = 1,\ldots,p), \\ \nabla H_i(x^*)^T d^w = 0 & (i \in I_{0+}), \\ \nabla H_i(x^*)^T d^w \geq 0 & (i \in I_{0-}), \\ \nabla G_i(x^*)^T d^w \leq 0 & (i \in I_{+0}), \\ \nabla G_i(x^*)^T d^w - \nu_i^w = 0 & (i \in I_{00}), \\ \nabla H_i(x^*)^T d^w - \rho_i^w = 0 & (i \in I_{00}), \\ \rho^v_{\alpha_1} + \rho^w_{\alpha_1} \geq 0, & \\ \rho^v_{\alpha_2} + \rho^w_{\alpha_2} = 0, & \\ \nu^v_{\alpha_1} + \nu^w_{\alpha_1} \leq 0\Big\}, & \end{array}$$

gphΦ is the union of finitely many polyhedral convex sets. Hence the assertion follows. □

The previous results allow us to state the following main result of this section.

Theorem 3.4 *Let x^* be a local minimizer of* (1) *such that MPVC-GCQ holds. Then there exist scalars $\lambda_i \in \mathbb{R}\,(i = 1,\ldots,m), \mu_j \in \mathbb{R}\,(j = 1,\ldots,p), \eta_i^H, \eta_i^G \in \mathbb{R}\,(i = 1,\ldots,l)$ such that*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{j=1}^p \mu_j \nabla h_j(x^*) - \sum_{i=1}^l \eta_i^H \nabla H_i(x^*) + \sum_{i=1}^l \eta_i^G \nabla G_i(x^*) = 0 \quad (22)$$

and

$$\begin{aligned} &\lambda_i \geq 0, \;\; g_i(x^*) \leq 0, \;\; \lambda_i g_i(x^*) = 0 \quad \forall i = 1,\ldots,m, \\ &\eta_i^H = 0\,(i \in I_+), \;\; \eta_i^H \geq 0\,(i \in I_{0-}), \;\; \eta_i^H \textit{ free } (i \in I_{0+}), \\ &\eta_i^G = 0\,(i \in I_{+-} \cup I_{0-} \cup I_{0+}), \;\; \eta_i^G \geq 0\,(i \in I_{+0} \cup I_{00}), \\ &\eta_i^H \eta_i^G = 0\,(i \in I_{00}). \end{aligned} \quad (23)$$

Proof. Since x^* is a local minimizer of (1), standard results from optimization imply that $\nabla f(x^*)^T d \geq 0$ for all $d \in \mathcal{T}(x^*)$, see, e.g., [24]. Since MPVC-GCQ holds at x^*, it therefore follows that $\nabla f(x^*) \in \mathcal{T}(x^*)^* = \mathcal{L}_{MPVC}(x^*)^*$. Consequently, we have $\nabla f(x^*)^T d \geq 0$ for all $d \in \mathcal{L}_{MPVC}(x^*)$. This is equivalent to $d^* = 0$ being a minimizer of

$$\min_d \; \nabla f(x^*)^T d \quad \text{s.t.} \quad d \in \mathcal{L}_{MPVC}(x^*). \quad (24)$$

Now, $d^* = 0$ being a minimizer of (24) is equivalent to $(d^*, \nu^*, \rho^*) := (0,0,0)$ being a minimizer of

$$\min_{d,\nu,\rho} \nabla f(x^*)^T d \quad \text{s.t.} \quad (d,\nu,\rho) \in \mathcal{D} := \mathcal{D}_1 \cap \mathcal{D}_2 \quad (25)$$

with $\mathcal{D}_1$ and $\mathcal{D}_2$ as defined in (19) and (20), respectively. Once more, since $(0,0,0)$ is a minimizer of (25), we have $\big(\nabla f(x^*)^T, 0, 0\big)^T w \geq 0$ for all $w \in \mathcal{T}\big((0,0,0),\mathcal{D}\big)$, where $\mathcal{T}\big((0,0,0),\mathcal{D}\big)$ denotes the tangent cone of $\mathcal{D}$ at the origin. Using [33, Proposition 6.5], this implies

$$\big(-\nabla f(x^*)^T, 0, 0\big)^T \in \mathcal{T}\big((0,0,0),\mathcal{D}\big)^\circ = \hat{N}\big((0,0,0),\mathcal{D}\big) \subseteq N\big((0,0,0),\mathcal{D}\big). \quad (26)$$

Since Φ, as defined in (21), is a polyhedral multifunction by Lemma 3.3, [31, Proposition 1] may be invoked to show that Φ is locally upper Lipschitz at every point $v \in \mathbb{R}^{n+2q}$. In particular, it is therefore calm at every $(v,w) \in \mathrm{gph}\Phi$ in the sense of [16]. Invoking [16, Corollary 4.2], we see that (26) implies

$$\big(-\nabla f(x^*)^T, 0, 0\big)^T \in N\big((0,0,0), \mathcal{D}_1\big) + N\big((0,0,0), \mathcal{D}_2\big).$$

Since $\mathcal{D}_1$ is polyhedral convex, the limiting normal cone of $\mathcal{D}_1$ is equal to the standard normal cone from convex analysis, and standard results on the representation of this normal cone (see, e.g., [6, 11]) yield the existence of certain vectors $\lambda, \mu, \mu^H, \mu^G$ such that

$$\begin{aligned}\begin{pmatrix} -\nabla f(x^*) \\ 0 \\ 0 \end{pmatrix} \in\ & \sum_{i \in I_g} \lambda_i \begin{pmatrix} \nabla g_i(x^*) \\ 0 \\ 0 \end{pmatrix} + \sum_{j=1}^{p} \mu_j \begin{pmatrix} \nabla h_j(x^*) \\ 0 \\ 0 \end{pmatrix} \\ & - \sum_{i \in I_{0+} \cup I_{0-}} \mu_i^H \begin{pmatrix} \nabla H_i(x^*) \\ 0 \\ 0 \end{pmatrix} + \sum_{i \in I_{+0}} \mu_i^G \begin{pmatrix} \nabla G_i(x^*) \\ 0 \\ 0 \end{pmatrix} \\ & - \sum_{i \in I_{00}} \mu_i^H \begin{pmatrix} \nabla H_i(x^*) \\ 0 \\ -e^i \end{pmatrix} + \sum_{i \in I_{00}} \mu_i^G \begin{pmatrix} \nabla G_i(x^*) \\ -e^i \\ 0 \end{pmatrix} \\ & + N\big((0,0,0), \mathcal{D}_2\big)\end{aligned} \tag{27}$$

with

$$\lambda_i \geq 0 \ (i \in I_g), \quad \mu_i^H \geq 0 \ (i \in I_{0-}), \quad \mu_i^G \geq 0 \ (i \in I_{+0}), \tag{28}$$

where e^i denotes the compatible unit vector in $\mathbb{R}^q$.

Using [33, Proposition 6.41] and Lemma 3.2, we get the following explicit representation of the remaining normal cone:

$$\begin{aligned} N((0,0,0), \mathcal{D}_2) &= N(0, \mathbb{R}^n) \times N\big((0,0), \{(\nu, \rho) \mid \rho_i \geq 0, \rho_i \nu_i \leq 0 \ \forall i = 1, \ldots, q\}\big) \\ &= \{0\}^n \times \big\{(u,v) \mid u_i \geq 0, u_i v_i = 0 \ \forall i = 1, \ldots, q\big\}. \end{aligned}$$

Applying the above equality to (27) yields

$$\mu_i^G \geq 0 \ \wedge \ \mu_i^G \mu_i^H = 0 \quad \forall i \in I_{00}. \tag{29}$$

Putting $\lambda_i := 0$ for $i \notin I_g$, $\eta_i^H := 0$ for $i \in I_+$, $\eta_i^G := 0$ for $i \in I_{0+} \cup I_{0-} \cup I_{+-}$, $\eta_i^G := \mu_i^G$ and $\eta_i^H := \mu_i^H$ for all other indices, we see from (28), (29) and the first row of (27) that (22) and (23) are satisfied. $\square$

Motivated by a corresponding terminology for MPECs (where it was introduced in [35]) and based on the fact that the optimality conditions (22), (23) from Theorem 3.4 were derived using the Mordukhovich normal cone, we call them the *M-stationary conditions* of an MPVC. They are slightly weaker than the standard KKT conditions (10), (11) from

Theorem 2.5. In fact, in the latter we have $\eta_i^H \geq 0$ and $\eta_i^G = 0$ for all $i \in I_{00}$, whereas now we only have $\eta_i^G \geq 0$ and $\eta_i^H \eta_i^G = 0$ for all $i \in I_{00}$. Geometrically, this means that, for every index $i \in I_{00}$, the pair (η_i^G, η_i^H) lies on the nonnegative η_i^H-axis for a KKT point, where it belongs to the union of the η_i^H-axis and the nonnegative η_i^G-axis for an M-stationary point.

4 Sufficient Conditions for MPVC-ACQ

It is the goal of this section to provide some relatively simple sufficient conditions for MPVC-ACQ. Thus, we automatically obtain sufficient conditions for MPVC-GCQ, too, since MPVC-ACQ implies MPVC-GCQ. Some more refined sufficient conditions for MPVC-ACQ will be discussed in the next section.

The first result of this section is an immediate consequence of Lemma 2.4 and states that MPVC-ACQ holds if ACQ is satisfied for $NLP_*(\beta_1, \beta_2)$, for any $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$.

Lemma 4.1 *Let x^* be feasible for* (1). *If, for any partition $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, the Abadie constraint qualification holds for $NLP_*(\beta_1, \beta_2)$, then MPVC-ACQ holds for* (1).

Proof. Using our assumption and Lemma 2.4, we obtain

$$\mathcal{T}(x^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \mathcal{T}_{NLP_*(\beta_1,\beta_2)}(x^*) = \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(I_{00})} \mathcal{L}_{NLP_*(\beta_1,\beta_2)}(x^*) = \mathcal{L}_{MPVC}(x^*),$$

which gives the assertion. $\square$

Note that the assumption in Lemma 4.1 is equivalent to assumption (A1) in [17]. An immediate consequence is the following theorem.

Theorem 4.2 *Let x^* be feasible for* (1) *and assume that all functions g_i, h_j, G_i, and H_i are linear. Then MPVC-ACQ holds at x^*.*

Proof. Since all constraints of $NLP_*(\beta_1, \beta_2)$ are linear for any $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, it follows from a well-known result in optimization that ACQ holds for each $NLP_*(\beta_1, \beta_2)$, $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Lemma 4.1 therefore gives the desired result. $\square$

To clarify the relationship between MPVC-MFCQ and MPVC-ACQ, we need the following auxiliary result.

Lemma 4.3 *Let x^* be feasible for* (1) *such that MPVC-MFCQ is satisfied. Then, for any $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$, MFCQ holds at x^* for $NLP_*(\beta_1, \beta_2)$.*

Proof. Let $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$ be given arbitrarily. We have to show that the gradients

$$\begin{array}{ll} \nabla h_j(x^*) & \forall j = 1, \ldots, p, \\ \nabla H_i(x^*) & \forall i \in I_{0+} \cup \beta_2 \end{array} \tag{30}$$

are linearly independent, and that there exists a vector $\tilde{d}$ such that

$$\begin{array}{ll} \nabla g_i(x^*)^T\tilde{d} < 0 & \forall i \in I_g, \\ \nabla H_i(x^*)^T\tilde{d} > 0 & \forall i \in I_{0-} \cup \beta_1, \\ \nabla G_i(x^*)^T\tilde{d} < 0 & \forall i \in I_{+0} \cup \beta_1, \\ \nabla h_j(x^*)^T\tilde{d} = 0 & \forall j = 1, \dots, p, \\ \nabla H_i(x^*)^T\tilde{d} = 0 & \forall i \in I_{0+} \cup \beta_2. \end{array} \tag{31}$$

The linear independence of (30) is trivially satisfied, as we have $\beta_2 \subseteq I_{00}$ and MPVC-MFCQ holds, cf. (13).

Since the occurring gradients are linearly independent, the linear system

$$\begin{pmatrix} \nabla h_j(x^*)^T & (j = 1, \dots, p) \\ \nabla H_i(x^*)^T & (i \in I_{0+} \cup \beta_2) \\ \nabla H_i(x^*)^T & (i \in \beta_1) \end{pmatrix} d = \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix}$$

has a solution $\hat{d}$, where $e \in \mathbb{R}^{|\beta_1|}$ denotes the vector of all ones. Now, choose d such that (14) is satisfied, and put

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

Then, for all $\delta > 0$, we have

$$\begin{array}{ll} \nabla h_j(x^*)^T d(\delta) = 0 & \forall j = 1, \dots, p, \\ \nabla H_i(x^*)^T d(\delta) = 0 & \forall i \in I_{0+} \cup \beta_2, \\ \nabla H_i(x^*)^T d(\delta) > 0 & \forall i \in \beta_1. \end{array}$$

Furthermore, for $\delta > 0$ sufficiently small, we have

$$\begin{array}{ll} \nabla g_i(x^*)^T d(\delta) < 0 & \forall i \in I_g, \\ \nabla H_i(x^*)^T d(\delta) > 0 & \forall i \in I_{0-}, \\ \nabla G_i(x^*)^T d(\delta) < 0 & \forall i \in \beta_1 \cup I_{+0}. \end{array}$$

This concludes the proof. □

The next theorem states that MPVC-MFCQ is a sufficient condition for MPVC-ACQ. An immediate consequence of this result is the chain of implications already given in (15).

Theorem 4.4 *Let x^* be feasible for* (1) *such that MPVC-MFCQ holds. Then MPVC-ACQ is satisfied.*

Proof. Lemma 4.3 shows that standard MFCQ holds for every program $NLP_*(\beta_1, \beta_2)$ with $(\beta_1, \beta_2) \in \mathcal{P}(I_{00})$. Hence standard ACQ holds for each program $NLP_*(\beta_1, \beta_2)$. The statement therefore follows from Lemma 4.1. □

In particular, it follows from Theorem 4.4 and (15) that MPVC-LICQ implies MPVC-ACQ.

5 More MPVC-tailored Constraint Qualifications

The goal of this section is to provide further MPVC-tailored constraint qualifications and to investigate their relationships. The analysis is motivated by similar considerations for MPECs in [37] and bilevel programs in [38], for example, see also the treatment for standard optimization problems in [21] and elsewhere.

In order to state these constraint qualifications, we first recall the definition of two well-known cones from, e.g., [5]. Given a feasible point $x \in X$ of (1), we call

$$\begin{aligned} \mathcal{A}(x) := \big\{ d \in \mathbb{R}^n \mid \exists \delta > 0, \exists \alpha : \mathbb{R} \to \mathbb{R}^n : \; \alpha(\tau) \in X \; \forall \tau \in (0, \delta), \\ \alpha(0) = x, \; \lim_{\tau \downarrow 0} \frac{\alpha(\tau) - \alpha(0)}{\tau} = d \big\} \end{aligned} \tag{32}$$

the *cone of attainable directions* of X at x, and

$$\mathcal{F}(x) := \big\{ d \in \mathbb{R}^n \setminus \{0\} \mid \exists \delta > 0 : \; x + \tau d \in X \; \forall \tau \in (0, \delta) \big\} \tag{33}$$

the *cone of feasible directions* of X at x. Then the following chain of inclusions

$$cl\big(\mathcal{F}(x)\big) \subseteq cl\big(\mathcal{A}(x)\big) \subseteq \mathcal{T}(x) \subseteq \mathcal{L}_{MPVC}(x) \subseteq \mathcal{L}(x) \tag{34}$$

holds, cf. [5, Lemma 5.2.1] and Lemma 2.7. Now, the standard *Zangwill constraint qualification* (ZCQ for short) is said to hold at x if $\mathcal{L}(x) \subseteq cl\big(\mathcal{F}(x)\big)$, and the standard *Kuhn-Tucker constraint qualification* (KTCQ for short) is satisfied at x if $\mathcal{L}(x) \subseteq cl\big(\mathcal{A}(x)\big)$. Using (34), we immediately see that

$$\text{ZCQ} \Longrightarrow \text{KTCQ} \Longrightarrow \text{ACQ}. \tag{35}$$

Since ACQ is already too strong for MPVCs, we therefore cannot expect ZCQ or KTCQ to hold for our program (1). However, similar to the definition of MPVC-ACQ and MPVC-GCQ, we obtain MPVC-tailored variants of these constraint qualifications by using the MPVC-linearized cone instead of the linearized cone itself.

Definition 5.1 *Let x^* be feasible for* (1). *Then*

(a) the MPVC-ZCQ *holds at x^* if $\mathcal{L}_{MPVC}(x^*) \subseteq cl(\mathcal{F}(x^*))$.*

(b) the MPVC-KTCQ *holds at x^* if $\mathcal{L}_{MPVC}(x^*) \subseteq cl(\mathcal{A}(x^*))$.*

An immediate consequence of the above definition and (34) are the implications

$$\text{MPVC-ZCQ} \Longrightarrow \text{MPVC-KTCQ} \Longrightarrow \text{MPVC-ACQ},$$

which are the counterparts of (35). Moreover, standard ZCQ (standard KTCQ) implies MPVC-ZCQ (MPVC-KTCQ).

In classical optimization, the case of a convex program, where all equality constraints are supposed to be (affine) linear and all the inequality constraints (as well as the objective function) are supposed to be convex, is often considered. Very popular constraint qualifications to be used in this context are the Slater-type constraint qualifications (SCQ for short), see, for example, [21].

Since the G_iH_i-restrictions in (1), being a product of two non-constant functions, are very likely to be nonconvex, these standard Slater-type constraint qualifications will rather often fail to hold in the case of an MPVC. Thus, it is our goal to find suitable variants for MPVCs. To this end, let us introduce the following terminology.

Definition 5.2 *The program* (1) *is called* MPVC-convex *if the functions h_j, G_i, H_i are (affine) linear and all components g_i are convex.*

The next definition states the MPVC-tailored versions of two Slater-type constraint qualifications.

Definition 5.3 *Let the program* (1) *be MPVC-convex. Then this program is said to satisfy*

(a) weak MPVC-SCQ *or* MPVC-WSCQ *at a feasible point x^* if there exists a vector $\hat{x}$ such that*

$$\begin{array}{ll} g_i(\hat{x}) < 0 & \forall i \in I_g, \\ h_j(\hat{x}) = 0 & \forall j = 1, \dots, p, \\ G_i(\hat{x}) \leq 0 & \forall i \in I_{+0} \cup I_{00}, \\ H_i(\hat{x}) = 0 & \forall i \in I_{0+} \cup I_{00}, \\ H_i(\hat{x}) \geq 0 & \forall i \in I_{0-}. \end{array} \tag{36}$$

(b) MPVC-SCQ *if there exists a vector $\hat{x}$ such that*

$$\begin{array}{ll} g_i(\hat{x}) < 0 & \forall i = 1, \dots, m, \\ h_j(\hat{x}) = 0 & \forall j = 1, \dots, p, \\ G_i(\hat{x}) \leq 0 & \forall i = 1, \dots, l, \\ H_i(\hat{x}) = 0 & \forall i = 1, \dots, l. \end{array}$$

Note that MPVC-SCQ obviously implies MPVC-WSCQ, whereas MPVC-SCQ has the advantage that it can be checked without knowledge of the feasible point x^*. With these definitions, we are now in a position to state the next theorem which tells us that MPVC-WSCQ implies MPVC-ZCQ and thus, in view of our previous results, we also see that MPVC-WSCQ and MPVC-SCQ are sufficient conditions for MPVC-ACQ.

Theorem 5.4 *Let x^* be feasible for the MPVC-convex program such that MPVC-WSCQ is satisfied. Then MPVC-ZCQ holds at x^*.*

Proof. Let $d \in \mathcal{L}_{MPVC}(x^*)$. We need to show that there is a sequence $d^k \in \mathcal{F}(x^*)$ such that d^k converges to d. To this end, choose $\hat{x}$ satisfying (36), a positive sequence $\{t_k\} \downarrow 0$, and put $d^k := d + t_k\hat{d} := d + t_k(\hat{x} - x^*)$. Then d^k obviously converges to d.

Now, let k be fixed for the time being. In order to see that d^k is an element of $\mathcal{F}(x^*)$, we need to prove that $x^* + \tau d^k$ is feasible for (1) for all $\tau > 0$ sufficiently small.

First of all, note that, since the functions g_i ($i = 1, \dots, l$) are convex, we have

$$\nabla g_i(x^*)^T \hat{d} = \nabla g_i(x^*)^T (\hat{x} - x^*) \le g_i(\hat{x}) - g_i(x^*) < 0 \quad \forall i \in I_g. \tag{37}$$

Furthermore, we also have

$$\nabla g_i(x^*)^T d \le 0 \quad \forall i \in I_g, \tag{38}$$

since d is an element of $\mathcal{L}_{MPVC}(x^*)$. Together, (37) and (38) imply

$$\nabla g_i(x^*)^T d^k < 0 \quad \forall i \in I_g.$$

Invoking Taylor's formula, it follows that, for all $\tau > 0$ sufficiently small, we have

$$g_i(x^* + \tau d^k) = g_i(x^*) + \tau \nabla g_i(x^*)^T d^k + o(\tau) = \tau \nabla g_i(x^*)^T d^k + o(\tau) < 0 \quad \forall i \in I_g. \tag{39}$$

By continuity, we also have $g_i(x^* + \tau d^k) < 0$ for all $i \notin I_g$ and all $\tau > 0$ sufficiently small, which together with (39) yields

$$g_i(x^* + \tau d^k) \le 0 \quad \forall i = 1, \dots, l, \tag{40}$$

for all $\tau > 0$ sufficiently small. In order to check the remaining constraints, we put $u := \tau t_k$ and note that $u > 0$ becomes arbitrarily small for $\tau \to 0$. The definition of u implies $x^* + \tau d^k = (1-u)x^* + u\hat{x} + \tau d$. Invoking the linearity of the respective functions and exploiting the fact that $d \in \mathcal{L}_{MPVC}(x^*)$, we thus obtain, for $\tau > 0$ sufficiently small,

$$\begin{aligned} h_j(x^* + \tau d^k) &= h_j((1-u)x^* + u\hat{x}) + \tau \underbrace{\nabla h_j(x^*)^T d}_{=0} \\ &= (1-u) \underbrace{h_j(x^*)}_{=0} + u \underbrace{h_j(\hat{x})}_{=0} = 0 \quad \forall j = 1, \dots, p. \end{aligned} \tag{41}$$

Similarly, we can compute that, for $\tau > 0$ sufficiently small, we have

$$\begin{aligned} H_i(x^* + \tau d^k) &= H_i((1-u)x^* + u\hat{x}) + \tau \nabla H_i(x^*)^T d \\ &= (1-u)H_i(x^*) + uH_i(\hat{x}) + \tau \nabla H_i(x^*)^T d \begin{cases} > 0, & \text{if } i \in I_+, \\ = 0, & \text{if } i \in I_{0+}, \\ \ge 0, & \text{if } i \in I_{0-} \cup I_{00}, \end{cases} \end{aligned} \tag{42}$$

which, in particular, implies

$$H_i(x^* + \tau d^k) \ge 0 \quad \forall i = 1, \dots, l. \tag{43}$$

Furthermore, for $\tau > 0$ sufficiently small, we also have

$$G_i(x^* + \tau d^k) = (1-u)G_i(x^*) + uG_i(\hat{x}) + \tau \nabla G_i(x^*)^T d \begin{cases} < 0, & \text{if } i \in I_{+-} \cup I_{0-}, \\ > 0, & \text{if } i \in I_{0+}, \\ \le 0, & \text{if } i \in I_{+0}. \end{cases} \tag{44}$$

Together, we obtain $G_i(x^* + \tau d_k)H_i(x^* + \tau d_k) \leq 0$ for all $i \in \{1, \dots, l\} \setminus I_{00}$ and for all $\tau > 0$ sufficiently small. Thus, it remains to check the G_iH_i-restriction for $i \in I_{00}$. First, let $i \in I_{00}$ such that $\nabla G_i(x^*)^T d > 0$. Since we have $d \in \mathcal{L}_{MPVC}(x^*)$, this implies $\nabla H_i(x^*)^T d = 0$ and thus $H_i(x^* + \tau d^k) = 0$, in view of (42), that is we have $G_i(x^* + \tau d_k)H_i(x^* + \tau d_k) = 0$. Second, let $i \in I_{00}$ such that $\nabla G_i(x^*)^T d \leq 0$. Then we have $G_i(x^* + \tau d^k) \leq 0$ in view of (44), and thus $G_i(x^* + \tau d_k)H_i(x^* + \tau d_k) \leq 0$, which concludes the proof. $\square$

In the figure below, the relationships among the different MPVC-tailored constraint qualifications are summarized:

$$
\begin{tikzcd}[column sep=small]
\text{MPVC-(W)SCQ} \arrow[d, Rightarrow] & & \\
\text{MPVC-ZCQ} \arrow[d, Rightarrow] & & \text{MPVC-LICQ} \arrow[d, Rightarrow] \\
\text{MPVC-KTCQ} \arrow[dr, Rightarrow] & & \text{MPVC-MFCQ} \arrow[dl, Rightarrow] \\
\text{MPVC-affine} \arrow[r, Rightarrow] & \text{MPVC-ACQ} \arrow[d, Rightarrow] & \\
& \text{MPVC-GCQ} &
\end{tikzcd}
$$

Here MPVC-affine refers to the situation from Theorem 4.2 where all mappings g_i, h_j, G_i, H_i are linear. The above figure summarizes the results which were actually shown in this paper. Some other implications also hold, for example, it was shown in [17] that MPVC-LICQ is a sufficient condition for standard Guignard CQ and, therefore, stronger stationary conditions hold under MPVC-LICQ. In general, however, these stronger stationary conditions do not hold under any of the other MPVC-tailored CQs.

6 Final Remarks

Motivated by the fact that most standard constraint qualifications are violated for mathematical programs with vanishing constraints, we introduced several new constraint qualifications which take the particular structure of the program into account. The weakest among these new constraint qualifications still guarantees an optimality condition to hold at a local minimum which is only slightly weaker than the standard KKT conditions. Several sufficient conditions and other constraint qualifications are also presented. In particular, some of these sufficient conditions are very simple and can be checked a priori without knowledge of the particular solution point.

In our future work, we plan to exploit some of the constraint qualifications introduced here in order to get second order conditions for mathematical programs with vanishing constraints. Moreover, we would like to see under which additional assumptions our necessary optimality condition is also a sufficient condition for local optimality under convexity-type assumptions.

References

[1] W. Achtziger: *On non-standard problem formulations in structural optimization.* In: J. Herskovits, S. Mazorche, and A. Canelas (eds.): *Proceedings of the Sixth World Congress of Structural and Multidisciplinary Optimization (WCSMO-6).* CD-ROM, ISBN 85-285-0070-5, 2005, Paper 1481, 1–6.

[2] W. Achtziger and C. Kanzow: *Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications.* Mathematical Programming, to appear.

[3] W. Achtziger, T. Hoheisel, and C. Kanzow: *A smoothing-regularization approach to mathematical programs with vanishing constraints.* Preprint, Institute of Mathematics, University of Würzburg, forthcoming.

[4] M. Anitescu: *Global convergence of an elastic mode approach for a class of mathematical programs with equilibrium constraints.* SIAM Journal on Optimization, 16 (2005), pp. 120–145.

[5] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty: *Nonlinear Programming. Theory and Algorithms.* John Wiley & Sons, 1993 (second edition).

[6] M. S. Bazaraa and C. M. Shetty: *Foundations of Optimization.* Lecture Notes in Economics and Mathematical Systems, Vol. 122, Springer, Berlin, Heidelberg, New York, 1976.

[7] V. DeMiguel, M.P. Friedlander, F.J. Nogales, and S. Scholtes: *A two-sided relaxation scheme for mathematical programs with equilibrium constraints.* SIAM Journal on Optimization, 16 (2005), pp. 587–609.

[8] F. Facchinei, H. Jiang, and L. Qi: *A smoothing method for mathematical programs with equilibrium constraints.* Mathematical Programming, 85 (1999), pp. 107–134.

[9] M. L. Flegel and C. Kanzow: *Abadie-type constraint qualification for mathematical programs with equilibrium constraints.* Journal of Optimization Theory and Applications, 124 (2005), pp. 595–614.

[10] M. L. Flegel and C. Kanzow: *On the Guignard constraint qualification for mathematical programs with equilibrium constraints.* Optimization, 54 (2005), pp. 517–534.

[11] M. L. Flegel and C. Kanzow: *A direct proof for M-stationarity under MPEC-ACQ for mathematical programs with equilibrium constraints.* In: S. Dempe and V. Kalashnikov (eds.): *Optimization with Multivalued Mappings: Theory, Applications and Algorithms.* Springer–Verlag, New York, 2006, pp. 111–122.

[12] M. L. Flegel, C. Kanzow, and J. V. Outrata: *Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints.* Set-Valued Analysis, to appear.

[13] R. Fletcher and S. Leyffer: *Solving mathematical program with complementarity constraints as nonlinear programs.* Optimization Methods and Software, 19 (2004), pp. 15–40.

[14] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes: *Local convergence of SQP methods for mathematical programs with equilibrium constraints.* SIAM Journal Optimization, 17 (2006), pp. 259–286.

[15] 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.

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

[17] T. Hoheisel and C. Kanzow: *On the Abadie and Guignard constraint qualification for mathematical progams with vanishing constraints.* Preprint 272, Institute of Mathematics, University of Würzburg, Würzburg, September 2006.

[18] X.M. Hu and D. Ralph: *Convergence of a penalty method for mathematical programming with complementarity constraints.* Journal of Optimization Theory and Applications, 123 (2004), pp. 365–390.

[19] S. Leyffer, G. Lopez-Calva, and J. Nocedal: *Interior methods for mathematical programs with complementarity constraints.* SIAM Journal on Optimization, 17 (2006), pp. 52–77.

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

[21] O. L. Mangasarian: *Nonlinear Programming.* McGraw Hill, New York, NY, 1969 (reprinted by SIAM, Philadelphia, PA, 1994).

[22] B. S. Mordukhovich: *Variational Analysis and Generalized Differentiation I. Basic Theory.* A Series of Comprehensive Studies in Mathematics, Vol. 330, Springer, Berlin, Heidelberg, 2006.

[23] B. S. Mordukhovich: *Variational Analysis and Generalized Differentiation II. Applications.* A Series of Comprehensive Studies in Mathematics, Vol. 331, Springer, Berlin, Heidelberg, 2006.

[24] J. Nocedal and S. J. Wright: *Numerical Optimization.* Springer Series in Operations Research, Springer, New York, NY, 1999.

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

[26] J. V. Outrata: *A generalized mathematical program with equilibrium constraints.* SIAM Journal of Control and Optimization, 38 (2000), pp. 1623–1638.

[27] 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.

[28] 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.

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

[30] D. Ralph and S.J. Wright: *Some properties of regularization and penalization schemes for MPECs.* Optimization Methods and Software, 19 (2004), pp. 527–556.

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

[32] R. T. Rockafellar: *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

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

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

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

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

[37] J. J. Ye: *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints.* Journal of Mathematical Analysis and Applications, 307 (2005), pp. 350–369.

[38] J. J. Ye: *Constraint qualifications and KKT conditions for bilevel programming problems.* Mathematics of Operations Research, to appear.