MATHEMATICAL PROGRAMS WITH VANISHING CONSTRAINTS: A NEW REGULARIZATION APPROACH WITH STRONG CONVERGENCE PROPERTIES[1]

Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz

Preprint 301 Dezember 2010

University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany
e-mail: hoheisel@mathematik.uni-wuerzburg.de
kanzow@mathematik.uni-wuerzburg.de
schwartz@mathematik.uni-wuerzburg.de

December 3, 2010

[1]This research was partially supported by a grant from the international doctorate program "Identification, Optimization, and Control with Applications in Modern Technologies" within the Elite-Network of Bavaria.

Abstract. Motivated by a recent method introduced by two of the authors in [16] for mathematical programs with equilibrium constraints (MPECs), we present a related regularization scheme for the solution of mathematical programs with vanishing constraints (MPVCs). This new regularization method has stronger convergence properties than existing ones. In particular, it is shown that every limit point is at least M-stationary under a linear independence-type constraint qualification. If, in addition, an asymptotic weak nondegeneracy assumption holds, the limit point is shown to be S-stationary. Second-order conditions are not needed to obtain these results. Furthermore, some results are given which state that the regularized subproblems satisfy suitable standard constraint qualifications such that existing software can be applied to these regularized problems.

Key Words: Mathematical programs with vanishing constraints, Constraint qualification, Regularization method, Global convergence.

1 Introduction

We consider a constrained optimization problem of the form

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \quad \forall i = 1, \ldots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \ldots, p, \\ & H_i(x) \geq 0 \quad \forall i = 1, \ldots, l, \\ & G_i(x)H_i(x) \leq 0 \quad \forall i = 1, \ldots, l \end{array} \tag{1}$$

that we call a *mathematical program with vanishing constraints*, MPVC for short, where we assume throughout that the functions $f, g_i, h_i, H_i, G_i : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable. The MPVC was introduced to the mathematical community in the recent paper [1], where it was extracted as a mathematical model of various applications including optimal topology design problems in mechanical structures. Several theoretical properties and different numerical approaches for MPVCs can be found in [1, 2, 7, 8, 9, 10, 11, 12, 13, 14].

The MPVC is closely related to a class of problems that is called a mathematical program with equilibrium (or complementarity) constraints (MPECs). Such an MPEC is a constrained optimization problem of the form

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \quad \forall i = 1, \ldots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \ldots, p, \\ & G_i(x) \geq 0 \quad \forall i = 1, \ldots, l, \\ & H_i(x) \geq 0 \quad \forall i = 1, \ldots, l, \\ & G_i(x)H_i(x) = 0 \quad \forall i = 1, \ldots, l. \end{array}$$

We refer the interested reader to the monographs [18, 19, 5] for more details on MPECs. In principle, it is possible to formulate an MPVC as an MPEC (and vice versa), and, indeed, there exist different ways how this can be accomplished, cf. the corresponding discussion in [1, 11] for some suitable reformulations. So far, however, all these reformulations have certain disadvantages like the introduction of additional (nonunique) solutions or the violation of certain constraint qualifications. This observation motivates to take into account the special structure of the MPVC and to consider this as an independent class of interesting optimization problems.

Nevertheless, the similarities between MPVCs on the one hand and MPECs on the other hand are obvious, hence it is natural to use existing ideas for the solution of one class of problems also for the solution of the other class of problems. This is precisely what is done in this paper: Based on a recent regularization method for the solution of MPECs from [16], we adapt the main idea and modify the approach in such a way that it exploits the structure of an MPVC. The result is a regularization method for MPVCs which has stronger convergence properties than existing (regularization) methods for MPVCs, cf. [2, 12, 14].

The organization of this paper is as follows: Section 2 summarizes some background material regarding stationary points and constraint qualifications for MPVCs. Section 3 gives the details and some preliminary properties of the new relaxation scheme. The convergence properties of the new regularization method are investigated in Section 4. We conclude with some final remarks in Section 5.

Some words regarding our notation: Most of the notation used is standard. For a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x) \in \mathbb{R}^n$ denotes the gradient of f at x which is understood as a column vector. Moreover, for $\lambda \in \mathbb{R}^n$ we denote by

$$\operatorname{supp}(\lambda) := \big\{i \in \{1, \ldots, n\} \mid \lambda_i \neq 0\big\}$$

the support of this vector.

2 Preliminaries

2.1 Constraint Qualifications for Standard Nonlinear Programs

Since the idea behind a relaxation method is to replace the difficult MPVC by a sequence of (hopefully simpler) standard nonlinear programs, we begin our preliminary section by recalling some constraint qualifications for this problem class. Consider the following nonlinear program

$$\begin{aligned} \min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 \quad \forall i = 1, \ldots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \ldots, p \end{aligned} \tag{2}$$

and define the set of active inequalities as

$$I_g(x^*) := \{i \mid g_i(x^*) = 0\}$$

for any $x^* \in \mathbb{R}^n$ feasible for the nonlinear program (2). Let Z denote the set of feasible points of (2) and $x^* \in Z$ be arbitrarily given. The *(Bouligand) tangent cone of Z at x^** is then defined as

$$\mathcal{T}_Z(x^*) := \Big\{d \in \mathbb{R}^n \mid \exists\{x^k\} \subseteq Z, \exists\{\tau_k\} \downarrow 0 \text{ such that } x^k \to x^* \text{ and } \frac{x^k - x^*}{\tau_k} \to d\Big\},$$

and the *linearized cone of Z at x^** is given by

$$\mathcal{L}_Z(x^*) := \big\{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0 \ (i \in I_g(x^*)),\ \nabla h_i(x^*)^T d = 0 \ (i = 1, \ldots, p)\big\}.$$

Furthermore, the *polar cone* to an arbitrary cone $\mathcal{C} \subseteq \mathbb{R}^n$ is defined as

$$\mathcal{C}^\circ := \{s \in \mathbb{R}^n \mid \forall d \in \mathcal{C} : \ s^T d \leq 0\}.$$

One of the constraint qualifications we are going to state uses positive-linearly dependent vectors. We therefore first recall the definition of positive-linear dependence.

Definition 2.1 *A set of vectors*

$$\{a_i \mid i \in I_1\} \cup \{b_i \mid i \in I_2\}$$

is said to be positive-linearly dependent *if there exist scalars* α_i $(i \in I_1)$ *and* β_i $(i \in I_2)$, *not all of them being zero, with* $\alpha_i \geq 0$ *for all* $i \in I_1$ *and*

$$\sum_{i \in I_1} \alpha_i a_i + \sum_{i \in I_2} \beta_i b_i = 0.$$

Otherwise, we say that these vectors are positive-linearly independent.

With these definitions, we are now able to define some constraint qualifications for nonlinear programs.

Definition 2.2 *A feasible point* x^* *for* (2) *is said to satisfy the*

(a) linear independence constraint qualification (LICQ) *if the gradients*

$$\{\nabla g_i(x^*) \mid i \in I_g(x^*)\} \cup \{\nabla h_i(x^*) \mid i = 1, \ldots, p\}$$

are linearly independent;

(b) constant positive-linear dependence constraint qualification *(CPLD) if, for any subsets* $I_1 \subseteq I_g(x^*)$ *and* $I_2 \subseteq \{1, \ldots, p\}$ *such that the gradients*

$$\{\nabla g_i(x^*) \mid i \in I_1\} \cup \{\nabla h_i(x^*) \mid i \in I_2\}$$

are positive-linearly dependent, there exists a neighbourhood $N(x^*)$ *of* x^* *such that the gradients*

$$\{\nabla g_i(x) \mid i \in I_1\} \cup \{\nabla h_i(x) \mid i \in I_2\}$$

are linearly dependent for all $x \in N(x^*)$*;*

(c) Abadie constraint qualification (ACQ) *if* $\mathcal{T}_Z(x^*) = \mathcal{L}_Z(x^*)$*;*

(d) Guignard constraint qualification (GCQ) *if* $\mathcal{T}_Z(x^*)^\circ = \mathcal{L}_Z(x^*)^\circ$.

The following relations hold between these four constraint qualifications:

$$\text{LICQ} \Longrightarrow \text{CPLD} \Longrightarrow \text{ACQ} \Longrightarrow \text{GCQ}.$$

The second implication was proven in [3], whereas the first and the third implication follow directly from the definitions. It is well known that every local minimum x^* of (2), such that GCQ holds in x^*, admits multipliers λ_i $(i = 1, \ldots, m)$ and μ_i $(i = 1, \ldots, p)$ such that the triple (x^*, λ, μ) is a *KKT point*, i.e.,

$$0 = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*)$$

with $\text{supp}(\lambda) \subseteq I_g(x^*)$ and $\lambda_i \geq 0$ $(i = 1, \ldots, m)$.

2.2 Stationary Points for MPVCs

While the KKT conditions are, more or less, the only prominent stationarity concept for standard nonlinear programs, there are several stationarity concepts in use when it comes to MPVCs. In order to state these, we need the following index sets: Let x^* be an arbitrary feasible point of (1). Then let $I_g = \{i \mid g_i(x^*) = 0\}$ be defined as before, and consider the additional index sets

$$I_+ := \big\{i \,\big|\, H_i(x^*) > 0\big\}, \quad I_0 := \big\{i \,\big|\, H_i(x^*) = 0\big\}. \tag{3}$$

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

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

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

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

Note that the first subscript indicates the sign of $H_i(x^*)$, whereas the second subscript stands for the sign of $G_i(x^*)$. We would also like to point out that the above index sets substantially depend on the chosen point x^*. Throughout this section, it will always be clear from the context which point these index sets refer to.

Definition 2.3 *Let x^* be feasible for the MPVC* (1). *Then x^* is called*

(a) weakly stationary *if there exist multipliers $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p, \eta^H, \eta^G \in \mathbb{R}^l$ such that*

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

and

$$\begin{array}{l} \lambda_i \geq 0 \ (i \in I_g), \ \lambda_i = 0 \ (i \notin I_g), \\ \eta_i^H = 0 \, (i \in I_+), \ \eta_i^H \geq 0 \, (i \in I_{0-}), \ \eta_i^H \ \textit{free} \ (i \in I_{0+} \cup I_{00}), \\ \eta_i^G = 0 \, (i \in I_{+-} \cup I_{0-} \cup I_{0+}), \ \eta_i^G \geq 0 \, (i \in I_{+0} \cup I_{00}). \end{array}$$

(b) T-stationary *if x^* is weakly stationary and $\eta_i^G \eta_i^H \leq 0$ for all $i \in I_{00}$.*

(c) M-stationary *if x^* is weakly stationary and $\eta_i^G \eta_i^H = 0$ for all $i \in I_{00}$.*

(d) S-stationary *if x^* is weakly stationary and $\eta_i^H \geq 0, \eta_i^G = 0$ for all $i \in I_{00}$.*

Obviously, the following implications hold for these stationarity concepts:

S-stationarity $\Longrightarrow$ M-stationarity $\Longrightarrow$ T-stationarity $\Longrightarrow$ weak stationarity.

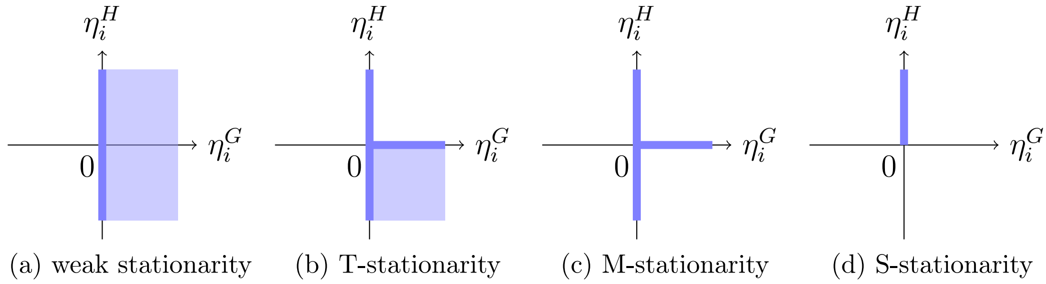

Figure 1: Geometric illustration of weak, T-, M-, and S-stationarity for an index $i \in I_{00}$

The only difference between these four stationarity concepts lies in the conditions on the multipliers corresponding to the bi-active set I_{00}. These conditions are illustrated in Figure 1. Hence, if the bi-active set is empty, all four stationary concepts coincide.

The notion of weak stationarity for MPVCs was introduced in [14], whereas M-stationarity for MPVCs is due to [9] and S-stationarity, which is in fact equivalent to the KKT conditions of (1), was first mentioned in [1]. T-stationarity was only recently brought on in [7]. In an MPEC setting, the counterpart of T-stationarity is usually called C-stationarity, cf. [20].

2.3 MPVC-tailored Constraint Qualifications

Since it is known that most standard constraint qualifications are violated by the vanishing constraints, special MPVC constraint qualifications have been developed to guarantee that a local solution of (1) is stationary in one of the above senses. Although there are much more MPVC constraint qualifications known by now, we will confine ourselves to only two of them. Whereas the first one is based on the famous linear independence constraint qualification for standard nonlinear programs and has been used for MPVCs quite some time now, the second one is an adaption of the less known constant positive-linear dependence constraint qualification which was introduced in [23] and further investigated in [3] and only recently applied to MPVCs in [12].

Definition 2.4 *A feasible point x^* of the MPVC* (1) *is said to satisfy the*

(a) MPVC-linear independence constraint qualification *(MPVC-LICQ) if the gradients*

$$\begin{aligned}&\{\nabla g_i(x^*) \mid i \in I_g\} \;\cup\; \{\nabla h_i(x^*) \mid i = 1, \dots, p\}\\ \cup\;\; &\{\nabla G_i(x^*) \mid i \in I_{00} \cup I_{+0}\} \;\cup\; \{\nabla H_i(x^*) \mid i \in I_0\}\end{aligned}$$

are linearly independent;

(b) MPVC-constant positive-linear dependence constraint qualification *(MPVC-CPLD) if, for all subsets $I_1 \subseteq I_g$, $I_2 \subseteq I_{0-}$, $I_3 \subseteq I_{+0} \cup I_{00}$, $I_4 \subseteq \{1, \dots, p\}$, $I_5 \subseteq I_{0+} \cup I_{00}$,*

the following implication holds true: If the gradients

$$\Big\{\{\nabla g_i(x) \mid i \in I_1\} \cup \{-\nabla H_i(x) \mid i \in I_2\} \cup \{\nabla G_i(x) \mid i \in I_3\}\Big\}$$
$$\cup \Big\{\{\nabla h_i(x) \mid i \in I_4\} \cup \{\nabla H_i(x) \mid i \in I_5\}\Big\}$$

are positive-linearly dependent in x^*, they remain linearly dependent in a whole neighbourhood of x^*.

In the definition of MPVC-CPLD, we use double face brackets to separate the gradients for which there are sign constraints in the definition of positive linear dependence from those without sign constraints.

Apart from those defined above, there exist a number of constraint qualifications tailored to MPVCs like MPVC-MFCQ, MPVC-CRCQ and MPVC-ACQ as variants of the standard MFCQ (Mangasarian-Fromovitz constraint qualification), standard CRCQ (constant rank constraint qualification), and standard ACQ. Some of the relations between these constraint qualifications are displayed in the diagram below, see [12] and the reference therein for more information about these constraint qualifications.

$$\begin{tikzcd}
 & \text{MPVC-MFCQ} \arrow[dr, Rightarrow] & & \\
\text{MPVC-LICQ} \arrow[ur, Rightarrow] \arrow[dr, Rightarrow] & & \text{MPVC-CPLD} \arrow[r, Rightarrow] & \text{MPVC-ACQ} \\
 & \text{MPVC-CRCQ} \arrow[ur, Rightarrow] & &
\end{tikzcd}$$

Analogous to the standard case, MPVC-LICQ is the strongest constraint qualification of the five mentioned here and MPVC-ACQ is the weakest. MPVC-CPLD relaxes both MPVC-MFCQ and MPVC-CRCQ, whereas it is known that neither MPVC-MFCQ implies MPVC-CRCQ nor vice versa.

In [1], it was proven that a local minimum of (1) satisfying MPVC-LICQ is S-stationary. However, the following example proves that we cannot expect local or even global minima to be S-stationary if MPVC-LICQ is violated.

Example 2.5 Consider the following 3-dimensional MPVC which is based on an example originally designed for MPECs from [20]:

$$\begin{array}{ll}
\min & f(x) = x_1 + x_2 - x_3 \\
\text{s.t.} & g_1(x) = -4x_1 + x_3 \leq 0, \\
 & g_2(x) = -4x_2 + x_3 \leq 0, \\
 & H(x) = x_2 \geq 0, \\
 & G(x)H(x) = x_1 x_2 \leq 0
\end{array}$$

One can easily verify that $x^* = (0,0,0)^T$ is the global minimum and that MPVC-LICQ is violated in x^* whereas MPVC-MFCQ and MPVC-CRCQ are satisfied. Although x^* is the global minimum, it is M-stationary but not S-stationary. In fact, there exist two multipliers corresponding to x^*, namely $(\lambda_1, \lambda_2, \eta^H, \eta^G) = (\frac{1}{4}, \frac{3}{4}, -2, 0)$ and $(\lambda_1, \lambda_2, \eta^H, \eta^G) = (\frac{3}{4}, \frac{1}{4}, 0, 2)$ which both satisfy the requirements for M-stationarity, whereas none of them gives S-stationarity. $\Diamond$

It is known that every local minimum of (1) satisfying MPVC-ACQ (or any stronger MPVC constraint qualification such as the ones introduced above) is M-stationary, see [9] for the proof where even less (MPVC-GCQ) is needed. The example above implies, however, that we cannot hope for more unless we assume significantly stronger constraint qualifications.

3 New Relaxation Approach

As their name indicates, the reason why MPVCs are considered complicated problems lies in the vanishing constraints. Illustrating both the theoretical and numerical nature of the difficulties, we consider the following most simple MPVC:

$$\begin{array}{ll} \min_{x_1,x_2} & f(x) \\ \text{s.t.} & H(x) = x_2 \geq 0, \\ & G(x)H(x) = x_1 x_2 \leq 0 \end{array} \tag{6}$$

The feasible set of (6) is depicted in Figure 2.

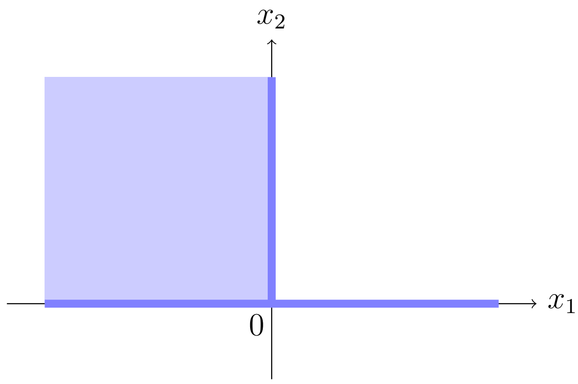

Figure 2: Feasible set for Example 6

Obviously, the feasible set consists of the union of the second quadrant with the half-axis $\mathbb{R}_+ \times \{0\}$. The latter part is lower-dimensional and violates both standard LICQ and MFCQ since here both constraints $H(x) \geq 0$ and $G(x)H(x) \leq 0$ are active. Thus, the perhaps most intuitive approach to overcome these deficiencies is to enlarge the feasible area around this half-axis such that at least one of these two constraints is non-active in every feasible point. This idea leads to a so-called relaxation or regularization method where a sequence of standard nonlinear programs is solved such that the feasible sets of the nonlinear programs converge to the feasible set of the original MPVC. These methods are very popular for the related problem class of MPECs, see for example [21, 6, 17, 22, 15]. So far, only two of these approaches were applied to MPVCs, cf. [2, 12]. In this paper, we want to adapt the idea from [16] which exhibits strong convergence properties when applied to MPECs.

To this end, we consider the function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\varphi(a,b) = \begin{cases} ab, & \text{if } a+b \geq 0, \\ -\frac{1}{2}(a^2+b^2), & \text{if } a+b < 0. \end{cases}$$

In [16] it was shown that this function has the following properties:

Lemma 3.1 *(a) φ is an NCP-function, i.e. $\varphi(a,b)=0$ if and only if $a \geq 0, b \geq 0, ab = 0$.*

(b) φ is continuously differentiable with gradient

$$\nabla\varphi(a,b) = \begin{cases} \begin{pmatrix} b \\ a \end{pmatrix}, & \text{if } a+b \geq 0, \\ \begin{pmatrix} -a \\ -b \end{pmatrix}, & \text{if } a+b < 0. \end{cases}$$

(c) φ has the property that

$$\varphi(a,b) \begin{cases} > 0, & \text{if } a > 0 \text{ and } b > 0, \\ < 0, & \text{if } a < 0 \text{ or } b < 0. \end{cases}$$

Based on this function, we define the function $\Phi(\cdot;t) : \mathbb{R}^2 \to \mathbb{R}$ given by

$$\Phi(a,b;t) := \begin{cases} a(b-t), & \text{if } \ a+b \geq t, \\ -\frac{1}{2}(a^2+(b-t)^2), & \text{if } \ a+b < t \end{cases} \tag{7}$$

for arbitrary values $t \geq 0$ and the corresponding relaxed problem NLP(t) for $t \geq 0$ as

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leq 0 & \forall i = 1,\dots,m, \\ & h_i(x) = 0 & \forall i = 1,\dots,p, \\ & H_i(x) \geq 0 & \forall i = 1,\dots,l, \\ & \Phi(G_i(x),H_i(x);t) \leq 0 & \forall i = 1,\dots,l. \end{array} \tag{8}$$

Hence, we replace the vanishing constraints

$$H_i(x) \geq 0, G_i(x)H_i(x) \leq 0$$

by

$$H_i(x) \geq 0, \Phi(G_i(x),H_i(x);t) \leq 0,$$

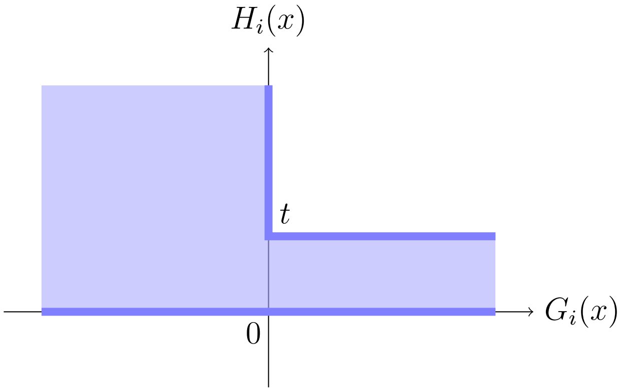

Figure 3: Geometric interpretation of the relaxation

see Figure 3 for a geometric illustration. Note that, unlike [16], we do not relax the constraint $G_i(x)H_i(x) \leq 0$ symmetrically in G_i and H_i since we consider MPVCs where there is no constraint $G_i(x) \geq 0$. This has the advantage that every local minimum x^* of (1), where $G_i(x^*) \leq 0$ for all i and $H_i(x^*) > 0$ for all i with $G_i(x^*) = 0$ also is a local minimum of NLP(t) if $t > 0$ is chosen sufficiently small. The following properties of NLP(t) can be proven completely analogous to [16]:

Lemma 3.2 *Let X be the feasible set of the MPVC* (1) *and $X(t)$ the feasible set of NLP(t) for $t \geq 0$. Then the following three statements hold:*

(a) $X(0) = X$.

(b) $X(t_1) \subseteq X(t_2)$ *for all* $0 \leq t_1 \leq t_2$.

(c) $\bigcap_{t \geq 0} X(t) = X$.

Thus, the relaxed problems NLP(t) have the requested properties that their feasible sets contain the one of the original MPVC and are converging to the original feasible set for $t \downarrow 0$.

To simplify the notation in the subsequent convergence analysis, we introduce the index sets

$$\begin{aligned} I_g(x) &:= \{i \mid g_i(x) = 0\}, \\ I_H(x) &:= \{i \mid H_i(x) = 0\}, \\ I_\Phi(x;t) &:= \{i \mid \Phi(G_i(x), H_i(x); t) = 0\} \end{aligned}$$

for $t \geq 0$ and x feasible for NLP(t). We also employ the following partition of the index set I_Φ:

$$I_\Phi^{00}(x;t) \;:=\; \{i \in I_\Phi \mid H_i(x) - t = 0, G_i(x) = 0\},$$

$$
\begin{array}{rcl}
I_\Phi^{+0}(x;t) & := & \{i \in I_\Phi \mid H_i(x) - t > 0, G_i(x) = 0\}, \\
I_\Phi^{0+}(x;t) & := & \{i \in I_\Phi \mid H_i(x) - t = 0, G_i(x) > 0\}.
\end{array}
$$

Note that these sets indeed form a partition of $I_\Phi(x;t)$ since the definition of Φ together with Lemma 3.1 (a) implies that

$$
\Phi(G_i(x), H_i(x); t) = 0 \Longleftrightarrow G_i(x) \geq 0,\ H_i(x) - t \geq 0,\ G_i(x)\big(H_i(x) - t\big) = 0.
$$

This relation will be used several times in our subsequent analysis without an explicit reference.

4 Convergence Results

Concentrating on the complicated constraints, we assume in this section that there are no standard inequality and equality constraints. However, extensions of the results and proofs to the case with standard constraints are straightforward.

The following theorem is our main convergence result. Basically, it states that a limit of KKT points of the regularized problems gives at least an M-stationary point of the original MPVC under fairly mild assumptions. Note that this result is stronger than the corresponding convergence results of existing relaxation methods where (under similar or even stronger assumptions) convergence can be shown only to weakly stationary points or to T-stationary points, cf. [12, 14]. In fact, so far none of the existing regularization methods for MPVCs can be shown to converge to M-stationary points without requiring additional assumptions to hold.

Theorem 4.1 *Let $\{t_k\} \downarrow 0$ such that $\{(x^k, \nu^k, \rho^k)\}$ is a sequence of KKT points of $NLP(t_k)$ with $x^k \to x^*$, where MPVC-CPLD is satisfied at x^*. Then x^* is an M-stationary point of (1).*

Proof. The feasibility of x^* for (1) is a direct consequence of Lemma 3.2. As (x^k, ν^k, ρ^k) is a KKT point of $NLP(t_k)$, we have

$$
0 = \nabla f(x^k) - \sum_{i=1}^{l} \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^{l} \rho_i^k \nabla[\Phi(G_i(x^k), H_i(x^k); t_k)],
$$

and

$$
\begin{array}{llll}
\nu_i^k \geq 0, & H_i(x^k) \geq 0, & \nu_i^k H_i(x^k) = 0 & (i = 1, \dots, l), \\
\rho_i^k \geq 0, & \Phi(G_i(x^k), H_i(x^k); t_k) \leq 0, & \rho_i^k \Phi(G_i(x^k), H_i(x^k); t_k) = 0 & (i = 1, \dots, l).
\end{array}
$$

Putting

$$
\delta_i^{G,k} := \begin{cases} \rho_i^k (H_i(x^k) - t_k), & i \in I_\Phi^{+0}(x^k; t_k), \\ 0, & \text{else}, \end{cases}
\qquad
\delta_i^{H,k} := \begin{cases} \rho_i^k G_i(x^k), & i \in I_\Phi^{0+}(x^k; t_k), \\ 0, & \text{else}, \end{cases}
$$

and noticing that

$$\nabla[\Phi(G_i(x^k), H_i(x^k); t_k)] = (H_i(x^k) - t_k)\nabla G_i(x^k) + G_i(x^k)\nabla H_i(x^k) = 0$$

for $i \in I_\Phi^{00}(x^k; t_k)$ we hence get

$$0 = \nabla f(x^k) - \sum_{i=1}^{l} \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^{l} \delta_i^{G,k} \nabla G_i(x^k) + \sum_{i=1}^{l} \delta_i^{H,k} \nabla H_i(x^k),$$

where $\delta^{G,k}, \delta^{H,k} \geq 0$ due to the definition of the index sets $I_\Phi^{0+}(x^k; t_k)$ and $I_\Phi^{+0}(x^k; t_k)$. In view of [22, Lem. A.1] we may assume that the gradients

$$\{\nabla H_i(x^k) \mid i \in \operatorname{supp}(\nu^k) \cup \operatorname{supp}(\delta^{H,k})\} \cup \{\nabla G_i(x^k) \mid i \in \operatorname{supp}(\delta^{G,k})\} \tag{9}$$

are linearly independent.

We claim that the sequence $\{(\nu^k, \delta^{G,k}, \delta^{H,k})\}$ is bounded. In order to prove this, suppose that

$$\|(\nu^k, \delta^{G,k}, \delta^{H,k})\| \to_{k\to\infty} \infty.$$

Then there exists $(\tilde{\nu}, \tilde{\delta}^G, \tilde{\delta}^H) \neq 0$ and an infinite subset $K \subseteq \mathbb{N}$ such that

$$\frac{(\nu^k, \delta^{G,k}, \delta^{H,k})}{\|(\nu^k, \delta^{G,k}, \delta^{H,k})\|} \to_{k\in K} (\tilde{\nu}, \tilde{\delta}^G, \tilde{\delta}^H).$$

By inserting this limiting process in the equation above, we get

$$\begin{aligned}
0 &= -\sum_{i\in\operatorname{supp}(\tilde{\nu})} \tilde{\nu}_i \nabla H_i(x^*) + \sum_{i\in\operatorname{supp}(\tilde{\delta}^G)} \tilde{\delta}_i^G \nabla G_i(x^*) + \sum_{i\in\operatorname{supp}(\tilde{\delta}^H)} \tilde{\delta}_i^H \nabla H_i(x^*) \\
&= \sum_{i\in\operatorname{supp}(\tilde{\nu})\cap I_{0-}} \nu_i(-\nabla H_i(x^*)) + \sum_{i\in\operatorname{supp}(\tilde{\delta}^G)} \tilde{\delta}_i^G \nabla G_i(x^*) \\
&\quad + \sum_{i\in\operatorname{supp}(\tilde{\delta}^H)} \tilde{\delta}_i^H \nabla H_i(x^*) + \sum_{i\in\operatorname{supp}(\tilde{\nu})\setminus I_{0-}} (-\nu_i)\nabla H_i(x^*).
\end{aligned}$$

Hence the vectors

$$\begin{aligned}
&\Big\{\{-\nabla H_i(x^*) \mid i \in \operatorname{supp}(\tilde{\nu}) \cap I_{0-}\} \cup \{\nabla G_i(x^*) \mid i \in \operatorname{supp}(\tilde{\delta}^G)\}\Big\} \\
\cup \;\; &\{\nabla H_i(x^*) \mid i \in \operatorname{supp}(\tilde{\delta}^H) \cup (\operatorname{supp}(\tilde{\nu}) \setminus I_{0-})\}
\end{aligned}$$

are positive-linearly dependent. In view of the inclusions

$$\begin{aligned}
\operatorname{supp}(\tilde{\nu}) \cap I_{0-} &\subseteq I_{0-}, \\
\operatorname{supp}(\tilde{\delta}^G) &\subseteq I_{00} \cup I_{+0}, \\
\operatorname{supp}(\tilde{\delta}^H) \cup (\operatorname{supp}(\tilde{\nu}) \setminus I_{0-}) &\subseteq I_{00} \cup I_{0+},
\end{aligned}$$

and MPVC-CPLD this contradicts the linear independence of the gradients in (9). Hence, the sequence $\{(\nu^k, \delta^{G,k}, \delta^{H,k})\}$ is bounded and therefore, at least on a subsequence, convergent to some limit $(\nu, \delta^G, \delta^H)$. By means of that, we define some new multipliers (η^G, η^H) by

$$\eta_i^G := \begin{cases} \delta_i^G, & i \in \operatorname{supp}(\delta^G), \\ 0, & \text{else}, \end{cases} \qquad \eta_i^H := \begin{cases} \nu_i, & i \in \operatorname{supp}(\nu), \\ -\delta_i^H, & i \in \operatorname{supp}(\delta^H), \\ 0, & \text{else}. \end{cases}$$

Note that η_i^H is well defined since we have $\operatorname{supp}(\nu^k) \cap \operatorname{supp}(\delta^{H,k}) = \emptyset$ for all k sufficiently large (otherwise, it is easy to see that there would exist an index i such that both $H_i(x^k) = 0$ and $H_i(x^k) = t_k > 0$, a contradiction). Then it follows that x^* with the multipliers (η^G, η^H) is at least weakly stationary. In order to verify that it is even M-stationary we assume the contrary. Then there exists an index $j \in I_{00}$ such that $\eta_j^G > 0$ and $\eta_j^H \neq 0$. Since $\eta_j^G > 0$, necessarily $j \in I_\Phi^{+0}(x^k; t_k)$ for k sufficiently large. Hence, $H_j(x^k) > t_k > 0$ and therefore, $\nu_j^k = 0$ and hence $\nu_j = 0$. Due to $\operatorname{supp}(\delta^H) \subseteq I_\Phi^{0+}(x^k; t_k)$, this yields $\eta_j^H = \nu_j = 0$, in contradiction to our assumption. This concludes the proof. $\square$

In order to obtain strong stationarity of the limit point, we need an additional assumption. To this end, we employ the notion of asymptotic weak nondegeneracy as defined below.

Definition 4.2 *Let $\{t_k\} \downarrow 0$ and $\{x^k\}$ be a sequence of feasible points of NLP(t_k) with $x^k \to x^*$. If for all k sufficiently large*

$$I_\Phi^{+0}(x^k; t_k) \cap I_{00} = \emptyset \quad \text{and} \quad I_\Phi^{0+}(x^k; t_k) \cap I_{00} = \emptyset$$

the sequence $\{x^k\}$ is called asymptotically weakly nondegenerate.

In [16], asymptotic weak nondegeneracy was defined differently (and for MPECs). However, is is easy to see that both definitions are equivalent. In combination with MPVC-CPLD, this condition is strong enough to guarantee S-stationarity of a limit point generated by the relaxation method.

Theorem 4.3 *Let $\{t_k\} \downarrow 0$ and $\{(x^k, \nu^k, \rho^k)\}$ be a sequence of KKT points of NLP(t_k) with $x^k \to x^*$. If MPVC-CPLD holds in x^* and the sequence $\{x^k\}$ is asymptotically weakly nondegenerate, then x^* is a strongly stationary point of* (1).

Proof. To prove this result, we only have to verify that the multipliers η^G, η^H we constructed in the proof of Theorem 4.1 satisfy $\eta_i^G = 0$ and $\eta_i^H \geq 0$ for all $i \in I_{00}$ under the additional assumption of asymptotic weak nondegeneracy. However, this assumption directly implies $\operatorname{supp}(\delta^G) \cap I_{00} = \emptyset$ and $\operatorname{supp}(\delta^H) \cap I_{00} = \emptyset$ and thus S-stationarity of the limit point. $\square$

Both convergence results rely on the assumption that the regularized problems admit a sequence of KKT points. The most intuitive approach to guarantee this would be to prove that the relaxed problems inherit a constraint qualification from the original problem (1). However, as the following example illustrates, this is not always the case.

Example 4.4 Consider again the example (6) from the beginning of Section 3:

$$\begin{array}{ll} \min_{x_1,x_2} & f(x) \\ \text{s.t.} & H(x) = x_2 \geq 0, \\ & G(x)H(x) = x_1x_2 \leq 0, \end{array} \tag{10}$$

and let $x^* = (0,0)^T$. Then it is easy to see that MPVC-LICQ is satisfied in x^*. Let us now consider the sequences $t_k = \frac{1}{k}$ and $x^k = (0, \frac{1}{k})^T$ for $k \in \mathbb{N}$. Then obviously $t_k \downarrow 0$ and $x^k \to x^*$, and the point x^k is feasible for NLP(t_k) for all $k \in \mathbb{N}$. The only active constraint is $\Phi(G(x^k), H(x^k); t_k) \leq 0$ and the corresponding gradient is

$$\nabla_x \Phi(G(x^k), H(x^k); t_k) = \begin{pmatrix} x_2^k - t_k \\ x_1^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

thus NLP(t_k) does not satisfy LICQ in x^k for all $k \in \mathbb{N}$. In fact, not even ACQ holds in x^k. Only GCQ, which is, in a sense, the weakest constraint qualification known for nonlinear programs, is valid at x^k. $\Diamond$

We are going to prove that this is not a mere coincidence but that MPVC-LICQ in a feasible point x^* of (1) always implies GCQ for all points x feasible for NLP(t), with $t > 0$ sufficiently small, that belong to a certain neighbourhood of x^*. To this end, we need two auxiliary results. The first one is stated in the following lemma and can be found in [4, Theorem 3.2.2].

Lemma 4.5 *Consider the cones*

$$\mathcal{C}_1 := \{d \in \mathbb{R}^n \mid a_i^T d \leq 0,\ \forall i = 1, \ldots, m,\quad b_i^T d = 0\ \forall i = 1, \ldots, p\}$$

and

$$\mathcal{C}_2 := \{s \in \mathbb{R}^n \mid s = \sum_{i=1}^m \alpha_i a_i + \sum_{i=1}^p \beta_i b_i,\quad \alpha_i \geq 0\ \forall i = 1, \ldots, m\}.$$

Then $\mathcal{C}_2 = \mathcal{C}_1^\circ$ *and* $\mathcal{C}_1 = \mathcal{C}_2^\circ$.

To simplify the notation in the proof of Theorem 4.7, we would like to introduce some nonlinear programs that are very similar to the relaxed problems NLP(t) but have better properties when it comes to constraint qualifications. Let $t > 0$ and $\hat{x}$ be feasible for NLP(t). Then choose an arbitrary subset $I \subseteq I_{00}^\Phi(\hat{x}; t)$ and denote its complement by $\bar{I} := I_{00}^\Phi \backslash I$. We then define the nonlinear program NLP(t, I) as

$$\begin{array}{lll} \min f(x) & \text{s.t.} & H_i(x) \geq 0, G_i(x) \leq 0 \quad \forall i \in I_\Phi^{+0}(\hat{x}; t) \cup I, \\ & & H_i(x) \geq 0, H_i(x) \leq t \quad \forall i \in I_\Phi^{0+}(\hat{x}; t) \cup \bar{I}, \\ & & H_i(x) \geq 0, \Phi(G_i(x), H_i(x); t) \leq 0 \quad \forall i \notin I_\Phi(\hat{x}; t) \end{array} \tag{11}$$

and denote its feasible set by $X(t, I)$. Then it is easy to see that

$$X(t, I) \subseteq X(t)$$

and that $\hat{x}$ is feasible for NLP(t, I), too. The following auxiliary lemma can be proven analogously to [16, Lemma 4.6]. Therefore, we do not state its proof here.

Lemma 4.6 *For all $t > 0$ and all $\hat{x}$ feasible for NLP(t),*

$$\mathcal{T}_{X(t)}(\hat{x}) = \bigcup_{I \subseteq I_\Phi^{00}(\hat{x};t)} \mathcal{T}_{X(t,I)}(\hat{x}).$$

These two auxiliary results enable us to prove the announced result about MPVC-LICQ for the original problem implying GCQ for NLP(t) in a neighbourhood.

Theorem 4.7 *Let x^* be feasible for the MPVC* (1) *such that MPVC-LICQ holds in x^*. Then there is a $T > 0$ and a neighbourhood $U(x^*)$ such that the following holds for all $t \in (0, T]$: If $x \in U(x^*)$ is feasible for NLP(t), then standard GCQ for NLP(t) holds in x.*

Proof. By MPVC-LICQ, the gradients

$$\{\nabla G_i(x^*) \mid i \in I_{00} \cup I_{+0}\} \cup \{\nabla H_i(x^*) \mid i \in I_0\} \tag{12}$$

are linearly independent and due to the continuity of the derivatives, they remain linearly independent in a neighbourhood. Hence, we can choose a $T > 0$ and a neighbourhood $U(x^*)$ such that for all $t \in (0, T]$ and all $\hat{x} \in U(x^*)$ feasible for NLP(t) the gradients in (12) are linearly independent and the following inclusions hold:

$$\begin{array}{rcl} I_H(\hat{x}) & \subseteq & I_0, \\ I_\Phi^{00}(\hat{x};t) \cup I_\Phi^{0+}(\hat{x};t) & \subseteq & I_{00} \cup I_{0+}, \\ I_\Phi^{00}(\hat{x};t) \cup I_\Phi^{+0}(\hat{x};t) & \subseteq & I_{00} \cup I_{+0}. \end{array} \tag{13}$$

Then for all $I \subseteq I_\Phi^{00}(\hat{x};t)$, the gradients corresponding to active constraints in NLP(t, I) are

$$\{\nabla G_i(\hat{x}) \mid i \in I_\Phi^{+0}(\hat{x};t) \cup I\} \cup \{\nabla H_i(\hat{x}) \mid i \in I_\Phi^{0+}(\hat{x};t) \cup \bar{I} \cup I_H(\hat{x})\}.$$

Thanks to the construction of T and $U(x^*)$, these gradients are linearly independent, i.e. LICQ is satisfied in $\hat{x}$ for NLP(t, I) (to this end, note that $i \in I_H(\hat{x})$ implies $i \notin I_\Phi(\hat{x};t)$). However, LICQ immediately implies ACQ, thus we know

$$\mathcal{T}_{X(t,I)}(\hat{x}) = \mathcal{L}_{X(t,I)}(\hat{x})$$

for all $I \subseteq I_\Phi^{00}(\hat{x};t)$. Applying Lemma 4.6 then yields

$$\mathcal{T}_{X(t)}(\hat{x}) = \bigcup_{I \subseteq I_\Phi^{00}(\hat{x};t)} \mathcal{L}_{X(t,I)}(\hat{x})$$

and passing over to the polar cone, we obtain

$$\mathcal{T}_{X(t)}(\hat{x})^\circ = \bigcap_{I \subseteq I_\Phi^{00}(\hat{x};t)} \mathcal{L}_{X(t,I)}(\hat{x})^\circ,$$

see [4, Theorem 3.1.9]. GCQ for NLP(t) is satisfied if $\mathcal{T}_{X(t)}(\hat{x})^\circ = \mathcal{L}_{X(t)}(\hat{x})^\circ$ holds. Since the inclusion $\mathcal{T}_{X(t)}(\hat{x})^\circ \supseteq \mathcal{L}_{X(t)}(\hat{x})^\circ$ is always true, it suffices to verify the opposite inclusion.

To do so, we will exploit the representation of $\mathcal{T}_{X(t)}(\hat{x})^\circ$ derived above. By definition, the linearized tangent cone to NLP(t, I) in $\hat{x}$ is given by

$$\begin{aligned}\mathcal{L}_{X(t,I)}(\hat{x}) = \{d \in \mathbb{R}^n \quad | \quad & \nabla G_i(\hat{x})^T d \leq 0 \ \forall i \in I_\Phi^{+0}(\hat{x};t) \cup I, \\ & \nabla H_i(\hat{x})^T d \leq 0 \ \forall i \in I_\Phi^{0+}(\hat{x};t) \cup \bar{I}, \\ & \nabla H_i(\hat{x})^T d \geq 0 \ \forall i \in I_H(\hat{x})\}.\end{aligned}$$

To calculate its polar cone, we can apply Lemma 4.5 and obtain

$$\begin{aligned}\mathcal{L}_{X(t,I)}(\hat{x})^\circ = \{s \in \mathbb{R}^n \quad | \quad & s = \sum_{i \in I_\Phi^{+0}(\hat{x};t) \cup I} \eta_i^G \nabla G_i(\hat{x}) + \sum_{i \in I_\Phi^{0+}(\hat{x};t) \cup \bar{I}} \eta_i^H \nabla H_i(\hat{x}) - \sum_{i \in I_H(\hat{x})} \nu_i \nabla H_i(\hat{x}), \\ & \eta^G, \eta^H, \nu \geq 0\}.\end{aligned}$$

Now consider an arbitrary element $s \in \mathcal{T}_{X(t)}(\hat{x})^\circ$. Due to our representation of this polar-cone as an intersection of the cones $\mathcal{L}_{X(t,I)}(\hat{x})^\circ$, we can choose an arbitrary $I \subseteq I_\Phi^{00}(\hat{x};t)$ and obtain

$$s = \sum_{i \in I_\Phi^{+0}(\hat{x};t) \cup I} \eta_i^G \nabla G_i(\hat{x}) + \sum_{i \in I_\Phi^{0+}(\hat{x};t) \cup \bar{I}} \eta_i^H \nabla H_i(\hat{x}) - \sum_{i \in I_H(\hat{x})} \nu_i \nabla H_i(\hat{x}),$$

where $\eta^G, \eta^H, \nu \geq 0$. On the other hand, we could also choose $\bar{I}$ instead of I, which would lead us to

$$s = \sum_{i \in I_\Phi^{+0}(\hat{x};t) \cup \bar{I}} \bar{\eta}_i^G \nabla G_i(\hat{x}) + \sum_{i \in I_\Phi^{0+}(\hat{x};t) \cup I} \bar{\eta}_i^H \nabla H_i(\hat{x}) - \sum_{i \in I_H(\hat{x})} \bar{\nu}_i \nabla H_i(\hat{x}),$$

where again $\bar{\eta}^G, \bar{\eta}^H, \bar{\nu} \geq 0$. The construction of $U(x^*)$ ensures that the gradients

$$\{\nabla G_i(\hat{x}) \mid i \in I_\Phi^{+0}(\hat{x};t) \cup I_\Phi^{00}(\hat{x};t)\} \cup \{\nabla H_i(\hat{x}) \mid i \in I_\Phi^{0+}(\hat{x};t) \cup I_\Phi^{00}(\hat{x};t) \cup I_H(\hat{x})\}$$

are linearly independent which implies that s has a unique representation based on these vectors. As a consequence, we now know $\eta_i^G = 0$ for all $i \in I$ and $\eta_i^H = 0$ for all $i \in \bar{I}$, i.e.

$$s = \sum_{i \in I_\Phi^{+0}(\hat{x};t)} \eta_i^G \nabla G_i(\hat{x}) + \sum_{i \in I_\Phi^{0+}(\hat{x};t)} \eta_i^H \nabla H_i(\hat{x}) - \sum_{i \in I_H(\hat{x})} \nu_i \nabla H_i(\hat{x}),$$

still with $\eta^G, \eta^H, \nu \geq 0$. However, an elementary calculation yields

$$\begin{aligned}\mathcal{L}_{X(t)}(\hat{x}) = \{d \in \mathbb{R}^n \quad | \quad & \nabla G_i(\hat{x})^T d \leq 0 \ \forall i \in I_\Phi^{+0}(\hat{x};t), \\ & \nabla H_i(\hat{x})^T d \leq 0 \ \forall i \in I_\Phi^{0+}(\hat{x};t), \\ & \nabla H_i(\hat{x})^T d \geq 0 \ \forall i \in I_H(\hat{x})\}\end{aligned}$$

and thus, by Lemma 4.5, we have proven $s \in \mathcal{L}_{X(t)}(\hat{x})^\circ$. Since $s \in \mathcal{T}_{X(t)}(\hat{x})^\circ$ was chosen arbitrarily, this verifies the inclusion $\mathcal{T}_{X(t)}(\hat{x})^\circ \subseteq \mathcal{L}_{X(t)}(\hat{x})^\circ$ and we have thus proven that GCQ for NLP(t) holds in $\hat{x}$. $\square$

The existence of Lagrange multipliers in local minima of NLP(t) is now a direct consequence of Theorem 4.7.

Theorem 4.8 *Let x^* be feasible for the MPVC (1) such that MPVC-LICQ holds in x^*. Then there is a $T > 0$ and a neighbourhood $U(x^*)$ such that the following holds for all $t \in (0, T]$: If $x \in U(x^*)$ is a local minimizer feasible for NLP(t), then there exist Lagrange multipliers such that x together with these multipliers is a KKT point of NLP(t).*

At first glance, Theorem 4.7 strikes the impression that the relaxed problems NLP(t) are not much better conditioned than the original MPVC. However, recall that LICQ is violated in every feasible point x^* of (1) where there is at least one index i such that $G_i(x^*) \geq 0$ and $H_i(x^*) = 0$, even if MPVC-LICQ holds in this point. In contrast to this, we are now going to prove that MPVC-LICQ in x^* implies LICQ for NLP(t) for most of the points out of a neighbourhood of x^*. A similar result also holds for the constraint qualification MPVC-CPLD, which was used in the previous convergence theorems.

Theorem 4.9 *Let x^* be feasible for the MPVC (1) such that MPVC-LICQ (MPVC-CPLD) holds in x^*. Then there is a $T > 0$ and a neighbourhood $U(x^*)$ such that the following holds for all $t \in (0, T]$: If $x \in U(x^*)$ is feasible for NLP(t) with $I_\Phi^{00}(x; t) = \emptyset$, then standard LICQ (CPLD) for NLP(t) holds in x.*

Proof. Let us first verify the assertion for MPVC-LICQ. To this end, choose $T > 0$ and $U(x^*)$ as in the poof of Theorem 4.7. Then for every $t \in (0, T]$ and every $\hat{x} \in U(x^*)$ feasible for NLP(t) with $I_\Phi^{00}(\hat{x}; t) = \emptyset$, the gradients corresponding to active constraints are

$$\{(H_i(\hat{x})-t)\nabla G_i(\hat{x}) \mid i \in I_\Phi^{+0}(\hat{x};t)\} \cup \{G_i(\hat{x})\nabla H_i(\hat{x}) \mid i \in I_\Phi^{0+}(\hat{x};t)\} \cup \{\nabla H_i(\hat{x}) \mid i \in I_H(\hat{x})\}.$$

By the definition of T and $U(x^*)$, these gradients are linearly independent, hence LICQ is satisfied in $\hat{x}$.

To prove the assertion under MPVC-CPLD, assume that there were sequences $t_k \downarrow 0$ and $x^k \to x^*$ with x^k feasible for NLP(t_k) and $I_\Phi^{00}(x^k; t_k) = \emptyset$ for all $k \in \mathbb{N}$ such that standard CPLD is not satisfied in x^k for all $k \in \mathbb{N}$. Since CPLD is violated, there have to be subsets $I_1^k \subseteq I_\Phi^{+0}(x^k; t_k)$, $I_2^k \subseteq I_\Phi^{0+}(x^k; t_k)$, and $I_3^k \subseteq I_H(x^k)$ such that the gradients

$$\Big\{\{(H_i(x^k) - t_k)\nabla G_i(x^k) \mid i \in I_1^k\} \cup \{G_i(x^k)\nabla H_i(x^k) \mid i \in I_2^k\} \cup \{-\nabla H_i(x^k) \mid i \in I_3^k\}\Big\} \cup \emptyset$$

are positive-linearly dependent in x^k, but linearly independent in points arbitrary close to x^k. Since there are only finitely many index sets, we can assume without loss of generality $I_i^k = I_i$ for all $i = 1, 2, 3$. For all k sufficiently large, we know $I_\Phi^{+0}(x^k; t_k) \subseteq I_{00} \cup I_{+0}$ and thus $I_1 \subseteq I_{00} \cup I_{+0}$. Analogously, one can verify $I_2 \subseteq I_{00} \cup I_{0+}$ and $I_3 \subseteq I_0$ for all $k \in \mathbb{N}$ sufficiently large. Positive-linear dependence in x^k as stated above also implies positive-linear dependence of the gradients

$$\{-\nabla H_i(x^k) \mid i \in I_3 \cap I_{0-}\} \cup \Big\{\{\nabla G_i(x^k) \mid i \in I_1\} \cup \{\nabla H_i(x^k) \mid i \in I_2 \cup (I_3 \backslash I_{0-})\}\Big\},$$

and because of the violation of CPLD, we can find a sequence $y^k \to x^*$ such that these gradients are linearly independent in y^k. If these gradients were positive-linearly independent in x^*, by continuity they would remain positive-linearly independent in a whole neighbourhood. This, however, contradicts the existence of the sequence $x^k \to x^*$. On the other hand, if they were positive-linearly dependent in x^*, MPVC-CPLD would imply that they remain linearly dependent in a neighbourhood, which contradicts the existence of $y^k \to x^*$. This concludes the proof. $\square$

5 Final Remarks

We introduced a new relaxation scheme for the solution of mathematical programs with vanishing constraints. This relaxation scheme has stronger convergence properties than existing ones, and the relaxed problems were shown to satisfy suitable standard constraint qualifications which allow the application of standard software for the solution of the relaxed (sub-) problems.

We believe that the theoretical advantages of the current method will also be very useful from a numerical point of view when applied to difficult instances of MPVCs which, in particular, have weak or T-stationary points that are attracted by other methods but not by our scheme. At the moment, however, the number of test examples for MPVCs is somewhat limited, therefore a numerical comparison of these different approaches does not really provide enough insight in order to figure out the pros and cons of these methods.

References

[1] W. Achtziger and C. Kanzow: *Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications.* Mathematical Programming 114, 2008, pp. 69–99.

[2] W. Achtziger, T. Hoheisel, and C. Kanzow: *A smoothing-regularization approach to mathematical programs with vanishing constraints.* Preprint 284, Institute of Mathematics, University of Würzburg, Würzburg, 2008.

[3] R. Andreani, J.M. Martínez, and M.L. Schuverdt: *The CPLD condition of Qi and Wei implies the quasinormality constraint qualification.* Journal of Optimization Theory and Applications 125, 2005 , pp. 473–485.

[4] M.S. Bazaraa and C.M. Shetty: *Foundations of Optimization.* Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin/Heidelberg, 1976.

[5] S. Dempe: *Foundations of Bilevel Programming.* Nonconvex Optimization and Its Applications, Vol. 61, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[6] A.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.

[7] D. Dorsch, V. Shikhman, and O. Stein: *MPVC: Critical point theory.* Preprint No. 138, Department of Mathematics - C, RWTH Aachen University, June 2010.

[8] T. Hoheisel and C. Kanzow: *On the Abadie and Guignard constraint qualification for mathematical progams with vanishing constraints.* Optimization 58, 2009, pp. 431–448.

[9] T. Hoheisel and C. Kanzow: *Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualifications.* Journal of Mathematical Analysis and Applications 337, 2008, pp. 292–310.

[10] T. Hoheisel and C. Kanzow: *First- and second-order optimality conditions for mathematical programs with vanishing constraints.* Applications of Mathematics 52, 2007, pp. 495–514 (special issue dedicated to J.V. Outrata's 60. birthday)

[11] T. Hoheisel, C. Kanzow, and J.V. Outrata: *Exact penalty results for mathematical programs with vanishing constraints.* Nonlinear Analysis: Theory, Methods, and Applications 72, 2010, pp. 2514–2526.

[12] T. Hoheisel, C. Kanzow, and A. Schwartz: *Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints.* Optimization Methods and Software, to appear.

[13] A.F. Izmailov and A.L. Pogosyan: *Optimality conditions and Newton-type methods for mathematical programs with vanishing constraints.* Computational Mathematics and Mathematical Physics 49, 2009, pp. 1128–1140.

[14] A.F. Izmailov and M.V. Solodov: *Mathematical programs with vanishing constraints: optimality conditions, sensitivity and a relaxation method.* Journal of Optimization Theory and Applications 142, 2009, pp. 501–532.

[15] A. Kadrani, J.-P. Dussault, and A. Benchakroun: *A new regularization scheme for mathematical programs with complementarity constraints.* SIAM Journal on Optimization 20, 2009, pp. 78–103.

[16] C. Kanzow and A. Schwartz: *A new regularization method for mathematical programs with complementarity constraints with strong convergence properties.* Preprint 296, Institute of Mathematics, University of Würzburg, Würzburg, 2010.

[17] G.H. Lin and M. Fukushima: *A modified relaxation scheme for mathematical programs with complementarity constraints.* Annals of Operations Research 133, 2005, pp. 63–84.

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

[19] 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.

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

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

[22] S. Steffensen and M. Ulbrich: *A new regularization scheme for mathematical programs with equilibrium constraints.* SIAM Journal on Optimization 20, 2010, pp. 2504–2539.

[23] L. Qi and Z. Wei: *On the constant positive linear dependence condition and its applications to SQP methods.* SIAM Journal on Optimization, 10, 2000, pp. 963–981.