CONVERGENCE OF A LOCAL REGULARIZATION APPROACH FOR MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY OR VANISHING CONSTRAINTS[1]

Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz

Preprint 293 February 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

February 10, 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. Mathematical programs with equilibrium or vanishing constraints (MPECs or MPVCs) are both known to be difficult optimization problems which typically violate all standard constraint qualifications. A number of methods try to exploit the particular structure of MPECs and MPVCs in order to overcome these difficulties. In a recent paper by Ulbrich and Veelken [37], this was done for MPECs by a local regularization idea that may be viewed as a modification of the popular global regularization technique by Scholtes [33]. The aim of this paper is twofold: First, we improve the convergence theory from [37] in the MPEC setting, and second we translate this local regularization idea to MPVCs and obtain a new solution method for this class of optimization problems for which several convergence results are given.

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

1 Introduction

We consider two classes of constrained optimization problems. The first one is of the form

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

and is called a *mathematical program with equilibrium (or complementarity) constraints*, MPEC for short, whereas the second one has a similar structure given by

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & H_i(x) \geq 0,\ G_i(x)H_i(x) \leq 0 \quad \forall i = 1, \dots, l \end{array} \tag{2}$$

and is called a *mathematical program with vanishing constraints*, abbreviavted by MPVC. Throughout this paper, we assume that all functions $f, g_i, h_i, H_i, G_i : \mathbb{R}^n \to \mathbb{R}$ are at least once continuously differentiable.

While the MPEC is well-known in the community, see, for example, the two monographs [25, 28], the MPVC has been formally introduced in [1] only recently. The main motivation for studying the theoretical and numerical properties of MPVCs comes from their background in topology optimization where important applications can be formulated as MPVCs.

Both problems, the MPEC and the MPVC, are viewed as difficult optimization problems due to the fact that they typically violate all standard constraint qualifications. Hence, standard KKT theory cannot be applied in general to get suitable optimality conditions. For this reason, there is a vast literature on new problem-dependent constraint qualifications, see, in particular, [10, 11, 12, 13, 22, 26, 27, 29, 34, 38, 39, 40] for the case of MPECs, and [15, 16, 17] for the case of MPVCs.

Moreover, these MPEC- and MPVC-tailored constraint qualifications are also the basis for the convergence of several specialized algorithms for the solution of these two difficult optimization problems. Here, we refer the reader to [5, 6, 8, 9, 14, 18, 20, 24, 31, 32, 33, 35, 37], where a large number of different algorithmic approaches for the solution of MPECs are studied, as well as to [2, 3, 19] for some related methods being applied to MPVCs.

One of the most popular methods for the solution of MPECs is the regularization approach by Scholtes [33] which replaces the original MPEC by a sequence of regularized optimization problems given by

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

for some $t \downarrow 0$. In order to distinguish this approach from the subsequent one, we call this the *global regularization method* for MPECs, since the kink in the complementarity condition $a \geq 0, b \geq 0, ab = 0$ is replaced by the set $a \geq 0, b \geq 0, ab \leq t$ which may be viewed as a relaxation that changes the feasible set not only around the kink at the origin, but globally in the whole space.

This observation has been exploited recently by Ulbrich and Veelken [37] who suggest a new *local regularization approach* for MPECs where the central idea is to relax the kink only locally, whereas further away from the origin, the feasible set is not changed. Hence their method changes the feasible set only in a small part, namely exactly in those points which typically cause trouble from a theoretical (violation of constraint qualifications) and numerical point of view.

Another difference between the global and local regularization approach is that Scholtes [33], among other things, proves convergence of his global method to so-called C-stationary points under the MPEC-LICQ assumption, whereas Ulbrich and Veelken [37] are able to show convergence of their local method under the weaker MPEC-CRCQ condition. For the precise definitions of MPEC-LICQ, MPEC-CRCQ, and C-stationarity, we refer the reader to the subsequent sections. Furthermore, Ulbrich and Veelken [37] present extensive numerical results for their local regularization approach being applied to the MacMPEC test problem library [23] which indicate that their method is quite competitive.

The aim of this paper is twofold: First, we will show that the local regularization method for MPECs converges to C-stationary points even under weaker conditions than those given in [37]. In particular, it follows that this method converges to C-stationary points also under the MPEC-MFCQ condition, a result that is not covered by [37]. Second, we will adapt the idea of the local regularization approach to MPVCs and state several convergence results for this class of optimization problems. Numerical results are not the scope of this paper, especially since [37] already shows that the method is quite successful at least in the MPEC context.

The organization of the paper is as follows: We begin with some preliminaries in Section 2 where we recall several (more or less) known constraint qualifications from standard optimization as well as the basic idea of the local regularization approach from [37]. In Section 3, we then consider MPECs and show that the local regularization method from [37] convergences to C-stationary points under a very weak assumption. Section 4 considers MPVCs, presents the adaptation of the local regularization method to this class of mathematical programs which results in a new solution scheme for MPVCs, and gives several convergence results. We close with some final remarks in Section 5.

Most of the notation used in this paper is standard. For a function $f : \mathbb{R}^n \to \mathbb{R}$, we write $\nabla f(x)$ for the gradient of f at $x \in \mathbb{R}^n$, where this gradient is interpreted as a column vector. For a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$, the partial derivatives are denoted by $D_1 f$ and $D_2 f$, respectively. Similarly, for f being twice continuously differentiable, the symbols $D_{11} f, D_{12} f, D_{22} f$ are used for the corresponding second-order derivatives. Given a vector $a \in \mathbb{R}^n$ and an index set $I \subseteq \{1, \dots, n\}$, we denote by a_I the vector given by

$a_I := (a_i)_{i\in I} \in \mathbb{R}^{|I|}$. The support of a vector $a \in \mathbb{R}^n$ is defined by

$$\operatorname{supp}(a) := \{i \in \{1, \dots, n\} \mid a_i \neq 0\}.$$

2 Preliminaries

2.1 Constraint Qualifications for Standard Problems

Let us consider the standard nonlinear program

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p \end{array} \tag{3}$$

with continuously differentiable functions $f, g_i, h_i : \mathbb{R}^n \to \mathbb{R}$. Furthermore, let X denote the feasible set of (3), and let $x^* \in X$ be an arbitrary feasible point. Recall that the (Bouligand) *tangent cone* of X at x^* is defined by

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

whereas the *linearized cone* of X at x^* is given by

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

where

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

denotes the set of active inequality constraints. Then x^* satisfies the

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

  $$\nabla g_i(x^*) \ (i \in I_g), \quad \nabla h_i(x^*) \ (i = 1, \dots, p)$$

  are linearly independent;

- *Mangasarian-Fromovitz constraint qualification* (MFCQ) if the gradients $\nabla h_i(x^*)$ $(i = 1, \dots, p)$ are linearly independent, and there exists a vector $d \in \mathbb{R}^n$ such that

  $$\nabla g_i(x^*)^T d < 0 \ (i \in I_g), \quad \nabla h_i(x^*)^T d = 0 \ (i = 1, \dots, p);$$

- *constant rank constraint qualification* (CRCQ) if there is a neighbourhood $N(x^*)$ of x^* such that for all subsets $I_1 \subseteq I_g$ and $I_2 \subseteq \{1, \dots, p\}$, the gradient vectors

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

  have the same rank (which depends on I_1, I_2) for all $x \in N(x^*)$;

- *Abadie constraint qualification* (ACQ) if $\mathcal{T}_X(x^*) = \mathcal{L}_X(x^*)$.

The reader might be less familiar with the CRCQ condition that was introduced in [21] and subsequently used successfully to weaken the assumptions (especially LICQ) in several theoretical results.

In order to formulate another, much more recent, constraint qualification that will play a fundamental role in our analysis, we have to introduce the notion of positive-linearly dependent vectors.

Definition 2.1 *Let x^* be feasible for* (3) *and $I_1 \subseteq I_g, I_2 \subseteq \{1, \dots, p\}$ be arbitrarily given. Then the set of gradients*

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

is called positive-linearly dependent *if there exist scalars $\{\alpha_i\}_{i \in I_1}$ and $\{\beta_i\}_{i \in I_2}$ with $\alpha_i \geq 0$ for all $i \in I_1$, not all of them being zero, such that*

$$\sum_{i \in I_1} \alpha_i \nabla g_i(x^*) + \sum_{i \in I_2} \beta_i \nabla h_i(x^*) = 0.$$

Otherwise, we say that this set of gradient vectors is positive-linearly independent.

Note that positive-linearly dependent vectors are, in particular, linearly dependent. This notion allows us to formulate the following condition from [30]: The feasible point x^* satisfies the

- *constant positive linear dependence condition* (CPLD) if, for any subsets $I_1 \subseteq I_g$ and $I_2 \subseteq \{1, \dots, p\}$ such that the gradients

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

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

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

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

Then the following implications hold between these different constraint qualifications:

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

The fact that LICQ implies both MFCQ and CRCQ is obvious. Furthermore, it is not difficult to see that CPLD is implied by MFCQ and CRCQ, whereas MFCQ and CRCQ are not related to each other and are therefore independent constraint qualifications. Finally, the fact that CPLD implies ACQ follows from results in [4, 7]. In particular, this means that CPLD is a constraint qualification for (3), which was not clear when this condition was introduced originally in [30].

Now, let

$$L(x, \lambda, \mu) := f(x) + \lambda^T g(x) + \mu^T h(x)$$

be the *Lagrangian* of the optimization probem (3). Then, given a local minimum x^* of (3) such that a suitable constraint qualification holds at x^*, there exist multipliers $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that (x^*, λ, μ) satisfies the corresponding *KKT conditions*

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= 0, \\ h(x) &= 0, \\ g(x) \leq 0,\ \lambda \geq 0,\ \lambda^T g(x) &= 0. \end{aligned}$$

Every triple (x^*, λ, μ) satisfying these KKT conditions is called a *KKT point*, whereas the x-part itself is called a *stationary point*. Furthermore, for twice continuously differentiable functions f, g, h, we say that a KKT point (x^*, λ, μ) satisfies the *strong second-order sufficiency condition* (SSOSC) for problem (3) if $d^T \nabla^2_{xx} L(x^*, \lambda, \mu) d > 0$ holds for all nonzero elements d from the set

$$\mathcal{C}(x^*) := \left\{ d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d = 0 \ (i : \lambda_i > 0),\ \nabla h_i(x^*)^T d = 0 \ (i = 1, \dots, p) \right\}$$

that we call the *critical cone*.

2.2 Regularization Functions

The central idea for both MPECs and MPVCs is to relax the feasible set in a suitable way. For both problems, it will be enough to relax the complementarity conditions $a \geq 0, b \geq 0, ab = 0$. The idea that is used here is due to [37] and will be the basis for the subsequent methods. Geometrically, the complementarity conditons are given by the two nonnegative half-axes in the two-dimensional space. If we rotate this set by 45 degrees to the left, we obtain the absolute value function which is nondifferentiable in the origin. The idea is to approximate this absolute value function locally (say, within the interval $[-1, 1]$ though this will later be scaled to smaller neighbourhoods) by a suitable smooth function in such a way that it coincides with the absolute value function outside this local neighbourhood. This is made more precise in the following definition going back to [37].

Definition 2.2 *$\theta : [-1, 1] \to \mathbb{R}$ is called a* regularization function *if it satisfies the following conditions:*

(a) θ is twice continuously differentiable on $[-1, 1]$;

(b) $\theta(-1) = \theta(1) = 1$*;*

(c) $\theta'(-1) = -1$ *and* $\theta'(1) = 1$*;*

(d) $\theta''(-1) = \theta''(1) = 0$*;*

(e) $\theta''(x) > 0$ *for all* $x \in (-1, 1)$*.*

Note that condition (e) implies that θ is strictly convex on $[-1, 1]$. The following result taken from [37, Lemma 3.1] reveals an immediate but crucial property of all regularization functions.

Lemma 2.3 *Let* $\theta : [-1, 1] \to \mathbb{R}$ *be a regularization function. Then it holds that* $\theta(x) > |x|$ *for all* $x \in (-1, 1)$*.*

Two simple examples of suitable regularization functions are

$$\theta(x) := \frac{2}{\pi} \sin\left(\frac{\pi}{2}x + \frac{3\pi}{2}\right) + 1 \quad \text{and} \quad \theta(x) := \frac{1}{8}\left(-x^4 + 6x^2 + 3\right),$$

cf. [37, 36]. The second function is the Hermite interpolation polynomial satisfying the requirements from Definition 2.2.

3 Convergence of Local Regularization for MPECs

3.1 Preliminaries on MPECs

Here we first recall some standard terminology for MPECs and then present the details of the local regularization method from [37].

Let us consider the MPEC from (1), and let x^* be a feasible point for (1). Recall that $I_g = \{i = 1, \ldots, m \mid g_i(x^*) = 0\}$, and let us define the following additional index sets:

$$\begin{aligned} I_{00} &:= \{i = 1, \ldots, l \mid H_i(x^*) = G_i(x^*) = 0\}, \\ I_{0+} &:= \{i = 1, \ldots, l \mid H_i(x^*) = 0, G_i(x^*) > 0\}, \\ I_{+0} &:= \{i = 1, \ldots, l \mid H_i(x^*) > 0, G_i(x^*) = 0\}. \end{aligned}$$

Note that the first (second) subscript indicates whether $H_i(x^*)$ ($G_i(x^*)$) is positive or zero (and not vice versa to keep the notation consistent with the corresponding standard notation for MPVCs, see Section 4).

Next, we recall some well-known stationarity concepts for MPECs.

Definition 3.1 *Let* x^* *be feasible for the MPEC* (1)*. Then* x^* *is said to be* *(a)* weakly stationary, *if there are multipliers* $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p, \gamma, \nu \in \mathbb{R}^l$ *such that*

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

and

$$\begin{aligned} &\lambda_i \geq 0,\ \lambda_i g_i(x^*) = 0\ (i = 1, \ldots, l) \\ &\gamma_i = 0\ (i \in I_{0+}),\ \nu_i = 0\ (i \in I_{+0}); \end{aligned}$$

(b) C-stationary, *if it is weakly stationary and* $\gamma_i \nu_i \geq 0$ *for all* $i \in I_{00}$*;*

(c) M-stationary, *if it is weakly stationary and* $\gamma_i, \nu_i > 0$ *or* $\gamma_i \nu_i = 0$ *for all* $i \in I_{00}$.

For these stationary concepts to hold, we need some constraint qualifications which are modifications of some standard constraint qualifications. More precisely, in order to define suitable MPEC-tailored constraint qualifications, one often uses the auxiliary problem

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leq 0 & \forall i = 1, \ldots, m, \\ & h_i(x) = 0 & \forall i = 1, \ldots, p, \\ & H_i(x) \geq 0, G_i(x) = 0 & \forall i \in I_{+0}, \\ & H_i(x) = 0, G_i(x) \geq 0 & \forall i \in I_{0+}, \\ & H_i(x) = 0, G_i(x) = 0 & \forall i \in I_{00}, \end{array}$$

which is called the *tightened nonlinear program* TNLP(x^*) for MPECs. Note that TNLP(x^*) substantially depends on the chosen point x^*. We now use this tightened nonlinear program in order to define suitable MPEC-tailored constraint qualifications.

Definition 3.2 *We say that MPEC-LICQ (MPEC-MFCQ, MPEC-CRCQ, MPEC-CPLD) is satisfied in a feasible point* x^* *of* (1) *if standard LICQ (MFCQ, CRCQ, CPLD) is satisfied for the corresponding tightened nonlinear program TNLP(*x^**).*

Let us write down at least the MPEC-CPLD explicitly: The feasible point x^* of (1) satisfies MPEC-CPLD if, for all subsets $I_1 \subseteq I_g$, $I_2 \subseteq \{1, \ldots, p\}$, $I_3 \subseteq I_{00} \cup I_{+0}$, and $I_4 \subseteq I_{00} \cup I_{0+}$, the following implication is true: If the gradients

$$\{\nabla g_i(x) \mid i \in I_1\} \cup \Big\{\{\nabla h_i(x) \mid i \in I_2\} \cup \{\nabla G_i(x) \mid i \in I_3\} \cup \{\nabla H_i(x) \mid i \in I_4\}\Big\} \qquad (4)$$

are positive-linearly dependent in x^*, they remain linearly dependent in a whole neighbourhood of x^*. Note that we use double curly brackets in (4) in order to group together those gradients for which there are no sign constraints in the definition of positive linear dependence.

We next show that MPEC-CPLD implies MPEC-ACQ and, therefore, is a constraint qualification for MPECs. Since MPEC-ACQ will not be used elsewhere in this paper, we do not give the precise definition here and refer the interested reader to [10].

Lemma 3.3 *MPEC-CPLD implies MPEC-ACQ, hence every local minimizer x^* of* (1) *satisfying MPEC-CPLD is M-stationary.*

Proof. In [11], it was shown that MPEC-ACQ is a sufficient condition for M-stationarity of local optima. Consequently, it suffices to prove that MPEC-CPLD implies MPEC-ACQ. To do so, let x^* be a feasible point for (1) satisfying MPEC-CPLD and consider the nonlinear programs $\mathrm{NLP}(J_1, J_2)$

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leq 0 & \forall i = 1, \dots, m, \\ & h_i(x) = 0 & \forall i = 1, \dots, p, \\ & H_i(x) \geq 0, G_i(x) = 0 & \forall i \in I_{+0} \cup J_1, \\ & H_i(x) = 0, G_i(x) \geq 0 & \forall i \in I_{0+} \cup J_2, \end{array}$$

where (J_1, J_2) is an arbitrary partition of I_{00}. It is easy to see that MPEC-CPLD, which was defined as standard CPLD for $\mathrm{TNLP}(x^*)$, implies standard CPLD for all $\mathrm{NLP}(J_1, J_2)$. However, according to our discussion in Section 2.1, CPLD implies ACQ. Hence, we know that standard ACQ holds for all $\mathrm{NLP}(J_1, J_2)$, where (J_1, J_2) is a partition of I_{00}. This is a sufficient condition for MPEC-ACQ, cf. [10] for a proof. $\square$

Lemma 3.3 together with the fact that the the MPEC variants of LICQ, MFCQ, CRCQ, and CPLD are defined via the corresponding standard constraint qualifications for the tightened nonlinear program $\mathrm{TNLP}(x^*)$, we obtain from Section 2.1 that the following implications hold:

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

3.2 Convergence Analysis of Local Regularization for MPECs

In order to solve the MPEC (1), Ulbrich and Veelken [37] propose to solve a series of relaxed problems $\mathrm{NLP}(t)$, $t > 0$, defined by

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

with

$$\Phi : \mathbb{R}^2 \to \mathbb{R}, \quad \Phi(x_1, x_2; t) := x_1 + x_2 - \varphi(x_1 - x_2; t),$$

and

$$\varphi(\cdot; t) : \mathbb{R} \to \mathbb{R}, \quad \varphi(a; t) := \begin{cases} |a|, & \text{if } |a| \geq t, \\ t\theta(\frac{a}{t}), & \text{if } |a| < t. \end{cases}$$

One of the main results in [37] states that, given a sequence $\{t_k\} \downarrow 0$ and a corresponding sequence $\{x^k\}$ of stationary points of NLP(t_k) with $x^k \to x^*$ such that MPEC-CRCQ holds in x^*, then this limit point x^* is a C-stationary point of (1). The following result shows that this statement actually holds under the weaker MPEC-CPLD assumption. In particular, this result therefore also holds under the MPEC-MFCQ condition in view of our previous discussion.

Theorem 3.4 *Let $\{t_k\}$ be a sequence with $t_k \downarrow 0$ and let $\{x^k\}$ be a sequence of stationary points of NLP(t_k) with $x^k \to x^*$ and such that MPEC-CPLD holds in x^*. Then x^* is a C-stationary point of* (1).

Here we skip the proof of this result since it can be obtained by a simple modification of the one from [37], taking into account that MPEC-CPLD is indeed sufficient to verify the statement. Alternatively, we refer the reader to the corresponding proof of Theorem 4.12 in the MPVC setting.

Recall that our result is slightly more general than the one given in [37] as both MPEC-CRCQ and MPEC-MFCQ imply MPEC-CPLD, whereas MPEC-MFCQ does not imply MPEC-CRCQ. However, in the absence of standard inequality constraints $g_i(x) \leq 0$, MPEC-LICQ and MPEC-MFCQ coincide and so do MPEC-CRCQ and MPEC-CPLD. Hence, the difference between MPEC-CRCQ and MPEC-CPLD is not too big in the MPEC setting. Interestingly enough, this is not the case for MPVCs as we will see in the following section.

We close this section with a simple example illustrating the previous theory.

Example 3.5 Consider the following two-dimensional MPEC:

$$\min f(x) = 2x_2 \quad \text{s.t.} \quad \begin{aligned} g_1(x) &= x_1^2 + x_2 \leq 0, \\ g_2(x) &= x_1 \leq 0, \\ G(x) &= x_2 \geq 0, \\ H(x) &= x_1 + x_2 \geq 0, \\ G(x)H(x) &= x_2(x_1 + x_2) = 0. \end{aligned} \tag{5}$$

It is easy to see that the feasible region is $X = \{x \in \mathbb{R}^2 \mid x_1 \in [-1, 0], x_2 = -x_1\}$ and the strict global minimum is $x^* = (0, 0)$. All constraints are active in x^* and the gradients are

$$\nabla g_1(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \nabla g_2(x^*) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \nabla G(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \nabla H(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For MPEC-MFCQ to hold, one would have to find a vector $d \in \mathbb{R}^2$ with $\nabla G(x^*)^T d = 0, \nabla H(x^*)^T d = 0$ and $\nabla g_i(x^*)^T d < 0$ for $i = 1, 2$. This is obviously impossible, hence MPEC-MFCQ does not hold in x^*. MPEC-CRCQ also is violated in x^* because the gradients $\nabla g_1(x), \nabla g_2(x)$ are linearly dependent in x^* but linearly independent everywhere else. On the other hand, the weaker constraint qualification MPEC-CPLD is satisfied in x^*. To see this, note first that every subset of gradients that does not include $\nabla g_1(x)$, is independent of x. Thus, we only have to consider those subsets of gradients that include $\nabla g_1(x)$. Now, it is easy to see that there are only two subsets such that the included gradients are positive-linearly dependent in x^*, namely $\{\nabla g_1(x)\} \cup \{\nabla G(x), \nabla H(x)\}$ and $\{\nabla g_1(x), \nabla g_2(x)\} \cup \{\nabla G(x), \nabla H(x)\}$, and that those remain linearly dependent in a whole neighbourhood.

Now consider a sequence of the corresponding relaxed problems NLP(t_k), $t_k \downarrow 0$, where the condition $G(x)H(x) = 0$ is replaced by $\Phi(G(x), H(x); t_k) \leq 0$. One can easily verify that x^* is also the global minimum of NLP(t_k) and that standard CPLD holds in x^* for all $k \in \mathbb{N}$. Thus, $\{x^k\}$ with $x^k := x^*$ for all $k \in \mathbb{N}$ is a sequence of stationary points of NLP(t_k) that trivially converges to x^*. Hence, (5) is an example for an MPEC where the relaxation method converges although only MPEC-CPLD is satisfied. $\Diamond$

4 Convergence of Local Regularization for MPVCs

4.1 MPVCs and a New Local Regularization Approach

Let us consider the MPVC from (2), and let x^* be an arbitrary feasible point of (2). Then let $I_g = \{i \mid g_i(x^*) = 0\}$ be defined as before, and consider the additional index sets

$$I_+ := \{i \,|\, H_i(x^*) > 0\}, \quad I_0 := \{i \,|\, H_i(x^*) = 0\}. \tag{6}$$

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

$$\begin{array}{rcl} I_{+0} & := & \{i \,|\, H_i(x^*) > 0, G_i(x^*) = 0\}, \\ I_{+-} & := & \{i \,|\, H_i(x^*) > 0, G_i(x^*) < 0\}. \end{array} \tag{7}$$

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

$$\begin{array}{rcl} I_{0+} & := & \{i \,|\, H_i(x^*) = 0, G_i(x^*) > 0\}, \\ I_{00} & := & \{i \,|\, H_i(x^*) = 0, G_i(x^*) = 0\}, \\ I_{0-} & := & \{i \,|\, H_i(x^*) = 0, G_i(x^*) < 0\}. \end{array} \tag{8}$$

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 4.1 *Let x^* be feasible for the MPVC (2). 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

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

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

The notion of weak stationarity for MPVCs was introduced in [19], whereas M-stationarity for MPVCs is due to [16]. We note that, in contrast to MPECs, there is currently no concept like C-stationarity available for MPVCs. We stress, however, that the terminology is somewhat misleading: While there are no sign constraints for the multipliers corresponding to the biactive constraints in a weakly stationary point of an MPEC, we have such constraints in a weakly stationary point of an MPVC. Hence weak stationarity for MPVCs is, in comparison, a much stronger concept than weak stationarity for MPECs, i.e., weak stationarity for MPVCs corresponds to something like C-stationarity in the MPEC setting.

Similar to the exposition of the previous section, we now introduce a number of MPVC-tailored constraint qualifications via a tightened nonlinear program. To this end, let x^* be any feasible point of the MPVC. Then the auxiliary problem

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

is called the *tightened nonlinear program* of the MPVC and denoted by TNLP(x^*) (note that we use the same symbol here as for the tightened nonlinear program for MPECs, but this should not cause any troubles since, throughout this section, TNLP(x^*) always denotes the tightened nonlinear program of the MPVC from (2)).

Definition 4.2 *We say that MPVC-LICQ (MPVC-MFCQ, MPVC-CRCQ, MPVC-CPLD) is satisfied in a feasible point* x^* *of* (1) *if standard LICQ (MFCQ, CRCQ, CPLD) is satisfied for the corresponding tightened nonlinear program TNLP(x^*).*

Let us write down at least MPVC-LICQ and MPVC-CPLD explicitly: The MPVC-LICQ holds at x^* if the gradients

$$\nabla g_i(x^*)\ (i \in I_g), \quad \nabla h_i(x^*)\ (i = 1, \ldots, p), \quad \nabla H_i(x^*)\ (i \in I_0), \quad \nabla G_i(x^*)\ (i \in I_{00} \cup I_{+0})$$

are linearly independent, whereas the MPVC-CPLD holds at x^* if for all subsets $I_1 \subseteq I_g$, $I_2 \subseteq I_{0-}$, $I_3 \subseteq I_{+0} \cup I_{00}$, $I_4 \subseteq \{1, \ldots, p\}$, $I_5 \subseteq I_{0+} \cup I_{00}$, 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^*. Here, again, we use double face brackets to separate the gradients from the inequalities from those of the equality constraints.

Note that, in contrast to MPECs, there is also a difference between MPVC-CRCQ and MPVC-CPLD in the nonstandard constraints defined by G_i and H_i.

Similar to the proof of Lemma 3.3, we can also verify the following result by exploiting some techniques from [16], where the reader can also find the precise definition of MPVC-ACQ.

Lemma 4.3 *MPVC-CPLD implies MPVC-ACQ, hence every local minimizer x^* of* (2) *satisfying MPVC-CPLD is M-stationary.*

Similar to Section 3.1, the previous definitions and Lemma 4.3 together show that the following implications hold:

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

A keystone for our approach is the fact that the constraints of the MPVC (2) can be equivalently reformulated using the absolute-value function. In fact, it is easy to see that, for all $x_1, x_2 \in \mathbb{R}$, we have

$$x_1 x_2 \leq 0,\ x_2 \geq 0 \quad \Longleftrightarrow \quad x_1 + x_2 \leq |x_1 - x_2|,\ x_2 \geq 0\,. \tag{9}$$

This observation implies that the MPVC (2) is equivalent to the following program:

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

Note, however, that this is a *nonsmooth* reformulation of the problem, due to the absolute-value function.

Now, the idea of the new local regularization approach for MPVCs is to approximate the abolute-value function by a regularization function, cf. Definition 2.2. In the sequel, let θ always denote such a regularization function. We then define the parameter-dependent function $\varphi(\cdot;t) : \mathbb{R} \to \mathbb{R}$ given by

$$\varphi(a;t) := \begin{cases} |a|, & \text{if } |a| \geq t, \\ t\theta(\frac{a}{t}), & \text{if } |a| < t, \end{cases} \tag{10}$$

where $t > 0$. Some properties of φ are given below.

Lemma 4.4 *Let φ be defined as in* (10). *Then we have*

(a) $\varphi(a;t) > |a|$ for all $a \in (-t,t)$ and for all $t > 0$;

(b) $\varphi(a;t) = |a|$ for $|a| \geq t$ and for all $t > 0$;

(c) $\lim_{t\to 0} \varphi(a;t) = |a|$ for all $a \in \mathbb{R}$;

(d) $\varphi(\cdot;t)$ is twice continuously differentiable for all $t > 0$.

Proof. (a) Let $a \in (-t,t)$ for some $t > 0$. Then $\frac{|a|}{t} < 1$, and the definition of φ therefore implies

$$\varphi(a;t) = t\theta\left(\frac{a}{t}\right) > t\frac{|a|}{t} = |a|,$$

where the (strict) inequality comes from Lemma 2.3.

(b), (d) These statements follow directly from the definition of θ.

(c) Let $a \in \mathbb{R}$. If $a = 0$, then the boundedness of θ immediately gives

$$\varphi(0;t) = t\theta\left(\frac{a}{t}\right) \to 0 = a \quad \text{for } t \to 0.$$

On the other hand, if $a \neq 0$, we have $|a| \geq t$ for all $t > 0$ sufficiently small. Hence we obtain $\varphi(a;t) = |a|$ for all $t > 0$ sufficiently small, and this gives our assertion also for $a \neq 0$. $\square$

With the aid of $\varphi(\cdot;t)$, we are in a position to define $\Phi(\cdot;t) : \mathbb{R}^2 \to \mathbb{R}$ by

$$\Phi(x_1, x_2;t) := x_1 + x_2 - \varphi(x_1 - x_2;t)\,. \tag{11}$$

Some useful properties of the function $\Phi(\cdot;t)$ are subsumed in the following results.

Lemma 4.5 *For $t > 0$ let $\Phi(\cdot;t)$ be the function given in* (11). *Then $\Phi(\cdot;t)$ is twice continuously differentiable with gradient*

$$\nabla\Phi(x_1,x_2;t) = \begin{cases} \binom{2}{0}, & \text{if } x_1 - x_2 \le -t, \\ \binom{0}{2}, & \text{if } x_1 - x_2 \ge t, \\ \binom{1-\theta'(\frac{x_1-x_2}{t})}{1+\theta'(\frac{x_1-x_2}{t})}, & \text{if } |x_1 - x_2| < t, \end{cases}$$

and Hessian

$$\nabla^2\Phi(x_1,x_2;t) = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, & \text{if } |x_1 - x_2| \ge t, \\ \frac{1}{t}\begin{pmatrix} -\theta''(\frac{x_1-x_2}{t}) & \theta''(\frac{x_1-x_2}{t}) \\ \theta''(\frac{x_1-x_2}{t}) & -\theta''(\frac{x_1-x_2}{t}) \end{pmatrix}, & \text{if } |x_1 - x_2| < t. \end{cases}$$

Proof. The proof follows immediately from the definition of Φ; cf. also [37]. □

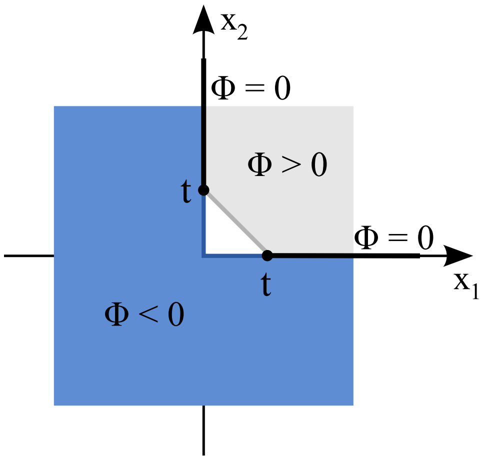


Figure 1: Illustration of Lemma 4.6

Lemma 4.6 *Let $\Phi(\cdot;t)$ be given by* (11). *Then the following holds true:*

$$\Phi(x_1,x_2;t) \begin{cases} < 0, & \text{if} \quad x_1 < 0 \quad \text{or} \quad x_2 < 0, \\ < 0, & \text{if} \quad x_1, x_2 \ge 0 \quad \text{and} \quad x_1 \cdot x_2 = 0 \quad \text{and} \quad |x_1 - x_2| < t, \\ = 0, & \text{if} \quad x_1, x_2 \ge 0 \quad \text{and} \quad x_1 \cdot x_2 = 0 \quad \text{and} \quad |x_1 - x_2| \ge t, \\ > 0, & \text{if} \quad x_1, x_2 > 0 \quad \text{and} \quad |x_1 - x_2| \ge t, \\ \text{free}, & \text{if} \quad x_1, x_2 > 0 \quad \text{and} \quad |x_1 - x_2| < t. \end{cases}$$

Proof. The proof follows immediately from considering the following cases:

i) For $x_1, x_2 \leq 0$ we have

$$\Phi(x_1, x_2; t) = \underbrace{x_1 + x_2}_{\leq 0} - \underbrace{\varphi(x_1 - x_2; t)}_{>0 \text{ by Lemma 4.4}} < 0.$$

ii) For $x_1 > 0, x_2 < 0$ we obtain from Lemma 4.4

$$\Phi(x_1, x_2; t) = x_1 + x_2 - \varphi(x_1 - x_2; t) \leq x_1 + x_2 - |x_1 - x_2| = 2x_2 < 0\,.$$

iii) For $x_1 < 0, x_2 > 0$ it follows again from Lemma 4.4 that

$$\Phi(x_1, x_2; t) = x_1 + x_2 - \varphi(x_1 - x_2; t) \leq x_1 + x_2 - |x_1 - x_2| = 2x_1 < 0\,.$$

iv) For $x_1 > 0, x_2 = 0$ we obtain from Lemma 4.4

$$\Phi(x_1, x_2; t) = x_1 - \varphi(x_1; t) \begin{cases} = 0, & \text{if} \quad x_1 \geq t, \\ < 0, & \text{if} \quad x_1 < t. \end{cases}$$

v) For $x_1 = 0, x_2 > 0$ it holds that

$$\Phi(x_1, x_2; t) = x_2 - \varphi(-x_2; t) \begin{cases} = 0, & \text{if} \quad x_2 \geq t, \\ < 0, & \text{if} \quad x_2 < t. \end{cases}$$

vi) For $x_1, x_2 > 0$ with $|x_1 - x_2| \geq t$ the definition of φ implies

$$\Phi(x_1, x_2; t) = x_1 + x_2 - \varphi(x_1 - x_2; t) = x_1 + x_2 - |x_1 - x_2| > 0.$$

The previous cases also show that, by continuity, the sign of Φ can be both positive and negative in the remaining case where $x_1, x_2 > 0$ and $|x_1 - x_2| < t$. $\square$

For a regularization parameter $t > 0$, the regularized problem $P(t)$ labels the following program:

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

Let the feasible set of $P(t)$ be denoted by $X(t)$.

Proposition 4.7 *For the feasible sets X of (2) and $X(t)$ of $P(t)$, we have the following relations:*

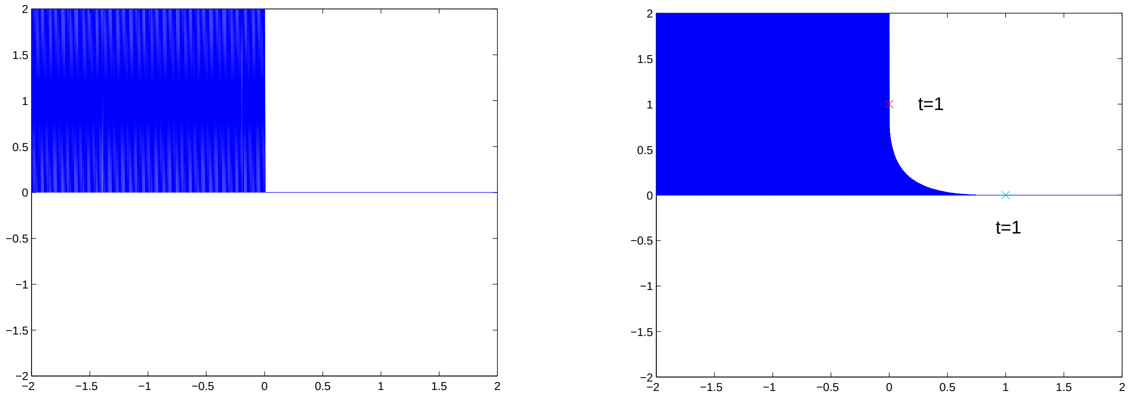
(a) For $t > 0$, we have $X \subset X(t)$;

(b) $X(0) = X$*;*

(c) For $t_1 < t_2$ *we have* $X(t_1) \subset X(t_2)$*.*

Proof. The proof of (a) follows immediately from Lemma 4.6. Statement (b) is due to (9) and the fact that $\Phi(x_1, x_2; 0) = x_1 + x_2 - |x_1 - x_2|$ for all $x_1, x_2 \in \mathbb{R}$, whereas statement (c) can be verified similarly to [37, Lemma 3.2]. $\square$

Item (a) from above is illustrated in the below example.



(a) X for Example 4.8 (b) $X(t)$ for Example 4.8

Figure 2: Feasible sets for Example 4.8

Example 4.8 Consider the MPVC

$$\min \; x_1^2 + x_2^2 \quad \text{s.t.} \quad x_2 \geq 0, \quad x_1 x_2 \leq 0.$$

Then its feasible set X and the feasible set $X(t)$ of the regularized problem for $t = 1$ are depicted in Figure 4.1. $\Diamond$

4.2 Convergence Analysis

Before we investigate some convergence properties of our regularization scheme, it is necessary to study some relationships between certain active sets, which arise naturally in this context. To this end, let x be an arbitrary feasible point of $P(t)$, and let us introduce the following index sets similar to those previously defined for a feasible point x^* of the MPVC itself in (6)–(8):

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

The set $I_+(x)$ is then partitioned, again, by

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

Analogously, $I_0(x)$ is partitioned into three subsets $I_{0+}(x), I_{00}(x), I_{0-}(x)$. Furthermore, put

$$M(x,t) := \{i \mid \Phi(G_i(x), H_i(x); t) = 0\}.$$

There are some elementary inclusions to hold for the above index sets and those from (6)–(8).

Lemma 4.9 *Let x^* be feasible for (2) and let $t > 0$. Then for all points x which are feasible for $P(t)$ and that are close enough to x^*, we have*

(a) $I_0(x) \subseteq I_0$;

(b) $M(x,t) \subseteq I_{00} \cup I_{0+} \cup I_{+0}$;

(c) $I_0(x) \cap M(x,t) \subseteq I_{0+} \cup I_{00}$.

Proof. (a) Let $i \notin I_0$. Thus, for x sufficiently close to x^* we have $H_i(x) > 0$ and hence, $i \notin I_0(x)$.

(b) Let $i \notin I_{00} \cup I_{0+} \cup I_{+0}$. Then we have $i \in I_{+-} \cup I_{0-}$ and thus, $G_i(x) < 0$ for x sufficiently close to x^*. Thus, in view of Lemma 4.6 we have $i \notin M(x,t)$.

(c) Follows immediately from (a) and (b). □

We are now in a position to state our three main convergence results. We begin with a result which may be viewed as the counterpart of [37, Theorem 5.1].

Theorem 4.10 *Let $\{(x^k, \lambda^k, \mu^k, \rho^k, \nu^k)\}$ be a sequence of KKT points of $P(t_k)$ for $t_k \downarrow 0$. Furthermore, let x^k converge to the point x^* satisfying MPVC-LICQ. Then the sequence $\{(\lambda^k, \mu^k, \eta^{G,k}, \eta^{H,k})\}$ with $\eta^{G,k}$ and $\eta^{H,k}$ given by*

$$\begin{array}{rcll} \eta_i^{G,k} & := & \nu_i^k D_1\Phi(G_i(x^k), H_i(x^k); t_k) & \forall i = 1, \dots, l, \\ \eta_i^{H,k} & := & \rho_i^k - \nu_i^k D_2\Phi(G_i(x^k), H_i(x^k); t_k) & \forall i = 1, \dots, l, \end{array} \tag{12}$$

converges to a limit $(\lambda, \mu, \eta^G, \eta^H)$ such that $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is a weakly stationary point of (2).

Proof. To simplify the notation within this proof, we assume that the MPVC does not contain any standard constraints given by g and h. The proof can easily be generalized to this case. Hence we concentrate our analysis on the difficult (vanishing) constraints.

The KKT conditions of $P(t_k)$ yield multipliers (ρ^k, ν^k) such that

$$0 \;=\; \nabla f(x^k) - \sum_{i=1}^{l} \rho_i^k \nabla H_i(x^k) + \sum_{i=1}^{l} \nu_i^k \nabla\big(\Phi(G_i(x^k), H_i(x^k); t_k)\big) \tag{13}$$

and

$$\begin{array}{llll} \rho_i^k \geq 0, & H_i(x^k) \geq 0, & \rho_i^k H_i(x^k) = 0 & (i = 1, \dots, l), \\ \nu_i^k \geq 0, & \Phi(G_i(x^k), H_i(x^k); t_k) \leq 0, & \nu_i^k \Phi(G_i(x^k), H_i(x^k); t_k) = 0 & (i = 1, \dots, l). \end{array} \tag{14}$$

Using the definitions of the multipliers $\eta^{G,k}$ and $\eta^{H,k}$ from (12), we may rewrite (13) as

$$0 = \nabla f(x^k) - \sum_{i=1}^{l} \eta_i^{H,k} \nabla H_i(x^k) + \sum_{i=1}^{l} \eta_i^{G,k} \nabla G_i(x^k)\,. \tag{15}$$

It is now our goal to investigate the signs of $\eta_i^{G,k}$ and $\eta_i^{H,k}$ with respect to the different index sets.

To this end, we begin with $\eta_i^{G,k}$. In view of Lemma 4.9 (b) and (14), for $i \in I_{+-} \cup I_{0-}$, we have $\nu_i^k = 0$ and thus, $\eta_i^{G,k} = 0$, for all k sufficiently large.

We now show that we have $\eta_i^{G,k} = 0$ for k sufficiently large for $i \in I_{0+}$, too. To this end, it obviously suffices to see that this holds for all $i \in I_{0+} \cap M(x^k, t_k)$, as for $i \in I_{0+} \setminus M(x^k, t_k)$ we have $\eta_i^{G,k} = 0$ anyway. Now, for $i \in I_{0+} \cap M(x^k, t_k)$ it clearly holds that $i \in I_{0+}(x^k) \cup I_{++}(x^k)$. More precisely, we have $i \in I_{0+}(x^k)$, because $i \in I_{++}(x^k) \cap M(x^k, t_k)$ would imply that $H_i(x^k), G_i(x^k) > 0$ and $G_i(x^k) + H_i(x^k) = \varphi(G_i(x^k) - H_i(x^k); t_k)$, and hence, by the definition of φ, we necessarily have $|G_i(x^k) - H_i(x^k)| \leq t_k \to 0$, which is a contradiction to $|G_i(x^k) - H_i(x^k)| \to G_i(x^*) > 0$, and so this case cannot occur. Consequently we have $i \in I_{0+}(x^k) \cap M(x^k, t_k)$ which means that $H_i(x^k) = 0, G_i(x^k) > 0$, and $\Phi(G_i(x^k), H_i(x^k); t_k) = 0$. This implies $D_1\Phi(G_i(x^k), H_i(x^k); t_k) = 0$ and thus $\eta_i^{G,k} = 0$ for all k sufficiently large.

Moreover, for $i \in I_{+-}$ we see from Lemma 4.6 that, for k sufficiently large, we have $\Phi(G_i(x^k), H_i(x^k); t_k) < 0$ and hence, $\nu_i^k = 0$. Furthermore, by analogous reasoning we also have $\rho_i^k = 0$. This yields $\eta_i^{H,k} = 0$ for k sufficiently large.

Eventually, it can also be argued that for $i \in I_{+0}$ we also have $\eta_i^{H,k} = 0$ for k sufficiently large. For these purposes, it obviously suffices to see that this holds true for $i \in I_{+0} \cap M(x^k, t_k)$, since otherwise, we have $\eta_i^{H,k} = 0$ anyway. So, for such an index i we have $\rho_i^k = 0$ for k sufficiently large. Moreover, it can be shown that $i \in I_{+0}(x^k)$ for k sufficiently large as follows: From Lemma 4.6 it is clear that $i \in I_{+0}(x^k) \cup I_{++}(x^k)$, but actually the case $i \in I_{++}(x^k)$ cannot occur, which can be verified just as above. Thus, we have $H_i(x^k) > 0$ and $G_i(x^k) = 0$. But as we also have $0 = \Phi(G_i(x^k), H_i(x^k); t_k) = H_i(x^k) - \varphi(-H_i(x^k); t_k)$, Lemma 4.4 yields $-H_i(x^k) \leq -t_k$ and thus, Lemma 4.5 gives $D_2\Phi(G_i(x^k), H_i(x^k); t_k) = 0$.

Together, thus far, we have shown that for k sufficiently large we have

$$\eta_i^{G,k} = 0 \; (i \in I_{0-} \cup I_{+-} \cup I_{0+}) \quad \text{and} \quad \eta_i^{H,k} = 0 \; (i \in I_{+})\,. \tag{16}$$

Thus, (15) can be written as

$$0 = \nabla f(x^k) - \sum_{i\in I_0} \eta_i^{H,k} \nabla H_i(x^k) + \sum_{i\in I_{+0}\cup I_{00}} \eta_i^{G,k} \nabla G_i(x^k).$$

This may be expressed in matrix-vector notation as $f^k = A_k^T \eta_I^k$, where we put

$$f^k := \nabla f(x^k), \quad \eta_I^k := \begin{pmatrix} \eta^{G,k}_{I_{00}\cup I_{+0}} \\ \eta^{H,k}_{I_0} \end{pmatrix} \quad \text{and} \quad A_k := \begin{pmatrix} -\nabla G_i(x^k) & (i \in I_{+0}\cup I_{00}) \\ \nabla H_i(x^k) & (i\in I_0) \end{pmatrix}.$$

Now, A_k converges to a matrix with full rank by the MPVC-LICQ assumption and as f^k is convergent, too, necessarily the sequence η_I^k is also convergent, say to a limit point $\eta_I = (\eta^G_{I_{00}\cup I_{+0}}, \eta^H_{I_0})$. In view of (16), we have $\eta_i^G = 0$ for all $i \in I_{0-}\cup I_{+-}\cup I_{0+}$ and $\eta_i^H = 0$ $(i \in I_+)$. Hence it remains to show that $\eta_i^G \geq 0$ for all $i \in I_{+0}\cup I_{00}$ and $\eta_i^H \geq 0$ for all $i \in I_{0-}$

From (14) we see that $\nu_i^k \geq 0$ for all $i = 1,\ldots,l$. Furthermore, Lemma 4.5 yields $D_1\Phi(G_i(x^k), H_i(x^k); t_k)) \geq 0$ and hence, $\eta_i^{G,k} = \nu_i^k D_1\Phi(G_i(x^k), H_i(x^k); t_k) \geq 0$ for all i. In particular, we infer that $\eta_i^G \geq 0$ for $i \in I_{00}\cup I_{+0}$.

In turn, for $i \in I_{0-}$, due to Lemma 4.9, we know that, for k sufficiently large, $i \notin M(x^k, t_k)$. Thus, we have $\eta_i^{H,k} = \rho_i^k \geq 0$ and hence, $\eta_i^H \geq 0$.

Altogether, this shows that (x^*, η^G, η^H) is a weakly stationary point of (2). □

We note that the corresponding result in [37] shows C-stationarity for the limit point in the MPEC setting, whereas, here, we only have weak stationarity. However, recall that this is mainly a problem of terminology, since weak stationarity for MPVCs is much stronger than weak stationarity for MPECs, and there exists no counterpart of C-stationarity for MPVCs.

In order to obtain more than weak stationarity, we need to assume SSOSC for the subproblems $P(t_k)$.

Theorem 4.11 *Let the assumptions and definitions of Theorem 4.10 hold and assume in addition that SSOSC holds at x^k for $P(t_k)$ for all $k \in \mathbb{N}$ (sufficiently large). Then the point $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is M-stationary for* (2).

Proof. Again, we assume throughout this proof that the MPVC (2) contains no standard constraints in order to focus on the main difficulties.

Recall that due to Theorem 4.10 the sequences $\eta_i^{G,k}$ and $\eta_i^{H,k}$ are convergent to limit points η_i^G and η_i^H, respectively, such that (x^*, η^G, η^H) is weakly stationary. Now, suppose that it is not M-stationary. Then there exists an index $j \in I_{00}$ such that $\eta_j^G > 0$ and $\eta_j^H \neq 0$. The fact that $\eta_j^G > 0$, implies that $\eta_j^{G,k} = \nu_j^k D_1\Phi(G_j(x^k), H_j(x^k); t_k) > 0$ for all k sufficiently large. Thus, in particular, we see that $\nu_j^k > 0$ and hence, from the complementarity in (14), we infer that

$$\Phi(G_j(x^k), H_j(x^k); t_k) = 0 \tag{17}$$

for k sufficiently large, that is $j \in M(x^k, t_k)$.

Furthermore, since we also have $D_1\Phi(G_j(x^k), H_j(x^k); t_k) > 0$, Lemma 4.5 yields

$$G_j(x^k) - H_j(x^k) < t_k \tag{18}$$

for k sufficiently large.

We now argue that the case $\eta_j^H > 0$ cannot occur: To this end, note that $\eta_j^H > 0$ implies $\eta_j^{H,k} = \rho_j^k - \nu_j^k D_2\Phi(G_j(x^k), H_j(x^k); t_k) > 0$ and hence, also invoking Lemma 4.5 and Definition 2.2, we have $\rho_j^k > \nu_j^k D_2\Phi(G_j(x^k), H_j(x^k); t_k) \geq 0$, and thus, in view of (14), we get $H_j(x^k) = 0$ for all k sufficiently large. Using (18), this yields $G_j(x^k) < t_k$. Invoking (17) and $H_j(x^k) = 0$, we obtain

$$0 = \Phi(G_j(x^k), H_j(x^k); t_k) = G_j(x^k) - \varphi(G_j(x^k); t_k)$$

Together with Lemma 4.4, this implies $G_j(x^k) \geq t_k > 0$ in contrast to (18) and $H_j(x^k) = 0$.

Thus, we only need to consider the case $\eta_j^H < 0$. At this, we clearly have $\eta_j^{H,k} = \rho_j^k - \nu_j^k D_2\Phi(G_j(x^k), H_j(x^k); t_k) < 0$ for all k sufficiently large. Hence, we obtain $0 \leq \rho_j^k < \nu_j^k D_2\Phi(G_j(x^k), H_j(x^k); t_k)$ and thus, in particular, $D_2\Phi(G_j(x^k), H_j(x^k); t_k) > 0$ which, by Lemma 4.5, gives $G_j(x^k) - H_j(x^k) > -t_k$. Together with (18) this yields

$$|G_j(x^k) - H_j(x^k)| < t_k\,. \tag{19}$$

Thus,

$$0 = \Phi(G_j(x^k), H_j(x^k); t_k) \overset{(19)}{=} G_j(x^k) + H_j(x^k) - t_k\theta\Big(\frac{G_j(x^k) - H_j(x^k)}{t_k}\Big).$$

Hence, Lemma 4.5 and an easy calculation yields $G_j(x^k), H_j(x^k) > 0$, that is $j \in I_{++}(x^k)$. In particular, we then have $\rho_j^k = 0$ and hence $\eta_j^{H,k} = -\nu_j^k D_2\Phi(G_j(x^k), H_j(x^k); t_k)$. This, invoking (19), Lemma 4.5 and the definition of $\eta_j^{G,k}, \eta_j^{H,k}$ gives

$$\begin{aligned} \eta_j^{G,k} &= \nu_j^k\Bigg(1 - \theta'\Big(\frac{G_j(x^k) - H_j(x^k)}{t_k}\Big)\Bigg), \\ \eta_j^{H,k} &= -\nu_j^k\Bigg(1 + \theta'\Big(\frac{G_j(x^k) - H_j(x^k)}{t_k}\Big)\Bigg) \end{aligned}$$

for all k sufficiently large. Putting $u_k := 1 - \theta'\Big(\frac{G_j(x^k) - H_j(x^k)}{t_k}\Big)$, this can be rewritten as

$$\eta_j^{G,k} = \nu_j^k u_k, \quad \eta_j^{H,k} = -\nu_j^k(2 - u_k). \tag{20}$$

Since we have $\eta_j^{G,k} \to \eta_j^G > 0$ and $\eta_j^{H,k} \to \eta_j^H < 0$, from (20) we can quickly deduce that the sequence $\{\nu_j^k\}$ is bounded, $\nu_j^k, u_k \not\to 0$ and $\frac{2-u_k}{u_k} \to -\frac{\eta_j^H}{\eta_j^G} > 0$.

Now, for $k \in \mathbb{N}$, choose a vector $d^k \in \mathbb{R}^n$ with the following properties:

$$\begin{array}{rcll} \nabla G_i(x^k)^T d^k & = & 0 & (i \in (I_{00} \cup I_{+0}) \setminus \{j\}), \\ \nabla H_i(x^k)^T d^k & = & 0 & (i \in I_0 \setminus \{j\}), \\ \nabla G_j(x^k)^T d^k & = & -\frac{2-u_k}{u_k}, & \\ \nabla H_j(x^k)^T d^k & = & 1, & \end{array} \tag{21}$$

which is possible due to MPVC-LICQ. Then, the sequence $\{d^k\}$ can be chosen such that it is bounded due to the convergence of the right hand side in (21) and the linear independence of the gradients occuring on the left hand side. Moreover, we have

$$d^k \in \mathcal{C}(x^k) \;=\; \{d \in \mathbb{R}^n \mid \begin{array}{lcll} \nabla H_i(x^k)^T d & = & 0 & (i : \rho_i^k > 0), \\ \nabla(\Phi(G_i(x^k), H_i(x^k); t_k))^T d & = & 0 & (i : \nu_i^k > 0)\} \end{array} \tag{22}$$

for all $k \in \mathbb{N}$, which can be seen as follows: The first defining condition is satisfied since we have $\{i \mid \rho_i^k > 0\} \subseteq I_0(x^k) \subseteq I_0$ for k sufficiently large and $\rho_j^k = 0$ as was stated earlier. To verify the second condition we first note that

$$\begin{array}{rcl} \nabla(\Phi(G_i(x^k), H_i(x^k); t_k))^T d^k & = & D_1\Phi(G_i(x^k), H_i(x^k); t_k)\nabla G_i(x^k)^T d^k \\ & & +D_2\Phi(G_i(x^k), H_i(x^k); t_k)\nabla H_i(x^k)^T d^k \,. \end{array} \tag{23}$$

Furthermore, we recall that, due to Lemma 4.9, we have $\{i \mid \nu_i^k > 0\} \subseteq M(x^k, t_k) \subseteq I_{00} \cup I_{0+} \cup I_{+0}$. In view of the choice of d^k and (23) it is immediately clear that for $i \in I_{00} \cap M(x^k, t_k)$ with $i \neq j$ we have $\nabla(\Phi(G_i(x^k), H_i(x^k); t_k))^T d^k = 0$. For the case of $i \in I_{0+} \cap M(x^k, t_k)$ it was argued in the proof of Theorem 4.10 that $D_1\Phi(G_i(x^k), H_i(x^k); t_k) = 0$ and thus, in view of (23), the second condition reduces to $D_2\Phi(G_i(x^k), H_i(x^k); t_k)\nabla H_i(x^k)^T d^k{=}0$, which is then, due to its choice, satisfied for d^k. An analogous reasoning will work for $i \in I_{+0} \cap M(x^k, t_k)$: The proof of Theorem 4.10 shows that $D_2\Phi(G_i(x^k), H_i(x^k); t_k) = 0$ in this case, hence the second condition reduces to $D_1\Phi(G_i(x^k), H_i(x^k); t_k)\nabla G_i(x^K)^T d^k = 0$ which holds by the choice of d^k in (21). Thus, it remains to verify the second condition for the index j. For these purposes, note that we have $D_1\Phi(G_j(x^k), H_j(x^k); t_k) = u_k$ and $D_2\Phi(G_j(x^k), H_j(x^k); t_k) = 2 - u_k$ due to (19) and Lemma 4.5, and hence, invoking the choice of d^k and (23), we obtain

$$\nabla\Phi(G_j(x^k), H_j(x^k); t_k)^T d^k = u_k(-\frac{2-u_k}{u_k}) + (2 - u_k) = 0,$$

which eventually yields $d^k \in \mathcal{C}(x^k)$ for all k sufficiently large. Now, we compute the Hessian of the Lagrangian $\mathcal{L}_t$ of the program $P(t)$:

$$
\begin{aligned}
&\nabla^2_{xx}\mathcal{L}_t(x^k,\rho^k,\nu^k)\\
&= \nabla^2 f(x^k) - \sum_{i=1}^{l}\rho_i^k\nabla^2 H_i(x^k) + \sum_{i=1}^{l}\nu_i^k\nabla^2(\Phi(G_i(x^k),H_i(x^k);t_k))\\
&= \nabla^2 f(x^k) - \sum_{i=1}^{l}\eta_i^{H,k}\nabla^2 H_i(x^k) + \sum_{i=1}^{l}\eta_i^{G,k}\nabla^2 G_i(x^k)\\
&\quad + \sum_{i\in M(x^k,t_k)}\nu_i^k D_{11}\Phi(G_i(x^k),H_i(x^k);t_k)\nabla G_i(x^k)\nabla G_i(x^k)^T\\
&\quad + \sum_{i\in M(x^k,t_k)}\nu_i^k D_{12}\Phi(G_i(x^k),H_i(x^k);t_k)(\nabla G_i(x^k)\nabla H_i(x^k)^T + \nabla H_i(x^k) G_i(x^k)^T)\\
&\quad + \sum_{i\in M(x^k,t_k)}\nu_i^k D_{22}\Phi(G_i(x^k),H_i(x^k);t_k)\nabla H_i(x^k)\nabla H_i(x^k)^T\\
&= \nabla^2 f(x^k) - \sum_{i=1}^{l}\eta_i^{H,k}\nabla^2 H_i(x^k) + \sum_{i=1}^{l}\eta_i^{G,k}\nabla^2 G_i(x^k)\\
&\quad + \sum_{i\in M(x^k,t_k)\cap I_{00}}\nu_i^k D_{11}\Phi(G_i(x^k),H_i(x^k);t_k)\nabla G_i(x^k)\nabla G_i(x^k)^T\\
&\quad + \sum_{i\in M(x^k,t_k)\cap I_{00}}\nu_i^k D_{12}\Phi(G_i(x^k),H_i(x^k);t_k)(\nabla G_i(x^k)\nabla H_i(x^k)^T + \nabla H_i(x^k)\nabla G_i(x^k)^T)\\
&\quad + \sum_{i\in M(x^k,t_k)\cap I_{00}}\nu_i^k D_{22}\Phi(G_i(x^k),H_i(x^k);t_k)\nabla H_i(x^k)\nabla H_i(x^k)^T\,,
\end{aligned}
\tag{24}
$$

where the last equality can be seen as follows: From Lemma 4.9 we have the inclusion $M(x^k,t_k)\subseteq I_{0+}\cup I_{00}\cup I_{+0}$. Then it can be argued, cf. the proof of Theorem 4.10, that we have $|G_i(x^k)-H_i(x^k)|\geq t_k$ for $i\in M(x^k,t_k)\cap I_{0+}$ and hence, due to Lemma 4.5, all second-order partial derivatives of $\Phi(G_i(x^k),H_i(x^k);t_k)$ are zero for k sufficiently large. The same observation holds for all $i\in M(x^k,t_k)\cap I_{+0}$ and all k sufficiently large, which shows the last equality in (24).

From (24) and the properties of d^k we thus infer that

$$
\begin{aligned}
&(d^k)^T\nabla^2_{xx}\mathcal{L}_t(x^k,\rho^k,\nu^k)d^k\\
&= (d^k)^T\nabla^2 f(x^k)(d^k) - \sum_{i=1}^{l}\eta_i^k(d^k)^T\nabla^2 H_i(x^k)d^k + \sum_{i=1}^{l}\eta_i^{G,k}(d^k)^T\nabla^2 G_i(x^k)d^k\\
&\quad +\nu_j^k D_{11}\Phi(G_j(x^k),H_j(x^k);t_k)\Big(\frac{2-u_k}{u_k}\Big)^2 - 2\nu_j^k D_{12}\Phi(G_j(x^k),H_j(x^k);t_k)\Big(\frac{2-u_k}{u_k}\Big)\\
&\quad +\nu_j^k D_{22}\Phi(G_j(x^k),H_j(x^k);t_k)\,.
\end{aligned}
$$

At this, the first three summands are bounded for $k\in\mathbb{N}$, due to the convergence of the multipliers, the boundedness of $\{d^k\}$ and the continuity of $\nabla^2 f,\nabla^2 H_i$ and $\nabla^2 G_i$ for $i=1,\dots,l$. In turn, for the remaining summands we compute, by means of Lemma 4.5, that

$$\begin{aligned}
&\nu_j^k D_{11}\Phi(G_j(x^k),H_j(x^k);t_k)\Big(\tfrac{2-u_k}{u_k}\Big)^2-2\nu_j^k D_{12}\Phi(G_j(x^k),H_j(x^k);t_k)\tfrac{2-u_k}{u_k}+\nu_j^k D_{22}\Phi(G_j(x^k),H_j(x^k);t_k)\\
&= \ \nu_j^k\Big(-\tfrac{1}{t_k}\theta''\Big(\tfrac{G_j(x^k)-H_j(x^k)}{t_k}\Big)\Big(\tfrac{2-u_k}{u_k}\Big)^2-2\tfrac{1}{t_k}\theta''\Big(\tfrac{G_j(x^k)-H_j(x^k)}{t_k}\Big)\tfrac{2-u_k}{u_k}-\tfrac{1}{t_k}\theta''\Big(\tfrac{G_j(x^k)-H_j(x^k)}{t_k}\Big)\Big)\\
&= \ -\tfrac{\nu_j^k}{t_k}\theta''\Big(\tfrac{G_j(x^k)-H_j(x^k)}{t_k}\Big)\Big(\Big(\tfrac{2-u_k}{u_k}\Big)^2+2\tfrac{2-u_k}{u_k}+1\Big)\to_{k\to\infty}-\infty\,,
\end{aligned}$$

where the asymptotic behaviour is due to the boundedness of $\{\nu_j^k\}(\not\to 0)$, the continuity of θ'', the fact that $\frac{2-u_k}{u_k}\to-\frac{\eta_j^H}{\eta_j^G}>0$ and $t_k\to 0$. In particular, this yields that for k sufficiently large we have $(d^k)^T\nabla^2_{xx}\mathcal{L}_t(x^k,\rho^k,\nu^k)d^k<0$ in contradiction to the fact that $d^k\in\mathcal{C}(x^k)$, which is actually the critical cone of SSOSC for $P(t_k)$ at x^k, which is supposed to hold for all k. □

We now present our third main result. It is similar to Theorem 4.10 in the sense that we get a weakly stationary point for our MPVC under the weaker MPVC-CPLD assumption (whereas in Theorem 4.10, MPVC-LICQ was assumed). However, we do not get convergence of the multipliers. The proof shows boundedness of certain scalars which then can be used to construct suitable multipliers showing that the limit point is indeed weakly stationary.

Theorem 4.12 *Let $(x^k,\lambda^k,\mu^k,\rho^k,\nu^k)$ be a sequence of KKT points of $P(t_k)$ for $t_k\downarrow 0$. Furthermore, let x^k converge to the point x^* satisfying MPVC-CPLD. Then x^* is a weakly stationary point of the MPVC* (2).

Proof. Similar to the previous proofs, we assume without loss of generality that no standard constraints are involved in our MPVC from (2).

The feasibility of x^* for the MPVC follows from the definition of $P(t_k)$ and $t_k\downarrow 0$. Furthermore, the KKT conditions of $P(t_k)$ yield the following properties of (x^k,ρ^k,ν^k) for all $k\in\mathbb{N}$:

$$0=\nabla f(x^k)-\sum_{i=1}^{l}\rho_i^k\nabla H_i(x^k)+\sum_{i=1}^{l}\nu_i^k(\alpha_i^k\nabla G_i(x^k)+(2-\alpha_i^k)\nabla H_i(x^k))$$

and

$$\begin{aligned}
\operatorname{supp}(\rho^k)\subseteq I_0(x^k),&\qquad \rho_i^k>0\ \forall i\in\operatorname{supp}(\rho^k),\\
\operatorname{supp}(\nu^k)\subseteq M(x^k,t_k),&\qquad \nu_i^k>0\ \forall i\in\operatorname{supp}(\nu^k),
\end{aligned}$$

where, according to Lemma 4.5, we have

$$\alpha_i^k=\begin{cases}0, & G_i(x^k)\geq H_i(x^k)+t_k,\\ 1-\theta'\left(\frac{G_i(x^k)-H_i(x^k)}{t_k}\right), & |G_i(x^k)-H_i(x^k)|<t_k,\\ 2, & H_i(x^k)\geq G_i(x^k)+t_k.\end{cases}$$

Now, define the multipliers

$$\delta_i^{G,k} := \nu_i^k \alpha_i^k, \quad \delta_i^{H,k} := (2-\alpha_i^k)\nu_i^k \quad \forall i = 1\ldots,l$$

for all $k \in \mathbb{N}$. Thus, it follows that

$$\begin{aligned} \operatorname{supp}(\delta^{G,k}) \subseteq M(x^k,t_k)\backslash I_0(x^k), \qquad & \delta_i^{G,k} > 0 \ \forall i \in \operatorname{supp}(\delta^{G,k}), \\ \operatorname{supp}(\delta^{H,k}) \subseteq M(x^k,t_k)\backslash\big(I_{00}(x^k) \cup I_{+0}(x^k)\big), \qquad & \delta_i^{H,k} > 0 \ \forall i \in \operatorname{supp}(\delta^{H,k}). \end{aligned}$$

The first statement is due to the fact that $\Phi(G_i(x^k), H_i(x^k); t_k) = 0 = H_i(x^k)$ would imply (cf. Lemma 4.6) that $G_i(x^k) \geq t_k$ and hence $\alpha_i^k = 0$. The second statement can be seen in a similar fashion.

Now, for all $k \in \mathbb{N}$ the following inclusions hold true:

$$\begin{aligned} &\operatorname{supp}(\delta^{G,k}) \subseteq M(x^k,t_k)\backslash I_0(x^k) \subseteq I_{00} \cup I_{+0}, \\ &\operatorname{supp}(\delta^{H,k}) \subseteq M(x^k,t_k)\backslash\big(I_{00}(x^k) \cup I_{+0}(x^k)\big) \subseteq I_{00} \cup I_{0+}. \end{aligned}$$

The first statement is a consequence of Lemma 4.9 (b) and the fact that $i \in I_{0+}$ would imply $G_i(x^k) - H_i(x^k) > t_k$ for k sufficiently large and hence, in view of Lemma 4.6, we would have $H_i(x^k) = 0$ to stay feasible. The second statement can be verified analogously.

By modifying the multipliers $(\rho^k, \delta^{G,k}, \delta^{H,k})$ in a suitable way, we can find multipliers $(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})$ such that $\bar{\rho}^k \geq 0$, $\bar{\delta}^{G,k} \geq 0$, $\bar{\delta}^{H,k} \geq 0$,

$$\begin{aligned} \operatorname{supp}(\bar{\rho}^k) \ &\subseteq \ \operatorname{supp}(\rho^k) \subseteq I_0(x^k), \\ \operatorname{supp}(\bar{\delta}^{G,k}) \ &\subseteq \ \operatorname{supp}(\delta^{G,k}) \subseteq M(x^k,t_k)\backslash I_0(x^k), \\ \operatorname{supp}(\bar{\delta}^{H,k}) \ &\subseteq \ \operatorname{supp}(\delta^{H,k}) \subseteq M(x^k,t_k)\backslash\big(I_{00}(x^k) \cup I_{+0}(x^k)\big), \end{aligned}$$

and

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

and all gradients from the equation above with nonvanishing multipliers are linearly independent, cf. [37, Lemma A.1]. Obviously, this implies $\operatorname{supp}(\bar{\rho}^k) \cap \operatorname{supp}(\bar{\delta}^{H,k}) = \emptyset$ for all $k \in \mathbb{N}$. In order to prove that the sequence $\{(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})\}$ is bounded, we assume the contrary. Then there exists a subset $K \subseteq \mathbb{N}$ such that $\big\|\{(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})\}\big\| \to_K \infty$ and

$$\frac{\{(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})\}}{\big\|\{(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})\}\big\|} \to_K (\tilde{\rho}, \tilde{\delta}^G, \tilde{\delta}^H) \neq (0,0,0).$$

By continuity of f, G, and H, this implies

$$0 = -\sum_{i=1}^{l} \tilde{\rho}_i \nabla H_i(x^*) + \sum_{i=1}^{l} \tilde{\delta}_i^G \nabla G_i(x^*) + \sum_{i=1}^{l} \tilde{\delta}_i^H \nabla H_i(x^*), \tag{25}$$

where for all $k \in K$ sufficiently large one has

$$\operatorname{supp}(\tilde{\rho}) \ \subseteq \ \operatorname{supp}(\bar{\rho}^k) \subseteq I_0(x^k) \subseteq I_0,$$

$$\begin{aligned}\mathrm{supp}(\tilde{\delta}^G) &\subseteq \mathrm{supp}(\bar{\delta}^{G,k}) \subseteq M(x^k,t_k)\backslash I_0(x^k) \subseteq \big(I_{00}\cup I_{+0}\big),\\ \mathrm{supp}(\tilde{\delta}^H) &\subseteq \mathrm{supp}(\bar{\delta}^{H,k}) \subseteq M(x^k,t_k)\backslash\big(I_{00}(x^k)\cup I_{+0}(x^k)\big) \subseteq I_{00}\cup I_{0+},\end{aligned}$$

as well as $\mathrm{supp}(\tilde{\rho})\cap \mathrm{supp}(\tilde{\delta}^H) \subseteq \mathrm{supp}(\bar{\rho}^k)\cap \mathrm{supp}(\bar{\delta}^{H,k}) = \emptyset$. Moreover, we have $\tilde{\delta}^G_i > 0$ for all $i \in \mathrm{supp}(\tilde{\delta}^G)$ and

$$I_{0-}\cap \big(\mathrm{supp}(\tilde{\rho})\cup \mathrm{supp}(\tilde{\delta}^H)\big) = I_{0-}\cap \mathrm{supp}(\tilde{\rho}),$$

because $G_i(x^*) < 0$ implies $i \notin M(x^k,t_k)$ for all k sufficiently large. In view of (25), it follows that the set

$$\Big\{\{-\nabla H_i(x^*) \mid i \in \mathrm{supp}(\tilde{\rho})\cap I_{0-}\} \cup \{\nabla G_i(x^*) \mid i \in \mathrm{supp}(\tilde{\delta}^G)\}\Big\}$$
$$\cup\big\{\nabla H_i(x^*) \mid i \in \mathrm{supp}(\tilde{\delta}^H)\cup(\mathrm{supp}(\tilde{\rho})\backslash I_{0-})\big\}$$

is positive-linearly dependent (note that, here, we use the splitting $\mathrm{supp}(\tilde{\rho}) = \big(\mathrm{supp}(\tilde{\rho}\cap I_{0-}\big) \cup \big(\mathrm{supp}(\tilde{\rho}) \setminus I_{0-}\big)$ and moved those terms from (25) belonging to the index set $\mathrm{supp}(\tilde{\rho}) \setminus I_{0-}$ from the first to the last sum which is possible since MPVC-CPLD requires no sign constraints for the corresponding multipliers). Hence the MPVC-CPLD implies linear dependence of these gradients in x^k for all k sufficiently large. This, however, is a contradiction to our choice of $(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})$. Thus, our assumption was wrong and the sequence $(\bar{\rho}^k, \bar{\delta}^{G,k}, \bar{\delta}^{H,k})$ is bounded and consequently has at least one accumulation point $(\bar{\rho}, \bar{\delta}^G, \bar{\delta}^H)$. This point then satisfies

$$0 = \nabla f(x^*) - \sum_{i=1}^{l} \bar{\rho}_i \nabla H_i(x^*) + \sum_{i=1}^{l} \bar{\delta}^G_i \nabla G_i(x^*) + \sum_{i=1}^{l} \bar{\delta}^H_i \nabla H_i(x^*)$$

and

$$\begin{aligned}&\mathrm{supp}(\bar{\delta}^G) \subseteq I_{00}\cup I_{+0}, \bar{\delta}^G_i > 0\ \forall i \in \mathrm{supp}(\bar{\delta}^G),\\ &\mathrm{supp}(\bar{\rho})\cup \mathrm{supp}(\bar{\delta}^H) \subseteq I_0,\\ &\mathrm{supp}(\bar{\rho})\cap \mathrm{supp}(\bar{\delta}^H) = \emptyset,\\ &I_{0-}\cap\big(\mathrm{supp}(\bar{\rho})\cup \mathrm{supp}(\bar{\delta}^H)\big) = I_{0-}\cap \mathrm{supp}(\bar{\rho}).\end{aligned}$$

Hence, the multipliers $\eta^G := \bar{\delta}^G$ and

$$\eta^H_i := \begin{cases} \bar{\rho}_i, & \text{if } i \in \mathrm{supp}(\bar{\rho}),\\ -\bar{\delta}^H, & \text{if } i \in \mathrm{supp}(\bar{\delta}^H),\\ 0, & \text{else}\end{cases}$$

are well defined and (x^*, η^H, η^G) is a weakly stationary point of the MPVC (2) (note that $\bar{\delta}^H_i = 0$ for all $i \in I_{0-}$). □

The following corollary is an immediate consequence of the latter result and the fact that both MPVC-MFCQ and MPVC-CRCQ imply MPVC-CPLD.

Corollary 4.13 *Let $\{(x^k, \lambda^k, \mu^k, \rho^k, \nu^k)\}$ be a sequence of KKT points of $P(t_k)$ for $t_k \downarrow 0$. Furthermore, let x^k converge to the point x^* satisfying MPVC-MFCQ or MPVC-CRCQ. Then x^* is a weakly stationary point of the MPVC* (2)*.*

We close this section with a simple observation: Since the local regularization changes the feasible set of the MPVC only locally around the origin, it follows that a local minimum x^* of the MPVC satisfying strict complementarity (i.e. either $G_i(x^*) > 0$ or $H_i(x^*) > 0$ holds for all $i = 1, \ldots, l$) is also a local minimum of the regularized problem $P(t)$ for all $t > 0$ sufficiently small, and vice versa. In this case, it is therefore not necessary to push the regularization parameter t down to zero.

5 Final Remarks

The results presented here are twofold. On the one hand, an existing convergence theorem for a local regularization approach for MPECs has been improved using a problem-tailored variant of the so-called CPLD condition. On the other hand, based on the ideas for MPECs, a local regularization scheme for MPVCs has been introduced and a number of convergence theorems could be proven.

An open problem and interesting future research topic is the question whether the convergence results under the CPLD-type conditions can also be shown to hold for the global regularization method by Scholtes [33] in the MPEC setting or for a corresponding method in the MPVC setting.

Acknowledgement: The second author would like to thank Gabriel Haeser from the University of Campinas for pointing his attention to the CPLD condition as a new constraint qualification for nonlinear programs.

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.* Technical Report, Institute of Mathematics, University of Würzburg, November 2008.

[3] W. Achtziger, T. Hoheisel, and C. Kanzow: *On a relaxation method for mathematical programs with vanishing constraints and its application in topology optimization.* Technical Report, Institute of Mathematics, University of Würzburg, forthcoming.

[4] 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.

[5] M. Anitescu: *On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints.* SIAM Journal on Optimization 15, 2005, pp. 1203–1236.

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

[7] D.P. Bertsekas and A.E. Ozdaglar: *Pseudonormality and a Lagrange multiplier theory for constrained optimization.* Journal of Optimization Theory and Applications 114, 2002, pp. 187–343.

[8] 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.

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

[10] 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.

[11] M.L. Flegel and C. Kanzow: *On M-stationary points for mathematical programs with equilibrium constraints.* Journal of Mathematical Analysis and Applications 310, 2005, pp. 286–302.

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

[13] M.L. Flegel, C. Kanzow, and J.V. Outrata: *Optimality conditions for disjunctive programs with applications to mathematical programs with equilibrium constraints.* Set-Valued Analysis 15, 2007, pp. 139-162.

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

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

[16] 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.

[17] 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)

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

[19] 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.

[20] A.F. Izmailov, A.L. Pogosyan, and M.V. Solodov: *Semismooth Newton method for the lifted reformulation of mathematical programs with complementarity constraints.* Technical Report, Institute of Pure and Applied Mathematics, Rio de Janeiro, Brazil, December 2009.

[21] R. Janin: *Directional derivative of the marginal function in nonlinear programming.* Mathematical Programming Study 21, 1984, pp. 110–126.

[22] C. Kanzow and A. Schwartz: *Mathematical programs with equilibrium constraints: Enhanced Fritz John-conditions, new constraint qualifications and improved exact penalty results.* Technical Report, Institute of Mathematics, University of Würzburg, Würzburg, October 2009.

[23] S. Leyffer: *MacMPEC: AMPL collection of MPECs.* Argonne National Laboratory, 2000, www.mcs.anl.gov/leyffer/MacMPEC.

[24] S. Leyffer, G. López-Calva, and J. Nocedal: *Interior methods for mathematical programs with complementarity constraints.* SIAM Journal on Optimization 17, 2007, pp. 52–77.

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

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

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

[28] 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.

[29] 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.

[30] 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.

[31] A.U. Raghunathan and L.T. Biegler: *An interior point method for mathematical programs with complementarity constraints (MPCCs).* SIAM Journal on Optimization 15, 2005, pp. 720–750.

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

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

[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] O. Stein: *Lifting mathematical programs with complementarity constraints.* Technical Report, Institute of Operations Research, Karlsruhe Institute of Technology, Germany, 2008.

[36] S. Veelken: *A New Relaxation Scheme for Mathematical Programs with Equilibrium Constraints: Theory and Numerical Experience.* Ph.D. Thesis, Faculty of Mathematics, Technical University of Munich, 2009.

[37] S. Veelken and M. Ulbrich: *A new regularization scheme for mathematical programs with equilibrium constraints.* Technical Report, Faculty of Mathematics, Technical University of Munich, February 2009.

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

[39] 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.

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