

**A NEW REGULARIZATION METHOD
FOR MATHEMATICAL PROGRAMS WITH
COMPLEMENTARITY CONSTRAINTS WITH
STRONG CONVERGENCE PROPERTIES¹**

Christian Kanzow and Alexandra Schwartz

Preprint 296

June 2010

University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany

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

June 30, 2010

¹This 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 complementarity) constraints, MPECs for short, form a difficult class of optimization problems. The feasible set has a very special structure and violates most of the standard constraint qualifications. Therefore, one typically applies specialized algorithms in order to solve MPECs. One very prominent class of specialized algorithms are the regularization (or relaxation) methods. The first regularization method for MPECs is due to Scholtes [*SIAM Journal on Optimization* 11, 2001, pp. 918–936], but in the meantime, there exist a number of different regularization schemes which try to relax the difficult constraints in different ways. However, almost all regularization methods converge to C-stationary points only, which is a very weak stationarity concept. An exception is a recent method by Kadrani, Dus-sault, and Benchakroun [*SIAM Journal on Optimization* 20, 2009, pp. 78–103] whose limit points are shown to be M-stationary. Here we provide a new regularization method which also converges to M-stationary points. The assumptions to prove this result are significantly weaker than for all other relaxation schemes. Furthermore, our relaxed problem has a much more favourable geometric shape than the one proposed by Kadrani et al.

Key Words: Mathematical programs with complementarity constraints, regularization method, global convergence, M-stationarity, strong stationarity, constraint qualifications.

1 Introduction

A mathematical program with complementarity (or equilibrium) constraints, MPEC for short, is a constrained optimization problem of the form

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

where $f, g_i, h_i, G_i, H_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are assumed to be continuously differentiable functions. Hence an MPEC consists of an objective function f which is to be minimized subject to some standard inequality and equality constraints defined by the mappings g_i and h_i , respectively, as well as by some additional complementarity-type constraints. In many applications, these complementarity constraints either arise directly from an equilibrium condition, or they are part of the optimality conditions from a (convex) lower level problem. The interested reader is referred to the two monographs [27, 32] for an introduction to and many applications of MPECs, as well as to the book [7] on the closely related class of bilevel programs.

The main problem, both from a theoretical and a numerical point of view, for the solution of MPECs comes from the complementarity constraints. In fact, these constraints imply that almost all of the constraint qualifications known for standard nonlinear programs are violated. This, in turn, means that the convergence assumptions for basically all standard methods for the solution of constrained optimization problems are not satisfied. During the last decade, several authors therefore proposed different solution algorithms which take into account the particular structure of an MPEC and try to avoid the problems arising from the complementarity constraints in one or another way. We refer the reader to [3, 4, 6, 8, 13, 19, 20, 21, 24, 25, 34, 35, 37, 39, 38] and references therein, where a number of different algorithmic ideas like penalization, smoothing, lifting, and regularization are used to overcome the inherent difficulty of an MPEC.

One of the most popular approaches for the solution of an MPEC is certainly the regularization scheme by Scholtes [37]. Besides this particular method, there are, in the meantime, a number of other regularization schemes available which try to relax the complementarity constraints in a different way. The regularization methods we are currently aware of are the following ones:

- the global regularization method by Scholtes [37],
- the two-sided regularization method by DeMiguel et al. [6],
- the smooth regularization method by Lin and Fukushima [25],
- the local regularization method by Steffensen and Ulbrich [38],
- the nonsmooth regularization method by Kadrani et al. [21].

The convergence results for these methods show that the first three methods from [37, 6, 25] converge to a C-stationary point under the MPEC-LICQ assumption (precise definitions of the different stationary concepts and MPEC-tailored constraint qualifications are given in Section 2), whereas the fourth method from [38] gives convergence to a C-stationary point under the weaker MPEC-CRCQ condition (this assumption was further relaxed in the recent paper [18]). Finally, the last method shows convergence to M-stationary points, again under the MPEC-LICQ assumption. This is a very interesting property since M-stationary is a much stronger optimality criterion than C-stationarity. Convergence to M-stationary points can also be shown for the other methods, but only under additional assumptions that are not required in [21].

The aim of this paper is to introduce a new regularization scheme with stronger properties than those methods considered previously. In particular, we show that our new method has the following nice features:

- the limit points are at least M-stationary points,
- convergence to M- and even strongly stationary points can be shown under conditions that are much weaker than those used by other methods,
- the shape of the feasible set of our regularized problem is much nicer than the one of the corresponding method by Kadrani et al. [21].

Hence we get the best convergence result that is currently only known for the very recent method by Kadrani et al. [21], but under significantly weaker assumptions and for a regularization that we believe is much easier to handle from a numerical point of view than the nonsmooth regularization from [21].

To this end, we organize our paper as follows: The next section recapitulates some stationarity concepts and constraint qualifications for MPECs as well as for standard nonlinear programs. Section 3 introduces our new relaxation and states some useful properties of the relaxed problem, whereas Section 4 is concerned with the convergence properties of our method. Some numerical results are given in Section 5 and a conclusion is drawn in Section 6.

Most of the notation used in this paper is standard. For a continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, 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. The support of a vector $a \in \mathbb{R}^n$ is defined by $\text{supp}(a) := \{i \in \{1, \dots, n\} \mid a_i \neq 0\}$. Furthermore, given two vectors $x, y \in \mathbb{R}^q$, we write $0 \leq x \perp y \geq 0$ as a shorthand for $x \geq 0, y \geq 0, x^T y = 0$.

2 Preliminaries

2.1 Constraint Qualifications for Standard Nonlinear Programs

Although the main topic of this paper are MPECs, the relaxed problems are standard nonlinear programs. Hence, we need some standard constraint qualifications to guarantee the existence of Lagrange multipliers in local minima of the relaxed problems. By now, there is a whole variety of constraint qualifications for nonlinear programs, thus we are

going to mention only those needed later in this work. Consider the following nonlinear program

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

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

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

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

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

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*

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

are linearly independent;

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

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

The second implication was proven in [2], 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^* is a *stationary point* of (2), i.e. there are multipliers λ_i ($i = 1, \dots, m$) and μ_i ($i = 1, \dots, p$) such that the triple (x^*, λ, μ) is a *KKT-point* meaning that

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

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

2.2 Stationary Points for MPECs

In contrast to standard nonlinear programs, several stationarity concepts are known for MPECs. To state them, we need the following index sets: Let x^* be feasible for the MPEC (1). Then we define

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

Note that these index sets depend on the chosen point x^* . However, it will always be clear from the context, which point they refer to. Although there are more stationarity concepts known for MPECs, we will restrict ourselves to the most common ones.

Definition 2.3. Let x^* be feasible for the MPEC (1). Then x^* is said to be

(a) weakly stationary, if there are multipliers $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p, \gamma, \nu \in \mathbb{R}^q$ such that

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

and

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

- (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 either $\gamma_i > 0, \nu_i > 0$ or $\gamma_i \nu_i = 0$ for all $i \in I_{00}$;
- (d) strongly stationary, if it is weakly stationary and $\gamma_i \geq 0, \nu_i \geq 0$ for all $i \in I_{00}$.

These four stationary concepts are illustrated in Figure 1.

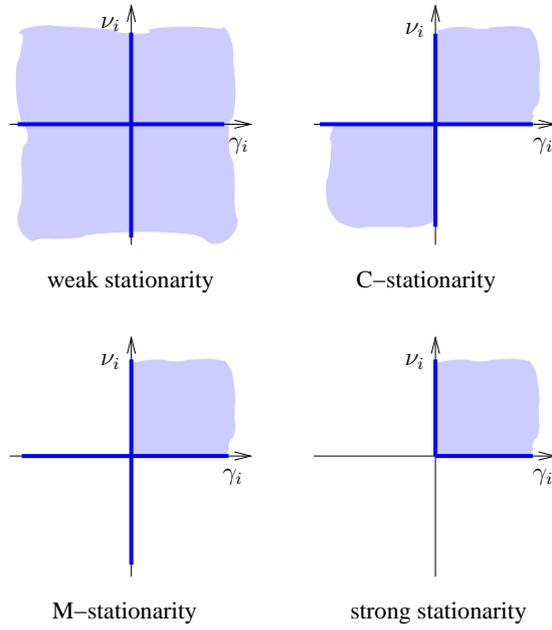


Figure 1: Geometric illustration of weak, C-, M-, and strong stationarity for an index i from the bi-active set I_{00}

Obviously, the following implications hold between these stationarity concepts:

$$\text{strong stationarity} \implies \text{M-stationarity} \implies \text{C-stationarity} \implies \text{weak stationarity}.$$

Differences between these stationary concepts arise only in the bi-active set I_{00} . If this set is empty, all four stationary concepts coincide. Weak and C-stationarity were introduced in [36], M-stationarity independently in [41, 30, 31, 40], and strong stationarity may already be found in [27]. We note, however, that strong stationarity can be shown to be equivalent to the standard KKT conditions of an MPEC, cf. [11].

2.3 MPEC-tailored Constraint Qualifications

To guarantee that a local minimizer of the MPEC (1) is stationary in one of the above senses, special MPEC-tailored constraint qualifications are used. By now, there is a whole zoo of constraint qualifications for MPECs but we are going to state only the ones used explicitly in the subsequent analysis. These are variants of some known constraint qualifications for standard nonlinear programs to the MPEC-setting, with the second one (constant positive-linear dependence constraint qualification) being widely unknown in the literature. It was first introduced in [33] and further investigated in [2] for standard nonlinear programs, and very recently modified for MPECs in [18].

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

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

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

are linearly independent;

(b) MPEC-constant positive-linear dependence constraint qualification (*MPEC-CPLD*) if, for any subsets $I_1 \subseteq I_g$, $I_2 \subseteq \{1, \dots, p\}$, $I_3 \subseteq I_{00} \cup I_{0+}$ and $I_4 \subseteq I_{00} \cup I_{+0}$ such that the gradients

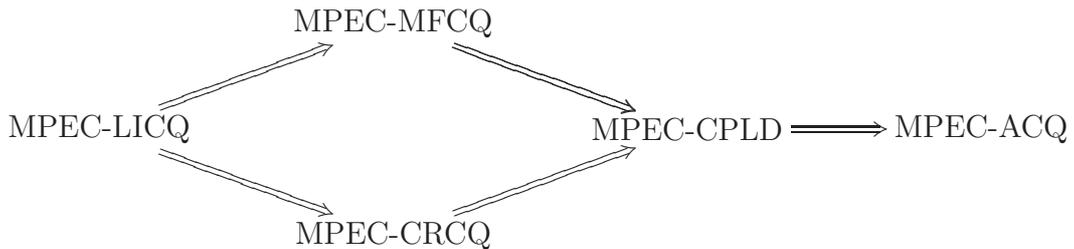
$$\{\nabla g_i(x^*) \mid i \in I_1\} \cup \{\nabla h_i(x^*) \mid i \in I_2\} \cup \{\nabla G_i(x^*) \mid i \in I_3\} \cup \{\nabla H_i(x^*) \mid i \in I_4\}$$

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

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

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

Note that, in the definition of MPEC-CPLD, we grouped those vectors together with extra curly brackets for which no sign restrictions apply. Apart from those defined above, there exist a number of other MPEC-tailored constraint qualifications like MPEC-MFCQ, MPEC-CRCQ and MPEC-ACQ as variants of the standard MFCQ (Mangasarian-Fromovitz constraint qualification), standard CRCQ (constant rank constraint qualification), and standard ACQ, see, e.g., [10, 12, 38, 18]. The relation of MPEC-LICQ and MPEC-CPLD to these other constraint qualifications is summarized in the following picture, cf. [18] and references therein for more details.



We see that MPEC-LICQ is the strongest constraint qualification among all, whereas MPEC-CPLD is much weaker and may be viewed as a common relaxation of both MPEC-MFCQ and MPEC-CRCQ. It is a well known fact that a local minimum x^* of the MPEC is a strongly stationary point if MPEC-LICQ holds at x^* , whereas neither MPEC-MFCQ nor MPEC-CRCQ implies strong stationarity. However, given a local minimum x^* which satisfies a relatively weak MPEC constraint qualification like, for example, MPEC-CPLD or MPEC-ACQ, it follows that x^* is at least an M-stationary point. Simple examples of MPECs show, however, that even a global minimum of the MPEC might not be strongly stationary, hence, in general, M-stationary is the best one can hope for unless relatively strong assumptions hold.

3 Regularization

3.1 Regularization by Kadrani et al.

Among the different regularization methods that exist for the solution of MPECs, there is only one very recent approach from Kadrani et al. [21] which converges to an M-stationary point. All the other regularization methods we are aware of and that were mentioned in the introduction converge, in general, only to C-stationary points. We therefore take a closer look to this particular regularization in this section.

Kadrani et al. [21] suggest to replace the complementarity conditions

$$G_i(x) \geq 0, \quad H_i(x) \geq 0, \quad G_i(x)H_i(x) = 0 \quad \forall i = 1, \dots, l$$

by

$$G_i(x) \geq -t, \quad H_i(x) \geq -t, \quad (G_i(x) - t)(H_i(x) - t) \leq 0 \quad \forall i = 1, \dots, l$$

for some parameter $t > 0$. The geometric interpretation of this particular regularization is given in Figure 2.

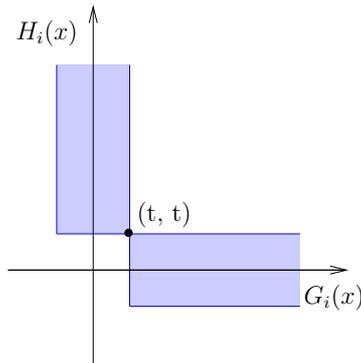


Figure 2: Illustration of the Kadrani et al. [21] regularization

The objective function and the other constraints are not modified. As shown in [21], the corresponding regularization method converges to an M-stationary point under the MPEC-LICQ assumption. Theoretically, this result is therefore much better than what is known for all the other regularization methods. Nevertheless, Figure 2 clearly shows a

potential drawback of this regularization: The feasible set of the regularized problem is almost disconnected, so one has to expect severe problems when solving the regularized problems by a standard optimization method. Moreover, it turns out that the feasible set of the original MPEC is not contained in the feasible set of the regularized problem, regardless of the choice of $t > 0$.

Our aim is therefore to construct a new regularization scheme which also converges to M-stationary points and which does not have these disadvantages. Moreover, we will show that convergence to M-stationary points is obtained under the much weaker MPEC-CPLD condition instead of the MPEC-LICQ condition.

3.2 New Regularization

Our relaxation is based on the function $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

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

This function has the following elementary 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

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

Proof. (a) First suppose that $a \geq 0, b \geq 0, ab = 0$. Then $a + b \geq 0$, and the definition of φ therefore gives $\varphi(a, b) = ab = 0$. Conversely, assume that $\varphi(a, b) = 0$. If $a + b \geq 0$, it then follows that $ab = 0$ and thus $a \geq 0, b = 0$ or $a = 0, b \geq 0$. On the other hand, if $a + b < 0$, we have $-\frac{1}{2}(a^2 + b^2) = 0$ which, in turn, implies $a = b = 0$, a contradiction to $a + b < 0$.

(b) This statement follows immediately from standard calculus rules.

(c) Using the continuity of φ together with the NCP-function property of part (a), it follows that φ has the same sign in all points of the positive orthant, as well as the same sign in all points in the other three orthants. Since $\varphi(1, 1) = 1 > 0$ and $\varphi(-1, -1) = -1 < 0$, the statement follows. \square

Based on this function, we define a continuously differentiable mapping $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^q$ componentwise by

$$\begin{aligned}\Phi_i(x; t) &:= \varphi(G_i(x) - t, H_i(x) - t) \\ &= \begin{cases} (G_i(x) - t)(H_i(x) - t), & \text{if } G_i(x) + H_i(x) \geq 2t, \\ -\frac{1}{2}((G_i(x) - t)^2 + (H_i(x) - t)^2), & \text{if } G_i(x) + H_i(x) < 2t, \end{cases}\end{aligned}$$

where $t \geq 0$ is an arbitrary parameter. With this function, we can formulate the relaxed or regularized problem NLP(t) for $t \geq 0$ as

$$\begin{aligned}\min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & G_i(x) \geq 0 \quad \forall i = 1, \dots, q, \\ & H_i(x) \geq 0 \quad \forall i = 1, \dots, q, \\ & \Phi_i(x; t) \leq 0 \quad \forall i = 1, \dots, q.\end{aligned}\tag{3}$$

Hence, in our approach, we replace the complementarity conditions

$$G_i(x) \geq 0, \quad H_i(x) \geq 0, \quad G_i(x)H_i(x) = 0$$

by

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

which, from a geometric point of view, gives a set of the form shown in Figure 3.

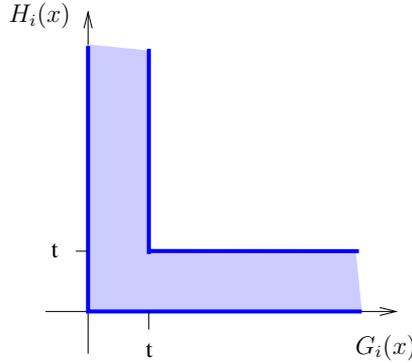


Figure 3: Geometric interpretation of the new regularization

Similar to the index sets used before, we define

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

for $t \geq 0$ and x feasible for $\text{NLP}(t)$. We also use a partition of the index set $I_\Phi(x; t)$ into the following three subsets:

$$\begin{aligned} I_\Phi^{00}(x; t) &:= \{i \in I_\Phi(x; t) \mid G_i(x) - t = 0, H_i(x) - t = 0\}, \\ I_\Phi^{0+}(x; t) &:= \{i \in I_\Phi(x; t) \mid G_i(x) - t = 0, H_i(x) - t > 0\}, \\ I_\Phi^{+0}(x; t) &:= \{i \in I_\Phi(x; t) \mid G_i(x) - t > 0, H_i(x) - t = 0\}. \end{aligned}$$

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

$$\Phi_i(x; t) = 0 \iff G_i(x) - t \geq 0, H_i(x) - t \geq 0, (G_i(x) - t)(H_i(x) - t) = 0.$$

In view of Lemma 3.1, the function Φ is continuously differentiable with its gradient given by

$$\nabla \Phi_i(x; t) = \begin{cases} (H_i(x) - t)\nabla G_i(x) + (G_i(x) - t)\nabla H_i(x), & \text{if } G_i(x) + H_i(x) \geq 2t, \\ -(G_i(x) - t)\nabla G_i(x) - (H_i(x) - t)\nabla H_i(x), & \text{if } G_i(x) + H_i(x) < 2t \end{cases} \quad (4)$$

for all $i = 1, \dots, q$.

The following result summarizes some simple properties of the regularized program $\text{NLP}(t)$.

Lemma 3.2. *Let X be the feasible set of the MPEC (1) and $X(t)$ the feasible set of $\text{NLP}(t)$ 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) $\bigcap_{t \geq 0} X(t) = X$.

Proof. (a) Taking into account the properties of φ and the definition of Φ , the condition

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

is equivalent to

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

for all $i = 1, \dots, q$. This proves $X(0) = X$.

(b) Let $0 \leq t_1 \leq t_2$ and x be an arbitrary element of $X(t_1)$. To prove $x \in X(t_2)$, we only have to show $\Phi_i(x; t_2) \leq 0$ for all $i = 1, \dots, q$. Let i be one of these indices. If $G_i(x) + H_i(x) < 2t_2$, we immediately obtain $\Phi_i(x; t_2) \leq 0$ since $\Phi_i(x; t)$ is always nonpositive in this case. Hence, the only case to consider is $G_i(x) + H_i(x) \geq 2t_2$. We want to prove $\Phi_i(x; t_2) = (G_i(x) - t_2)(H_i(x) - t_2) \leq 0$. Assume this is not true. Then either both values $G_i(x) - t_2$ and $H_i(x) - t_2$ would have to be positive or both negative. However, if both values were negative, we would have $G_i(x) + H_i(x) < 2t_2$, a contradiction. If both values were positive, $G_i(x) - t_1$ and $H_i(x) - t_1$ also were both positive and thus $\Phi_i(x; t_1) > 0$, a contradiction to $x \in X(t_1)$.

(c) According to part (a) and (b), we know $X = X(0) \subseteq X(t)$ for all $t \geq 0$ and thus $X \subseteq \bigcap_{t \geq 0} X(t)$. Now let $x \in \bigcap_{t \geq 0} X(t)$ be an arbitrary element. To prove $x \in X$, we only have to show $\Phi_i(x; 0) \leq 0$ for all $i = 1, \dots, q$. Assume that there is an i such that $\Phi_i(x; 0) > 0$. This implies (in fact is equivalent to) $G_i(x) > 0$ and $H_i(x) > 0$. Now choose an arbitrary $\bar{t} > 0$ with $\bar{t} < \min\{G_i(x), H_i(x)\}$. This definition of \bar{t} yields $G_i(x) + H_i(x) > 2\bar{t}$ and thus $\Phi_i(x; \bar{t}) = (G_i(x) - \bar{t})(H_i(x) - \bar{t}) > 0$. Consequently, $x \notin X(\bar{t})$ which is a contradiction to $x \in \bigcap_{t \geq 0} X(t)$. \square

The previous result shows, in particular, that the feasible set X of the original MPEC is always contained in the feasible set $X(t)$ of the regularized program $\text{NLP}(t)$ (in contrast to the approach by Kadrani et al. [21]), and that our relaxation exhibits the desired behaviour $\lim_{t \downarrow 0} X(t) = X$. Note also that, from a geometric point of view, our regularized problem has a much nicer feasible set than the one by Kadrani et al. [21] which, we recall, consists of almost disconnected pieces.

Remark 3.3. (a) The particular NCP-function φ used here can be replaced by other suitable NCP-functions. However, we stress that we cannot use an arbitrary NCP-function that, geometrically, gives the same feasible set for the regularized problem $\text{NLP}(t)$ since the stationary point properties that will be shown in the subsequent section heavily depend on the particular representation of this feasible set. Nevertheless, one particular alternative is the mapping

$$\varphi(a, b) := \theta(a) + \theta(b) - \theta(|a - b|)$$

with $\theta : \mathbb{R} \rightarrow \mathbb{R}$ being given by

$$\theta(\tau) := \begin{cases} -\frac{1}{2}\tau^2, & \text{if } \tau < 0, \\ \frac{1}{2}\tau^2, & \text{if } \tau \geq 0 \end{cases} \quad \text{or} \quad \theta(\tau) = \frac{1}{3}\tau^3.$$

This function is a particular member of a class of NCP-functions introduced in [29]. It is not difficult to see that our analysis goes through also for this mapping.

(b) The regularization used in this paper enlarges the feasible region coming from the complementarity constraints to the north-eastern direction. Alternatively, we may also use a regularization to the south-western direction by replacing the complementarity conditions

$$G_i(x) \geq 0, \quad H_i(x) \geq 0, \quad G_i(x)H_i(x) = 0$$

by

$$G_i(x) \geq -t, \quad H_i(x) \geq -t, \quad \Phi_i(x; 0) \leq 0$$

We may also combine the two relaxations and regularize with respect to the north-eastern and the south-western direction simultaneously. Figure 4 illustrates the three possible regularizations.

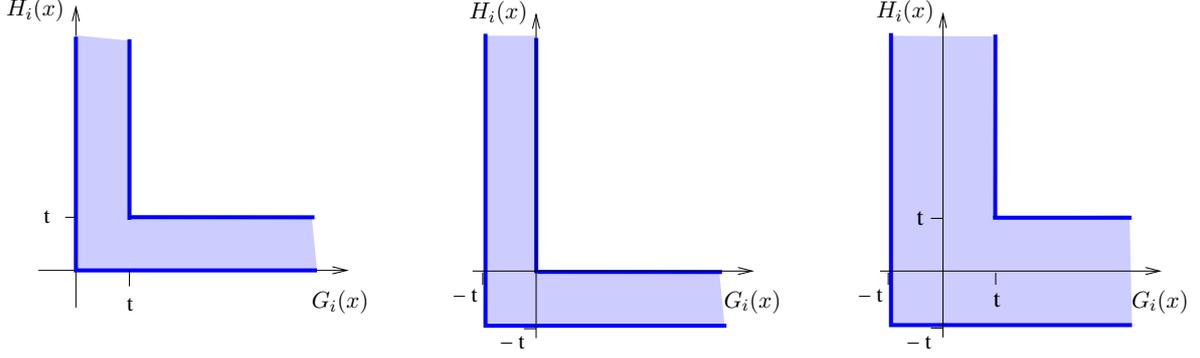


Figure 4: The three possible relaxations: The picture on the left-hand side shows the relaxation used in this paper, the picture in the middle shows an alternative relaxation, and the picture on the right combines the two relaxations.

4 Convergence Properties

4.1 Convergence to M- and Strongly Stationary Points

In this section, we are concerned with stationarity properties of limit points of our relaxation method. If we solve $NLP(t_k)$ for a sequence $\{t_k\} \downarrow 0$ and obtain KKT-points $(x^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$ of $NLP(t_k)$, where $x^k \rightarrow x^*$, what kind of MPEC-stationarity can we expect in x^* ? The next theorem gives an answer to this question.

Theorem 4.1. *Let $\{t_k\} \downarrow 0$ and $\{(x^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)\}$ be a sequence of KKT-points of $NLP(t_k)$ with $x^k \rightarrow x^*$. If MPEC-CPLD holds in x^* , then x^* is an M-stationary point of the MPEC (1).*

Proof. Obviously, x^* is feasible for the MPEC (1) and for all $k \in \mathbb{N}$ sufficiently large, we have

$$\begin{aligned} I_g(x^k) &\subseteq I_g, \\ I_G(x^k) \cup I_{\Phi}^{00}(x^k; t_k) \cup I_{\Phi}^{0+}(x^k; t_k) &\subseteq I_{00} \cup I_{0+}, \\ I_H(x^k) \cup I_{\Phi}^{00}(x^k; t_k) \cup I_{\Phi}^{+0}(x^k; t_k) &\subseteq I_{00} \cup I_{+0}. \end{aligned} \quad (5)$$

Since all $(x^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$ are KKT-points of $NLP(t_k)$, we have

$$\begin{aligned} 0 &= \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) \\ &\quad + \sum_{i=1}^q \delta_i^k \nabla \Phi_i(x^k; t_k) \end{aligned}$$

with

$$\begin{aligned} \lambda_i^k &= 0 \quad \forall i \notin I_g(x^k) \quad \text{and} \quad \lambda_i^k \geq 0 \quad \forall i \in I_g(x^k), \\ \gamma_i^k &= 0 \quad \forall i \notin I_G(x^k) \quad \text{and} \quad \gamma_i^k \geq 0 \quad \forall i \in I_G(x^k), \\ \nu_i^k &= 0 \quad \forall i \notin I_H(x^k) \quad \text{and} \quad \nu_i^k \geq 0 \quad \forall i \in I_H(x^k), \\ \delta_i^k &= 0 \quad \forall i \notin I_{\Phi}(x^k; t) \quad \text{and} \quad \delta_i^k \geq 0 \quad \forall i \in I_{\Phi}(x^k; t). \end{aligned}$$

Since the representation of $\nabla\Phi_i$ immediately gives $\nabla\Phi_i(x^k; t_k) = 0$ for all $i \in I_{\Phi}^{00}(x^k; t_k)$ and all $k \in \mathbb{N}$, we may also assume $\delta_i^k = 0$ for all $i \in I_{\Phi}^{00}(x^k; t_k)$ and all $k \in \mathbb{N}$. Thus, we can rewrite the equation above as

$$\begin{aligned} 0 &= \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) \\ &\quad + \sum_{i=1}^q \delta_i^{G,k} \nabla G_i(x^k) + \sum_{i=1}^q \delta_i^{H,k} \nabla H_i(x^k) \end{aligned}$$

where

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

Note that the multipliers $\delta^{G,k}$ and $\delta^{H,k}$ are nonnegative, too. According to [38, Lem. A.1], we may assume without loss of generality that the gradients corresponding to nonvanishing multipliers in this equation are linearly independent for all $k \in \mathbb{N}$ (note that this may change the multipliers, but a previously positive multiplier will stay at least nonnegative and a vanishing multiplier will remain zero).

Our next step is to prove that the sequence $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k})$ is bounded. If we assume the contrary, we can find a subsequence K such that

$$\frac{(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k})}{\|(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k})\|} \rightarrow_K (\lambda, \mu, \gamma, \nu, \delta^G, \delta^H) \neq 0.$$

Dividing by $\|(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k})\|$ and taking this limit in the equation above yields

$$\begin{aligned} 0 &= \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \mu_i \nabla h_i(x^*) - \sum_{i=1}^q \gamma_i \nabla G_i(x^*) - \sum_{i=1}^q \nu_i \nabla H_i(x^*) \\ &\quad + \sum_{i=1}^q \delta_i^G \nabla G_i(x^*) + \sum_{i=1}^q \delta_i^H \nabla H_i(x^*), \end{aligned}$$

i.e. the gradients

$$\begin{aligned} &\{\nabla g_i(x^*) \mid i \in \text{supp}(\lambda)\} \cup \{\nabla h_i(x^*) \mid i \in \text{supp}(\mu)\} \\ &\cup \{\nabla G_i(x^*) \mid i \in \text{supp}(\gamma) \cup \text{supp}(\delta^G)\} \cup \{\nabla H_i(x^*) \mid i \in \text{supp}(\nu) \cup \text{supp}(\delta^H)\} \end{aligned} \quad (6)$$

are positive-linearly dependent. MPEC-CPLD guarantees that they remain linearly dependent in a whole neighbourhood. This, however, is a contradiction to the linear independence of these gradients in x^k . Here, we used

$$\text{supp}(\lambda, \mu, \gamma, \nu, \delta^G, \delta^H) \subseteq \text{supp}(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k})$$

for all k sufficiently large and (5).

Consequently, our assumption was wrong and the sequence $\{(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k})\}$ is bounded. Therefore, we can assume without loss of generality that the whole sequence is convergent to some limit $(\lambda^*, \mu^*, \gamma^*, \nu^*, \delta^{G,*}, \delta^{H,*})$. Since $I_G(x^k) \cap I_{\Phi}^{0+}(x^k; t_k) = \emptyset$ and $I_H(x^k) \cap I_{\Phi}^{+0}(x^k; t_k) = \emptyset$ for all $k \in \mathbb{N}$, it is easy to see that the multipliers

$$\hat{\gamma}_i = \begin{cases} \gamma_i^* & \text{if } i \in \text{supp}(\gamma^*), \\ -\delta_i^{G,*} & \text{if } i \in \text{supp}(\delta^{G,*}), \\ 0 & \text{else,} \end{cases}$$

$$\hat{\nu}_i = \begin{cases} \nu_i^* & \text{if } i \in \text{supp}(\nu^*), \\ -\delta_i^{H,*} & \text{if } i \in \text{supp}(\delta^{H,*}), \\ 0 & \text{else} \end{cases}$$

are well defined, and we obtain

$$0 = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^p \mu_i^* \nabla h_i(x^*) - \sum_{i=1}^q \hat{\gamma}_i \nabla G_i(x^*) - \sum_{i=1}^q \hat{\nu}_i \nabla H_i(x^*).$$

Here, $\lambda^* \geq 0$ and

$$\begin{aligned} \text{supp}(\lambda^*) &\subseteq I_g(x^k) \subseteq I_g, \\ \text{supp}(\hat{\gamma}) &= \text{supp}(\gamma^*) \cup \text{supp}(\delta^{G,*}) \subseteq I_G(x^k) \cup I_{\Phi}^{0+}(x^k; t_k) \subseteq I_{00} \cup I_{0+}, \\ \text{supp}(\hat{\nu}) &= \text{supp}(\nu^*) \cup \text{supp}(\delta^{H,*}) \subseteq I_H(x^k) \cup I_{\Phi}^{+0}(x^k; t_k) \subseteq I_{00} \cup I_{+0} \end{aligned}$$

for all k sufficiently large. Consequently, we have $\hat{\gamma}_i = 0$ for all $i \in I_{+0}$ and $\hat{\nu}_i = 0$ for all $i \in I_{0+}$, i.e. $(x^*, \lambda^*, \mu^*, \hat{\gamma}, \hat{\nu})$ is at least a weakly stationary point of the MPEC (1). To prove M-stationarity, assume that there is an $i \in I_{00}$ with $\hat{\gamma}_i < 0$ and $\hat{\nu}_i \neq 0$ (the case $\hat{\gamma}_i \neq 0$ and $\hat{\nu}_i < 0$ can be treated in a symmetric way). The condition $\hat{\gamma}_i < 0$ implies $i \in \text{supp}(\delta^{G,*}) \subseteq I_{\Phi}^{0+}(x^k; t_k)$ for all k sufficiently large. Because of

$$I_{\Phi}^{0+}(x^k; t_k) \cap (I_H(x^k) \cup I_{\Phi}^{+0}(x^k; t_k)) = \emptyset$$

for all $k \in \mathbb{N}$, this yields $\hat{\nu}_i = 0$ in contradiction to our assumption. \square

Under stronger assumptions like the one defined below, we can even obtain strong stationarity of the limit point.

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

$$\begin{aligned} \frac{G_i(x^k)}{H_i(x^k)} &\leq 1 \quad \text{for all } i \in I_{\Phi}^{+0}(x^k; t_k) \cap I_{00}, \quad \text{and} \\ \frac{H_i(x^k)}{G_i(x^k)} &\leq 1 \quad \text{for all } i \in I_{\Phi}^{0+}(x^k; t_k) \cap I_{00} \end{aligned}$$

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

Related asymptotic weak nondegeneracy conditions were also used in [14, 26, 21]. A direct comparison of the different nondegeneracy conditions is not possible in general since they depend on the particular regularization. However, our feeling is that our definition is a relatively weak assumption that will often be satisfied in practice.

The next result shows that MPEC-CPLD together with the asymptotic weak nondegeneracy condition already guarantees that the limit point is strongly stationary.

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

Proof. Using Theorem 4.1, we know that x^* is at least M-stationary. To verify strong stationarity, we use the proof of Theorem 4.1 again. The only change is that, in the very end, we now additionally apply the asymptotic weak nondegeneracy condition:

Assume that $(x^*, \lambda^*, \mu^*, \hat{\gamma}, \hat{\nu})$ is not a strongly stationary point of the MPEC (1). Then we can find an $i \in I_{00}$, where $\hat{\gamma}_i < 0$ or $\hat{\nu}_i < 0$. Let us assume $\hat{\gamma}_i < 0$ without loss of generality, the second case can be treated in the same way. Then, by construction, $i \in \text{supp}(\delta^{G,*}) \subseteq I_{\Phi}^{0+}(x^k; t_k)$ and consequently $G_i(x^k) = t_k$, $H_i(x^k) > t_k$ for all k sufficiently large. This however implies $\frac{H_i(x^k)}{G_i(x^k)} > 1$ for all those k in contradiction to the assumption of asymptotic weak nondegeneracy. \square

We note that both Theorem 4.1 and Theorem 4.3 require significantly weaker assumptions than those which are used in the corresponding convergence results of existing regularization methods which typically need MPEC-LICQ (instead of MPEC-CPLD) as well as a second-order condition (not needed here). On the other hand, in this context, we also refer to the discussion in the following section.

4.2 Existence of Multipliers

There is an implicit assumption used in the previous two convergence results, namely that there exists a sequence of KKT points for the regularized problems $NLP(t_k)$. In particular, we therefore require the existence of Lagrange multipliers. The aim of this section is to show that these Lagrange multipliers indeed exist under suitable assumptions. The most natural idea would be to show that the regularized problems $NLP(t_k)$ (at least for $t_k > 0$ sufficiently small) inherit some constraint qualification from the original MPEC. However, this is not true in general. In fact, the following example shows that the MPEC itself might satisfy MPEC-LICQ at a feasible point x^* , whereas the corresponding regularized problem violates standard LICQ.

Example 4.4 Consider the standard two-dimensional MPEC

$$\min f(x) \quad \text{s.t.} \quad 0 \leq x_1 \perp x_2 \geq 0.$$

Obviously, MPEC-LICQ holds at $x^* = (0, 0)$. Now consider the sequences $t_k = \frac{1}{k}$ and $x^k = (\frac{1}{k}, \frac{1}{k})$ for $k \in \mathbb{N}$. It is easy to see that $t_k \downarrow 0$ and $x^k \rightarrow x^*$. Furthermore, x^k is feasible for $NLP(t_k)$ for all $k \in \mathbb{N}$. However, for all $k \in \mathbb{N}$ the only active gradient is

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

hence LICQ for the nonlinear program $NLP(t_k)$ does not hold in x^k for all $k \in \mathbb{N}$. In fact, not even ACQ, one of the weakest constraint qualifications known for standard nonlinear programs, holds in x^k . However, the even weaker GCQ is satisfied. \diamond

Inspired by this example, we are going to prove that, whenever MPEC-LICQ holds in a point x^* which is feasible for (1), there is a neighbourhood of x^* such that for all $t > 0$ sufficiently small and all x in this neighbourhood which are feasible for $NLP(t)$, standard GCQ holds. To do so, we need some auxiliary results. The first one is a lemma that facilitates the calculation of polar cones to linearized cones. It can be found, for example, in [5, Theorem 3.2.2].

Lemma 4.5. *Consider the cones*

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

and

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

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

In the proof of Theorem 4.7 we are going to work with some nonlinear programs that are closely related to $NLP(t)$ but have better properties concerning constraint qualifications. Let $t > 0$ and \hat{x} be feasible for $NLP(t)$. Let I be an arbitrary subset of $I_\Phi^{00}(\hat{x}; t)$ and $\bar{I} := I_\Phi^{00}(\hat{x}; t) \setminus I$ its complement. We define the nonlinear program $NLP(t, I)$ as

$$\begin{aligned} \min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & G_i(x) \geq 0, H_i(x) \geq 0, G_i(x) \leq t \quad \forall i \in I_\Phi^{0+}(\hat{x}; t) \cup I, \\ & G_i(x) \geq 0, H_i(x) \geq 0, H_i(x) \leq t \quad \forall i \in I_\Phi^{+0}(\hat{x}; t) \cup \bar{I}, \\ & G_i(x) \geq 0, H_i(x) \geq 0, \Phi_i(x; t) \leq 0 \quad \forall i \notin I_\Phi(\hat{x}; t) \end{aligned} \quad (7)$$

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 lemma sheds some light on the relation between the tangent cone of $NLP(t)$ and the tangent cones of $NLP(t, I)$.

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

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

Proof. To prove the first inclusion, let d be an arbitrary element of $\mathcal{T}_{X(t)}(\hat{x})$. This implies that there exists a sequence $x^k \rightarrow_{X(t)} \hat{x}$ and a sequence $\tau_k \downarrow 0$ such that

$$d = \lim_{k \rightarrow \infty} \frac{x^k - \hat{x}}{\tau_k}.$$

If we can find an $I \subseteq I_{\Phi}^{00}(\hat{x}; t)$ such that $x^k \in X(t, I)$ for infinitely many $k \in \mathbb{N}$, we have proven $d \in \bigcup_{I \subseteq I_{\Phi}^{00}(\hat{x}; t)} \mathcal{T}_{X(t, I)}(\hat{x})$. However, for every $i \in I_{\Phi}^{00}(\hat{x}; t)$ and all $k \in \mathbb{N}$, either $G_i(x^k) \leq t$ or $H_i(x^k) \leq t$. Hence, by choosing an appropriate subsequence $K \subseteq \mathbb{N}$ and defining I as the set of all indices i where $G_i(x^k) \leq t$ for all $k \in K$, we can construct such a set I .

To prove the second inclusion, choose an arbitrary $I \subseteq I_{\Phi}^{00}(\hat{x}, t)$ and an arbitrary $d \in \mathcal{T}_{X(t, I)}(\hat{x})$. This implies the existence of sequences $x^k \rightarrow_{X(t, I)} \hat{x}$ and $\tau_k \downarrow 0$ such that

$$d = \lim_{k \rightarrow \infty} \frac{x^k - \hat{x}}{\tau_k}.$$

Because of $X(t, I) \subseteq X(t)$, this yields $d \in \mathcal{T}_{X(t)}(\hat{x})$. \square

Now, we are in a position to state and prove the main result of this section.

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

Proof. Since MPEC-LICQ holds in x^* , the gradients

$$\{\nabla g_i(x) \mid i \in I_g\} \cup \{\nabla h_i(x) \mid i = 1, \dots, p\} \cup \{\nabla G_i(x) \mid i \in I_{00} \cup I_{0+}\} \cup \{\nabla H_i(x) \mid i \in I_{00} \cup I_{+0}\} \quad (8)$$

are linearly independent in x^* . Because of the continuity of the derivatives, they remain linearly independent in a whole neighbourhood. Hence, we can choose $\bar{t} > 0$ and $U(x^*)$ such that for all $t \in (0, \bar{t}]$ and all $x \in U(x^*)$ feasible for NLP(t) the gradients (8) are linearly independent in x , and the following inclusions hold, cf. (5):

$$\begin{aligned} I_g(x) &\subseteq I_g, \\ I_G(x) &\subseteq I_{00} \cup I_{0+}, \\ I_H(x) &\subseteq I_{00} \cup I_{+0}, \\ I_{\Phi}^{00}(x; t) \cup I_{\Phi}^{0+}(x; t) &\subseteq I_{00} \cup I_{0+}, \\ I_{\Phi}^{00}(x; t) \cup I_{\Phi}^{+0}(x; t) &\subseteq I_{00} \cup I_{+0}. \end{aligned}$$

Now choose an arbitrary $t \in (0, \bar{t}]$ and $\hat{x} \in U(x^*)$ such that \hat{x} is feasible for NLP(t). Then \hat{x} is also feasible for NLP(t, I) for all $I \subseteq I_{\Phi}^{00}(\hat{x}, t)$ and the active gradients are

$$\begin{aligned} &\{\nabla g_i(\hat{x}) \mid i \in I_g(\hat{x})\} \cup \{\nabla h_i(\hat{x}) \mid i = 1, \dots, p\} \cup \\ &\{\nabla G_i(\hat{x}) \mid i \in I_G(\hat{x}) \cup I_{\Phi}^{0+}(\hat{x}; t) \cup I\} \cup \{\nabla H_i(\hat{x}) \mid i \in I_H(\hat{x}) \cup I_{\Phi}^{+0}(\hat{x}; t) \cup \bar{I}\}. \end{aligned}$$

Thus, by construction of \bar{t} and $U(x^*)$, standard LICQ for NLP(t, I) holds in \hat{x} . Since LICQ implies ACQ, we have

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

for all $I \subseteq I_{\Phi}^{00}(\hat{x}; t)$. Together with Lemma 4.6, this yields

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

Passing to the polar cone, we obtain

$$\mathcal{T}_{X(t)}(\hat{x})^\circ = \bigcap_{I \subseteq I_{\Phi}^{00}(\hat{x}, t)} \mathcal{L}_{X(t, I)}(\hat{x})^\circ, \quad (9)$$

see [5, Theorem 3.1.9]. To prove that GCQ for NLP(t) holds in \hat{x} , we only have to prove the inclusion

$$\mathcal{T}_{X(t)}(\hat{x})^\circ \subseteq \mathcal{L}_{X(t)}(\hat{x})^\circ,$$

the opposite inclusion is always true. By definition, the linearized tangent cone of NLP(t, I) in \hat{x} is given by

$$\begin{aligned} \mathcal{L}_{X(t, I)}(\hat{x}) = \{d \in \mathbb{R}^n \mid & \nabla g_i(\hat{x})^T d \leq 0 \ \forall i \in I_g(\hat{x}), \\ & \nabla h_i(\hat{x})^T d = 0 \ \forall i = 1, \dots, p, \\ & \nabla G_i(\hat{x})^T d \geq 0 \ \forall i \in I_G(\hat{x}), \\ & \nabla H_i(\hat{x})^T d \geq 0 \ \forall i \in I_H(\hat{x}), \\ & \nabla G_i(\hat{x})^T d \leq 0 \ \forall i \in I_{\Phi}^{0+}(\hat{x}; t) \cup I, \\ & \nabla H_i(\hat{x})^T d \leq 0 \ \forall i \in I_{\Phi}^{+0}(\hat{x}; t) \cup \bar{I}\}. \end{aligned}$$

Therefore, Lemma 4.5 yields

$$\begin{aligned} \mathcal{L}_{X(t, I)}(\hat{x})^\circ = \{s \in \mathbb{R}^n \mid s = & \sum_{i \in I_g(\hat{x})} \lambda_i \nabla g_i(\hat{x}) + \sum_{i=1}^p \mu_i \nabla h_i(\hat{x}) - \sum_{i \in I_G(\hat{x})} \gamma_i \nabla G_i(\hat{x}) \\ & - \sum_{i \in I_H(\hat{x})} \nu_i \nabla H_i(\hat{x}) + \sum_{i \in I_{\Phi}^{0+}(\hat{x}; t) \cup I} \delta_i \nabla G_i(\hat{x}) + \sum_{i \in I_{\Phi}^{+0}(\hat{x}; t) \cup \bar{I}} \sigma_i \nabla H_i(\hat{x}) \\ & \lambda, \gamma, \nu, \delta, \sigma \geq 0\}. \end{aligned}$$

Now let s be an arbitrary element of $\mathcal{T}_{X(t)}(\hat{x})^\circ$. The representation of $\mathcal{T}_{X(t)}(\hat{x})^\circ$ in (9) then implies $s \in \mathcal{L}_{X(t, I)}(\hat{x})^\circ$ for all $I \subseteq I_{\Phi}^{00}(\hat{x}, t)$. If we fix such an index set I , we obtain

$$\begin{aligned} s = & \sum_{i \in I_g(\hat{x})} \lambda_i \nabla g_i(\hat{x}) + \sum_{i=1}^p \mu_i \nabla h_i(\hat{x}) - \sum_{i \in I_G(\hat{x})} \gamma_i \nabla G_i(\hat{x}) - \sum_{i \in I_H(\hat{x})} \nu_i \nabla H_i(\hat{x}) \\ & + \sum_{i \in I_{\Phi}^{0+}(\hat{x}; t) \cup I} \delta_i \nabla G_i(\hat{x}) + \sum_{i \in I_{\Phi}^{+0}(\hat{x}; t) \cup \bar{I}} \sigma_i \nabla H_i(\hat{x}) \end{aligned}$$

with some multipliers $\mu \in \mathbb{R}^p$ and $\lambda, \gamma, \nu, \delta, \sigma \geq 0$. On the other hand, $s \in \mathcal{L}_{X(t, \bar{I})}(\hat{x})^\circ$ also holds, thus we also have

$$\begin{aligned} s = & \sum_{i \in I_g(\hat{x})} \bar{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^p \bar{\mu}_i \nabla h_i(\hat{x}) - \sum_{i \in I_G(\hat{x})} \bar{\gamma}_i \nabla G_i(\hat{x}) - \sum_{i \in I_H(\hat{x})} \bar{\nu}_i \nabla H_i(\hat{x}) \\ & + \sum_{i \in I_{\Phi}^{0+}(\hat{x}; t) \cup \bar{I}} \bar{\delta}_i \nabla G_i(\hat{x}) + \sum_{i \in I_{\Phi}^{+0}(\hat{x}; t) \cup I} \bar{\sigma}_i \nabla H_i(\hat{x}) \end{aligned}$$

with some multipliers $\bar{\mu} \in \mathbb{R}^p$ and $\bar{\lambda}, \bar{\gamma}, \bar{\nu}, \bar{\delta}, \bar{\sigma} \geq 0$. However, by construction of \bar{t} and $U(x^*)$, the gradients

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

are linearly independent, hence the multipliers have to be the same. In particular, this implies

$$\delta_i = 0 \quad \forall i \in I \quad \text{and} \quad \sigma_i = 0 \quad \forall i \in \bar{I}.$$

Since an elementary calculation shows that

$$\begin{aligned} \mathcal{L}_{X(t)}(\hat{x}) = \{d \in \mathbb{R}^n \mid & \nabla g_i(\hat{x})^T d \leq 0 \quad \forall i \in I_g(\hat{x}), \\ & \nabla h_i(\hat{x})^T d = 0 \quad \forall i = 1, \dots, p, \\ & \nabla G_i(\hat{x})^T d \geq 0 \quad \forall i \in I_G(\hat{x}), \\ & \nabla H_i(\hat{x})^T d \geq 0 \quad \forall i \in I_H(\hat{x}), \\ & \nabla G_i(\hat{x})^T d \leq 0 \quad \forall i \in I_{\Phi}^{0+}(\hat{x}; t), \\ & \nabla H_i(\hat{x})^T d \leq 0 \quad \forall i \in I_{\Phi}^{+0}(\hat{x}; t)\}, \end{aligned}$$

application of Lemma 4.5 yields $s \in \mathcal{L}_{X(t)}(\hat{x})^\circ$. Note that the representation of $\mathcal{L}_{X(t)}(\hat{x})$ above exploits the structure of $\nabla \Phi(x; t)$ as given in (4). Since $s \in \mathcal{T}_{X(t)}(\hat{x})^\circ$ was chosen arbitrarily, we have proven

$$\mathcal{T}_{X(t)}(\hat{x})^\circ \subseteq \mathcal{L}_{X(t)}(\hat{x})^\circ,$$

i.e. GCQ for $\text{NLP}(t)$ holds in \hat{x} . □

The existence of Lagrange multipliers in local minima of $\text{NLP}(t)$ is a direct consequence of Theorem 4.7.

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

Note that Theorem 4.8 implies the existence of multipliers at a local minimum of the regularized problem $\text{NLP}(t)$ since Theorem 4.7 shows that the standard GCQ holds for the regularized problem under the MPEC-LICQ assumption. Moreover, recall that Example 4.4 indicates that we cannot expect a stronger result, even the ACQ may not hold for $\text{NLP}(t)$ under MPEC-LICQ. In a sense, this is similar to some results that are known for the MPEC itself, cf. [11]. However, the following result shows that there is a significant difference between MPECs themselves and our regularized problem $\text{NLP}(t)$. In fact, it is known that the MPEC does not satisfy standard LICQ (or even the weaker MFCQ) at an *arbitrary feasible point*. On the other hand, the next result shows that standard LICQ holds for $\text{NLP}(t)$ if MPEC-LICQ is satisfied and, in addition, the index set $I_{\Phi}^{00}(x; t)$ is empty. The latter assumption excludes only points where $(G_i(x), H_i(x)) = (t, t)$ for at least one i . In fact, this result also shows that MPEC-CPLD for the original MPEC implies standard CPLD for the regularized subproblems $\text{NLP}(t)$.

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

Proof. We first verify the assertion for MPEC-LICQ. Since MPEC-LICQ holds in x^* , the gradients

$$\{\nabla g_i(x) \mid i \in I_g\} \cup \{\nabla G_i(x) \mid i \in I_{00} \cup I_{0+}\} \cup \{\nabla H_i(x) \mid i \in I_{00} \cup I_{+0}\} \cup \{\nabla h_i(x) \mid i = 1, \dots, p\} \quad (10)$$

are linearly independent in $x = x^*$. Because of the continuity of the derivatives, they remain linearly independent in a whole neighbourhood. Thus, we can choose $\bar{t} > 0$ and $U(x^*)$ such that for all $t \in (0, \bar{t}]$ and all $x \in U(x^*)$ feasible for NLP(t), the gradients (10) are linearly independent in x , and the following inclusions hold:

$$\begin{aligned} I_g(x) &\subseteq I_g, \\ I_G(x) &\subseteq I_{00} \cup I_{0+}, \\ I_H(x) &\subseteq I_{00} \cup I_{+0}, \\ I_{\Phi}^{00}(x; t) \cup I_{\Phi}^{0+}(x; t) &\subseteq I_{00} \cup I_{0+}, \\ I_{\Phi}^{00}(x; t) \cup I_{\Phi}^{+0}(x; t) &\subseteq I_{00} \cup I_{+0}. \end{aligned} \quad (11)$$

Now choose an arbitrary $t \in (0, \bar{t}]$. When $x \in U(x^*)$ is feasible for NLP(t) with $I_{\Phi}^{00}(x; t) = \emptyset$, the active gradients in x are

$$\begin{aligned} &\{\nabla g_i(x) \mid i \in I_g(x)\} \cup \{\nabla h_i(x) \mid i = 1, \dots, p\} \cup \\ &\cup \{-\nabla G_i(x) \mid i \in I_G(x)\} \cup \{(H_i(x) - t)\nabla G_i(x) \mid i \in I_{\Phi}^{0+}(x; t)\} \\ &\cup \{-\nabla H_i(x) \mid i \in I_H(x)\} \cup \{(G_i(x) - t)\nabla H_i(x) \mid i \in I_{\Phi}^{+0}(x; t)\}, \end{aligned}$$

where $G_i(x) - t > 0$ for $i \in I_{\Phi}^{+0}(x; t)$ and $H_i(x) - t > 0$ for $i \in I_{\Phi}^{0+}(x; t)$. Hence, the choice of \bar{t} and $U(x^*)$ implies that these gradients are linearly independent, too. Therefore, standard LICQ holds in x .

It remains to prove the assertion under MPEC-CPLD. To this end, assume that there were sequences $t_k \downarrow 0$ and $x^k \rightarrow x^*$ 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 $k \in \mathbb{N}$. Violation of CPLD means that there are subsets $I_1^k \subseteq I_g(x^k)$, $I_2^k \subseteq \{1, \dots, p\}$, $I_3^k \subseteq I_G(x^k)$, $I_4^k \subseteq I_H(x^k)$, $I_5^k \subseteq I_{\Phi}^{0+}(x^k; t_k)$, $I_6^k \subseteq I_{\Phi}^{+0}(x^k; t_k)$ such that the gradients

$$\begin{aligned} &\left\{ \{\nabla g_i(x^k) \mid i \in I_1^k\} \cup \{-\nabla G_i(x^k) \mid i \in I_3^k\} \cup \{-\nabla H_i(x^k) \mid i \in I_4^k\} \right. \\ &\left. \cup \{(H_i(x^k) - t_k)\nabla G_i(x^k) \mid i \in I_5^k\} \cup \{(G_i(x^k) - t_k)\nabla H_i(x^k) \mid i \in I_6^k\} \right\} \cup \{\nabla h_i(x^k) \mid i \in I_2^k\} \end{aligned}$$

are positive-linearly dependent in x^k , but linearly independent in points arbitrary close to x^k . We may assume without loss of generality $I_i^k = I_i$ for all $i = 1, \dots, 6$. For all k sufficiently large, we know $I_g(x^k) \subseteq I_g$ and thus $I_1 \subseteq I_g$. Analogously, we obtain $I_3 \cup I_5 \subseteq I_{00} \cup I_{0+}$ and $I_4 \cup I_6 \subseteq I_{00} \cup I_{+0}$. Positive-linear dependence in x^k as we stated it above also implies positive-linear dependence of the gradients

$$\{\nabla g_i(x^k) \mid i \in I_1\} \cup \left\{ \{\nabla h_i(x^k) \mid i \in I_2\} \cup \{\nabla G_i(x^k) \mid i \in I_3 \cup I_5\} \cup \{\nabla H_i(x^k) \mid i \in I_4 \cup I_6\} \right\},$$

and because of the violation of CPLD, we can find a sequence $y^k \rightarrow x^*$ 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 $x^k \rightarrow x^*$. On the other hand, if they were positive-linearly dependent in x^* , MPEC-CPLD would imply that they remain linearly dependent in a neighbourhood, which contradicts the existence of $y^k \rightarrow x^*$. This concludes the proof. \square

We close this section by noting that the previous result also holds for some other constraint qualifications. In fact, it is possible to show that MPEC-MFCQ for the original MPEC implies standard MFCQ for the regularized problem. Furthermore, MPEC-CRCQ for the MPEC itself also implies standard CRCQ for the regularized problem $\text{NLP}(t)$. The corresponding proofs are very similar to the one of Theorem 4.9, so we skip the details (also because MPEC-MFCQ and MPEC-CRCQ are neither defined in this paper nor used somewhere else).

5 Numerical Results

Algorithm 5.1 Relaxation algorithm (x_0, t_0, σ)

Require: a starting vector x^0 , an initial relaxation parameter t_0 , and a parameter $\sigma \in (0, 1)$

Set $k := 0$.

while stopping criterion is not satisfied **do**

 Find a solution x^{k+1} of $\text{NLP}(t_k)$. To solve $\text{NLP}(t_k)$, use x^k as starting vector.

 Let $t_{k+1} \leftarrow t_k \cdot \sigma$ and $k \leftarrow k + 1$.

end while

Return: the final iterate $x_{opt} := x^k$, the corresponding function value $f(x_{opt})$ and the relaxation parameter t_{k-1} used in the last iteration

The relaxation method proposed in Section 3 leads to Algorithm 5.1. We implemented this algorithm in MATLAB 7.8.0. As stopping criterion, we used the following condition: If either $t_k < 10^{-8}$ or the maximum constraint violation in x^k (for $k \geq 1$) is less than 10^{-6} , the algorithm terminates. Here, the maximum constraint violation in a point x is defined as

$$\text{maxVio}(x) = \max\{\max\{0, g(x)\}, |h(x)|, |\min\{G(x), H(x)\}|\}.$$

The reason for the first condition is that the relaxed problem $\text{NLP}(t_k)$ is very similar to the original MPEC for extremely small relaxation parameters t_k , and thus standard NLP solvers might have trouble finding a solution. The second condition is motivated by the fact that a local minimum of $\text{NLP}(t_k)$, which is feasible for the original MPEC, also is a local minimum of (1). Hence, we can stop immediately in this case.

In order to illustrate the positive influence of our regularization method, we first consider a two-dimensional toy problem. Using sophisticated NLP-solvers, this problem can actually be solved easily, but since this does not show the potential improvements

that can be obtained by relaxation for larger or more complicated problems, we take the MATLAB routine `fmincon` to solve $NLP(t_k)$, which is a reasonable and widely available solver, but certainly not comparable to some of the more recent software.

The particular toy problem that we consider here is

$$\min(x_1 - 1)^2 + (x_2 - 1)^2 \quad \text{s.t.} \quad 0 \leq x_1 \perp x_2 \leq 0, \quad (12)$$

which has two strongly stationary points in $(1, 0)^T$ and $(0, 1)^T$ and one C-stationary point in $(0, 0)^T$. We choose a grid of starting points in the rectangle $[-1, 2] \times [-1, 2]$ and apply several relaxation algorithms (with `fmincon` as NLP solver and the parameters $(t_0, \sigma) = (0.5, 0.1)$) to all these starting points. If $\|x_{opt} - (1, 0)^T\|_2 < 10^{-5}$, we mark the corresponding starting point with a *, if $\|x_{opt} - (0, 1)^T\|_2 < 10^{-5}$, we use a +, and if $\|x_{opt} - (0, 0)^T\|_2 < 10^{-5}$, we mark the starting point with a o. Starting points with no entry indicate that the corresponding method fails to converge to anything reasonable from this particular starting point.

For comparison, we implemented the relaxation scheme of Scholtes [37], where

$$\Phi_i^S(x; t) = G_i(x)H_i(x) - t,$$

the relaxation scheme of Steffensen and Ulbrich [38], where

$$\Phi_i^{VU}(x; t) = \begin{cases} G_i(x) + H_i(x) - |G_i(x) - H_i(x)|, & \text{if } |G_i(x) - H_i(x)| \geq t, \\ G_i(x) + H_i(x) - t\theta\left(\frac{G_i(x) - H_i(x)}{t}\right), & \text{if } |G_i(x) - H_i(x)| < t \end{cases}$$

with

$$\theta(x) = \frac{2}{\pi} \sin\left(\frac{\pi}{2}x + 3\frac{\pi}{2}\right),$$

as well as the relaxation scheme by Kadrani et al., see [21] and the corresponding discussion in Section 3.1. We also applied `fmincon` directly to the problem. The results are displayed in Figure 5.

The results are as they had to be expected: Both, the direct application of `fmincon` to the MPEC and the local relaxation method by Steffensen and Ulbrich are attracted by the C-stationary point $(0, 0)^T$ for some starting points with $x_1 = x_2$ and fail to converge or do not reach the required accuracy for some other starting points. In contrast to that, our relaxation and the one proposed by Kadrani et al. converge to one of the two strongly stationary solutions for all starting points. The results for Scholtes' relaxation are not displayed in the figure as the optimal solutions found by this method were close to $(1, 0)^T$ or $(0, 1)^T$ for all starting points, but only up to an accuracy of 10^{-3} .

We next want to illustrate the behaviour of our method when applied to a variety of more serious test problems. To this end, we choose 126 of the 180 problems from the MacMPEC collection [22]; 41 problems were discarded because of their size, 2 because an error occurred during the evaluation of the objective function or constraints by AMPL, and 11 problems were not of the form considered in this paper. As some of the problems considered have over 100 variables or constraints, we now use the TOMLAB 7.4.0 solver `snopt` instead of the MATLAB routine `fmincon`. Communication between AMPL and MATLAB is achieved using the mex function `amplfunc` [28], see also [15, 9] for more information on `amplfunc` and on complementarity constraints in AMPL. We use the

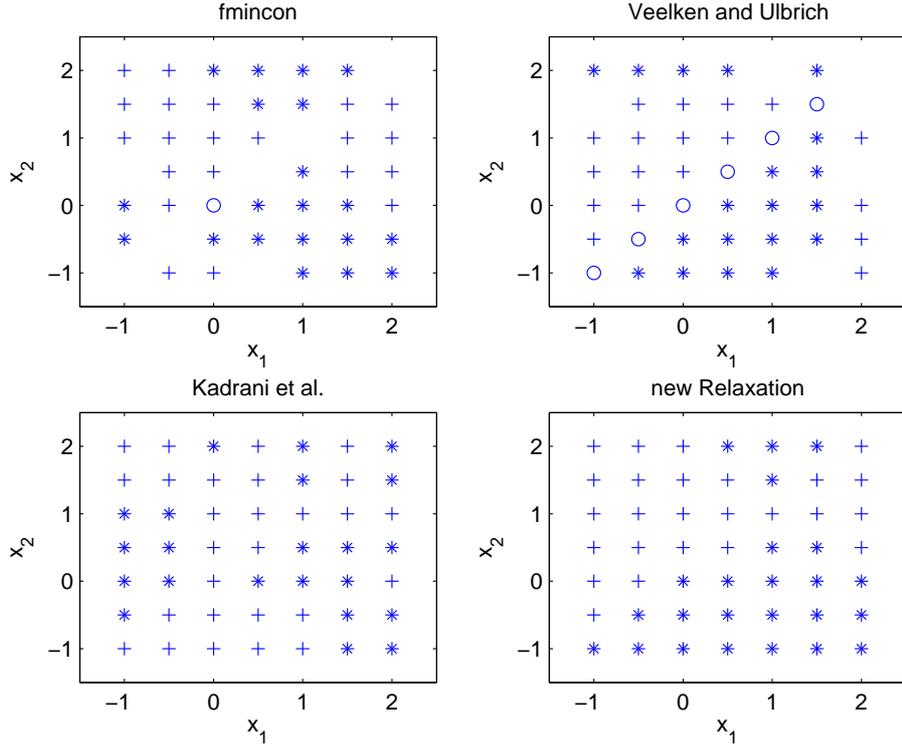


Figure 5: Results corresponding to Example (12)

starting vector suggested by AMPL and the parameters $(t_0, \sigma) = (1, 0.1)$ for all test examples. To determine these parameters, we did some testing using the parameter combinations $(t_0, \sigma) \in \{1, 10\} \times \{0.01, 0.1\}$, and $(t_0, \sigma) = (1, 0.1)$ was the pair that did best, although we could not observe significant differences for the different pairs (t_0, σ) .

For every test problem the data given in Table 2 has to be interpreted according to Table 1. Note that we changed the signs of some objective functions such that now all test problems are minimization problems. Thus the results can be compared more easily.

Problem	name of the test problem
(n, m, p, q)	number of variables, inequality, equality, and complementarity constraints
$f_{MacMPEC}$	best known objective function value according to MacMPEC site
f_{opt}	optimal objective function value found by our algorithm
$\max \text{Vio}(x_{opt})$	maximal constraint violation in our solution (defined as above)
t_{fin}	relaxation parameter used in the last iteration

Table 1: Explanation of results for MacMPEC collection

Obviously, our algorithm is able to find an optimal solution for almost all test problems with an objective function value which is equal to the best function value known so far for the large majority of all test problems. Sometimes the optimal function values are difficult to compare since the feasibility measure plays a central role, and we do not know which condition was used for those methods which obtained the best function value known so

far. We also note that there are a few problems where the algorithm has trouble achieving feasibility. However, for some of these problems, this had to be expected: `design-cent-3` is known to be infeasible and `ex9.2.2`, `ralph1`, and `scholtes4` are known to have B-stationary solutions that are not strongly stationary, cf. [23].

We also note that, for a number of test problems, the final value of t_k is equal to its initial value $t_0 = 1$. This means that we found the MPEC-solution by solving just a single regularized problem. This indicates some kind of finite termination and, very likely, corresponds to the case where, geometrically speaking, the minimum of the original MPEC is somewhere in the south-western direction from the origin since, in this direction, we do not relax the feasible set of the original MPEC.

Finally, we also made some preliminary numerical experiments with the alternative formulations suggested in Remark 3.3. Basically, using the other NCP-function with the quadratic θ -term does not seem to make a big difference. Using, however, the cubic θ -terms seems less favourable probably due to the higher degree of nonlinearity, although for a few test examples such as `bilin` for example, the solution found this way was better than for the other methods. The other two relaxations suggested in Remark 3.3 (b) seem to work well, although the overall behaviour is slightly worse than for the relaxation in the north-eastern direction. But this might be problem-dependent.

Problem	(n, m, p, q)	$f_{MacMPEC}$	f_{opt}	$\max \text{Vio}(x_{opt})$	t_{fin}
bar-truss-3	(35, 7, 28, 6)	1.01666e+04	1.01666e+04	1.00e-07	1e-07
bard1	(5, 2, 1, 3)	1.70000e+01	1.70000e+01	1.00e-07	1e-07
bard2	(12, 18, 5, 3)	-6.59800e+03	-6.59800e+03	7.02e-14	1e+00
bard3	(6, 6, 3, 1)	-1.26787e+01	-1.26787e+01	1.78e-15	1e+00
bard1m	(6, 3, 1, 3)	1.70000e+01	1.70000e+01	1.00e-07	1e-07
bard2m	(12, 18, 5, 3)	-6.59800e+03	-6.59800e+03	6.93e-14	1e+00
bard3m	(6, 4, 1, 3)	-1.26787e+01	-1.26787e+01	1.00e-06	1e-08
bilevel1	(10, 5, 2, 6)	-6.00000e+01	5.00000e+00	1.00e-06	1e-08
bilevel2	(16, 13, 4, 8)	-6.60000e+03	-6.60000e+03	9.98e-07	1e-07
bilevel3	(11, 4, 6, 3)	-1.26787e+01	-1.26787e+01	1.00e-07	1e-07
bilevel2m	(16, 13, 4, 8)	-6.60000e+03	-6.60000e+03	9.98e-07	1e-07
bilin	(8, 3, 0, 6)	-1.84000e+01	-1.60000e+01	1.00e-07	1e-07
dempe	(3, 0, 1, 1)	3.12500e+01	2.82501e+01	1.00e-06	1e-08
design-cent-1	(12, 3, 6, 3)	-1.86065e+00	-1.86065e+00	1.00e-06	1e-08
design-cent-2	(13, 7, 6, 3)	-3.48382e+00	-3.25025e+17	2.48e+02	1e-08
design-cent-21	(13, 7, 6, 3)	-3.48382e+00	-3.48382e+00	1.00e-06	1e-08
design-cent-3	(15, 3, 6, 3)	infeasible	-2.46727e+15	1.00e+00	1e-08
design-cent-31	(15, 3, 6, 3)	-3.72337e+00	-3.72337e+00	2.73e-07	1e-07
design-cent-4	(22, 7, 10, 8)	-3.07920e+00	-3.07920e+00	1.00e-07	1e-07
desilva	(6, 4, 2, 2)	-1.00000e+00	-1.00000e+00	0.00e+00	1e+00
df1	(2, 4, 0, 1)	0.00000e+00	1.23260e-32	0.00e+00	1e+00
ex9.1.1	(13, 1, 7, 5)	-1.30000e+01	-1.30000e+01	8.44e-15	1e-05
ex9.1.2	(8, 5, 5, 2)	-6.25000e+00	-6.25000e+00	1.00e-06	1e-08
ex9.1.3	(23, 8, 15, 6)	-2.92000e+01	-6.00000e+00	2.00e-06	1e-08
ex9.1.4	(8, 4, 5, 2)	-3.70000e+01	-3.70000e+01	6.37e-09	1e+00
ex9.1.5	(13, 3, 7, 5)	-1.00000e+00	-9.99999e-01	1.00e-06	1e-08
ex9.1.6	(14, 2, 7, 6)	-1.50000e+01	-4.90000e+01	5.00e-06	1e-08
ex9.1.7	(17, 5, 9, 6)	-6.00000e+00	-6.00003e+00	1.00e-06	1e-06
ex9.1.8	(11, 6, 5, 3)	-3.25000e+00	-3.25000e+00	1.11e-16	1e+00
ex9.1.9	(12, 2, 6, 5)	3.11111e+00	3.11111e+00	1.00e-06	1e-08
ex9.1.10	(11, 6, 5, 3)	-3.25000e+00	-3.25000e+00	1.11e-16	1e+00
ex9.2.1	(10, 2, 5, 4)	2.50000e+01	1.70000e+01	1.00e-06	1e-06
ex9.2.2	(9, 5, 4, 3)	1.00000e+02	9.99900e+01	1.00e-03	1e-08
ex9.2.3	(14, 9, 8, 4)	-5.50000e+01	5.00000e+00	1.42e-14	1e-01
ex9.2.4	(8, 3, 5, 2)	5.00000e-01	4.99999e-01	1.00e-06	1e-06
ex9.2.5	(8, 2, 4, 3)	9.00000e+00	5.00000e+00	1.00e-07	1e-07
ex9.2.6	(16, 4, 6, 6)	-1.00000e+00	-1.00000e+00	1.00e-06	1e-07
ex9.2.7	(10, 2, 5, 4)	2.50000e+01	1.70000e+01	1.00e-06	1e-06
ex9.2.8	(6, 3, 3, 2)	1.50000e+00	1.50000e+00	1.00e-06	1e-08
ex9.2.9	(9, 4, 5, 3)	2.00000e+00	2.00000e+00	1.00e-06	1e-06
flp2	(4, 4, 0, 2)	0.00000e+00	1.50998e-11	0.00e+00	1e-06
flp4-1	(80, 30, 0, 30)	0.00000e+00	5.52513e-15	0.00e+00	1e+00
flp4-2	(110, 50, 0, 60)	0.00000e+00	1.13150e-13	0.00e+00	1e+00
flp4-3	(140, 100, 0, 70)	0.00000e+00	9.79526e-09	0.00e+00	1e+00

Problem	(n, m, p, q)	$f_{MacMPEC}$	f_{opt}	$\max \text{Vio}(x_{opt})$	t_{fin}
flp4-4	(200, 150, 0, 100)	0.00000e+00	1.12391e-12	0.00e+00	1e+00
gauvin	(3, 2, 0, 2)	2.00000e+01	2.00000e+01	1.00e-06	1e-06
gnash10	(13, 2, 4, 8)	-2.30823e+02	-2.30823e+02	1.00e-06	1e-06
gnash11	(13, 2, 4, 8)	-1.29912e+02	-1.29912e+02	1.00e-06	1e-06
gnash12	(13, 2, 4, 8)	-3.69331e+01	-3.69331e+01	1.00e-06	1e-06
gnash13	(13, 2, 4, 8)	-7.06178e+00	-7.06179e+00	1.00e-06	1e-08
gnash14	(13, 2, 4, 8)	-1.79046e-01	-1.79053e-01	1.00e-05	1e-08
gnash15	(13, 2, 4, 8)	-3.54699e+02	-3.54699e+02	1.00e-06	1e-08
gnash16	(13, 2, 4, 8)	-2.41442e+02	-2.41442e+02	1.00e-06	1e-08
gnash17	(13, 2, 4, 8)	-9.07491e+01	-9.07491e+01	1.00e-06	1e-08
gnash18	(13, 2, 4, 8)	-2.56982e+01	-2.56982e+01	1.00e-06	1e-08
gnash19	(13, 2, 4, 8)	-6.11671e+00	-6.11671e+00	1.00e-06	1e-08
hs044-i	(20, 12, 4, 10)	1.56178e+01	2.80046e+01	1.00e+00	1e-08
incid-set1-8	(118, 77, 50, 49)	3.81600e-17	4.68375e-16	4.69e-08	1e+00
incid-set1c-8	(117, 82, 49, 49)	3.81600e-17	3.81639e-17	8.14e-07	1e+00
incid-set2-8	(117, 75, 49, 49)	4.51800e-03	5.04269e-03	2.23e-09	1e+00
incid-set2c-8	(117, 82, 49, 49)	5.47100e-03	5.63006e-03	5.12e-07	1e+00
jr1	(2, 0, 0, 1)	5.00000e-01	5.00000e-01	0.00e+00	1e+00
jr2	(2, 0, 0, 1)	5.00000e-01	4.99999e-01	1.00e-06	1e-08
kth1	(2, 0, 0, 1)	0.00000e+00	0.00000e+00	0.00e+00	1e+00
kth2	(2, 0, 0, 1)	0.00000e+00	0.00000e+00	0.00e+00	1e+00
kth3	(2, 0, 0, 1)	5.00000e-01	5.00000e-01	0.00e+00	1e-06
liswet1-050	(152, 51, 52, 50)	1.39900e-02	1.39943e-02	7.16e-14	1e+00
monteiro	(136, 72, 62, 62)	-3.75300e+01	-3.75300e+01	3.14e-06	1e-08
monteiroB	(136, 72, 62, 62)	-8.27859e+02	-8.27860e+02	3.25e-06	1e-08
nash1a	(6, 4, 2, 2)	7.88861e-30	2.54196e-16	1.00e-05	1e-08
nash1b	(6, 4, 2, 2)	7.88861e-30	1.91080e-22	1.00e-05	1e-08
nash1c	(6, 4, 2, 2)	7.88861e-30	4.66144e-24	1.00e-05	1e-08
nash1d	(6, 4, 2, 2)	7.88861e-30	9.11312e-15	1.00e-05	1e-08
nash1e	(6, 4, 2, 2)	7.88861e-30	6.34667e-22	1.00e-05	1e-08
outrata31	(5, 2, 0, 4)	3.20770e+00	3.20770e+00	1.00e-06	1e-08
outrata32	(5, 2, 0, 4)	3.44940e+00	3.44940e+00	1.00e-06	1e-06
outrata33	(5, 2, 0, 4)	4.60425e+00	4.60425e+00	1.00e-06	1e-08
outrata34	(5, 2, 0, 4)	6.59268e+00	6.59268e+00	1.00e-06	1e-08
pack-comp1-8	(107, 49, 49, 49)	6.00000e-01	6.00000e-01	1.01e-05	1e-08
pack-comp1c-8	(107, 56, 49, 49)	6.00000e-01	6.00000e-01	5.92e-07	1e-07
pack-comp1p-8	(107, 34, 49, 49)	6.00000e-01	6.00000e-01	1.07e-05	1e-08
pack-comp2-8	(107, 49, 49, 49)	6.73117e-01	6.73117e-01	3.09e-09	1e+00
pack-comp2c-8	(107, 56, 49, 49)	6.73458e-01	6.73458e-01	4.18e-10	1e+00
pack-comp2p-8	(107, 34, 49, 49)	6.73117e-01	6.73901e-01	3.90e-08	1e+00
pack-rig1-8	(87, 59, 46, 32)	7.87932e-01	7.87932e-01	1.91e-17	1e-01
pack-rig1c-8	(87, 66, 46, 32)	7.88300e-01	7.88300e-01	4.86e-17	1e-01
pack-rig1p-8	(105, 62, 49, 47)	7.87932e-01	7.87932e-01	2.63e-05	1e-08
pack-rig2-8	(85, 59, 46, 30)	7.80404e-01	7.80404e-01	6.61e-10	1e+00

Problem	(n, m, p, q)	$f_{MacMPEC}$	f_{opt}	$\max \text{Vio}(x_{opt})$	t_{fin}
pack-rig2c-8	(85, 66, 46, 30)	7.99306e-01	7.99306e-01	2.47e-09	1e+00
pack-rig2p-8	(103, 62, 49, 45)	7.80404e-01	7.80404e-01	1.87e-08	1e+00
pack-rig3-8	(85, 59, 46, 30)	7.35202e-01	7.35202e-01	1.17e-13	1e-01
pack-rig3c-8	(85, 66, 46, 30)	7.53473e-01	7.53473e-01	5.68e-06	1e-08
portfl-i-1	(87, 62, 13, 12)	1.50200e-05	1.50203e-05	1.00e-07	1e-07
portfl-i-2	(87, 62, 13, 12)	1.45700e-05	1.45696e-05	1.00e-07	1e-07
portfl-i-3	(87, 62, 13, 12)	6.26500e-06	6.26305e-06	1.00e-07	1e-07
portfl-i-4	(87, 62, 13, 12)	2.17700e-06	2.16125e-06	9.95e-05	1e-08
portfl-i-6	(87, 62, 13, 12)	2.36100e-06	2.34116e-06	1.21e-04	1e-08
qpec-100-1	(105, 2, 0, 100)	9.90028e-02	9.90009e-02	1.00e-07	1e-07
qpec-100-2	(110, 2, 0, 100)	-6.26049e+00	-6.43006e+00	1.00e-07	1e-07
qpec-100-3	(110, 4, 0, 100)	-5.48287e+00	-5.48058e+00	1.00e-07	1e-07
qpec-100-4	(120, 4, 0, 100)	-3.60073e+00	-3.98077e+00	1.00e-07	1e-07
qpec-200-1	(210, 4, 0, 200)	-1.93483e+00	-1.93484e+00	1.01e-07	1e-07
qpec-200-2	(220, 4, 0, 200)	-2.40299e+01	-2.36642e+01	1.00e-07	1e-07
qpec-200-3	(220, 8, 0, 200)	-1.95341e+00	-1.95341e+00	1.00e-07	1e-07
qpec-200-4	(240, 8, 0, 200)	-6.19323e+00	-6.18707e+00	1.00e-07	1e-07
qpec1	(30, 0, 0, 20)	8.00000e+01	8.00000e+01	0.00e+00	1e+00
qpec2	(30, 0, 0, 20)	4.50000e+01	4.49867e+01	3.32e-04	1e-08
ralph1	(2, 1, 0, 1)	0.00000e+00	-1.00028e-03	1.00e-03	1e-08
ralph2	(2, 0, 0, 1)	0.00000e+00	-3.00005e-08	1.00e-04	1e-08
ralphmod	(104, 8, 0, 100)	-6.83033e+02	-6.83033e+02	6.94e-05	1e-08
scale1	(2, 0, 0, 1)	1.00000e+00	1.00000e+00	0.00e+00	1e-06
scale2	(2, 0, 0, 1)	1.00000e+00	9.99998e-01	1.00e-06	1e-06
scale3	(2, 0, 0, 1)	1.00000e+00	1.00000e+00	0.00e+00	1e-06
scale4	(2, 0, 0, 1)	1.00000e+00	9.80100e-01	1.00e-04	1e-08
scale5	(2, 0, 0, 1)	1.00000e+02	9.99998e+01	1.00e-06	1e-06
scholtes1	(3, 1, 0, 1)	2.00000e+00	2.00000e+00	0.00e+00	1e+00
scholtes2	(3, 1, 0, 1)	1.50000e+01	1.50000e+01	0.00e+00	1e+00
scholtes3	(2, 0, 0, 1)	5.00000e-01	5.00000e-01	0.00e+00	1e-06
scholtes4	(3, 2, 0, 1)	-3.07336e-07	-6.95884e-04	3.48e-04	1e-08
scholtes5	(3, 0, 0, 2)	1.00000e+00	1.00000e+00	0.00e+00	1e+00
sl1	(8, 6, 2, 3)	1.00000e-04	1.00000e-04	4.09e-16	1e+00
stackelberg1	(3, 2, 1, 1)	-3.26667e+03	-3.26667e+03	1.42e-14	1e+00
tap-09	(86, 40, 32, 32)	1.09143e+02	1.21566e+02	2.00e-07	1e-07
tap-15	(194, 94, 68, 83)	1.84295e+02	1.87258e+02	2.00e-07	1e-07
water-fl	(213, 310, 116, 44)	3.41192e+03	3.34422e+03	1.37e-05	1e-08
water-net	(66, 96, 36, 14)	9.29169e+02	9.27138e+02	7.93e-03	1e-08

Table 2: Results for MacMPEC collection

6 Final Remarks

This paper proposes a new regularization scheme for the solution of mathematical programs with complementarity constraints. The method was shown to converge at least to M-stationary points which is a much stronger property than what is known for the majority of other regularization methods. Moreover, convergence to these M-stationary points (and also to strong stationary points under an additional condition) is shown under significantly weaker assumptions than those used in related approaches. The numerical results indicate that the method works quite well in practice, even without special tuning of the particular NLP-solver that is applied to the regularized problems. This is in contrast, for example, to the methods by Kadrani et al. [21] and Steffensen and Ulbrich [38] where special care has to be taken in order to overcome some difficulties arising from the particular regularization used in these two approaches.

Finally, we believe that the new regularization idea used in this paper can also be adapted to the class of mathematical programs with vanishing constraints in order to get stronger convergence properties for relaxation schemes for this class of methods, see, e.g., [1, 16, 17, 18] for some relevant literature regarding this problem class.

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] R. ANDREANI, J.M. MARTÍNEZ, AND M.L. SCHUVERDT: *The CPLD condition of Q_i and Wei implies the quasinormality constraint qualification*. Journal of Optimization Theory and Applications, 125, 2005, pp. 473–485.
- [3] 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.
- [4] 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.
- [5] M.S. BAZARAA AND C.M. SHETTY: *Foundations of Optimization*. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin/Heidelberg, 1976.
- [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] S. DEMPE: *Foundations of Bilevel Programming*. Nonconvex Optimization and Its Applications, Vol. 61, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

- [8] F. FACCHINEI, H. JIANG, AND L. QI: *A smoothing method for mathematical programs with equilibrium constraints*. Mathematical Programming 85, 1999, pp. 107–134.
- [9] M.C. FERRIS, R. FOURER, AND D.M. GAY: *Expressing complementarity problems in an algebraic modeling language and communicating them to solvers*. SIAM Journal on Optimization 9, 1999, pp. 991–1009.
- [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 the Guignard constraint qualification for mathematical programs with equilibrium constraints*. Optimization 54, 2005, pp. 517–534.
- [12] 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.
- [13] R. FLETCHER AND S. LEYFFER: *Solving mathematical programs with complementarity constraints as nonlinear programs*. Optimization Methods and Software 19, 2004, pp. 15–40.
- [14] M. FUKUSHIMA AND J.-S. PANG: *A globally convergent sequential quadratic programming algorithm for mathematical programming with linear complementarity constraints*. In M. THÉRA AND R. TICHATSCHKE (Eds.): *Ill-posed Variational Problems and Regularization Technique*. Lecture Notes in Economics and Mathematical Systems 477, Springer-Verlag, Berlin/Heidelberg, 1999, pp. 99–110.
- [15] D.M. GAY: *Hooking your solver to AMPL*. Technical Report 97-4-06, Computing Sciences Research Center, Bell Laboratories, Murray Hill, New Jersey 07974, see also www.ampl.com/REFS/hooks2.pdf.
- [16] T. HOHEISEL AND C. KANZOW: *On the Abadie and Guignard constraint qualification for mathematical programs with vanishing constraints*. Optimization 58, 2009, pp. 431–448.
- [17] 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.
- [18] T. HOHEISEL, C. KANZOW, AND A. SCHWARTZ: *Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints*. Technical Report 293, Institute of Mathematics, University of Würzburg, Würzburg, February 2010.
- [19] 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.

- [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] 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.
- [22] S. LEYFFER: *MacMPEC: AMPL collection of MPECs*, www.mcs.anl.gov/~leyffer/MacMPEC, 2000
- [23] S. LEYFFER: *Complementarity constraints as nonlinear equations: Theory and numerical experience*. In S. DEMPE AND V. KALASHNIKOV (Eds.): *Optimization and Multivalued Mappings*. Springer-Verlag, 2006, pp. 169–208.
- [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] 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.
- [26] X.-W. LIU AND J. SUN: *Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints*. Mathematical Programming, Series B 101, 2004, pp. 231–26.
- [27] Z.-Q. LUO, J.-S. PANG, AND D. RALPH: *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, New York, Melbourne, 1996.
- [28] www.netlib.org/ampl/solvers/examples/amplfunc.c
- [29] O.L. MANGASARIAN: *Equivalence of the complementarity problem to a system of nonlinear equations*. SIAM Journal on Applied Mathematics 31, 1976, pp. 89–92.
- [30] J.V. OUTRATA: *Optimality conditions for a class of mathematical programs with equilibrium constraints*. Mathematics of Operations Research 24, 1999, pp. 627–644.
- [31] J.V. OUTRATA: *A generalized mathematical program with equilibrium constraints*. SIAM Journal on Control and Optimization 38, 2000, pp. 1623–1638.
- [32] 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.
- [33] 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.

- [34] 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.
- [35] D. RALPH AND S.J. WRIGHT: *Some properties of regularization and penalization schemes for MPECs*. Optimization Methods and Software 19, 2004, pp. 527–556.
- [36] H. SCHEEL AND S. SCHOLTES: *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*. Mathematics of Operations Research 25, 2000, pp. 1–22.
- [37] S. SCHOLTES: *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*. SIAM Journal on Optimization 11, 2001, pp. 918–936.
- [38] S. STEFFENSEN AND M. ULBRICH: *A new regularization scheme for mathematical programs with equilibrium constraints*. SIAM Journal on Optimization, to appear.
- [39] O. STEIN: *Lifting mathematical programs with complementarity constraints*. Mathematical Programming, to appear.
- [40] J.J. YE: *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*. SIAM Journal on Optimization 10, 2000, pp. 943–962.
- [41] 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.