THEORETICAL AND NUMERICAL COMPARISON OF RELAXATION METHODS FOR MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS[1]

Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz

Preprint 299 September 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

September 3, 2010

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

Abstract. Mathematical programs with equilibrium constraints (MPECs) are difficult optimization problems whose feasible sets do not satisfy most of the standard constraint qualifications. Hence MPECs cause difficulties both from a theoretical and a numerical point of view. As a consequence, a number of MPEC-tailored solution methods have been suggested during the last decade which are known to converge under suitable assumptions. Among these MPEC-tailored solution schemes, the relaxation methods are certainly one of the most prominent class of solution methods. Several different relaxation schemes are available in the meantime, and the aim of this paper is to provide a theoretical and numerical comparison of these schemes. More precisely, in the theoretical part, we improve the convergence theorems of several existing relaxation methods. There, we also take a closer look at the properties of the feasible sets of the relaxed problems and show which standard constraint qualifications are satisfied for these relaxed problems. Finally, the numerical comparison is based on the MacMPEC test problem collection.

Key Words: Mathematical programs complementarity constraints, Relaxation method, Constraint qualification, Global convergence, Performance profiles

Mathematics Subject Classification: 65K05, 90C30, 90C31

1 Introduction

We consider a nonlinear optimization problem of the form

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

where $f, g_i, h_i, G_i, H_i : \mathbb{R}^n \to \mathbb{R}$ are assumed to be at least continuously differentiable functions. It is known under the label *mathematical program with complementarity constraints*, *MPCC* for short, but we prefer the more pronounceable moniker *MPEC*, which stands for *mathematical program with equilibrium constraints*. Its feasible set will be denoted by X.

The MPEC and the closely related class of *bilevel programs* has a number of important applications arising from different areas like Stackelberg games, robotics, or several optimal design problems. The reader is referred to the corresponding monographs [8, 26, 28] for more details and several properties of MPECs.

From a theoretical point of view, an MPEC is a highly difficult nonlinear program (NLP) since the standard Mangasarian-Fromovitz constraint qualification (and, therefore, also the stronger linear independence constraint qualification) is violated at any feasible point of an MPEC, cf. [36]. This, in turn, implies that standard algorithms for NLPs typically have severe difficulties in solving MPECs.

This observation was and still is the main motivation for the introduction of more specialized algorithms for the solution of MPECs that take into account the particular structure of an MPEC. Several different ideas are known in the literature, like smoothing, penalization, lifting, relaxation (or regularization), and suitable modifications of standard NLP solvers. We refer the interested reader to [1, 2, 7, 10, 14, 18, 19, 21, 24, 25, 30, 31, 33, 35, 34, 22] for more details.

The aim of this paper is to provide a theoretical and numerical comparison of one class of solutions methods, namely the class of relaxation methods. To this end, we take a closer look at the following relaxation schemes:

- the global relaxation method by Scholtes [33]
- the smooth relaxation method by Lin and Fukushima [25]
- the nonsmooth relaxation method by Kadrani et al. [21]
- the local relaxation method by Steffensen and Ulbrich [34]
- the relaxation method by Kanzow and Schwartz [22].

In addition to these five relaxation schemes, the authors are also aware of the two-sided relaxation method by Demiguel et al. [7] whose motivation, however, is slightly different: In all five relaxation schemes above, there is a single parameter t that is driven down to zero, whereas in the two-sided approach from [7], there are different parameters, and some of them should eventually stay fixed. Hence, although it would, in principle, be possible to include the two-sided relaxation within our comparison by simply taking a single parameter t that is driven to zero also in that approach, we believe that this is not consistent with the original idea from [7], hence we discard that method from our theoretical and numerical comparison.

Several convergence properties are known for the above-mentioned five relaxation methods that can be found in the corresponding references [33, 25, 21, 34, 22] as well as in some subsequent works [31, 16, 17]. However, at least for some of these relaxation schemes, it turns out that the existing convergence results can still be improved significantly. Moreover, a numerical comparison among these methods is still missing.

The main motivation and contribution of this paper is therefore to

- improve the convergence theorems of several existing relaxation methods for MPECs
- show which (standard) constraint qualifications are satisfied by the relaxed problems (this, in particular, guarantees the existence of Lagrange multipliers at a local minimum of the relaxed problem), and to
- present a numerical comparison of all relaxation schemes based on the MacMPEC test problem collection from Leyffer [23].

To this end, we organize the paper in the following way: Section 2 states some background material on constraint qualifications and stationarity concepts. Section 3 then contains the main theoretical contribution of this paper, with each of the five relaxation schemes being discussed in a separate subsection. The numerical comparison is given in Section 4, and we close with some final remarks in Section 5.

The notation used within this paper is standard. The gradient of a differentiable mapping $f : \mathbb{R}^n \to \mathbb{R}$ will be denoted by $\nabla f(x)$ and is always viewed as a column vector. Given an index set I, its cardinality is indicated by $|I|$. Finally, we use the symbol $\mathrm{supp}(z) := \{i \mid z_i \neq 0\}$ for the support of a vector $z \in \mathbb{R}^n$.

2 Preliminaries

2.1 Constraint Qualifications for Standard Problems

Let us consider the standard nonlinear program

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

with continuously differentiable functions $f, g_i, h_i : \mathbb{R}^n \to \mathbb{R}$. Furthermore, let $\mathcal{X}$ denote the feasible set of (2), and let $x^* \in \mathcal{X}$ be an arbitrary feasible point. Recall that the

(Bouligand) *tangent cone* of $\mathcal{X}$ at x^* is defined by

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

whereas the *linearized cone* of $\mathcal{X}$ at x^* is given by

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

where

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

denotes the set of active inequality constraints. Moreover, for an arbitrary nonempty set $C \subseteq \mathbb{R}^n$, the set

$$C^\circ := \{v \in \mathbb{R}^n \mid v^T d \le 0 \ \forall d \in C\}$$

is called the *polar cone* of C.

We are now in a position to define the standard constraint qualifications needed. We say that $x^* \in \mathcal{X}$ 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 neighborhood $N(x^*)$ 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 $\mathcal{T}_{\mathcal{X}}(x^*) = \mathcal{L}_{\mathcal{X}}(x^*)$;

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

The reader might be less familiar with the CRCQ condition that was introduced in [20]. Moreover, also GCQ, which goes back to [15], maybe less known, but it has turned out to be the only standard constraint qualification which is still useful in an MPEC context, see, e.g., [11].

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* (2) *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.

Remark 2.2 A point x^* feasible for (2) satisfies MFCQ if and only if the set of vectors

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

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 [29]: 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

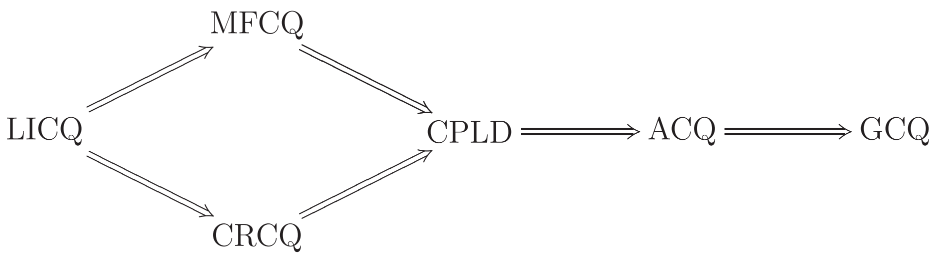
  $$\big\{\nabla g_i(x^*) \mid i \in I_1\big\} \cup \big\{\nabla h_i(x^*) \mid i \in I_2\big\}$$

  are positive-linearly dependent, there exists a neighborhood $N(x^*)$ 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:



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. Moreover, it is an immediate consequence of the definitions that ACQ implies GCQ. Finally, the fact that CPLD implies ACQ follows from results in [3, 6]. In particular, this means that CPLD is in fact a constraint qualification for (2), which was not clear when this condition was introduced originally in [29].

Now, let

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

be the *Lagrangian* of the optimization problem (2). Then, given a local minimum x^* of (2) such that a suitable constraint qualification, like, e.g., GCQ (or any of the other constraint qualifications mentioned before), holds at x^*, then x^* is a *stationary point* of (2), i.e., there exist multipliers $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ such that (x^*, λ, μ) is a *KKT point* of (2), i.e., this triple satisfies the corresponding *KKT conditions*

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

hence we call the x-part of a KKT point a stationary point.

2.2 MPEC-tailored Constraint Qualifications and Stationarity Concepts

Due to the well-known fact that all standard constraint qualifications, except for GCQ, are far too restrictive for MPECs, cf. [11], a whole bunch of more problem-tailored constraint qualifications has been introduced in the past. In this section we recall some of them.

Moreover, due to the fact that most standard CQs are likely to be violated at a local minimizer of an MPEC, the KKT conditions cannot readily be considered as the proper stationarity concept. In view of that, a couple of more appropriate and, hence, necessarily weaker stationarity notions have arisen. In order to state these, it is helpful to define some crucial index sets that will occur frequently in the subsequent analysis.

Let x^* be feasible for (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.

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

(a) weakly stationary, *if there are multipliers $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p, \gamma, \nu \in \mathbb{R}^l$ 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 $\lambda_i \geq 0,\ \lambda_i g_i(x^*) = 0\ (i = 1, \dots, l),\ \gamma_i = 0\ (i \in I_{+0}),\ \nu_i = 0\ (i \in I_{0+})$;

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

Note that strong stationarity implies M-stationarity, M-stationarity implies C-stationarity, and C-stationarity, in turn, implies weak stationarity. Furthermore, it can be shown that strong stationarity is equivalent to the standard KKT conditions of an MPEC, cf. [11]. However, even for very simple MPECs, strong stationarity may not hold at a global minimum, see [32] for a counterexample. Therefore, in general, M-stationarity is the strongest stationary concept that one can expect to hold at a local minimum under suitable assumptions.

In order to define the MPEC constraint qualifications needed for our analysis, we make use of the auxiliary program

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

which is called the *tightened nonlinear program* TNLP(x^*). Note that TNLP(x^*) substantially depends on the chosen point x^*. This program leads to the following handy definition.

Definition 2.4 *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^*).*

Since we frequently employ it in the following, let us write down MPEC-MFCQ and MPEC-CPLD explicitly.

A point $x^* \in X$ satisfies MPEC-MFCQ if and only if the gradients

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

are linearly independent, and there exists a vector $d \in \mathbb{R}^n$ such that

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

On the other hand, $x^* \in X$ satisfies MPEC-CPLD if and only 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

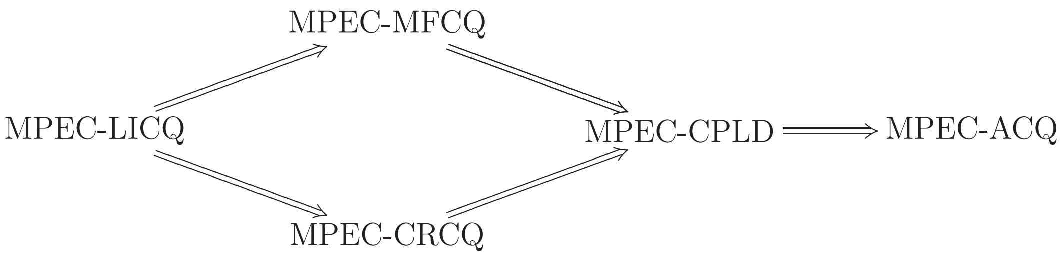
$$\{\nabla g_i(x^*) \mid i \in I_1\} \cup \Big\{ \{\nabla h_i(x^*) \mid i \in I_2\} \cup \{\nabla G_i(x^*) \mid i \in I_3\} \cup \{\nabla H_i(x^*) \mid i \in I_4\} \Big\} \quad (3)$$

are positive-linearly dependent, there exists a neighborhood $N(x^*)$ 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 (3), we put an extra pair of curly brackets around those vectors where no sign constraints occur in the definition of positive linear dependence.

The relation between these MPEC-CQs are shown in the following diagram:



Note that this diagram also includes the MPEC-ACQ whose definition is different from the other MPEC-CQs, so we do not include it here, especially since it will not be used in our analysis. The reader is referred to [12] for a precise definition of MPEC-ACQ, as well as to [16] and references therein for a justification of the above implications. It is known that, given a local minimizer x^* of the MPEC satisfying MPEC-ACQ (or any of the other MPEC-CQs defined above), then x^* is an M-stationary point, see [11, 12, 13].

3 Convergence Properties of Relaxation Schemes

The basic idea of all relaxation schemes is to get rid of the complicated complementarity constraints

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

by replacing these conditions in a suitable way such that the corresponding relaxed problem has nicer properties. The relaxed problem depends on a parameter $t > 0$ which has to be driven to zero in order to re-obtain the underlying MPEC.

For all relaxation schemes discussed here (we will discuss them in chronological order of their date of publication), suitable convergence results are already known. Typically, the most basic convergence results are as follows: Given a sequence $t_k \to 0$ and a corresponding sequence of stationary points x^k of the relaxed problems $R(t_k)$ such that x^k converges to x^* and such that a suitable MPEC CQ holds at x^*, then x^* is a C-stationary point (for three of the methods to be discussed below) or an M-stationary point (for the remaining two methods). Furthermore, under additional conditions, one can verify that the limit point x^* has further properties, like being M- or even strongly stationary.

In our subsequent analysis, we try to improve only the most basic convergence results by relaxing the corresponding MPEC CQs. The additional results which guarantee stronger properties of the limit point x^* are not discussed here since a corresponding generalization of the existing results are usually straightforward (in the sense that the techniques from the corresponding papers can be used also in our weaker framework). However, we also show which standard constraint qualification holds for the relaxed problems (hence these relaxed problems then possess multipliers at a local minimum). In some cases, our results generalize existing ones, in other cases, we prove completely new results.

3.1 The Global Relaxation Scheme by Scholtes

Probably the first attempt to use a relaxation idea for solving MPECs goes back to Scholtes [33]. It is closely related to the smoothing-type method by Facchinei et al. [10]. Some local properties of Scholtes' approach around a strongly stationary point can also be found in Ralph and Wright [31].

The basic idea of the relaxation scheme by Scholtes is to replace the MPEC by a sequence of the parametrized NLPs of the form (see Figure 1)

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

We denote the feasible set by $X^S(t)$. Since, geometrically, this is a global relaxation of the complementarity conditions, we call this approach the *global relaxation method*.

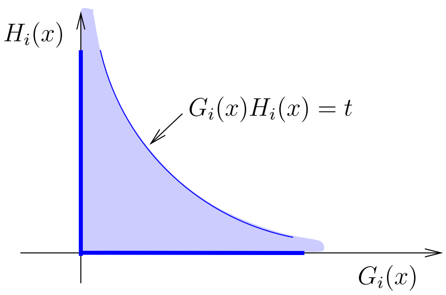


Figure 1: Geometric interpretation of the relaxation method by Scholtes [33]

For the convergence analysis, some index sets are needed:

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

The following is the most basic convergence result for the global relaxation method.

Theorem 3.1 *Let $\{t_k\} \downarrow 0$ and let x^k be a stationary point of $R^S(t_k)$ with $x^k \to x^*$ such that MPEC-MFCQ holds at x^*. Then x^* is a C-stationary point of* (1).

Proof. Since x^k is a KKT point of $R^S(t_k)$ there exist multipliers $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$ such that

$$\begin{aligned} 0 = {} & \nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^{l} \gamma_i^k \nabla G_i(x^k) \\ & - \sum_{i=1}^{l} \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^{l} \delta_i^k [H_i(x^k)\nabla G_i(x^k) + G_i(x^k)\nabla H_i(x^k)] \end{aligned} \tag{4}$$

with

$$\begin{aligned} \lambda^k \geq 0 \quad &\text{and} \quad \text{supp}(\lambda^k) \subseteq I_g(x^k), \\ \gamma^k \geq 0 \quad &\text{and} \quad \text{supp}(\gamma^k) \subseteq I_G(x^k), \\ \nu^k \geq 0 \quad &\text{and} \quad \text{supp}(\nu^k) \subseteq I_H(x^k), \\ \delta^k \geq 0 \quad &\text{and} \quad \text{supp}(\delta^k) \subseteq I_{GH}(x^k; t_k) \end{aligned}$$

for all $k \in \mathbb{N}$. This implies

$$\text{supp}(\gamma^k) \cap \text{supp}(\delta^k) = \emptyset, \quad \text{supp}(\nu^k) \cap \text{supp}(\delta^k) = \emptyset \tag{5}$$

for all $k \in \mathbb{N}$. Moreover, for all $k \in \mathbb{N}$ sufficiently large, we have $I_g(x^k) \subseteq I_g$, $I_G(x^k) \subseteq I_{00} \cup I_{0+}$, and $I_H(x^k) \subseteq I_{00} \cup I_{+0}$.

Our next step is to define suitable new multipliers

$$\tilde{\gamma}_i^k = \begin{cases} \gamma_i^k, & \text{if } i \in \operatorname{supp}(\gamma^k), \\ -\delta_i^k H_i(x^k), & \text{if } i \in \operatorname{supp}(\delta^k) \backslash I_{+0}, \\ 0, & \text{else,} \end{cases}$$

and

$$\tilde{\nu}_i^k = \begin{cases} \nu_i^k, & \text{if } i \in \operatorname{supp}(\nu^k), \\ -\delta_i^k G_i(x^k), & \text{if } i \in \operatorname{supp}(\delta^k) \backslash I_{0+}, \\ 0, & \text{else.} \end{cases}$$

With these multipliers, we can rewrite (4) as

$$\begin{aligned} 0 \;=\; & \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^l \tilde{\gamma}_i^k \nabla G_i(x^k) - \sum_{i=1}^l \tilde{\nu}_i^k \nabla H_i(x^k) \\ & + \sum_{i \in I_{+0}} \delta_i^k H_i(x^k) \nabla G_i(x^k) + \sum_{i \in I_{0+}} \delta_i^k G_i(x^k) \nabla H_i(x^k). \end{aligned}$$

If we assume that the sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k, \delta^k_{I_{+0} \cup I_{0+}})\}$ is unbounded, then one can find a subsequence K such that the normed sequence converges:

$$\frac{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k, \delta^k_{I_{+0} \cup I_{0+}})}{\|(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k, \delta^k_{I_{+0} \cup I_{0+}})\|} \to_K (\lambda, \mu, \tilde{\gamma}, \tilde{\nu}, \delta_{I_{+0} \cup I_{0+}}) \neq 0.$$

The equation above then yields

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

where $\lambda \geq 0$ and for all $k \in K$ sufficiently large

$$\begin{aligned} \operatorname{supp}(\lambda) &\subseteq I_g(x^k) \subseteq I_g, \\ \operatorname{supp}(\tilde{\gamma}) &\subseteq I_G(x^k) \cup I_{GH}(x^k; t_k) \backslash I_{+0} \subseteq I_{00} \cup I_{0+}, \\ \operatorname{supp}(\tilde{\nu}) &\subseteq I_H(x^k) \cup I_{GH}(x^k; t_k) \backslash I_{0+} \subseteq I_{00} \cup I_{+0}. \end{aligned}$$

Additionally, $(\lambda, \mu, \tilde{\gamma}, \tilde{\nu}) \neq 0$ has to hold. Otherwise, $\delta_i > 0$ would have to hold for at least one $i \in I_{+0} \cup I_{0+}$. Assume without loss of generality $\delta_i > 0$ for an $i \in I_{+0}$. This implies $\delta_i^k > 0$ for all k sufficiently large and consequently $\tilde{\nu}_i^k = -\delta_i^k G_i(x^k)$ for those k. Because of $i \in I_{+0}$, this yields $\tilde{\nu}_i = \lim_{k \in K} -\delta_i^k G_i(x^k) < 0$, a contradiction to our assumption $\tilde{\nu} = 0$.

However, $(\lambda, \mu, \tilde{\gamma}, \tilde{\nu}) \neq 0$ is a contradiction to the prerequisite that MPEC-MFCQ holds in x^*. Thus, we may assume without loss of generality that the sequence is convergent to some vector $(\lambda^*, \mu^*, \tilde{\gamma}^*, \tilde{\nu}^*, \delta^*_{I_{+0} \cup I_{0+}})$. It is easy to see that $\lambda^* \geq 0$ and $\operatorname{supp}(\lambda^*) \subseteq I_g$. According to the definition of $\tilde{\gamma}^k$ and $\tilde{\nu}^k$, we have

$$\operatorname{supp}(\tilde{\gamma}^*) \subseteq I_{00} \cup I_{0+}, \quad \operatorname{supp}(\tilde{\nu}^*) \subseteq I_{00} \cup I_{+0}.$$

The continuous differentiability of f, g, h, G, H then implies

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

To prove the C-stationarity of x^*, it remains to show that $\gamma_i^* \nu_i^* \geq 0$ for all $i \in I_{00}$. Assume that there is an $i \in I_{00}$ with $\gamma_i^* < 0$ and $\nu_i^* > 0$ or with $\nu_i^* > 0$ and $\gamma_i^* < 0$. We consider only the first case, the second one can be treated similarly. Because of $\gamma_i^k \geq 0$, the condition $\gamma_i^* < 0$ implies $i \in \operatorname{supp}(\delta^k)$ for all $k \in \mathbb{N}$ sufficiently large. This implies $i \notin \operatorname{supp}(\nu^k)$ in view of (5) and, therefore, $\nu_i^* \leq 0$ in contradiction to our assumption. □

Note that the corresponding result in Scholtes [33] assumes MPEC-LICQ and shows that the sequence of multipliers corresponding to the stationary points x^k converges, whereas here we assume the weaker MPEC-MFCQ which, obviously, does not guarantee convergence of the corresponding sequence of multipliers, but the proof shows that one can extract a sequence of multipliers which stays bounded and is, therefore, convergent at least on a subsequence.

The assumption of x^k being a stationary point of the relaxed problem $R^S(t_k)$ assumes the existence of multipliers. A priori, it is not clear that these multipliers really exist. The following result essentially guarantees the existence of these multipliers by showing that MPEC-MFCQ at a feasible point x^* of the original MPEC implies that standard MFCQ holds for the relaxed problems $R^S(t)$, at least locally around x^*.

Theorem 3.2 *Let x^* be feasible for (1) such that MPEC-MFCQ is satisfied at x^*. Then there exists a neighborhood N of x^* and $\bar{t} > 0$ such that standard MFCQ for $R^S(t)$ is satisfied at all $x \in N \cap X^S(t)$.*

Proof. First note that, by continuity, for all $x \in X^S(t)$ sufficiently close to x^*, we have

$$\begin{array}{rcl} I_g(x) & \subseteq & I_g, \\ I_G(x) & \subseteq & I_{00} \cup I_{0+}, \\ I_H(x) & \subseteq & I_{00} \cup I_{+0}, \\ I_{GH}(x) \cap I_G(x) & = & \emptyset, \\ I_{GH}(x) \cap I_H(x) & = & \emptyset. \end{array} \qquad (6)$$

Since MPEC-MFCQ holds, the gradients

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

are positive-linearly independent. In view of [29, Prop. 2.2], this implies that the set of gradients

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

is also positive-linearly independent for all $x \in X^S(t)$ sufficiently close to x^*. Taking into account that

$$I_G(x) \cup \big(I_{GH}(x) \cap I_{0+}\big) \cup \big(I_{GH}(x) \cap I_{00}\big) \subseteq I_{00} \cup I_{0+}$$

and

$$I_H(x) \cup \big(I_{GH}(x) \cap I_{+0}\big) \cup \big(I_{GH}(x) \cap I_{00}\big) \subseteq I_{00} \cup I_{+0}$$

for all $x \in X^S(t)$ sufficiently close to x^* and using the fact that $G_i(x) > 0, H_i(x) \approx 0$ for all $i \in I_{+0}$ as well as $G_i(x) \approx 0, H_i(x) > 0$ for all $i \in I_{0+}$ whenever x is close to x^*, it follows that there is a neighbourhood $N(x^*)$ such that the set of vectors

$$\begin{array}{rl} \nabla g_i(x) & (i \in I_g(x)), \\ \nabla h_i(x) & (i = 1, \ldots, p), \\ \nabla G_i(x) & (i \in I_G(x)), \\ \nabla H_i(x) & (i \in I_H(x)), \\ G_i(x)\nabla H_i(x) + H_i(x)\nabla G_i(x) & (i \in I_{GH}(x) \cap I_{0+}), \\ G_i(x)\nabla H_i(x) + G_i(x)\nabla H_i(x) & (i \in I_{GH}(x) \cap I_{+0}), \\ \nabla G_i(x) & (i \in I_{GH}(x) \cap I_{00}), \\ \nabla H_i(x) & (i \in I_{GH}(x) \cap I_{00}) \end{array} \tag{7}$$

is positive-linearly independent for all $x \in X^S(t) \cap N(x^*)$.

We now claim that standard MFCQ holds for the relaxed program $R^S(t)$ whenever $x \in X^S(t) \cap N(x^*)$. To this end, take an arbitrary $x \in X^S(t) \cap N(x^*)$. In view of Remark 2.2, we have to show that

$$\begin{aligned} 0 \;=\; & \sum_{i \in I_g(x)} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \mu_i \nabla h_i(x) - \sum_{i \in I_G(x)} \alpha_i \nabla G_i(x) \\ & - \sum_{i \in I_H(x)} \beta_i \nabla H_i(x) + \sum_{i \in I_{GH}(x)} \gamma_i (G_i(x)\nabla H_i(x) + H_i(x)\nabla G_i(x)) \end{aligned} \tag{8}$$

with $\mu \in \mathbb{R}^p$ and $\lambda, \alpha, \beta, \gamma \geq 0$ holds only for the trivial vector. To see this, we rewrite (8) as

$$\begin{aligned} 0 \;=\; & \sum_{i \in I_g(x)} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \mu_i \nabla h_i(x) - \sum_{i \in I_G(x)} \alpha_i \nabla G_i(x) - \sum_{i \in I_H(x)} \beta_i \nabla H_i(x) \\ & + \sum_{i \in I_{GH}(x) \cap (I_{0+} \cup I_{+0})} \gamma_i (G_i(x)\nabla H_i(x) + H_i(x)\nabla G_i(x)) \\ & + \sum_{i \in I_{00} \cap I_{GH}(x)} (\gamma_i G_i(x)) \nabla H_i(x) + \sum_{i \in I_{00} \cap I_{GH}(x)} (\gamma_i H_i(x)) \nabla G_i(x). \end{aligned} \tag{9}$$

Applying the positive-linear independence of the vectors from (7) to (9) and using (6), we immediately obtain that all coefficients from (8) are zero, and this completes the proof. $\square$

3.2 The Relaxation Scheme by Lin and Fukushima

The relaxation scheme proposed by Lin and Fukushima in [25] employs the following relaxation, see Figure 2:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0, \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & G_i(x)H_i(x) - t^2 \le 0 \quad \forall i = 1, \dots, l, \\ & (G_i(x) + t)(H_i(x) + t) - t^2 \ge 0 \quad \forall i = 1, \dots, l. \end{array} \qquad R^{LF}(t)$$

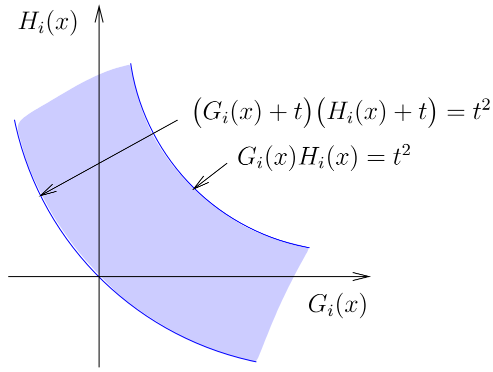


Figure 2: Geometric interpretation of the relaxation method by Lin and Fukushima [25]

Its feasible set is denoted by $X^{LF}(t)$.

The corresponding convergence theorem is given below, see [17] for a proof.

Theorem 3.3 *Let $\{t_k\} \downarrow 0$ and let x^k be a stationary point of $R^{LF}(t)$ with $x^k \to x^*$ such that MPEC-MFCQ holds in x^*. Then x^* is C-stationary.*

Note that a result similar to the previous theorem was shown in [25], but under the stronger MPEC-LICQ condition.

We next see that MPEC-MFCQ at a feasible point x^* of the underlying MPEC implies that standard MFCQ holds for the corresponding regularized problems at an arbitrary point $x \in X^{LF}(t)$ sufficiently close to x^*. Again, we refer to [17] for a proof.

Theorem 3.4 *Let x^* be feasible for (1) such that MPEC-MFCQ is satisfied at x^*. Then there exists a neighborhood N of x^* such that standard MFCQ for $R^{LF}(t)$ is satisfied at all $x \in N \cap X^{LF}(t)$.*

Altogether, it follows that the relaxation scheme by Lin and Fukushima has the same theoretical properties as the previous relaxation method by Scholtes.

3.3 The Relaxation Scheme by Kadrani et al.

The approach in [21] proposes the following relaxation, see Figure 3:

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

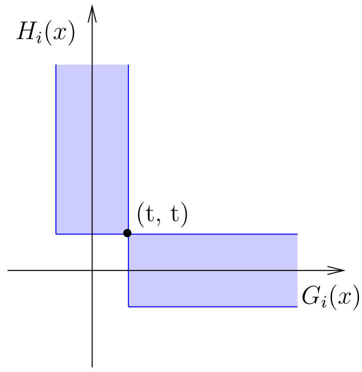


Figure 3: Geometric interpretation of the relaxation method by Kadrani et al. [21]

Its feasible set is denoted by $X^{KDB}(t)$.

For a refined convergence result, we need to define certain index sets. To this end, let x be feasible for $R^{KDB}(t)$. Then we set

$$
\begin{array}{rcl}
I_G(x,t) & := & \{i \mid G_i(x) + t = 0\}, \\
I_H(x,t) & := & \{i \mid H_i(x) + t = 0\}, \\
I_\Phi(x,t) & := & \{i \mid (G_i(x) - t)(H_i(x) - t) = 0\}, \\
I_\Phi^{0*}(x,t) & := & \{i \in I_\Phi(x,t) \mid G_i(x) - t = 0\}, \\
I_\Phi^{0+}(x,t) & := & \{i \in I_\Phi^{0*}(x,t) \mid H_i(x) - t > 0\}, \\
I_\Phi^{0-}(x,t) & := & \{i \in I_\Phi^{0*}(x,t) \mid H_i(x) - t < 0\}, \\
I_\Phi^{00}(x,t) & := & \{i \in I_\Phi(x,t) \mid G_i(x) - t = H_i(x) - t = 0\}, \\
I_\Phi^{*0}(x,t) & := & \{i \in I_\Phi(x,t) \mid H_i(x) - t = 0\}, \\
I_\Phi^{+0}(x,t) & := & \{i \in I_\Phi^{*0}(x,t) \mid G_i(x) - t > 0\}, \\
I_\Phi^{-0}(x,t) & := & \{i \in I_\Phi^{*0}(x,t) \mid G_i(x) - t < 0\}.
\end{array}
\qquad (10)
$$

The following is the main convergence result for the relaxation method by Kadrani et al. [21]. It generalizes a corresponding result from [21] by replacing the MPEC-LICQ assumption by the much weaker MPEC-CPLD condition.

Theorem 3.5 *Let $\{t_k\} \downarrow 0$ and assume that x^k is a stationary point of $R^{KDB}(t_k)$ for all $k \in \mathbb{N}$. Moreover, suppose that $x^k \to x^*$ such that MPEC-CPLD holds at x^*. Then x^* is an M-stationary point of* (1).

Proof. Note that in this proof we skip the standard constraints to keep the notation as compact as possible.

Since x^k is a KKT point of $R^{KDB}(t_k)$ for all k, there exist multipliers $\alpha^k, \beta^k, \gamma^k \geq 0$ such that

$$0 = \nabla f(x^k) - \sum_{i=1}^{l} \alpha_i^k \nabla G_i(x^k) - \sum_{i=1}^{l} \beta_i^k \nabla H_i(x^k) + \sum_{i=1}^{l} \gamma_i^k [(H_i(x^k) - t_k) \nabla G_i(x^k) + (G_i(x^k) - t_k) \nabla H_i(x^k)]$$

and

$$\alpha_i^k (G_i(x^k) + t_k) = 0, \quad \beta_i^k (H_i(x^k) + t_k) = 0, \quad \gamma_i^k (G_i(x^k) - t_k)(H_i(x^k) - t_k) = 0.$$

Now, put

$$\eta_i^{G,k} := -\gamma_i^k (H_i(x^k) - t_k), \quad \eta_i^{H,k} := -\gamma_i^k (G_i(x^k) - t_k).$$

Then we infer from the equations above that

$$0 = \nabla f(x^k) - \sum_{i=1}^{l} \alpha_i^k \nabla G_i(x^k) - \sum_{i=1}^{l} \beta_i^k \nabla H_i(x^k) - \sum_{i=1}^{l} \eta_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^{l} \eta_i^{H,k} \nabla H_i(x^k) \quad (11)$$

and, for all k sufficiently large,

$$\begin{array}{rcccl} \mathrm{supp}(\alpha^k) & \subseteq & I_G(x^k, t_k) & \subseteq & I_{00} \cup I_{0+}, \\ \mathrm{supp}(\beta^k) & \subseteq & I_H(x^k, t_k) & \subseteq & I_{00} \cup I_{+0}, \\ \mathrm{supp}(\eta^{G,k}) & \subseteq & I_\Phi^{0*}(x^k, t_k) & \subseteq & I_{00} \cup I_{0+}, \\ \mathrm{supp}(\eta^{H,k}) & \subseteq & I_\Phi^{*0}(x^k, t_k) & \subseteq & I_{00} \cup I_{+0}. \end{array} \quad (12)$$

Moreover, one sees that

$$\begin{array}{rcl} \mathrm{supp}(\alpha^k) \cap \mathrm{supp}(\eta^{G,k}) & = & \emptyset, \\ \mathrm{supp}(\beta^k) \cap \mathrm{supp}(\eta^{H,k}) & = & \emptyset, \\ \mathrm{supp}(\eta^{G,k}) \cap \mathrm{supp}(\eta^{H,k}) & = & \emptyset. \end{array} \quad (13)$$

In addition, we have

$$\begin{array}{rcl} i \in \mathrm{supp}(\eta^{G,k}) \cap \mathrm{supp}(\beta^k) & \Longrightarrow & \eta_i^{G,k} > 0, \\ i \in \mathrm{supp}(\eta^{H,k}) \cap \mathrm{supp}(\alpha^k) & \Longrightarrow & \eta_i^{H,k} > 0. \end{array} \quad (14)$$

Without loss of generality, cf. [34, Lem. A.1], we may assume that the following gradients

$$\left\{ \nabla G_i(x^k) \mid \mathrm{supp}(\alpha^k) \cup \mathrm{supp}(\eta^{G,k}) \right\} \cup \left\{ \nabla H_i(x^k) \mid \mathrm{supp}(\beta^k) \cup \mathrm{supp}(\eta^{H,k}) \right\} \quad (15)$$

are linearly independent. Now, we want to prove that the sequence $\{(\alpha^k, \beta^k, \eta^{G,k}, \eta^{H,k})\}$ is bounded. For these purposes, we assume the contrary. Nevertheless, we may suppose, without loss of generality, that there is a vector $(\tilde{\alpha}, \tilde{\beta}, \tilde{\eta}^G, \tilde{\eta}^H)$ such that

$$\frac{(\alpha^k, \beta^k, \eta^{G,k}, \eta^{H,k})}{\|(\alpha^k, \beta^k, \eta^{G,k}, \eta^{H,k})\|} \to (\tilde{\alpha}, \tilde{\beta}, \tilde{\eta}^G, \tilde{\eta}^H) \neq 0,$$

and, clearly, for all k (sufficiently large) one has

$$\begin{array}{ll} \operatorname{supp}(\tilde{\alpha}) \subseteq \operatorname{supp}(\alpha^k), & \operatorname{supp}(\tilde{\beta}) \subseteq \operatorname{supp}(\beta^k), \\ \operatorname{supp}(\tilde{\eta}^G) \subseteq \operatorname{supp}(\eta^{G,k}), & \operatorname{supp}(\tilde{\eta}^H) \subseteq \operatorname{supp}(\eta^{H,k}). \end{array} \tag{16}$$

By passing to the limit, (11) therefore yields

$$0 = \sum_{i=1}^{l} \tilde{\alpha}_i \nabla G_i(x^*) + \sum_{i=1}^{l} \tilde{\beta}_i \nabla H_i(x^*) + \sum_{i=1}^{l} \tilde{\eta}_i^G \nabla G_i(x^*) + \sum_{i=1}^{l} \tilde{\eta}_i^H \nabla H_i(x^*),$$

i.e., the gradients

$$\left\{ \nabla G_i(x^*) \mid \operatorname{supp}(\tilde{\alpha}) \cup \operatorname{supp}(\tilde{\eta}^G) \right\} \cup \left\{ \nabla H_i(x^*) \mid \operatorname{supp}(\tilde{\beta}) \cup \operatorname{supp}(\tilde{\eta}^H) \right\}$$

are (positive-) linearly dependent. This, in view of MPEC-CPLD, remains true for x^k instead of x^*. But in view of (16), this contradicts the linear independence in (15). Thus, we can infer that $\{(\alpha^k, \beta^k, \eta^{G,k}, \eta^{H,k})\}$ is bounded, that is, at least on a subsequence, we have

$$(\alpha^k, \beta^k, \eta^{G,k}, \eta^{H,k}) \to (\alpha, \beta, \eta^G, \eta^H),$$

for some vectors $\alpha, \beta, \eta^G, \eta^H$ satisfying

$$\begin{array}{ll} \operatorname{supp}(\alpha) \subseteq \operatorname{supp}(\alpha^k), & \operatorname{supp}(\beta) \subseteq \operatorname{supp}(\beta^k), \\ \operatorname{supp}(\eta^G) \subseteq \operatorname{supp}(\eta^{G,k}), & \operatorname{supp}(\eta^H) \subseteq \operatorname{supp}(\eta^{H,k}). \end{array} \tag{17}$$

Now, for $i = 1, \ldots, l$, put

$$\lambda_i^G := \begin{cases} \alpha_i, & \text{if } i \in \operatorname{supp}(\alpha), \\ \eta_i^G, & \text{if } i \in \operatorname{supp}(\eta^G), \\ 0, & \text{else,} \end{cases} \quad \text{and} \quad \lambda_i^H := \begin{cases} \beta_i, & \text{if } i \in \operatorname{supp}(\beta), \\ \eta_i^H, & \text{if } i \in \operatorname{supp}(\eta^H), \\ 0, & \text{else.} \end{cases}$$

In view of (17) and (13), λ^G and λ^H are at least well-defined. We now show that x^*, together with the multipliers λ^G, λ^H, is an M-stationary point. To this end, first note that (11) implies

$$0 = \nabla f(x^*) - \sum_{i=1}^{l} \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^{l} \lambda_i^H \nabla H_i(x^*).$$

Furthermore, note that, for $i \in I_{+0}$, we have $i \notin \operatorname{supp}(\alpha^k) \cup \operatorname{supp}(\eta^{G,k})$ in view of (12). Using (17), this implies $i \notin \operatorname{supp}(\alpha) \cup \operatorname{supp}(\eta^G)$, hence the definition of λ^g gives $\lambda_i^G = 0$.

A symmetric argument shows that $\lambda_i^H = 0$ for all $i \in I_{0+}$. This means that x^* is at least weakly stationary. Furthermore, if either $\lambda_i^G = 0$ or $\lambda_i^H = 0$, the M-stationarity conditions are satisfied for such an index i. Consequently, taking into account the definitions of λ_i^G, λ_i^H, it remains to consider four cases.

Case 1: $i \in \text{supp}(\alpha) \cap \text{supp}(\beta)$. Then $\lambda_i^G = \alpha_i \geq 0$ and $\lambda_i^H = \beta_i \geq 0$, so that the M-stationarity conditions hold for such an index.

Case 2: $i \in \text{supp}(\alpha) \cap \text{supp}(\eta^H)$. Then $i \in \text{supp}(\alpha^k) \cap \text{supp}(\eta^{H,k})$ for all $k \in \mathbb{N}$ sufficiently large, cf. (17). Hence (14) implies $\eta_i^{H,k} > 0$ for all $k \in \mathbb{N}$ sufficiently large which, in turn, gives $\eta_i^H \geq 0$, hence $\lambda_i^H \geq 0$. Furthermore, since $i \in \text{supp}(\eta^{H,k})$, we have $i \notin \text{supp}(\eta^{G,k})$ by (13), hence $i \notin \text{supp}(\eta^G)$ by (14). This implies $\lambda_i^G \geq 0$ and shows that the M-stationarity conditions also holds for an index i from Case 2.

Case 3: $i \in \text{supp}(\eta^G) \cap \text{supp}(\beta)$. Here a symmetric reasoning to Case 2 shows that the M-stationary conditions also hold in this case.

Case 4: $i \in \text{supp}(\eta^G) \cap \text{supp}(\eta^H)$. Then (14) implies that $i \in \text{supp}(\eta^{G,k}) \cap \text{supp}(\eta^{H,k})$ for all $k \in \mathbb{N}$ sufficiently large. In view of (13), we see that this case cannot occur.

Altogether, this shows that x^* is an M-stationary point. $\square$

The question regarding the existence of KKT multipliers for the relaxed problems, as needed in the above convergence result, cannot be answered as satisfactory and quickly as for the foregoing approaches. To illustrate this, let us consider the following example.

Example 3.6 Consider the two-dimensional MPEC

$$\min x_1^2 + x_2^2 \quad \text{s.t.} \quad x_1 \geq 0,\, x_2 \geq 0,\, x_1 x_2 = 0. \tag{18}$$

Obviously, this MPEC satisfies MPEC-LICQ, and hence in particular MPEC-CPLD, at any feasible point. Now, choose $t_k := \frac{1}{k}$ and $x^k := (t_k, t_k)$ for $k \in \mathbb{N}$. Then $x^k \to x^* := (0,0)$ and geometric arguments and a quick calculation show that $\mathcal{T}_{X^{KDB}(t_k)}(x^k) = \{d \in \mathbb{R}^2 \mid d_1 d_2 \leq 0\} \neq \mathbb{R}^2 = \mathcal{L}_{X^{KDB}(t_k)}(x^k)$. Hence, ACQ is violated at x^k, in particular, all stronger concepts like CPLD, MFCQ or LICQ are also violated at x^k. On the other hand, it is easy to see that GCQ is satisfied.

Despite this counterexample, Kadrani et al. [21] were able to verify existence of KKT multipliers for the relaxed problem under the MPEC-LICQ assumption. The following result is a refinement of their observation and partly motivated by Example 3.6.

Theorem 3.7 *Let x^* be feasible for* (1) *such that MPEC-LICQ holds at x^*. Then there exists $\bar{t} > 0$ and a neighborhood N of x^* such that (standard) GCQ for $R^{KDB}(t)$ is fulfilled at all $\hat{x} \in N \cap X^{KDB}(t)$ for all $t \in (0, \bar{t})$.*

Proof. Again, we skip the standard constraints from the proof without loss of generality.

Let $t > 0$ and $\hat{x} \in X^{KDB}(t)$. Furthermore, let $I \subseteq I_\Phi^{00}(\hat{x}, t)$ and put $\bar{I} := I_\Phi^{00}(\hat{x}, t) \setminus I$. Herewith, define the program $NLP(I)$

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & G_i(x) + t \geq 0 \quad (i = 1, \dots, l), \\ & H_i(x) + t \geq 0 \quad (i = 1, \dots, l), \\ & G_i(x) - t \leq 0 \quad (i \in I_\Phi^{0+}(\hat{x}, t) \cup I), \\ & G_i(x) - t \geq 0 \quad (i \in I_\Phi^{0-}(\hat{x}, t) \cup \bar{I}), \\ & H_i(x) - t \leq 0 \quad (i \in I_\Phi^{+0}(\hat{x}, t) \cup \bar{I}), \\ & H_i(x) - t \geq 0 \quad (i \in I_\Phi^{-0}(\hat{x}, t) \cup I), \end{array}$$

and denote its feasible set by $\hat{X}(I)$. Then we have $\hat{x} \in \hat{X}(I)$ and, locally around $\hat{x}$, we have $\hat{X}(I) \subseteq X^{KDB}(t)$. We now claim that

$$\mathcal{T}_{X^{KDB}(t)}(\hat{x}) = \bigcup_{I \subseteq I_\Phi^{00}(\hat{x}, t)} \mathcal{T}_{\hat{X}(I)}(\hat{x}). \tag{19}$$

The $\supseteq$-inclusion is obvious. For the converse direction let $d \in \mathcal{T}_{X^{KDB}(t)}(\hat{x})$, i.e., there exists $\{x^k\} \subseteq X^{KDB}(t)$ with $x^k \to \hat{x}$ and $\{t_k\} \downarrow 0$ such that $\frac{x^k - \hat{x}}{t_k} \to d$. By continuity, for k sufficiently large, we have

$$\begin{array}{ll} G_i(x^k) - t \leq 0 & (i \in I_\Phi^{0+}(\hat{x}, t)), \\ G_i(x^k) - t \geq 0 & (i \in I_\Phi^{0-}(\hat{x}, t)), \\ H_i(x^k) - t \leq 0 & (i \in I_\Phi^{+0}(\hat{x}, t)), \\ H_i(x^k) - t \geq 0 & (i \in I_\Phi^{-0}(\hat{x}, t)), \end{array}$$

since $x^k \in X^{KDB}(t)$. Moreover, we also have $H_i(x^k) + t \geq 0, G_i(x^k) + t \geq 0$ $(i = 1, \dots, l)$ anyway. Due to the fact that $I_\Phi^{00}(\hat{x}, t)$ is finite, there exists an infinite subset $K \subseteq \mathbb{N}$ and $I \subseteq I_\Phi^{00}(\hat{x}, t)$ such that

$$\begin{array}{ll} G_i(x^k) - t \leq 0 & (i \in I_\Phi^{0+}(\hat{x}, t) \cup I), \\ G_i(x^k) - t \geq 0 & (i \in I_\Phi^{0-}(\hat{x}, t) \cup \bar{I}), \\ H_i(x^k) - t \leq 0 & (i \in I_\Phi^{+0}(\hat{x}, t) \cup \bar{I}), \\ H_i(x^k) - t \geq 0 & (i \in I_\Phi^{-0}(\hat{x}, t) \cup I) \end{array}$$

for all $k \in K$. Therefore, $\{x^k\}_K \subseteq \hat{X}(I)$ and hence, $d \in \mathcal{T}_{\hat{X}(I)}(\hat{x})$, which gives the desired inclusion.

Now, for an arbitrary $I \subseteq I_\Phi^{00}(\hat{x}, t)$, the active gradients of $NLP(I)$ at $\hat{x}$ are

$$\begin{array}{ll} \nabla G_i(\hat{x}) & (i \in I_G(\hat{x}, t) \subseteq I_{00} \cup I_{0+}), \\ \nabla H_i(\hat{x}) & (i \in I_H(\hat{x}, t) \subseteq I_{00} \cup I_{+0}), \\ \nabla G_i(\hat{x}) & (i \in I_\Phi^{0+}(\hat{x}, t) \cup I \subseteq I_{00} \cup I_{0+}), \\ \nabla G_i(\hat{x}) & (i \in I_\Phi^{0-}(\hat{x}, t) \cup \bar{I} \subseteq I_{00} \cup I_{0+}), \\ \nabla H_i(\hat{x}) & (i \in I_\Phi^{+0}(\hat{x}, t) \cup \bar{I} \subseteq I_{00} \cup I_{+0}), \\ \nabla H_i(\hat{x}) & (i \in I_\Phi^{-0}(\hat{x}, t) \cup I \subseteq I_{00} \cup I_{+0}), \end{array}$$

where for the index set inclusions, in particular, we exploit the fact that $I_\Phi^{00}(\hat{x}, t) \subseteq I_{00}$.

The above gradients are, in view of MPEC-LICQ at x^*, linearly independent if $\hat{x}$ is sufficiently close to x^*, i.e., LICQ and hence ACQ holds for $NLP(I)$ at $\hat{x}$. This means that $\mathcal{T}_{\hat{X}(I)}(\hat{x}) = \mathcal{L}_{\hat{X}(I)}(\hat{x})$ and in view of (19) and invoking [4, Th. 3.1.9], this yields

$$\mathcal{T}_{X^{KDB}(t)}(\hat{x})^\circ = \bigcap_{I \subseteq I_\Phi^{00}(\hat{x}, t)} \mathcal{L}_{\hat{X}(I)}(\hat{x})^\circ, \tag{20}$$

where, by means of [4, Th. 3.2.2], we have

$$\begin{aligned} \mathcal{L}_{\hat{X}(I)}(\hat{x})^\circ = \{v \in \mathbb{R}^n \mid \exists_{\alpha,\beta,\gamma,\delta,\epsilon,\rho \geq 0} : v \ = \ & - \sum_{i \in I_G(\hat{x},t)} \alpha_i \nabla G_i(\hat{x}) - \sum_{i \in I_H(\hat{x},t)} \beta_i \nabla H_i(\hat{x}) \\ & + \sum_{i \in I_\Phi^{0+}(\hat{x},t) \cup I} \gamma_i \nabla G_i(\hat{x}) - \sum_{i \in I_\Phi^{0-}(\hat{x},t) \cup \bar{I}} \delta_i \nabla G_i(\hat{x}) \\ & + \sum_{i \in I_\Phi^{+0}(\hat{x},t) \cup \bar{I}} \epsilon_i \nabla H_i(\hat{x}) - \sum_{i \in I_\Phi^{-0}(\hat{x},t) \cup I} \rho_i \nabla H_i(\hat{x})\}. \end{aligned}$$

In order to verify GCQ for $R^{KDB}(t)$ at $\hat{x}$, i.e., $\mathcal{T}_{X^{KDB}(t)}(\hat{x})^\circ \subseteq \mathcal{L}_{X^{KDB}(t)}(\hat{x})^\circ$, let $v \in \mathcal{T}_{X^{KDB}(t)}(\hat{x})^\circ$. Using (20), it then follows that, for some $I \subseteq I_\Phi^{00}(\hat{x}, t)$, we have both $d \in \mathcal{L}_{\hat{X}(I)}(\hat{x})^\circ$ and $d \in \mathcal{L}_{\hat{X}(\bar{I})}(\hat{x})^\circ$. Then exploiting the representation above and the linear independence of the occuring gradients, it is quickly argued that the multipliers with indices in I and $\bar{I}$ must vanish and hence, v can be expressed as

$$\begin{aligned} v \ = \ & - \sum_{i \in I_G(\hat{x},t)} \alpha_i \nabla G_i(\hat{x}) - \sum_{i \in I_H(\hat{x},t)} \beta_i \nabla H_i(\hat{x}) + \sum_{i \in I_\Phi^{0+}(\hat{x},t)} \gamma_i \nabla G_i(\hat{x}) \\ & - \sum_{i \in I_\Phi^{0-}(\hat{x},t)} \delta_i \nabla G_i(\hat{x}) + \sum_{i \in I_\Phi^{+0}(\hat{x},t)} \epsilon_i \nabla H_i(\hat{x}) - \sum_{i \in I_\Phi^{-0}(\hat{x},t)} \rho_i \nabla H_i(\hat{x}) \end{aligned}$$

for some $\alpha, \beta, \gamma, \delta, \epsilon, \rho \geq 0$, and this means that $v \in \mathcal{L}_{X^{KDB}(t)}(\hat{x})^\circ$, again by [4, Th. 3.2.2]. This concludes the proof. □

The following result shows that $R^{KDB}(t)$ satisfies stronger constraint qualifications in all points where $I_\Phi^{00}(x, t) = \emptyset$ holds.

Theorem 3.8 *Let x^* be feasible for the MPEC* (1) *such that MPEC-CPLD (MPEC-LICQ) holds at x^*. Then there is a $\bar{t} > 0$ and a neighborhood N of x^* such that the following holds for all $t \in (0, \bar{t}]$: If $x \in N \cap X^{KDB}(t)$ with $I_\Phi^{00}(x, t) = \emptyset$ then standard CPLD (LICQ) for $R^{KDB}(t)$ holds at x.*

Proof. Note that we only need to prove the CPLD part, since the assertion on LICQ follows from [21, Th. 2.4].

Now, suppose our assertion is false. Then there exist sequences $\{t_k\} \downarrow 0$ and $\{x^k\} \subseteq X^{KDB}(t_k)$ with $x^k \to x^*$ and $I_\Phi^{00}(x, t) = \emptyset$ such that CPLD for $R^{KDB}(t_k)$ is violated at

x^k. This yields subsets $I_1^k \subseteq I_g(x^k), I_2^k \subseteq \{1,\dots,p\}, I_3^k \subseteq I_G(x^k,t_k), I_4^k \subseteq I_H(x^k,t_k), I_5^k \subseteq I_\Phi^{0*}(x^k,t_k), I_6^k \subseteq I_\Phi^{*0}(x^k,t_k)$ such that the gradients

$$\{\nabla h_i(x) \mid i \in I_2^k\} \quad \cup \quad \Big\{\{\nabla g_i(x) \mid i \in I_1^k\} \cup \{-\nabla G_i(x) \mid i \in I_3^k\} \cup \{-\nabla H_i(x) \mid i \in I_4^k\} \\ \cup \{(H_i(x)-t_k)\nabla G_i(x) \mid i \in I_5^k\} \cup \{(G_i(x)-t_k)\nabla H_i(x) \mid i \in I_6^k\}\Big\}$$

are positive-linearly dependent at $x = x^k$, but linearly independent in x arbitrary close to x^k. Moreover, by a finiteness argument, we can assume without loss of generality that $I_i^k = I_i$ for $i = 1,\dots,6$ and all $k \in \mathbb{N}$. Then it is easy to see that $I_1 \subseteq I_g, I_3 \cup I_5 \subseteq I_{00} \cup I_{0+}$, and $I_4 \cup I_6 \subseteq I_{00} \cup I_{+0}$. The positive-linear dependence of the above gradients at x^k immediately implies the positive-linear dependence of the gradients

$$\{\nabla g_i(x) \mid i \in I_1\} \cup \Big\{\{\nabla h_i(x) \mid i \in I_2\} \cup \{\nabla G_i(x) \mid i \in I_3 \cup I_5\} \cup \{\nabla H_i(x) \mid i \in I_4 \cup I_6\}\Big\}$$

at $x = x^k$. Due to the violation of CPLD at x^k this yields a sequence $\{y^k\} \to x^*$ such that these gradients are linearly independent at $x = y^k$. If they were positive-linearly independent at x^*, by [29, Theorem 2.2], they would remain positive-linearly independent nearby, which contradicts the existence of $\{x^k\}$. On the other hand, if they were positive-linearly dependent, by MPEC-CPLD, they would remain linearly dependent in a whole neighborhood, which contradicts the existence of $\{y^k\}$. This concludes the proof. □

3.4 The Local Relaxation Scheme by Steffensen and Ulbrich

The scheme introduced by Steffensen and Ulbrich in [34] is based on the relaxation (cf. Figure 4)

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

with

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

where

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

and $\theta : [-1,1] \to \mathbb{R}$ is a function satisfying:

(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 $x \in (-1, 1)$.

Let $X^{SU}(t)$ denote the feasible set of $R^{SU}(t)$.

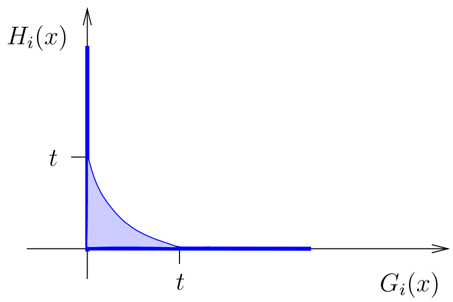


Figure 4: Geometric interpretation of the relaxation method by Steffensen and Ulbrich [34]

The original convergence result from [34] states that, given a convergent sequence $x^k \to x^*$ of stationary points of the relaxed problems $R^{SU}(t_k)$ with $\{t_k\} \downarrow 0$, then the limit point x^* is C-stationary provided that MPEC-CRCQ holds at x^*. Very recently, it has been shown that the same statement holds under the weaker MPEC-CPLD assumption, see [16], which is therefore the result that we restate here. Note that the assertion holds, in particular, under MPEC-MFCQ.

Theorem 3.9 *Let $\{t_k\} \downarrow 0$ and let x^k be a stationary point of $R^{SU}(t_k)$ with $x^k \to x^*$ such that MPEC-CPLD holds in x^*. Then x^* is a C-stationary point of* (1).

The local relaxation scheme discussed in this section has another advantage, namely that it might not be necessary for the sequence $\{t_k\}$ to go down to zero under suitable assumptions, in particular, when $G_i(x^*) + H_i(x^*) > 0$ holds for all $i = 1, \dots, l$. This follows immediately from the observation that, in this case, the feasible sets of the MPEC itself and of the relaxed problem $R^{SU}(t)$ coincide locally.

On the other hand, the question under which assumptions one may expect to get multipliers for the relaxed problem has not been discussed thus far. To this end, let us first consider a simple example.

Example 3.10 Consider again the MPEC from Example 3.6. Given a sequence $\{t_k\} \downarrow 0$, we define $\{x^k\}$ by $x^k := (t_k, 0)$. Then the active gradients at x^k are $-\binom{0}{1}, \nabla\Phi(t_k, 0; t_k) =$

$\binom{0}{2}$, which are obviously positive-linearly dependent. On the other hand, for $\varepsilon > 0$ sufficiently small, the above gradients evaluated at $x_\varepsilon^k := (t_k - \varepsilon, 0)$ become $-\binom{0}{1}, \binom{1-\theta'(\frac{t_k-\varepsilon}{t_k})}{1+\theta'(\frac{t_k-\varepsilon}{t_k})}$, which are obviously linearly independent. Hence, CPLD is violated at x^k for all k, although MPEC-LICQ holds at $x^* = (0,0)$. However, it is easy to see that ACQ is fulfilled.

In order to prove our result on constraint qualifications, some index sets need to be defined. For these purposes, let $t > 0$ and $\hat{x} \in X^{SU}(t)$. Then we put

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

Theorem 3.11 *Let x^* be feasible for* (1) *such that MPEC-LICQ holds at x^*. Then there exists a neighborhood N of x^* and $\bar{t} > 0$ such that for all $t \in (0, \bar{t})$ and $\hat{x} \in X^{SU}(t) \cap N$ (standard) ACQ for $R^{SU}(t)$ is satisfied at $\hat{x}$.*

Proof. Note that, again, we skip the standard constraints from the proof for clarity's sake.

Now, if $I_{0+} \cup I_{+0} \neq \emptyset$, first put $\bar{t} := \frac{1}{2} \min\{G_i(x^*)\,(i \in I_{+0}), H_i(x^*)\,(i \in I_{0+})\}$. Then, in particular, one has $\bar{t} > 0$. Otherwise choose $\bar{t} > 0$ arbitrarily. Now, let $t \in (0, \bar{t})$ and $\hat{x} \in X^{SU}(t)$. Then we define the program $NLP(\hat{x})$ by

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & G_i(x) \geq 0 & (i \notin I_\Phi(\hat{x}, t)), \\ & H_i(x) \geq 0 & (i \notin I_\Phi(\hat{x}, t)), \\ & \Phi(G_i(x), H_i(x); t) = 0 & (i \in I_\Phi(\hat{x}, t) \cap \{I_G(\hat{x}) \cup I_H(\hat{x})\}), \\ & \Phi(G_i(x), H_i(x); t) \leq 0 & (i \notin I_\Phi(\hat{x}, t) \cap \{I_G(\hat{x}) \cup I_H(\hat{x})\}), \end{array}$$

and denote its feasible region by $\hat{X}$. Then, clearly, $\hat{x} \in \hat{X}$. Moreover, using [16, Lem. 4.5/4.6], the gradients for the active constraints of $NLP(\hat{x})$ at $\hat{x}$ read to

$$\begin{array}{ll} \nabla G_i(\hat{x}) & (i \in I_G(\hat{x}) \setminus I_\Phi(\hat{x}, t)), \\ \nabla H_i(\hat{x}) & (i \in I_H(\hat{x}) \setminus I_\Phi(\hat{x}, t)), \\ 2\nabla G_i(\hat{x}) & (i \in I_G(\hat{x}) \cap I_\Phi(\hat{x}, t)), \\ 2\nabla H_i(\hat{x}) & (i \in I_H(\hat{x}) \cap I_\Phi(\hat{x}, t)), \\ \alpha_i \nabla G_i(\hat{x}) + \beta_i \nabla H_i(\hat{x}) & (i \in I_\Phi(\hat{x}, t) \setminus \{I_G(\hat{x}) \cup I_H(\hat{x})\}), \end{array} \tag{21}$$

where $\alpha_i = 1 - \theta'(\frac{G_i(\hat{x}) - H_i(\hat{x})}{t}), \beta_i = 1 + \theta'(\frac{G_i(\hat{x}) - H_i(\hat{x})}{t})$. Now, for $\hat{x}$ sufficiently close to x^*, we have the inclusions

$$I_G(\hat{x}) \subseteq I_{00} \cup I_{0+}, \quad I_H(\hat{x}) \subseteq I_{00} \cup I_{+0},$$

and by the choice of $\bar{t}$, we also get

$$I_\Phi(\hat{x}, t) \setminus \{I_G(\hat{x}) \cup I_H(\hat{x})\} \subseteq I_{00}.$$

Hence, in view of MPEC-LICQ at x^*, the gradients in (21) are linearly independent for $\hat{x}$ sufficiently close to x^*, and thus, LICQ and in particular ACQ for $NLP(\hat{x})$ holds at $\hat{x}$.

Moreover, we have

$$\begin{aligned}\mathcal{L}_{\hat{X}}(\hat{x}) &= \Big\{ d \in \mathbb{R}^n \mid \begin{array}{rl} \nabla G_i(\hat{x})^T d \geq 0 & (i \in I_G(\hat{x}) \setminus I_\Phi(\hat{x},t)), \\ \nabla H_i(\hat{x})^T d \geq 0 & (i \in I_H(\hat{x}) \setminus I_\Phi(\hat{x},t)), \\ \nabla G_i(\hat{x})^T d = 0 & (i \in I_G(\hat{x}) \cap I_\Phi(\hat{x},t)), \\ \nabla H_i(\hat{x})^T d = 0 & (i \in I_H(\hat{x}) \cap I_\Phi(\hat{x},t)), \\ \nabla(\Phi(G_i(\hat{x}), H_i(\hat{x});t))^T d \leq 0 & (i \in I_\Phi(\hat{x},t) \setminus \{I_G(\hat{x}) \cup I_H(\hat{x})\}) \end{array} \Big\} \\ &= \mathcal{L}_{X^{SU}(t)}(\hat{x}),\end{aligned} \tag{22}$$

where the last equality can easily be verified by direct calculation. We now want to show that, locally around $\hat{x}$, we have $\hat{X} \subseteq X^{SU}(t)$.

For these purposes, it remains to see that, for $x \in \hat{X}$ sufficiently close to $\hat{x}$, we have

$$G_i(x) \geq 0 \ (i \in I_\Phi(\hat{x},t)) \quad \text{and} \quad H_i(x) \geq 0 \ (i \in I_\Phi(\hat{x},t)).$$

To this end, consider first the case of $i \in I_\Phi(\hat{x},t) \cap \{I_G(\hat{x}) \cup I_H(\hat{x})\}$. Then, in particular, $\Phi(G_i(x), H_i(x);t) = 0$, which in view of [16, Lemma 4.6] gives $G_i(x), H_i(x) \geq 0$. If otherwise $i \notin I_G(\hat{x}) \cup I_H(\hat{x})$, we get $G_i(x), H_i(x) > 0$ by continuity.

The local inclusion $\hat{X} \subseteq X(t)$ yields $\mathcal{T}_{\hat{X}}(\hat{x}) \subseteq \mathcal{T}_{X^{SU}(t)}(\hat{x})$ and hence, by ACQ for $NLP(\hat{x})$ at $\hat{x}$ and (22), we have

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

which gives the assertion. □

The following result shows that a stronger constraint qualification holds at all points of $X^{SU}(t)$ where the local relaxation is active.

Theorem 3.12 *Let x^* be feasible for* (1) *such that MPEC-LICQ holds at x^*. Then there exists a neighborhood N of x^* such that the following holds: If $x \in N \cap X^{SU}(t)$ with $I_\Phi(x,t) \cap \{I_G(x) \cup I_H(x)\} = \emptyset$ then standard LICQ holds for $R^{SU}(t)$ at x.*

Proof. Let $x \in X^{SU}(t)$. Then, if $I_\Phi(x,t) \cap \{I_G(x) \cup I_H(x)\} = \emptyset$, the active gradients for $R^{SU}(t)$ at x are,

$$\begin{array}{rl} \nabla g_i(x) & (i \in I_g(x)), \\ \nabla h_i(x) & (i = 1, \dots, p), \\ \nabla G_i(x) & (i \in I_G(x)), \\ \nabla H_i(x) & (i \in I_H(x)), \\ \alpha_i \nabla G_i(x) + \beta_i \nabla H_i(x) & (i \in I_\Phi(x,t) \setminus \{I_G(x) \cup I_H(x)\}), \end{array}$$

where $\alpha_i = 1 - \theta'(\frac{G_i(\hat{x})-H_i(\hat{x})}{t}), \beta_i = 1 + \theta'(\frac{G_i(\hat{x})-H_i(\hat{x})}{t})$. Moreover, cf. also the proof of Theorem 3.11, for x sufficiently close to x^*, we have the inclusions

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

and in view of MPEC-LICQ the active gradients from above are linearly independent, i.e., LICQ holds at x. □

3.5 The Relaxation Scheme by Kanzow and Schwartz

The relaxation scheme established by Kanzow and Schwartz in [22] is given by (see Figure 5)

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

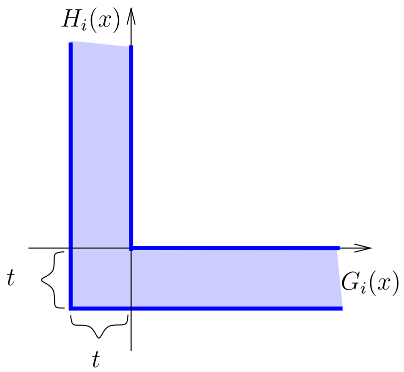


Figure 5: Geometric interpretation of the relaxation method by Kanzow and Schwartz [22]

where $\Phi : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ is defined by $\Phi(x_1, x_2; t) := \phi(x_1 - t, x_2 - t)$ and $\phi : \mathbb{R}^2 \to \mathbb{R}$ is given by

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

Let $X^{KS}(t)$ be the feasible set of $R^{KS}(t)$. Given $x \in X^{KS}(t)$, we put $I_\Phi^{00}(x;t) := \{i \mid G_i(x) - t = H_i(x) - t = 0\}$. The following is the main convergence result for this relaxation, see [22] for a proof.

Theorem 3.13 *Let $\{t_k\} \downarrow 0$ and x^k be a stationary point of NLP(t_k) with $x^k \to x^*$ such that MPEC-CPLD holds in x^*. Then x^* is an M-stationary point of the MPEC* (1)*.*

Existence of multipliers can also be guaranteed under suitable assumptions.

Theorem 3.14 *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 N of x^* such that GCQ holds for $R^{KS}(t)$ at all $\hat{x} \in N \cap X^{KS}(t)$ for all $t \in (0, \bar{t})$.*

Theorem 3.15 *Let x^* be feasible for the MPEC* (1) *such that MPEC-CPLD (MPEC-LICQ) holds at x^*. Then there is a $\bar{t} > 0$ and a neighborhood N of x^* such that the following holds for all $t \in (0, \bar{t}]$: If $x \in U(x^*) \cap X^{KS}(t)$ with $I^{00}_{\Phi}(x; t) = \emptyset$, then standard CPLD (LICQ) for $R^{KS}(t)$) holds at x.*

In Table 1, we try to summarize the results of this section in a very concise way. The columns contain the five relaxation schemes discussed here. The first two lines then state under which MPEC constraint qualification a limit point of a sequence generated by one of these methods is either C- or M-stationary. The second part of the table indicates under which MPEC constraint qualification the corresponding regularized problem satisfies one of the standard NLP constraint qualifications. Of course, this part only holds locally around a given feasible point x^* of the MPEC.

Relaxation	Scholtes	Lin–Fukush.	Kadrani et al.	Steff.–Ulbrich	K.–Schwartz
	stationary point results				
Assuming limit pts. are	MPEC-MFCQ C–stationary	MPEC-MFCQ C–stationary	MPEC-CPLD M–stationary	MPEC-CPLD C–stationary	MPEC-CPLD M–stationary
	existence of Lagrange multipliers				
Assuming NLP(t) satisf.	MPEC-MFCQ MFCQ	MPEC-MFCQ MFCQ	MPEC-LICQ GCQ	MPEC-LICQ ACQ	MPEC-LICQ GCQ

Table 1: Summary of results regarding stationary points and constraint qualifications for the different relaxation methods

Note, however, that this second part does not cover the complete story. To this end, let us first note that MPEC-LICQ implies that the MPEC itself satisfies standard GCQ, cf. [11]. Therefore, when simply looking at the table, it seems that the feasible sets of some of the relaxation methods do not have better properties than the underlying MPEC, hence one might wonder why using such a regularization. In fact, GCQ or the slightly stronger ACQ are relatively weak conditions which, however, guarantee the existence of Lagrange multipliers at a local minimum. The main difference is that standard MFCQ (hence also standard LICQ) is violated at *any* feasible point of the MPEC itself, while the corresponding results for the three regularization methods by Kadrani et al. [21], Steffensen and Ulbrich [34], and Kanzow and Schwartz [22] typically satisfy LICQ (hence also MFCQ) in many points under the MPEC-LICQ condition.

Let us discuss this point is more detail. The two relaxation methods by Scholtes [33] and Lin and Fukushima [25], besides being only convergent to C-stationary points, have no problems regarding constraint qualifications: MPEC-LICQ implies LICQ for the corresponding regularized problems, as shown in the original references [33, 25], and MPEC-MFCQ also implies MFCQ, as shown in this section for the Scholtes-relaxation and in [16] for the Lin-Fukushima-relaxation. Moreover, it also seems that MPEC-CRCQ (MPEC-CPLD) implies CRCQ (CPLD) for the corresponding relaxed problem, i.e., basically any MPEC constraint qualification implies that the corresponding standard CQ holds for the relaxed problem (locally, of course).

The situation is completely different with the other three relaxation schemes. These other three schemes have stronger convergence properties than the first two methods, more precisely, the two relaxation schemes by Kadrani et al. [21] and Kanzow and Schwartz [22] converge to M-stationary points, which is a much stronger property than C-stationarity, whereas the local regularization approach by Steffensen and Ulbrich [34] only converge to C-stationary points, but has a nice finite termination property in the sense that it is not always necessary that the relaxation parameter t has to be driven down to zero. On the other hand, our analysis and the corresponding (counter-) examples show that the relaxed problems of any of these three methods usually do not inherit the corresponding standard CQ from an MPEC CQ.

In fact, it is easy to see that the relaxed problem by Steffensen and Ulbrich [34] not only violates LICQ, but also CRCQ and CPLD, whereas ACQ (hence GCQ) is satisfied under MPEC-LICQ. Similarly, the relaxed problems by Kadrani et al. [21] and Kanzow and Schwartz [22] do not even satisfy ACQ, whereas GCQ holds under MPEC-LICQ. Hence, from this point of view, it seems that the Steffensen-Ulbrich regularization is slightly better than the other two relaxations. However, also this is not true in general since, speaking in the $(G_i(x), H_i(x))$-space, both the Kadrani et al.- and the Kanzow-Schwartz-regularization satisfy standard LICQ in all points except for one (locally and assuming MPEC-LICQ, of course), whereas the Steffensen-Ulbrich relaxation violates standard LICQ in many points, namely in all points on the G_i- and H_i-axes where the feasible set of the MPEC is not changed by the local relaxation.

4 Numerical Comparison

We implemented all five methods in MATLAB 7.10.0 and ran some tests using the MacMPEC collection [23]. The basic algorithm is Algorithm 1, where the maximum violation of all constraints

$$\text{maxVio}(x_{opt}) = \max\{\max\{0, g(x_{opt})\}, |h(x_{opt})|, |\min\{G(x_{opt}), H(x_{opt})\}|\} \qquad (23)$$

is used to measure the feasibility of the final iterate x_{opt}.

However, before we delve into the numerical details, let us clarify the aim of this section. So far, we have discussed the different theoretical properties of these five relaxation methods. Now, we want to find out what differences there are in the numerical behaviour.

Algorithm 1 Basic relaxation algorithm $(x_0, t_0, \sigma, t_{\min}, \varepsilon)$

Require: a starting vector x^0, an initial relaxation parameter t_0, and parameters $\sigma \in (0,1)$, $t_{\min} > 0$, and $\varepsilon > 0$.

Set $k := 0$.
while $t_k \geq t_{\min}$ **do**
Find an approximate solution x^{k+1} of the relaxed problem $R(t_k)$. To solve $R(t_k)$, use x^k as starting vector.
Let $t_{k+1} \leftarrow \sigma \cdot t_k$ and $k \leftarrow k+1$.
end while

Return: the final iterate $x_{opt} := x^k$, the corresponding function value $f(x_{opt})$, and the maximum constraint violation $\text{maxVio}(x_{opt})$.

Therefore, we tried to implement all five methods as similar as possible to ensure that different numerical results are caused by the different properties of the relaxations and not by algorithmic differences. As a consequence, we did not optimize our implementation individually for every relaxation, i.e. it is possible to obtain better results by tailoring the algorithms to the characteristics of the relaxations. For example, Steffensen and Ulbrich [34] and Kadrani, Dussault, and Benchakroun [21] proposed numerical approaches that deal with the specific characteristics of their relaxations.

All relaxations except for the one by Kadrani et al. have the property that the following inclusion holds for all $0 \leq t_1 < t_2$

$$X(t_1) \subset X(t_2),$$

where $X(t)$ is the feasible area of the relaxed problem $R(t)$ and $X = X(0)$ is the feasible area of the MPEC (1). This can be used in the numerical implementation in the following way: If a relaxed problem $R(t_k)$ is infeasible, the MPEC is infeasible, too, and we can terminate the algorithm immediately. We can also terminate the algorithm early if the solution x^{k+1} of an iteration k is feasible for the MPEC because in this case, x^{k+1} is also a solution of the MPEC itself. Finally, if the solution x^{k+1} also is feasible for $R(t_{k+1})$, is also is a solution of $R(t_{k+1})$. Thus, we can skip the next iteration and reduce the relaxation parameter further until x^{k+1} is not feasible for the next iteration anymore. These changes are incorporated into Algorithm 2, where feasibility of the iterate x^k for the original MPEC is measured by the violation of the complementarity constraints

$$\text{compVio}(x^k) = \| \min\{G(x^k), H(x^k)\}\|_\infty.$$

Note, that the standard constraints $g(x) \leq 0$ and $h(x) = 0$ are part of the relaxed problems $R(t)$ and therefore do not need to be checked here.

We used the parameters $t_{\min} = 10^{-15}$ and $\varepsilon = 10^{-6}$ for all relaxations and the TOMLAB 7.4.0 solver `snopt` to solve the relaxed problems $R(t_k)$. The remaining parameters and the

Algorithm 2 Improved relaxation algorithm $(x_0, t_0, \sigma, t_{\min}, \varepsilon)$

Require: a starting vector x^0, an initial relaxation parameter t_0, and parameters $\sigma \in (0,1)$, $t_{\min} > 0$, and $\varepsilon > 0$.

Set $k := 0$.
while ($t_k \geq t_{\min}$ and compVio$(x^k) > \varepsilon$) or $k = 0$ **do**
  Find an approximate solution x^{k+1} of the relaxed problem $R(t_k)$. To solve $R(t_k)$, use x^k as starting vector.
  If $R(t_k)$ is infeasible, terminate the algorithm.
  Let $t_{k+1} \leftarrow \max_{l=1,2,3,\ldots}\{\sigma^l \cdot t_k \mid x^{k+1} \notin X(\sigma^l \cdot t_k) \text{ and } \sigma^l \cdot t \geq t_{\min}\}$ and $k \leftarrow k+1$.
end while

Return: the final iterate $x_{opt} := x^k$, the corresponding function value $f(x_{opt})$, and the maximum constraint violation maxVio(x_{opt}).

used algorithm are given in Table 2. Note that we cannot use the improved Algorithm 2 for the relaxation method by Kadrani et al. since the feasible area of the MPEC (1) is not included in the feasible area of the relaxed problems used in this method.

Relaxation	Scholtes	Lin–Fukush.	Kadrani et al.	Steff.–Ulbrich	K.–Schwartz
Algorithm	Algorithm 2	Algorithm 2	Algorithm 1	Algorithm 2	Algorithm 2
t_0	1	1	1	$\frac{2\pi}{\pi-2}$	1
σ	0.0001	0.01	0.01	0.01	0.01

Table 2: Parameters and algorithms used for the different relaxation methods

The parameters t_0 and σ are chosen such that in the k-th iteration (assume that no iterations are skipped) the point $(0.01^k, 0.01^k)$ lies on the boundary of the relaxed feasible area, for example $\Phi_i^{SU}(0.01^k, 0.01^k; t_k) = 0$ for all $k = 0, 1, \ldots$ with $t_k = t_0\sigma^k$. Thus, it is guaranteed that the relaxed area shrinks with the same speed for all relaxation methods.

To illustrate the different behaviour of the five relaxation methods, we chose 126 problems from the MacMPEC collection [23]. The other problems were discarded partly due to their size or form and partly because errors occured during the evaluation of the objective function or the constraints by AMPL. Communication between AMPL and MATLAB is achieved using the mex function `amplfunc` [27]. To present the results, we use performance profiles as introduced by Dolan and Moré in [9]. In Figure 6, the performance profile for the maximum violation of all constraints as defined in (23) is depicted.

It can be seen that the relaxation method by Steffensen and Ulbrich produces the smallest constraint violation, followed by the relaxation of Scholtes and the one of Kanzow and Schwartz. We consider a problem as solved if the maximum violation of all constraints is less than or equal to 10^{-6} (independent of the optimal function value found by the

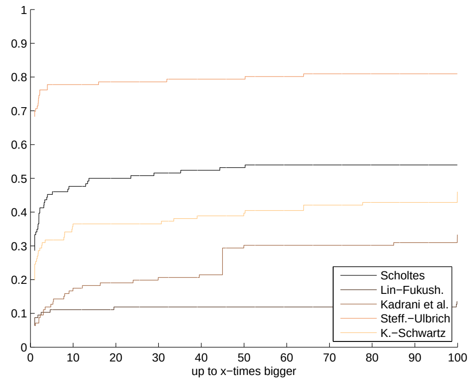


Figure 6: Comparison of constraint violation

corresponding method). According to this criterion, the relaxation method by Scholtes did not solve 12 problems, the relaxation method by Lin and Fukushima 62 problems, Kadrani et al. 19 problems. The relaxation method by Steffensen and Ulbrich failed to solve 18 problems and the method by Kanzow and Schwartz did not solve 23 problems. We would like to mention at this point that the number of unsolved problems depends on the chosen parameters. If a bigger σ is chosen, say $\sigma = 0.1$, the method by Kanzow and Schwartz solves most problems followed by Steffensen-Ulbrich and then Scholtes. However, the performance profiles do not change significantly and, all in all, more problems are solved with the parameters given in Table 2. Trouble had to be expected for at least some test problems: `design-cent-1` is known to be infeasible, `ralphmod` has an unbounded set of feasible solutions and `ex9.2.2, qpec2, ralph1`, and `scholtes4` do not have strongly stationary solutions, see for example [5, 30]. In fact, many of them are among the unsolved problems for all approaches.

We expected the relaxation method by Scholtes to be quite successful since it is the relaxation with the most regular subproblems. For the same reason, we are somewhat surprised by the results of the relaxation method by Steffensen and Ulbrich. Although the corresponding relaxed problems satisfy only very weak constraint qualifications, this method produces highly feasible solutions. This might be due to the fact that the relaxed feasible area is much smaller than for all other methods since it is only relaxed locally around the origin. However, in those cases where the maximum constraint violation is not less than or equal to 10^{-6}, it is mostly between 1 and 10, i.e. this small relaxed area sometimes leads to problems. This is very much in contrast to some of the other methods, e.g., for the majority of the 23 "unsolved" problems by the method of Kanzow and Schwartz, the feasibility is around 10^{-5}, and the iterates are usually close to the solution.

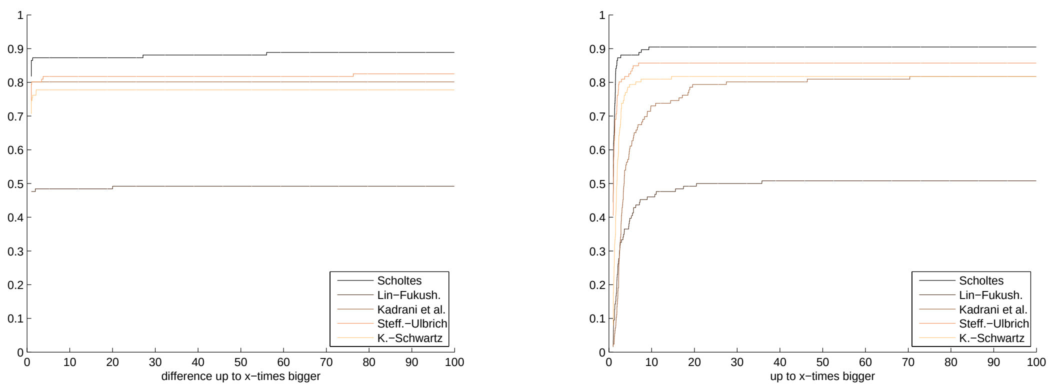
The two sided relaxation by Lin and Fukushima seems to cause serious numerical trouble. Here, the constraint violation in the unsolved problems covers the whole spectrum from 10^{-5} up to 10. This might be due to the fact that the functions $G_i(x)H_i(x) - t^2$ and

$(G_i(x) + t)(H_i(x) + t) - t^2$ nearly coincide for small relaxation parameters $t > 0$. Perhaps a different approach, where the relaxation parameters are individually adjusted such that for every pair of functions relaxing one complementarity constraint only one relaxation parameter tends to zero, would work better for this method.

In the following two performance profiles, we set the values corresponding to unsolved problems to $+\infty$. These two performance profiles compare the value of the objective function in the final iterate and the time needed for the calculation of this iterate. We plotted performance profiles comparing the number of objective function evaluations and gradient evaluations as well, but they look almost exactly like the one comparing the time. Thus, we do not include them here. A few words on the performance profile comparing the optimal function value: As the optimal function value is negative for some test problems, we have to normalize the corresponding data slightly different than Dolan and Moré. Let f_R^k be the optimal function value for test problem k found by the relaxation method R, $R \in \{S, LF, KDB, SU, KS\}$. We then define the normalized data for the method by Scholtes as

$$\bar{f}_S^k := \frac{f_S^k - \min\{f_R \mid R \in \{S, LF, KDB, SU, KS\}\}}{|\min\{f_R \mid R \in \{S, LF, KDB, SU, KS\}\}|}$$

and analogously for all other methods, i.e. we consider the difference to the best value found by any of the five methods normalized by the absolute value of this best value.



(a) Comparison of function value (b) Comparison of elapsed time

Figure 7: Comparison of function value and time

Note that the highest possible value for a relaxation method in Figure 7 is the percentage of solved problems, e.g. 90.5% for the method by Scholtes. Figure 7b indicates that the relaxation method by Scholtes needs the least time, followed by the one by Steffensen and Ulbrich. The relaxation method by Kanzow and Schwartz is slightly faster that the one by Kadrani et al. It had to be expected that the order here is about the same as the one in Figure 6 as we terminate the relaxation algorithm early if a solution feasible for the MPEC (1) is found. The only exception to this rule is the relaxation by Kadrani et al., see the discussion corresponding to Algorithm 2. If we take a look at Figure 7a, we see that the relaxation by Scholtes finds the best function values, again followed by Steffensen and Ulbrich, Kadrani et al., and Kanzow and Schwartz. However, these last three methods are very close to each other, and differences in the optimal function value might simply be explained by different feasibilities, i.e., a point less feasible for one approach is likely to have a better function value than a point that was computed by another method and is closer to the feasible set (or even feasible).

All in all, we have seen that the oldest and simplest relaxation, namely the one by Scholtes, still yields the best numerical results although most of the other methods have better theoretical properties. The relaxation method proposed by Lin and Fukushima is theoretically equivalent to the one by Scholtes and has the advantage of needing less constraints. However, it has serious numerical problems if one uses only one relaxation parameter. Thus, we would propose to combine this method with an active set strategy like to the one used by Demiguel et al. in [7]. The relaxation methods by Kadrani et al. and Kanzow and Schwartz have similar theoretical properties and also behave similarly when it comes to numerical results. The method by Kanzow and Schwartz is slightly faster and the final iterates have a smaller constraint violation, whereas the method by Kadrani et al. sometimes finds slightly smaller objective function values. As it had to be expected, both exhibit slow convergence for those test problems where it is known that the solutions are not strongly stationary. Finally, the relaxation method by Steffensen and Ulbrich works surprisingly well without special tuning of the solver as it was proposed in the original work [34] although the feasible area of the relaxed problems does not have a strictly feasible interior and most constraint qualifications are violated. However, in contrast to the other methods, a certain instability can be observed since the final iterates either have an extremely small maximum constraint violation or they are far away from being feasible.

5 Final Remarks

In this paper, we gave a theoretical and numerical comparison of five different relaxation schemes for the solution of mathematical programs with equilibirum constraints. In the theoretical part, we were able to improve a number of existing convergence results, and also added some completely new results regarding the satisfaction of standard constraint qualifications for the relaxed problems.

Despite some theoretical advantages of some of the newer relaxation schemes, the numerical comparison favours the oldest relaxation scheme due to Scholtes [33]. This is not completely surprising since it is the simplest regularization among all schemes investigated here. However, one should also take into account that the situation might be different when these solvers are applied to highly difficult MPECs where C-stationary points attract those methods which, in general, converge to C-stationary points only. A corresponding (and sufficiently large) test suite of such problems is currently not available and, therefore, special tests on these kind of problems are not included.

We believe that the method from [21] and, especially, the method from [22] will eventually outperform the other methods on these kind of difficult MPECs, but leave this point as one of our future research topics.

References

[1] 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.

[2] 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.

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

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

[5] H.Y. Benson, D.F. Shanno, and R.J. Vanderbei: *Interior point methods for nonconvex nonlinear programming: complementarity constraints.* Tech. Report ORFE-02-02, Operations Research and Financial Engineering, Princeton University, Princeton, NJ, 2002.

[6] 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.

[7] 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.

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

[9] E.D. Dolan and J.J. Moré: *Benchmarking optimization software with performance profiles.* Mathematical Programming Series A 91, 2002, pp. 201–213.

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

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

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

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

[16] T. Hoheisel, C. Kanzow, and A. Schwartz: *Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints.* Preprint 293, Institute of Mathematics, University of Würzburg, Würzburg, February 2010.

[17] T. Hoheisel, C. Kanzow, and A. Schwartz: *Improved convergence properties of the Lin-Fukushima-regularization method for mathematical programs with complementarity constraints.* Preprint 297, Institute of Mathematics, University of Würzburg, Würzburg, September 2010.

[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, 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.

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

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

[23] S. Leyffer: *MacMPEC: AMPL collection of MPECs*, www.mcs.anl.go/˜leyffer/MacMPEC, 2000

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

[27] www.netlib.org/ampl/solvers/examples/amplfunc.c

[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] 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.

[30] 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.

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

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

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

[35] O. Stein: *Lifting mathematical programs with complementarity constraints.* Mathematical Programming, to appear.

[36] J.J. Ye and D.L. Zhu: *Optimality conditions for bilevel programming problems.* Optimization 33, 1995, pp. 9–27.