

**THE PRICE OF INEXACTNESS:
CONVERGENCE PROPERTIES OF RELAXATION
METHODS FOR MATHEMATICAL PROGRAMS
WITH EQUILIBRIUM CONSTRAINTS REVISITED**

Christian Kanzow and Alexandra Schwartz

Preprint 315

March 2013

University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany

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

March 11, 2013

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 prominent class of specialized algorithms are the relaxation (or regularization) methods. The first relaxation 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. Among the most recent examples for such methods are the ones from Kadrani, Dussault, and Benchakroun [*SIAM Journal on Optimization* 20, 2009, pp. 78–103] and Kanzow and Schwartz [*SIAM Journal on Optimization*, to appear]. Surprisingly, although these recent methods have better theoretical properties than Scholtes’ relaxation, numerical comparisons show that this method is still among the fastest and most reliable ones, see for example Hoheisel et al. [*Mathematical Programming* 137, 2013, pp. 257–288]. To give a possible explanation for this, we consider the fact that, numerically, the regularized subproblems are not solved exactly. In this light, we analyze the convergence properties of a number of relaxation schemes and study the impact of inexactly solved subproblems on the kind of stationarity we can expect in a limit point.

Key Words: Mathematical programs with complementarity constraints; Mathematical programs with equilibrium constraints; Global convergence; KKT-points; Stationary points; Strong stationarity; M-stationarity; C-stationarity; Weak stationarity; Inexact relaxation methods; Inexact regularization methods.

1 Introduction

In this paper, we consider *mathematical programs with complementarity (or equilibrium) constraints*, MPECs for short. These are constrained optimization problems of the form

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

where $f, g_i, h_i, G_i, H_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are assumed to be continuously differentiable, and where the notation $0 \leq x \perp y \geq 0$ for two vectors $x, y \in \mathbb{R}^q$ is a shorthand for the conditions $x \geq 0, y \geq 0, x^T y = 0$. For several applications and theoretical issues of MPECs, we refer to the two monographs [25, 29] as well as the related book [7] on bilevel programming.

In principle, MPECs may be viewed as standard nonlinear optimization problems (NLPs), but they have a special structure since, apart from the usual equality and inequality constraints, they have the additional complementarity constraints given by the functions G_i, H_i which may equivalently be rewritten as

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

a formulation that has been exploited, e.g., in [12]. Whatever formulation is used for the complementarity constraints, however, these constraints cause some troubles both from a theoretical and a numerical point of view, especially because it is easy to see that most standard constraint qualifications are violated at any feasible point of the MPEC, cf. [25, 42].

On the other hand, the special structure of the complementarity constraints can also be exploited in order to develop a number of specialized algorithms for the solution of MPECs. In fact, during the last 15 years, a number of different methods have been suggested in order to deal with the inherent difficulty of an MPEC, including penalty, smoothing, interior-point, lifting, and relaxation methods. We refer the reader to [1, 2, 6, 8, 12, 17, 18, 19, 21, 23, 24, 31, 32, 34, 35, 36] and references therein for more details.

This manuscript deals with the class of relaxation (often also called regularization) methods for MPECs. In particular, the focus will be on the following relaxation methods:

- the global relaxation method by Scholtes [34]
- the smooth relaxation method by Lin and Fukushima [24]
- the local relaxation method by Steffensen and Ulbrich [35]
- the nonsmooth relaxation method by Kadrani et al. [19]
- the L-shaped relaxation method by the authors from [21].

More precisely, the method by Lin and Fukushima [24] is treated in a separate paper [22], hence the corresponding results will only be summarized here, whereas a detailed analysis and discussion for the other four methods are included in this paper. The basic idea of all these methods is to replace the original MPEC by a

sequence of potentially simpler NLPs depending on a certain parameter t such that, for $t \rightarrow 0$, the relaxed problems approach the original one. Algorithmically, one therefore tries to solve the sequence of NLPs by standard software, and hopes that this sequence converges to a solution of the MPEC for $t \rightarrow 0$.

Of course, it is usually not possible to “solve” the relaxed NLPs, hence most of the previous papers assume that a KKT-point can be found, and then consider the behaviour of the corresponding sequence of KKT-points. Under suitable assumptions, the first three methods from [34, 24, 35] show convergence to C-stationary points, whereas the last two methods from [19, 21] have the stronger and interesting property that they converge to M-stationary points (for precise definitions of C-, M- and related stationarities, we refer to Section 2).

However, in the majority of the papers on relaxation methods, the authors suppose that they can compute the KKT-points of the relaxed NLPs exactly. This assumption is unrealistic from a numerical point of view and does not coincide with the usual termination criteria used in standard NLP software. The best one can hope for is that a suitable solver is able to find an approximate KKT-point of the relaxed NLPs. But then the question is whether the above-mentioned convergence results still hold if we consider a sequence of approximate KKT-points only. The answer is obviously negative, since we lose the sign structure of certain multipliers if we do not consider exact KKT-points, hence one has to expect weaker convergence results.

More precisely, this paper shows that, without any additional assumptions, three of the above relaxation methods converge to weakly stationary points only, whereas two of them converge to C-stationary points. The result is, however, quite surprising, since, in particular, the two best methods from [19, 21], where we obtain M-stationarity as a limit of exact KKT-points, converge to weakly stationary points only, whereas the method by Scholtes [34], for example, still converges to a C-stationary point, i.e. for this method, we do not lose anything by replacing exact KKT-points by approximate KKT-points. At least for the authors, this result was quite astonishing, since originally we thought that the limit points were only weakly stationary also for this method. However, since, in contrast to most of the other relaxation schemes, we were not able to find a corresponding example, we tried to prove convergence to C-stationary points and eventually succeeded. This convergence property may partially explain the results from [16], where, among other results, the authors provided a numerical comparison of the five relaxation methods considered here based on a collection of test problems. And although some of the other methods have better theoretical properties, the relaxation introduced by Scholtes was the best with respect to the criteria number of solved problems, function value in the solution, and elapsed time. Apart from this, however, we also give additional conditions for the methods from [19, 21, 35] under which they converge to C-, M-, or strongly stationary points.

The paper is organized in the following way: Section 2 contains some basic definitions like suitable constraint qualifications, stationarity concepts as well as our notion of an approximate KKT-point. The subsequent sections then consider the convergence behaviour of the different inexact relaxation methods. We begin with the global relaxation by Scholtes in Section 3 and proceed with a summary of the results for the smooth relaxation by Lin and Fukushima in Section 4 which also

converges to C-stationary points when approximate KKT-points are considered. Section 5 then considers the local relaxation by Steffensen and Ulbrich, and the subsequent two Sections 6 and 7 deal with the nonsmooth and L-shaped relaxation by Kadrani et al. and by the authors, respectively. We close this paper with some final remarks in Section 8.

A few words concerning the notation: For a smooth function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, we denote the *gradient* of f at a point x by $\nabla f(x)$ and assume that it is a column vector. The *support* of a vector $x \in \mathbb{R}^n$ is abbreviated as $\text{supp}(x) := \{i \in \{1, \dots, n\} \mid x_i \neq 0\}$.

2 Preliminaries

2.1 Standard Nonlinear Programs

As mentioned previously, the idea of a relaxation method is to relax the complementarity constraints and thus obtain a sequence of standard nonlinear programs. For this reason, we need some notation and a few basic facts about NLPs. Consider the following nonlinear program

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

where $f, g_i, h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are assumed to be continuously differentiable. We denote the feasible set by $X \subseteq \mathbb{R}^n$ and define the set of active inequalities $I_g(x^*) := \{i \in \{1, \dots, m\} \mid g_i(x^*) = 0\}$ for an $x^* \in X$.

Now let $x^* \in X$ be a local minimum of (2) and assume that a suitable constraint qualification holds in x^* . Under these assumptions, it is well known that x^* is a *stationary point*, i.e. there exist multipliers $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ such that (x^*, λ, μ) is a *KKT-point*. This means that the triple (x^*, λ, μ) satisfies the equation

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

and the conditions $\lambda \geq 0, \text{supp}(\lambda) \subseteq I_g(x^*)$, see, e.g., [3, 26] for more details.

Unfortunately, when NLPs are solved numerically, one rarely ends up in a KKT-point. The termination criteria used in standard software like IPOPT [38], SNOPT [13], KNITRO [4, 5], or filterSQP [11] checks whether an approximate KKT-point has been found (in addition to other stopping criteria). If we neglect performance improving details such as slack variables and scaling, these methods produce approximate solutions (x^*, λ, μ) satisfying the following conditions.

Definition 2.1 *Let $x^* \in \mathbb{R}^n$ and $\varepsilon > 0$ be given. If there exist vectors $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ such that*

$$\begin{aligned} & \left\| \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{i=1}^p \mu_i \nabla h_i(x^*) \right\|_{\infty} \leq \varepsilon, \\ & g_i(x^*) \leq \varepsilon, \quad \lambda_i \geq -\varepsilon, \quad |g_i(x^*)\lambda_i| \leq \varepsilon \quad \forall i = 1, \dots, m, \\ & |h_i(x^*)| \leq \varepsilon \quad \forall i = 1, \dots, p, \end{aligned} \tag{3}$$

x^* is called an ε -stationary point of the NLP (2).

It is clear that the single ε used in the previous definition can be replaced by different ones for the different parts of the KKT-conditions. In order to keep the notation simple, we decided to take the same ε for all parts. Furthermore, it is also clear that an appropriate definition of an ε -stationary point depends on the particular class of methods or even on the particular solver. For example, interior-point methods generate iterates where nonnegativity constraints (on multipliers or slack variables) automatically hold, hence the condition $\lambda_i \geq -\varepsilon$ could be replaced by the stronger condition $\lambda_i \geq 0$ in this case. SQP-type methods typically also guarantee that nonnegativity constraints are satisfied, but due to numerical inaccuracies, there is usually no guarantee for this. Finally, semismooth Newton methods applied to a suitable reformulation of the KKT-conditions do not take into account any box constraints (in general). Since our subsequent analysis should be independent of a particular solver, we use the framework from Definition 2.1 that we believe is general enough to cover all situations of interest.

2.2 Mathematical Programs with Complementarity Constraints

Now let us return to the MPEC (1). Again, we denote the set of feasible points by $X \subseteq \mathbb{R}^n$ and define the following index sets for an $x^* \in X$:

$$\begin{aligned} I_g(x^*) &:= \{i \in \{1, \dots, m\} \mid g_i(x^*) = 0\}, \\ I_{0+}(x^*) &:= \{i \in \{1, \dots, q\} \mid G_i(x^*) = 0, H_i(x^*) > 0\}, \\ I_{00}(x^*) &:= \{i \in \{1, \dots, q\} \mid G_i(x^*) = 0, H_i(x^*) = 0\}, \\ I_{+0}(x^*) &:= \{i \in \{1, \dots, q\} \mid G_i(x^*) > 0, H_i(x^*) = 0\}. \end{aligned}$$

Obviously, $I_g(x^*)$ is the set of active inequalities as defined for NLPs and the sets $I_{0+}(x^*)$, $I_{00}(x^*)$, and $I_{+0}(x^*)$ form a partition of the set of complementarity constraints. If the point x^* is clear from the context, we sometimes abbreviate the index sets by I_g , I_{0+} , I_{00} , and I_{+0} , respectively.

In contrast to NLPs, where KKT-points are the most common stationarity concept, a number of different stationarity definitions for MPECs have emerged over the last few years. Here, we will restrict ourselves to those important for our analysis.

Definition 2.2 *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 the equation*

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

and the conditions

$$\lambda_i \geq 0, \quad (i \in I_g), \quad \lambda_i = 0 \quad (i \notin I_g), \quad \gamma_i = 0 \quad (i \in I_{+0}), \quad \nu_i = 0 \quad (i \in I_{0+})$$

are satisfied;

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

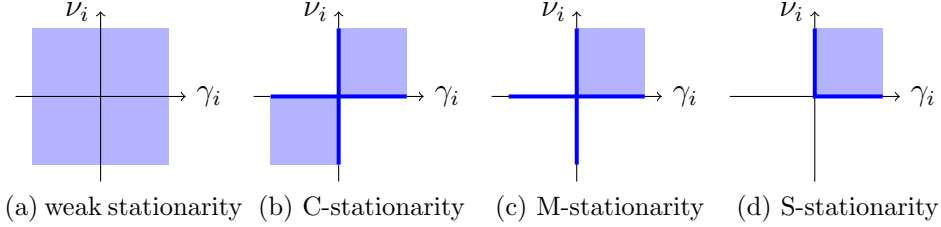


Figure 1: Geometric illustration of weak, C-, M-, and strong stationarity for an index i from the bi-active set $I_{00}(x^*)$

The different conditions on the multipliers γ_i, ν_i with $i \in I_{00}(x^*)$ are illustrated in Figure 1. Obviously, the stationarity concepts differ only in the conditions on these multipliers and thus coincide when the biactive set $I_{00}(x^*)$ is empty. Otherwise, the following implications hold:

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

The notion of weak and C-stationarity comes from the seminal paper [33], whereas M-stationarity was introduced independently in [41, 27, 28, 39], and the concept of strong stationarity may already be found in [25]. As pointed out in [9], strong stationarity is equivalent to the standard KKT conditions of an MPEC.

In order to guarantee that a local minimum x^* of (1) is a stationary point in any of the previous senses, one needs to assume that a constraint qualification is satisfied in x^* . Since most standard CQs are violated in feasible points of (1), many MPEC-analogues of these CQs have been developed. Here, we mention only those needed later.

Definition 2.3 A feasible point x^* of the MPEC (1) is said to satisfy the

- (a) MPEC-linear independence CQ (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-Mangasarian Fromovitz CQ (MPEC-MFCQ), 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 positively linearly independent;

(c) MPEC-constant rank CQ (MPEC-CRCQ), 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 linearly dependent, there exists a neighborhood $N(x^*)$ of x^* such that the gradients

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

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

(d) MPEC-constant positive linear dependence CQ (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 positively linearly dependent, there exists a neighborhood $N(x^*)$ of x^* such that the gradients

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

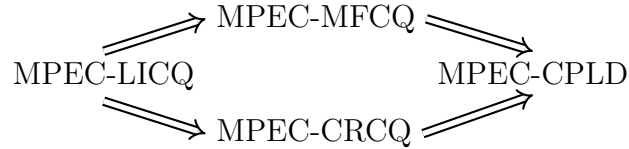
are linearly dependent for all $x \in N(x^*)$.

In the definition of MPEC-MFCQ and MPEC-CPLD, we use the notion of positive linear dependent vectors. A set of vectors $a_i, b_j \in \mathbb{R}^n$, $i \in I$, $j \in J$ is called *positively linearly dependent*, if there exist scalars $\{\alpha_i\}_{i \in I}$ and $\{\beta_j\}_{j \in J}$ with $\alpha_i \geq 0$ for all $i \in I$, not all of them being zero, such that

$$\sum_{i \in I} \alpha_i a_i + \sum_{j \in J} \beta_j b_j = 0.$$

Otherwise, the set of vectors is called *positively linearly independent*.

Obviously, linear independence implies positive linear independence. Hence, the following implications hold between the MPEC-CQs:



Note that each of these CQs imply that a local minimum is M-stationary, see, e.g., [10, 40], but only MPEC-LICQ is sufficient to guarantee strong stationarity of a local minimum, cf. [9, 25, 30] and references therein. The MPEC-LICQ is among the first MPEC-tailored constraint qualifications and may already be found in [25, 33], the MPEC-MFCQ was introduced in [33], MPEC-CRCQ arises for the first time in [35], and MPEC-CPLD is a more recent conditions from [15, 16]. Note that MPEC-MFCQ is typically defined in a different way, but can easily be seen to be equivalent to the above condition.

3 The Global Relaxation by Scholtes

In [34], Scholtes suggested to replace the complementarity conditions by the inequalities

$$G_i(x) \geq 0, H_i(x) \geq 0, \Phi_i^S(x; t) := G_i(x)H_i(x) - t \leq 0 \quad \forall i = 1, \dots, q$$

with a relaxation parameter $t > 0$. This transforms the feasible set of the complementarity constraints to the shape depicted in Figure 2.

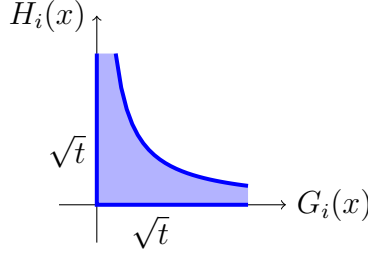


Figure 2: Geometric interpretation of the relaxation by Scholtes

The relaxation leads to a sequence of relaxed nonlinear programs

$$\begin{aligned} \min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 & \forall i = 1, \dots, m, \\ & h_i(x) = 0 & \forall i = 1, \dots, p, \\ & G_i(x) \geq 0 & \forall i = 1, \dots, q, \\ & H_i(x) \geq 0 & \forall i = 1, \dots, q, \\ & \Phi_i^S(x; t) = G_i(x)H_i(x) - t \leq 0 & \forall i = 1, \dots, q \end{aligned} \quad (4)$$

with $t \downarrow 0$, which we will denote by $\text{NLP}^S(t)$. From the original paper [34], the following convergence result is known.

Theorem 3.1 *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 $\text{NLP}^S(t_k)$. If $x^k \rightarrow x^*$ and MPEC-LICQ holds in x^* , then x^* is a C-stationary point of the MPEC (1).*

In [16], it was shown that MPEC-MFCQ is in fact enough to guarantee C-stationarity of a limit point. But what happens, if we consider ε_k -stationary points of $\text{NLP}(t_k)$ instead of stationary points? In the case of Scholtes' relaxation, the following result shows that this does not have any effect as long as we take care about the speed of convergence of $\varepsilon_k \downarrow 0$.

Theorem 3.2 *Let $\{t_k\} \downarrow 0$, $\varepsilon_k = o(t_k)$, $\{x^k\}$ be a sequence of ε_k -stationary points of $\text{NLP}^S(t_k)$ with multipliers $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$, and assume that $x^k \rightarrow x^*$. If MPEC-MFCQ holds in x^* , then x^* is a C-stationary point of the MPEC.*

Proof: Since all x^k are ε_k -stationary points of $\text{NLP}(t_k)$, we have

$$\left\| \begin{aligned} & \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) \\ & - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^q \delta_i^k \nabla \Phi_i^S(x^k; t_k) \end{aligned} \right\|_{\infty} \leq \varepsilon_k$$

with

$$\begin{aligned}
g_i(x^k) &\leq \varepsilon_k, & \lambda_i^k &\geq -\varepsilon_k, & |\lambda_i^k g_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, m, \\
|h_i(x^k)| &\leq \varepsilon_k & & & & & \forall i = 1, \dots, p, \\
G_i(x^k) &\geq -\varepsilon_k, & \gamma_i^k &\geq -\varepsilon_k, & |\gamma_i^k G_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\
H_i(x^k) &\geq -\varepsilon_k, & \nu_i^k &\geq -\varepsilon_k, & |\nu_i^k H_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\
\Phi_i^S(x^k; t_k) &\leq \varepsilon_k, & \delta_i^k &\geq -\varepsilon_k, & |\delta_i^k \Phi_i^S(x^k; t_k)| &\leq \varepsilon_k & \forall i = 1, \dots, q,
\end{aligned}$$

where $\nabla \Phi_i^S(x^k; t_k) = H_i(x^k) \nabla G_i(x^k) + G_i(x^k) \nabla H_i(x^k)$. Obviously, the limit x^* is feasible for the MPEC (1). We define the multipliers

$$\begin{aligned}
\delta_i^{G,k} &:= \begin{cases} \delta_i^k H_i(x^k) & \text{if } i \in I_{00}(x^*) \cup I_{0+}(x^*), \\ 0 & \text{if } i \in I_{+0}(x^*), \end{cases} \\
\delta_i^{H,k} &:= \begin{cases} \delta_i^k G_i(x^k) & \text{if } i \in I_{00}(x^*) \cup I_{+0}(x^*), \\ 0 & \text{if } i \in I_{0+}(x^*). \end{cases}
\end{aligned}$$

Then we have

$$\begin{aligned}
&\left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) \right. \\
&\quad - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) + \sum_{i=1}^q \delta_i^{G,k} \nabla G_i(x^k) + \sum_{i \in I_{+0}} \delta_i^k H_i(x^k) \nabla G_i(x^k) \\
&\quad \left. - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^q \delta_i^{H,k} \nabla H_i(x^k) + \sum_{i \in I_{0+}} \delta_i^k G_i(x^k) \nabla H_i(x^k) \right\|_{\infty} \leq \varepsilon_k.
\end{aligned}$$

We claim that the multipliers $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k}, \delta_{I_{+0} \cup I_{0+}}^k)$ are bounded. If the sequence were unbounded, we could assume without loss of generality convergence of the sequence

$$\frac{(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k}, \delta_{I_{+0} \cup I_{0+}}^k)}{\|(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k}, \delta_{I_{+0} \cup I_{0+}}^k)\|} \rightarrow (\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}, \bar{\delta}^G, \bar{\delta}^H, \bar{\delta}_{I_{+0} \cup I_{0+}}) \neq 0.$$

Then the ε_k -stationarity of x^k yields

$$\begin{aligned}
&\sum_{i=1}^m \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^p \bar{\mu}_i \nabla h_i(x^*) - \sum_{i=1}^q \bar{\gamma}_i \nabla G_i(x^*) - \sum_{i=1}^q \bar{\nu}_i \nabla H_i(x^*) \\
&\quad + \sum_{i=1}^q \bar{\delta}_i^G \nabla G_i(x^*) + \sum_{i=1}^q \bar{\delta}_i^H \nabla H_i(x^*) = 0,
\end{aligned} \tag{5}$$

where we took into account that $H_i(x^k) \rightarrow 0$ ($i \in I_{+0}(x^*)$) and $G_i(x^k) \rightarrow 0$ ($i \in I_{0+}(x^*)$). For all $i = 1, \dots, m$, the ε_k -stationarity implies $\bar{\lambda}_i \geq 0$. If $\bar{\lambda}_i > 0$, we have $\lambda_i^k > c$ for some constant $c > 0$ and all k sufficiently large. This yields

$$0 \leq \frac{|g_i(x^k)|}{t_k} \leq \frac{\varepsilon_k}{t_k |\lambda_i^k|} \leq \frac{\varepsilon_k}{t_k c} \rightarrow 0$$

due to $\varepsilon_k = o(t_k)$ and thus $i \in I_g(x^*)$. Analogously, we have $\bar{\gamma}_i \geq 0$ for all $i = 1, \dots, q$ and $\bar{\gamma}_i > 0$ implies $\frac{G_i(x^k)}{t_k} \rightarrow 0$ and thus $i \in I_{00}(x^*) \cup I_{0+}(x^*)$, and also $\bar{\nu}_i \geq 0$ for all $i = 1, \dots, q$ with $\bar{\nu}_i > 0$ implying $\frac{H_i(x^k)}{t_k} \rightarrow 0$ and thus $i \in I_{00}(x^*) \cup I_{+0}(x^*)$.

Now assume $\bar{\delta}_i^G < 0$ for some i . This implies $i \in I_{00}(x^*) \cup I_{0+}(x^*)$ and $\delta_i^{G,k} = \delta_i^k H_i(x^k) < -c$ for some constant $c > 0$ and all k sufficiently large. If $i \in I_{0+}(x^*)$ it would follow that $\delta_i^k < -\frac{c}{2H_i(x^*)} < 0$ for all k sufficiently large, a contradiction to the ε_k -stationarity of x^k . Hence, we know $i \in I_{00}(x^*)$, which implies $|\delta_i^k| \rightarrow \infty$. Therefore, we can find a constant $c > 0$ such that $|\delta_i^k| > c$ for all k sufficiently large. Using this, we get

$$0 \leq \frac{|G_i(x^k)H_i(x^k) - t_k|}{t_k} \leq \frac{\varepsilon_k}{t_k|\delta_i^k|} \leq \frac{\varepsilon_k}{t_k c} \rightarrow 0. \quad (6)$$

Similarly, if we assume $\bar{\delta}_i^G > 0$ for some i , we know $i \in I_{00}(x^*) \cup I_{0+}(x^*)$ and $\delta_i^{G,k} = \delta_i^k H_i(x^k) > c$ for some constant $c > 0$ and all k sufficiently large. If $i \in I_{00}(x^*)$, this implies $|\delta_i^k| \rightarrow \infty$, whereas for $i \in I_{0+}(x^*)$ we get $|\delta_i^k| > \frac{c}{2H_i(x^*)}$ for all k sufficiently large. In any case, we can find a constant $c > 0$ with $|\delta_i^k| > c$ for all k sufficiently large and proceed as we did in the previous case in order to see that (6) holds.

Consequently, $\bar{\delta}_i^G \neq 0$ implies $\frac{G_i(x^k)H_i(x^k)}{t_k} \rightarrow 1$ and thus $\frac{G_i(x^k)}{t_k} \rightarrow 0$ and $\frac{H_i(x^k)}{t_k} \rightarrow 0$. Using a completely symmetric argument, the same holds for all indices i with $\bar{\delta}_i^H \neq 0$. From these observations, we can conclude that

$$\text{supp}(\bar{\gamma}) \cap \text{supp}(\bar{\delta}^G) = \emptyset \quad \text{and} \quad \text{supp}(\bar{\nu}) \cap \text{supp}(\bar{\delta}^H) = \emptyset$$

Furthermore, we have

$$\text{supp}(\bar{\gamma}) \cup \text{supp}(\bar{\delta}^G) \subseteq I_{00} \cup I_{0+} \quad \text{and} \quad \text{supp}(\bar{\nu}) \cup \text{supp}(\bar{\delta}^H) \subseteq I_{00} \cup I_{+0}.$$

Hence, if $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}, \bar{\delta}^G, \bar{\delta}^H) \neq 0$, (5) yields a contradiction to MPEC-MFCQ.

If, on the other hand, $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}, \bar{\delta}^G, \bar{\delta}^H) = 0$, there has to be an $i \in I_{0+}(x^*) \cup I_{+0}(x^*)$ with $\bar{\delta}_i \neq 0$. First consider the case $i \in I_{0+}(x^*)$. Then by definition

$$\bar{\delta}_i^G = \lim_{k \rightarrow \infty} \frac{\delta_i^k H_i(x^k)}{\|(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k}, \delta_{I_{+0} \cup I_{0+}}^k)\|} = \bar{\delta}_i \underbrace{H_i(x^*)}_{>0} \neq 0,$$

a contradiction to the assumption $\bar{\delta}_i^G = 0$. In an analogous way, we obtain a contradiction in the case $i \in I_{+0}(x^*)$.

Consequently, the sequence $\{(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^{G,k}, \delta^{H,k}, \delta_{I_{+0} \cup I_{0+}}^k)\}$ is bounded and therefore converges to some limit $(\lambda^*, \mu^*, \tilde{\gamma}, \tilde{\nu}, \tilde{\delta}^G, \tilde{\delta}^H, \tilde{\delta}_{I_{+0} \cup I_{0+}})$ at least on a subsequence. By passing to this subsequence, we can assume convergence on the whole sequence. It is easy to see that the support of this limit has the same properties we derived before for $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}, \bar{\delta}^G, \bar{\delta}^H, \bar{\delta}_{I_{+0} \cup I_{0+}})$. Thus, the following multipliers are well defined:

$$\gamma_i^* = \begin{cases} \tilde{\gamma}_i & \text{if } i \in \text{supp}(\tilde{\gamma}), \\ -\tilde{\delta}_i^G & \text{if } i \in \text{supp}(\tilde{\delta}^G), \\ 0 & \text{else} \end{cases} \quad \text{and} \quad \nu_i^* = \begin{cases} \tilde{\nu}_i & \text{if } i \in \text{supp}(\tilde{\nu}), \\ -\tilde{\delta}_i^H & \text{if } i \in \text{supp}(\tilde{\delta}^H), \\ 0 & \text{else} \end{cases}$$

and x^* together with the multipliers $(\lambda^*, \mu^*, \gamma^*, \nu^*)$ is a weakly stationary point of the MPEC (1).

In order to prove C-stationarity of x^* , assume that there were an $i \in I_{00}(x^*)$ such that $\gamma_i^* \nu_i^* < 0$. We consider without loss of generality only the case $\gamma_i^* < 0, \nu_i^* > 0$, the other one can be treated the same way. Since we know that $\tilde{\gamma}_i \geq 0$, this implies $i \in \text{supp}(\tilde{\delta}^G)$ and consequently $\delta_i^{G,k} = \delta_i^k H_i(x^k) > 0$ for all k sufficiently large. Since $i \in \text{supp}(\tilde{\delta}^G)$ also implies $\frac{G_i(x^k)H_i(x^k)}{t_k} \rightarrow 1$ (by a similar argument leading to (6)), we can conclude $i \in \text{supp}(\tilde{\delta}^H)$ because $i \in \text{supp}(\tilde{\nu})$ would imply $\frac{H_i(x^k)}{t_k} \rightarrow 0$ in view of the ε_k -stationarity. Thus, we also have $\delta_i^{H,k} = \delta_i^k G_i(x^k) < 0$ for all k sufficiently large. Hence, $G_i(x^k)$ and $H_i(x^k)$ have different signs, a contradiction to $\frac{G_i(x^k)H_i(x^k)}{t_k} \rightarrow 1$. Thus, x^* with the multipliers $(\lambda^*, \mu^*, \gamma^*, \nu^*)$ is a C-stationary point of the MPEC (1). \square

4 The Smooth Relaxation by Lin and Fukushima

In contrast to the previous approach, Lin and Fukushima suggest in [24] to replace the complementarity conditions by only two inequalities

$$\Phi_i^{LF}(x; t) := (G_i(x) + t)(H_i(x) + t) - t^2 \geq 0 \quad \forall i = 1, \dots, q$$

and

$$\Phi_i^S(x; t^2) := G_i(x)H_i(x) - t^2 \leq 0 \quad \forall i = 1, \dots, q$$

with a relaxation parameter $t > 0$. This transforms the feasible set of the complementarity constraints to the shape depicted in Figure 3.

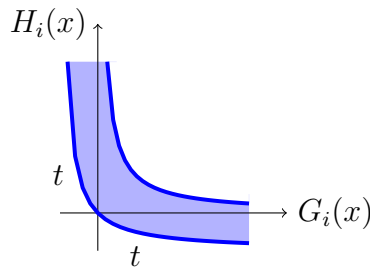


Figure 3: Geometric interpretation of the relaxation by Lin and Fukushima

Note that Φ_i^S is the function from Scholtes but here the parameter t is replaced by t^2 , whereas the potentially simple nonnegativity conditions are replaced by the new function Φ_i^{LF} . This relaxation leads to a different sequence of relaxed nonlinear programs

$$\begin{aligned} \min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 & \forall i = 1, \dots, m, \\ & h_i(x) = 0 & \forall i = 1, \dots, p, \\ & \Phi_i^{LF}(x; t) = (G_i(x) + t)(H_i(x) + t) - t^2 \geq 0 & \forall i = 1, \dots, q, \\ & \Phi_i^S(x; t^2) = G_i(x)H_i(x) - t^2 \leq 0 & \forall i = 1, \dots, q \end{aligned} \quad (7)$$

with $t \downarrow 0$, which we will denote by $NLP^{LF}(t)$. From the original paper [24], the following convergence result is known.

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

Reference [14] shows that the above result remains true under the weaker MPEC-MFCQ condition. The next result shows that we still get C-stationary points in the limit if we compute only ε_k -stationary points of the nonlinear programs $NLP^{LF}(t_k)$ provided that ε_k goes to zero sufficiently fast. A proof of this result is given in the accompanying paper [22].

Theorem 4.2 *Let $\{t_k\} \downarrow 0$, $\varepsilon_k = o(t_k^2)$, $\{x^k\}$ be a sequence of ε_k -stationary points of $NLP^{LF}(t_k)$, and assume that $x^k \rightarrow x^*$. If MPEC-MFCQ holds in x^* , then x^* is a C-stationary point of the MPEC.*

The assumptions in Theorem 4.2 are essentially identical to those from Theorem 3.2 for the Scholtes approach. Formally, the condition $\varepsilon_k = o(t_k^2)$ looks stronger than $\varepsilon_k = o(t_k)$, but recall that the Scholtes function Φ_i^S is parameterized by t^2 instead of t in the Lin-Fukushima-regularization, hence these conditions coincide.

Since the proof of Theorem 4.2 is not given here, let us add some comments why a result of this kind had to be expected, taking into account that we already know from the previous section that the Scholtes-approach converges to C-stationary points: The complementarity conditions $G_i(x) \geq 0, H_i(x) \geq 0, G_i(x)H_i(x) = 0$ consist of two parts, namely the two nonnegativity constraints and the pure complementarity relation. What makes the MPEC a complicated problem are these pure complementarity relations. But these pure complementarity conditions are treated by essentially the same function Φ_i^S as in the approach by Scholtes, whereas the less complicated nonnegativity constraints are replaced by the new function Φ_i^{LF} . Hence one can expect that the latter does not cause severe troubles, whereas the former can be dealt with as in the Scholtes approach. In principle, these heuristics justify the convergence to C-stationary points for the inexact Lin-Fukushima relaxation method. Nevertheless, the detailed analysis is quite involved and needs some extra techniques, hence we decided to treat this method in a separate paper [22], also taking into account that the current one is already quite long.

5 The Local Relaxation by Steffensen and Ulbrich

In contrast to the previous two methods, Steffensen and Ulbrich suggest in [35] to relax the feasible area for the complementarity constraints only around the kink in the origin. To this end, they use a relaxation function Φ_i^{SU} defined by

$$\Phi_i^{SU}(x; t) := G_i(x) + H_i(x) - \varphi(G_i(x) - H_i(x); t)$$

for a function $\varphi(\cdot; t) : \mathbb{R} \rightarrow \mathbb{R}$ given by

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

where θ denotes a regularization function as defined in [15, 35]. In particular, θ has the following properties:

- θ is twice continuously differentiable,
- $\theta(-1) = \theta(1) = 1$,
- $\theta'(-1) = -1$ and $\theta'(1) = 1$,
- $\theta(a) > |a|$ for all $a \in (-1, +1)$,
- $|\theta'(a)| < 1$ for all $a \in (-1, +1)$.

This function Φ^{SU} is used to relax the complementarity constraints by

$$G_i(x) \geq 0, \quad H_i(x) \geq 0, \quad \Phi_i^{SU}(x; t) \leq 0 \quad \forall i = 1, \dots, q.$$

The resulting shape of the feasible set of the complementarity constraints is depicted in Figure 4.

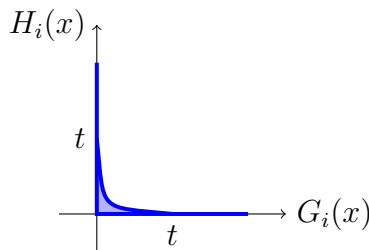


Figure 4: Geometric interpretation of the relaxation by Steffensen and Ulbrich

Hence, we consider the following sequence of nonlinear programs

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

with $t \downarrow 0$, which we will denote by $NLP^{SU}(t)$. From [35], the following result is known.

Theorem 5.1 *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 $NLP^{SU}(t_k)$. If $x^k \rightarrow x^*$ and MPEC-CRCQ holds in x^* , then x^* is a C-stationary point of the MPEC (1).*

In [15] this result was proven to hold under the weaker assumption of MPEC-CPLD in the limit point; in particular, the statement therefore holds under MPEC-MFCQ. Again, we want to analyze the price we have to pay if we replace the sequence of KKT-points by a sequence of ε_k -stationary points. To do so, we need the following auxiliary result, which is a direct consequence of the definition of Φ_i , see also [15, 35].

Lemma 5.2 *The function Φ_i^{SU} is continuously differentiable with gradient*

$$\nabla \Phi_i^{SU}(x^k; t_k) = \alpha_i^k \nabla G_i(x^k) + \beta_i^k \nabla H_i(x^k),$$

where

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

Throughout this section, the scalars α_i^k, β_i^k will always denote the numbers defined in Lemma 5.2.

The following is the first main convergence result for the inexact local regularization method from [35].

Theorem 5.3 *Let $\{t_k\} \downarrow 0$, $\{\varepsilon_k\} \downarrow 0$, $\{x^k\}$ be a sequence of ε_k -stationary points of $NLP^{SU}(t_k)$, and assume that $x^k \rightarrow x^*$. Then x^* is a weakly stationary point of the MPEC provided that MPEC-MFCQ holds at x^* .*

Proof: Since all x^k are ε_k -stationary points of $NLP^{SU}(t_k)$, there exist multipliers $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$ such that

$$\begin{aligned} & \left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) \right. \\ & \left. + \sum_{i=1}^q \delta_i^k \nabla \Phi_i^{SU}(x^k; t_k) \right\|_{\infty} \leq \varepsilon_k \end{aligned}$$

with

$$\begin{aligned} g_i(x^k) &\leq \varepsilon_k, & \lambda_i^k &\geq -\varepsilon_k, & |\lambda_i^k g_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, m, \\ |h_i(x^k)| &\leq \varepsilon_k, & & & & & \forall i = 1, \dots, p, \\ G_i(x^k) &\geq -\varepsilon_k, & \gamma_i^k &\geq -\varepsilon_k, & |\gamma_i^k G_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\ H_i(x^k) &\geq -\varepsilon_k, & \nu_i^k &\geq -\varepsilon_k, & |\nu_i^k H_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\ \Phi_i^{SU}(x^k; t_k) &\leq \varepsilon_k, & \delta_i^k &\geq -\varepsilon_k, & |\delta_i^k \Phi_i^{SU}(x^k; t_k)| &\leq \varepsilon_k & \forall i = 1, \dots, q. \end{aligned}$$

This obviously implies that the limit x^* is feasible for the MPEC (1). Taking into account the expression for the gradient $\nabla \Phi_i(x^k; t_k)$ from Lemma 5.2 and using the abbreviations

$$\tilde{\gamma}_i^k := \gamma_i^k - \delta_i^k \alpha_i^k, \quad \tilde{\nu}_i^k := \nu_i^k - \delta_i^k \beta_i^k \quad \forall i = 1, \dots, q,$$

we can rewrite the sum as

$$\left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \tilde{\gamma}_i^k \nabla G_i(x^k) - \sum_{i=1}^q \tilde{\nu}_i^k \nabla H_i(x^k) \right\|_{\infty} \leq \varepsilon_k.$$

We claim that the sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ stays bounded. Otherwise, we may assume without loss of generality that

$$\frac{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)}{\|(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\|} \rightarrow (\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) \neq 0.$$

Dividing the previous inequality by $\|(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\|$ and taking the limit $k \rightarrow \infty$, we then obtain

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

From the ε_k -stationarity, we can infer $\bar{\lambda}_i \geq 0$ for all $i = 1, \dots, m$, and $\bar{\lambda}_i > 0$ implies $g_i(x^k) \rightarrow 0$ and, consequently, $i \in I_g(x^*)$.

We next claim that $\text{supp}(\bar{\gamma}) \subseteq I_{00}(x^*) \cup I_{0+}(x^*)$. Suppose this is not true, i.e., there exists an index i with $\bar{\gamma}_i \neq 0$ and $i \in I_{+0}(x^*)$. Then we have $G_i(x^*) > 0$ and $H_i(x^*) = 0$. In particular, the inequality $G_i(x^k) - H_i(x^k) \geq t_k$ then holds for all k sufficiently large. Hence Lemma 5.2 gives $\alpha_i^k = 0$ and, therefore, $\tilde{\gamma}_i^k = \gamma_i^k$. Consequently, $\bar{\gamma}_i \neq 0$ yields $|\gamma_i^k| = |\tilde{\gamma}_i^k| \rightarrow 0$. It therefore follows from the ε_k -stationarity that $G_i(x^k) \rightarrow 0$, a contradiction to $G_i(x^*) > 0$ and the continuity of G_i . By a symmetric argument, we get $\text{supp}(\bar{\nu}) \subseteq I_{00}(x^*) \cup I_{+0}(x^*)$.

Hence, our equation reduces to

$$\sum_{i \in I_g} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^p \bar{\mu}_i \nabla h_i(x^*) - \sum_{i \in I_{0+} \cup I_{00}} \bar{\gamma}_i \nabla G_i(x^*) - \sum_{i \in I_{+0} \cup I_{00}} \bar{\nu}_i \nabla H_i(x^*) = 0.$$

Since $\bar{\lambda}_i \geq 0$ for all $i \in I_g(x^*)$, the assumed MPEC-MFCQ therefore gives $\bar{\lambda}_i = 0$ ($i \in I_g$), $\bar{\mu}_i = 0$ ($i = 1, \dots, p$), $\bar{\gamma}_i = 0$ ($i \in I_{0+} \cup I_{00}$), $\bar{\nu}_i = 0$ ($i \in I_{+0} \cup I_{00}$). Since the remaining multipliers were already shown to be zero, we get a contradiction to $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) \neq 0$. This shows that the sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ is indeed bounded.

Subsequencing if necessary, we may therefore assume that

$$(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k) \rightarrow (\lambda^*, \mu^*, \gamma^*, \nu^*).$$

By the same arguments used above, this implies $\lambda_i^* \geq 0$ for all $i = 1, \dots, m$ and $\text{supp}(\lambda^*) \subseteq I_g(x^*)$, as well as $\text{supp}(\gamma^*) \subseteq I_{00}(x^*) \cup I_{0+}(x^*)$ and $\text{supp}(\nu^*) \subseteq I_{00}(x^*) \cup I_{+0}(x^*)$. This shows that x^* , together with the multipliers $(\lambda^*, \mu^*, \gamma^*, \nu^*)$, is a weakly stationary point of the MPEC. \square

In order to verify C-stationarity, we need a stronger assumption regarding the choice of the sequence $\{\varepsilon_k\}$ and, more importantly, a condition regarding the way the sequence $\{x^k\}$ is computed as an inexact solution of $\text{NLP}(t_k)$. The second main convergence result is now formulated in the following theorem, where we recall that the numbers α_i^k, β_i^k are those defined in Lemma 5.2.

Theorem 5.4 *Let $\{t_k\} \downarrow 0$, $\varepsilon_k = o(t_k)$, $\{x^k\}$ be a sequence of ε_k -stationary points of $\text{NLP}^{SU}(t_k)$, and assume that $x^k \rightarrow x^*$. Suppose further that there is a constant $c > 0$ such that, for all $i \in I_{00}(x^*)$, we have*

$$\alpha_i^k \geq c \quad \text{and} \quad \beta_i^k \geq c \quad \text{for all sufficiently large } k \text{ with } |G_i(x^k) - H_i(x^k)| < t_k. \quad (8)$$

Then x^ is a C-stationary point of the MPEC provided that MPEC-MFCQ holds at x^* .*

Proof: We know from Theorem 5.3 that the limit point x^* , together with the multipliers $(\lambda^*, \mu^*, \gamma^*, \nu^*)$ constructed in the previous proof, is at least weakly stationary. In order to verify C-stationarity, assume that there exists an index $i \in I_{00}(x^*)$ such that $\gamma_i^* \nu_i^* < 0$. By symmetry, and subsequencing if necessary, we may assume without loss of generality that

$$\gamma_i^* = \lim_{k \rightarrow \infty} (\gamma_i^k - \alpha_i^k \delta_i^k) < 0 \quad \text{and} \quad \nu_i^* = \lim_{k \rightarrow \infty} (\nu_i^k - \beta_i^k \delta_i^k) > 0.$$

From the ε_k -stationarity of x^k and the properties of θ , we know

$$\liminf_{k \rightarrow \infty} \gamma_i^k \geq 0, \quad \liminf_{k \rightarrow \infty} \nu_i^k \geq 0, \quad \liminf_{k \rightarrow \infty} \delta_i^k \geq 0 \quad \text{and} \quad \alpha_i^k, \beta_i^k \in [0, 2] \quad \forall k \in \mathbb{N}.$$

Hence, $\gamma_i^* < 0$ and $\nu_i^* > 0$ implies $\alpha_i^k \delta_i^k > C$ and $\nu_i^k > C$ for some constant $C > 0$ and all k sufficiently large. As in the proof of Theorem 3.2, the latter yields $\frac{H_i(x^k)}{t_k} \rightarrow 0$ from the ε_k -stationarity. We now distinguish three cases and derive a contradiction for each of these cases.

Case 1: There is a subsequence such that $G_i(x^k) - H_i(x^k) \leq -t_k$ holds for all $k \in K$ with some index set $K \subseteq \mathbb{N}$. Division by t_k shows that

$$-1 \geq \frac{G_i(x^k)}{t_k} - \frac{H_i(x^k)}{t_k} \geq \frac{-\varepsilon_k}{t_k} - \frac{H_i(x^k)}{t_k} \rightarrow_K 0,$$

a contradiction.

Case 2: There is a subsequence such that $G_i(x^k) - H_i(x^k) \geq t_k$ holds for all $k \in K$ for some $K \subseteq \mathbb{N}$. Then Lemma 5.2 implies $\alpha_i^k = 0$ for all $k \in K$, a contradiction to $\alpha_i^k \delta_i^k > C$ for all k sufficiently large.

Case 3: There is a subsequence such that $|G_i(x^k) - H_i(x^k)| < t_k$ for all $k \in K$ and some index set $K \subseteq \mathbb{N}$. Due to $\frac{H_i(x^k)}{t_k} \rightarrow 0$, this implies $\frac{G_i(x^k)}{t_k} \rightarrow_K \rho^* \in [-1, 1]$, at least on a subsequence. More precisely, $\rho^* \in [0, 1]$ because, for negative $G_i(x^k)$, the ε_k -stationarity implies $G_i(x^k) \in [-\varepsilon_k, 0)$ and thus $\frac{G_i(x^k)}{t_k} \rightarrow 0$. Due to $\alpha_i^k \delta_i^k > C$ for all k sufficiently large and the boundedness of α_i^k , we also know $\delta_i^k \rightarrow 0$. Hence, the ε_k -stationarity together with $\varepsilon_k = o(t_k)$ implies

$$\begin{aligned} 0 &= \lim_{k \rightarrow K\infty} \frac{|\delta_i^k| |\Phi_i^{SU}(x^k; t_k)|}{t_k} = \lim_{k \rightarrow K\infty} \frac{|\delta_i^k| \left| G_i(x^k) + H_i(x^k) - t_k \theta \left(\frac{G_i(x^k) - H_i(x^k)}{t_k} \right) \right|}{t_k} \\ &= \lim_{k \rightarrow K\infty} |\delta_i^k| \cdot \left(\left| \frac{G_i(x^k)}{t_k} + \frac{H_i(x^k)}{t_k} - \theta \left(\frac{G_i(x^k)}{t_k} - \frac{H_i(x^k)}{t_k} \right) \right| \right), \end{aligned}$$

i.e. $\rho^* - \theta(\rho^*) = 0$. According to the properties of θ , for $\rho^* \in [0, 1]$, this is only true for $\rho^* = 1$. Consequently, we know $\theta'(\rho^*) = 1$. However, by Lemma 5.2, this implies $\alpha_i^k \rightarrow_K 0$, a contradiction to the assumption (8).

Since one of these three cases must occur, it follows that the limit point x^* , together with the multipliers $(\lambda^*, \mu^*, \gamma^*, \nu^*)$, is a C-stationary point. \square

Let us take a closer look at condition (8). By Lemma 5.2, the coefficients α_i^k and β_i^k are given by

$$\alpha_i^k = 1 - \theta' \left(\frac{G_i(x^k) - H_i(x^k)}{t_k} \right) \quad \text{and} \quad \beta_i^k = 1 + \theta' \left(\frac{G_i(x^k) - H_i(x^k)}{t_k} \right)$$

if $|G_i(x^k) - H_i(x^k)| < t_k$. Due to the properties of θ , we have $\alpha_i^k, \beta_i^k \in (0, 2)$ with α_i^k going to zero if and only if $\frac{G_i(x^k) - H_i(x^k)}{t_k} \rightarrow 1$, and β_i^k approaching zero if and only if $\frac{G_i(x^k) - H_i(x^k)}{t_k} \rightarrow -1$. Hence, condition (8) is equivalent to the following condition: There is a constant $\tilde{c} \in (0, 1)$ such that, for all $i \in I_{00}(x^*)$, we have

$$\frac{|G_i(x^k) - H_i(x^k)|}{t_k} \leq \tilde{c} \quad \text{for all sufficiently large } k \text{ with } |G_i(x^k) - H_i(x^k)| < t_k.$$

A look at the proof of Theorem 5.4 (Case 3) reveals that it is necessary to keep α_i^k bounded away from zero only when $\frac{H_i(x^k)}{t_k} \rightarrow 0$, in which case we eventually have $H_i(x^k) \in (-\tilde{c}t_k, \tilde{c}t_k)$. A symmetrical argument shows that we only need to take care of β_i^k in case $\frac{G_i(x^k)}{t_k} \rightarrow 0$. From a geometrical point of view, this means that, for all $i \in I_{00}(x^*)$, the pairs $(G_i(x^k), H_i(x^k))$ must, at least for k sufficiently large, not lie in any of the two areas marked in Figure 5.

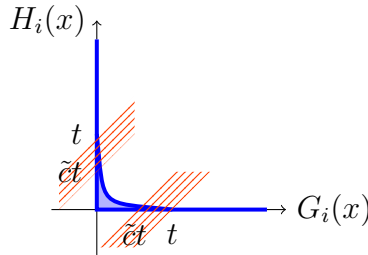


Figure 5: Geometric interpretation of condition (8)

The following examples show that the two additional assumptions (namely $\varepsilon_k = o(t_k)$ and condition (8)) are, in general, necessary in order to obtain a C-stationary point. The first counterexample shows that the condition $\varepsilon_k = o(t_k)$ had to be imposed in Theorem 5.4 (whereas this condition is not needed in Theorem 5.3 to prove weak stationarity of a limit point).

Example 5.5 Consider the simple MPEC

$$\min x_2 - x_1 \quad \text{s.t.} \quad x_1 \geq 0, x_2 \geq 0, x_1 x_2 = 0. \quad (9)$$

Furthermore, consider the particular regularization function

$$\theta(y) := \frac{1}{8}(-y^4 + 6y^2 + 3) \quad (10)$$

suggested in [37]. Note that it satisfies $\theta'(0) = 0$. Now, the conditions for ε_t -

stationarity of $\text{NLP}^{SU}(t)$ read as follows:

$$\begin{aligned} \left\| \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \gamma^t \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \nu^t \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \delta^t \alpha^t \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \delta^t \beta^t \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\|_{\infty} &\leq \varepsilon_t, \\ x_1^t \geq -\varepsilon_t, \gamma^t &\geq -\varepsilon_t, |\gamma^t x_1^t| \leq \varepsilon_t, \\ x_2^t \geq -\varepsilon_t, \nu^t &\geq -\varepsilon_t, |\nu^t x_2^t| \leq \varepsilon_t, \\ \Phi^{SU}(x^t; t) \leq \varepsilon_t, \delta^t &\geq -\varepsilon_t, |\delta^k \Phi^{SU}(x^t; t)| \leq \varepsilon_t. \end{aligned}$$

Let us choose an arbitrary $t > 0$ and let $x^t = (x_1^t, x_2^t)$ be the unique point in the positive orthant satisfying $x_1^t = x_2^t$ and $\Phi^{SU}(x^t; t) = 0$. Then it is easy to see that $x^t \geq 0, x^t \rightarrow 0$ for $t \downarrow 0$, and $x_1^t, x_2^t \leq t$. Furthermore, taking into account that $\theta'(0) = 0$, it follows that $\alpha^t = \beta^t = 1$ for all $t > 0$. Therefore, if we take $\delta^t := 1, \gamma^t := 0, \nu^t := 2$, and $\varepsilon_t := 2t$ for all $t > 0$, it is easy to see that the tuple $(x^t, \gamma^t, \nu^t, \delta^t)$ satisfies all ε_t -stationarity conditions. Furthermore, assumption (8) also holds at the origin $x^* := (0, 0)$. We have $x^t \rightarrow x^*$ as $t \downarrow 0$ and x^* is easily seen to be weakly stationary, but not C-stationary. The only assumption that is violated from Theorem 5.4 is the condition $\varepsilon_t = o(t)$ since here we only have $\varepsilon_t = O(t)$. \diamond

The following counterexample shows that we also have to use an additional requirement like (8) in order to obtain convergence to C-stationary points.

Example 5.6 Let us consider once again the MPEC from (9). To define $\text{NLP}^{SU}(t)$, we take once again the regularization function θ from (10). The construction of this counterexample is a bit more tricky, and we therefore begin with some preliminary observations. To this end, let us define the mapping

$$h(y) := \frac{y - \theta(y)}{1 - \theta'(y)} \quad \text{for } y \in [0, 1].$$

Formally, this mapping is not well defined for $y = 1$ since $\theta'(1) = 1$. However, an easy calculation shows that we have

$$h(y) = \frac{(y+3)(y-1)}{4(y+2)},$$

so that the zero of the denominator cancels out. The last expression of h implies that

$$h'(y) = \frac{1}{4} + \frac{3}{4(y+2)^2} > 0 \quad \forall y \in [0, 1].$$

Hence $h : [0, 1] \rightarrow [-3/8, 0]$ is strictly increasing. Consequently, for each $t \in (0, 3/8)$, there exists exactly one $y^t \in (0, 1)$ such that $h(y^t) = -t$, and it follows that $\lim_{t \downarrow 0} y^t = 1$.

Now, let us come back to the example from (9). For any $t \in (0, 3/8)$ and the corresponding y^t constructed as above, we define

$$x^t := (x_1^t, x_2^t) := (ty^t, 0), \quad \gamma^t := 0, \quad \nu^t := 1 + \frac{1 + \theta'(y^t)}{1 - \theta'(y^t)}, \quad \delta^t := \frac{1}{1 - \theta'(y^t)}, \quad \varepsilon_t := t^2.$$

Then we always have $|x_1^t - x_2^t| = ty^t < t$ and, therefore

$$\begin{aligned}
x_1^t &\geq 0, & \gamma^t &\geq 0, & x_1^t \gamma^t &\leq \varepsilon_t, \\
x_2^t &\geq 0, & \nu^t &\geq 1 \geq 0, & x_2^t \nu^t &\leq \varepsilon_t, \\
\Phi^{SU}(x^t; t) &= x_1^t + x_2^t - t\theta\left(\frac{x_1^t - x_2^t}{t}\right) = t(y^t - \theta(y^t)) &\leq 0, \\
&& \delta^t &\geq 0, \\
|\delta^t \Phi^{SU}(x^t; t)| &= \left| \delta^t \left(x_1^t + x_2^t - t\theta\left(\frac{x_1^t - x_2^t}{t}\right) \right) \right| \\
&= \left| \frac{t(y^t - \theta(y^t))}{1 - \theta'(y^t)} \right| = |th(y^t)| = t^2 \leq \varepsilon_t, \\
\left| -1 - \gamma^t + \delta^t \left(1 - \theta'\left(\frac{x_1^t - x_2^t}{t}\right) \right) \right| &= \left| -1 + \frac{1 - \theta'(y^t)}{1 - \theta'(y^t)} \right| \leq \varepsilon_t, \\
\left| 1 - \nu^t + \delta^t \left(1 + \theta'\left(\frac{x_1^t - x_2^t}{t}\right) \right) \right| &= \left| 1 - 1 - \frac{1 + \theta'(y^t)}{1 - \theta'(y^t)} + \frac{1 + \theta'(y^t)}{1 - \theta'(y^t)} \right| \leq \varepsilon_t.
\end{aligned}$$

In particular, all ε_t -stationarity conditions are satisfied. Furthermore, we have $\varepsilon_t = t^2 = o(t)$. Nevertheless, $\lim_{t \downarrow 0} (x_1^t, x_2^t) = (0, 0)$ is not a C-stationary point. This does not contradict Theorem 5.4 since condition (8) is violated for the particular sequence $\{x^t\}$ constructed in this example: It is easy to verify that $\alpha^t \rightarrow 0$ for $t \rightarrow 0$. \diamond

6 The Nonsmooth Relaxation by Kadrani et al.

While the previous approaches all proposed to smooth the kink in the feasible set, Kadrani et al. [19] suggest to replace the complementarity conditions by the inequalities

$$G_i(x) + t \geq 0, \quad H_i(x) + t \geq 0, \quad \Phi_i^{KDB}(x; t) := (G_i(x) - t)(H_i(x) - t) \leq 0 \quad \forall i = 1, \dots, q$$

with a relaxation parameter $t > 0$. Note that Φ_i^{KDB} is continuously differentiable with gradient

$$\nabla \Phi_i^{KDB}(x; t) := (G_i(x) - t)\nabla H_i(x) + (H_i(x) - t)\nabla G_i(x) \quad \forall i = 1, \dots, q.$$

This relaxation transforms the feasible set of the complementarity constraints to the nonsmooth shape depicted in Figure 6.

The relaxation leads to the following sequence of relaxed nonlinear programs

$$\begin{aligned}
\min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 && \forall i = 1, \dots, m, \\
& h_i(x) = 0 && \forall i = 1, \dots, p, \\
& G_i(x) + t \geq 0 && \forall i = 1, \dots, q, \\
& H_i(x) + t \geq 0 && \forall i = 1, \dots, q, \\
& \Phi_i^{KDB}(x; t) = (G_i(x) - t)(H_i(x) - t) \leq 0 && \forall i = 1, \dots, q
\end{aligned} \tag{11}$$

with $t \downarrow 0$, which we will denote by $\text{NLP}^{KDB}(t)$. The following convergence result is known from reference [19].

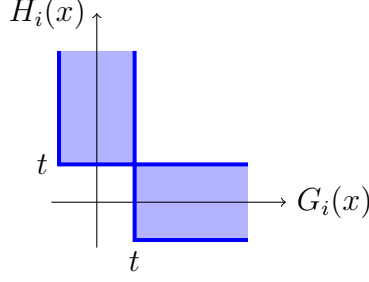


Figure 6: Geometric interpretation of the relaxation by Kadrani et al.

Theorem 6.1 *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 $NLP^{KDB}(t_k)$. If $x^k \rightarrow x^*$ and MPEC-LICQ holds in x^* , then x^* is an M-stationary point of the MPEC (1).*

The same result was subsequently shown to hold under the much weaker MPEC-CPLD condition in [16]. Hence, this relaxation method is among those with the strongest convergence properties known so far: All limit points are M-stationary, whereas all previously discussed relaxation methods are guaranteed to converge to C-stationary points only. However, when we replace KKT-points with ε -stationary points, we lose most of this advantage. Strictly speaking, this is in contrast to the statement in [19] where the authors claim that each limit point of certain ε_k -stationary points (ε_k -stationarity in [19] is defined in a slightly different way, requiring stronger conditions than in our setting) are M-stationary. However, as already observed in [21], the proof is erroneous, and a counterexample given in [21] shows that the statement itself is indeed not true even under the stronger ε_k -stationarity conditions from [19]. In fact, without any additional assumptions, we only get convergence to weakly stationary points, as shown in the following result.

Theorem 6.2 *Let $\{t_k\} \downarrow 0$, $\{\varepsilon_k\} \downarrow 0$, $\{x^k\}$ be a sequence of ε_k -stationary points of $NLP^{KDB}(t_k)$, and assume that $x^k \rightarrow x^*$. Then x^* is a weakly stationary point of the MPEC provided that MPEC-MFCQ holds at x^* .*

Proof: Since x^k is an ε_k -stationary point of $NLP^{KDB}(t_k)$, there exist multipliers $(\lambda^k, \mu^k, \gamma^k, \nu^k, \delta^k)$ such that

$$\left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^q \delta_i^k \nabla \Phi_i^{KDB}(x^k; t_k) \right\|_{\infty} \leq \varepsilon_k$$

with

$$\begin{aligned} g_i(x^k) &\leq \varepsilon_k, & \lambda_i^k &\geq -\varepsilon_k, & |\lambda_i^k g_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, m, \\ |h_i(x^k)| &\leq \varepsilon_k & & & & & \forall i = 1, \dots, p, \\ G_i(x^k) + t_k &\geq -\varepsilon_k, & \gamma_i^k &\geq -\varepsilon_k, & |\gamma_i^k (G_i(x^k) + t_k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\ H_i(x^k) + t_k &\geq -\varepsilon_k, & \nu_i^k &\geq -\varepsilon_k, & |\nu_i^k (H_i(x^k) + t_k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\ \Phi_i^{KDB}(x^k; t_k) &\leq \varepsilon_k, & \delta_i^k &\geq -\varepsilon_k, & |\delta_i^k \Phi_i^{KDB}(x^k; t_k)| &\leq \varepsilon_k & \forall i = 1, \dots, q. \end{aligned}$$

Hence, the limit x^* is at least feasible for the MPEC. Taking into account the expression for the gradient of Φ_i^{KDB} and setting

$$\tilde{\gamma}_i^k := \gamma_i^k - \delta_i^k(H_i(x^k) - t_k), \quad \tilde{\nu}_i^k := \nu_i^k - \delta_i^k(G_i(x^k) - t_k) \quad \forall i = 1, \dots, q,$$

we can rewrite the first inequality as

$$\left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \tilde{\gamma}_i^k \nabla G_i(x^k) - \sum_{i=1}^q \tilde{\nu}_i^k \nabla H_i(x^k) \right\|_{\infty} \leq \varepsilon_k.$$

We claim that the sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ stays bounded. Otherwise, we may assume without loss of generality that

$$\frac{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)}{\|(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\|} \rightarrow (\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) \neq 0.$$

Dividing the inequality by $\|(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\|$ and taking $k \rightarrow \infty$ then yields

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

From the ε_k -stationarity, we obtain $\bar{\lambda}_i \geq 0$ for all $i = 1, \dots, m$, and $\bar{\lambda}_i > 0$ implies $g_i(x^*) = 0$, hence $\text{supp}(\bar{\lambda}) \subseteq I_g(x^*)$.

We next want to show that $\text{supp}(\bar{\gamma}) \subseteq I_{00}(x^*) \cup I_{0+}(x^*)$. Suppose that there exists an index i such that $\bar{\gamma}_i \neq 0$ and $i \in I_{+0}(x^*)$. Then we have $G_i(x^*) > 0$ and $H_i(x^*) = 0$. This implies $G_i(x^k) \pm t_k \rightarrow G_i(x^*) > 0$ and thus, by the ε_k -stationarity, $\gamma_i^k \rightarrow 0$. At the same time, the ε_k -stationarity implies

$$|\delta_i^k \Phi_i^{KDB}(x^k; t_k)| = |\delta_i^k (H_i(x^k) - t_k)(G_i(x^k) - t_k)| \rightarrow 0$$

and consequently $\delta_i^k (H_i(x^k) - t_k) \rightarrow 0$. Combined, this gives us $\tilde{\gamma}_i^k \rightarrow 0$. Consequently, we have $\bar{\gamma}_i = 0$ in contrast to our choice of the index i .

By a symmetric argument, we can show $\text{supp}(\bar{\nu}) \subseteq I_{00}(x^*) \cup I_{+0}(x^*)$. The above equation therefore reduces to

$$\sum_{i \in I_g} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^p \bar{\mu}_i \nabla h_i(x^*) - \sum_{i \in I_{00} \cup I_{0+}} \bar{\gamma}_i \nabla G_i(x^*) - \sum_{i \in I_{00} \cup I_{+0}} \bar{\nu}_i \nabla H_i(x^*) = 0.$$

MPEC-MFCQ now implies $\bar{\lambda}_i = 0$ ($i \in I_g$), $\bar{\mu}_i = 0$ ($i = 1, \dots, p$), $\bar{\gamma}_i = 0$ ($i \in I_{00} \cup I_{0+}$), and $\bar{\nu}_i = 0$ ($i \in I_{00} \cup I_{+0}$). Since all the other components were already shown to be zero, we get a contradiction to the fact that $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) \neq 0$.

Hence the sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ is bounded. Without loss of generality, we may assume that the entire sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ converges to a limit $(\lambda^*, \mu^*, \gamma^*, \nu^*)$. But this limit is weakly stationary since it is easy to see that the multipliers $\lambda^*, \mu^*, \gamma^*, \nu^*$ have the same properties as $\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}$. \square

As pointed out in [21], it is indeed possible to provide examples where a sequence of ε_k -stationary points generated by the relaxation method by Kadrani et al. converges to a weakly stationary point which is not even C-stationary. Hence it is not possible to expect more than what is shown in the previous proof.

We next want to verify that one may get stronger stationarity of the limit under additional assumptions regarding the way the sequence $\{x^k\}$ is computed.

Theorem 6.3 Let $\{t_k\} \downarrow 0$, $\varepsilon_k = o(t_k)$, $\{x^k\}$ be a sequence of ε_k -stationary points of $NLP^{KDB}(t_k)$, and assume that $x^k \rightarrow x^*$ with MPEC-MFCQ holding in x^* . Suppose further that there is a constant $c > 0$ such that, for all $i \in I_{00}(x^*)$ and all k sufficiently large,

(a) the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin \begin{aligned} &(t_k, (1+c)t_k) \times ((1-c)t_k, t_k) \\ &\cup ((1-c)t_k, t_k) \times (t_k, (1+c)t_k). \end{aligned} \quad (12)$$

Then x^* is a C-stationary point of the MPEC.

(b) the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin \begin{aligned} &(t_k, (1+c)t_k) \times ((1-c)t_k, t_k) \\ &\cup ((1-c)t_k, t_k) \times (t_k, (1+c)t_k) \\ &\cup (t_k, (1+c)t_k)^2. \end{aligned} \quad (13)$$

Then x^* is an M-stationary point of the MPEC.

(c) the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin \begin{aligned} &((1-c)t_k, (1+c)t_k) \times (t_k, \infty) \\ &\cup (t_k, \infty) \times ((1-c)t_k, (1+c)t_k). \end{aligned} \quad (14)$$

Then x^* is an S-stationary point of the MPEC.

Proof: We know from Theorem 6.2 that x^* , together with suitable multipliers $(\lambda^*, \mu^*, \gamma^*, \nu^*)$, is at least weakly stationary.

(a) In order to verify C-stationarity, it remains to show that $\gamma_i^* \nu_i^* \geq 0$ holds for all $i \in I_{00}(x^*)$. By contradiction, without loss of generality, let us assume that there is an index $i \in I_{00}(x^*)$ such that $\gamma_i^* < 0$ and $\nu_i^* > 0$. From the previous proof, it then follows that

$$\gamma_i^* = \lim_{k \rightarrow \infty} (\gamma_i^k - \delta_i^k (H_i(x^k) - t_k)) < 0$$

(possibly on a subsequence). Since $\liminf_k \gamma_i^k \geq 0$ due to the ε_k -stationarity, it follows that there is a constant $C > 0$ with $\delta_i^k (H_i(x^k) - t_k) \geq C$ for all k sufficiently large. Due to $H_i(x^k) - t_k \rightarrow 0$, this implies $|\delta_i^k| \rightarrow \infty$. From the ε_k -stationarity, we know that only $\delta_i^k \rightarrow \infty$ is possible and hence $H_i(x^k) > t_k$ holds for all k sufficiently large. From this and $\varepsilon_k = o(t_k)$, we can infer

$$\nu_i^k \left(\frac{H_i(x^k)}{t_k} + 1 \right) = \nu_i^k \left(\frac{H_i(x^k) + t_k}{t_k} \right) \rightarrow 0 \implies \nu_i^k \rightarrow 0.$$

The same reasoning combined with $\delta_i^k (H_i(x^k) - t_k) \geq C$ gives us

$$\delta_i^k (H_i(x^k) - t_k) \left(\frac{G_i(x^k)}{t_k} - 1 \right) = \frac{\delta_i^k \Phi_i^{KDB}(x^k; t_k)}{t_k} \rightarrow 0 \implies \frac{G_i(x^k)}{t_k} \rightarrow 1,$$

i.e. $G_i(x^k) \in ((1-c)t_k, (1+c)t_k)$ for all k sufficiently large. If we take a look at the same limit from a different angle, we also get

$$\delta_i^k (G_i(x^k) - t_k) \left(\frac{H_i(x^k)}{t_k} - 1 \right) = \frac{\delta_i^k \Phi_i^{KDB}(x^k; t_k)}{t_k} \rightarrow 0.$$

Now, let us use assumption (12). Due to $H_i(x^k) > t_k$ and $G_i(x^k) \in ((1-c)t_k, (1+c)t_k)$ for all k sufficiently large, either $H_i(x^k) \geq (1+c)t_k$ or $G_i(x^k) \geq t_k$ or both has to hold for all k sufficiently large. At least one of the two possibilities has to be satisfied on a whole subsequence $K \subseteq \mathbb{N}$. In the first case, it follows that $\delta_i^k(G_i(x^k) - t_k) \rightarrow_K 0$ and thus

$$\nu_i^* = \lim_{k \rightarrow_K \infty} (\nu_i^k - \delta_i^k(G_i(x^k) - t_k)) = 0.$$

In the second case, we get at least $\nu_i^* \leq 0$. In both cases, we therefore get a contradiction to the assumption $\nu_i^* > 0$. Hence x^* is a C-stationary point of the MPEC.

(b) Since (13) implies (12), it follows from statement (a) that x^* is at least C-stationary. To see that x^* is M-stationary, it remains to show that the following two implications hold for every index $i \in I_{00}(x^*)$:

$$(\gamma_i^* < 0 \implies \nu_i^* = 0) \quad \text{and} \quad (\nu_i^* < 0 \implies \gamma_i^* = 0).$$

Without loss of generality, let us only consider the first case $\gamma_i^* < 0$ for some index $i \in I_{00}(x^*)$. Following the proof of part (a), we can deduce that

$$\begin{aligned} H_i(x^k) > t_k, \quad \nu_i^k \rightarrow 0, \quad \delta_i^k(G_i(x^k) - t_k) \left(\frac{H_i(x^k)}{t_k} - 1 \right) \rightarrow 0, \\ G_i(x^k) \in ((1-c)t_k, (1+c)t_k) \quad \text{for all } k \text{ sufficiently large.} \end{aligned}$$

Now, using assumption (13), there are only two possibilities left, namely $H_i(x^k) \geq (1+c)t_k$, which implies $\nu_i^* = 0$ in the same way as in the proof of part (a), and $G_i(x^k) = t_k$, which also gives $\nu_i^* = 0$. Thus, in each of these two cases, x^* is an M-stationary point of the MPEC.

(c) Here we have to show that $\gamma_i^* \geq 0$ and $\nu_i^* \geq 0$ holds for all $i \in I_{00}(x^*)$. Assume, again without loss of generality, that we have $\gamma_i^* < 0$ for some index $i \in I_{00}(x^*)$. Following the proof of part (a), this implies $H_i(x^k) > t_k$ for all k sufficiently large and $\frac{G_i(x^k)}{t_k} \rightarrow 1$. This, however, contradicts assumption (14). Consequently, x^* is an S-stationary point of the MPEC. \square

Figure 7 illustrates the additional conditions from Theorem 6.3, which require that the iterates $(G_i(x^k), H_i(x^k))$ do not lie in any of the marked areas for all $i \in I_{00}(x^*)$ and all k sufficiently large.

7 The L-Shaped Relaxation from [21]

To keep the strong convergence properties of the relaxation by Kadrani et al. but obtain a more favorable feasible set, the authors suggested in [21] to replace the complementarity conditions by the inequalities

$$G_i(x^k) \geq 0, \quad H_i(x^k) \geq 0, \quad \Phi_i^{KS}(x; t) \leq 0 \quad \forall i = 1, \dots, q$$

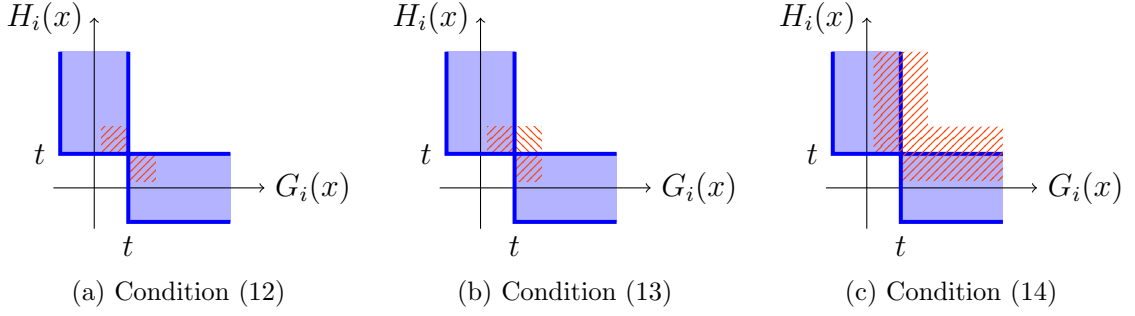


Figure 7: Geometric interpretation of the conditions from Theorem 6.3

with a relaxation parameter $t > 0$. The function Φ_i^{KS} is defined piecewise as follows

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

This transforms the feasible set of the complementarity constraints to the shape depicted in Figure 8.

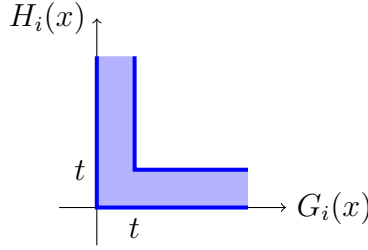


Figure 8: Geometric interpretation of the relaxation by Kanzow and Schwartz

The relaxation leads to the following sequence of relaxed nonlinear programs

$$\begin{aligned} \min f(x) \quad \text{s.t.} \quad & g_i(x) \leq 0 & \forall i = 1, \dots, m, \\ & h_i(x) = 0 & \forall i = 1, \dots, p, \\ & G_i(x) \geq 0 & \forall i = 1, \dots, q, \\ & H_i(x) \geq 0 & \forall i = 1, \dots, q, \\ & \Phi_i^{KS}(x; t) \leq 0 & \forall i = 1, \dots, q \end{aligned} \tag{15}$$

with $t \downarrow 0$, which we will denote by $\text{NLP}^{KS}(t)$. From the report version [20] of the paper [21], the following convergence result is known.

Theorem 7.1 *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 $\text{NLP}^{KS}(t_k)$. If $x^k \rightarrow x^*$ and MPEC-CPLD holds in x^* , then x^* is an M-stationary point of the MPEC (1).*

Additionally, the following result for inexactly solved subproblems was derived in [21]. Note, however, that this paper uses a slightly stronger definition of ε -stationary points than the one applied here.

Theorem 7.2 Let $\{t_k\} \downarrow 0$, $\varepsilon_k = o(t_k)$, $\{x^k\}$ be a sequence of ε_k -stationary points of $NLP^{KS}(t_k)$, and assume that $x^k \rightarrow x^*$. If MPEC-LICQ holds in x^* , then x^* is a weakly stationary point of the MPEC.

If, additionally, there is a subsequence $K \subseteq \mathbb{N}$ such that

$$G_i(x^k) \leq t_k, \quad H_i(x^k) \leq t_k \quad \forall k \in K, \quad \forall i \in I_{00}(x^*) \quad (16)$$

holds, then x^* is a C -stationary point of the MPEC.

Now, we will improve this result by proving it for the more general definition of ε -stationary points used here and by refining the additional assumption needed to guarantee that the limit point is more than weakly stationary.

Theorem 7.3 Let $\{t_k\} \downarrow 0$, $\{\varepsilon_k\} \downarrow 0$, $\{x^k\}$ be a sequence of ε_k -stationary points of $NLP^{KS}(t_k)$, and assume that $x^k \rightarrow x^*$ with MPEC-MFCQ holding in x^* . Then x^* is a weakly stationary point of the MPEC.

Suppose further that $\varepsilon_k = o(t_k)$ and there is a constant $c > 0$ such that, for all $i \in I_{00}(x^*)$ and all k sufficiently large,

(a) the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin (t_k, (1+c)t_k) \times ((1-c)t_k, t_k) \cup ((1-c)t_k, t_k) \times (t_k, (1+c)t_k). \quad (17)$$

Then x^* is a C -stationary point of the MPEC.

(b) the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin (t_k, (1+c)t_k) \times ((1-c)t_k, t_k) \cup ((1-c)t_k, t_k) \times (t_k, (1+c)t_k) \cup (t_k, (1+c)t_k)^2 \cup ((1-c)t_k, t_k)^2. \quad (18)$$

Then x^* is an M -stationary point of the MPEC.

(c) the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin (((1-c)t_k, (1+c)t_k) \times ((1-c)t_k, \infty) \cup ((1-c)t_k, \infty) \times ((1-c)t_k, (1+c)t_k)) \setminus \{(t_k, t_k)\}. \quad (19)$$

Then x^* is an S -stationary point of the MPEC.

Proof: Since all x^k are ε_k -stationary points of $NLP(t_k)$, we have

$$\left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \gamma_i^k \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^k \nabla H_i(x^k) + \sum_{i=1}^q \delta_i^k \nabla \Phi_i^{KS}(x^k; t_k) \right\|_{\infty} \leq \varepsilon_k$$

with

$$\begin{aligned} g_i(x^k) &\leq \varepsilon_k, & \lambda_i^k &\geq -\varepsilon_k, & |\lambda_i^k g_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, m, \\ |h_i(x^k)| &\leq \varepsilon_k & & & & & \forall i = 1, \dots, p, \\ G_i(x^k) &\geq -\varepsilon_k, & \gamma_i^k &\geq -\varepsilon_k, & |\gamma_i^k G_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\ H_i(x^k) &\geq -\varepsilon_k, & \nu_i^k &\geq -\varepsilon_k, & |\nu_i^k H_i(x^k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \\ \Phi_i^{KS}(x^k; t_k) &\leq \varepsilon_k, & \delta_i^k &\geq -\varepsilon_k, & |\delta_i^k \Phi_i^{KS}(x^k; t_k)| &\leq \varepsilon_k & \forall i = 1, \dots, q, \end{aligned}$$

where we $\Phi_i^{KS}(x^k; t_k)$ is defined as before with the gradient

$$\nabla \Phi_i^{KS}(x^k; t_k) = \begin{cases} (H_i(x^k) - t_k) \nabla G_i(x^k) + (G_i(x^k) - t_k) \nabla H_i(x^k) & \text{if } G_i(x^k) + H_i(x^k) \geq 2t_k, \\ -(G_i(x^k) - t_k) \nabla G_i(x^k) - (H_i(x^k) - t_k) \nabla H_i(x^k) & \text{else.} \end{cases}$$

Hence, the limit x^* is obviously feasible for the MPEC (1). We define the multipliers

$$\begin{aligned} \tilde{\gamma}_i^k &:= \begin{cases} \gamma_i^k - \delta_i^k (H_i(x^k) - t_k) & \text{if } G_i(x^k) + H_i(x^k) \geq 2t_k, \\ \gamma_i^k + \delta_i^k (G_i(x^k) - t_k) & \text{else,} \end{cases} \\ \tilde{\nu}_i^k &:= \begin{cases} \nu_i^k - \delta_i^k (G_i(x^k) - t_k) & \text{if } G_i(x^k) + H_i(x^k) \geq 2t_k, \\ \nu_i^k + \delta_i^k (H_i(x^k) - t_k) & \text{else.} \end{cases} \end{aligned}$$

Then we have

$$\left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^q \tilde{\gamma}_i^k \nabla G_i(x^k) - \sum_{i=1}^q \tilde{\nu}_i^k \nabla H_i(x^k) \right\|_{\infty} \leq \varepsilon_k.$$

We claim that the multipliers $(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)$ are bounded. If the sequence were unbounded, we could assume without loss of generality convergence of the sequence

$$\frac{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)}{\|(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\|} \rightarrow (\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) \neq 0.$$

Then the ε_k -stationarity of x^k yields

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

Additionally, the ε_k -stationarity yields $\bar{\lambda}_i \geq 0$ for all $i = 1, \dots, m$, and $\bar{\lambda}_i > 0$ implies $g_i(x^*) = 0$, hence $\text{supp}(\bar{\lambda}) \subseteq I_g(x^*)$.

Now consider an $i \in I_{+0}(x^*)$. This implies $G_i(x^k) + H_i(x^k) \geq 2t_k$ for all k sufficiently large and thus $\tilde{\gamma}_i^k = \gamma_i^k - \delta_i^k (H_i(x^k) - t_k)$. The ε_k -stationarity yields $\gamma_i^k G_i(x^k) \rightarrow 0$, hence $\gamma_i^k \rightarrow 0$, and

$$\delta_i^k \Phi_i^{KS}(x^k; t_k) = \delta_i^k (H_i(x^k) - t_k) (G_i(x^k) - t_k) \rightarrow 0,$$

thus $\delta_i^k (H_i(x^k) - t_k) \rightarrow 0$. Consequently, we have $\bar{\gamma}_i = 0$. This shows that $\text{supp}(\bar{\gamma}) \subseteq I_{00}(x^*) \cup I_{0+}(x^*)$. By a symmetric argument, we obtain $\text{supp}(\bar{\nu}) \subseteq I_{00}(x^*) \cup I_{+0}(x^*)$. Thus, the equation above reduces to

$$\sum_{i \in I_g} \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^p \bar{\mu}_i \nabla h_i(x^*) - \sum_{i \in I_{00} \cup I_{0+}} \bar{\gamma}_i \nabla G_i(x^*) - \sum_{i \in I_{00} \cup I_{+0}} \bar{\nu}_i \nabla H_i(x^*) = 0$$

with $\bar{\lambda}_i \geq 0$ for all $i \in I_g(x^*)$. Hence MPEC-MFCQ implies $\bar{\lambda}_i = 0$ ($i \in I_g$), $\bar{\mu}_i = 0$ ($i = 1, \dots, p$), $\bar{\gamma}_i = 0$ ($i \in I_{00} \cup I_{0+}$), and $\bar{\nu}_i = 0$ ($i \in I_{00} \cup I_{+0}$). Altogether, we get a contradiction to the fact that $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) \neq 0$.

Hence the sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ is bounded. Without loss of generality, we can assume that the entire sequence $\{(\lambda^k, \mu^k, \tilde{\gamma}^k, \tilde{\nu}^k)\}$ converges to a limit

$(\lambda^*, \mu^*, \gamma^*, \nu^*)$. The limit is then weakly stationary since the multipliers $\lambda^*, \mu^*, \gamma^*, \nu^*$ have the same properties as $\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}$.

(a) To show that x^* is C-stationary under the additional assumption from (a), it suffices to show that $\gamma_i^* \nu_i^* \geq 0$ for all $i \in I_{00}(x^*)$. To this end, it is enough to consider the case where $\gamma_i^* < 0$ or $\nu_i^* < 0$. Without loss of generality, let us consider the case where $\gamma_i^* < 0$ for some index $i \in I_{00}(x^*)$. We will show that this implies $\nu_i^* \leq 0$ by considering the following two possible cases separately.

Case 1: There is a subsequence $K \subseteq \mathbb{N}$ with $G_i(x^k) + H_i(x^k) \geq 2t_k$ for all $k \in K$. Then we have $\gamma_i^* = \lim_{k \rightarrow K\infty} (\gamma_i^k - \delta_i^k (H_i(x^k) - t_k)) < 0$, which, due to $\liminf_{k \rightarrow \infty} \gamma_i^k \geq 0$, implies $\delta_i^k (H_i(x^k) - t_k) \geq C$ for a constant $C > 0$ and all $k \in K$ sufficiently large. Because of $H_i(x^k) - t_k \rightarrow 0$, it follows that $|\delta_i^k| \rightarrow_K \infty$. Thanks to the ε_k -stationarity, only $\delta_i^k \rightarrow_K \infty$ and thus $H_i(x^k) > t_k$ for all $k \in K$ sufficiently large is possible. Using this, the ε_k -stationarity, and $\varepsilon_k = o(t_k)$, we get

$$\nu_i^k \frac{H_i(x^k)}{t_k} \rightarrow 0 \quad \implies \quad \nu_i^k \rightarrow_K 0$$

and

$$\delta_i^k (H_i(x^k) - t_k) \left(\frac{G_i(x^k)}{t_k} - 1 \right) = \frac{\delta_i^k \Phi_i^{KS}(x^k; t_k)}{t_k} \rightarrow 0 \quad \implies \quad \frac{G_i(x^k)}{t_k} \rightarrow_K 1.$$

By assumption (17), $H_i(x^k) > t_k$ and $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$ is possible only if $H_i(x^k) \geq (1+c)t_k$ or $G_i(x^k) \geq t_k$. In the first case, the ε_k -stationarity implies

$$\delta_i^k (G_i(x^k) - t_k) \left(\frac{H_i(x^k)}{t_k} - 1 \right) \rightarrow 0 \quad \implies \quad \delta_i^k (G_i(x^k) - t_k) \rightarrow_K 0$$

and thus

$$\nu_i^* = \lim_{k \rightarrow K\infty} (\nu_i^k - \delta_i^k (G_i(x^k) - t_k)) = 0.$$

In the second case, we get at least $\nu_i^* \leq 0$.

Case 2: There is a subsequence $K \subseteq \mathbb{N}$ with $G_i(x^k) + H_i(x^k) < 2t_k$ for all $k \in K$. Then we have $\gamma_i^* = \lim_{k \rightarrow K\infty} (\gamma_i^k + \delta_i^k (G_i(x^k) - t_k)) < 0$, which, due to $\liminf_{k \rightarrow \infty} \gamma_i^k \geq 0$, implies $\delta_i^k (G_i(x^k) - t_k) \leq -C$ for a constant $C > 0$ and all $k \in K$ sufficiently large. Analogously to the previous case, this implies $\delta_i^k \rightarrow_K \infty$ and $G_i(x^k) < t_k$ for all $k \in K$ sufficiently large. Furthermore, we note that the ε_k -stationarity and $\varepsilon_k = o(t_k)$ implies, in this case,

$$\begin{aligned} \left| \delta_i^k (G_i(x^k) - t_k) \left(\frac{G_i(x^k)}{t_k} - 1 \right) \right| &= \left| \delta_i^k \frac{(G_i(x^k) - t_k)^2}{t_k} \right| \\ &\leq \left| \delta_i^k \frac{(G_i(x^k) - t_k)^2 + (H_i(x^k) - t_k)^2}{t_k} \right| \quad (20) \\ &= 2 \left| \delta_i^k \frac{\Phi_i^{KS}(x^k; t_k)}{t_k} \right| \rightarrow 0, \end{aligned}$$

so that the previous discussion immediately yields $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$. We claim that this implies $\frac{H_i(x^k)}{t_k} \rightarrow_K 1$. Due to $G_i(x^k) + H_i(x^k) < 2t_k$ for all $k \in K$, we know already

$$\frac{H_i(x^k)}{t_k} \leq 2 - \frac{G_i(x^k)}{t_k} \rightarrow_K 2 - 1 = 1$$

and from the ε_k -stationarity and $\varepsilon_k = o(t_k)$ we get

$$\frac{H_i(x^k)}{t_k} \geq -\frac{\varepsilon_k}{t_k} \rightarrow 0.$$

Now assume that there was a subsequence $K' \subseteq K$ with $\frac{H_i(x^k)}{t_k} \rightarrow_{K'} C'$ with $C' \in [0, 1)$. This would imply $H_i(x^k) \leq G_i(x^k)$ for all $k \in K'$ sufficiently large. Since we can also deduce

$$\delta_i^k(H_i(x^k) - t_k) \left(\frac{H_i(x^k)}{t_k} - 1 \right) \rightarrow 0$$

in a way analogous to (20), it then follows that $\delta_i^k(H_i(x^k) - t_k) \rightarrow_{K'} 0$. By combining these facts and $\delta_i^k \rightarrow_K \infty$, we would get

$$-C \geq \delta_i^k(G_i(x^k) - t_k) \geq \delta_i^k(H_i(x^k) - t_k) \rightarrow_{K'} 0,$$

a contradiction. Thus, we have proven $\frac{H_i(x^k)}{t_k} \rightarrow_K 1$. Next, let us consider $\nu_i^* = \lim_{k \rightarrow \infty} (\nu_i^k + \delta_i^k(H_i(x^k) - t_k))$. From the ε_k -stationarity, $\varepsilon_k = o(t_k)$, and $\frac{H_i(x^k)}{t_k} \rightarrow_K 1$, we can infer $\nu_i^k \rightarrow_K 0$. Due to assumption (17) and $G_i(x^k) < t_k$, $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$, we know $H_i(x^k) \leq t_k$ for all $k \in K$ sufficiently large. Together, this implies

$$\nu_i^* = \lim_{k \rightarrow \infty} (\nu_i^k + \delta_i^k(H_i(x^k) - t_k)) \leq 0.$$

Thus, both cases guarantee $\nu_i^* \leq 0$, hence $\gamma_i^* \nu_i^* \geq 0$ holds, i.e. x^* is a C-stationary point of the MPEC.

(b) To prove M-stationarity, it suffices to show that $\gamma_i^* < 0$ implies $\nu_i^* = 0$ for any index $i \in I_{00}(x^*)$, since a similar argument can be used to see that $\nu_i^* < 0$ implies $\gamma_i^* = 0$. To verify this statement, we can follow the proof of part (a) and therefore only state the differences.

Case 1: Under assumption (18), $H_i(x^k) > t_k$ and $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$ is possible only if either $H_i(x^k) \geq (1+c)t_k$, which immediately gives $\nu_i^* = 0$, or $G_i(x^k) = t_k$, which also implies $\nu_i^* = \lim_{k \rightarrow \infty} (\nu_i^k - \delta_i^k(G_i(x^k) - t_k)) = 0$.

Case 2: Under assumption (18), $G_i(x^k) < t_k$, $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$, and $\frac{H_i(x^k)}{t_k} \rightarrow_K 1$ is possible only if $H_i(x^k) = t_k$, which implies

$$\nu_i^* = \lim_{k \rightarrow \infty} (\nu_i^k + \delta_i^k(H_i(x^k) - t_k)) = 0.$$

Thus, both cases guarantee $\nu_i^* = 0$, i.e. x^* is an M-stationary point of the MPEC.

(c) Finally, to prove S-stationarity, we show that assumption (19) contradicts $\gamma_i^* < 0$ in both cases (as well as $\nu_i^* < 0$ by a symmetric argument).

Case 1: In the first case, $\gamma_i^* < 0$ implies $H_i(x^k) > t_k$ and $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$, a contradiction to (19).

Case 2: Here, $\gamma_i^* < 0$ implies $G_i(x^k) < t_k$, $\frac{G_i(x^k)}{t_k} \rightarrow_K 1$, and $\frac{H_i(x^k)}{t_k} \rightarrow_K 1$, a contradiction to (19). \square

Figure 9 illustrates the additional conditions from Theorem 7.3, which require that the iterates $(G_i(x^k), H_i(x^k))$ do not lie in any of the marked areas for all $i \in I_{00}(x^*)$ and all k sufficiently large.

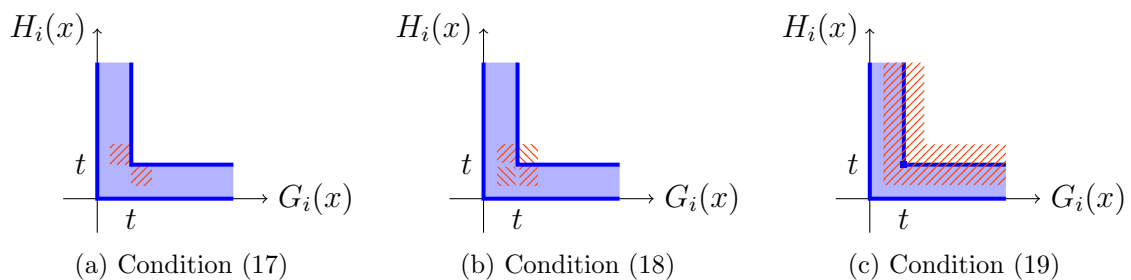


Figure 9: Geometric interpretation of the conditions from Theorem 7.3

Obviously, the assumptions needed here for M- and S-stationarity differ slightly from those used in Theorem 6.3. The necessity of this is illustrated in the following example.

Example 7.4 Consider the two-dimensional MPEC

$$\min -x_1 - x_2 \quad \text{s.t.} \quad 0 \leq x_1 \perp x_2 \geq 0$$

and sequences $t \downarrow 0$, $\varepsilon_t = t^2$. Then it is easy to verify that the points $x^t = ((1-t)t, (1-t)t)^T$ are ε_t -stationary points of $\text{NLP}^{KS}(t)$ with the multipliers $\gamma^t = 0$, $\nu^t = 0$, $\delta^t = \frac{1}{\varepsilon_t}$. On the other hand $x^t \rightarrow (0,0)^T$, which is a C-stationary point of the MPEC and satisfies even MPEC-LICQ, but is not an M-stationary point. \diamond

8 Final Remarks

This paper shows that the relaxation methods by Scholtes [34] and by Lin and Fukushima [24] (the latter is formally treated in the accompanying paper [22]) converge to C-stationary points even if the corresponding NLP-subproblems are only solved inexactly in the sense that an ε -stationary point is computed, whereas a corresponding analysis for the other relaxation methods from [19, 21, 35] yields weakly stationary points only. This is surprising since some of the other relaxation methods, namely [19, 21], have stronger convergence properties than the ones from [34] and [24] when exact KKT-points are computed for the NLP-subproblems.

Of course, the kind of stationarity that one can obtain depends very much on the notion of ε -stationarity. In this paper, we use a rather weak formulation of ε -stationarity since this seems to cover all interesting termination criteria used in existing NLP software, therefore making our results independent of any particular NLP-solver. Hence one might expect to get better convergence results for some inexact relaxation methods if a stronger notion of ε -stationarity is employed. We believe, however, that this is not true as long as this stronger notion of ε -stationarity is still a realistic one that can be implemented by a suitable solver as an inexact termination criterion. To this end, we note that the definition of ε -stationarity used in the two previous papers [19, 21] is, in fact, stronger, but that counterexample given in [21] indicates that both the nonsmooth relaxation scheme by Kadrani et al. and the the L-shaped relaxation method by the authors converge to weakly stationary points only even under this stronger notion of ε -stationarity. Moreover, we stress that the other counterexamples provided in this paper also satisfy much stronger conditions than those required by the definition of ε -stationarity here.

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] M.S. BAZARAA AND C.M. SHETTY: *Foundations of Optimization*. Lecture Notes in Economics and Mathematical Systems, Springer, 1976.
- [4] R. BYRD, M.E. HRIBAR, AND J. NOCEDAL: *An interior point method for large-scale nonlinear programming*. SIAM Journal on Optimization 9, 1999, pp. 877–900.
- [5] R. BYRD, J. NOCEDAL, AND R.A. WALTZ: *KNITRO: An integrated package for nonlinear optimization*. In G. DI PILLO AND M. ROMA (eds.): *Large-Scale Nonlinear Optimization*. Springer, 2006, pp. 35–59.
- [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.L. FLEGEL AND C. KANZOW: *On the Guignard constraint qualification for mathematical programs with equilibrium constraints*. Optimization 54, 2005, pp. 517–534.
- [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] R. FLETCHER AND S. LEYFFER: *Nonlinear programming without a penalty function*. Mathematical Programming 91, 2002, pp. 239–269.
- [12] R. FLETCHER AND S. LEYFFER: *Solving mathematical programs with complementarity constraints as nonlinear programs*. Optimization Methods and Software 19, 2004, pp. 15–40.
- [13] P.E. GILL, W. MURRAY, AND M.A. SAUNDERS: *SNOPT: An SQP algorithm for large-scale constrained optimization*. SIAM Journal on Optimization 12, 2002, pp. 979–1006.
- [14] T. HOHEISEL, C. KANZOW, AND A. SCHWARTZ: *Improved convergence properties of the Lin-Fukushima-regularization method for mathematical programs with complementarity constraints*. Numerical Algebra, Control, and Optimization 1, 2011, pp. 49–60.
- [15] T. HOHEISEL, C. KANZOW, AND A. SCHWARTZ: *Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints*. Optimization Methods and Software 27, 2012, pp. 483–512.
- [16] T. HOHEISEL, C. KANZOW, AND A. SCHWARTZ: *Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints*. Mathematical Programming 137, 2013, pp. 257–288.
- [17] 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.
- [18] A.F. IZMAILOV, A.L. POGOSYAN, AND M.V. SOLODOV: *Semismooth Newton method for the lifted reformulation of mathematical programs with complementarity constraints*. Computational Optimization and Applications 51, 2012, pp. 199–221.
- [19] 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.
- [20] 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, Germany, June 2010.

- [21] C. KANZOW AND A. SCHWARTZ: *A new regularization method for mathematical programs with complementarity constraints with strong convergence properties*. SIAM Journal on Optimization, to appear.
- [22] C. KANZOW AND A. SCHWARTZ: *Convergence properties of the inexact Lin-Fukushima relaxation method for mathematical programs with equilibrium constraints*. Preprint 314, Institute of Mathematics, University of Würzburg, Würzburg, Germany, March 2013.
- [23] 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.
- [24] 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.
- [25] Z.-Q. LUO, J.-S. PANG, AND D. RALPH: *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, New York, Melbourne, 1996.
- [26] J. NOCEDAL AND S.J. WRIGHT: *Numerical Optimization*. Springer, second edition 2006.
- [27] J.V. OUTRATA: *Optimality conditions for a class of mathematical programs with equilibrium constraints*. Mathematics of Operations Research 24, 1999, pp. 627–644.
- [28] J.V. OUTRATA: *A generalized mathematical program with equilibrium constraints*. SIAM Journal on Control and Optimization 38, 2000, pp. 1623–1638.
- [29] 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.
- [30] J.-S. PANG AND M. FUKUSHIMA: *Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints*. Computational Optimization and Applications 13, 1999, pp. 111–136.
- [31] A.U. RAGHUNATHAN AND L.T. BIEGLER: *An interior point method for mathematical programs with complementarity constraints (MPCCs)*. SIAM Journal on Optimization 15, 2005, pp. 720–750.
- [32] D. RALPH AND S.J. WRIGHT: *Some properties of regularization and penalization schemes for MPECs*. Optimization Methods and Software 19, 2004, pp. 527–556.
- [33] H. SCHEEL AND S. SCHOLTES: *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*. Mathematics of Operations Research 25, 2000, pp. 1–22.

- [34] S. SCHOLTES: *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*. SIAM Journal on Optimization 11, 2001, pp. 918–936.
- [35] S. STEFFENSEN AND M. ULBRICH: *A new regularization scheme for mathematical programs with equilibrium constraints*. SIAM Journal on Optimization 20, 2010, pp. 2504–2539.
- [36] O. STEIN: *Lifting mathematical programs with complementarity constraints*. Mathematical Programming 131, 2012, pp. 71–94.
- [37] S. VEELKEN: *A new relaxation scheme for mathematical programs with equilibrium constraints: Theory and numerical experience*. Ph.D. Thesis, Technical University of Munich, Germany, 2009.
- [38] A. WÄCHTER AND L.T. BIEGLER: *On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming*. Mathematical Programming 106, 2006, pp. 25–57.
- [39] 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.
- [40] J.J. YE: *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*. Journal of Mathematical Analysis and Applications 307, 2005, pp. 350–369.
- [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.
- [42] J.J. YE AND D.L. ZHU: *Optimality conditions for bilevel programming problems*. Optimization 33, 1995, pp. 9–27.