CONVERGENCE PROPERTIES OF THE INEXACT LIN-FUKUSHIMA RELAXATION METHOD FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

Christian Kanzow and Alexandra Schwartz

Preprint 314 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

Dedicated to Masao Fukushima, in great respect, on the occasion of his 65th birthday.

Abstract. Mathematical programs with equilibrium (or complementarity) constraints, MPECs for short, form a difficult class of optimization problems. The feasible set of MPECs is described by standard equality and inequality constraints as well as additional complementarity constraints that are used to model equilibrium conditions in different applications. But these complementarity constraints imply that MPECs violate most of the standard constraint qualifications. Therefore, more specialized algorithms are typically applied to MPECs that take into account the particular structure of the complementarity constraints. One popular class of these specialized algorithms are the relaxation (or regularization) methods. They replace the MPEC by a sequence of nonlinear programs NLP(t) depending on a parameter t, then compute a KKT-point of each NLP(t), and try to get a suitable stationary point of the original MPEC in the limit $t \to 0$. For most relaxation methods, one can show that a C-stationary point is obtained in this way, a few others even get M-stationary points, which is a stronger property. So far, however, these results have been obtained under the assumption that one is able to compute exact KKT-points of each NLP(t). But this assumption is not implementable, hence a natural question is: What kind of stationarity do we get if we only compute approximate KKT-points? It turns out that most relaxation methods only get a weakly stationary point under this assumption, while in this paper, we show that the smooth relaxation method by Lin and Fukushima [*Annals of Operations Research 133, 2005, pp. 63–84*] still yields a C-stationary point, i.e. the inexact version of this relaxation scheme has the same convergence properties as the exact counterpart.

Key Words: Mathematical programs with complementarity constraints, mathematical programs with equilibrium constraints, global convergence, KKT-points, stationary points, C-stationarity, weak stationarity, inexact relaxation methods, inexact regularization methods

AMS subject classifications. 65K05, 90C30, 90C31

1 Introduction

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

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

where $f, g_i, h_i, G_i, H_i : \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable. The MPEC has received a lot of attention during the last 15 years since it provides a useful model for several applications coming from different areas like economics, game theory, mechanics etc. For more details, we refer the interested reader to the two monographs [19, 24] and to the book [6] on bilevel programming, a subject that is closely related to MPECs.

In principle, the MPEC may be viewed as a standard nonlinear program, NLP for short. However, the feasible set has a very special structure since, apart from the usual equality and inequality constraints, there are additional complementarity constraints. The existence of the complementarity constraints cause some troubles since they imply that most constraint qualifications, typically used in the context of NLPs, are violated at any feasible point, see [19, 35]. This, in turn, means that one has to expect difficulties when solving MPECs by standard software.

For this reason, a number of MPEC-tailored solution methods have been devised that try to take into account the special structure of the additional complementarity constraints. Among the different approaches are penalty, smoothing, interior-point, lifting, and relaxation methods. The interested reader is refered to [1, 2, 5, 7, 9, 12, 13, 14, 15, 17, 18, 26, 27, 29, 30, 31] and references therein for more details.

Here our focus is on the class of relaxation (or regularization) methods for the solution of MPECs. The first relaxation method is due to Scholtes [29]. In the meantime, a number of other relaxation methods exist, among them are the smooth relaxation method by Lin and Fukushima [18], the local relaxation method by Steffensen and Ulbrich [30], the so-called nonsmooth relaxation method by Kadrani et al. [14], and the L-shaped relaxation method by the authors [15]. The basic idea of all these relaxation methods is the same: They approximate (usually enlarge) the feasible set of the MPEC in a suitable way to get a nonlinear program NLP(t) depending on a certain parameter t such that the relaxed programs NLP(t) converge to the original MPEC when $t \to 0$.

Algorithmically, one then considers a sequence $t_k \to 0$ and computes a sequence of KKT-points of the nonlinear programs NLP(t_k), hoping that this sequence then converges to a suitable stationary point of the underlying MPEC. Since different stationarity concepts exist for MPECs, one has to expect that different relaxation methods may converge to different kind of stationarities. This is indeed the case, since one can show (technical assumptions ahead) that the three methods by Scholtes [29], by Lin and Fukushima [18], and by Steffensen and Ulbrich [30] converge to so-called C-stationary points, whereas the two methods by Kadrani et al. [14] and by the authors [15] were designed in such a way that they converge to M-stationary points, a concept that is stronger than C-stationary. Hence these two methods have better convergence properties than the other three methods.

But the convergence to C- and M-stationary points depends on the assumption that exact KKT-points can be computed for each NLP(t_k). This assumption is, of course, unrealistic. Hence, a practically more important question is: What happens if we compute only approximate KKT-points of NLP(t_k)? Doing this means that the precise sign structure of the multipliers within the KKT-conditions of the relaxed problems NLP(t_k) is lost, so one has to expect a weaker stationarity in the limit $t_k \to 0$.

Most of the previous papers on relaxation methods for the solution of MPECs are dealing with exact KKT-conditions. An exception are the two papers [14] by Kadrani et al. and [15] by the authors. The former claims to get M-stationarity also in the case when only approximate KKT-points are computed, but the proof is erroneous, and a counterexample in [15] shows that one gets only weakly stationary points. The same happens for the inexact version of the L-shaped method from [15].

Though this result had to be expected, it is somewhat disappointing. We therefore took a closer look at all relaxation methods mentioned previously and investigated their limiting behaviour when only inexact KKT-points are computed. The overall result is quite surprising: The two best (for the case of exact KKT-points) methods from [14, 15] as well as the local relaxation method from Steffensen and Ulbrich [30] converge to weakly stationary points only, whereas for the original method from Scholtes [29] and the Lin-Fukushima relaxation method [18] one still obtains C-stationary points in the limit. Hence the three methods from [14, 15, 18] lose quite a bit of their convergence properties when moving from exact to inexact KKT-points, while nothing is lost by the two methods from [18, 29]. The proofs for most of these statements, together with some additional results on the two methods from [14, 15], can be found in the accompanying paper [16]. Since that paper is already quite long, and since, in any case, the treatment of the Lin-Fukushima relaxation needs some extra work, we decided to separate this method from the other ones and present the corresponding convergence result in this paper.

The organization is as follows: Section 2 presents some basic definitions like suitable constraint qualifications, stationarity concepts as well as our notion of an approximate KKT-point. Section 3 investigates the convergence behaviour of an inexact version of the smooth relaxation method by Lin and Fukushima [18] and proves that C-stationary points are obtained in the limit. Extensive numerical results for this method may already be found in [11] and are therefore not part of this paper. We then close with some final remarks in Section 4.

Notation: For a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, 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 $\operatorname{supp}(x) := \{i \in \{1, \dots, n\} \mid x_i \neq 0\}$.

2 Preliminaries

This section contains some background material from standard NLPs and MPECs. Let us begin with standard NLPs. As mentioned previously, the idea of any relaxation method is to relax the complementarity constraints and thus obtain a sequence of standard NLPs. For this reason, we need some notation and a few basic facts about NLPs. Therefore, let us consider the following nonlinear program

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

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

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 $\lambda \geq 0$, $\operatorname{supp}(\lambda) \subseteq I_g(x^*)$, see, e.g., [3, 21] for more details.

Unfortunately, when NLPs are solved numerically, one rarely ends up in a KKT-point. The termination criteria used in standard software checks (in addition to other things) whether the KKT-conditions are satisfied approximately. The precise way this is done might depend on the class of methods and also on the particular solver used. However, the following definition should be sufficiently general in order to cover all situations that occur in practice.

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

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

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

It is clear that the single ε in the previous definition can be replaced by different ones for different constraints, but to keep the notation as simple as possible, we decided to take the same ε everywhere. It is not difficult to see, however, that our main result still holds with different ε provided that the corresponding assumption is adapted in a suitable (straightforward) way.

We further stress that our definition of ε-stationarity is rather general and, in particular, less restrictive than the corresponding ε-stationarity concepts used in the related papers [14, 15]. Despite its generality, however, we will be able to get C-stationarity in the limit of the Lin-Fukushima-relaxation method, in particular, this convergence result then also holds under any other ε-stationarity condition that is stronger than ours.

We next introduce some notation and basic definitions for the MPEC (1). Again, we denote the set of feasible points by $X \subseteq \mathbb{R}^n$ and define the following index sets for any $x \in X$:

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

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 $x = x^*$ for some point x^* that will be clear from the context, we sometimes abbreviate the index sets $I_g(x^*), I_{0+}(x^*), I_{00}(x^*)$, and $I_{+0}(x^*)$ 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 that are important in our context.

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*

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

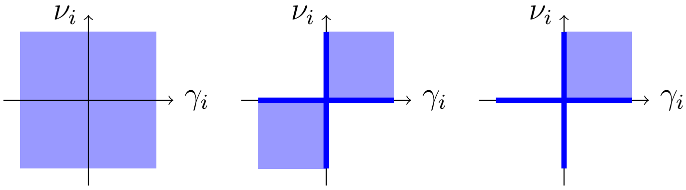
and the conditions

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

are satisfied;

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

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



(a) weak stationarity (b) C-stationarity (c) M-stationarity

Figure 1: Geometric illustration of weak, C-, and M-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, it is clear that weak stationarity is the weakest concept, while M-stationarity is the strongest concept within these three stationarity conditions. For our analysis, only weak and C-stationarity will be used, the definition of M-stationarity is included only for the sake of completeness since it has been mentioned in the introduction. The notion of weak and C-stationarity comes from the paper [28], whereas M-stationarity was introduced independently in [22, 23, 32, 34].

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, similar to standard NLPs, that a suitable 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*

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

are linearly independent;

(b) MPEC-Mangasarian Fromovitz CQ *(MPEC-MFCQ), if the gradients*

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

are positively linearly independent;

In the definition of MPEC-MFCQ, we use the notion of positively linearly independent vectors. Here, a set $\{a_i \mid i \in I\} \cup \{b_j \mid j \in J\}$ of vectors $a_i, b_j \in \mathbb{R}^n$ 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*. Note that MPEC-MFCQ is typically defined in a different way, but can easily be seen, by a theorem of the alternative, cf. [20], to be equivalent to the above condition.

It is clear that MPEC-LICQ is stronger than MPEC-MFCQ. Further note that MPEC-MFCQ imply that a local minimum is M-stationary, see, e.g., [8, 33], whereas MPEC-LICQ implies even a stronger stationarity concept, see [8, 19, 25] and references therein for more details.

3 The Inexact Smooth Relaxation Method

The relaxation method suggested by Lin and Fukushima [18] replaces the complementarity conditions $G_i(x) \geq 0,\ H_i(x) \geq 0,\ G_i(x)H_i(x) = 0$ 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

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

with a relaxation parameter $t > 0$. This transforms the feasible set of the complementarity constraints to the shape depicted in Figure 2. Here the superscript in Φ_i^S is used to remind the reader that this function is essentially identical to the one used by Scholtes [29], whereas the superscript in Φ_i^{LF} is used as an abbreviation for Lin and Fukushima since this is the new function introduced in [18].

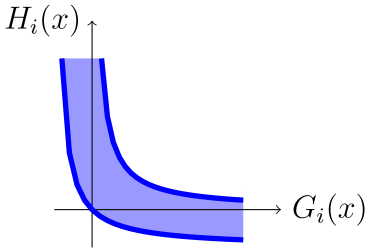


Figure 2: Geometric interpretation of the relaxation by Lin and Fukushima

The Lin-Fukushima-relaxation leads to a sequence of relaxed nonlinear programs

$$\begin{array}{lll} \min f(x) & \text{s.t.} & g_i(x) \le 0 & \forall\, i = 1,\dots,m, \\ & & h_i(x) = 0 & \forall\, i = 1,\dots,p, \\ & & \Phi_i^{LF}(x;t) = (G_i(x)+t)(H_i(x)+t) - t^2 \ge 0 & \forall\, i = 1,\dots,q, \\ & & \Phi_i^S(x;t^2) = G_i(x)H_i(x) - t^2 \le 0 & \forall\, i = 1,\dots,q \end{array} \tag{4}$$

with $t \downarrow 0$, which we will denote by $\text{NLP}^{LF}(t)$. From the original paper [18], the following convergence result is known.

Theorem 3.1 *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 $x^k \to x^*$ and MPEC-LICQ holds in x^*, then x^* is a C-stationary point of the MPEC* (1)*.*

Reference [10] 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 $\text{NLP}^{LF}(t_k)$ provided that ε_k goes to zero sufficiently fast.

Theorem 3.2 *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 $x^k \to x^*$. If MPEC-MFCQ holds in x^*, then x^* is a C-stationary point of the MPEC.*

Proof: Since x^k is an ε_k-stationary point of $\text{NLP}^{LF}(t_k)$, there exist multipliers λ^k, μ^k, τ^k, and δ^k such that

$$\Bigg\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) \\ - \sum_{i=1}^q \tau_i^k \nabla \Phi_i^{LF}(x^k;t_k) + \sum_{i=1}^q \delta_i^k \nabla \Phi_i^S(x^k;t_k^2) \Bigg\|_\infty \le \varepsilon_k$$

and

$$\begin{array}{l}
g_i(x^k) \leq \varepsilon_k, \quad \lambda_i^k \geq -\varepsilon_k, \quad |\lambda_i^k g_i(x^k)| \leq \varepsilon_k, \quad \forall i = 1, \dots, m, \\
|h_i(x^k)| \leq \varepsilon_k \quad \forall i = 1, \dots, p, \\
\Phi_i^{LF}(x^k; t_k) \geq -\varepsilon_k, \quad \tau_i^k \geq -\varepsilon_k, \quad \left|\tau_i^k \Phi_i^{LF}(x^k; t_k)\right| \leq \varepsilon_k \quad \forall i = 1, \dots, q, \\
\Phi_i^{S}(x^k; t_k^2) \leq \varepsilon_k, \quad \delta_i^k \geq -\varepsilon_k, \quad \left|\delta_i^k \Phi_i^{S}(x^k; t_k^2)\right| \leq \varepsilon_k \quad \forall i = 1, \dots q.
\end{array}$$

It is easy to see that the limit point x^* is feasible for the MPEC. Furthermore, we have

$$\begin{aligned}
\nabla \Phi_i^{LF}(x^k; t_k) &= \big(H_i(x^k) + t_k\big)\nabla G_i(x^k) + \big(G_i(x^k) + t_k\big)\nabla H_i(x^k), \\
\nabla \Phi_i^{S}(x^k; t_k^2) &= H_i(x^k)\nabla G_i(x^k) + G_i(x^k)\nabla H_i(x^k).
\end{aligned}$$

From the ε_k-stationarity of x^k, we therefore obtain

$$\begin{aligned}
\Bigg\| &\nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) \\
&- \sum_{i=1}^q \tau_i^k \big((H_i(x^k) + t_k)\nabla G_i(x^k) + (G_i(x^k) + t_k)\nabla H_i(x^k)\big) \\
&+ \sum_{i=1}^q \delta_i^k \big(H_i(x^k)\nabla G_i(x^k) + G_i(x^k)\nabla H_i(x^k)\big)\Bigg\|_\infty \leq \varepsilon_k.
\end{aligned}$$

Defining the new multipliers

$$\begin{aligned}
\gamma_i^k &:= \begin{cases} \tau_i^k\big(H_i(x^k) + t_k\big) - \delta_i^k H_i(x^k) & \text{if } i \in I_{00}(x^*) \cup I_{0+}(x^*), \\ 0 & \text{if } i \in I_{+0}(x^*), \end{cases} \\
\nu_i^k &:= \begin{cases} \tau_i^k\big(G_i(x^k) + t_k\big) - \delta_i^k G_i(x^k) & \text{if } i \in I_{00}(x^*) \cup I_{+0}(x^*), \\ 0 & \text{if } i \in I_{0+}(x^*). \end{cases}
\end{aligned}$$

this can be rewritten as

$$\begin{aligned}
\Bigg\| &\nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla h_i(x^k) - \sum_{i \in I_{00} \cup I_{0+}} \gamma_i^k \nabla G_i(x^k) \\
&- \sum_{i \in I_{00} \cup I_{+0}} \nu_i^k \nabla H_i(x^k) - \sum_{i \in I_{+0}} \big[\tau_i^k\big(H_i(x^k) + t_k\big) - \delta_i^k H_i(x^k)\big]\nabla G_i(x^k) \\
&- \sum_{i \in I_{0+}} \big[\tau_i^k\big(G_i(x^k) + t_k\big) - \delta_i^k G_i(x^k)\big]\nabla H_i(x^k)\Bigg\|_\infty \leq \varepsilon_k.
\end{aligned}$$

We claim that the sequence of multipliers $\big\{(\lambda^k, \mu^k, \gamma^k, \nu^k, \tau^k_{I_{+0} \cup I_{0+}}, \delta^k_{I_{+0} \cup I_{0+}})\big\}$ is bounded. Suppose this sequence is unbounded. Then we may assume without loss of generality that

$$\frac{(\lambda^k, \mu^k, \gamma^k, \nu^k, \tau^k_{I_{+0} \cup I_{0+}}, \delta^k_{I_{+0} \cup I_{0+}})}{\|(\lambda^k, \mu^k, \gamma^k, \nu^k, \tau^k_{I_{+0} \cup I_{0+}}, \delta^k_{I_{+0} \cup I_{0+}})\|} \to (\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}, \bar{\tau}_{I_{+0} \cup I_{0+}}, \bar{\delta}_{I_{+0} \cup I_{0+}}) \neq 0. \qquad (5)$$

Using the ε_k-stationarity, we would obtain

$$\sum_{i=1}^m \bar{\lambda}_i \nabla g_i(x^*) + \sum_{i=1}^p \bar{\mu}_i \nabla h_i(x^*) - \sum_{i=1}^q \bar{\gamma} \nabla G_i(x^*) - \sum_{i=1}^q \bar{\nu} \nabla H_i(x^*) = 0, \qquad (6)$$

where we used the fact that $t_k \to 0, G_i(x^k) \to 0$ for all $i \in I_{0+}(x^*)$, $H_i(x^k) \to 0$ for all $i \in I_{+0}(x^*)$, and $\{\nabla G_i(x^k)\}, \{\nabla H_i(x^k)\}$ are bounded by the continuous differentiability of the mappings G_i, H_i. Note that $\operatorname{supp}(\bar{\gamma}) \subseteq I_{00}(x^*) \cup I_{0+}(x^*)$, and $\operatorname{supp}(\bar{\nu}) \subseteq I_{00}(x^*) \cup I_{+0}(x^*)$ holds. The ε_k-stationarity also implies $\bar{\lambda} \geq 0$, and $g_i(x^*) = 0$ for all i with $\bar{\lambda}_i \neq 0$: If $\bar{\lambda}_i > 0$, we have $\lambda_i^k > c$ for some constant $c > 0$ and all k sufficiently large. This yields

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

due to $\varepsilon_k \downarrow 0$ and thus we have $\operatorname{supp}(\bar{\lambda}) \subseteq I_g(x^*)$. But then (6) and the assumed MPEC-MFCQ implies that $(\bar{\lambda}, \bar{\mu}, \bar{\gamma}, \bar{\nu}) = 0$ which, in turn, gives $\left(\bar{\tau}_{I_{+0} \cup I_{0+}}, \bar{\delta}_{I_{+0} \cup I_{0+}}\right) \neq 0$. Hence $\bar{\tau}_{I_{+0} \cup I_{0+}} \neq 0$ or $\bar{\delta}_{I_{+0} \cup I_{0+}} \neq 0$.

We claim that this is not possible. To this end, first consider the subvector $\bar{\tau}_{I_{+0} \cup I_{0+}}$ and suppose that $\bar{\tau}_i \neq 0$ for some index $i \in I_{+0}$ (a similar argument can be used for an index $i \in I_{0+}$). The ε_k-stationarity yields $\bar{\tau}_i > 0$, and using $\varepsilon_k = o(t_k^2)$ gives

$$\tau_i^k \left(\frac{(G_i(x^k) + t_k)(H_i(x^k) + t_k)}{t_k^2} - 1 \right) \to 0.$$

Consequently, we have

$$\frac{(G_i(x^k) + t_k)(H_i(x^k) + t_k)}{t_k^2} \to 1. \tag{7}$$

On the other hand, we have

$$0 = \bar{\nu}_i = \lim_{k \to \infty} \left[\underbrace{\frac{\tau_i^k}{\|(\cdots)\|}}_{\to \bar{\tau} > 0} \underbrace{\left(G_i(x^k) + t_k\right)}_{\to G_i(x^*) > 0} - \underbrace{\frac{\delta_i^k}{\|(\cdots)\|}}_{\to \bar{\delta}_i} \underbrace{G_i(x^k)}_{\to G_i(x^*) > 0} \right],$$

where $\|(\cdots)\|$ stands for the denominator from (5). Hence we necessarily have $\bar{\delta}_i > 0$. Therefore, using once more the ε_k-stationarity together with the assumption $\varepsilon_k = o(t_k^2)$, we obtain

$$\delta_i^k \left[\frac{G_i(x^k) H_i(x^k)}{t_k^2} - 1 \right] \to 0 \quad \Longrightarrow \quad \frac{G_i(x^k) H_i(x^k)}{t_k^2} \to 1.$$

Combining both limits yields

$$\underbrace{\frac{G_i(x^k) H_i(x^k)}{t_k^2}}_{\to 1} + \underbrace{\frac{G_i(x^k)}{t_k}}_{\to \infty \text{ since } i \in I_{+0}} + \frac{H_i(x^k)}{t_k} + 1 = \frac{\left(G_i(x^k) + t_k\right)\left(H_i(x^k) + t_k\right)}{t_k^2} \to 1, \tag{8}$$

hence it follows that $H_i(x^k)/t_k \to -\infty$, whereas $G_i(x^k) + t_k \to G_i(x^*) > 0$ yields $H_i(x^k) + t_k > 0$ for all k sufficiently large in view of (7), which means that $H_i(x^k) > -t_k$ and therefore gives the contradiction $\frac{H_i(x^k)}{t_k} > -1$ for all k sufficiently large.

In a similar way, we obtain a contradiction to the assumption $\bar{\delta}_i \neq 0$ for some index $i \in I_{0+}(x^*) \cup I_{+0}(x^*)$. Altogether, we therefore have $(\bar{\tau}_{I_{+0} \cup I_{0+}}, \bar{\delta}_{I_{+0} \cup I_{0+}}) = 0$, but this contradicts (5).

Consequently, the sequence $\{(\lambda^k, \mu^k, \gamma^k, \nu^k, \tau^k_{I_{+0}\cup I_{0+}}, \delta^k_{I_{+0}\cup I_{0+}})\}$ is bounded and can therefore be assumed, without loss of generality, to converge to some limit point $(\lambda^*, \mu^*, \gamma^*, \nu^*, \tau^*_{I_{+0}\cup I_{0+}}, \delta^*_{I_{+0}\cup I_{0+}})$. The ε_k-stationarity then implies

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

as well as $\lambda^* \geq 0, \operatorname{supp}(\lambda^*) \subseteq I_g(x^*)$, and

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

This shows that x^* is a weakly stationary point.

To verify C-stationarity, it remains to prove that $\gamma_i^* \nu_i^* \geq 0$ for all $i \in I_{00}(x^*)$. Assume that there exists an index $i \in I_{00}(x^*)$ such that $\gamma_i^* \nu_i^* < 0$. Without loss of generality, we may assume that $\gamma_i^* < 0$ and $\nu_i^* > 0$. Since $\gamma_i^* = \lim_{k\to\infty} \big(\tau_i^k (H_i(x^k) + t_k) - \delta_i^k H_i(x^k)\big) < 0$, it follows that there is a suitable constant $c > 0$ and a subsequence such that either $\tau_i^k (H_i(x^k) + t_k) < -c < 0$ or $\delta_i^k H_i(x^k) > c > 0$ for all sufficiently large k on this subsequence. Using $H_i(x^k) \to 0$ and $t_k \to 0$, we have either $|\tau_i^k| \to \infty$ or $|\delta_i^k| \to \infty$, and the ε_k-stationarity therefore yields either $\tau_i^k \to \infty, H_i(x^k) + t_k < 0$ or $\delta_i^k \to \infty, H_i(x^k) > 0$ for all k sufficiently large on a suitable subsequence. Using the ε_k-stationarity and $\varepsilon_k = o(t_k^2)$ once more, we obtain from $\tau_i^k \to \infty$ resp. $\delta_i^k \to \infty$ that either

$$\tau_i^k \left[\frac{(G_i(x^k) + t_k)(H_i(x^k) + t_k)}{t_k^2} - 1 \right] \to 0 \quad \Longrightarrow \quad \frac{(G_i(x^k) + t_k)(H_i(x^k) + t_k)}{t_k^2} \to 1$$

or

$$\delta_i^k \left[\frac{G_i(x^k) H_i(x^k)}{t_k^2} - 1 \right] \to 0 \quad \Longrightarrow \quad \frac{G_i(x^k) H_i(x^k)}{t_k^2} \to 1$$

on a suitable subsequence.

This implies that, on the subsequence, only one of the multipliers τ_i^k, δ_i^k can be unbounded. In fact, if both were unbounded, we would have

$$\frac{(G_i(x^k) + t_k)(H_i(x^k) + t_k)}{t_k^2} \to 1 \quad \text{and} \quad \frac{G_i(x^k) H_i(x^k)}{t_k^2} \to 1$$

for some index $i \in I_{00}(x^*)$. Then, we obtain a contradiction in the following way: If, in addition, we have $H_i(x^k) + t_k < 0$, then $(G_i(x^k) + t_k)(H_i(x^k) + t_k)/t_k^2 \to 1$ yields $G_i(x^k) + t_k < 0$ and thus $(G_i(x^k) + H_i(x^k))/t_k < -2t_k/t_k = -2 < -1$, whereas, on the other hand, we have $(G_i(x^k) + H_i(x^k))/t_k \to -1$ (see (8)), which gives the desired contradiction in this case. On the other hand, if, in addition, we have $H_i(x^k) > 0$, we obtain from $(G_i(x^k) H_i(x^k))/t_k^2 \to 1$ that $G_i(x^k) > 0$ and thus $(G_i(x^k) + H_i(x^k))/t_k > (0 + 0)/t_k = 0 > -1$, a contradiction to $(G_i(x^k) + H_i(x^k))/t_k \to -1$ (see again (8)).

First, consider the case $\tau_i^k (H_i(x^k) + t_k) < -c$ on the subsequence. Since this implies that δ_i^k is bounded on the subsequence, it follows that

$$\gamma_i^* = \lim_k \underbrace{\tau_i^k}_{\to\infty} \underbrace{\big(H_i(x^k) + t_k\big)}_{<0} < 0 \quad \text{and} \quad \nu_i^* = \lim_k \underbrace{\tau_i^k}_{\to\infty} \big(G_i(x^k) + t_k\big) > 0.$$

Hence $G_i(x^k) + t_k > 0$ and thus all entries of the sequence $\big(G_i(x^k) + t_k\big)\big(H_i(x^k) + t_k\big)/t_k^2$ are negative, a contradiction to the fact that this subsequence converges to 1.

Next, consider the case $\delta_i^k H_i(x^k) > c$ on the subsequence, implying that the corresponding subsequence of $\{\tau_i^k\}$ is bounded. Then

$$\gamma_i^* = \lim_k \big(- \underbrace{\delta_i^k}_{\to\infty} \underbrace{H_i(x^k)}_{>0}\big) < 0 \quad \text{and} \quad \nu_i^* = \lim_k \big(- \underbrace{\delta_i^k}_{\to\infty} G_i(x^k)\big) > 0.$$

Consequently, we have $G_i(x^*) < 0$, but then the entries of the corresponding subsequence $\big(G_i(x^k)H_i(x^k)\big)/t_k^2$ are negative, a contradiction to the fact that this subsequence converges to 1. $\square$

The following example shows that it is necessary to adjust the speed of $\varepsilon_k \downarrow 0$ with respect to $t_k \downarrow 0$, for example by $\varepsilon_k = o(t_k^2)$, in order to obtain C-stationary points in the limit.

Example 3.3 Consider the two-dimensional MPEC

$$\min_x \; -x_1 + x_2 \quad \text{s.t.} \quad x_1 \geq 0,\;\; x_2 \geq 0,\;\; x_1 x_2 = 0.$$

Let $t \downarrow 0$ arbitrarily given, and consider the corresponding sequences

$$\varepsilon_t := 6t, \quad x^t := \big(x_1^t, x_2^t\big) := \big(\frac{1}{2}t, 2t\big), \quad \tau^t := \frac{5}{3t}, \quad \delta^t := \frac{3}{t}.$$

Then x^t is an ε_t-stationary point of $\text{NLP}^{LF}(t)$ since an elementary calculation shows that

$$\begin{aligned}
&\Big\|\nabla f(x^t) - \tau^t \nabla \Phi^{LF}(x^t; t) + \delta^t \nabla \Phi^S(x^t; t^2)\Big\|_\infty = \\
&\qquad \left\| \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \frac{5}{3t}\begin{pmatrix} 3t \\ 1.5t \end{pmatrix} + \frac{3}{t}\begin{pmatrix} 2t \\ 0.5t \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| = 0 \leq \varepsilon_t, \\
&\Phi^{LF}(x^t; t) = \frac{7}{2}t^2 \geq 0 \geq -\varepsilon_t, \\
&\tau^t = \frac{5}{3t} \geq 0 \geq -\varepsilon_t \\
&\big|\tau^t \Phi^{LF}(x^t; t)\big| = \frac{35}{6}t \leq 6t = \varepsilon_t, \\
&\Phi^S(x^t; t^2) = t^2 - t^2 = 0 \leq \varepsilon_t, \\
&\delta^t = \frac{3}{t} \geq 0 \geq -\varepsilon_t \\
&\big|\delta^t \Phi^S(x^t; t^2)\big| = 0 \leq \varepsilon_t.
\end{aligned}$$

Furthermore, we have $x^t \to (0,0)$, but the origin is only weakly stationary. Since this example satisfies MPEC-LICQ, hence also MPEC-MFCQ, the only assumption that is violated in Theorem 3.2 is the condition $\varepsilon = o(t^2)$. $\Diamond$

4 Final Remarks

This paper shows that the relaxation method by Lin and Fukushima [18] converges to C-stationary points even if the corresponding relaxed nonlinear programs are solved only inexactly in the sense that ε-stationary points are computed. This property is in contrast to most of the other existing relaxation schemes whose inexact versions converge to weakly stationary points only, see the corresponding results and discussion in the accompanying paper [16].

Compared to the other relaxation schemes, the numerical behaviour of the Lin-Fukushima-approach was less favourable in the comparison given in [11]. However, in view of the theoretical results stated here, one might get re-interested into this approach, possibly by using an NLP-solver that is able to deal with the particular constraints arising in this context in a better way. To this end, let us note that the ε-stationary condition used in our framework should be sufficiently general in order to deal with essentially all kinds of NLP-solvers and corresponding NLP-termination criteria to cover also this more specialized subproblem-solver.

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] 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.

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

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

[8] 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.

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] C. Kanzow and A. Schwartz: *The price of inexactness: Convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited.* Preprint 315, Institute of Mathematics, University of Würzburg, Würzburg, Germany, March 2013.

[17] 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.

[18] 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.

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

[20] O.L. Mangasarian: *Nonlinear Programming.* McGraw-Hill, New York, NY, 1969 (reprinted by SIAM, Philadelphia, PA, 1994).

[21] J. Nocedal and S.J. Wright: *Numerical Optimization.* Springer, second edition 2006.

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

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

[24] 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.

[25] 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.

[26] 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.

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

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

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

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

[31] O. Stein: *Lifting mathematical programs with complementarity constraints.* Mathematical Programming 131, 2012, pp. 71–94.

[32] 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.

[33] 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.

[34] 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.

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