IMPROVED CONVERGENCE PROPERTIES OF THE LIN-FUKUSHIMA-REGULARIZATION METHOD FOR MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS[1]

Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz

Preprint 297 September 2010

University of Würzburg
Institute of Mathematics
Am Hubland
97074 Würzburg
Germany

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

September 2, 2010

Dedicated to Masao Fukushima, in great respect, on the occasion of his 60th birthday.

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

Abstract. We consider a regularization method for the numerical solution of mathematical programs with complementarity constraints (MPCC) introduced by Gui-Hua Lin and Masao Fukushima. Existing convergence results are improved in the sense that the MPCC-LICQ assumption is replaced by the weaker MPCC-MFCQ. Moreover, some preliminary numerical results are presented in order to illustrate the theoretical improvements.

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

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

1 Introduction

We consider an optimization problem of the form

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

where $f, g_i, h_i, G_i, H_i : \mathbb{R}^n \to \mathbb{R}$ are assumed to be at least continuously differentiable functions. It is known under the label *mathematical program with complementarity constraints* (MPCC for short). The interested reader is referred to the monographs [14, 16, 4] for several applications and a theoretical background of MPCCs.

In principle, one may view an MPCC as a standard nonlinear program, where the feasible set is described by a finite number of equality and inequality constraints. However, the particular structure of these constraints causes some difficulties. In fact, it is known [22] and easy to see that the standard Mangasarian-Fromovitz constraint qualification, MFCQ for short (and therefore also the stronger linear independence constraint qualification, LICQ) is violated at any feasible point of (1). This, in turn, implies that one has to expect severe difficulties when using standard algorithms like SQP- or interior-point-type methods for the solution of MPCCs.

This observation is the main motivation for the construction of more specialized algorithms that take into account the particular structure of an MPCC. There are different approaches available, but one of the most popular idea is certainly the relaxation method which, basically, enlarges the feasible set, especially the complementarity conditions, in order to avoid problems arising from this part.

Historically, the first relaxation method is due to Scholtes [19], followed by a modified relaxation scheme from Lin and Fukushima [12]. In the meantime, a number of further relaxation (or regularization) methods are available, see the two-sided relaxation method by Demiguel et al. [3], the nonsmooth relaxation scheme by Kadrani et al. [9], the local relaxation algorithm by Steffensen and Ulbrich [20], and the very recent relaxation method by Kanzow and Schwartz [10].

The paper [7] shows that the assumptions for convergence of the Steffensen-Ulbrich relaxation method can be relaxed. Furthermore, the very recent work [8], among other things, also improves the convergence properties of the two regularization methods by Scholtes and Kadrani et al.

Here, our main focus is on the relaxation scheme by Lin and Fukushima [12]. The most basic result in [12] states that, under a relatively strong assumption called MPCC-LICQ, a sequence of KKT points of the relaxed problems exists and converges to a limit point that is C-stationary. Here, the MPCC-LICQ assumption is used three times, namely to

- prove the existence of Lagrange multipliers for the relaxed problems
- to verify convergence of the sequence of these multipliers, and

- to guarantee that the limit point is C-stationary.

In this paper, we show that these properties, more or less, still hold under the weaker MPCC-MFCQ condition: The relaxed problems satisfy standard MFCQ which guarantees the existence of Lagrange multipliers at local minima, the sequence of multipliers is bounded (hence convergent at least on a subsequence), and the limit point is still C-stationary.

The organization of this paper is as follows: In Section 2, we first recall the definitions of MFCQ and LICQ and introduce the corresponding MPCC counterparts of these constraint qualifications. Also the concept of a C-stationary point is introduced there. Section 3 first recalls the modified relaxation scheme by Lin and Fukushima [12] and then shows that MPCC-MFCQ implies that, locally, standard MFCQ holds for the relaxed problem. The main convergence result is contained in Section 4 where we show that any limit point of a sequence of stationary points of the relaxed programs is a C-stationary point of the MPCC.

Some words regarding notation: Most of the notation used is standard. The real and natural numbers are denoted by $\mathbb{R}$ and $\mathbb{N}$, respectively, whereas $\mathbb{R}_+$ is the set of the nonnegative real numbers. In addition, for a (finite index) set I, its cardinality is indicated by $|I|$. Moreover, we use the symbol

$$\mathrm{supp}(y) := \{i \mid y_i \neq 0\}$$

for the support of a vector $y \in \mathbb{R}^n$.

2 Preliminaries

Let us consider the standard nonlinear program

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

with continuously differentiable functions $f, g_i, h_i : \mathbb{R}^n \to \mathbb{R}$. Let x^* be feasible for (2). Then the well-known *Karush-Kuhn-Tucker (KKT) conditions*, see, e.g., [1], are said to hold at x^* if there exist multipliers (λ, μ) such that

$$0 = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) \quad \text{and} \quad \lambda_i \geq 0, \quad \lambda_i g_i(x^*) = 0 \quad \forall i = 1, \ldots, m.$$

At this, the triple (x^*, λ, μ) is called a *KKT point*, whereas x^* is named *stationary point*.

Assumptions about the constraint functions that make the KKT conditions provide a necessary optimality criterion for (2) are called *constraint qualifications*. Two of the strongest ones, in the sense that in fact much less is needed to yield KKT conditions at a local minimizer of (2), are the *linear independence constraint qualification* and the *Mangasarian-Fromovitz constraint qualification*, which are defined below. For these purposes, put

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

if x^* is feasible for (2).

Definition 2.1 *Let x^* be feasible for* (2). *Then we say that*

(a) the linear independence constraint qualification *(LICQ) is fulfilled at x^* if the gradients*

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

are linearly independent.

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

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

Remark 2.2 (a) Obviously, we have the implication: LICQ $\implies$ MFCQ.

(b) It can be shown, using, e.g., *Motzkin's theorem of the alternative* [15], that MFCQ for (2) is satisfied at x^* if and only if the system

$$0 = \sum_{i \in I_g} \alpha_i \nabla g_i(x^*) + \sum_{i=1}^{p} \beta_i \nabla h_i(x^*), \quad \alpha \in \mathbb{R}_+^{|I_g|}, \ \beta \in \mathbb{R}^p,$$

only has the trivial solution. In this situation the gradients $\{\nabla g_i(x^*) \mid i \in I_g\} \cup \{\nabla h_i(x^*) \mid i = 1, \dots, p\}$ are called *positive-linearly independent*. Note, however, that there is no sign constraint for the multipliers of the equality constraints.

It is well known, see [2], that MFCQ, and hence in particular LICQ, is violated at any feasible point of the MPCC (1). To this end, some problem-tailored analogues have been established in the past. In order to define these, we need to introduce some crucial index sets. To this end, let x^* be feasible for (1). Then we put

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

Apparently, these index sets depend substantially on the respective point x^*, but we believe that no ambiguity arises from that, since it will always be clear which point we refer to.

Definition 2.3 *Let x^* be feasible for the MPCC* (1). *Then we say that*

(a) MPCC-LICQ is satisfied at x^* if the gradients

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

are linearly independent.

(b) MPCC-MFCQ is satisfied at x^* if the gradients

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

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

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

Similar to the standard case, MPCC-LICQ implies MPCC-MFCQ, and the converse direction is, in general, false. In fact, there is quite a gap between MPCC-LICQ on the one hand and MPCC-MFCQ on the other in the following sense: While under MPCC-LICQ a local minimizer of (1) is known to admit KKT multipliers, cf. [18, 5], it can only be shown to be *M-stationary* under MPCC-MFCQ, see [21, 6]. Hence, MPCC-LICQ must be viewed as a substantially stronger assumption.

The stationarity context applicable for our purposes is *C-stationarity*, see the Definition below.

Definition 2.4 *Let x^* be feasible for the MPCC* (1). *Then x^* is called a C-*stationary point *of* (1) *if there exist multipliers $(\lambda, \mu, \gamma, \nu)$ such that*

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

and

$$\begin{array}{l}\lambda_i \geq 0,\ \lambda_i g_i(x^*) = 0\ (i = 1, \dots, l) \\ \gamma_i = 0\ (i \in I_{+0}),\ \nu_i = 0\ (i \in I_{0+}), \\ \gamma_i \nu_i \geq 0\ (i \in I_{00}).\end{array}$$

It is well known, cf. [18], that C-statonarity is a necessary optimality condition for (1) under, e.g. MPCC-MFCQ. Note, however, that in fact much less than MPCC-MFCQ is needed to yield C-stationarity at a local minimizer of an MPCC. The importance of the C-stationarity concept comes from the fact that several regularization methods are known to converge to C-stationary points under suitable assumptions, including the methods from [19, 12, 3, 20]. This also includes the regularization scheme by Lin and Fukushima. The aim of the following sections is to show that convergence to such a C-stationary point holds under conditions that are weaker than those presented in [12].

3 The Modified Relaxation Scheme

The relaxation scheme proposed by Lin and Fukushima in [12] employs the following relaxation:

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

Its feasible set is denoted by $X(t)$. A geometric interpretation of the relaxation can be seen in Figure 1.

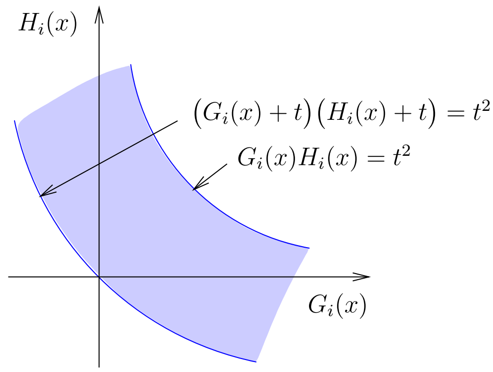

Figure 1: Illustration of the relaxation method

Figure 1 also displays that $R(t)$ is actually a relaxation of the MPCC (1). This is also formally confirmed in the following proposition.

Proposition 3.1 *We have*

$$X = \bigcap_{t>0} X(t).$$

Proof. See [12, Th. 2.1]. □

In order to investigate the modified relaxation scheme in depth, we need to introduce some further index sets. To this end, let $x \in X(t)$ for $t > 0$. Then we put:

$$\begin{aligned} I_\Phi^+(x;t) &:= \{i \mid G_i(x)H_i(x) - t^2 = 0\}, \\ I_\Phi^-(x;t) &:= \{i \mid (G_i(x) + t)(H_i(x) + t) - t^2 = 0\}. \end{aligned}$$

To facilitate the following proofs, we want to take a closer look at these index sets. Let x be feasible for $R(t)$ and $i \in I_\Phi^+(x, t)$. This implies $G_i(x)H_i(x) = t^2 > 0$, i.e. $G_i(x), H_i(x) \neq 0$

and both have the same sign. Now, assume that both values were negative. This would imply

$$G_i(x)H_i(x) + t(G_i(x) + H_i(x)) < t^2$$

in contradiction to the feasibility of x. Thus, we have the following implication:

$$i \in I_\Phi^+(x;t) \quad \Longrightarrow \quad G_i(x) > 0, H_i(x) > 0. \tag{3}$$

Now consider the case where $i \in I_\Phi^-(x;t)$. This implies $(G_i(x)+t)(H_i(x)+t) = t^2 > 0$, so either both values $G_i(x)+t, H_i(x)+t$ are strictly greater or smaller than zero. Assume that both values are negative. This implies $G_i(x), H_i(x) < 0$ and thus

$$G_i(x)H_i(x) - t^2 = -t(G_i(x) + H_i(x)) > 0$$

in contradiction to the feasibility of x. Hence, we obtain the following implication:

$$i \in I_\Phi^-(x;t) \quad \Longrightarrow \quad G_i(x) + t > 0, H_i(x) + t > 0. \tag{4}$$

The subsequent result shows that the relaxed problem $R(t)$ is in fact less ill-posed with respect to constraint qualifications than the original MPCC (1).

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

Proof. The LICQ part is due to [12, Th. 2.3].

Due to MPCC-MFCQ at x^*, in view of [17, Prop. 2.2], one can see that the following set of vectors is positive-linearly independent for all $x \in X(t)$ sufficiently close to x^*:

$$\begin{array}{rl}
\nabla g_i(x) & \big(i \in I_g(x)\big), \\
\nabla h_i(x) & (i = 1, \dots, p), \\
G_i(x)\nabla H_i(x) + H_i(x)\nabla G_i(x) & \big(i \in I_\Phi^+(x;t) \cap (I_{+0} \cup I_{0+})\big), \\
\nabla H_i(x) & \big(i \in I_\Phi^+(x;t) \cap I_{00}\big), \\
\nabla G_i(x) & \big(i \in I_\Phi^+(x;t) \cap I_{00}\big), \\
\big(G_i(x)+t\big)\nabla H_i(x) + \big(H_i(x)+t\big)\nabla G_i(x) & \big(i \in I^-(x;t) \cap (I_{+0} \cup I_{0+})\big), \\
\nabla H_i(x) & \big(i \in I_\Phi^-(x;t) \cap I_{00}\big), \\
\nabla G_i(x) & \big(i \in I_\Phi^-(x;t) \cap I_{00}\big).
\end{array} \tag{5}$$

We now claim that standard MFCQ holds for the relaxed program $R(t)$ whenever $x \in X(t) \cap N(x^*)$ for some sufficiently small neighborhood $N(x^*)$ of x^*. To this end, let x be such an element. In view of Remark 2.2, we have to show that

$$\begin{aligned}
0 \quad = \quad & \sum_{i \in I_g(x)} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \mu_i \nabla h_i(x) + \sum_{i \in I_\phi^+(x,t)} \alpha_i (G_i(x)\nabla H_i(x) + H_i(x)\nabla G_i(x)) \\
& - \sum_{i \in I_\Phi^-(x,t)} \beta_i [(G_i(x)+t)\nabla H_i(x) + (H_i(x)+t)\nabla G_i(x)],
\end{aligned}$$

with suitable multipliers $\mu \in \mathbb{R}^p$ and $\lambda, \alpha, \beta \geq 0$ holds only for the zero vector. In order to see this, let us rewrite the above equation as

$$\begin{aligned} 0 \;=\; & \sum_{i \in I_g(x)} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \mu_i \nabla h_i(x) \\ & + \sum_{i \in I_\Phi^+(x;t) \cap (I_{+0} \cup I_{0+})} \alpha_i (G_i(x) \nabla H_i(x) + H_i(x) \nabla G_i(x)) \\ & + \sum_{i \in I_\Phi^+(x;t) \cap I_{00}} \alpha_i G_i(x) \nabla H_i(x) + \sum_{i \in I_\Phi^+(x;t) \cap I_{00}} \alpha_i H_i(x) \nabla G_i(x) \\ & - \sum_{i \in I_\Phi^-(x;t) \cap (I_{+0} \cup I_{0+})} \beta_i [(G_i(x) + t) \nabla H_i(x) + (H_i(x) + t) \nabla G_i(x)] \\ & - \sum_{i \in I_\Phi^-(x;t) \cap I_{00}} \beta_i (G_i(x) + t) \nabla H_i(x) - \sum_{i \in I_\Phi^-(x;t) \cap I_{00}} \beta_i^k (H_i(x) + t) \nabla G_i(x). \end{aligned} \tag{6}$$

Now, using the positive-linear independence of the vectors from (5) and applying this observation to (6), taking into account (3) and (4), we immediately see that $(\lambda, \mu, \alpha, \beta) = 0$, and this completes the proof. □

4 Convergence Result

In this section we state the main convergence result which can be viewed as a refinement of [12, Theorem 3.3]. It shows that MPCC-MFCQ is sufficient to guarantee that a limit point of a sequence of stationary points of the relaxed programs $R(t)$ is C-stationary. The corresponding result in [12] requires the stronger MPCC-LICQ condition in order to obtain this statement.

Theorem 4.1 *Let $\{t_k\} \downarrow 0$ and let x^k be a stationary point of $R(t)$ with $x^k \to x^*$ such that MPCC-MFCQ holds in x^*. Then x^* is C-stationary.*

Proof. Since x^k is a stationary point of $R(t_k)$, we have multipliers $\lambda^k, \mu^k, \delta^{+,k}, \delta^{-,k}$ such that

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

with

$$\begin{aligned} \lambda^k \geq 0 \quad & \text{and} \quad \operatorname{supp}(\lambda^k) \subseteq I_g(x^k), \\ \delta^{+,k} \geq 0 \quad & \text{and} \quad \operatorname{supp}(\delta^{+,k}) \subseteq I_\Phi^+(x^k; t_k), \\ \delta^{-,k} \geq 0 \quad & \text{and} \quad \operatorname{supp}(\delta^{-,k}) \subseteq I_\Phi^-(x^k; t_k) \end{aligned}$$

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

$$\operatorname{supp}(\delta^{+,k}) \cap \operatorname{supp}(\delta^{-,k}) = \emptyset \tag{7}$$

for all $k \in \mathbb{N}$. Hence the following new multipliers are at least well-defined:

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

and

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

With these multipliers, we can rewrite the equation from the beginning as

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

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

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

The equation above then yields

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

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

$$\begin{aligned} \operatorname{supp}(\lambda) &\subseteq I_g(x^k) \subseteq I_g, \\ \operatorname{supp}(\gamma) &\subseteq I_{00} \cup I_{0+}, \\ \operatorname{supp}(\nu) &\subseteq I_{00} \cup I_{+0}. \end{aligned}$$

Additionally, $(\lambda, \mu, \gamma, \nu) \neq 0$ has to hold. Otherwise, $\delta_i^+ > 0$ or $\delta_i^- > 0$ would have to hold for at least one index $i \in I_{+0} \cup I_{0+}$. Assume first without loss of generality that $\delta_i^+ > 0$ for an $i \in I_{+0}$. This implies $\delta_i^{+,k} > 0$ for all k sufficiently large and consequently

$\nu_i^k = -\delta_i^{+,k} G_i(x^k)$ for those k. Because of $i \in I_{+0}$, this yields $\nu_i = \lim_{k \in K} -\delta_i^{+,k} G_i(x^k) < 0$, a contradiction to our assumption $\nu = 0$. Now assume $\delta_i^- > 0$ for an $i \in I_{+0}$. This implies $\delta_i^{-,k} > 0$ for all K sufficiently large and thus $\nu_i^k = \delta_i^{-,k}(G_i(x^k) + t_k)$ for those k. Because of $i \in I_{+0}$, this yields $\nu_i = \lim_{k \in K} \delta_i^{-,k}(G_i(x^k) + t_k) > 0$, again a contradiction to our assumption $\nu = 0$.

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

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

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

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

To prove the C-stationarity of x^*, it remains to show $\gamma_i^* \nu_i^* \geq 0$ for all $i \in I_{00}$. Assume that there is an $i \in I_{00}$ with $\gamma_i^* < 0$ and $\nu_i^* > 0$ or with $\gamma_i^* > 0$ and $\nu_i^* < 0$. We consider only the first case, the second one can be treated similarly. Since $\gamma_i^* < 0$, we have $\gamma_i^k < 0$ for all $k \in \mathbb{N}$ sufficiently large. But this implies $i \in \operatorname{supp}(\delta^{+,k})$ since, otherwise, the definition of γ_i^k would imply $i \in \operatorname{supp}(\delta^{-,k})$, hence $i \in I_\Phi^-(x^k; t_k)$, hence $\delta_i^{-,k} > 0$ and $H_i(x^k) + t_k > 0$ by (4) and, therefore, $\gamma_i^k > 0$ due to the definition of γ_i^k. Knowing that $i \in \operatorname{supp}(\delta^{+,k})$, we have $i \notin \operatorname{supp}(\delta^{-,k})$ in view of (7). This implies that either $\nu_i^k = 0$ or $\nu_i^k = -\delta_i^{+,k} G_i(x^k)$ for all $k \in \mathbb{N}$ sufficiently large. However, $i \in \operatorname{supp}(\delta^{+,k})$ gives $i \in I_\Phi^+(x^k; t_k)$ and, therefore, $G_i(x^k) > 0$ by (3). This shows that, in any case, we have $\nu_i^k \leq 0$ which, in turn, gives the contradiction $\nu_i^* \leq 0$. □

Note that the previous proof actually shows that a suitable sequence of multipliers remains bounded and, therefore, converges at least on a subsequence under the MPCC-MFCQ condition. The related result in [12], on the other hand, shows convergence of a corresponding sequence of multipliers under the stronger MPCC-LICQ assumption.

5 Numerical Experience

As the original work [12] does not include any numerical results, we want to illustrate the behaviour of this relaxation method on a few problems taken from the MacMPEC test problem collection [11]. A more complete view is currently in preparation by the authors in [8], which includes both a theoretical and a numerical comparison of basically all existing regularization methods for MPCCs.

Our implementation of the Lin-Fukushima regularization methods uses the framework from Algorithm 1. The implementation is done in MATLAB 7.10.0:

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

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

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

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

Before we present the numerical results, we would like to discuss a few details of the algorithm. The stopping criterion basically consists of two conditions. The condition $t \geq t_{\min}$ is based on the fact that, for extremely small relaxation parameters t, the relaxed problem $R(t)$ is very similar to the original MPCC. Thus, we expect standard NLP solvers to have difficulties finding a solution of $R(t)$ if the relaxation parameter becomes too small. The term $\text{compVio}(x)$ in the stopping criterion is defined as

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

and measures the violation of the complementarity constraints in the current iterate x^k. Since the standard constraints $g(x) \leq 0$ and $h(x) = 0$ are incorporated into the relaxed problem $R(t)$, a local minimum of $R(t_k)$ with $\text{compVio}(x^k) = 0$ is a local minimum of the original MPCC, too, and we can stop immediately in this case. Here, we used Proposition 3.1. This Proposition also implies that we can terminate the algorithm if one of the relaxed problems $R(t_k)$ is infeasible since the feasible area of the MPCC is a subset of the feasible area of $R(t)$. The update rule for the relaxation parameter is designed to guarantee that the solution of the current iteration x^{k+1} is infeasible for the next iteration because otherwise, it would be a solution of the next iteration, too. Eventually, we return the final iterate x_{opt} with the corresponding function value $f(x_{opt})$, and, to measure the feasibility of x_{opt}, the maximum violation of all constraints

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

In our implementation of Algorithm 1 we use the TOMLAB 7.4.0 solver `knitro` to solve the relaxed problems $R(t_k)$ and the parameters $(\sigma, t_{\min}, \varepsilon) = (0.01, 10^{-15}, 10^{-5})$.

To illustrate the behaviour of this algorithm, we first examine the following problem

$$\min_{x_1,x_2} \frac{1}{2}\left((x_1-1)^2 + (x_2-1)^2\right) \quad \text{s.t.} \quad x_1 \geq 0, x_2 \geq 0, x_1 x_2 = 0$$

which corresponds to `scholtes3` in the MacMPEC collection. This problem has two solutions $(1,0)^T$ and $(0,1)^T$, and a C-stationary point $(0,0)$ which is obviously not a solution. However, it is possible that Algorithm 1 converges towards the C-stationary point since (t,t) is a KKT-point of the relaxed problem $\mathrm{R}(t)$ for all $t \in (0,1)$. Thus, we did two experiments: First, we started the algorithm in the unconstrained global minimum $(1,1)^T$ and chose an initial relaxation $t_0 = 0.9$ such that the global minimum is infeasible for $R(t_0)$. Otherwise, nothing would happen in the first iteration. Table 1 gives the corresponding results.

k	x^{k+1}	$f(x^{k+1})$	compVio(x^{k+1})	t_k
0	(0.90000 0.90000)	0.01000	0.9000000	9e-01
1	(0.00008 0.99992)	0.49992	0.0000810	9e-03
2	(0.00000 1.00000)	0.50000	0.0000000	9e-05

Table 1: Results for `scholtes3` with initial point $(1,1)^T$

Obviously, the algorithm stays on the (t,t)-line only for two iterations and then converges to one of the two solutions. The same behavior can be observed when we start directly in the C-stationary point $(0,0)^T$ with a smaller initial relaxation $t_0 = 0.1$, see Table 2.

k	x^{k+1}	$f(x^{k+1})$	compVio(x^{k+1})	t_k
0	(0.01010 0.98990)	0.49000	0.0101021	1e-01
1	(0.00000 1.00000)	0.50000	0.0000010	1e-03

Table 2: Results for `scholtes3` with initial point $(0,0)^T$

Thus, although it is theoretically possible that the algorithm converges towards the undesirable C-stationary point, this does not happen in praxis.

The second example we would like to examine is `scholtes4` from the MacMPEC collection. This example looks as follows:

$$\begin{aligned}\min_{x_1,x_2,x_3} x_1 + x_2 - x_3 \quad \text{s.t.} \quad & -4x_1 + x_3 \leq 0, \\ & -4x_2 + x_3 \leq 0, \\ & x_1 \geq 0, x_2 \geq 0, x_1 x_2 = 0.\end{aligned}$$

This problem has only one solution $(0,0,0)^T$ where MPEC-LICQ is violated but MPEC-MFCQ holds. The results corresponding to the initial point $(0,1,0)^T$ and an initial relaxation of $t_0 = 1$ are given in Table 3.

It can be seen that the iterates still converge towards the known solution although MPCC-LICQ is violated. However, the convergence is slower than in the first example where MPCC-LICQ is satisfied.

k	x^{k+1}	$f(x^{k+1})$	compVio(x^{k+1})	t_k
0	(1.00000 1.00000 4.00000)	-2.00000	1.0000000	1e+00
1	(0.01000 0.01000 0.04000)	-0.02000	0.0100000	1e-02
2	(0.00010 0.00010 0.00040)	-0.00020	0.0001005	1e-04
3	(0.00002 0.00002 0.00008)	-0.00004	0.0000196	1e-06
4	(0.00002 0.00002 0.00008)	-0.00004	0.0000199	1e-08
5	(0.00002 0.00002 0.00008)	-0.00004	0.0000199	1e-10
6	(0.00001 0.00001 0.00005)	-0.00003	0.0000133	1e-12
7	(0.00000 0.00000 0.00002)	-0.00001	0.0000050	1e-14

Table 3: Results for `scholtes4` with initial point $(0, 1, 0)^T$

These small-dimensional examples nicely show the behaviour of the regularization scheme and illustrate our theory, they cannot be used, however, in order to obtain a complete picture regarding the behaviour of the Lin-Fukushima regularization method on more complicated or higher-dimensional examples. As mentioned in the beginning of this section, this is part of a more extensive numerical testing that can be found in [8] which, among other things, includes a comparison of several relaxation schemes for MPCCs.

6 Final Remarks

In this paper, we have shown that the MPCC-LICQ assumption can be replaced by the weaker MPCC-MFCQ condition in order to get a C-stationary point for the Lin-Fukushima regularization method [12]. We have also shown that MPCC-MFCQ implies that the regularized problems satisfy standard MFCQ (locally). While it seems possible to prove that many other MPCC-tailored constraint qualification imply that (locally) the corresponding standard constraint qualification holds for the regularized problem, it is an open question whether one can further relax the MPCC-MFCQ assumption to get C-stationary points in the limit.

References

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

[2] Y. Chen and M. Florian: *The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions.* Optimization, 32 (1995), pp. 193-209.

[3] 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.

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

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

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

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

[8] T. Hoheisel, C. Kanzow, and A. Schwartz: *Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints.* Preprint, Institute of Mathematics, University of Würzburg, Würzburg, forthcoming.

[9] 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.

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

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

[12] 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.

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

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

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

[16] 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.

[17] L. Qi and Z. Wei: *On the constant positive linear dependence condition and its applications to SQP methods.* SIAM Journal on Optimization, 10, 2000, pp. 963–981.

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

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

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

[21] 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.

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