

# A NEW RELAXATION SCHEME FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

SONJA VEELKEN\* AND MICHAEL ULBRICH†

**Abstract.** We present a new relaxation scheme for mathematical programs with equilibrium constraints (MPEC), where the complementarity constraints are replaced by a reformulation that is exact for the complementarity conditions corresponding to sufficiently non-degenerate complementarity components and relaxes only the remaining complementarity conditions. A positive parameter determines to what extent the complementarity conditions are relaxed. The relaxation scheme is such that a strongly stationary solution of the MPEC is also a solution of the relaxed problem if the relaxation parameter is chosen sufficiently small. We discuss the properties of the resulting parameterized nonlinear programs and compare stationary points and solutions. We further prove that a limit point of a sequence of stationary points of a sequence of relaxed problems is C-stationary if it satisfies a so-called MPEC-constant rank constraint qualification and it is M-stationary if it satisfies the MPEC-linear independence constraint qualification and the stationary points satisfy a second order sufficient condition. From this relaxation scheme, a numerical approach is derived that is applied to a comprehensive testset. The numerical results show that the approach combines good efficiency with high robustness.

**1. Introduction.** We consider mathematical programs with equilibrium constraints (MPEC) of the form

$$(1.1) \quad \begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) \geq 0 \\ & h(x) = 0 \\ & 0 \leq x_1 \perp x_2 \geq 0, \end{aligned}$$

where  $x = (x_0, x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p$ . Here, all inequalities are meant componentwise. Throughout, we assume that  $f : \mathbb{R}^{n+2p} \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^{n+2p} \rightarrow \mathbb{R}^m$ , and  $h : \mathbb{R}^{n+2p} \rightarrow \mathbb{R}^q$  are twice continuously differentiable functions. The MPEC is thus a nonlinear program (NLP) that includes the complementarity constraint

$$(1.2) \quad 0 \leq x_1 \perp x_2 \geq 0,$$

which equivalently can be written as  $x_1 \geq 0$ ,  $x_2 \geq 0$ ,  $x_1^T x_2 = 0$ . This constraint is the source of all the special properties of MPECs that distinguish them from standard NLPs, and thus needs to be handled with special care. Sometimes, MPECs are formulated with a seemingly more general complementarity condition  $0 \leq G(x) \perp H(x) \geq 0$ , where  $G$  and  $H$  are twice continuously differentiable functions mapping  $\mathbb{R}^{n+2p}$  to  $\mathbb{R}^p$ . Note, however, that an MPEC of this form can be transformed to the form (1.1) by introducing slack variables.

Although (1.1) is the most common standard form of an MPEC, the original definition is actually more general. Strictly speaking, MPECs are NLPs that contain a parametric *variational inequality* (VI) or a parametric optimization problem as additional constraint [FP03]. In the latter case, the MPEC is usually called *bilevel program* and the optimization problem in the constraints is the *lower level* optimization problem. Stackelberg games and other types of bilevel programs are very important in economics. We refer to [Bar98, Dem02] for an in depth investigation of bilevel programming. Variational inequalities are powerful tools for modelling various kinds of system equilibria, such as traffic equilibria, Nash equilibria, equilibria of forces etc., and optimization of such systems then results in MPECs. Further details on MPECs and their importance in applications can be found in [FP97, LPR96, KOZ98] and the references therein. Other applications, e.g., in structural design [FTL99] and robotics [AAP04] result directly in problems of the form

---

\*Chair of Mathematical Optimization, Fakultät für Mathematik, Boltzmannstr. 3, D-85747 Garching b. München, Germany (mulbrich@ma.tum.de).

†TU München, Chair of Mathematical Optimization, Fakultät für Mathematik, Boltzmannstr. 3, D-85747 Garching b. München, Germany (mulbrich@ma.tum.de).

(1.1). If it is necessary to avoid possible ambiguity with other types of MPECs, the problem (1.1) is called *mathematical program with complementarity constraints* (MPCC). The close connection between MPECs involving a VI and the MPEC (1.1) is given by the first order stationarity conditions for VIs. Under suitable regularity conditions, every solution of the VI satisfies stationarity conditions that can be written as a system consisting of equations and complementarity conditions. By replacing the VI with this system, an MPEC of the form (1.1) is obtained, which then usually is a relaxation of the original MPEC. Under additional monotonicity assumptions on the VI, the stationarity conditions become sufficient, and then the original MPEC is equivalent to the MPEC (1.1) that was derived via the stationarity conditions. For MPECs involving an optimization problem as constraint, the first order optimality conditions of the lower level problem can be used to derive an MPEC of the form (1.1).

Recently, significant progress has been made in the numerical solution of MPECs and the current paper contributes to these investigations. Since, as already said, the main issue in MPECs is the complementarity constraint, it is crucial to be familiar with its meaning and to know the following equivalent formulations of (1.2):

- (i)  $x_1 \geq 0, \quad x_2 \geq 0, \quad x_1^T x_2 = 0,$
- (ii)  $x_{1j} \geq 0, \quad x_{2j} \geq 0, \quad x_{1j} x_{2j} = 0, \quad j = 1, \dots, p,$
- (iii)  $x_{1j} \geq 0, \quad x_{2j} \geq 0, \quad x_{1j} = 0 \text{ or } x_{2j} = 0, \quad j = 1, \dots, p,$
- (iv)  $\min(x_{1j}, x_{2j}) = 0, \quad j = 1, \dots, p.$

Any formulation of (1.2) by smooth constraints results in positively linearly dependent gradients of the active constraints and thus in a violation of the Mangasarian-Fromovitz constraint qualification at every feasible point  $x$  of (1.1). Hence, a straightforward application of standard optimality theory for nonlinear programming is not possible for MPECs. Moreover, as the standard necessary and sufficient conditions form the basis of most nonlinear programming methods, they cannot be directly used to solve MPECs, which makes these problems difficult to solve.

Recent numerical approaches for solving MPECs follow different directions. Besides using different algorithms on the NLP level, e.g., SQP or interior point methods, they have the common property that they work with special treatment of the complementarity constraints, either by relaxation or by taking care of its degeneracy, which has very distinct structure. Regularization schemes introduce a parameter  $t \geq 0$  and formulate a family of NLPs such that, for  $t = 0$ , the MPEC is recovered, while for  $t > 0$  more regular NLPs are obtained. The smaller  $t > 0$ , the better the NLP approximates the MPEC. Hence, by solving a sequence of parametrized NLPs, one obtains a sequence of approximate solutions that, under suitable assumptions, converges to a solution of the MPEC. Approaches of this type are discussed, e.g., in [FL05, RW04, Sch01]. A prototype of such a regularization is [Sch01]

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_1^T x_2 \leq t.$$

A similar method using penalization techniques is suggested in [HR04]. Other approaches that concern exact penalization of the MPEC are described in [LPRW96, SS99]. Fletcher et al. [FL04, FLRS06] investigated the direct application of an SQP solver to the MPEC. An analysis of the global convergence of this approach was carried out by [Ani05], and superlinear local convergence is proved in [FLRS06]. Recently, the application of interior point methods to MPECs receives increasing attention [LLCN06, BR05, DFNS05].

The relaxation method we propose in this paper can be regarded as a combination of ideas from regularization schemes with the relaxation-free approach by Fletcher et al. in [FLRS06]. For  $t > 0$ , our relaxation of the complementarity condition for each pair  $(x_{1j}, x_{2j}) \in \mathbb{R}^2$  is done only on a subset of the triangle with the vertices  $(0, 0)$ ,  $(t, 0)$ , and  $(0, t)$ . Therefore, if the relaxation parameter is sufficiently small then around a local

solution  $x^*$  of (1.1) our relaxed problem only modifies the complementarity constraints that correspond to *degenerate* components of the vector  $x^*$ , i.e., for indices  $j$  with  $x_{1j}^* = x_{2j}^* = 0$ . Hence, we merge the overall relaxation for all components of [Sch01] with the exactness for the strictly complementary components of the approach of [FLRS06]. If  $t = 0$ , then the parametrized nonlinear program corresponds to the original MPEC. Therefore, we will be able to show that, under reasonable conditions, by solving a sequence of such NLPs parametrized by  $t$ , we will find a solution of (1.1) as soon as  $t > 0$  is sufficiently small.

**2. Preliminaries.** Now we review some notations and known results from MPEC theory, which we will use in the subsequent analysis of our relaxation method. For further details concerning the theory of MPECs, we refer to [FP99, SS00, Ye05].

In the following, we will work with the Lagrangian of (1.1),

$$(2.1) \quad \mathcal{L}_{\text{MPEC}}(x, \lambda, \mu, \hat{\nu}_1, \hat{\nu}_2) = f(x) - \sum_{j=1}^m \lambda_j g_j(x) - \sum_{i=1}^q \mu_i h_i(x) - \hat{\nu}_1^T x_1 - \hat{\nu}_2^T x_2.$$

Also, we define several index sets of active constraints for problem (1.1).

$$\begin{aligned} I_g(x) &= \{i \in \{1, \dots, m\} : g_i(x) = 0\}, \\ I_1(x) &= \{j \in \{1, \dots, p\} : x_{1j} = 0\}, \\ I_2(x) &= \{j \in \{1, \dots, p\} : x_{2j} = 0\}. \end{aligned}$$

There exist several constraint qualifications for MPECs. Since we will only make use of the MPEC-LICQ, we confine ourselves to review only the definition of this particular constraint qualification and omit others that exist in the literature for MPECs.

**DEFINITION 2.1.** Let  $x^*$  be a feasible point of (1.1). Then the MPEC-LICQ (*MPEC-Linear Independence Constraint Qualification*) is said to hold at  $x^*$  if the gradients

$$\begin{aligned} \nabla h_i(x^*) & \quad i \in \{1, \dots, q\}, \\ \nabla g_j(x^*) & \quad j \in I_g(x^*), \\ e_{n+j} & \quad j \in I_1(x^*), \\ e_{n+p+j} & \quad j \in I_2(x^*), \end{aligned}$$

are linearly independent. Here,  $e_j$  denotes the  $j$ -th unit vector in  $\mathbb{R}^{n+2p}$ .

Note that the definition of the MPEC-LICQ differs from the standard LICQ, as the gradients of the constraints corresponding to the complementarity condition  $x_1 \perp x_2$  are left out. However, the MPEC-LICQ represents the standard LICQ for the following related optimization problem known as *Relaxed Nonlinear Program* [FL04, FP99, SS00].

$$\begin{array}{ll} \text{RNLP} & \min \quad f(x) \\ & \text{s.t.} \quad g(x) \geq 0 \\ & \quad \quad h(x) = 0 \\ & \quad \quad x_{1j} = 0, \quad x_{2j} \geq 0 \quad j \in (I_1 \setminus I_2)(x^*) \\ & \quad \quad x_{2j} = 0, \quad x_{1j} \geq 0 \quad j \in (I_2 \setminus I_1)(x^*) \\ & \quad \quad x_{1j} \geq 0, \quad x_{2j} \geq 0 \quad j \in (I_1 \cap I_2)(x^*). \end{array}$$

Here we use the notation  $(I_1 \setminus I_2)(x^*)$  for  $I_1(x^*) \setminus I_2(x^*)$  and  $(I_1 \cap I_2)(x^*)$  for  $I_1(x^*) \cap I_2(x^*)$ , respectively.

Note that the feasible set of RNLP in this case is determined by the point  $x^*$ . Sometimes, it is useful to determine the feasible set of RNLP not by  $x^*$  but by a different point, e.g., by the current iterate  $x^k$  of an iterative method. This, however, needs to be clearly indicated to avoid ambiguity.

In the convergence analysis of our regularization method we will further use another constraint qualification that we call MPEC-CRCQ, since it corresponds to the original Constant Rank Constraint Qualification for NLPs, which was introduced by Janin in [Jan84], applied to RNLP (estimated at  $\hat{x}$ ).

DEFINITION 2.2. Let  $\hat{x}$  be feasible for the MPEC. Then the MPEC-CRCQ (*MPEC-Constant Rank Constraint Qualification*) is said to hold in  $\hat{x}$ , if for every  $\mathcal{K}_g \subset I_g(\hat{x})$ ,  $\mathcal{K}_1 \subset I_1(\hat{x})$ ,  $\mathcal{K}_2 \subset I_2(\hat{x})$  and  $\mathcal{K}_h \subset \{1, \dots, q\}$  there exists a neighbourhood  $\mathcal{U}(\hat{x})$  such that for every  $y \in \mathcal{U}(\hat{x})$  the family of gradient vectors

$$\{\nabla g_j(y) : j \in \mathcal{K}_g\} \cup \{\nabla h_j(y) : j \in \mathcal{K}_h\} \cup \{e_{1j} : j \in \mathcal{K}_1\} \cup \{e_{2j} : j \in \mathcal{K}_2\},$$

has the same rank as the family

$$\{\nabla g_j(\hat{x}) : j \in \mathcal{K}_g\} \cup \{\nabla h_j(\hat{x}) : j \in \mathcal{K}_h\} \cup \{e_{1j} : j \in \mathcal{K}_1\} \cup \{e_{2j} : j \in \mathcal{K}_2\},$$

where  $e_{1j}$  denotes  $e_{n+j} \in \mathbb{R}^{n+2p}$  and  $e_{2j}$  denotes  $e_{n+p+j} \in \mathbb{R}^{n+2p}$ , respectively.

The B-stationarity condition we mention next is the most fundamental stationarity concept for MPECs [FP99, SS00, Ye05]. It uses the tangent cone  $\mathcal{T}(\mathcal{Z}, x^*)$  of the feasible set  $\mathcal{Z}$  at  $x^*$  to state first order optimality condition.

DEFINITION 2.3. Let  $\mathcal{M} \subset \mathbb{R}^\ell$  denote a nonempty set and let  $x \in \mathcal{M}$ . The *tangent cone* of  $\mathcal{M}$  at  $x$  is defined by

$$\mathcal{T}(\mathcal{M}, x) = \{d \in \mathbb{R}^\ell : \exists \{x^k\} \subset \mathcal{M}, \exists \eta_k > 0, \eta_k \rightarrow 0 \text{ with } x^k \rightarrow x \text{ and } (x^k - x)/\eta_k \rightarrow d\}.$$

DEFINITION 2.4. Let  $\mathcal{Z}$  be the feasible set of (1.1) and let  $x^* \in \mathcal{Z}$ . Then  $x^*$  is called B-(*Bouligand*)-stationary, if

$$\nabla f(x^*)^T d \geq 0 \quad \forall d \in \mathcal{T}(\mathcal{Z}, x^*).$$

Hence, B-stationarity expresses the necessary optimality condition that there does not exist a feasible descent direction at a local optimum  $x^*$ . However, as this condition is difficult to verify, we will make use of the following stationarity concepts.

DEFINITION 2.5.

1. A point  $x^*$  is called C-(*Clarke*)-stationary if there exist multipliers  $\lambda^* \in \mathbb{R}^m$ ,  $\mu^* \in \mathbb{R}^q$ ,  $\hat{\nu}_1^* \in \mathbb{R}^p$  and  $\hat{\nu}_2^* \in \mathbb{R}^p$ , such that the following conditions

$$(2.2) \quad \begin{aligned} \nabla f(x^*) - \nabla g(x^*)\lambda^* - \nabla h(x^*)\mu^* - \begin{pmatrix} 0 \\ \hat{\nu}_1^* \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ \hat{\nu}_2^* \end{pmatrix} &= 0 \\ h(x^*) &= 0 \\ g(x^*) &\geq 0 \\ \lambda^* &\geq 0 \\ g_i(x^*)\lambda_i^* &= 0, \quad (1 \leq i \leq m) \\ x_1^* &\geq 0 \\ x_2^* &\geq 0 \\ x_{1j}^* = 0 \text{ or } x_{2j}^* &= 0, \quad (1 \leq j \leq p) \\ x_{1j}^*\hat{\nu}_{1j}^* &= 0, \quad (1 \leq j \leq p) \\ x_{2j}^*\hat{\nu}_{2j}^* &= 0, \quad (1 \leq j \leq p) \end{aligned}$$

are satisfied and

$$\forall j \in (I_1 \cap I_2)(x^*) : \hat{\nu}_{1j}^*\hat{\nu}_{2j}^* \geq 0$$

holds.

2. A point  $x^*$  is called M-(Mordukhovich)-stationary, if there exist multipliers  $\lambda^*$ ,  $\mu^*$ ,  $\hat{\nu}_1^*$ , and  $\hat{\nu}_2^*$ , such that (2.2) is satisfied and

$$\forall j \in (I_1 \cap I_2)(x^*) : \hat{\nu}_{1j}^* \hat{\nu}_{2j}^* > 0 \quad \text{or} \quad \hat{\nu}_{1j}^* \hat{\nu}_{2j}^* = 0$$

holds.

3. A point  $x^*$  is called strongly stationary, if there exist multipliers  $\lambda^*$ ,  $\mu^*$ ,  $\hat{\nu}_1^*$ , and  $\hat{\nu}_2^*$ , such that (2.2) is satisfied and

$$\forall j \in (I_1 \cap I_2)(x^*) : \hat{\nu}_{1j}^* \geq 0 \quad \text{and} \quad \hat{\nu}_{2j}^* \geq 0.$$

holds.

Note that the stationarity concepts differ only in the additional condition on the multipliers  $\hat{\nu}_{1j}^*$  and  $\hat{\nu}_{2j}^*$  for indices  $j \in (I_1 \cap I_2)(x^*)$ , i.e., for the degenerate components. Hence, if  $x^*$  satisfies strict complementarity, the stationarity conditions in Definition 2.5 are all equal. Otherwise we have the implications: *strong stationarity*  $\Rightarrow$  *M-stationarity*  $\Rightarrow$  *C-stationarity*.

It can be shown that strong stationarity corresponds to the standard stationarity for NLPs applied to the RNLP [SS00]. Furthermore, if the MPEC-LICQ holds in  $x^*$ , B-stationarity of  $x^*$  implies strong stationarity. Hence, under MPEC-LICQ strong stationarity is a necessary optimality condition.

Moreover, it is obvious that if  $x^*$  is C- (or M- or strongly) stationary and the MPEC-LICQ holds in  $x^*$ , then the multiplier tuple  $(\lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$  is unique. This follows from  $\lambda_j^* = 0$  for  $j \notin I_g(x^*)$ ,  $\hat{\nu}_{1j}^* = 0$  for  $j \notin I_1(x^*)$ ,  $\hat{\nu}_{2j}^* = 0$  for  $j \notin I_2(x^*)$  and from the fact that the rest of the multipliers is uniquely determined from

$$\sum_{j \in I_g(x^*)} \lambda_j^* \nabla g_j(x^*) + \sum_{i=1}^q \mu_i^* \nabla h_i(x^*) + \sum_{j \in I_1(x^*)} \hat{\nu}_{1j}^* e_{n+j} + \sum_{j \in I_2(x^*)} \hat{\nu}_{2j}^* e_{n+p+j} = \nabla f(x^*).$$

Next, before defining the second order optimality conditions, we introduce the following three sets of critical directions in  $x^*$ , following the definitions in [RW04]. Let

$$\begin{aligned} \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) = \{ (d_0, d_1, d_2) \in \mathbb{R}^{n+2p} \setminus \{0\} : \\ \nabla h_i(x^*)^T d = 0, \quad i \in \{1, \dots, q\} \\ \nabla g_j(x^*)^T d = 0, \quad j \in I_g(x^*) \quad \text{and} \quad \lambda_j^* > 0 \\ \nabla g_j(x^*)^T d \geq 0, \quad j \in I_g(x^*) \quad \text{and} \quad \lambda_j^* = 0 \\ d_{1j} = 0, \quad j \in (I_1 \setminus I_2)(x^*) \\ d_{1j} = 0, \quad j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{1j}^* > 0 \\ d_{1j} \geq 0, \quad j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{1j}^* = 0 \\ d_{2j} = 0, \quad j \in (I_2 \setminus I_1)(x^*) \\ d_{2j} = 0, \quad j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{2j}^* > 0 \\ d_{2j} \geq 0, \quad j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{2j}^* = 0 \} \end{aligned}$$

and let

$$\mathcal{S}^*(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) = \{ (d_0, d_1, d_2) \in \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) : \\ \min(d_{1j}, d_{2j}) = 0, \quad j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{1j}^* = \hat{\nu}_{2j}^* = 0 \}.$$

These two sets differ only in the additional condition on the components of a direction  $d \in \mathbb{R}^{n+2p} \setminus \{0\}$  corresponding to indices  $j \in (I_1 \cap I_2)(x^*)$ , with  $\hat{\nu}_{1j}^* = \hat{\nu}_{2j}^* = 0$ . Therefore the defined two sets satisfy the relationship  $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) \subset \mathcal{S}^*(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ .

DEFINITION 2.6.

1. Let  $x^*$  be a B-stationary point of (1.1) and suppose the MPEC-LICQ holds in  $x^*$ . Let  $(\lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$  be the unique multipliers of  $x^*$  satisfying (2.2). Then  $x^*$  satisfies the MPEC-SOSC (*MPEC-Second Order Sufficient Condition*), if

$$(2.3) \quad d^T \nabla_{xx} \mathcal{L}_{\text{MPEC}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) d > 0$$

holds for all  $d = (d_0, d_1, d_2) \in \mathcal{S}^*(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ .

2. If (2.3) holds for all  $d \in \bar{S}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ , then  $x^*$  is said to satisfy the RNLP-SOSC.

Considering the definitions of the sets, we obtain for the second order conditions the relationship RNLP-SOSC  $\Rightarrow$  MPEC-SOSC. Note that the RNLP-SOSC corresponds to the standard SOSC [Fle00] for NLPs applied to the RNLP.

Finally we state a result, concerning the second order sufficient conditions which is a simple application of the appropriate result in [SS00].

**THEOREM 2.1.** *If  $x^*$  is a strongly stationary point of the MPEC (1.1) that satisfies the MPEC-LICQ, as well as the MPEC-SOSC, then  $x^*$  is a strict local minimum of (1.1).*

Hence the MPEC-SOSC is sufficient to guarantee the local optimality for (1.1) of a strongly stationary point  $x^*$ .

**3. Relaxation.** To derive the relaxation scheme proposed in this work, we first consider the scalar complementarity condition

$$(3.1) \quad x_1 \geq 0, \quad x_2 \geq 0, \quad x_1 x_2 = 0.$$

Now if we introduce new Cartesian coordinates

$$y := x_1 + x_2 \quad \text{and} \quad z := x_1 - x_2,$$

then the constraints of (3.1) can be written as

$$y \geq -z, \quad y \geq z, \quad y \leq |z|.$$

To smooth out the kink of the absolute value function on the interval  $[-1, 1]$ , we define a function  $\theta : \mathcal{D} \subset \mathbb{R} \rightarrow \mathbb{R}$  where  $\mathcal{D}$  is an open subset of  $\mathbb{R}$  with  $[-1, 1] \subset \mathcal{D}$ , and  $\theta(z)$  satisfies the following conditions

ASSUMPTIONS 3.1.

1.  $\theta$  is twice continuously differentiable on  $[-1, 1]$ ,
2.  $\theta(1) = \theta(-1) = 1$ ,
3.  $\theta'(-1) = -1$  and  $\theta'(1) = 1$
4.  $\theta''(-1) = \theta''(1) = 0$  and
5.  $\theta$  is strictly convex on  $[-1, 1]$ .

Throughout this paper we will assume that Assumptions 3.1 hold. A suitable example for  $\theta$  is

$$(3.2) \quad \theta(z) := \frac{2}{\pi} \sin \left( z \frac{\pi}{2} + \frac{3\pi}{2} \right) + 1.$$

Combining  $\theta$  inside  $[-1, 1]$  with the absolute value function outside  $[-1, 1]$  and scaling the interval leads to a  $C^2$ -function

$$(3.3) \quad \phi(z, t) := \begin{cases} |z| & |z| \geq t \\ t \theta\left(\frac{z}{t}\right) & |z| < t, \end{cases}$$

which gives us a relaxation of the form

$$(3.4) \quad y \geq -z, \quad y \geq z, \quad y \leq \phi(z, t).$$

Switching back to our original coordinate system yields

$$x_1 + x_2 \geq x_2 - x_1, \quad x_1 + x_2 \geq x_1 - x_2, \quad x_1 + x_2 \leq \varphi(x_1, x_2, t),$$

with  $\varphi(x_1, x_2, t) := \phi(x_1 - x_2, t)$ . The first two conditions are equivalent to  $x_1 \geq 0$  and  $x_2 \geq 0$ . Hence, relaxing each of the complementarity constraints of the MPEC as described

above, we get a parametric nonlinear program  $R(t)$  of the form:

$$(3.5) \quad \begin{array}{ll} R(t) & \min \quad f(x) \\ & \text{s.t.} \quad g(x) \geq 0 \\ & \quad \quad h(x) = 0 \\ & \quad \quad x_1, x_2 \geq 0 \\ & \quad \quad \Phi(x_1, x_2, t) \leq 0, \end{array}$$

where  $\Phi(x_1, x_2, t) : \mathbb{R}^{2p} \times \mathbb{R}_0^+ \rightarrow \mathbb{R}^p$

$$\Phi_j(x_1, x_2, t) := x_{1j} + x_{2j} - \varphi(x_{1j}, x_{2j}, t)$$

and

$$\varphi(x_{1j}, x_{2j}, t) = \phi(x_{1j} - x_{2j}, t) = \begin{cases} |x_{1j} - x_{2j}| & |x_{1j} - x_{2j}| \geq t \\ t \theta\left(\frac{x_{1j} - x_{2j}}{t}\right) & |x_{1j} - x_{2j}| < t \end{cases}$$

For the discussion of  $R(t)$  we need the following facts concerning the function  $\theta$ .

LEMMA 3.1. *Let  $\theta$  satisfy the Assumption 3.1. Then for all  $z \in (-1, 1)$ , there holds*

$$\theta(z) > |z|, \quad |\theta'(z)| < 1.$$

Furthermore, for  $\eta > 0$  and  $|z| < \eta$  the function  $t \mapsto \phi(z, t)$  defined in (3.3) is strictly monotonically increasing on  $[\eta, \infty)$ .

*Proof.*

Let  $z \in (-1, 1)$  be arbitrary. Due to the strict convexity of  $\theta$  on  $[-1, 1]$ , there holds

$$\begin{aligned} \theta(z) &> \theta(1) + \theta'(1)(z - 1) = 1 + z - 1 = z, \\ \theta(z) &> \theta(-1) + \theta'(-1)(z + 1) = 1 + (-1) \cdot (z + 1) = -z, \end{aligned}$$

which shows  $\theta(z) > |z|$ . Furthermore,

$$1 = \theta(1) > \theta(z) + \theta'(z)(1 - z) > |z| + \theta'(z)(1 - z) \geq z + \theta'(z)(1 - z),$$

Hence,  $\theta'(z) < \frac{1 - z}{1 - z} = 1$ . In the same way,

$$1 = \theta(-1) > \theta(z) + \theta'(z)(-1 - z) > |z| - \theta'(z)(1 + z) \geq -z - \theta'(z)(1 + z),$$

Thus,  $-\theta'(z) < \frac{1 + z}{1 + z} = 1$ .

Now let  $\eta > 0$ ,  $|z| < \eta$ , and consider  $t \geq \eta$ . Then  $|z| < \eta \leq t$  and thus  $\phi(z, t) = t\theta(z/t)$ . Hence, the strict monotonicity of  $\phi(z, t)$  w.r.t.  $t$  follows from

$$\frac{d}{dt}\phi(z, t) = \theta(z/t) - t\theta'(z/t)(z/t^2) > \left|\frac{z}{t}\right| - 1 \cdot \left|\frac{z}{t}\right| = 0.$$

□

With this additional information in hand, we once again briefly summarize the derivation of the relaxation  $R(t)$ : The complementarity condition  $x_{1j} \geq 0$ ,  $x_{2j} \geq 0$ ,  $x_{1j}x_{2j} = 0$  is equivalent to

$$x_{1j} \geq 0, \quad x_{2j} \geq 0, \quad x_{1j} + x_{2j} - |x_{1j} - x_{2j}| \leq 0.$$

For  $|x_{1j} - x_{2j}| \geq t$  we use this formulation. For  $|x_{1j} - x_{2j}| < t$ , however, we relax the third constraint according to

$$\Phi_j(x_1, x_2, t) = x_{1j} + x_{2j} - t\theta((x_{1j} - x_{2j})/t) \leq 0.$$

This is a true relaxation since, by Lemma 3.1, for  $|x_{1j} - x_{2j}| < t$  there holds

$$t\theta((x_{1j} - x_{2j})/t) > |x_{1j} - x_{2j}|.$$

Figure 3 illustrates how the derived relaxation method combines the regularization scheme proposed by Scholtes in [Sch01] with the approach to solve the MPEC by using exact NCP-functions (for example the *Min-function*) analysed in [Ley06].

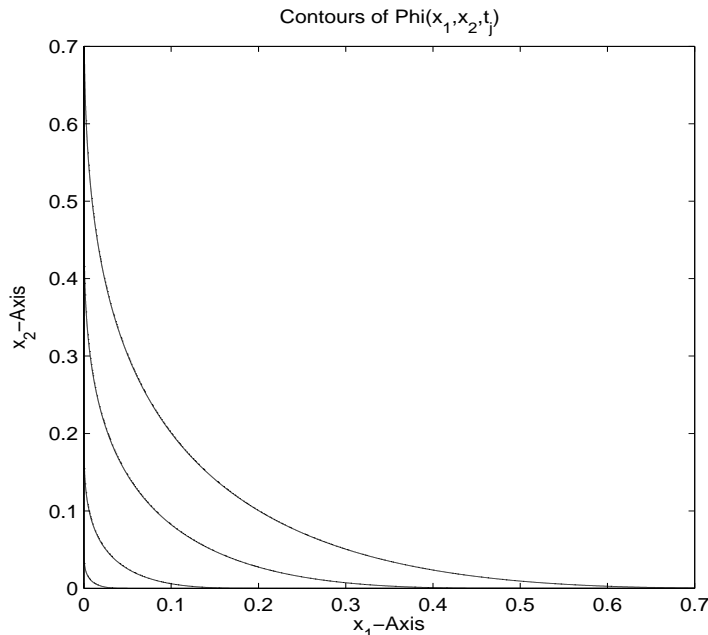


FIG. 3.1. contours of  $\Phi_j(x_1, x_2, t) = 0$  with  $\theta(z)$  defined as in (3.2) for different values of  $t > 0$ , namely  $t \in \{0.05, 0.2, 0.5, 0.8\}$

For decreasing  $t \rightarrow 0^+$ , the relaxation is increasingly localized to smaller and smaller neighborhoods of  $(x_{1j}, x_{2j}) = (0, 0)$ .

In the following two Lemmas 3.2 and 3.3 we essentially prove that the feasible sets  $\mathcal{Z}(t)$  of  $R(t)$  converge monotonically decreasing to the feasible set  $\mathcal{Z}$  of the MPEC (1.1) and that the feasible sets of (1.1) and  $R(0)$  coincide. Furthermore, only the complementarity constraints are relaxed, and the modification concerning the pair  $(x_{1j}, x_{2j})$  only takes place within the triangle with vertices  $(0, 0)$ ,  $(t, 0)$ ,  $(0, t)$ . In particular, the modification  $\mathcal{Z}(t)$  of the feasible set  $\mathcal{Z}$  only affects pairs  $(x_{1j}, x_{2j}) \geq 0$  with  $\max(x_{1j}, x_{2j}) < t$ . Hence, the modification of the feasible set is not felt by a strictly complementary pair  $(x_{1j}, x_{2j}) \geq 0$  as soon as  $t \leq \max(x_{1j}, x_{2j})$ .

We now prove these sketched results rigorously.

LEMMA 3.2. For  $t \geq 0$ , define the sets

$$\begin{aligned} C &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1 x_2 = 0\}, \\ C(t) &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 - \varphi(x_1, x_2, t) \leq 0\}, \\ T(t) &= \text{conv}\{(0, 0), (t, 0), (0, t)\}. \end{aligned}$$

Then the following holds

$$\begin{aligned} C &\subset C(t) \subset C \cup T(t) && \forall t \geq 0, \\ C(t_1) &\subsetneq C(t_2) && \forall t_2 > t_1 \geq 0, \\ \bigcap_{t>0} C(t) &= C(0) = C. \end{aligned}$$



*Proof.*

We first prove  $C \subset C(t)$ . Let  $(x_1, x_2) \in C$ .

If  $|x_1 - x_2| \geq t$  then

$$(3.6) \quad x_1 + x_2 - \varphi(x_1, x_2, t) = x_1 + x_2 - |x_1 - x_2| = 2 \min(x_1, x_2) = 0.$$

If  $|x_1 - x_2| < t$  then  $|(x_1 - x_2)/t| < 1$  and thus

$$(3.7) \quad \begin{aligned} x_1 + x_2 - \varphi(x_1, x_2, t) &= x_1 + x_2 - t\theta((x_1 - x_2)/t) \\ &< x_1 + x_2 - |x_1 - x_2| = 2 \min(x_1, x_2) = 0. \end{aligned}$$

Next, we prove  $C(t) \subset C \cup T(t)$ . Consider  $(x_1, x_2) \in C(t)$ .

If  $|x_1 - x_2| \geq t$  then

$$0 \geq x_1 + x_2 - \varphi(x_1, x_2, t) = x_1 + x_2 - |x_1 - x_2| = 2 \min(x_1, x_2),$$

which together with  $x_1, x_2 \geq 0$  implies  $x_1 x_2 = 0$  and thus  $(x_1, x_2) \in C$ .

If  $|x_1 - x_2| < t$  then  $(x_1 - x_2)/t \in (-1, 1)$ . Using that  $\theta$  is convex with  $\theta(-1) = \theta(1) = 1$ , we obtain  $\theta((x_1 - x_2)/t) \leq 1$  and thus

$$0 \geq x_1 + x_2 - \varphi(x_1, x_2, t) = x_1 + x_2 - t\theta((x_1 - x_2)/t) \geq x_1 + x_2 - t.$$

This shows that  $(x_1, x_2) \in T(t)$ .

Next, we prove  $C(t_1) \subsetneq C(t_2)$  for  $0 \leq t_1 < t_2$ . Consider  $(x_1, x_2)$  with  $x_1, x_2 \geq 0$ .

For  $|x_1 - x_2| \geq t_2$ , there holds

$$\varphi(x_1, x_2, t_1) = |x_1 - x_2| = \varphi(x_1, x_2, t_2).$$

Hence, in this case,  $(x_1, x_2) \in C(t_1)$  implies  $(x_1, x_2) \in C(t_2)$ .

For  $t_1 \leq |x_1 - x_2| < t_2$ , we obtain

$$(3.8) \quad \varphi(x_1, x_2, t_2) = t_2\theta((x_1 - x_2)/t_2) > |x_1 - x_2| = \varphi(x_1, x_2, t_1).$$

Thus, also in this case,  $(x_1, x_2) \in C(t_1)$  implies  $(x_1, x_2) \in C(t_2)$ .

Setting, e.g.,  $x_1 = \frac{1}{2}(t_2\theta(t_1/t_2) + t_1)$ ,  $x_2 = \frac{1}{2}(t_2\theta(t_1/t_2) - t_1)$  gives a pair  $(x_1, x_2)$  with  $x_1, x_2 > 0$ ,  $x_1 - x_2 = t_1$ , and  $x_1 + x_2 - \varphi(x_1, x_2, t_2) = 0$ . Here we have used  $t_2\theta(t_1/t_2) > t_1$ . Thus, (3.8) yields  $(x_1, x_2) \in C(t_2) \setminus C(t_1)$ .

Next, consider the case  $|x_1 - x_2| < t_1$ . Since by Lemma 3.1, the function  $t \mapsto \phi(x_1 - x_2, t)$  is strictly monotonically increasing for  $t \geq t_1$ , it follows that

$$\varphi(x_1, x_2, t_2) = \phi(x_1 - x_2, t_2) > \phi(x_1 - x_2, t_1) = \varphi(x_1, x_2, t_1).$$

We conclude that in the case  $|x_1 - x_2| < t_1$  under consideration,  $(x_1, x_2) \in C(t_1)$  implies  $(x_1, x_2) \in C(t_2)$ . In addition, the third constraint in  $C(t_1)$  is stricter than the one in  $C(t_2)$ .

The last assertion follows from

$$C \subset C(0) \subset \bigcap_{t>0} C(t) \subset C \cup \bigcap_{t>0} T(t) = C \cup T(0) = C.$$

□

From this, we can obtain relations between the feasible sets  $\mathcal{Z}(t)$  of  $R(t)$  for different values of  $t$  and also between  $\mathcal{Z}(t)$  and the feasible set  $\mathcal{Z}$  of (1.1).

LEMMA 3.3. *Denote by  $\mathcal{Z}$  the feasible set of (1.1) and by  $\mathcal{Z}(t)$ ,  $t \geq 0$ , the feasible set of  $R(t)$ . Then for all  $0 \leq t_1 < t_2$ , there holds*

$$(3.9) \quad \mathcal{Z}(t_1) \subset \mathcal{Z}(t_2), \quad \mathcal{Z} = \mathcal{Z}(0) = \bigcap_{t>0} \mathcal{Z}(t).$$

Furthermore, if there exists  $x = (x_0, x_1, x_2) \in \mathcal{Z}(t_2)$  with  $\Phi_j(x_1, x_2, t_2) = 0$  and  $|x_{1j} - x_{2j}| < t_1$  for at least one  $j \in \{1, \dots, p\}$ , then  $x \notin \mathcal{Z}(t_1)$ .

*Proof.* For  $0 \leq t_1 < t_2$ , the feasible sets  $\mathcal{Z}$ ,  $\mathcal{Z}(t_1)$ , and  $\mathcal{Z}(t_2)$  only differ in the constraints concerning (relaxed) complementarity:

$$\mathcal{Z} : (x_{1j}, x_{2j}) \in C, \quad \mathcal{Z}(t_1) : (x_{1j}, x_{2j}) \in C(t_1), \quad \mathcal{Z}(t_2) : (x_{1j}, x_{2j}) \in C(t_2).$$

Here we have used the notation of Lemma 3.2. There, it was shown that

$$C = C(0) \subset C(t_1) \subsetneq C(t_2), \quad C = C(0) = \bigcap_{t>0} C(t).$$

This proves all assertions in (3.9).

Now, let  $(x_0, x_1, x_2) \in \mathcal{Z}(t_2)$  with  $\Phi_j(x_1, x_2, t_2) = 0$  and  $|x_{1j} - x_{2j}| < t_1$  for at least one  $j \in \{1, \dots, p\}$ . Then, by Lemma 3.1,  $\phi(x_{1j} - x_{2j}, t)$  is strictly monotonically increasing w.r.t.  $t$  for  $t \geq t_1$ , and thus

$$x_{1j} + x_{2j} = \varphi(x_{1j}, x_{2j}, t_2) = \phi(x_{1j} - x_{2j}, t_2) > \phi(x_{1j} - x_{2j}, t_1) = \varphi(x_{1j}, x_{2j}, t_1).$$

Therefore  $\Phi_j(x_1, x_2, t_1) > 0$  and consequently  $(x_0, x_1, x_2) \notin \mathcal{Z}(t_1)$ . □

The following lemma makes assertions on the activity of the constraints  $\Phi_j(x_1, x_2, t) \leq 0$  at feasible points of the MPEC (1.1).

LEMMA 3.4. *Let  $x$  be feasible for (1.1) and  $t > 0$ . Then  $x$  is feasible for  $R(t)$  and there holds:*

1. *If  $\max(x_{1j}, x_{2j}) < t$ , then  $\Phi_j(x_1, x_2, t) < 0$ .*
2. *If  $\max(x_{1j}, x_{2j}) \geq t$  then  $\Phi_j(x_1, x_2, t) = 0$ .*

*Proof.* The feasibility of  $x \in \mathcal{Z}$  for  $R(t)$  follows from Lemma 3.3.

To proceed, we note that for  $x \in \mathcal{Z}$  there holds  $x_{1j}, x_{2j} \geq 0$ ,  $x_{1j}x_{2j} = 0$ , which implies  $\max(x_{1j}, x_{2j}) = |x_{1j} - x_{2j}|$ . Now, if  $\max(x_{1j}, x_{2j}) < t$ , then, by Lemma 3.1,

$$\begin{aligned} \Phi_j(x_1, x_2, t) &= x_{1j} + x_{2j} - t\theta \left( \frac{x_{1j} - x_{2j}}{t} \right) < x_{1j} + x_{2j} - t \left| \frac{x_{1j} - x_{2j}}{t} \right| \\ &= x_{1j} + x_{2j} - |x_{1j} - x_{2j}| = 2 \min(x_{1j}, x_{2j}) = 0. \end{aligned}$$

On the other hand, if  $\max(x_{1j}, x_{2j}) \geq t$  then  $|x_{1j} - x_{2j}| \geq t$  and since  $x_{1j} \geq 0$ ,  $x_{2j} \geq 0$  and  $x_{1j}x_{2j} = 0$  it follows by (3.6) that  $\Phi_j(x_1, x_2, t) = 0$ . □

From Lemma 3.3 it follows that for every sequence  $(t_k)$  with  $0 \leq t_{k+1} < t_k < \dots < t_0$  we obtain

$$\mathcal{Z}(0) \subset \mathcal{Z}(t_{k+1}) \subset \mathcal{Z}(t_k) \subset \dots \subset \mathcal{Z}(t_0).$$

Using this information we can prove the next result, which compares the strict minima of  $R(t)$  for different values of the parameter  $t$ .

LEMMA 3.5. *Let  $x^*$  be a strict minimum of  $R(\hat{t})$  in an  $\varepsilon$ -neighbourhood  $\mathcal{B}_\varepsilon(x^*)$  of  $x^*$  that satisfies the complementarity constraints  $0 \leq x_1 \perp x_2 \geq 0$ . Then  $x^*$  is a strict minimum of  $R(t)$  for every  $t \in [0, \hat{t}]$  in the same  $\varepsilon$ -neighbourhood  $\mathcal{B}_\varepsilon(x^*)$ .*

*Proof.* Since  $x^*$  is a strict local minimum of  $R(\hat{t})$  that satisfies the complementarity constraints, it follows that  $x^* \in \mathcal{Z}(0) \subset \mathcal{Z}(t)$  for all  $t > 0$ . Moreover, by Lemma 3.3, we have

$$(3.10) \quad \forall x \in (\mathcal{Z}(t) \cap \mathcal{B}_\varepsilon(x^*)) \subset (\mathcal{Z}(\hat{t}) \cap \mathcal{B}_\varepsilon(x^*)) : f(x) > f(x^*),$$

for every  $t \in [0, \hat{t}]$ , i.e.,  $x^*$  is also a strict minimum for  $R(t)$  for every  $t \in [0, \hat{t}]$  in the same  $\varepsilon$ -neighbourhood  $\mathcal{B}_\varepsilon(x^*)$  of  $x^*$ .  $\square$

Let  $\mathcal{L}_{R(t)}$  denote the corresponding Lagrangian function for  $R(t)$ , with

$$(3.11) \quad \mathcal{L}_{R(t)}(x, \lambda, \mu, \nu_1, \nu_2, \xi) = f(x) - \sum_{j \in I_g} \lambda_j g_j(x) - \sum_{i=1}^q \mu_i h_i(x) - \nu_1^T x_1 - \nu_2^T x_2 + \sum_{j=1}^p \xi_j \Phi_j(x_1, x_2, t).$$

In the sequel  $x$  is said to be a stationary point for  $R(t)$  if there exist multipliers  $\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*$  such that

$$(3.12) \quad \begin{aligned} \nabla f(x) - \nabla g(x)\lambda^* - \nabla h(x)\mu^* - \begin{pmatrix} 0 \\ \nu_1^* \\ \nu_2^* \end{pmatrix} + \begin{pmatrix} 0 \\ \nabla_{x_1} \Phi(x_1, x_2, t) \xi^* \\ \nabla_{x_2} \Phi(x_1, x_2, t) \xi^* \end{pmatrix} &= 0 \\ h(x) &= 0 \\ g(x) &\geq 0 \\ \lambda^* &\geq 0 \\ g_j(x)\lambda_j^* &= 0, \quad \forall j \\ x_1 \geq 0, \quad x_2 &\geq 0 \\ \Phi(x_1, x_2, t) &\leq 0 \\ \nu_1^* \geq 0, \quad \nu_2^* &\geq 0 \\ \xi^* &\geq 0 \\ x_{1j}\nu_{1j}^* = 0, \quad x_{2j}\nu_{2j}^* &= 0, \quad \forall j \\ \xi_j^* \Phi_j(x_1, x_2, t) &= 0, \quad \forall j. \end{aligned}$$

with  $\nabla_{x_1} \Phi(x_1, x_2, t)$  denoting the transposed Jacobian of  $\Phi(x_1, x_2, t)$  with respect to  $x_1$  (the same holds accordingly for  $\nabla_{x_2} \Phi(x_1, x_2, t)$ ,  $\nabla g(x)$  and  $\nabla h(x)$ ). In view of the definition of  $\Phi_j(x_1, x_2, t)$ , the matrix  $\nabla \Phi(x_1, x_2, t)$  is very sparse. Its entries  $\partial \Phi_i(x_1, x_2, t) / \partial x_{1j}$ ,  $\partial \Phi_i(x_1, x_2, t) / \partial x_{2j}$  all vanish except for the two diagonals

$$\alpha_j := \frac{\partial \Phi_j(x_1, x_2, t)}{\partial x_{1j}} = \begin{cases} 0 & x_{1j} \geq x_{2j} + t \\ 2 & x_{1j} \leq x_{2j} - t \\ 1 - \theta'(\frac{x_{1j} - x_{2j}}{t}) & \text{otherwise} \end{cases}$$

$$\beta_j := \frac{\partial \Phi_j(x_1, x_2, t)}{\partial x_{2j}} = \begin{cases} 0 & x_{2j} \geq x_{1j} + t \\ 2 & x_{2j} \leq x_{1j} - t \\ 1 + \theta'(\frac{x_{1j} - x_{2j}}{t}) & \text{otherwise} \end{cases},$$

i.e.,

$$\nabla \Phi(x_1, x_2, t) = \begin{pmatrix} 0 \\ D_1 \\ D_2 \end{pmatrix}$$

with  $D_1 = \text{diag}(\alpha_j)_{j \in \{1, \dots, p\}}$  and  $D_2 = \text{diag}(\beta_j)_{j \in \{1, \dots, p\}}$ . Note that due to the conditions on  $\theta(z)$  for all feasible points  $x \in \mathbb{R}^{n+2p}$  we obtain  $0 \leq \alpha_j \leq 2$  and  $0 \leq \beta_j \leq 2$ . Moreover, the values of  $\alpha_j$  and  $\beta_j$  are strictly monotone increasing for  $z := (x_{1j} - x_{2j})/t$  in  $(-1, 1)$ .

LEMMA 3.6. Let  $\theta(z)$  satisfy Assumption 3.1 and let  $\alpha(z) := 1 - \theta'(z)$  and  $\beta(z) := 1 + \theta'(z)$ . Then  $\alpha(z)$  is strictly monotone decreasing and  $\beta(z)$  is strictly monotone increasing for  $z \in (-1, 1)$ .

*Proof.* If  $\theta(z)$  satisfies the Assumptions 3.1, then for all  $z \in (-1, 1)$

$$\frac{\partial \alpha}{\partial z} = -\theta''(z) < 0 \quad \text{and} \quad \frac{\partial \beta}{\partial z} = \theta''(z) > 0.$$

□

We will denote the active set of  $R(t)$  for the constraints  $\Phi_j(x_1, x_2, t) \leq 0$  by

$$I_\Phi(x, t) = \{j \in \{1, \dots, p\} : \Phi_j(x_1, x_2, t) = 0\}.$$

The following lemma relates the active index sets  $I_1(x)$  and  $I_2(x)$  to  $I_\Phi(x, t)$ . These relations will later be used in the analysis of the stationary points and the local minima of the MPEC and of  $R(t)$ .

LEMMA 3.7. Let  $t > 0$  and let  $x$  be a feasible point of  $R(t)$ , then it follows that

1. if  $j \in I_1(x)$  and  $x_{2j} \geq t$  or  $j \in I_2(x)$  and  $x_{1j} \geq t$ , then  $j \in I_\Phi(x, t)$  and the corresponding gradients of the active constraints are positively linearly dependent.
2. If  $x_{1j} < t$  and  $x_{2j} < t$ , then  $I_\Phi(x, t) \cap (I_1 \cup I_2)(x) = \emptyset$ .

*Proof.* To prove the first part suppose  $x$  is feasible for  $R(t)$  and  $j \in I_1(x)$  and  $x_{2j} \geq t$  or  $j \in I_2(x)$  and  $x_{1j} \geq t$ , then  $|x_{1j} - x_{2j}| \geq t$ . Hence, since  $x_{1j} \geq 0$ ,  $x_{2j} \geq 0$  and  $x_{1j}x_{2j} = 0$  it follows by (3.6) that  $\Phi_j(x_1, x_2, t) = 0$ , such that  $j \in I_\Phi(x, t)$ .

By the special structure of  $\nabla_x \Phi_j(x_1, x_2, t)$  and the values of  $\alpha_j$  and  $\beta_j$ , respectively, for feasible  $x$  with  $|x_{1j} - x_{2j}| \geq t$ , it follows that

$$\nabla \Phi_j(x_1, x_2, t) = 2e_{1j} \quad \text{or} \quad \nabla \Phi_j(x_1, x_2, t) = 2e_{2j},$$

where  $e_{1j}$  and  $e_{2j}$  are defined as in Definition 2.2. Therefore, we obtain positive linear dependence with the gradient  $-e_{1j}$  of the constraint  $x_{1j} \leq 0$  or with the gradient  $-e_{2j}$  of the constraint  $x_{2j} \leq 0$ , respectively.

Now let  $x_{1j} < t$  and  $x_{2j} < t$  then  $|x_{1j} - x_{2j}| < t$ . Assume there exists an index  $j \in I_\Phi(x, t) \cap (I_1 \cup I_2)(x)$ . Then  $\min(x_{1j}, x_{2j}) = 0$  and  $\Phi_j(x_1, x_2, t) = 0$  in contradiction to what we have shown in (3.7). □

Combining the two statements of Lemma 3.7 leads to the fact that either the constraints  $x_{1j} \geq 0$  or  $x_{2j} \geq 0$  and  $\Phi_j(x_1, x_2, t) \leq 0$  are both active and the corresponding gradients are positively linearly dependent or at most one of them appears in the KKT-system with a possibly non-vanishing multiplier.

**4. Properties of the Relaxation.** In this section we relate the stationary points of the MPEC to the stationary points of  $R(t)$ . However a crucial assumption we need concerns the parameter  $t \geq 0$ , which is assumed to be small enough. As for some results the maximal value for  $t$  depends on the stationary point  $x^*$  of the MPEC, we need the following definition.

DEFINITION 4.1. Let  $x^*$  be a strongly stationary point of (1.1), then we define

$$\tau(x^*) := \min\{x_{ij}^* : i \in \{1, 2\}, j \in \{1, \dots, p\} \text{ and } x_{ij}^* > 0\}.$$

If  $x_{ij}^* = 0$  for all  $j \in \{1, \dots, p\}$  and  $i \in \{1, 2\}$ , then we set  $\tau(x^*) := +\infty$ .

We are now able to state the relation between the stationary points of both (1.1) and  $R(t)$ .

LEMMA 4.1. Let  $x^*$  be a strongly stationary point of (1.1) with multipliers  $\lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*$ . Choose

$$(4.1) \quad \begin{aligned} \nu_{1j}^* &\geq \max(0, \hat{\nu}_{1j}^*), & \xi_j^* &= \frac{\nu_{1j}^* - \hat{\nu}_{1j}^*}{2}, & \nu_{2j}^* &= \hat{\nu}_{2j}^* = 0 & (j \in (I_1 \setminus I_2)(x^*)) \\ \nu_{2j}^* &\geq \max(0, \hat{\nu}_{2j}^*), & \xi_j^* &= \frac{\nu_{2j}^* - \hat{\nu}_{2j}^*}{2}, & \nu_{1j}^* &= \hat{\nu}_{1j}^* = 0 & (j \in (I_2 \setminus I_1)(x^*)) \\ \nu_{1j}^* &= \hat{\nu}_{1j}^*, & \nu_{2j}^* &= \hat{\nu}_{2j}^*, & \xi_j^* &= 0 & (j \in (I_1 \cap I_2)(x^*)) \end{aligned}$$

Then  $(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$  satisfies the stationarity conditions (3.12) of  $R(t)$  for every  $t \in (0, \tau(x^*))$ .

*Proof.* First note that since  $x^*$  is feasible for (1.1), it follows by Lemma 3.3 that  $x^*$  is feasible for  $R(t)$  for all  $t > 0$ . Now consider the multipliers defined by (4.1). First of all, since strong stationarity implies  $\hat{\nu}_{1j}^* \geq 0$  and  $\hat{\nu}_{2j}^* \geq 0$  for all  $j \in (I_1 \cap I_2)(x^*)$ , we obtain

$$\nu_1^* \geq 0, \quad \nu_2^* \geq 0, \quad \text{and} \quad \xi^* \geq 0.$$

Moreover, since

$$x_{1j}^* \hat{\nu}_{1j}^* = 0 \quad \text{and} \quad x_{2j}^* \hat{\nu}_{2j}^* = 0 \quad \text{for all} \quad j \in \{1, \dots, p\}$$

it follows that

$$x_{1j}^* \nu_{1j}^* = 0 \quad \text{and} \quad x_{2j}^* \nu_{2j}^* = 0 \quad \text{for all} \quad j \in \{1, \dots, p\}.$$

Now let  $t \in (0, \tau(x^*))$ . Then by the definition of  $\tau(x^*)$ , it follows that  $x_{2j}^* \geq t$  for all  $j \in (I_1 \setminus I_2)(x^*)$  and  $x_{1j}^* \geq t$  for all  $j \in (I_2 \setminus I_1)(x^*)$ . Therefore  $\max(x_{1j}^*, x_{2j}^*) \geq t$  if  $j \in (I_1 \setminus I_2)(x^*)$  or  $j \in (I_2 \setminus I_1)(x^*)$ . By Lemma 3.4, this implies

$$(4.2) \quad \Phi_j(x_{1j}^*, x_{2j}^*, t) = 0 \quad \text{for all} \quad j \in (I_1 \setminus I_2)(x^*) \cup (I_2 \setminus I_1)(x^*).$$

By definition,  $\xi_j^*$  can only be strictly positive if either  $j \in (I_1 \setminus I_2)(x^*)$  or  $j \in (I_2 \setminus I_1)(x^*)$ . Hence by (4.2) there holds  $\xi_j^* \Phi_j(x_{1j}^*, x_{2j}^*, t) = 0$  for all  $j \in \{1, \dots, p\}$ .

Furthermore, (4.2) implies

$$(4.3) \quad \begin{array}{ll} \alpha_j = 0 & \text{and} \quad \beta_j = 2 & \text{for} \quad j \in (I_2 \setminus I_1)(x^*) \\ \alpha_j = 2 & \text{and} \quad \beta_j = 0 & \text{for} \quad j \in (I_1 \setminus I_2)(x^*). \end{array}$$

Now notice that the multipliers  $\nu_1^*, \nu_2^*$  and  $\xi^*$  defined by (4.1) satisfy the conditions

$$(4.4) \quad \begin{array}{lll} \hat{\nu}_{1j}^* = \nu_{1j}^* - 2\xi_j^* & \text{and} \quad \hat{\nu}_{2j}^* = \nu_{2j}^* & \text{for } j \in (I_1 \setminus I_2)(x^*) \\ \hat{\nu}_{2j}^* = \nu_{2j}^* - 2\xi_j^* & \text{and} \quad \hat{\nu}_{1j}^* = \nu_{1j}^* & \text{for } j \in (I_2 \setminus I_1)(x^*) \\ \hat{\nu}_{1j}^* = \nu_{1j}^* - \alpha_j \xi_j^* = \nu_{1j}^* & \text{and} \quad \hat{\nu}_{2j}^* = \nu_{2j}^* - \beta_j \xi_j^* = \nu_{2j}^* & \text{for } j \in (I_1 \cap I_2)(x^*) \end{array}$$

which are (in view of (4.3)) exactly the conditions that we obtain for the multipliers  $\nu_1^*, \nu_2^*$  and  $\xi^*$  by comparing the first  $n + 2p$  equations of (2.2) with those of the stationarity conditions (3.12) of  $R(t)$ .  $\square$

Note that by examining the proof of Lemma 4.1 we get the answer to the question why we have to assume  $t \leq \tau(x^*)$ . If  $t \leq \tau(x^*)$  does not hold, then there can occur cases where we cannot define suitable multipliers by (4.1). In fact for  $0 < x_{1j}^* < t$  there holds  $j \in I_2(x^*) \setminus (I_1(x^*) \cup I_\Phi(x^*, t))$ . Now, if  $\hat{\nu}_{2j}^* < 0$ , it would be necessary to choose  $\xi_j^* > 0$ , which is not possible since  $j \notin I_\Phi(x^*, t)$ . In the same way, if  $0 < x_{2j}^* < t$ , then  $j \in I_1(x^*) \setminus (I_2(x^*) \cup I_\Phi(x^*, t))$ . Hence, if  $\hat{\nu}_{1j}^* < 0$ , it would be necessary to choose  $\xi_j^* > 0$ , which again is not possible.

The multipliers defined in Lemma 4.1 are non-unique even if the MPEC multipliers are unique. However, as we now show, we can achieve uniqueness under certain assumptions if we choose the multipliers  $\nu_{ij}^* \geq 0$  and  $\xi_j^* \geq 0$  smallest possible. This leads to the following special choice:

$$(4.5) \quad \begin{array}{lll} \nu_{1j}^* = \max(0, \hat{\nu}_{1j}^*), & \xi_j^* = \frac{\nu_{1j}^* - \hat{\nu}_{1j}^*}{2}, & \nu_{2j}^* = \hat{\nu}_{2j}^* = 0 \quad (j \in (I_1 \setminus I_2)(x^*)) \\ \nu_{2j}^* = \max(0, \hat{\nu}_{2j}^*), & \xi_j^* = \frac{\nu_{2j}^* - \hat{\nu}_{2j}^*}{2}, & \nu_{1j}^* = \hat{\nu}_{1j}^* = 0 \quad (j \in (I_2 \setminus I_1)(x^*)) \\ \nu_{1j}^* = \hat{\nu}_{1j}^*, & \nu_{2j}^* = \hat{\nu}_{2j}^*, & \xi_j^* = 0 \quad (j \in (I_1 \cap I_2)(x^*)) \end{array}$$

If we now assume that the MPEC-LICQ holds in  $x^*$ , then we can prove the uniqueness for the multipliers  $(\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$ , where  $\nu_1^*, \nu_2^*$  and  $\xi^*$  are defined by (4.1).

LEMMA 4.2. Let  $x^*$  be a strongly stationary point with multipliers  $\lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*$  that satisfies the MPEC-LICQ and let  $t \in (0, \tau(x^*)]$ . Then the multipliers  $(\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$  defined by (4.1) are unique.

*Proof.* Since the MPEC-LICQ holds at  $x^*$ , the MPEC multipliers  $\lambda^*, \mu^*, \hat{\nu}_1^*$ , and  $\hat{\nu}_2^*$  are unique. The uniqueness of the multipliers  $(\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$  then directly follows from the definition in (4.5).  $\square$

To prove the opposite direction of Lemma 4.1 we obviously need the feasibility of  $x^*$  for the MPEC. However, provided  $x^*$  is feasible for (1.1) and  $t$  is sufficiently small, the stationarity of  $x^*$  for  $R(t)$  implies that  $x^*$  is strongly stationary.

LEMMA 4.3. Suppose  $t > 0$  and  $(x^*(t), \lambda^*(t), \mu^*(t), \nu_1^*(t), \nu_2^*(t), \xi^*(t))$  satisfies (3.12) and let  $x^*(t)$  be feasible for (1.1). Then  $x^*(t)$  is strongly stationary with multipliers  $\lambda^* = \lambda^*(t), \mu^* = \mu^*(t)$  and

$$\begin{aligned}\hat{\nu}_{1j}^* &= \nu_{1j}^*(t) - \alpha_j \xi_j^*(t) \\ \hat{\nu}_{2j}^* &= \nu_{2j}^*(t) - \beta_j \xi_j^*(t).\end{aligned}$$

*Proof.* Due to the feasibility of  $x^*(t)$  for the MPEC (1.1), the choice of the multipliers, and the values of  $\alpha_j, \beta_j$  and  $\xi_j^*(t)$ , respectively, the conditions of (2.2) are a direct consequence of the conditions of (3.12). Furthermore, the nonnegativity of  $\hat{\nu}_{ij}^*$  ( $i = 1, 2$ ) for  $j \in (I_1 \cap I_2)(x^*)$  is implied by  $\xi_j^*(t) = 0$  and the nonnegativity of  $\nu_{ij}^*(t)$  ( $i = 1, 2$ ).  $\square$

Finally, having related the stationary points of both problems (1.1) and  $R(t)$ , we can now describe the relation between the strict local minima of the two problems. Therefore we compare the second order optimality conditions.

If  $x^*$  is a strongly stationary point of (1.1) or a stationary point of  $R(t)$  with  $t \in (0, \tau(x^*)]$  and if we choose the multipliers as described in Lemma 4.1 or Lemma 4.3 respectively, then we obtain  $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) = \mathcal{S}_t(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$  and  $\nabla_{xx} \mathcal{L}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*) = \nabla_{xx} \mathcal{L}_{\text{MPEC}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ , hence in this case the RNLP-SOSC for (1.1) and the SOSC for  $R(t)$  are identical conditions.

THEOREM 4.1. Let  $x^*$  be feasible for (1.1), then

1. if  $x^*$  is a strongly stationary point of the MPEC (1.1) that satisfies the RNLP-SOSC, then  $x^*$  is a stationary point of  $R(t)$  for every  $t \in (0, \tau(x^*)]$  that satisfies the SOSC. Thus,  $x^*$  is a strict local minimum of  $R(t)$ .
2. if  $x^*$  is a stationary point for  $R(t)$  that satisfies the SOSC, then  $x^*$  is also a strongly stationary point of the MPEC (1.1) that satisfies the RNLP-SOSC. Therefore, it is also a strict local minimum of (1.1).

*Proof.* In 1. the stationarity of  $x^*$  for  $R(t)$  follows from Lemma 4.1. Furthermore, in 2. the stationarity of  $x^*$  for (1.1) follows from Lemma 4.3.

Hence, it remains to show that the second order conditions at  $x^*$  imply each other. Therefore, we compare the second derivative with respect to  $x^*$  of the Lagrangian functions of the MPEC (2.1) with the one of  $R(t)$  (3.11):

$$\begin{aligned}\nabla_{xx}^2 \mathcal{L}_{\text{MPEC}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*) &= \nabla_{xx}^2 f(x^*) - \sum_{j \in I_g} \lambda_j^* \nabla_{xx}^2 g_j(x^*) - \sum_{i=1}^q \mu_i^* \nabla_{xx}^2 h_i(x^*) \\ \nabla_{xx}^2 \mathcal{L}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*) &= \nabla_{xx}^2 f(x^*) - \sum_{j \in I_g} \lambda_j^* \nabla_{xx}^2 g_j(x^*) - \sum_{i=1}^q \mu_i^* \nabla_{xx}^2 h_i(x^*) \\ &\quad + \sum_{j=1}^p \xi_j^* \begin{pmatrix} 0 & 0 \\ 0 & \nabla_{yy}^2 \Phi_j(x_1^*, x_2^*, t) \end{pmatrix}.\end{aligned}$$

with  $y := (x_1, x_2)$ . We note that they only differ in the additional term

$$M_t(x^*) = \sum_{j=1}^p \xi_j^* \begin{pmatrix} 0 & 0 \\ 0 & \nabla_{yy}^2 \Phi_j(x_1^*, x_2^*, t) \end{pmatrix}.$$

From Lemma 3.4 and the special structure of  $\nabla \Phi_j(x_1, x_2, t)$  it follows that either  $\xi_j^* = 0$  or  $\nabla_{yy}^2 \Phi_j(x_1^*, x_2^*, t) = 0$ , such that  $M_t(x^*)$  vanishes completely. Hence, the Hessians of both Lagrangian functions are equal.

Next, we prove that if  $x^*$  is a strongly stationary point with multipliers  $\lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*$ , and  $t \in (0, \tau(x^*)]$ , then the set of critical directions in  $x^*$  of  $R(t)$ , denoted by

$$\begin{aligned} \mathcal{S}_t(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*) = \{ d \in \mathbb{R}^{n+2p} \setminus \{0\} : \\ \nabla h_i(x^*)^T d = 0, \quad i \in \{1, \dots, q\} \\ \nabla g_j(x^*)^T d = 0, \quad j \in I_g(x^*), \quad \lambda_j^* > 0 \\ \nabla g_j(x^*)^T d \geq 0, \quad j \in I_g(x^*), \quad \lambda_j^* = 0 \\ \nabla \Phi_j(x_1^*, x_2^*, t)^T d = 0, \quad j \in I_\Phi(x^*, t), \quad \xi_j^* > 0 \\ \nabla \Phi_j(x_1^*, x_2^*, t)^T d \leq 0, \quad j \in I_\Phi(x^*, t), \quad \xi_j^* = 0 \\ d_{1j} = 0, \quad j \in I_1(x^*), \quad \nu_{1j}^* > 0 \\ d_{1j} \geq 0, \quad j \in I_1(x^*), \quad \nu_{1j}^* = 0 \\ d_{2j} = 0, \quad j \in I_2(x^*), \quad \nu_{2j}^* > 0 \\ d_{2j} \geq 0, \quad j \in I_2(x^*), \quad \nu_{2j}^* = 0 \}. \end{aligned}$$

is a subset of  $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ , where  $\nu_1^*, \nu_2^*$ , and  $\xi^*$  are given by (4.1). These multipliers satisfy (4.4).

Let  $x^*$  be feasible for (1.1), let  $0 < t \leq \tau(x^*)$ , and consider  $d \in \mathcal{S}_t(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$ . For  $j \in (I_1 \setminus I_2)(x^*)$  we obtain  $d_{1j} \geq 0$ . Since  $t \leq \tau(x^*)$ , by Lemma 3.4 there holds  $j \in I_\Phi(x^*, t)$  and furthermore  $\alpha_j = 2$  and  $\beta_j = 0$ . Hence,

$$0 \geq \nabla \Phi_j(x_1^*, x_2^*, t)^T d = \alpha_j d_{1j} + \beta_j d_{2j} = 2d_{1j}.$$

Therefore,  $d_{1j} = 0$  for all  $j \in (I_1 \setminus I_2)(x^*)$ . In the same way it can be shown that  $d_{2j} = 0$  for all  $j \in (I_2 \setminus I_1)(x^*)$ .

Next consider  $j \in (I_1 \cap I_2)(x^*)$ . Then there holds  $d_{1j} \geq 0$  and  $d_{2j} \geq 0$ . Furthermore, if  $j \in (I_1 \cap I_2)(x^*)$  and  $\hat{\nu}_{1j}^* > 0$ , then  $\alpha_j \geq 0$  and  $\xi_j^* \geq 0$  imply

$$0 < \hat{\nu}_{1j}^* = \nu_{1j}^* - \alpha_j \xi_j^* \leq \nu_{1j}^*,$$

and thus  $d_{1j} = 0$ . Similarly,  $j \in (I_1 \cap I_2)(x^*)$  and  $\hat{\nu}_{2j}^* > 0$  yield

$$0 < \hat{\nu}_{2j}^* = \nu_{2j}^* - \beta_j \xi_j^* \leq \nu_{2j}^*$$

and thus  $d_{2j} = 0$ . It follows that  $d \in \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ .

Conversely, we now consider a stationary point  $x^*$  for  $R(t)$  that is feasible for (1.1) and assume that  $t > 0$ . We will prove that in this case  $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$  is a subset of  $\mathcal{S}_t(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$ . Therefore, consider now a direction  $d \in \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1^*, \hat{\nu}_2^*)$ .

For  $j \in I_\Phi(x^*, t)$  we have by Lemma 3.4 that either  $x_{1j}^* = 0$  and  $x_{2j}^* \geq t$  or  $x_{2j}^* = 0$  and  $x_{1j}^* \geq t$ .

In the case  $x_{1j}^* = 0$  and  $x_{2j}^* \geq t$  we conclude  $j \in (I_1 \setminus I_2)(x^*)$  and  $\alpha_j = 2, \beta_j = 0$ . Hence,  $d_{1j} = 0$  and

$$\nabla \Phi_j(x_1^*, x_2^*, t)^T d = \alpha_j d_{1j} + \beta_j d_{2j} = 2d_{1j} = 0.$$

If  $x_{2j}^* = 0$  and  $x_{1j}^* \geq t$  then we conclude  $j \in (I_2 \setminus I_1)(x^*)$  and  $\alpha_j = 0, \beta_j = 2$ . Therefore,  $d_{2j} = 0$  and

$$\nabla \Phi_j(x_1^*, x_2^*, t)^T d = \alpha_j d_{1j} + \beta_j d_{2j} = 2d_{2j} = 0.$$

Hence, we have shown  $\nabla\Phi_j(x_1^*, x_2^*, t)^T d = 0$  for all  $j \in I_\Phi(x^*, t)$ .

Next, for  $j \in I_1(x^*)$ , there holds  $d_{1j} \geq 0$ . Now, consider the case  $j \in I_1(x^*)$  and  $\nu_{1j}^* > 0$ . If  $j \in I_1(x^*) \cap I_\Phi(x^*, t)$ , then we have already shown that  $d_{1j} = 0$ . For  $j \in I_1(x^*) \setminus I_\Phi(x^*, t)$ , on the other hand, we have  $\xi_j^* = 0$  and thus

$$\hat{\nu}_{1j}^* = \nu_{1j}^* - \alpha_j \xi_j^* = \nu_{1j}^* > 0,$$

such that in this case  $d_{1j} = 0$  holds, too.

In the same way, for  $j \in I_2(x^*)$ , we obtain  $d_{2j} \geq 0$ . Considering the case  $j \in I_2(x^*)$  and  $\nu_{2j}^* > 0$ , we see from the considerations above that  $d_{2j} = 0$  if  $j \in I_2(x^*) \cap I_\Phi(x^*, t)$ . For  $j \in I_2(x^*) \setminus I_\Phi(x^*, t)$ , on the other hand, we have  $\xi_j^* = 0$ , hence

$$\hat{\nu}_{2j}^* = \nu_{2j}^* - \beta_j \xi_j^* = \nu_{2j}^* > 0,$$

such that  $d_{2j} = 0$ . Thus  $d \in \mathcal{S}_t(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$  is shown.  $\square$

By combining Theorem 4.1 with Lemma 3.5, we can further extend this result.

**COROLLARY 4.1.** *Let  $x^*$  be a strongly stationary point of (1.1) that satisfies the RNLPSOSC. Then there exist an  $\varepsilon > 0$ , such that for every  $t \in [0, \tau(x^*)]$ ,  $x^*$  is a strict minimum in  $\mathcal{B}_\varepsilon(x^*)$  of  $R(t)$ , which even satisfies the SOSC.*

*Proof.* In view of Theorem 4.1, it follows that  $x^*$  is a strict local minimum of  $R(t)$  for every  $t \in (0, \tau(x^*)]$ . Hence, if we choose  $t = \tau(x^*)$ , then there exist an  $\varepsilon > 0$ , such that  $x^*$  is a strict minimum in  $\mathcal{B}_\varepsilon(x^*)$  of  $R(\tau(x^*))$ . Next, if we apply Lemma 3.5, then  $x^*$  is a strict minimum in  $\mathcal{B}_\varepsilon(x^*)$  of  $R(t)$  for every  $t \in [0, \tau(x^*)]$ .  $\square$

**5. Convergence.** In the previous section we compared the local solutions of both problems (1.1) and  $R(t)$ . The starting point of the next theorem is a positive sequence of parameters  $t_k \rightarrow 0$  and a convergent sequence of stationary points  $x^k$  of the problems  $R(t_k)$ . If a limit point  $\bar{x}$  satisfies the MPEC-LICQ, then we can prove that it is a C-stationary point of (1.1). Moreover, if we tie together the multipliers of the two positively linearly dependent gradients, the sequence of multiplier vectors of the stationary points  $x^k$  converges to the multiplier vector of  $\bar{x}$ . If, in addition, each  $x^k$  satisfies the SOSC for  $R(t_k)$ , then we can prove that  $\bar{x}$  is even M-stationary.

**THEOREM 5.1.** *Let  $(t_k)_{k \in \mathbb{N}}$  be a sequence with  $t_k > 0$  for all  $k \in \mathbb{N}$  that satisfies  $t_k \rightarrow 0$ , further let  $(x^k)_{k \in \mathbb{N}}$  be a sequence of stationary points of  $R(t_k)$  that satisfies  $x^k \rightarrow \bar{x}$  and suppose the MPEC-LICQ holds at  $\bar{x}$ .*

1. *Then  $\bar{x}$  is a C-stationary point of the MPEC with unique multipliers  $\bar{\lambda}, \bar{\mu}, \hat{\nu}_1^*$  and  $\hat{\nu}_2^*$ . Moreover, they satisfy*

$$\begin{aligned} \bar{\lambda}_j &= \lim_{k \rightarrow \infty} \lambda_j^k \geq 0 & j \in I_g(\bar{x}) \\ \bar{\lambda}_j &= \lim_{k \rightarrow \infty} \lambda_j^k = 0 & j \notin I_g(\bar{x}) \\ \bar{\mu}_i &= \lim_{k \rightarrow \infty} \mu_i^k & i \in \{1, \dots, q\} \\ \hat{\nu}_{1j}^* &= \lim_{k \rightarrow \infty} (\nu_{1j}^k - \xi_j^k \alpha_j^k) & j \in I_1(\bar{x}) \\ \hat{\nu}_{1j}^* &= \lim_{k \rightarrow \infty} (\nu_{1j}^k - \xi_j^k \alpha_j^k) = 0 & j \notin I_1(\bar{x}) \\ \hat{\nu}_{2j}^* &= \lim_{k \rightarrow \infty} (\nu_{2j}^k - \xi_j^k (2 - \alpha_j^k)) & j \in I_2(\bar{x}) \\ \hat{\nu}_{2j}^* &= \lim_{k \rightarrow \infty} (\nu_{2j}^k - \xi_j^k (2 - \alpha_j^k)) = 0 & j \notin I_2(\bar{x}) \end{aligned}$$



2. If in addition the SOSOC holds in each stationary point  $x^k$  of  $R(t^k)$ , then  $\bar{x}$  is  $M$ -stationary.

*Proof.* Due to the MPEC-LICQ, the uniqueness of the MPEC multipliers follows. It remains to show that such MPEC multipliers exist and that the stated choice results in these MPEC multipliers. Moreover, we have to prove that they satisfy the C-stationarity condition at  $\bar{x}$ .

Since  $g(x)$  is continuous  $I_g(x^k) \subset I_g(\bar{x})$  for sufficiently large  $k \in \mathbb{N}$  and as each  $x^k$  is a stationary point of  $R(t_k)$ , it satisfies the KKT conditions (3.12). Hence,

$$(5.1) \quad \begin{aligned} \nabla f(x^k) &= \sum_{j \in I_g(\bar{x})} \lambda_j^k \nabla g_j(x^k) + \sum_{i=1}^q \mu_i^k \nabla h_i(x^k) \\ &+ \sum_{j=1}^p (\nu_{1j}^k - \xi_j^k \alpha_j^k) e_{1j} \\ &+ \sum_{j=1}^p (\nu_{2j}^k - \xi_j^k (2 - \alpha_j^k)) e_{2j} \end{aligned}$$

As  $x^k \rightarrow \bar{x}$  and  $t_k \rightarrow 0$ , then for sufficiently large  $k \in \mathbb{N}$  it holds  $x_{1j}^k \geq t_k$  for all  $j \notin I_1(\bar{x})$ . Hence, by the feasibility of  $x^k$  and the first part of Lemma 3.7, we have  $j \in I_2(x^k) \cap I_{\Phi}(x^k, t_k)$  for all  $j \notin I_1(\bar{x})$ . Accordingly, we get  $j \in I_1(x^k) \cap I_{\Phi}(x^k, t_k)$  for all  $j \notin I_2(\bar{x})$ . Hence, by Lemma 3.7 and the values of  $\alpha_j$  for sufficiently large  $k \in \mathbb{N}$ ,

$$\begin{aligned} \nu_{1j}^k &= 0 & \text{and} & & \alpha_j^k &= 0 & \text{for all } j \notin I_1(\bar{x}), \\ \nu_{2j}^k &= 0 & \text{and} & & 2 - \alpha_j^k &= 0 & \text{for all } j \notin I_2(\bar{x}), \end{aligned}$$

such that (5.1) can be rewritten as  $A_k^T \pi_k = \nabla f(x^k)$ , where  $\pi_k = ((\lambda^k)_{I_g(\bar{x})}, \mu^k, \gamma_1^k, \gamma_2^k)$  with  $\gamma_{1j}^k := (\nu_{1j}^k - \xi_j^k \alpha_j^k)$ ,  $j \in I_1(\bar{x})$ , and  $\gamma_{2j}^k := (\nu_{2j}^k - \xi_j^k (2 - \alpha_j^k))$ ,  $j \in I_2(\bar{x})$ , and  $A_k$  denotes the matrix consisting of the row vectors

$$\begin{aligned} \nabla g_j(x^k)^T & \quad j \in I_g(\bar{x}), \\ \nabla h_i(x^k)^T & \quad i \in \{1, \dots, q\}, \\ e_{1j}^T & \quad j \in I_1(\bar{x}), \\ e_{2j}^T & \quad j \in I_2(\bar{x}). \end{aligned}$$

Due to the continuous differentiability of  $g$  and  $h$  the row vectors  $\nabla g_j(x^k)^T$  and  $\nabla h_i(x^k)^T$  converge to  $\nabla g_j(\bar{x})^T$  and  $\nabla h_i(\bar{x})^T$ , respectively. Hence, the sequence  $(A_k)$  converges to the matrix  $A$  consisting of the row vectors

$$\begin{aligned} \nabla g_j(\bar{x})^T & \quad j \in I_g(\bar{x}) \\ \nabla h_i(\bar{x})^T & \quad i \in \{1, \dots, q\} \\ e_{1j}^T & \quad j \in I_1(\bar{x}) \\ e_{2j}^T & \quad j \in I_2(\bar{x}). \end{aligned}$$

Since the MPEC-LICQ is assumed to hold at  $\bar{x}$ , these vectors are linearly independent, such that  $A$  has full row rank. Hence, there exists a unique solution vector  $\pi$  solving  $A^T \pi = \nabla f(\bar{x})$ . Moreover, the full row rank of  $A$  implies that  $AA^T$  is invertible, such that by the convergence of  $(A_k)$  and the perturbation lemma  $A_k A_k^T$  is invertible for sufficiently large  $k \in \mathbb{N}$ . Hence there exists a unique solution vector  $\pi_k = (A_k A_k^T)^{-1} (A_k \nabla f(x^k))$ . Furthermore, since  $\nabla f(x^k)$  converges to  $\nabla f(\bar{x})$

$$\pi_k = (A_k A_k^T)^{-1} (A_k \nabla f(x^k)) \longrightarrow (AA^T)^{-1} (A \nabla f(\bar{x})) = \pi.$$

Therefore,  $(\lambda^k, \mu^k, \nu_{1j}^k - \xi_j^k \alpha_j^k, \nu_{2j}^k - \xi_j^k (2 - \alpha_j^k))$  converges to the unique multiplier vector  $(\bar{\lambda}, \bar{\mu}, \hat{\nu}_1^*, \hat{\nu}_2^*)$  of  $\bar{x}$  satisfying (2.2).

Since the feasibility of  $\bar{x}$  follows by Lemma 3.3, to prove that  $\bar{x}$  is C-stationary it remains to show that  $\hat{\nu}_{1j}^* \hat{\nu}_{2j}^* \geq 0$  holds for all  $j \in (I_1 \cap I_2)(\bar{x})$ . Suppose without loss of generality that there exists an index  $j \in (I_1 \cap I_2)(\bar{x})$  with  $\hat{\nu}_{1j}^* < 0$  and  $\hat{\nu}_{2j}^* > 0$ . It follows from the convergence  $(\nu_{1j}^k - \xi_j^k \alpha_j^k) \rightarrow \hat{\nu}_{1j}^*$  and  $(\nu_{2j}^k - \xi_j^k (2 - \alpha_j^k)) \rightarrow \hat{\nu}_{2j}^*$  that

$$(5.2) \quad \nu_{1j}^k - \xi_j^k \alpha_j^k < 0$$

and

$$(5.3) \quad \nu_{2j}^k - \xi_j^k (2 - \alpha_j^k) > 0$$

for sufficiently large  $k \in \mathbb{N}$ . As  $\xi_j^k (2 - \alpha_j^k) \geq 0$  for all  $k \in \mathbb{N}$  condition (5.3) implies that  $\nu_{2j}^k > 0$  for sufficiently large  $k \in \mathbb{N}$ . Therefore,  $j \in I_2(x^k)$  must hold. Hence, either  $j \in I_2(x^k) \setminus I_\Phi(x^k, t_k)$  or  $j \in I_2(x^k) \cap I_\Phi(x^k, t_k)$ . If  $j \in I_2(x^k) \setminus I_\Phi(x^k, t_k)$ , then  $\xi_j^k = 0$ . This however implies that  $\nu_{1j}^k - \xi_j^k \alpha_j^k = \nu_{1j}^k \geq 0$  which contradicts (5.2). If on the other hand  $j \in I_2(x^k) \cap I_\Phi(x^k, t_k)$ , then by the second part of Lemma 3.7  $\nu_{1j}^k = 0$  and  $\alpha_j^k = 0$  for sufficiently large  $k \in \mathbb{N}$ . Therefore,  $\nu_{1j}^k - \xi_j^k \alpha_j^k = 0$  which again contradicts (5.2).

To prove the second part of the theorem, assume that the SOSOC holds in each  $x^k$  and that  $\bar{x}$  is not M-stationary. Then, there exists at least one index  $j_0 \in (I_1 \cap I_2)(\bar{x})$  with  $\hat{\nu}_{1j_0}^* < 0$  and  $\hat{\nu}_{2j_0}^* < 0$ . By the convergence of the multipliers, we thus have

$$(5.4) \quad \nu_{1j_0}^k - \xi_{j_0}^k \alpha_{j_0}^k < 0 \quad \text{and} \quad \nu_{2j_0}^k - \xi_{j_0}^k (2 - \alpha_{j_0}^k) < 0,$$

for sufficiently large  $k \in \mathbb{N}$ . However, since  $\nu_{1j_0}^k \geq 0$  and  $\nu_{2j_0}^k \geq 0$ , by (5.4) it follows that

$$\xi_{j_0}^k > 0 \quad \text{and} \quad 0 < \alpha_{j_0}^k < 2$$

for every  $k \in \mathbb{N}$  that is large enough. Since Lemma 3.7 implies  $I_\Phi(x^k, t_k) \cap (I_1 \cap I_2)(x^k) = \emptyset$ , and considering the values of  $\alpha_j$  for  $j \in I_\Phi(x^k, t_k) \cap (I_1 \setminus I_2)(x^k)$  or  $j \in I_\Phi(x^k, t_k) \cap (I_2 \setminus I_1)(x^k)$ , respectively, for  $k$  large enough it follows that  $j_0 \in I_\Phi(x^k, t_k) \setminus (I_1 \cup I_2)(x^k)$ . Hence,  $\nu_{1j_0}^k = 0$  and  $\nu_{2j_0}^k = 0$  for all  $k \in \mathbb{N}$  being sufficiently large. Therefore,

$$(5.5) \quad 0 > \hat{\nu}_{1j_0}^* = - \lim_{k \rightarrow \infty} \xi_{j_0}^k \alpha_{j_0}^k \quad \text{and} \quad 0 > \hat{\nu}_{2j_0}^* = - \lim_{k \rightarrow \infty} \xi_{j_0}^k (2 - \alpha_{j_0}^k).$$

As  $0 \leq \alpha_{j_0}^k \leq 2$ , there cannot exist a subsequence  $(\xi_{j_0}^k)_{k \in \mathcal{K}}$  of  $(\xi_{j_0}^k)$  that converges to zero. Hence, there exist  $\varepsilon > 0$  such that

$$(5.6) \quad \xi_{j_0}^k \geq \varepsilon$$

for all  $k$  sufficiently large. Next we prove that there exists  $\hat{\varepsilon} > 0$  with

$$(5.7) \quad \hat{\varepsilon} < \alpha_{j_0}^k \leq 2 - \hat{\varepsilon}$$

for all  $k$  sufficiently large. To this end, assume that there exist a subsequence  $(\alpha_{j_0}^k)_{k \in \mathcal{K}}$  that converges to zero. Then, by (5.5), we obtain  $(\xi_{j_0}^k)_{k \in \mathcal{K}} \rightarrow -\hat{\nu}_{2j_0}^*/2$  and thus  $(\xi_{j_0}^k \alpha_{j_0}^k)_{k \in \mathcal{K}} \rightarrow 0 \neq -\hat{\nu}_{1j_0}^* > 0$ , which is a contradiction. In the same way, if we assume that there exist a subsequence  $(\alpha_{j_0}^k)_{k \in \mathcal{K}}$  that converges to 2, we obtain  $(\xi_{j_0}^k)_{k \in \mathcal{K}} \rightarrow -\hat{\nu}_{1j_0}^*/2$  and thus  $(\xi_{j_0}^k (2 - \alpha_{j_0}^k))_{k \in \mathcal{K}} \rightarrow 0 \neq -\hat{\nu}_{2j_0}^* > 0$ , which again is a contradiction. By Lemma 3.6 it therefore follows in addition that there exists  $\tilde{\varepsilon} > 0$  with

$$(5.8) \quad -1 + \tilde{\varepsilon} \leq \frac{x_{1j_0}^k - x_{2j_0}^k}{t_k} \leq 1 - \tilde{\varepsilon}$$

for all  $k$  sufficiently large. Also, from (5.5) and (5.7) we can conclude that  $(\xi_{j_0}^k)$  is bounded.

Because the MPEC-LICQ holds in  $\bar{x}$ , for sufficiently large  $k \in \mathbb{N}$  we can construct a sequence  $(d^k) \subset \mathbb{R}^{n+2p}$  such that

$$\begin{aligned} \nabla h_i(x^k)^T d^k &= 0, & i \in \{1, \dots, q\} \\ \nabla g_j(x^k)^T d^k &= 0, & j \in I_g(\bar{x}) \\ d_{1j}^k &= 0, & j \in I_1(\bar{x}) \setminus \{j_0\} \\ d_{2j}^k &= 0, & j \in I_2(\bar{x}) \setminus \{j_0\} \\ d_{1j_0}^k &= 1, \\ d_{2j_0}^k &= -\frac{\alpha_{j_0}^k}{2 - \alpha_{j_0}^k}. \end{aligned}$$

Due to (5.7), these  $d^k$  are well defined and bounded. Moreover, these directions are contained in the corresponding sets of critical directions  $\mathcal{S}_t(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ , as

$$\nabla \Phi_{j_0}(x_1^k, x_2^k, t_k)^T d^k = \alpha_{j_0}^k d_{1j_0} + (2 - \alpha_{j_0}^k) d_{2j_0} = 0.$$

The twice continuous differentiability of  $f, g$  and  $h$  and the convergence of  $\lambda^k$  and  $\mu^k$  imply that the first three parts (5.9), (5.10) and (5.11) of the right-hand side of

$$\begin{aligned} d^{kT} \nabla_{xx}^2 \mathcal{L}_{R(t_k)}(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k) d^k &= \\ (5.9) \quad &= d^{kT} \nabla_{xx}^2 f(x^k) d^k \\ (5.10) \quad &- \sum_{j \in I_g(\bar{x})} \lambda_j^k d^{kT} \nabla_{xx}^2 g_j(x^k) d^k \\ (5.11) \quad &- \sum_{i=1}^q \mu_i^k d^{kT} \nabla_{xx}^2 h_i(x^k) d^k \\ (5.12) \quad &+ \sum_{j=1}^p \xi_j^k d^{kT} \begin{pmatrix} 0 & 0 \\ 0 & \nabla_{yy}^2 \Phi_j(x_1^k, x_2^k, t_k) \end{pmatrix} d^k, \end{aligned}$$

where  $y := (x_1, x_2)$ , are bounded for  $k \rightarrow \infty$ . Furthermore, for (5.12), we have

$$\begin{aligned} \sum_{j=1}^p \xi_j^k d^{kT} \begin{pmatrix} 0 & 0 \\ 0 & \nabla_{yy}^2 \Phi_j(x_1^k, x_2^k, t_k) \end{pmatrix} d^k &= \sum_{j \in I_{\Phi}^k \setminus (I_1^k \cup I_2^k)} \xi_j^k c_j(x^k, t_k) (d_{1j}^k - d_{2j}^k)^2 \\ &= \xi_{j_0}^k c_{j_0}(x^k, t_k) (d_{1j_0}^k - d_{2j_0}^k)^2 \\ &= \xi_{j_0}^k c_{j_0}(x^k, t_k) \left(1 + \frac{\alpha_{j_0}^k}{2 - \alpha_{j_0}^k}\right)^2 \\ &< \xi_{j_0}^k c_{j_0}(x^k, t_k), \end{aligned}$$

where

$$c_j(x, t) = \begin{cases} 0 & |x_{1j} - x_{2j}| \geq t \\ -\frac{1}{t} \theta''\left(\frac{x_{1j} - x_{2j}}{t}\right) & |x_{1j} - x_{2j}| < t \end{cases}$$

and  $I_{\Phi}^k = I_{\Phi}(x^k, t_k)$  and  $I_1^k \cup I_2^k = (I_1 \cup I_2)(x^k)$ . Since we have shown (5.8) for large  $k$  and since  $\theta''$  is continuous and strictly positive on  $(-1, 1)$ , it follows that we can find a  $\delta > 0$  such that

$$\theta''\left(\frac{x_{1j_0}^k - x_{2j_0}^k}{t_k}\right) > \delta$$

for all  $k \in \mathbb{N}$  sufficiently large. Hence,

$$\lim_{k \rightarrow \infty} c_{j_0}(x^k, t_k) = - \lim_{k \rightarrow \infty} \frac{1}{t_k} \theta''\left(\frac{x_{1j_0}^k - x_{2j_0}^k}{t_k}\right) < - \lim_{k \rightarrow \infty} \frac{\delta}{t_k} = -\infty.$$

However, as (5.6) holds for  $k$  sufficiently large, this implies

$$\sum_{j=1}^p \xi_j^k d^{kT} \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2 \Phi_j(x_1^k, x_2^k, t_k) \end{pmatrix} d^k < \xi_{j_0}^k c_{j_0}(x^k, t_k) \rightarrow -\infty$$

for the last term (5.12) and hence

$$d^{kT} \nabla_{xx}^2 \mathcal{L}_{R(t_k)}(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k) d^k \rightarrow -\infty,$$

for  $k \rightarrow \infty$  which contradicts the assumption that the SOSOC holds in each  $x^k$ .  $\square$

Since every B-stationary point that satisfies the MPEC-LICQ is strongly stationary, we are further interested in what can be proved if we weaken the assumption of the MPEC-LICQ in Theorem 5.1. As we will see, assuming that a weaker condition, namely the MPEC-CRCQ, does hold we can still prove convergence to a C-stationary point.

First we introduce a further definition that we will use to abbreviate the notation of the proof.

DEFINITION 5.1. Let  $\lambda \in \mathbb{R}^m$  then the support of  $\lambda$  is defined as

$$\text{supp}(\lambda) := \{j \in \{1, \dots, m\} : \lambda_j \neq 0\}.$$

THEOREM 5.2. Let  $(t_k)_{k \in \mathbb{N}}$  be a sequence with  $t_k > 0$  for all  $k \in \mathbb{N}$  that satisfies  $t_k \rightarrow 0$ . Furthermore, let  $(x^k)_{k \in \mathbb{N}}$  be a sequence of stationary points of  $R(t_k)$  that satisfies  $x^k \rightarrow \bar{x}$  and suppose that the MPEC-CRCQ holds at  $\bar{x}$ . Then  $\bar{x}$  is a C-stationary point of the MPEC

To prove Theorem 5.2, we first show that under MPEC-CRCQ there exists a sequence of multipliers associated with the convergent sequence of stationary points  $x^k$  of  $R(t_k)$  which contains a bounded subsequence. This bounded subsequence then yields a convergent subsequence of multipliers whose limit satisfies the C-stationarity conditions for the limit point  $\bar{x}$ .

*Proof.* Let  $x^k$  be a stationary point of  $R(t_k)$ . Since  $g(x)$  is continuous we have for sufficiently large  $k$  that  $I_g(x^k) \subset I_g(\bar{x})$ ,  $I_1(x^k) \subset I_1(\bar{x})$ , and  $I_2(x^k) \subset I_2(\bar{x})$ . Hence, as  $x^k$  is stationary for  $R(t_k)$ , for  $k$  sufficiently large, there exist  $\lambda^k \geq 0$ ,  $\mu^k$ ,  $\nu_1^k \geq 0$ ,  $\nu_2^k \geq 0$ ,  $\xi^k \geq 0$  such that

$$\begin{aligned} (5.13) \quad & -\nabla f(x^k) + \sum_{j \in I_g(\bar{x})} \lambda_j^k \nabla g_j(x^k) + \sum_{j=1}^q \mu_j^k \nabla h_j(x^k) \\ & + \sum_{j \in I_1(\bar{x})} \nu_{1j}^k e_{1j} + \sum_{j \in I_2(\bar{x})} \nu_{2j}^k e_{2j} \\ & - \sum_{j \in I_\Phi(x^k, t_k)} \xi_j^k (\alpha_j^k e_{1j} + (2 - \alpha_j^k) e_{2j}) = 0. \end{aligned}$$

If we define for all  $k \in \mathbb{N}$

$$\gamma_{1j}^k := \alpha_j^k \xi_j^k \quad \text{and} \quad \gamma_{2j}^k := (2 - \alpha_j^k) \xi_j^k$$

for all  $j \in \{1, \dots, p\}$ , then  $\gamma_{1j}^k \geq 0$  and  $\gamma_{2j}^k \geq 0$ . Moreover,

$$(5.14) \quad \begin{aligned} \gamma_{1j}^k > 0 & \iff \alpha_j^k > 0 \quad \text{and} \quad \xi_j^k > 0 \implies j \in I_\Phi(x^k, t_k) \setminus I_2(x^k) \\ \gamma_{2j}^k > 0 & \iff \alpha_j^k < 2 \quad \text{and} \quad \xi_j^k > 0 \implies j \in I_\Phi(x^k, t_k) \setminus I_1(x^k) \end{aligned}$$

Hence, for large  $k$ ,

$$(5.15) \quad \begin{aligned} \text{supp}(\gamma_1^k) &\subset I_\Phi(x^k, t_k) \setminus I_2(x^k) \subset I_1(\bar{x}) \\ \text{supp}(\gamma_2^k) &\subset I_\Phi(x^k, t_k) \setminus I_1(x^k) \subset I_2(\bar{x}). \end{aligned}$$

Since, for large  $k$ ,

$$(5.16) \quad \begin{aligned} \text{supp}(\lambda^k) &\subset I_g(\bar{x}), \quad \text{supp}(\mu^k) \subset I_h(\bar{x}) \\ \text{supp}(\nu_1^k) &\subset I_1(\bar{x}), \quad \text{supp}(\nu_2^k) \subset I_2(\bar{x}), \end{aligned}$$

we can then write (5.13) as

$$(5.17) \quad \begin{aligned} -\nabla f(x^k) &+ \sum_{j \in \text{supp}(\lambda^k)} \lambda_j^k \nabla g_j(x^k) + \sum_{j \in \text{supp}(\mu^k)} \mu_j^k \nabla h_j(x^k) \\ &+ \sum_{j \in \text{supp}(\nu_1^k)} \nu_{1j}^k e_{1j} + \sum_{j \in \text{supp}(\nu_2^k)} \nu_{2j}^k e_{2j} \\ &+ \sum_{j \in \text{supp}(\gamma_1^k)} \gamma_{1j}^k (-e_{1j}) + \sum_{j \in \text{supp}(\gamma_2^k)} \gamma_{2j}^k (-e_{2j}) = 0. \end{aligned}$$

As a consequence of Lemma A.1, for every  $x^k$  we can find a multiplier vector  $(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)$ , such that the system

$$(5.18) \quad \begin{aligned} &\{\nabla g_j(x^k) : j \in \text{supp}(\bar{\lambda}^k)\} \cup \{\nabla h_j(x^k) : j \in \text{supp}(\bar{\mu}^k)\} \cup \{e_{1j} : j \in \text{supp}(\bar{\nu}_1^k)\} \\ &\cup \{e_{2j} : j \in \text{supp}(\bar{\nu}_2^k)\} \cup \{-e_{1j} : j \in \text{supp}(\bar{\gamma}_1^k)\} \cup \{-e_{2j} : j \in \text{supp}(\bar{\gamma}_2^k)\} \end{aligned}$$

is linearly independent and

$$(5.19) \quad \begin{aligned} \text{supp}(\bar{\lambda}^k) &\subset \text{supp}(\lambda^k), \quad \text{supp}(\bar{\mu}^k) \subset \text{supp}(\mu^k) \\ \text{supp}(\bar{\nu}_1^k) &\subset \text{supp}(\nu_1^k), \quad \text{supp}(\bar{\nu}_2^k) \subset \text{supp}(\nu_2^k), \\ \text{supp}(\bar{\gamma}_1^k) &\subset \text{supp}(\gamma_1^k), \quad \text{supp}(\bar{\gamma}_2^k) \subset \text{supp}(\gamma_2^k). \end{aligned}$$

The linear independence of the system implies that

$$(5.20) \quad \text{supp}(\bar{\nu}_1^k) \cap \text{supp}(\bar{\gamma}_1^k) = \emptyset \quad \text{and} \quad \text{supp}(\bar{\nu}_2^k) \cap \text{supp}(\bar{\gamma}_2^k) = \emptyset.$$

From (5.15), (5.16), and (5.19) we have that, for sufficiently large  $k$ ,

$$(5.21) \quad \text{supp}(\bar{\nu}_1^k) \cup \text{supp}(\bar{\gamma}_1^k) \subset I_1(\bar{x}) \quad \text{and} \quad \text{supp}(\bar{\nu}_2^k) \cup \text{supp}(\bar{\gamma}_2^k) \subset I_2(\bar{x}).$$

Since there are only finitely many configurations of different supports  $\text{supp}(\bar{\lambda}^k), \dots, \text{supp}(\bar{\gamma}_2^k)$ , There exists a configuration  $S_{\bar{\lambda}}, \dots, S_{\bar{\gamma}_2}$  of supports such that

$$(5.22) \quad \text{supp}(\bar{\lambda}^k) = S_{\bar{\lambda}}, \quad \dots, \quad \text{supp}(\bar{\gamma}_2^k) = S_{\bar{\gamma}_2}$$

for infinitely many  $k$ . Define

$$K_s = \{k \in \mathbb{N} : (5.22) \text{ holds for } k\}.$$

We first show that

$$(5.23) \quad \{(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)\}_{K_s} \text{ is bounded.}$$

Assume this is not true. Then there exists an unbounded subsequence  $\{(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)\}_{\hat{K}}$ ,  $\hat{K} \subset K_s$ . Without restriction, we can assume that  $(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k) \neq 0$  for all  $k \in \hat{K}$ . We now consider the normalized (i.e., bounded) sequence of multipliers

$$(\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k) = \frac{(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)}{\|(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)\|}, \quad k \in \hat{K}.$$

This sequence contains a convergent subsequence

$$(\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\gamma}_1^k, \tilde{\gamma}_2^k)_{k \in K} \longrightarrow (\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\gamma}_1, \tilde{\gamma}_2).$$

with

$$\begin{aligned} \text{supp}(\tilde{\lambda}) &\subset \text{supp}(\tilde{\lambda}^k) = \text{supp}(\bar{\lambda}^k) = S_{\bar{\lambda}} \subset \text{supp}(\lambda^k), \\ \text{supp}(\tilde{\mu}) &\subset \text{supp}(\tilde{\mu}^k) = \text{supp}(\bar{\mu}^k) = S_{\bar{\mu}} \subset \text{supp}(\mu^k), \\ \text{supp}(\tilde{\nu}_1) &\subset \text{supp}(\tilde{\nu}_1^k) = \text{supp}(\bar{\nu}_1^k) = S_{\bar{\nu}_1} \subset \text{supp}(\nu_1^k), \\ \text{supp}(\tilde{\nu}_2) &\subset \text{supp}(\tilde{\nu}_2^k) = \text{supp}(\bar{\nu}_2^k) = S_{\bar{\nu}_2} \subset \text{supp}(\nu_2^k), \\ \text{supp}(\tilde{\gamma}_1) &\subset \text{supp}(\tilde{\gamma}_1^k) = \text{supp}(\bar{\gamma}_1^k) = S_{\bar{\gamma}_1} \subset \text{supp}(\gamma_1^k), \\ \text{supp}(\tilde{\gamma}_2) &\subset \text{supp}(\tilde{\gamma}_2^k) = \text{supp}(\bar{\gamma}_2^k) = S_{\bar{\gamma}_2} \subset \text{supp}(\gamma_2^k) \end{aligned}$$

for all  $k \in K$ . Therefore, for all  $k \in K$ , we get by (5.17)

$$\begin{aligned} 0 &= \lim_{K \ni k \rightarrow \infty} \left( -\frac{\nabla f(x^k)}{\omega_k} + \sum_{j \in S_{\bar{\lambda}}} \frac{\bar{\lambda}_j^k}{\omega_k} \nabla g_j(x^k) + \sum_{j \in S_{\bar{\mu}}} \frac{\bar{\mu}_j^k}{\omega_k} \nabla h_j(x^k) \right. \\ &\quad + \sum_{j \in S_{\bar{\nu}_1}} \frac{\bar{\nu}_{1j}^k}{\omega_k} e_{1j} + \sum_{j \in S_{\bar{\nu}_2}} \frac{\bar{\nu}_{2j}^k}{\omega_k} e_{2j} \\ &\quad \left. + \sum_{j \in S_{\bar{\gamma}_1}} \frac{\bar{\gamma}_{1j}^k}{\omega_k} (-e_{1j}) + \sum_{j \in S_{\bar{\gamma}_2}} \frac{\bar{\gamma}_{2j}^k}{\omega_k} (-e_{2j}) \right) \\ &= \sum_{j \in S_{\bar{\lambda}}} \tilde{\lambda}_j \nabla g_j(\bar{x}) + \sum_{j \in S_{\bar{\mu}}} \tilde{\mu}_j \nabla h_j(\bar{x}) \\ (5.24) \quad &+ \sum_{j \in S_{\bar{\nu}_1}} \tilde{\nu}_{1j} e_{1j} + \sum_{j \in S_{\bar{\nu}_2}} \tilde{\nu}_{2j} e_{2j} \\ &+ \sum_{j \in S_{\bar{\gamma}_1}} \tilde{\gamma}_{1j} (-e_{1j}) + \sum_{j \in S_{\bar{\gamma}_2}} \tilde{\gamma}_{2j} (-e_{2j}) \end{aligned}$$

where  $\omega_k = \|(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)\|$ . As  $\|(\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\gamma}_1, \tilde{\gamma}_2)\| = 1$ , there have to exist some entries of  $(\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\gamma}_1, \tilde{\gamma}_2)$  that do not vanish. Hence, the system

$$(5.25) \quad \begin{aligned} &\{\nabla g_j(\bar{x}) : j \in S_{\bar{\lambda}}\} \cup \{\nabla h_j(\bar{x}) : j \in S_{\bar{\mu}}\} \cup \{e_{1j} : j \in S_{\bar{\nu}_1}\} \\ &\cup \{e_{2j} : j \in S_{\bar{\nu}_2}\} \cup \{-e_{1j} : j \in S_{\bar{\gamma}_1}\} \cup \{-e_{2j} : j \in S_{\bar{\gamma}_2}\} \end{aligned}$$

is linearly dependent. By (5.16), (5.19), and (5.22), these vectors correspond to constraints that are all active at  $\bar{x}$ . Now, since  $\bar{x}$  satisfies the MPEC-CRCQ, there exists a neighbourhood  $\mathcal{U}(\bar{x})$ , such that for all  $y \in \mathcal{U}(\bar{x})$  the system (5.25) evaluated at  $y$  has the same rank, i.e., is also linearly dependent. Furthermore, as  $x^k$  converges to  $\bar{x}$ , there exists a  $k_1 \in \mathbb{N}$ , such that  $x^k$  lies in  $\mathcal{U}(\bar{x})$  for all  $k \geq k_1$ . Due to (5.22), this implies that the system (5.18) is linearly dependent for all  $k \in K$ ,  $k \geq k_1$ , which contradicts the linear independence of the vectors in (5.18). Hence, (5.23) holds.

Due to boundedness, the sequence of multipliers  $\{(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)\}_{K_s}$  has a convergent subsequence

$$\{(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)\}_{K'_s} \rightarrow (\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \gamma_1^*, \gamma_2^*).$$

By continuity and (5.13), there holds

$$(5.26) \quad \begin{aligned} &-\nabla f(\bar{x}) + \sum_{j \in I_g(\bar{x})} \lambda_j^* \nabla g_j(\bar{x}) + \sum_{j=1}^q \mu_j^* \nabla h_j(\bar{x}) \\ &+ \sum_{j \in I_1(\bar{x})} \hat{\nu}_{1j}^* e_{1j} + \sum_{j \in I_2(\bar{x})} \hat{\nu}_{2j}^* e_{2j} = 0. \end{aligned}$$

with

$$(5.27) \quad \hat{\nu}_{1j}^* := \begin{cases} \nu_1^* & j \in \text{supp}(\nu_1^*) \\ -\gamma_1^* & j \in \text{supp}(\gamma_1^*) \\ 0 & \text{else} \end{cases} \quad \hat{\nu}_{2j}^* := \begin{cases} \nu_2^* & j \in \text{supp}(\nu_2^*) \\ -\gamma_2^* & j \in \text{supp}(\gamma_2^*) \\ 0 & \text{else} \end{cases} .$$

The definition (5.27) of the multipliers  $\hat{\nu}_1^*$  and  $\hat{\nu}_2^*$  is well-defined. In fact, due to convergence of the subsequence, we have that, for all  $k \in K'_s$ ,

$$\begin{aligned} \text{supp}(\nu_1^*) \subset S_{\bar{\nu}_1} &= \text{supp}(\bar{\nu}_1^k), & \text{supp}(\gamma_1^*) \subset S_{\bar{\gamma}_1} &= \text{supp}(\bar{\gamma}_1^k), \\ \text{supp}(\nu_2^*) \subset S_{\bar{\nu}_2} &= \text{supp}(\bar{\nu}_2^k), & \text{supp}(\gamma_2^*) \subset S_{\bar{\gamma}_2} &= \text{supp}(\bar{\gamma}_2^k) \end{aligned}$$

Hence, by (5.20) we conclude

$$\text{supp}(\nu_1^*) \cap \text{supp}(\gamma_1^*) = \emptyset \quad \text{and} \quad \text{supp}(\nu_2^*) \cap \text{supp}(\gamma_2^*) = \emptyset.$$

Thus,  $\bar{x}$  is weakly stationary. Suppose it is not C-stationary, then there exists at least one index  $j_0 \in I_1(\bar{x}) \cap I_2(\bar{x})$  with either  $\hat{\nu}_{1j_0}^* < 0$  and  $\hat{\nu}_{2j_0}^* > 0$  or  $\hat{\nu}_{1j_0}^* > 0$  and  $\hat{\nu}_{2j_0}^* < 0$ . Let (wlog)  $\hat{\nu}_{1j_0}^* < 0$  and  $\hat{\nu}_{2j_0}^* > 0$ . Then, by the convergence of the subsequence,

$$j_0 \in \text{supp}(\bar{\gamma}_1^k) \quad \text{and} \quad j_0 \in \text{supp}(\bar{\nu}_2^k)$$

for  $k \in K$  sufficiently large. Using (5.19), this implies

$$j_0 \in \text{supp}(\gamma_1^k) \quad \text{and} \quad j_0 \in \text{supp}(\nu_2^k)$$

for all  $k \in K$  being large enough. From this and (5.15) it follows that

$$j_0 \in I_{\Phi}(x^k, t_k) \setminus I_2(x^k) \quad \text{and} \quad j_0 \in I_2(x^k)$$

for all  $k \in K$  sufficiently large. Due to this contradiction,  $\bar{x}$  has to be C-stationary.  $\square$

The following example shows that we cannot improve the conclusion of Theorem 5.2 concerning the stationarity conditions that are satisfied in the limit point  $\bar{x}$  while maintaining the assumptions.

EXAMPLE 5.1.

$$\begin{aligned} \min \quad & \frac{1}{2}((x_1 - 1)^2 + (x_2 - 1)^2) \\ \text{s. t.} \quad & 0 \leq x_1 \perp x_2 \geq 0, \quad : \nu_1, \nu_2, \xi \end{aligned}$$

Consider the sequence  $(x_1^k, x_2^k) = (\vartheta_k, \vartheta_k)$  with  $\vartheta_k := 1/2t\theta(0) \rightarrow 0$ . Then  $\Phi(x_1^k, x_2^k, t_k) = 0$  for all  $k$  and  $(x_1^k, x_2^k)$  is a sequence of KKT-points for  $R(t_k)$  for a sequence  $(t_k)$  with  $t_k \searrow 0$ , as

$$\begin{pmatrix} \vartheta_k - 1 \\ \vartheta_k - 1 \end{pmatrix} + \xi^k \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

with  $\xi^k = 1 - \vartheta_k$ ,  $\nu_1^k = 0$  and  $\nu_2^k = 0$ .

The vector  $(x_1^k, x_2^k)$  converges to  $(\bar{x}_1, \bar{x}_2) = (0, 0)$  which clearly satisfies the MPEC-CRCQ. The MPEC multipliers  $(\hat{\nu}_1^*, \hat{\nu}_2^*)$  that we can construct according to Theorem 5.2 from the limits of  $\nu_1^k$ ,  $\nu_2^k$ ,  $\xi^k$  and  $\alpha^k$  are  $(\hat{\nu}_1^*, \hat{\nu}_2^*) = (-1, -1)$ . Since we cannot construct multipliers  $(\hat{\nu}_1^*, \hat{\nu}_2^*)$  that satisfy the M-stationarity condition  $\hat{\nu}_1^* \hat{\nu}_2^* = 0$  or  $\hat{\nu}_1^* > 0$  and  $\hat{\nu}_2^* > 0$ , it follows that  $(0, 0)$  is C-stationary but not M-stationary.

---

**Algorithm 1:** Basic Outer Algorithm

---

- 1 Choose an initial vector  $(x^0, \lambda^0, \mu^0, \nu_1^0, \nu_2^0, \xi^0)$ , an initial parameter  $t_0 > 0$ , an update parameter  $\sigma \in (0, 1)$ , tolerance parameters  $\varepsilon_C > 0$  and  $\varepsilon_{SQP} > 0$  and  $\delta_{\min}$
  - repeat**
  - 2   Solve  $R(t_k)$  up to the given accuracy determined by  $\varepsilon_{SQP}$ .
  - 3   Update the relaxation parameter  $t_k$  :  
     $t_{k+1} \leftarrow \max(\sigma t_k, \delta_{\min})$
  - 4   Set  $k \leftarrow k + 1$
  - until**  $\text{compl}(x^*(t_k)) \leq \varepsilon_C$
- 

**6. Numerical Results.** The theoretical results of the previous sections motivate the approach to solve an MPEC of the form (1.1) by solving a sequence of nonlinear programs  $R(t_k)$  for a positive, decreasing sequence  $(t_k)$  until the complementarity condition is sufficiently satisfied.

As a basic outer algorithm to solve (1.1) we therefore implemented Algorithm 1 as an AMPL script starting an NLP solver as a black box.

To solve the NLPs we chose the SQP solver `filterSQP`. As a measure for the feasibility of an iterate  $x$  with respect to the complementarity constraints we use

$$\text{compl}(x) = \sqrt{\sum_{j=1}^p \min(x_{1j}, x_{2j})^2}.$$

Moreover, we consider a problem to be solved correctly if the constraint violation and the KKT-residual for the last solved problem  $R(t_k)$  are sufficiently small, i.e., smaller than  $\varepsilon_{SQP}$ . Moreover, the complementarity condition has to be sufficiently satisfied for the solution  $x^*(t_k)$  of the last solved problem  $R(t_k)$ , i.e.  $\text{compl}(x^*(t_k)) \leq \varepsilon_C$ . For the numerical tests we set throughout  $\varepsilon_{SQP} = 10^{-8}$  and  $\varepsilon_C = 10^{-8}$ .

Concerning the function  $\theta(z)$  that we use to describe the relaxation, we tested the basic algorithm for two different functions: the first one is a suitably adapted sin-function  $s(z)$  and the second one is a polynomial  $p(z)$ :

$$s(z) = \frac{2}{\pi} \sin\left(z \frac{\pi}{2} + \frac{3\pi}{2}\right) + 1 \quad \text{and} \quad p(z) = \frac{1}{8}(-z^4 + 6z^2 + 3).$$

Both functions satisfy the Assumption 3.1. Our tests for a fixed combination of parameters  $(t_0, \sigma)$  reveal that the performance of Algorithm 1 with  $\theta(z) = s(z)$  performs better than the polynomial, although the difference is not significant. We therefore continued using  $s(z)$ .

For selecting suitable parameter settings, we compared the performance of the basic outer algorithm for a selection of parameter combinations. For this, we used a subset of test problems taken from MacMPEC, which is a collection of MPECs maintained by Sven Leyffer [Ley00].

The observations we made concerning the performance for the different parameter combinations  $(t_0, \sigma) \in \{0.00, 1.00, 10.0\} \times \{0.001, 0.010, 0.100\}$  led us to the conclusion that the parameter choice  $(t_0, \sigma) = (10.0, 0.10)$  performs best concerning robustness of the algorithm and in parts concerning small iteration numbers.

Looking at the failures occurring in these preliminary numerical tests, we observed two typical situations where the basic outer algorithm fails to find an appropriate solution of (1.1). We therefore added some simple modifications to the basic outer algorithm that help to handle these two situations:



The first situation in which Algorithm 1 might fail to solve the MPEC can be described as follows. Since we use an SQP method to solve  $R(t_k)$ , we solve a sequence of QPs, where we linearize the constraints of  $R(t_k)$ . However, the linearization

$$(6.1) \quad \Phi_j(x_1^k, x_2^k, t^k) + \nabla \Phi_j(x_1^k, x_2^k, t^k)^T d \leq 0$$

of the constraint  $\Phi_j(x_1^k + d_1, x_2^k + d_2, t^k) \leq 0$  in some cases might cause a sequence of smaller and smaller steps, which then causes the SQP algorithm to stop although we are not close enough to a solution yet:

Consider a positive parameter  $t_k > 0$  and a pair  $(x_{1j}^k, x_{2j}^k)$  of the current iterate  $x^k$  with  $x_{1j}^k < t_k$  and  $x_{2j}^k = 0$  (the case that  $x_{2j}^k < t_k$  and  $x_{1j}^k = 0$  is similar). Hence it satisfies strict complementarity and is feasible with respect to the constraint  $\Phi_j(x_1, x_2, t_k) \leq 0$ . If we linearize  $\Phi_j(x_1^k + d_1, x_2^k + d_2, t_k) \leq 0$ , then we obtain (6.1), which is equivalent to  $\alpha_j^k d_{1j} + (2 - \alpha_j^k) d_{2j} \leq -\Phi_j^k$ , where  $\alpha_j^k$  denotes the corresponding partial derivative of  $\Phi_j^k = \Phi_j(x_1^k, x_2^k, t_k)$ . Assume that a good step towards the solution would have the form  $t e_{1j}$  with  $t \geq t_k$ . Using Taylor expansion to determine the restriction of the length of a step  $d$  along the direction  $e_{1j}$ , we get

$$d_{1j} \leq -\frac{\Phi_j^k}{\alpha_j^k} \approx \frac{1}{3}(t_k - x_{1j}^k).$$

Hence, in this situation the QP solver produces a sequence  $(d^k)$  with  $\lim_{k \rightarrow \infty} d_{1j}^k = 0$  and  $x_{1j}^k < t_k$  for all  $k \in \mathbb{N}$ , i.e., the sequence  $(x_{1j}^k)$  will never “pass the barrier”  $t_k$ . In order to intervene in time when this situation occurs, we decrease  $t_k$  in between the SQP iterations, thus, if for the current iterate  $x^k$  a pair  $(x_{1j}^k, x_{2j}^k)$  approaches either  $(t_k, 0)$  or  $(0, t_k)$ , then we reduce  $t_k$  directly after the solution of the QP, such that  $x_{ij}^k > t_k$  ( $i = 1$  or  $2$ , respectively). Since we cannot use the solver as a black box any longer, we now incorporate the outer algorithm directly in the SQP solver `filterSQP`. Furthermore, as we do not want to interfere with the other constraints  $\Phi_i(x_1, x_2, t_k) \leq 0$  with  $i \in \{1, \dots, p\} \setminus \{j\}$ , we use a parameter vector  $t^k = (t_1^k, \dots, t_p^k)$  instead of a scalar parameter, such that we have  $\Phi_j(x_1, x_2, t_j^k) \leq 0$  for all  $j \in \{1, \dots, p\}$  and we can update  $t_j^k$  independently.

The second situation concerns the case that, since we use the solution vector  $x^*(t_k)$  as an initial point to solve  $R(t_{k+1})$ , which has a smaller feasible region compared to  $R(t_k)$ , our initial point  $x^*(t_k)$  might not be feasible for  $R(t_{k+1})$ . Hence, the SQP solver then first has to solve a feasibility problem. If this cannot be solved successfully, then the outer algorithm gets stuck in the infeasible point  $x^*(t_k)$ . We made the observation that if we decrease  $t_k$  more gently, then in most cases the solver is able to find a feasible point for  $R(t_{k+1})$ . In order to maintain the good performance concerning the iteration numbers for the problems that were solved without this feasibility problem for a more stringent parameter update, we first try a more aggressive parameter update and only in the case that  $R(t_{k+1})$  could not be solved successfully, we “re-update”  $t_k$  and use a more gentle decrease for  $t_{k+1}$  (thus we enlarge  $t_{k+1}$ ) and solve  $R(t_{k+1})$  again. If  $R(t_{k+1})$  was then successfully solved for the gentle decrease, then for the next parameter  $t_k$  we try again the more aggressive parameter update.

For the comparison of the numerical results for this modified algorithm with the results for the exact [FLRS06] and the relaxed [Sch01] bilinear solution approach for MPECs we used a set of 153 test problems. Table 6.1 displays a summary of some main features of the problems contained in our set of test problems. The abbreviations of the first column have the following meanings:

- LL : Linear objective function and linear constraint functions
- QL : Quadratic objective function and linear constraint functions
- L/QQ : Linear or quadratic objective function and quadratic constraint functions
- L/QU : Linear or quadratic objective function and no general constraints
- O : Other types of Programs.

type	nr. of variables			total nr. of constr.			nr. of compl. constr.		
	0-19	20-499	$\geq 500$	0-19	10-499	$\geq 500$	0-9	10-99	$\geq 100$
LL	15	23	29	15	23	29	24	14	29
QL	21	21	1	21	21	1	21	11	11
L/Q Q	3	1	–	3	1	–	4	–	–
L/Q U	8	1	–	8	1	–	8	–	1
O	12	7	11	12	7	11	12	7	11
sum	59	53	41	59	53	41	69	32	52

TABLE 6.1  
Short description of the test set

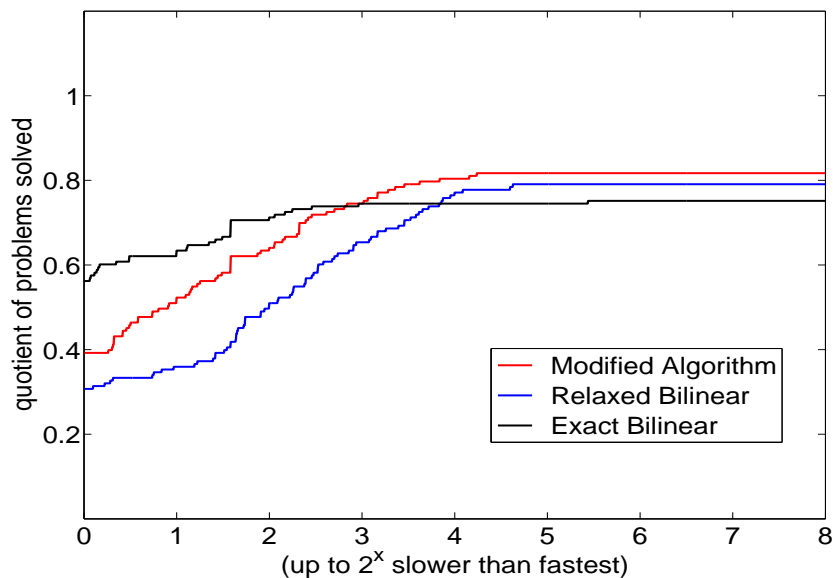


FIG. 6.1. Comparison of the new relaxation method to the exact and the relaxed bilinear approach

This classification has been taken from [Ley00], where one can find it in more detail. However, we sometimes slightly modified the problems dimension by adding slack variables to obtain an MPEC of the form (1.1).

We summarize the numerical results with regard to the iteration numbers that we obtained for all three approaches in Figure 6.1, using the performance profile that was introduced by Dolan and Moré in [DM02].

Figure 6.1 shows that the new relaxation method represents a compromise between small iteration numbers and robustness of the method: using the exact bilinear solution approach most problems are solved within comparably few iterations. For 56% of the test problems the exact bilinear approach produces the smallest iteration number. Concerning the new relaxation method 39% of the test problems are solved within the smallest number of iterations and the relaxed bilinear approach solves 31% of the test problems as best. However, the advantage of the exact bilinear approach concerning small iteration numbers might be due to the fact that most of the test problems have strongly stationary solutions (confer for example [DFNS05]).

The new relaxation fails to find an appropriate solution, i.e., a point where the KKT- and the feasibility residual do not exceed  $\varepsilon_{SQP}$  and the condition on the complementarity measure is fulfilled, for 18% of all 153 test problems. The relaxed bilinear approach fails for 20% of all test problems and the exact bilinear approach fails to find an appropriate solution for 25% of the test problems. Hence, concerning the robustness, the new relaxation method performs best. Moreover, according to Leyffer [Ley00] at least 4 failures are due

		Exact Bilinear	Relaxed Bilinear	New Relaxation
ex9.2.2	$f(x^*)$	1.00 E+02	1.00 E+02	1.00 E+02
	feas.-residual	3.84 E-17	7.11 E-15	0.00 E+00
	$\text{compl}(x^*)$	2.41 E-09	1.15 E-09	0.00 E+00
	KKT-residual	1.85 E+01	1.85 E+00	3.97 E -17
	$\xi_{\max}^*$	1.78 E+09	9.39 E+08	9.95 E+00
	inner(outer) Iterations	29(1)	91(20)	49(14)
ralph1	$f(x^*)$	-9.60 E-09	-1.00 E-06	-1.82 E-09
	feas.-residual	9.22 E-17	1.00E-12	0.00 E+00
	$\text{compl}(x^*)$	9.60 E-09	1.00 E-06	1.82 E-09
	KKT-residual	2.36 E-01	0.00 E+00	1.57 E -16
	$\xi_{\max}^*$	3.47 E+07	5.00 E+05	0.50 E+00
	inner(outer) Iterations	33(1)	86(20)	10(10)
scholtes4	$f(x^*)$	-3.84 E-09	-2.00E-06	-3.63 E-09
	feas.-residual	3.70 E -18	1.00E-12	0.00 E+00
	$\text{compl}(x^*)$	1.92E-09	1.00 E-06	1.82 E-09
	KKT-residual	4.71 E-01	0.00 E+00	4.34 E -18
	$\xi_{\max}^*$	3.47 E+08	1.00 E+06	1.00 E+00
	inner(outer) Iterations	32(1)	85(20)	11(10)

TABLE 6.2

Results for ex9.2.2, ralph1 and scholtes4

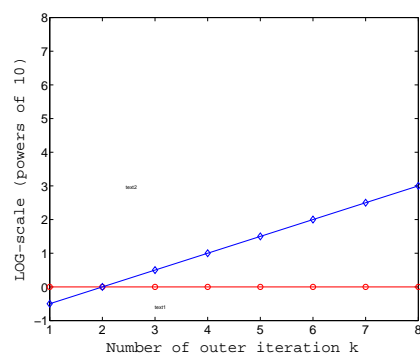
to infeasible problems, such that they are in fact no proper failures.

Finally, let us consider the numerical results of three examples of the test problem set in more detail. The test problems we consider are `ex9.2.2`, `ralph1` and `scholtes4` which are known to have B-stationary points that are not strongly stationary [Ley06].

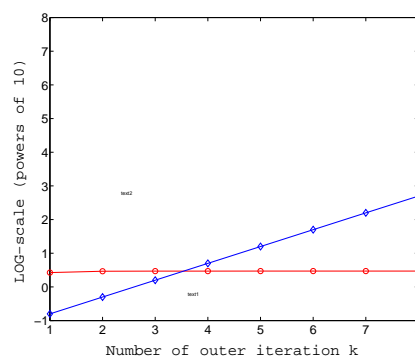
By the reported results in [Ley06, FLRS06] and some simple theoretical considerations, we expect that for the relaxed and the exact bilinear approach the generated multiplier sequences  $(\xi^k)$  of the complementarity conditions are unbounded. Furthermore, we expect this unboundedness of the multipliers to cause numerical difficulties. In Table 6.2 we summarize some characteristic numbers for these three examples. These results support our expectations. Although both methods, the relaxed and the exact bilinear approach, produce the same final objective function value as Leyffer reports in [Ley00], the same we obtain using the new relaxation method, they fail to find an appropriate solution for all three problems, as the conditions on the KKT-residual and in parts the complementarity constraints are not sufficiently satisfied. This failure is presumably due to the numerical values of  $\xi^*$ , i.e., the multiplier corresponding to the complementarity conditions, which become very large for both approaches. The numerical values for  $\xi^*$ ,  $\text{compl}(x^*)$  and the KKT-residual we obtain for the new relaxation method, by contrast, are of suitable sizes. This indicates that the new relaxation approach is more appropriate for solving this type of MPECs than the bilinear approaches. Figure 6.2, where we display the development of the significant complementarity multiplier  $\xi^k$  for the relaxed bilinear approach and the new relaxation method, emphasizes this observation: for the relaxed bilinear approach the values of  $\xi^k$  diverge, whereas they stay constant for the new relaxation method.

## REFERENCES

- [AAP04] S. Akella, M. Anitescu, and J. Peng. Mathematical Programs with Complementarity Constraints for Multiple Robot Coordination with Contact and Friction. *Proceedings of the IEEE International Conference in Robotics and Automation (ICRA 2004)*, New Orleans, LA, pages 5224–5232, 2004.
- [Ani05] M. Anitescu. Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints. *SIAM J. Optim.*, 16(1):120–145, 2005.



(a) Results for scholtes4



(b) Results for ex9.2.2

FIG. 6.2. Comparison of complementarity multiplier  $\xi^k$  for  $t^k = 0.1^{k-2}$

- [Bar98] J.F. Bard. *Practical Bilevel Optimization: Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, 1998.
- [BR05] L.T. Biegler and A.U. Raghunathan. An interior point method for mathematical programs with complementarity constraints (mpccs). *SIAM J. Optim.*, 15(3):720–750, 2005.
- [Dem02] S. Dempe. *Foundations of bilevel programming*. Nonconvex Optimization and Its Applications 61, Kluwer Academic Publishers, Dordrecht, 2002.
- [DFNS05] V. DeMiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes. A two-sided relaxation scheme for mathematical programs with equilibrium constraints. *SIAM J. Optim.*, 16(2):587–609, 2005.
- [DM02] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, 91:201–213, 2002.
- [FL04] R. Fletcher and S. Leyffer. Solving mathematical programs with complementary constraints as nonlinear programs. *Optim. Methods Softw.*, 19(1):15–40, 2004.
- [FL05] M. Fukushima and G.H. Lin. A modified relaxation scheme for mathematical programs with complementarity constraints. *Ann. Oper. Res.*, 133:63–84, 2005.
- [Fle00] R. Fletcher. *Practical Methods of Optimization, Second Edition*. John Wiley & Sons, New York, 2000.
- [FLRS06] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. Local convergence of SQP methods for mathematical programs with equilibrium constraints. *SIAM J. Optim.*, 17(1):259–286, 2006.
- [FP97] M.C. Ferris and J.S. Pang. Engineering and economic applications of complementarity problems. *SIAM Review*, 39(4):669–713, 1997.
- [FP99] M. Fukushima and J.-S. Pang. Complementarity constraint qualifications and simplified  $B$ -stationary conditions for mathematical programs with equilibrium constraints. *Comput. Optim. Appl.*, 13(1-3):111–136, 1999.
- [FP03] F. Facchinei and J.S. Pang. *Finite-dimensional variational inequalities and complementarity problems*, volume 1. Springer Series in Operations Research. New York, Springer, 2003.
- [FTL99] M.C. Ferris and F. Tin-Loi. On the solution of a minimum weight elastoplastic problem involving displacement and complementarity constraints. *Comput. Methods Appl. Mech. Engrg.*, 174:107–120, 1999.
- [HR04] X.M. Hu and D. Ralph. Convergence of a penalty method for mathematical programming with complementarity constraints. *J. Optim. Theory Appl.*, 123(2):365–390, 2004.
- [Jan84] R. Janin. Directional derivative of the marginal function in nonlinear programming. *Mathematical Programming Study*, 21:110–126, 1984.
- [KOZ98] M. Kočvara, J. Outrata, and J. Zowe. *Nonsmooth approach to optimization problems with equilibrium constraints. Theory, applications and numerical results*. Nonconvex Optimization and Its Applications. 28. Dordrecht: Kluwer Academic Publishers., 1998.
- [Ley00] S. Leyffer. MacMPEC: AMPL collection of MPECs. [www.mcs.anl.gov/leyffer/MacMPEC](http://www.mcs.anl.gov/leyffer/MacMPEC), 2000.
- [Ley06] S. Leyffer. Complementarity constraints as nonlinear equations: theory and numerical experience. Dempe, Stephan (ed.) et al., *Optimization with multivalued mappings*. New York, NY: Springer. Springer Optimization and Its Applications 2, 169-208 (2006)., 2006.
- [LLCN06] S. Leyffer, G. López-Calva, and J. Nocedal. Interior methods for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 17(1):52–77, 2006.
- [LPR96] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints*. Cambridge: Cambridge Univ. Press., 1996.
- [LPRW96] Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.Q. Wu. Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints. *Math. Program.*, 75(1):19–76, 1996.
- [Pad99] M. Padberg. *Linear Optimization and Extensions*. Springer, second edition, 1999.
- [RW04] D. Ralph and S.J. Wright. Some properties of regularization and penalization schemes for MPECs. *Optim. Methods Softw.*, 19(5):527–556, 2004.

- [Sch01] S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 11(4):918–936, 2001.
- [SS99] S. Scholtes and M. Stöhr. Exact penalization of mathematical programs with equilibrium constraints. *SIAM J. Control Optim.*, 37(2):617–652, 1999.
- [SS00] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. *Math. Oper. Res.*, 25(1):1–22, 2000.
- [Ye05] J.J. Ye. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *J. Math. Anal. Appl.*, 307(1):350–369, 2005.

### Appendix A.

The following Lemma is an auxiliary result that we use in the proof of Theorem 5.2.

LEMMA A.1. *Let  $(x, \lambda, \mu)$  be a KKT-triple of the NLP*

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & g(x) \geq 0 \\ & h(x) = 0. \end{aligned}$$

*Then there exist feasible multipliers  $(\bar{\lambda}, \bar{\mu})$ , such that*

$$\bar{\lambda}_j = 0 \quad \forall j : \lambda_j = 0, \quad \bar{\mu}_j = 0 \quad \forall j : \mu_j = 0,$$

*and the system of vectors*

$$\begin{aligned} \nabla g_j(x) & \quad j \in I_g(x), \bar{\lambda}_j > 0 \\ \nabla h_j(x) & \quad j \in \{i \in I_h \mid \bar{\mu}_i \neq 0\} \end{aligned}$$

*is linearly independent.*

*Proof.* Since  $(x, \lambda, \mu)$  is a KKT-triple, there holds

$$(A.1) \quad \nabla f(x) - \sum_{j \in I_g(x)} \lambda_j \nabla g_j(x) - \sum \mu_j \nabla h_j(x) = 0, \quad \lambda_j \geq 0.$$

If we substitute  $\mu_j := \mu_j^+ - \mu_j^-$  in (A.1) with  $\mu_j^+ = \max(0, \mu_j)$ ,  $\mu_j^- = \min(0, -\mu_j)$ , then  $\mu_j^+, \mu_j^- \geq 0$ , and finding multipliers satisfying (A.1) corresponds to finding a solution to

$$(A.2) \quad Az = b, \quad z \geq 0,$$

where  $b = \nabla f(x)$  and the columns of  $A$  are composed by the gradient vectors  $\nabla g_j(x)$  such that  $\lambda_j > 0$ ,  $\nabla h_j(x)$  such that  $\mu_j^+ > 0$ , and  $-\nabla h_j(x)$  such that  $\mu_j^- > 0$ . Applying a result of linear programming, see for example [Pad99], it is possible to find a  $\bar{z}$  solving (A.2) such that the columns of  $A$  corresponding to  $\{j : \bar{z}_j \neq 0\}$  are linearly independent. Now  $\bar{\lambda}$  and  $\bar{\mu}$  can be easily obtained from  $\bar{z}$ . □