Mathematical Programming manuscript No.
(will be inserted by the editor)

Xinwei Liu · Jie Sun

Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints

Received: date / Revised version: date

Abstract. Generalized stationary points of the mathematical program with equilibrium constraints (MPEC) are studied to better describe the limit points produced by interior point methods for MPEC. A primal-dual interior-point method is then proposed, which solves a sequence of relaxed barrier problems derived from MPEC. Global convergence results are deduced under fairly general conditions other than strict complementarity or the linear independence constraint qualification for MPEC (MPEC-LICQ). It is shown that every limit point of the generated sequence is a strong stationary point of MPEC if the penalty parameter of the merit function is bounded. Otherwise, a point with certain stationarity can be obtained. Preliminary numerical results are reported, which include a case analyzed by Leyffer for which the penalty interior-point algorithm failed to find a stationary point.

Key words. Global convergence, interior-point methods, mathematical programming with equilibrium constraints, stationary point

1. Introduction

Given functions $f : \Re^{n+m} \to \Re$, $c : \Re^{n+m} \to \Re^p$, $g : \Re^{n+m} \to \Re^\ell$, and $F : \Re^{n+m} \to \Re^m$, consider the mathematical program with equilibrium constraints (MPEC):

$$\min \quad f(x, y) \tag{1}$$
$$\text{s.t.} \quad c(x, y) \le 0, \tag{2}$$
$$y \in \mathcal{S}(x), \tag{3}$$

where $\mathcal{S}(x)$ is the solution set of a parametric variational inequality problem (PVI)

$$y \in \mathcal{S}(x) \iff \begin{cases} g(x, y) \le 0, \\ F(x, y)^\top (z - y) \ge 0, \ \forall \ z \text{ such that } g(x, z) \le 0. \end{cases} \tag{4}$$

Throughout the paper, we suppose that f, c, and F are twice continuously differentiable and that g is triply continuously differentiable. We note that if $\ell = m$

X.-W. Liu: Singapore-MIT Alliance, National University of Singapore and Department of Applied Mathematics, Hebei University of Technology, Tianjin, China, e-mail: smaliuxw@nus.edu.sg

J. Sun: School of Business and Singapore-MIT Alliance, National University of Singapore, Republic of Singapore. Fax: (65)6779-2621, e-mail: jsun@nus.edu.sg

Mathematics Subject Classification (1991): 90C30, 90C33, 90C55, 49M37, 65K10

and $g(x,y) = -y$, then the PVI is reduced to a parametric nonlinear complementarity problem, in which case MPEC is specifically called the mathematical program with complementarity constraints (MPCC).

MPEC includes the bilevel programming problem (e.g., [14,45]) as its special case and has extensive applications in practical areas such as traffic control, engineering design, and economic modeling, see [2,26,35,36]. Since there are variational inequalities in the constraints of the problem, the feasible region may be nonconvex, nonsmooth, disconnected, and non-closed even if all involved functions have very good analytical properties (see [35]). As such, MPEC is known to be a class of very difficult optimization problems [4,8,9].

There have been many papers dealing with MPEC in recent years. Some of them considered the existence and stationarity of its solution, for example [19, 25,27,35,37,39,44], while some other papers proposed algorithms for MPEC, see [13,18,20,21,28,35,36,38,41,43]. Upon the success of interior-point methods for linear and nonlinear programming (NLP), the interior-point approach has been extended to solve MPEC as well. The penalty interior-point algorithm (PIPA) developed by Luo, Pang and Ralph [35] is the first interior-point method for MPEC. Its global convergence requires the linear independence constraint qualification for MPEC (MPEC-LICQ) and strict complementarity. However, it was found recently by Leyffer [29] that some conditions required by PIPA for convergence may collapse at some iterates. As a result, PIPA may fail to find a stationary point for a simple MPCC.

An interesting idea of designing algorithms for MPEC, discussed in a series of recent studies, e.g., [5,6,10,11,13,15,31,40] [1], is to reformulate the MPEC as an NLP problem or a sequence of NLP problems and to solve the reformulated problem(s) by various algorithms of NLP. On one hand, this approach allows us to take advantage of certain NLP algorithms to obtain rapid local convergence (see Fletcher, Leyffer, Ralph, and Scholtes [16]) and, on the other hand, it brings new challenges to the NLP algorithm to be used. In particular, since MPECs violate the Mangasarian-Fromovitz constraint qualification [8,42], the usual conditions for global convergence of an NLP algorithm may not be met and the linearized complementarity constraints may be inconsistent even for iterates close to the solutions (see [31] for a detailed analysis). Therefore, a successful NLP algorithm should be able to handle the irregularity of the reformulated problem. An example of such algorithms is due to Fletcher and Leyffer [15] where their filter SQP method shows very good performance on a large collection of test problems called MacMPEC [30]. Benson, Shanno and Vanderbei [6] identified the possible difficulties in convergence when applying LOQO (an interior-point method for NLP) to MPEC, some heuristics for implementation were suggested to overcome difficulties in global convergence.

In this paper we present an interior-point method for MPEC. The method, together with its convergence theory, is an extension of a robust method [32,34] developed by the authors for NLP. The original motivation of that research was

[1] The papers [5,10,31,40] were available to us after the first draft of this paper was submitted.

to develop an algorithm that can handle some "bad" cases of NLP. The idea and analysis used in that research turn out to be useful in studying MPEC as well. In the context of MPEC, the PVI constraint in MPEC is first transformed to a group of complementarity constraints, the complementarity equations are then relaxed to inequalities with a relaxation parameter θ (see (28) below). We then solve the relaxed problem by an interior-point method in which the barrier parameter μ is a fixed fraction of θ, so it is decreased simultaneously with θ. To our knowledge, the relaxation scheme is known (see for instance Scholtes [41]) for some time, but is not fully explored and the way of reducing θ and μ appears to be new. The method has the following properties.

1. All linearized constraints including the linearized complementarity constraints are always consistent.
2. Global convergence results are derived under fairly general conditions other than the MPEC-LICQ or the strict complementarity condition.
3. Under certain conditions, the algorithm can find a point with certain stationarity. Specifically, it is shown that every limit point of the generated sequence is a strong stationary point of the MPEC if the penalty parameter of the merit function is bounded. Otherwise, one of the limit points could be a singular stationary point, an infeasible stationary point, or a weak stationary point (All of the related definitions will be given later).

The method has been implemented to solve the test problems of Facchinei, Jiang and Qi [13] with satisfactory results. We present two numerical examples to show the behavior of the algorithm when the problem is infeasible and when the MPEC-LICQ does not hold, respectively. In addition, we solve the example given by Leyffer [29] for which the PIPA fails to find a stationary point. Numerical experience with the MacMPEC test problems are also reported.

The solution of the relaxed barrier problem plays an important role in our method. The search direction is computed in two-steps. First, an auxiliary step is computed through a minimization problem. Then, the auxiliary step is used in a modified primal-dual Newton equation to calculate the search direction. In addition, the barrier function with ℓ_2-penalty is selected as the merit function where the penalty parameter is adjusted adaptively. Different steplengths for the primal and dual updates are used, while special care is taken to avoid rapid reduction in slack variables.

The paper is organized as follows. In Section 2, we define certain weak stationarities of MPEC used in our convergence analysis. In Section 3, we describe the relaxation scheme that paves a way of solving MPEC by interior-point methods. It is shown that, under certain conditions, the KKT points of the relaxed problems converge to strong stationary points of the MPEC as the relaxation parameter tends to zero. In Section 4, we present a primal-dual interior-point method and derive convergence results for the relaxed barrier problems. In Section 5, we describe our algorithm for MPEC and present global convergence results. In Section 6, we report our numerical results.

It is better to clarify some notations at this point. All vectors are column vectors except that for simplicity we write (x, y) to stand for the column vector $[x^\top \ y^\top]^\top$. A vector with superscript k is related to the k-th iterate; its subscript j means its j-th component. All matrices related to iterate k are indexed by subscript k. The norm $\|\cdot\|$ represents the Euclidean norm. $\nabla g_i(x,y) = (\nabla_x g_i(x,y), \nabla_y g_i(x,y)),\ i = 1,...,\ell$, and $\nabla g(x,y) = [\nabla g_1(x,y) \ \ldots \ \nabla g_\ell(x,y)]$, $\nabla g_{\mathcal{J}}(x,y) = [\nabla g_j(x,y)|j \in \mathcal{J}]$, where $\mathcal{J}$ is an index set. For functions involving x, y and other vectors such as $H(x,y,\lambda)$ used below, we use the notations $\nabla H(x,y,\lambda) = (\nabla_x H(x,y,\lambda), \nabla_y H(x,y,\lambda))$ and $\nabla_E H(x,y,\lambda) = (\nabla_x H(x,y,\lambda)$, $\nabla_y H(x,y,\lambda)$,$\nabla_\lambda H(x,y,\lambda))$ ("E" for "entire"). For any vector v, diag(v) stands for the diagonal matrix whose diagonal is the vector v.

We often have to deal with different index sets. Here is a list of them, in which λ_j is the multiplier associated with g_j.

$$\begin{aligned}
\mathcal{C}_0(x,y) &= \{j \in \{1,\ldots,p\}|c_j(x,y) = 0\} \\
\mathcal{G}_0(x,y) &= \{j \in \{1,\ldots,\ell\}|g_j(x,y) = 0\} \\
\mathcal{G}_0(\lambda) &= \{j \in \{1,\ldots,\ell\}|\lambda_j = 0\} \\
\mathcal{G}_{00}(x,y,\lambda) &= \{j \in \{1,\ldots,\ell\}|g_j(x,y) = 0, \lambda_j = 0\} \\
\mathcal{G}_{0+}(x,y,\lambda) &= \{j \in \{1,\ldots,\ell\}|g_j(x,y) = 0, \lambda_j > 0\}
\end{aligned}$$

Finally, we denote the feasible set of the MPEC by $\mathcal{F}$ and by **strict complementarity** we mean that $\mathcal{G}_{00}(x,y,\lambda) = \emptyset$.

2. Generalized stationary properties of MPEC

We make the following **blanket assumption** throughout this paper.

Assumption 1.

(1) For every $(x,y) \in \mathcal{F}$, the vectors $\{\nabla_y g_j(x,y)|\ j \in \mathcal{G}_0(x,y)\}$ are linearly independent.

(2) For any fixed $x \in \{x \in \Re^n | c(x,y) \le 0$ for some $y \in \Re^m\}$ and each $j \in \{1,\ldots,\ell\}$, $g_j(x,\cdot)$ is convex.

It should be noted that Assumption 1 always holds in the important special case of MPCC. Under Assumption 1, $y \in \mathcal{S}(x)$ if and only if there is a unique $\lambda \in \Re^\ell$ such that

$$\begin{cases} F(x,y) + \sum_{j=1}^{\ell} \lambda_j \nabla_y g_j(x,y) = 0, \\ \lambda \ge 0,\ g(x,y) \le 0,\ \lambda \circ g(x,y) = 0 \end{cases} \tag{5}$$

where $\circ$ denotes the Hadamard product. In general we designate the set of λ that satisfies (5) as $M(x,y)$. It is easy to show that if Assumption 1 holds and if (x,y) is bounded, then $M(x,y)$ is also bounded and problem (1)-(3) is equivalent to

$$\begin{aligned}
\min \quad & f(x,y) && (6)\\
\text{s.t.} \quad & c(x,y) \le 0, && (7)\\
& H(x,y,\lambda) = 0, && (8)\\
& \lambda \ge 0,\ g(x,y) \le 0,\ \lambda \circ g(x,y) = 0, && (9)
\end{aligned}$$

where $H(x, y, \lambda) = F(x, y) + \sum_{j=1}^{\ell} \lambda_j \nabla_y g_j(x, y)$. However, Assumption 1 does not imply strict complementarity. See the following example.

Example 1. Consider the MPEC

$$\min \quad f(x, y_1, y_2) = \frac{1}{2}(x-1)^2 + y_1 + y_2 \tag{10}$$

$$\text{s.t.} \quad x \geq 0, \tag{11}$$

$$y_1 \geq 0, \; y_2 \geq 0, \tag{12}$$

$$\begin{pmatrix} 2x \\ y_1 - y_2 \end{pmatrix}^\top \begin{pmatrix} z_1 - y_1 \\ z_2 - y_2 \end{pmatrix} \geq 0, \; \forall z_1 \geq 0, \; z_2 \geq 0. \tag{13}$$

Assumption 1 holds at the optimal point $(x^*, y_1^*, y_2^*) = (1, 0, 0)$ with $\lambda_1^* = 2$, $\lambda_2^* = 0$. However, the strict complementarity does not hold.

The following definition is well known.

Definition 1. *A point $(x, y) \in \mathcal{F}$ is a* **B-stationary point** *of MPEC if*

$$\nabla_x f(x, y)^\top d_x + \nabla_y f(x, y)^\top d_y \geq 0, \; \text{for all } (d_x, d_y) \in \mathcal{T}(x, y; \mathcal{F}), \tag{14}$$

where $\mathcal{T}(x, y; \mathcal{F})$ is the tangent cone of $\mathcal{F}$ at (x, y).

It is generally difficult to give an explicit expression of $\mathcal{T}(x, y; \mathcal{F})$. Instead, the following concepts of strong and weak stationary points of MPEC, due to Scholtes and Scheel [42], are often used in algorithmic design.

Definition 2.
(1) A point $(x, y) \in \mathcal{F}$ is a **strong stationary point** *of MPEC if for $\lambda \in M(x, y)$ there exist multipliers $\zeta \in \Re^p$, $\eta \in \Re^\ell$ and $\pi \in \Re^m$ such that*

$$\nabla f(x, y) + \nabla c(x, y)\zeta + \nabla g(x, y)\eta + \nabla H(x, y, \lambda)\pi = 0, \tag{15}$$

$$\zeta^\top c(x, y) = 0, \; \zeta \geq 0, \tag{16}$$

$$\pi^\top \nabla_y g_j(x, y) = 0, \; \forall \; j \in \mathcal{G}_{0+}(x, y, \lambda), \tag{17}$$

$$\eta_j = 0, \; \forall \; j \notin \mathcal{G}_0(x, y). \tag{18}$$

$$\pi^\top \nabla_y g_j(x, y) \geq 0, \; \forall \; j \in \mathcal{G}_{00}(x, y, \lambda), \tag{19}$$

$$\eta_j \geq 0, \; \forall \; j \in \mathcal{G}_{00}(x, y, \lambda). \tag{20}$$

hold.
(2) A point $(x, y) \in \mathcal{F}$ is called a **weak stationary point** *of MPEC if for $\lambda \in M(x, y)$ there exist $\zeta \in \Re^p$, $\eta \in \Re^\ell$, and $\pi \in \Re^m$ such that (15)-(18) hold.*

Obviously, a strong stationary point must be a weak stationary point, but not vice versa. However, under strict complementarity, the two concepts are equivalent since the set $\mathcal{G}_{00}(x, y, \lambda)$ is vacuous. In general, B-stationarity and strong stationarity do not imply each other unless the so-called MPEC-LICQ holds, see Proposition 1 below.

Definition 3. *For any* $(x,y)\in\mathcal{F}$ *and* $\lambda\in M(x,y)$*, the* **MPEC-LICQ** *holds at* (x,y) *if*

$$\begin{pmatrix} \nabla H(x,y,\lambda) & \nabla c_{\mathcal{C}_0(x,y)}(x,y) & \nabla g_{\mathcal{G}_0(x,y)}(x,y) & 0 \\ \nabla_\lambda H(x,y,\lambda) & 0 & 0 & [e_j, j\in\mathcal{G}_0(\lambda)] \end{pmatrix} \tag{21}$$

has full column rank, where e_j *is the* j*-th coordinate vector.*

Proposition 1. *If MPEC-LICQ holds at* $(x^*,y^*)\in\mathcal{F}$*, then* (x^*,y^*) *is a B-stationary point of MPEC if and only if it is a strong stationary point of MPEC.*

This proposition can be derived in a similar way to the derivation of Theorem 3.3.4 in [35], where the result has been proved in a more general setting. Similar results are obtained in [39,42].

To describe convergence properties of our algorithm, we need to consider other types of stationarity.

Definition 4.

(1) A point $(x,y)\in\mathcal{F}$ *is called a* **singular stationary point** *of MPEC if the MPEC-LICQ does not hold at* (x,y)*.*

(2) A point (x,y) *is called an* **infeasible stationary point** *of MPEC if* $(x,y)\notin\mathcal{F}$*, and for some* $\lambda\in\Re^\ell$ *and some scalar* $\theta>0$*,* (x,y,λ) *is a stationary point of the problem*

$$\min_{(x,y,\lambda)}\ \{\|c_+\|^2+\|H\|^2+\|g_+\|^2+\|\lambda_-\|^2+\|(\lambda\circ g+\theta e)_-\|^2\}, \tag{22}$$

that is, (x,y,λ) *satisfies the following equations*

$$\nabla c\ c_+ + \nabla H\ H + \nabla g\ g_+ + \nabla g\ \Lambda(\lambda\circ g+\theta e)_- = 0, \tag{23}$$

$$\nabla_y g^\top H + \Lambda\lambda_- + diag(g)(\lambda\circ g+\theta e)_- = 0, \tag{24}$$

where $\Lambda = diag(\lambda)$*,* $H=H(x,y,\lambda)$*,* $c_+=\max\{c(x,y),0\}$*,* $g_+=\max\{g(x,y),0\}$*,* $\lambda_-=\min\{\lambda,0\}$*,* $e=(1,\ldots,1)$ *and* $(\lambda\circ g+\theta e)_-=\min\{\lambda\circ g(x,y)+\theta e,0\}$*.*

These definitions were first given in the context of NLP in [32]. Here we extend them to MPEC. We note that the objective function (22) can be thought of as the ℓ_2-measure of the total infeasibility of problem (25)-(28) below, so the infeasible stationarity makes sense. Of course, if (x,y,λ) is a feasible point, then for any $\theta>0$, this quantity is zero.

3. An NLP relaxation of MPEC

Suppose $\theta>0$ is a parameter. By θ-relaxation of MPEC we mean the following nonlinear program

$$\text{(NLP}(\theta)) \qquad \min\ f(x,y) \tag{25}$$

$$\text{s.t.}\ \ c(x,y)\le 0, \tag{26}$$

$$H(x,y,\lambda)=0, \tag{27}$$

$$\lambda\ge 0,\ g(x,y)\le 0,\ -\lambda\circ g(x,y)\le\theta e, \tag{28}$$

where the complementarity constraints in the reformulated MPEC (6)-(9) are relaxed into inequalities. It is obvious that if $\theta = 0$ then (25)-(28) reduces to (6)-(9). The following result shows that the MPEC-LICQ implies the LICQ of NLP(θ) for all sufficiently small θ.

Proposition 2. *For $(x^*, y^*) \in \mathcal{F}$ and $\lambda^* \in M(x^*, y^*)$, if the MPEC-LICQ holds at (x^*, y^*), then there exists a neighborhood $\mathcal{N}$ of (x^*, y^*, λ^*) so that for sufficiently small $\theta > 0$, the LICQ holds at any $(\bar{x}, \bar{y}, \bar{\lambda}) \in \mathcal{N}$ feasible to NLP(θ).*

The proof of Proposition 2 is based on a continuity argument and similar results have been proved in [20,41]. For brevity, we omit it. To simplify the notation, let

$$\tilde{G}(x, y, \lambda) = (c(x, y),\ g(x, y),\ -\lambda), \tag{29}$$

$$G_\theta(x, y, \lambda) = (\tilde{G}(x, y, \lambda),\ -\lambda \circ g(x, y) - \theta e). \tag{30}$$

Then the constraints of NLP(θ) can be written as $G_\theta(x, y, \lambda) \le 0,\ H(x, y, \lambda) = 0$ and

$$\nabla G_\theta(x, y, \lambda) = [\nabla c(x, y)\ \nabla g(x, y)\ 0\ \ -[\nabla g(x, y)]\lambda]\,, \tag{31}$$

$$\nabla_\lambda G_\theta(x, y, \lambda) = [0\ 0\ \ -I\ \ -\mathrm{diag}\,(g(x, y))]\,, \tag{32}$$

where I is the $\ell \times \ell$ identity matrix. The Lagrange function of program (25)-(28) is

$$L_\theta(x, y, \lambda, u, v) = f(x, y) + u^\top G_\theta(x, y, \lambda) + v^\top H(x, y, \lambda), \tag{33}$$

where $u \in \Re_+^{p+3\ell}$ and $v \in \Re^m$ are the multipliers. Let $\bar{u} = (u_1, \ldots, u_p)$, $\hat{u} = (u_{p+1}, \ldots, u_{p+\ell})$ and $\tilde{u} = (u_{p+2\ell+1}, \ldots, u_{p+3\ell})$. By using (31)-(32) and noting that $\nabla_\lambda H(x, y, \lambda) = \nabla_y g(x, y)^\top$, the KKT conditions of NLP(θ) can be written as

$$\nabla f(\bar{x}, \bar{y}) + \nabla c(\bar{x}, \bar{y})\bar{u} + \nabla g(\bar{x}, \bar{y})\eta + \nabla H(\bar{x}, \bar{y}, \bar{\lambda})v = 0, \tag{34}$$

$$\nabla_y g(\bar{x}, \bar{y})^\top v \ge \mathrm{diag}\,(g(\bar{x}, \bar{y}))\tilde{u},\ \ \bar{\lambda} \circ (\nabla_y g(\bar{x}, \bar{y})^\top v - \mathrm{diag}\,(g(\bar{x}, \bar{y}))\tilde{u}) = 0, \tag{35}$$

$$\bar{\lambda} \ge 0,\ \ \tilde{u} \ge 0,\ \ -\bar{\lambda} \circ g(\bar{x}, \bar{y}) - \theta e \le 0,\ \ \tilde{u} \circ (-\bar{\lambda} \circ g(\bar{x}, \bar{y}) - \theta e) = 0, \tag{36}$$

$$\eta - (\hat{u} - \bar{\lambda} \circ \tilde{u}) = 0, \tag{37}$$

$$\hat{u} \ge 0,\ \ g(\bar{x}, \bar{y}) \le 0,\ \ \hat{u} \circ g(\bar{x}, \bar{y}) = 0, \tag{38}$$

$$\bar{u} \ge 0,\ \ c(\bar{x}, \bar{y}) \le 0,\ \ \bar{u} \circ c(\bar{x}, \bar{y}) = 0, \tag{39}$$

$$H(\bar{x}, \bar{y}, \bar{\lambda}) = 0. \tag{40}$$

Now we show that as $\theta \to 0$, the KKT points of NLP(θ) converge to a strong stationary point of MPEC if the primal and dual variables are bounded.

Proposition 3. *Suppose that $(\bar{x}, \bar{y}, \bar{\lambda})$ is a KKT point of NLP(θ), $(\bar{u}, \hat{u}, \tilde{u}, v)$ is the corresponding multiplier vector associated with constraint $(c, g, -\lambda \circ g - \theta e, H)$. If the sequence $\{(\bar{x}, \bar{y}, \bar{\lambda}, \bar{u}, \hat{u}, \tilde{u}, v)\}$ is uniformly bounded as $\theta \to 0$ and $(x^*, y^*, \lambda^*, \bar{u}^*, \hat{u}^*, \tilde{u}^*, v^*)$ is one of its limit points, then (x^*, y^*) is a strong stationary point of the MPEC (1)-(3).*

Proof. By using $\bar{u}^*$, η^* and v^* to replace ζ^*, η^* and π^*, it is easy to see that (15) and (16) hold because of (34) and (39). Assumption 1, (36)-(39), and (40) imply that $(x^*, y^*) \in \mathcal{F}$ and λ^* is the unique element of $M(x^*, y^*)$. For any $j \in \mathcal{G}_{00}(x^*, y^*, \lambda^*)$, one has $\lambda_j^* = 0$ and $g_j^* = 0$. Thus, by (35), $v^{*\top}\nabla_y g_j^* \geq g_j^* \tilde{u}_j^* = 0$, which proves (19). For $j \in \mathcal{G}_{0+}(x^*, y^*, \lambda^*)$, by (35), one has $v^{*\top}\nabla_y g_j^* = g_j^* \tilde{u}_j^* = 0$, which proves (17). Finally, we have (18) and (20) by (37). ∎

In the remainder of this section, we will use $(\bar{x}, \bar{y}, \bar{\lambda})$ to denote a KKT point of NLP(θ). The following example shows the usage of the θ-relaxation and the role of Proposition 3.

$$\min \; f(x,y) = \frac{1}{2}\left[(x-1)^2 + (y-1)^2\right] \tag{41}$$

$$\text{s.t.} \; 2x - \lambda = 0, \tag{42}$$

$$\lambda \geq 0, \; y \geq 0, \; \lambda y \leq \theta, \tag{43}$$

where $\theta > 0$. Its KKT conditions are as follows.

$$x - 1 + 2v = 0, \; y - 1 - \hat{u} + \lambda\tilde{u} = 0, \; 2x - \lambda = 0, \tag{44}$$

$$\lambda \geq 0, \; -v + y\tilde{u} \geq 0, \; \lambda(-v + y\tilde{u}) = 0, \tag{45}$$

$$\tilde{u} \geq 0, \; \lambda y - \theta \leq 0, \; \tilde{u}(\lambda y - \theta) = 0, \tag{46}$$

$$\hat{u} \geq 0, \; y \geq 0, \; \hat{u}y = 0. \tag{47}$$

For $\theta \leq \frac{1}{2}$, all solutions satisfy $\lambda y = \theta$ and they are

$$(i) \quad (\bar{x}, \bar{y}, \bar{\lambda}) = \frac{1}{2}(1,1,2) + \left(\sqrt{\frac{1}{4} - \frac{1}{2}\theta}\right)(1,-1,2), \; (\hat{u}, \tilde{u}) = \left(0, \frac{1}{2}\right); \tag{48}$$

$$(ii) \quad (\bar{x}, \bar{y}, \bar{\lambda}) = \frac{1}{2}(1,1,2) - \left(\sqrt{\frac{1}{4} - \frac{1}{2}\theta}\right)(1,-1,2), (\hat{u}, \tilde{u}) = \left(0, \frac{1}{2}\right); \tag{49}$$

$$(iii) \quad (\bar{x}, \bar{y}, \bar{\lambda}) = \frac{\sqrt{2}}{2}\theta(1,1,2), \; (\hat{u}, \tilde{u}) = \left(0, \frac{\sqrt{2} - \theta}{2\theta}\right). \tag{50}$$

It is easy to see that solutions (i) and (ii) converge to the strong stationary points $(1,0)$ and $(0,1)$ respectively, but solution (iii) does not where $\tilde{u}$ is unbounded. Satisfication of the MPEC-LICQ at (1,0) and (0,1) implies that they are also B-stationary points.

The example shows that the success of solving MPEC by NLP(θ) depends on the dual boundedness of NLP(θ). We next derive a sufficient condition for this property. We start with some definitions.

Definition 5. *A sequence* $\{(\bar{x}, \bar{y}, \bar{\lambda})\}$ *is* **asymptotically weakly nondegenerate,** *if* $(\bar{x}, \bar{y}, \bar{\lambda}) \to (x^*, y^*, \lambda^*)$ *as* $\theta \to 0$, *and there is a* $\bar{\theta} > 0$ *such that for* $\theta \in (0, \bar{\theta})$ *and all* $i \in \mathcal{G}_{00}(x^*, y^*, \lambda^*) \cap \mathcal{I}_\theta$, *there exist constants* $\varsigma_1 \geq \varsigma_2 > 0$ *such that* $\varsigma_1 \geq |g_i(\bar{x}, \bar{y})/\bar{\lambda}_i| \geq \varsigma_2$, *where* $\mathcal{I}_\theta = \{i \mid -\bar{\lambda}_i g_i(\bar{x}, \bar{y}) = \theta\}$.

This definition is of similar nature to that given by Fukushima and Pang [20], which requires that $\bar{\lambda}_i$ and $g_i(\bar{x}, \bar{y})$ tend to zero in the same order. It is noted that if the strict complementarity holds at (x^*, y^*), then the asymptotically weakly nondegenerate condition holds since $\mathcal{G}_{00}(x^*, y^*, \lambda^*) = \emptyset$, but not vice versa.

The following definition is well known in the theory of NLP.

Definition 6. *For $\theta > 0$, the* **second-order necessary optimality condition** *of NLP(θ) holds at $(\bar{x}, \bar{y}, \bar{\lambda})$ if $(\bar{x}, \bar{y}, \bar{\lambda})$ is feasible to NLP(θ), and there exist $u = (\bar{u}, \hat{u}, \underline{u}, \tilde{u}) \in \Re^{p+\ell+\ell+\ell}$ and $v \in \Re^m$ such that (34)-(40) are satisfied and*

$$d^\top \nabla^2 L_\theta(\bar{x}, \bar{y}, \bar{\lambda}, u, v) d \geq 0 \tag{51}$$

for all $d = (d_a, d_\lambda)$ satisfying

$$\nabla c_i(\bar{x}, \bar{y})^\top d_a = 0, \ \textit{for } i \in \mathcal{C}_0(\bar{x}, \bar{y}); \tag{52}$$
$$\nabla_E H(\bar{x}, \bar{y}, \bar{\lambda})^\top d = 0; \tag{53}$$
$$\nabla g_i(\bar{x}, \bar{y})^\top d_a = 0, \ \textit{for } i \in \mathcal{G}_0(\bar{x}, \bar{y}); \tag{54}$$
$$\bar{\lambda}_i \nabla g_i(\bar{x}, \bar{y})^\top d_a + g_i(\bar{x}, \bar{y}) d_{\lambda i} = 0, \ i \in \mathcal{I}_\theta. \tag{55}$$

We have the following sufficient conditions for the dual boundedness required by Proposition 3, which is similar in flavor to Theorem 3.1 of [20].

Proposition 4. *Suppose that $\{(\bar{x}, \bar{y}, \bar{\lambda})\}$ is bounded as $\theta \to 0$, Θ is an infinite set of θ in a sufficiently small neighborhood of zero such that $(\bar{x}, \bar{y}, \bar{\lambda}) \to (x^*, y^*, \lambda^*)$ as $\theta \in \Theta$ and $\theta \to 0$. Then $\{(\bar{u}, \tilde{u}, \hat{u}, v) \mid \theta \in \Theta\}$ is bounded if*

$$\begin{cases} \textit{the second order necessary optimality condition of NLP}(\theta) \textit{ holds at } (\bar{x}, \bar{y}, \bar{\lambda}) \\ \ \textit{for } \theta \in \Theta, \\ \{(\bar{x}, \bar{y}, \bar{\lambda}) \mid \theta \in \Theta\} \textit{ is asymptotically weakly nondegenerate, and} \\ \textit{the MPEC-LICQ holds at } (x^*, y^*, \lambda^*). \end{cases}$$

Proof. By the MPEC-LICQ and (34), $\{(\bar{u}, \eta, v) \mid \theta \in \Theta\}$ is uniformly bounded. Thus, for $\theta \in \Theta$, $\hat{u}_i$, where $i \in \mathcal{G}_0(\bar{x}, \bar{y})$, and $\bar{\lambda}_i \tilde{u}_i$, where $i \in \mathcal{I}_\theta) = \{i| - \bar{\lambda}_i g_i(\bar{x}, \bar{y}) = -\theta\}$, are also bounded. Suppose $\tilde{u}$ is unbounded. Let $\pi_\theta(x, y, \lambda) = u^\top G_\theta(x, y, \lambda)$, $d_a = (d_x, d_y)$ and $d = (d_a, d_\lambda)$, then

$$\begin{aligned} d^\top \nabla_E^2 \pi_\theta(\bar{x}, \bar{y}, \bar{\lambda}) d = & \sum_{i \in \mathcal{C}_0(\bar{x}, \bar{y})} \bar{u}_i d_a^\top \nabla^2 c_i(\bar{x}, \bar{y}) d_a + \sum_{i \in \mathcal{G}_0(\bar{x}, \bar{y})} \hat{u}_i d_a^\top \nabla^2 g_i(\bar{x}, \bar{y}) d_a \\ & - \sum_{i \in \mathcal{I}_\theta} \bar{\lambda}_i \tilde{u}_i d_a^\top \nabla^2 g_i(\bar{x}, \bar{y}) d_a - 2 \sum_{i \in \mathcal{I}_\theta} \tilde{u}_i d_{\lambda i} \nabla g_i(\bar{x}, \bar{y})^\top d_a. \end{aligned} \tag{56}$$

It follows from the MPEC-LICQ at (x^*, y^*), Proposition 2, and the asymptotical weak nondegeneracy that there exists a bounded sequence $\{d_\theta\}$ with $d_\theta = (d_a, d_\lambda)$ and $d_\theta \neq 0$ for all $\theta \leq \bar{\theta}$ such that

$$\nabla c_i(\bar{x}, \bar{y})^\top d_a = 0, \ \forall \ i \in \mathcal{C}_0(\bar{x}, \bar{y}); \tag{57}$$
$$\nabla_E H(\bar{x}, \bar{y}, \bar{\lambda})^\top d = 0; \tag{58}$$
$$\nabla g_i(\bar{x}, \bar{y})^\top d_a = 0, \ \forall \ i \in \mathcal{G}_0(\bar{x}, \bar{y}); \tag{59}$$
$$\nabla g_i(\bar{x}, \bar{y})^\top d_a = g_i(\bar{x}, \bar{y}) / \bar{\lambda}_i, \ d_{\lambda i} = -1, \ \forall \ i \in \mathcal{I}_\theta^1; \tag{60}$$
$$\nabla g_i(\bar{x}, \bar{y})^\top d_a = -1, \ d_{\lambda i} = \bar{\lambda}_i / g_i(\bar{x}, \bar{y}), \ \forall \ i \in \mathcal{I}_\theta^2 \backslash \mathcal{I}_\theta^1, \tag{61}$$

where $\mathcal{I}_\theta^1 = \{i \in \mathcal{I}_\theta \mid g_i(x^*, y^*) = 0\}$, $\mathcal{I}_\theta^2 = \{i \in \mathcal{I}_\theta \mid \lambda_i^* = 0\}$. By (60) and (61), we have $\bar{\lambda}_i \nabla g_i(\bar{x}, \bar{y})^\top d_a + g_i(\bar{x}, \bar{y}) d_{\lambda i} = 0,\ i \in \mathcal{I}_\theta$. By (56), there holds

$$d_\theta^\top \nabla_E^2 \pi_\theta(\bar{x}, \bar{y}, \bar{\lambda}) d_\theta \to -\infty \text{ as } \theta \to 0, \tag{62}$$

which contradicts the second-order necessary optimality condition since

$$d_\theta^\top \left[\nabla_E^2 f(\bar{x}, \bar{y}) + \sum_{i=1}^m v_i \nabla_E^2 H_i(\bar{x}, \bar{y}, \bar{\lambda}) \right] d_\theta$$

is bounded. The result follows immediately. ∎

Let us recall the example (41)-(43). For solution (i), it is apparent that the MPEC-LICQ and the strict complementarity hold at the limit $(1, 0, 2)$. Moreover, it is easy to verify that the second-order necessary optimality condition holds since for all $d = (d_x, d_y, d_\lambda)$ satisfying

$$2d_x \qquad - \ d_\lambda = 0, \tag{63}$$

$$\bar{\lambda} d_y + \bar{y} d_\lambda = 0, \tag{64}$$

we have

$$d^\top \nabla^2 L(\bar{x}, \bar{y}, \bar{\lambda}, \hat{u}, \tilde{u}, v) d = \left[\frac{1}{2} - \frac{\theta}{(1 + \sqrt{1 - 2\theta})^2} \right]^2 d_\lambda^2 \geq 0. \tag{65}$$

The same is true for solution (ii). Thus the three conditions in Proposition 4 hold for solutions (i) and (ii). Consequently, the dual variables of (i) and (ii) are bounded as $\theta \to 0$. On the other hand, at solution (iii) some of dual variables are unbounded and the second order necessary optimality condition does not hold. Hence, the results are consistent with Proposition 4.

4. The relaxed barrier problem

We note that applying interior-point approach to (6)-(9) directly will result in a conflict. This is because, by introducing slack variables, the barrier problem corresponding to (6)-(9) is

$$\min \ f(x, y) - \sum_{i=1}^p \mu \ln \bar{z}_i - \sum_{j=1}^\ell \mu \ln \hat{z}_j - \sum_{j=1}^\ell \mu \ln \hat{\lambda}_j \tag{66}$$

$$\text{s.t.} \ c(x, y) + \bar{z} = 0, \tag{67}$$

$$H(x, y, \lambda) = 0, \tag{68}$$

$$g(x, y) + \hat{z} = 0,\ -\lambda + \hat{\lambda} = 0,\ -\lambda \circ g(x, y) = 0, \tag{69}$$

where $\mu > 0$ is the barrier parameter, $\bar{z} > 0,\ \hat{z} > 0$ and $\hat{\lambda} > 0$ are slack variable vectors. Constraints (69) indicate that for $j = 1, \ldots, \ell$ we must have $\hat{\lambda}_j = 0$ if

$\hat{z}_j$ is bounded away from zero and vice versa, which conflicts with the objective function (66) that attempts to draw both $\hat{z}_j$ and $\hat{\lambda}_j$ away from boundary.

We therefore consider to apply the interior-point approach to the θ-relaxation of MPEC, which leads us to the following θ-relaxed log-barrier problem, henceforth referred as the **relaxed barrier problem**:

$$\min \ f(x,y) - \sum_{i=1}^{p} \mu \ln \bar{z}_i - \sum_{j=1}^{\ell} \mu \ln \hat{z}_j - \sum_{j=1}^{\ell} \mu \ln \hat{\lambda}_j - \sum_{j=1}^{\ell} \mu \ln \tilde{z}_j \tag{70}$$

$$\text{s.t.} \ c(x,y) + \bar{z} = 0, \tag{71}$$

$$H(x,y,\lambda) = 0, \tag{72}$$

$$g(x,y) + \hat{z} = 0, \tag{73}$$

$$-\lambda + \hat{\lambda} = 0, \tag{74}$$

$$-\lambda \circ g(x,y) + \tilde{z} = \theta e. \tag{75}$$

By using (30), (70)-(75) can simply be written as

$$\min \ f(s) - \sum_{i=1}^{q} \mu \ln z_i \tag{76}$$

$$\text{s.t.} \ G_\theta(s) + z = 0, \tag{77}$$

$$H(s) = 0, \tag{78}$$

where $s = (x,y,\lambda) \in \Re^{n+m+\ell}$ is the variable vector, $z = (\bar{z}, \hat{z}, \hat{\lambda}, \tilde{z})$ is the slack vector, $f(s) = f(x,y)$, $G_\theta(s) = G_\theta(x,y,\lambda)$, $H(s) = H(x,y,\lambda)$ and $q = p + 3\ell$.

In the following two subsections, we describe a primal-dual algorithm for solving problem (76)-(78) for fixed μ and derive global convergence results of the algorithm. The algorithm for MPEC is then presented in Section 5, which uses the algorithm in this section as the inner loop and decreases μ in the outer loop.

4.1. The algorithm for problem (76)-(78).

Define the merit function with ℓ_2 penalty

$$\phi(s,z;\rho) = f(s) - \sum_{i=1}^{q} \mu \ln z_i + \rho \|(G_\theta(s) + z, H(s))\|, \tag{79}$$

where $\rho > 0$ is the penalty parameter, the norm $\|\cdot\|$ is the Euclidian norm.

At the current iterate (s^k, z^k), suppose that $u^k \in \Re^q_+$ and $v^k \in \Re^m$ are the approximate multipliers corresponding to constraints (77) and (78), respectively. Let $Z_k = \operatorname{diag}(z^k)$, $U_k = \operatorname{diag}(u^k)$, $\nabla G^k_\theta = \nabla G_\theta(s^k)$, $\nabla H^k = \nabla H(s^k)$ and $\nabla f_k = \nabla f(s^k)$. Let B_k be a positive definite approximation to the Lagrangian Hessian

$$\nabla^2 L(s^k, u^k, v^k) = \nabla^2 f_k + \sum_{i=1}^{q} u^k_i \nabla^2 (G_\theta)^k_i + \sum_{j=1}^{m} v^k_j \nabla^2 H^k_j.$$

Suppose that $(\hat{d}_s^k, \hat{d}_z^k)$ is an approximate solution of the problem

$$\min\ \psi_k(d_s, d_z) = \frac{1}{2}(d_s^\top B_k d_s + d_z^\top Z_k^{-1} U_k d_z) + \rho_k \|(G_\theta^k + z^k + \nabla {G_\theta^k}^\top d_s + d_z, H^k + \nabla {H^k}^\top d_s)\| \quad (80)$$

such that some prescribed conditions (see the next subsection) hold. Then we compute the search direction $(d_s^k, d_z^k, d_u^k, d_v^k)$ by solving the modified primal-dual system of equations

$$B_k d_s + \nabla G_\theta^k d_u + \nabla H^k d_v = -(\nabla f_k + \nabla G_\theta^k u^k + \nabla H^k v^k), \quad (81)$$

$$U_k d_z + Z_k d_u = -(Z_k U_k e - \mu e), \quad (82)$$

$$\nabla {G_\theta^k}^\top d_s + d_z = \nabla {G_\theta^k}^\top \hat{d}_s^k + \hat{d}_z^k, \quad (83)$$

$$\nabla {H^k}^\top d_s = \nabla {H^k}^\top \hat{d}_s^k. \quad (84)$$

Note that the right-hand-sides of (83) and (84) are different from the traditional interior-point approach. For motivation of this modification the reader is referred to [33,34].

We are now ready to state our algorithm for the relaxed barrier problem with fixed θ and μ.

Algorithm 1. *(The algorithm for problem (76)-(78))*

Step 1 *Set* $\mu > 0$, $(s^0, z^0, u^0, v^0) \in \Re^{n+m+\ell} \times \Re^q_{++} \times \Re^q_{++} \times \Re^m$, $B_0 \in \Re^{(n+m+\ell)\times(n+m+\ell)}$ *and scalars* $\rho_0 > 0$, $\xi \in (0,1)$, $0 < \beta_1 < 1 < \beta_2$, $\sigma_0 \in (0, \frac{1}{2})$. *Let* $k := 0$;

Step 2 *Calculate the primal search direction* (d_s^k, d_z^k) *and the dual direction* (d_u^k, d_v^k) *by the primal-dual system of equations (81)-(84), where* $(\hat{d}_s^k, \hat{d}_z^k)$ *is derived by approximately minimizing (80);*

Step 3 *Let*

$$\pi_k(d^k; \rho_k) = \nabla f_k^\top d_s^k - \mu e^\top Z_k^{-1} d_z^k - \rho_k \delta(d_s^k, d_z^k),$$

where $d^k = (d_s^k, d_z^k)$ *and* $\delta(d_s^k, d_z^k) = \|(G_\theta^k + z^k, H^k)\| - \|(G_\theta^k + z^k + \nabla {G_\theta^k}^\top d_s^k + d_z^k, H^k + \nabla {H^k}^\top d_s^k)\|$. *If*

$$\pi_k(d^k; \rho_k) \le -\frac{1}{2} {d_s^k}^\top B_k d_s^k - \frac{1}{2} {d_z^k}^\top Z_k^{-1} U_k d_z^k, \quad (85)$$

let $\rho_{k+1} = \rho_k$; *Otherwise, we replace* ρ_k *by a larger* ρ_{k+1} *(for example* $\rho_{k+1} \ge 2\rho_k$*) such that (85) holds;*

Step 4 *Compute* $\hat{\alpha}_k \in (0,1]$ *such that* $z^k + \hat{\alpha}_k d_z^k \ge \xi z^k$, *and select firstly* $\sigma \in (0,1]$ *and then* $\gamma_k \in [0,1]$ *as large as possible such that*

$$\phi(s^k + \sigma\hat{\alpha}_k d_s^k, z^k + \sigma\hat{\alpha}_k d_z^k; \rho_{k+1}) - \phi(s^k, z^k; \rho_{k+1}) \le \sigma_0 \sigma \hat{\alpha}_k \pi_k(d^k; \rho_{k+1}), \quad (86)$$

$$\beta_1 \mu e \le (U_k + \gamma_k D_u^k) \max\{z^k + \sigma\hat{\alpha}_k d_z^k, -G_\theta(s^k + \sigma\hat{\alpha}_k d_s^k)\} \le \beta_2 \mu e, \quad (87)$$

where $D_u^k = diag(d_u^k)$. *Let* $\alpha_k = \sigma\hat{\alpha}_k$. *The new primal iterate is generated by*

$$s^{k+1} = s^k + \alpha_k d_s^k, \tag{88}$$

$$z^{k+1} = \max\{z^k + \alpha_k d_z^k, -G_\theta^{k+1}\}, \tag{89}$$

and the new dual iterate is generated by

$$u^{k+1} = u^k + \gamma_k d_u^k,\ \ v^{k+1} = v^k + d_v^k; \tag{90}$$

Step 5 If the stopping criterion holds, stop; else calculate values ∇G_θ^{k+1}, ∇H^{k+1}, ∇f_{k+1}, G_θ^{k+1} *and* H^{k+1}, *update the approximate Hessian* B_k *by* B_{k+1}, *let* $k := k+1$, *and go to Step 2.*

In practical implementations of the algorithm we may use some more flexible update for generating the dual iterate. Since Algorithm 1 is only taken as an inner loop of our algorithm for MPEC, we will give the stopping criterion in the algorithm for MPEC.

4.2. Convergence of Algorithm 1.

Notice that the relaxed barrier problem differs from the inequality-constrained NLP only by adding an equality constraint. To simplify the proofs, we will refer to the results in [32,33] if a proof is lengthy and is the same as in the references.

Suppose that an infinite sequence $\{(s^k, z^k, u^k, v^k)\}$ is produced by Algorithm 1. We need the following general assumptions.

Assumption 2.

(1) $\{s^k\}$ *is bounded. That is, there is an open and bounded set* $\Omega \subset \Re^{n+m+\ell}$ *such that* $s^k \in \Omega$ *for all nonnegative integers* k.

(2) There exist constants $\nu_1 \geq \nu_2 > 0$ *such that* $\nu_2\|d\|^2 \leq d^\top B_k d \leq \nu_1\|d\|^2$ *for all* $d \in \Re^{n+m+\ell}$.

(3) $\nabla H(s^k)$ *has full column rank for all* $k \geq 0$.

The following results can be derived similarly to Lemma 4.2 in [32] and Lemmas 3.2, 3.3, and 3.6 in [33].

Lemma 1. *Under Assumption 2, we have*

(1) $\{z^k\}$ *is bounded;*

(2) $\{u^k\}$ *is componentwise bounded away from zero.*

Furthermore, if $\{\rho_k\}$ *is bounded, then*

(3) $\{z^k\}$ *is componentwise bounded away from zero;*

(4) $\{u^k\}$ *is bounded;*

(5) if $\{(d_s^k, d_z^k, d_u^k)\}$ *is bounded, then there exists* $\alpha^* \in (0,1]$ *such that* $\alpha_k \geq \alpha^*$ *for all* $k \geq 0$.

Lemma 2. *Under Assumption 2, if* $(\hat{d}_s^k, \hat{d}_z^k)$ *solves problem (80) exactly, then* $(\hat{d}_s^k, \hat{d}_z^k)$ *satisfies the following conditions.*

(1) $(\nabla G_\theta^k(G_\theta^k+z^k)+\nabla H^k H^k, Z_k(G_\theta^k+z^k))\to 0$ *as* $(\hat{d}_s^k,\hat{d}_z^k)\to 0$.

(2) It holds that $\psi_k(\hat{d}_s^k,\hat{d}_z^k)\le\psi_k(0,0)$*, and there exist constants* $\hat\rho>0$ *and* $\varsigma>0$ *so that for all* $\rho_k\ge\hat\rho$*,*

$$\psi_k(\hat{d}_s^k,\hat{d}_z^k)-\psi_k(0,0)\le-\varsigma\rho_k\|(\nabla G_\theta^k(G_\theta^k+z^k)+\nabla H^k H^k, Z_k(G_\theta^k+z^k))\|^2.$$

(3) There exist $\nu\in(0,1)$*,* $\hat\rho>0$ *and* $\varpi>0$ *so that* $\forall\ \rho_k\ge\hat\rho$*,*

$$\|(\hat{d}_s^k, Z_k^{-1}\hat{d}_z^k)\|\le\varpi\|(G_\theta^k+z^k,H^k)\| \tag{91}$$

and

$$\psi_k(\hat{d}_s^k,\hat{d}_z^k)\le\nu\psi_k(0,0) \tag{92}$$

if one of the following conditions holds:

(i) $\{z^k\}$ *is componentwise bounded away from zero;*

(ii) The vectors $\nabla H_j^k,\ j=1,\ldots,m,\ \nabla(G_\theta)_i^k,\ i\in\mathcal{G}_0^k=\{i\mid z_i^k=0,\ i=1,\ldots,q\}$ *are linearly independent.*

(4) For all k*,* $(\hat{d}_s^k,Z_k^{-1}\hat{d}_z^k)/\sqrt{\rho_k}$ *are uniformly bounded.*

Proof. (1) Suppose that there is an infinite index subset $\mathcal{K}$ such that for $k\in\mathcal{K}$, $k\to\infty$, $(\nabla G_\theta^k(G_\theta^k+z^k)+\nabla H^k H^k, Z_k(G_\theta^k+z^k))\not\to 0$. Then as $(\hat{d}_s^k,\hat{d}_z^k)\to 0$,

$$\|(G_\theta^k+z^k+\nabla G_\theta^{k\top}\hat{d}_s^k+\hat{d}_z^k, H^k+\nabla H^{k\top}\hat{d}_s^k)\|\ne 0,\ \text{and are bounded}\ \forall\ k\in\mathcal{K}.$$

Since $(\hat{d}_s^k,\hat{d}_z^k)$ solves problem (80), we have $\nabla\psi(\hat{d}_s^k,\hat{d}_z^k)=0$. Hence

$$B_k\hat{d}_s^k+\rho_k\frac{\nabla G_\theta^k(G_\theta^k+z^k+\nabla G_\theta^{k\top}\hat{d}_s^k+\hat{d}_z^k)+\nabla H^k(H^k+\nabla H^{k\top}\hat{d}_s^k)}{\|(G_\theta^k+z^k+\nabla G_\theta^{k\top}\hat{d}_s^k+\hat{d}_z^k, H^k+\nabla H^{k\top}\hat{d}_s^k)\|}=0,$$
$$U_k\hat{d}_z^k+\rho_k\frac{Z_k(G_\theta^k+z^k+\nabla G_\theta^{k\top}\hat{d}_s^k+\hat{d}_z^k)}{\|(G_\theta^k+z^k+\nabla G_\theta^{k\top}\hat{d}_s^k+\hat{d}_z^k, H^k+\nabla H^{k\top}\hat{d}_s^k)\|}=0,$$

which contradicts the supposition as $(\hat{d}_s^k,\hat{d}_z^k)\to 0$. Thus, we have proved (1).

For simplicity of subsequent statements in this proof, we define

$$\tilde{B}_k=\begin{pmatrix}B_k & \\ & Z_kU_k\end{pmatrix},\ \tilde{C}_k=\begin{pmatrix}\nabla G_\theta^k & \nabla H^k\\ Z_k & \end{pmatrix},\ \tilde{c}^k=\begin{pmatrix}G_\theta^k+z^k\\ H^k\end{pmatrix},\ \tilde{d}^k=\begin{pmatrix}\hat{d}_s^k\\ Z_k^{-1}\hat{d}_z^k\end{pmatrix},$$

and $\psi_k(\tilde{d}^k)=\psi_k(\hat{d}_s^k,\hat{d}_z^k)$.

(2) The inequality $\psi_k(\hat{d}_s^k,\hat{d}_z^k)\le\psi_k(0,0)$ is obvious. By Proposition 2.2 in [32], letting

$$\bar{d}^k=-\min(1,\eta_k)\tilde{B}_k^{-1}\tilde{C}_k\tilde{c}^k,$$

where $\eta_k=(\tilde{c}^{k\top}(\tilde{C}_k^\top\tilde{B}_k^{-1}\tilde{C}_k)\tilde{c}^k)/(\tilde{c}^{k\top}(\tilde{C}_k^\top\tilde{B}_k^{-1}\tilde{C}_k)^2\tilde{c}^k)$, then we have

$$\begin{aligned}&\psi_k(\tilde{d}^k)-\psi_k(0)\le\psi_k(\bar{d}^k)-\psi_k(0)\\&\le\frac{1}{2}\left\{1-\rho_k\min[\frac{1}{\|\tilde{c}^k\|},\frac{\eta_k}{\|\tilde{c}^k\|}]\right\}\tilde{c}^{k\top}(\tilde{C}_k^\top\tilde{B}_k^{-1}\tilde{C}_k)\tilde{c}^k.\end{aligned} \tag{93}$$

If $\hat{\rho} \geq 2M_1 \geq 2\|\tilde{c}^k\|$ (M_1 is a constant), then for $\rho_k \geq \hat{\rho}$ we have $1-\rho_k/\|\tilde{c}^k\| \leq -\rho_k/(2M_1) < 0$. If $\eta_k \geq 1$, then it follows from (93), Assumption 2 (2), and (87) that

$$\psi_k(\tilde{d}^k) - \psi_k(0) \leq -[\rho_k/(4M_1)] \min\{\nu_1^{-1}, \beta_2^{-1}\mu^{-1}\}\|\tilde{C}_k\tilde{c}^k\|^2. \tag{94}$$

If $\eta_k < 1$, then by (93), we have

$$\begin{aligned}
&\psi_k(\tilde{d}^k) - \psi_k(0) \\
&\leq \frac{1}{2}\left\{1 - \rho_k\eta_k/\|\tilde{c}^k\|\right\} \tilde{c}^{k\top}(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)\tilde{c}^k \\
&= \frac{1}{2}\left\{1 - \rho_k \frac{\|(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{\frac{1}{2}}\tilde{c}^k\|^2}{\|(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)\tilde{c}^k\|^2\|\tilde{c}^k\|}\right\} \tilde{c}^{k\top}(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)\tilde{c}^k \\
&\leq \frac{1}{2}\left\{1 - \frac{\rho_k}{\|(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{\frac{1}{2}}\|^2\|\tilde{c}^k\|}\right\} \tilde{c}^{k\top}(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)\tilde{c}^k.
\end{aligned} \tag{95}$$

If $\rho_k > \hat{\rho} \geq 2M_2 \geq 2\|(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{\frac{1}{2}}\|^2\|\tilde{c}^k\|$ (M_2 is a constant), then

$$\psi_k(\tilde{d}^k) - \psi_k(0) \leq -[\rho_k/(4M_2)] \min\{\nu_1^{-1}, \beta_2^{-1}\mu^{-1}\}\|\tilde{C}_k\tilde{c}^k\|^2. \tag{96}$$

The result follows from (94) and (96) by selecting $\varsigma = \min\{1/(4M_1), 1/(4M_2)\} \cdot \min\{\nu_1^{-1}, \beta_2^{-1}\mu^{-1}\}$.

(3) Let $\tilde{d}_0^k = -\tilde{B}_k^{-1}\tilde{C}_k(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{-1}\tilde{c}^k$. Then

$$\psi_k(\tilde{d}^k) \leq \psi_k(\tilde{d}_0^k) = \frac{1}{2}\tilde{c}^{k\top}(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{-1}\tilde{c}^k. \tag{97}$$

If any one of conditions (i) and (ii) holds, then $\tilde{C}_k$ has full column rank, which implies that $\psi_k(\tilde{d}^k) \leq \varsigma_1\|\tilde{c}^k\|^2$ ($\varsigma_1 > 0$ is a constant). By Assumption 2 (2) and the boundedness of $Y_k\Lambda_k$, we have $\psi_k(\tilde{d}^k) \geq \varsigma_2\|\tilde{d}^k\|^2$ ($\varsigma_2 > 0$). Thus $\|\tilde{d}^k\| \leq \sqrt{(\varsigma_1/\varsigma_2)}\|\tilde{c}^k\|$. The result of (91) follows by letting $\varpi = \sqrt{\zeta_1/\zeta_2}$.

Under the given conditions, $\{(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{-1}\}$ is uniformly bounded, thus we have $\psi_k(\hat{d}_s^k, \hat{d}_z^k) \leq \nu\psi_k(0,0)$ for $\rho_k \geq \hat{\rho} \geq \frac{1}{2}\|(\tilde{C}_k^\top \tilde{B}_k^{-1}\tilde{C}_k)^{-1}\tilde{c}^k\|$. Thus, (92) is also valid.

(4) This result follows readily from the coerciveness of ψ_k. A detailed proof can be found from Lemma 4.10 in [32]. ∎

Remark. In practical implementations, we do not need the exact solution of problem (80). The approximate solutions which satisfy (1) – (4) of Lemma 2 can be computed very easily. We omit the details and refer the interested reader to [32].

Lemma 3. *Under Assumption 2, if $\{\rho_k\}$ is bounded, then $\{(d_s^k, d_z^k, d_u^k)\}$ and $\{v^k\}$ are bounded.*

Proof. By Lemma 2, $\psi_k(\hat{d}_s^k, \hat{d}_z^k) \le \psi_k(0,0)$, then it follows from the coerciveness of $\psi_k(\hat{d}_s^k, \hat{d}_z^k)$ that $(\hat{d}_s^k, \hat{d}_z^k)$ is bounded. The solution to the system of equations (81)-(84) can be written as

$$\begin{pmatrix} d_s^k \\ d_z^k \\ d_u^k \\ v^{k+1} \end{pmatrix} = Q_k^{-1} \begin{pmatrix} -(\nabla f_k + \nabla G_\theta^k u^k) \\ -(U_k e - \mu Z_k^{-1} e) \\ \nabla G_\theta^{k\top} \hat{d}_s^k + \hat{d}_z^k \\ \nabla H^{k\top} \hat{d}_s^k \end{pmatrix}, \tag{98}$$

as long as Q_k is invertible, where

$$Q_k = \begin{pmatrix} B_k & 0 & \nabla G_\theta^k & \nabla H^k \\ 0 & Z_k^{-1} U_k & I & 0 \\ \nabla G_\theta^{k\top} & I & 0 & 0 \\ \nabla H^{k\top} & 0 & 0 & 0 \end{pmatrix}. \tag{99}$$

We next show that Q_k is truly invertible and Q_k^{-1} is bounded. Let $Q_k^{11} = \begin{pmatrix} B_k & 0 \\ 0 & Z_k^{-1}U_k \end{pmatrix}$, $Q_k^{12} = \begin{pmatrix} \nabla G_\theta^k & \nabla H^k \\ I & 0 \end{pmatrix}$, then Q_k^{11} is positive definite, $Q_k^{11^{-1}}$ is bounded, and Q_k^{12} has full column rank by Lemma 1 and Assumption 2 (3). By doing some calculations, we have $Q_k^{-1} = \begin{pmatrix} Q_k^l & Q_k^d \\ (Q_k^d)^\top & Q_k^r \end{pmatrix}$, where $Q_k^l = Q_k^{11^{-1}} - Q_k^{11^{-1}} Q_k^{12} (Q_k^{12^\top} Q_k^{11^{-1}} Q_k^{12})^{-1} Q_k^{12^\top} Q_k^{11^{-1}}$, $Q_k^d = Q_k^{11^{-1}} Q_k^{12} (Q_k^{12^\top} Q_k^{11^{-1}} Q_k^{12})^{-1}$, $Q_k^r = -(Q_k^{12^\top} Q_k^{11^{-1}} Q_k^{12})^{-1}$. It can be seen that Q_k^{-1} exists and is bounded. The lemma follows from (98), the boundedness of Q_k^{-1} and Lemma 1 immediately. ∎

The following result shows that the algorithm converges to the KKT point of program (76)-(78) if $\{\rho_k\}$ is bounded.

Lemma 4. *Under Assumption 2, if ρ_k is bounded, then*

$$\lim_{k\to\infty} \|(d_s^k, d_z^k)\| = 0, \tag{100}$$

$$\lim_{k\to\infty} \|(G_\theta^{k+1} + z^{k+1}, H^{k+1})\| = 0, \tag{101}$$

$$\lim_{k\to\infty} \|Z_{k+1} U_{k+1} e - \mu e\| = 0, \tag{102}$$

$$\lim_{k\to\infty} \|\nabla f_{k+1} + \nabla G_\theta^{k+1} u^{k+1} + \nabla H^{k+1} v^{k+1}\| = 0. \tag{103}$$

Moreover, $\gamma_k = 1$ for all sufficiently large k.

Proof. It follows from (85), (87) and Lemma 2 (2) that

$$\pi_k(d^k; \rho_k) \le -\frac{1}{2}\nu_2 \|d_s^k\|^2 - \frac{1}{2}\beta_1 \|Z_k^{-1} d_z^k\|^2. \tag{104}$$

Since the sequence $\{\rho_k\}$ is monotonically increasing and is bounded, there exists a positive integer k_0 such that $\rho_k = \rho_{k_0}$ for all $k \ge k_0$. Thus, by (86), the sequence $\{\phi(s^k, z^k; \rho_k)\}$ is a monotonically decreasing sequence for $k \geq k_0$. Since $|\phi(s^k, z^k; \rho_k)|$ is bounded by Assumption 2 and Lemma 1 (3), $\phi(s^k, z^k; \rho_k)$ is bounded below. Hence, the limit of the sequence $\{\phi(s^k, z^k; \rho_k)\}$ exists, which by Lemma 1 (5) implies that

$$\pi_k(d^k; \rho_k) \to 0, \text{ as } k \to \infty. \tag{105}$$

Then (100) follows from (104), (105) and Lemma 1 (3).

By (100), (83) and (84), we have

$$\|(\nabla G_\theta^{k\top} \hat{d}_s^k + \hat{d}_z^k, \nabla H^{k\top} \hat{d}_s^k)\| \to 0 \text{ as } k \to \infty, \tag{106}$$

which implies that $\delta(\hat{d}_s^k, \hat{d}_z^k) \to 0$ as $k \to \infty$. Since $\psi_k(\hat{d}_s^k, \hat{d}_z^k) \leq \rho_k \|(G_\theta^k + z^k, H^k)\|$, we have

$$\frac{1}{2}\left[\nu_2 \|\hat{d}_s^k\|^2 + \beta_1 \|Z_k^{-1} \hat{d}_z^k\|^2\right] \leq \rho_k \delta(\hat{d}_s^k, \hat{d}_z^k). \tag{107}$$

Thus, $(\hat{d}_s^k, \hat{d}_z^k) \to 0$ as $k \to \infty$. It follows from Lemma 2 (1) that $\|(G_\theta^k + z^k, \nabla H^k H^k)\| \to 0$. Then (101) is obtained by (100) and Assumption 2 (3).

By (82), (100) and Lemma 1 we have $\lim_{k\to\infty} Z_k(u^k + d_u^k) - \mu e = 0$. Then by (100) again, we have

$$\lim_{k\to\infty} Z_{k+1}(u^k + d_u^k) - \mu e = 0, \tag{108}$$

which implies that $\gamma_k = 1$ in Step 5 of the algorithm. Thus (102) is derived.

The limit (103) follows from (81), (100), and the continuity of ∇f, ∇G_θ and ∇H. ∎

The following lemma addresses the case where $\{\rho_k\}$ is unbounded.

Lemma 5. *Under Assumption 2, if ρ_k is unbounded, then*

(1) $\{z^k\}$ is not componentwise bounded away from zero and there exists a convergent subsequence with $k \in \mathcal{K}$ such that $(s^k, z^k) \to (s^*, z^*)$ as $k \in \mathcal{K}$ and $k \to \infty$ with $\nabla G_{\theta i}^*,\ i \in \mathcal{G}_0^*,\ \nabla H_j^*,\ j = 1, \ldots, m$ being linearly dependent, where $\mathcal{G}_0^* = \{i \mid z_i^* = 0\}$;

(2) there is a subsequence $\{(s^k, z^k) \mid k \in \mathcal{K}\}$ such that

$$\lim_{k\in\mathcal{K}, k\to\infty} \left\| \begin{pmatrix} \nabla G_\theta^k & \nabla H^k \\ Z_k & 0 \end{pmatrix} \begin{pmatrix} G_\theta^k + z^k \\ H^k \end{pmatrix} \right\| = 0. \tag{109}$$

Proof. (1) Suppose that it is not the case. Then, for sufficiently large $k \in \mathcal{K}$, z^k is componentwise bounded away from zero, or $\nabla G_{\theta i}^k,\ i \in \mathcal{G}_0^k,\ \nabla H_j^k,\ j = 1, \ldots, m$, are linearly independent, where $\mathcal{G}_0^k = \{i \mid z_i^k = 0\}$. Thus by Lemma 2 (3), there exist constants $\varpi_i > 0$ $(i = 1, 2)$ such that

$$\|(\hat{d}_s^k, Z_k^{-1} \hat{d}_z^k)\| \leq \varpi_1 \|(G_\theta^k + z^k, H^k)\|, \tag{110}$$

$$\psi_k(\hat{d}_s^k, \hat{d}_z^k) - \psi_k(0, 0) \leq -\varpi_2 \rho_k \|(G_\theta^k + z^k, H^k)\|. \tag{111}$$

It follows from (81)-(84) that

$$\begin{aligned}&\pi_k(d^k;\rho_k)+\frac{1}{2}{d_s^k}^\top B_k d_s^k+\frac{1}{2}{d_z^k}^\top Z_k^{-1}U_k d_z^k\\ &\le \nabla f_k^\top \hat{d}_s^k-\mu e^\top Z_k^{-1}\hat{d}_z^k+\psi_k(\hat{d}_s^k,\hat{d}_z^k)-\psi_k(0,0).\end{aligned} \tag{112}$$

Thus, there exists a constant $\hat{\rho}>0$ such that $\pi_k(d^k;\rho_k)\le 0$ for $\rho_k\ge\hat{\rho}$, which, by Step 3 of Algorithm 1, implies that $\{\rho_k\}$ is bounded, a contradiction.

(2) Suppose that (109) does not hold. Then, for all sufficiently large k, there exists a constant $\chi_1>0$ such that

$$\left\|\begin{pmatrix}\nabla G_\theta^k & \nabla H^k\\ Z_k & 0\end{pmatrix}\begin{pmatrix}G_\theta^k+z^k\\ H^k\end{pmatrix}\right\|\ge\chi_1. \tag{113}$$

By Lemma 2 (2) and (4), for all $k\in\mathcal{K}$, there are positive constants χ_2 and ς such that $\|(\hat{d}_s^k, Z_k^{-1}\hat{d}_z^k)\|\le\chi_2\sqrt{\rho_k}$ and

$$\psi_k(\hat{d}_s^k,\hat{d}_z^k)-\psi_k(0,0)\le-\varsigma\chi_1^2\rho_k. \tag{114}$$

Thus, there exists a constant $\chi_3>0$ such that

$$\pi_k(\hat{d}^k;\rho_k)+\frac{1}{2}(\hat{d}_s^k)^\top B_k\hat{d}_s^k+\frac{1}{2}(\hat{d}_z^k)^\top Z_k^{-1}U_k\hat{d}_z^k\le\chi_3\sqrt{\rho_k}-\varsigma\chi_1^2\rho_k. \tag{115}$$

The inequality (115) indicates there exists a large $\hat{\rho}>0$ such that (85) holds for all $\rho_k\ge\hat{\rho}$, which is a contradiction to the assumption that $\rho_k\to\infty$. ∎

We summarize the results in the following theorem.

Theorem 1. *Under Assumption 2, suppose $\{(s^k,z^k)\}$ is an infinite sequence generated by Algorithm 1, $\{\rho_k\}$ is the penalty parameter sequence. Then one and only one of the following assertions is true:*

(A) The sequence $\{\rho_k\}$ is bounded. Then for every limit point (s^*,z^*), there exists (u^*,v^*) so that

$$\|(G_\theta^*+z^*,H^*)\|=0,\ Z^*U^*e=\mu e,\ \nabla f^*+\nabla G_\theta^* u^*+\nabla H^* v^*=0, \tag{116}$$

namely, (s^*,z^*) is a KKT point of (76)-(78).

(B) The sequence $\{\rho_k\}$ is unbounded and there is a limit point (s^*,z^*) which either satisfies that $\|((G_\theta^*)_+,H^*)\|=0$ and that $\nabla H_j^*(j=1,\dots,m)$, $\nabla G_{\theta i}^*(i\in\mathcal{I}=\{i\in\{1,\dots,q\}:G_{\theta i}^*=0\})$ are linearly dependent, or satisfies that $\|((G_\theta^*)_+,H^*)\|\ne 0$ and that

$$\nabla G_\theta^*(G_\theta^*)_+ +\nabla H^* H^*=0. \tag{117}$$

Proof. Conclusion (A) follows from Lemma 4. The first part of (B) is derived by Lemma 5 (1). Now suppose $\|((G_\theta^*)_+,H^*)\|\ne 0$, by Lemma 5 (2), $Z^*(G_\theta^*+z^*)=0$. Since $G_\theta^*+z^*\ge 0$ by (89), we have $G_\theta^*+z^*=(G_\theta^*)_+$. Hence, by Lemma 5 (2), the second part of (B) is obtained. ∎

5. The algorithm for MPEC and its global convergence

Based on the algorithm and analysis in previous sections, we now present our algorithm for MPEC and give its global convergence results.

A traditional approach is that we solve the relaxed barrier problem by letting $\mu \downarrow 0$ for each fixed θ. The process is then repeated as $\theta \downarrow 0$. For examples we can see [13, 41].

Unlike the traditional approach, our algorithm takes a shortcut to reduce μ and θ simultaneously. In particular, the barrier parameter μ is selected to be a fraction of θ (so θ is a multiple of μ). Thus, the barrier problem (76)-(78) is slightly different from its traditional counterpart in that the barrier parameter appears both in the constraints and in the objective function. All the convergence results in the last section would be still valid, however, since all those results were independent of how μ is specified.

Algorithm 2. *(The algorithm for the MPEC)*

Step 1 *Set the initial point* $(x^0, y^0, \lambda^0, z^0, u^0, v^0)$ *with* $(x^0, y^0, \lambda^0) \in \Re^{n+m+\ell}$, $z^0 \in \Re^{p+3\ell}_{++}$, $u^0 \in \Re^{p+3\ell}_{++}$ *and* $v^0 \in \Re^m$, *the initial barrier parameter* $\mu_0 > 0$, *penalty parameter* $\rho_0 > 0$, *constants* $\sigma > 0$, $\tau > 0$, $\gamma > 0$, $\kappa \in (0,1)$, *and the stopping tolerances* $\epsilon > 0$, $\epsilon' > 0$. *Let* $\theta_0 = \tau\mu_0$, $j := 0$;

Step 2 *Starting from* $(x^j, y^j, \lambda^j, z^j, u^j, v^j)$, *solve the barrier problem (76)-(78) by Algorithm 1. The Algorithm 1 is terminated when the iterate* $(x^{k_j}, y^{k_j}, \lambda^{k_j}, z^{k_j}, u^{k_j}, v^{k_j})$ *satisfies one of the following groups of conditions:*

$$(i) \quad \begin{cases} \|(G_{\theta_j}^{k_j} + z^{k_j}, H^{k_j})\| < \gamma\mu_j, \\ \|Z_{k_j} U_{k_j} e - \mu_j e\| < \gamma\mu_j, \\ \left\| \begin{pmatrix} \nabla_x f_{k_j} + \nabla_x G_{\theta_j}^{k_j} u^{k_j} + \nabla_x H^{k_j} v^{k_j} \\ \nabla_y f_{k_j} + \nabla_y G_{\theta_j}^{k_j} u^{k_j} + \nabla_y H^{k_j} v^{k_j} \\ \nabla_\lambda G_{\theta_j}^{k_j} u^{k_j} + \nabla_\lambda H^{k_j} v^{k_j} \end{pmatrix} \right\| < \gamma\mu_j; \end{cases} \tag{118}$$

$$(ii) \quad \begin{cases} \|((G_0^{k_j})_+, H^{k_j})\| \geq \gamma\epsilon, \\ \left\| \begin{pmatrix} \nabla_E G_{\theta_j}^{k_j} (G_{\theta_j}^{k_j} + z^{k_j}) + \nabla_E H^{k_j} H^{k_j} \\ Z_{k_j} (G_{\theta_j}^{k_j} + z^{k_j}) \end{pmatrix} \right\| < \epsilon; \end{cases} \tag{119}$$

$$(iii) \quad \begin{cases} \|((G_0^{k_j})_+, H^{k_j})\| < \epsilon, \\ \det\left(\begin{bmatrix} (\nabla_E (\tilde{G}^{k_j})_{\tilde{\mathcal{I}}_j})^\top \\ (\nabla_E H^{k_j})^\top \end{bmatrix} \begin{bmatrix} \nabla_E (\tilde{G}^{k_j})_{\tilde{\mathcal{I}}_j} & \nabla_E H^{k_j} \end{bmatrix} \right) \\ \qquad < \epsilon, \end{cases} \tag{120}$$

where $Z_{k_j} = diag(z^{k_j})$ *and* $U_{k_j} = diag(u^{k_j})$, $G_0^{k_j}$ *is the value of* $G_\theta^{k_j}$ *when* $\theta = 0$, $\det(\cdot)$ *is the determinant,* $\tilde{G}^{k_j} = (c^{k_j}, g^{k_j}, -\lambda^{k_j})$, $\tilde{\mathcal{I}}_j = \{i \| (\tilde{G}^{k_j})_i| < \epsilon\}$, $\nabla_E(\tilde{G}^{k_j})_{\tilde{\mathcal{I}}_j}$ *is the submatrix of* $\nabla_E(\tilde{G}^{k_j})$ *consisting of all columns indexed by* $i \in \tilde{\mathcal{I}}_j$.

Set

$$(x^{j+1}, y^{j+1}, \lambda^{j+1}) = (x^{k_j}, y^{k_j}, \lambda^{k_j}), \tag{121}$$
$$(z^{j+1}, u^{j+1}, v^{j+1}) = (z^{k_j}, u^{k_j}, v^{k_j}), \tag{122}$$

and

$$\rho_{j+1} = \max\{\rho_{k_j}, \|(u^{j+1}, v^{j+1})\| + \sigma\}. \tag{123}$$

If $\min(y^{j+1}) < \epsilon'$ *and Algorithm 1 terminates at (119) or (120), stop. If Algorithm 1 terminates at (118), go to the next step.*

Step 3 If $\mu_j < \epsilon$*, stop; else set* $\mu_{j+1} = \kappa\mu_j$, $\theta_{j+1} = \tau\mu_{j+1}$, $j := j+1$*, and go to Step 2.*

Different from the algorithm for general nonlinear programming in [32], we update the penalty parameter ρ_j by the information on multipliers, see (123), where we do not need scalar σ to be positive.

The stopping conditions (118), (119) and (120) are based on the results of last section. For any given μ_j, if Algorithm 1 has found an approximate KKT point of the relaxed barrier problem, then we proceed the algorithm to a new and smaller barrier parameter. Otherwise, by Theorem 1, if the tolerances are sufficiently small, then the algorithm will produce a sequence with a limit point (x^*, y^*, λ^*) for which one of the following assertions is true.

(1) $\|((G^*_{\theta_j})_+, H^*)\| = 0$*, and the vectors* $\nabla_E G^*_{\theta_j i}$ $(i \in \mathcal{I}_j = \{i \mid (G^*_{\theta_j})_i = 0\})$, $\nabla_E H^*_i$ $(i = 1, \ldots, m)$ *are linearly dependent. In this case,* (x^*, y^*, λ^*) *is a feasible point of NLP(*θ_j*) with* $\theta_j = \tau\mu_j$*. If* $\|((G^*_0)_+, H^*)\| = 0$*, that is,* (x^*, y^*, λ^*) *is also feasible to MPEC, then* $\mathcal{I}_j = \{i \mid \tilde{G}^*_i = 0\}$*, and* (x^*, y^*) *is a singular stationary point of MPEC since the vectors* $\nabla_E \tilde{G}^*_i$ $(i \in \{i \mid \tilde{G}^*_i = 0\})$, $\nabla_E H^*_i$ $(i = 1, \ldots, m)$ *are linearly dependent. In this case, Algorithm 2 terminates at (120); otherwise, the point* (x^*, y^*) *is an infeasible stationary point of MPEC and Algorithm 2 terminates at (119).*

(2) $\|((G^*_{\theta_j})_+, H^*)\| \neq 0$ *and* $\nabla_E G^*_{\theta_j}(G^*_{\theta_j})_+ + \nabla_E H^* H^* = 0$*, so* (x^*, y^*) *is an infeasible stationary point of MPEC. Algorithm 2 terminates at (119).*

Recall that these results require an assumption that $\nabla_E H_j(x^k, y^k, \lambda^k)$, $j = 1, \ldots, m$ are linearly independent, which is guaranteed if $F(x^k, \cdot)$ is strongly monotone and $g_j(x^k, \cdot)$, $j = 1, \ldots, \ell$ are convex for all $k \geq 0$.

To summarize, we have the following convergence results for the algorithm.

Theorem 2. *At termination, one of the two alternatives must hold.*

(A) For some μ_j*, Algorithm 2 does not proceed to Step 3. It terminates at an inner loop. Then the termination point is an approximate singular stationary point of MPEC if it is approximately feasible to the MPEC, otherwise it is an approximate infeasible stationary point.*

(B) For each μ_j*, Algorithm 2 proceeds to Step 3, the algorithm terminates at an outer loop. Then it terminates either at an approximate strong stationary point or at an approximate weak stationary point.*

Proof. (A) This result has been given by the above statements (1) and (2).

(B) We have this result since the point satisfying (118) for sufficiently small μ_j and θ_j is an approximate weak stationary point of MPEC, furthermore, if u^{k_j} and v^{k_j} are bounded, then it is also an approximate strong stationary point (Proposition 3). ∎

The case (B) of Theorem 2 can be further clarified as follows, which does not require a proof.

Theorem 3. *Assume that Algorithm 2 proceeds to Step 3 for each μ_j, $\epsilon = 0$ and an infinite sequence $\{(x^j, y^j, \lambda^j)\}$ is generated. Moreover, assume that $\{(x^j, y^j, \lambda^j)\}$ is uniformly bounded.*

(B1) If $\{\rho_j\}$ is bounded, then every limit point of $\{(x^j, y^j)\}$ is a strong stationary point of MPEC (1)-(3). If in addition the MPEC-LICQ holds at this limit point, then it is a B-stationary point of the MPEC.

(B2) If $\{\rho_j\}$ is unbounded, then every limit point of $\{(x^j, y^j)\}$ is a weak stationary point of MPEC, which may not be a strong stationary point of MPEC.

6. Numerical results

Algorithm 1 and Algorithm 2 have been coded in MATLAB. The initial parameters in Algorithm 1 are selected as $\sigma_0 = 0.1$, $\beta_1 = 0.01$, $\beta_2 = 100$, and $\xi = 0.005$. $B_0 = I$ is the identity matrix. The computation of $(\hat{d}_x^k, \hat{d}_y^k)$ is done by Algorithm 6.1 in [32]. In Algorithm 2, we select $\mu_0 = 0.1$, $\rho_0 = 1$, $\sigma = -10$, $\tau = 2$, $\kappa = 0.1$, $\gamma = 100$ and $\epsilon = 10^{-6}$, $\epsilon' = 10^{-15}$. The initial slack variables and the dual variables are given by

$$z^0 = e_{(p+3\ell)},\ u^0 = \mu_0 e_{(p+3\ell)},\ v^0 = 0, \tag{124}$$

where $e_{(p+3\ell)}$ is a $(p+3\ell)$-dimensional vector of ones. In implementing Algorithm 1, we terminate the algorithm if $\|(d_s^{k_j}, d_z^{k_j})\|$ is sufficiently small (less than $\epsilon\mu_j$) or if one of the stopping criteria (118)-(120) is met.

The approximate Hessian B_k is updated to B_{k+1} by the well-known damped BFGS update procedure.

6.1. The set of test problems in [13] and some special examples.

The numerical tests in this subsection are conducted on a COMPAQ personal computer with a Pentium-III 450MHz processor and WINDOWS98 operating system. We first applied our algorithms to the set of test problems listed in the Appendix of [13]. Some of the test problems were also used in [1,3,37,38,43] to test various algorithms developed for MPEC.

The test problem 7 has a nondifferentiable term $w = \max\{0, x_1 + x_2 + y_1 - 2y_2 - 40\}$ in the objective function, we reformulate it as a smooth problem with

$$f(x, y) = 2x_1 + 2x_2 - 3y_1 - 3y_2 + 2000w^2 - 60, \tag{125}$$

and $w \geq 0$, $w \geq x_1 + x_2 + y_1 - 2y_2 - 40$.

The initial x^0s are given by [13], but there is no information on how to select y^0 and λ^0. To our convenience we set

$$y^0 = \eta_1 e_m,\ \lambda^0 = \eta_2 e_\ell \tag{126}$$

where e_m and e_ℓ are respectively m-dimensional and ℓ-dimensional vectors of ones, and η_1 and η_2 are two constants (given in Table 1), which for most test problems are selected to be 0 or x_1^0 (the first component of x^0), depending which one gives a better numerical result.

The computational results are reported in Tables 1 and 2, in which we label the problem in the same way as in [13], for example, 1(a) represents the test problem 1 with the starting point (a), whereas 8(2) is the test problem 8 with the second group of data.

Table 1 includes the solutions and the optimal values obtained by our algorithm. Compared to Table 1 in [13], The optimal values are agreeable up to 10^{-6} with that given by [13]. In the test we also noted that the solution may not be unique for some test problems such as Problems 9 and 10 if we use different y^0 and λ^0.

We list the numbers of function evaluation (FN), gradient evaluation (GR), the number of total inner iterations (IT) in Table 1. The function evaluation includes the evaluation of the objective function and the constraint functions. Similarly, the gradient evaluation also include the evaluation of the gradients of the objective function and the constraint functions. For easy comparison with [13], the total numbers of evaluating nonlinear functions and their gradients are put in parentheses in the respective columns, where all components of $H(x, y, \lambda)$ are always treated as nonlinear functions. Note that we count the evaluation number in term of vectors while [13] counts each component separately, so the number in the parentheses equals the number out of the parentheses multiplied by the number of nonlinear functions in the corresponding problems. The evaluations on $\lambda \circ g(x, y)$ are not included because they can be derived directly. Comparing with the results of Table 2 in [13], we see that our algorithm generally requires fewer computations on the functions and their gradients except for a few problems such as problem 6.

While the relaxation parameter is linearly dependent on the barrier parameter, the barrier parameter is decreased by a fixed factor 0.1. Hence the number of outer iterations is 7 for all test problems. This way of updating the barrier parameter has been used in a number of interior-point algorithms for NLP such as Byrd, Hribar and Nocedal [7]. Some other strategies are reported in [23, 24]. Our preliminary experiment appears to show that there is no obvious impact on the number of iterations if we change from the fixed factor to a variable factor. One of the reasons may be that the solution to the MPEC is always feasible to the NLP relaxation NLP(θ) and can be reached even if θ is not sufficiently small. However, we agree with a referee on that this point may worth to be investigated in more details in future.

In Table 2, we report the optimal penalty parameter ρ^* and the residuals of first-order conditions, constraint violations and complementarity, where RD=

Table 1. Solutions and optimal values

Prob	(η_1, η_2)	x^*	f^*	IT	FN	GR
1(a)	(0,0)	4.06041	3.207700	16	17 (119)	17 (119)
(b)	(0,0)	4.06041	3.207700	15	16 (112)	16 (112)
2(a)	(0,0)	5.15361	3.449404	19	20 (140)	20 (140)
(b)	(0,0)	5.15360	3.449404	15	16 (112)	16 (112)
3(a)	(0,0)	2.38942	4.604254	16	18 (126)	17 (119)
(b)	(0,0)	2.38942	4.604254	19	20 (140)	20 (140)
4(a)	(0,0)	1.37313	6.592684	15	17 (119)	16 (112)
(b)	(10,10)	1.37313	6.592684	21	29 (203)	22 (154)
5(a)	(0,0)	(0.50018,0.50018)	-1.000000	12	13 (65)	13 (65)
6(a)	(0,5)	93.33333	-3266.666667	33	101 (202)	34 (68)
7(a)	(0,0)	(25.00125,30.00000)	4.999375	30	46 (138)	31 (93)
8(1)	(75,75)	55.55129	-343.345260	22	23 (115)	23 (115)
8(2)	(75,75)	42.53824	-203.155072	22	23 (115)	23 (115)
8(3)	(75,75)	24.14506	-68.135650	23	24 (120)	24 (120)
8(4)	(75,75)	12.37270	-19.154065	24	25 (125)	25 (125)
8(5)	(75,75)	4.75356	-3.161181	24	25 (125)	25 (125)
8(6)	(25,25)	50.00000	-346.893197	27	28 (140)	28 (140)
8(7)	(20,20)	39.79144	-224.037202	29	33 (165)	30 (150)
8(8)	(15,15)	24.25713	-80.785972	27	28 (140)	28 (140)
8(9)	(12.5,12.5)	13.01965	-22.837119	30	32 (160)	31 (155)
8(10)	(10,10)	6.00235	-5.349137	23	24 (120)	24 (120)
9(a)	(0,0)	(5.00000,9.00000)	1.640116e-12	15	17 (51)	16 (48)
9(b)	(0,0)	(5.00000,9.00000)	1.640719e-12	16	17 (51)	17 (51)
9(c)	(0,0)	(5.00000,9.00000)	1.640082e-12	17	18 (54)	18 (54)
9(d)	(0,0)	(5.00000,9.00000)	6.752681e-15	17	18 (54)	18 (54)
9(e)	(0,0)	(5.00000,9.00000)	1.845236e-14	15	16 (48)	16 (48)
10(a)	(0,5)	(7.00000,3.00000, 12.00000,18.00000)	-6600.000000	77	97 (485)	78 (390)
11(a)	(2,2)	(0.00038,2.00000)	-12.678711	24	25 (200)	25 (200)

$\|\nabla f^* + \nabla G^* u^* + \nabla H^* v^*\|$, RP= $\|(\tilde{G}^*_+, H^*)\|$ ($\tilde{G}$ is defined by (30)), RC= ${z^*}^\top u^*$ and CC= $\|\lambda^* \circ g^*\|_\infty$. These data are not reported in [13]. We include them for future reference. The results in Table 2 show that Algorithm 2 obtained approximate strong stationary points for all test problems including the problems without strict complementarity (e.g. Problem 1). As indicated in our analysis, the penalty parameter should be, and in fact it is, bounded for all those test problems.

Another observation from Table 1 is that for most test problems the number of function evaluations (FN) is not much larger than the number of iterations, which means that a full Newton step has been used in most iterations (especially for problems 1, 2, 3, 4(a), 5, 8, 9, 11), which is a condition to guarantee local superlinear convergence. The interested reader can find more details about Newton steps in [32]. In the computational test we observed that, when the parameter

Table 2. Residuals on KKT conditions

Prob	ρ^*	RD	RP	RC	CC
1(a)	1	3.91260e-06	9.43459e-13	1.40066e-06	1.08144e-07
(b)	1	3.88400e-06	9.31053e-13	1.40395e-06	1.08133e-07
2(a)	1	6.47515e-06	6.21528e-10	1.40000e-06	1.08331e-07
(b)	1	4.06954e-06	2.90570e-13	1.40004e-06	1.08346e-07
3(a)	2	2.54043e-06	7.87426e-11	1.40000e-06	1.32039e-07
(b)	2	2.45985e-06	6.17211e-11	1.40000e-06	1.32039e-07
4(a)	2	7.36826e-06	1.07658e-12	1.40000e-06	1.62779e-07
(b)	2	7.37578e-06	1.08232e-12	1.40000e-06	1.62779e-07
5(a)	1	3.21444e-07	4.71205e-14	1.06296e-06	3.62825e-11
6(a)	119.0367	9.64017e-06	1.42109e-14	4.30350e-07	6.64785e-08
7(a)	21.6355	2.94241e-06	4.94778e-15	2.39160e-06	1.04159e-07
8(1)	213.5024	5.07962e-08	3.08413e-15	2.60000e-06	1.06089e-07
8(2)	219.9075	6.38270e-07	4.94071e-13	2.60000e-06	1.04141e-07
8(3)	232.9427	3.92007e-07	4.84775e-15	2.59377e-06	1.01998e-07
8(4)	244.5515	2.77047e-07	1.72379e-15	2.60043e-06	1.00927e-07
8(5)	253.7753	8.98560e-07	1.79391e-15	2.60000e-06	1.00335e-07
8(6)	92.7429	1.66264e-10	6.35312e-15	2.60000e-06	1.92255e-07
8(7)	74.3471	4.44064e-09	2.70727e-13	2.60000e-06	1.94887e-07
8(8)	52.6502	3.36448e-11	2.41483e-14	2.60000e-06	1.88150e-07
8(9)	41.3860	5.10181e-08	3.26219e-13	2.60000e-06	1.93200e-07
8(10)	28.0840	9.71402e-06	1.34951e-13	2.60000e-06	1.58395e-07
9(a)	1	2.09588e-06	9.12282e-16	1.00001e-06	1.00012e-07
9(b)	1	2.40326e-06	3.09777e-15	1.00001e-06	1.00012e-07
9(c)	1	2.24411e-06	5.13123e-15	1.00001e-06	1.00012e-07
9(d)	2	3.31860e-07	3.66756e-15	1.00000e-06	1.00000e-07
9(e)	1	2.06686e-06	1.25382e-15	1.00000e-06	1.00000e-07
10(a)	311.7059	5.27949e-07	7.62533e-16	4.50000e-06	1.00000e-07
11(a)	2	8.59246e-07	9.58334e-08	1.47988e-06	1.44038e-07

μ is small enough, the algorithm tends to solve the relaxed barrier problem in very few iterations (not more than 3 iterations as $\mu = 10^{-7}$ for all test problems and only 1 iteration for most problems).

We then apply our algorithm to three special examples. The first example is presented by Leyffer in [29] to show that PIPA may not find a stationary point.

$$\min \ x + y \tag{127}$$
$$\text{s.t.} \ x \in [-1, 1], \tag{128}$$
$$-1 + x + \lambda = 0, \tag{129}$$
$$y \geq 0, \ \lambda \geq 0, \ y\lambda = 0. \tag{130}$$

The standard starting point is $(0, 0.02, 1)$, and the optimal solution is $(-1, 0, 2)$. Our algorithm solves it successfully after 8 iterations. FN=GR=9, $\rho^* = 1$, RD=1.67696e-06, RP=0, RC=3.65130e-07 and CC=7.52964e-09.

The second example is

$$\begin{aligned} \min \ & (x-2)^2 + y^2 && (131)\\ \text{s.t.}\ & x \geq 0, && (132)\\ & (1-x)^3 - \lambda = 0, && (133)\\ & y \geq 0,\ \lambda \geq 0,\ y\lambda = 0, && (134) \end{aligned}$$

of which the optimal point is $(1,0,0)$ and is also a singular stationary point of the problem. The initial point is $(1,1,1)$. The algorithm stops at (120) of Step 2 with residues 1.0967e-15 and 0, respectively. The solution is $(1.00000, 0.00707, 0.00000)$ and the multiplier vector is $(0.0141, -9.3213\text{e+}10,\ 9.3213\text{e+}10,\ 0.0001, 0.5000)$ after 42 iterations. The solution is an approximate singular stationary point. FN= 85, GR= 43, μ = 1.0000e-04, ρ^* = 1.4837e+11, RD= 1.08486, RP= 1.09669e-15, RC= 3.98811e-04 and CC= 7.49567e-18.

The third example is

$$\begin{aligned} \min \ & x + (y-1) && (135)\\ \text{s.t.}\ & x^2 + 1 \leq 0, && (136)\\ & x + y - \lambda = 0, && (137)\\ & y \geq 0,\ \lambda \geq 0,\ y\lambda = 0, && (138) \end{aligned}$$

which is obviously an infeasible MPEC. The point $(0,0,0)$ minimizes the ℓ_2-infeasibility of constraints. The initial point is $(-1,1,1)$. Our algorithm stops at (119) after 221 iterations with residues 1.0000 and 5.6985e-07. The solution is (9.06114e-06,−9.53364e-08,9.86731e-06), which is an approximate infeasible stationary point. FN= 323, GR= 222, μ = 1.0e-07, ρ^* = 1.3251e+20, RD= 2.40724e+15, RP= 1.00000, RC= 1.91654e-06, CC= 9.40713e-13. These results are interesting since they show that Algorithm 2 may obtain certain points with weak stationarity defined in this paper when some other methods may fail to find meaningful solutions.

6.2. The MacMPEC test problems using AMPL interface.

By hooking our MATLAB codes to AMPL, we apply Algorithm 2 to the MacMPEC test problems (see [15,30]) which are the same as [15], where *.nl* files are read by a mex file amplfunc.mexhp7*. However, we have a difficulty to use the sparse version spamfunc.c to derive the corresponding mex file. It happens that, when the problem is large, we were out of memory (partially due to the MATLAB environment). Those problems in Table 3 are marked by "-" in the corresponding columns. The last two problems are marked by "out of domain", meaning that certain iterate goes beyond the domain of some function used, so the solution process is adjourned. The details on using MATLAB with AMPL can be found in Gay's preprint [22] and in the book of Fourer et al. [17].

The computational experiments have been done on a Hewlett Packard C3600 workstation with the UNIX system. Many test problems in the former subsection are included in this test suite, it is still worthwhile to solve them again since these problems have been uniformly reformulated as MPECs in a "blackbox" format; namely we do not have the freedom to reformulate them into a form that is convenient for applying our algorithm. All starting points are standard and fixed by the suite or the default of AMPL. In this experiment, we have $\rho_0 = 10$ for all test problems if not specified.

The numerical results are given in Table 3, where n, m_i, m_e and p are the numbers of variables, inequality constraints (including bound constraints), general equality constraints, and complementarity constraints, respectively. "iter" represents the number of total inner iterations. f^* is the value of the objective function at the solution. "cc" is the maximum of residues of complementarity constraints. ρ^* is the value of the penalty parameter when the algorithm terminates. Due to memory limitation, we cannot solve some large problems with more than 1500 variables.

It can be noted from Table 3 that Algorithm 2 does well on almost all problems of relatively small sizes. According to the stopping criterion, and referring to the results given by [6,15], we have derived the approximate strong stationary points if they exist for the solved problems. It might be worthwhile to note that the algorithm has found the approximate optimal solutions for problems *ex9.2.2, qpec2, ralph1* and *scholtes4* which do not possess a strong stationary point. However, it has trouble with twelve MPECs ($\approx 9\%$ of the test problems). Six of them took 1000 iterations and the accuracy requirement was yet to meet. For the other six, from our observation, the trouble is that the direction-finding linear system (81)-(84) cannot be solved correctly. We noted that the algorithm may produce a rank-deficient coefficient matrix for the linear system. This kind of unsuccessful cases are marked "rank deficient" in the table. It has been pointed out by a referee that the optimal packaging problems *pack-comp1-*, pack-comp1c-*, pack-comp2-*, pack-comp2c-*, pack-rig1-*, pack-rig1c-*, pack-rig2-** and *pack-rig2c-** possess constraints other than the complementarity constraints which cause the problem to have no strict interior. Hence inspite of the relaxation of the complementarity constraints, NLP(θ) will still have no interior, which may be the reason of the peculiar behavior of the algorithm in solving those problems. The problems *pack-rig1p-*, pack-rig2p-** are penalty reformulations of *pack-rig1-*, pack-rig2-** respectively, which possess a strictly feasible interior, so the algorithm can solve these instances and can converge with bounded penalty parameter.

Lastly, we provide a comparison of the algorithm with FilterSQP and LOQO (see [6,15]) by showing their log scaling performance profiles (see Dolan and Moré [12]) respectively in Figure 1, where we use the 106 problems the algorithm solved successfully in less than 1000 iterations. The data for the filterSQP and LOQO are taken from [15] and [6]. The performance of the algorithms is measured by

$$\Pi_s(t) = \frac{1}{|P|}\left|\{p : \log_2(\frac{iter(s,p)}{best\_iter(p)}) \le t,\ p \in P\}\right|,$$

where P is the set of problems, $|\cdot|$ is the cardinality of a set $\cdot$, $iter(s,p)$ is the number of iterations the solver s took on problem p and $best\_iter(p)$ is

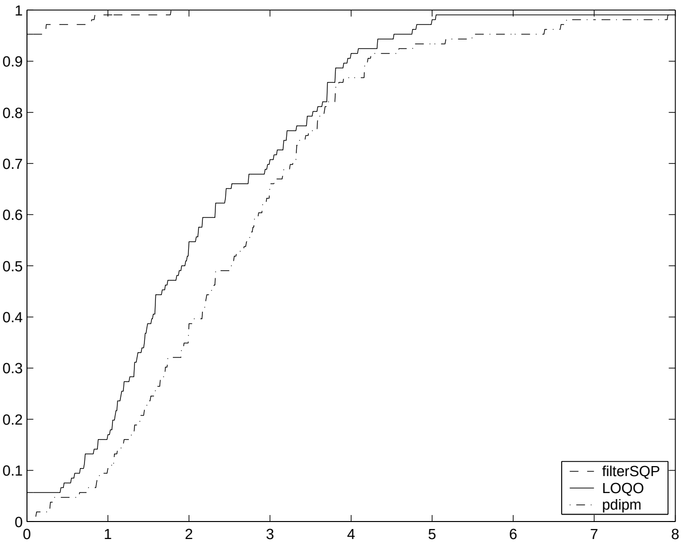

Fig. 1. Plot of the Performance Profile $\Pi_s(t)$

the smallest number of iterations any known solver took. Under this measure, the algorithm (named as pdipm in the figure) performs similarly to LOQO but inferior to filterSQP for the test problems (We have the same observation for data in [5]). However, as mentioned in Section 6.1, it appears that the algorithm has a strength in handling "irregular" problems such as the infeasible ones or the ones with dependent gradients at optimality.

Acknowledgements. The research is partially supported by Singapore-MIT Alliance and Grant RP314000-042/057-112 of National University of Singapore. The authors would like to thank Professor J.-S. Pang for proposing the research and introducing the related references during his fruitful visit to National University of Singapore. We are also grateful to the editors and two anonymous referees for their valuable comments in improving the manuscript.

References

1. Aiyoshi, E., Shimizu, K.: Hierarchical decentralized systems and its new solution by a barrier method, IEEE Trans. Syst., Man, Cybern., SMC, **11**, 444-449 (1981)
2. Anandalingam, G., Friesz, T.L., eds.: Hierarchical Optimization, Ann. Oper. Res., 1992
3. Bard, J.F.: Convex two-level optimization, Math. Program., **40**, 15-27 (1988)
4. Ben-Ayed, O., Blair, C.E.: Computational difficulties of bilevel linear programming, Oper. Res., **38**, 556-559 (1990)
5. Benson, H.Y., Sen, A., Shanno, D.F., Vanderbei, R.J.: Interior-point algorithms, penalty methods and equilibrium problems, Technical Report ORFE-03-02, Operations Research and Financial Engineering, Princeton University, 2003
6. Benson, H.Y., Shanno, D.F., Vanderbei, R.J.: Interior-point methods for nonconvex nonlinear programming: complementarity constraints, Research report ORFE-02-02, Operations Research and Financial Engineering, Princeton University, 2002
7. Byrd, R.H., Hribar, M.E., Nocedal, J.: An interior-point algorithm for large-scale nonlinear programming, SIAM J. Optim., **9**, 877-900 (1999)

Table 3. Numerical results on MacMPEC test suite

name	n	m_i	m_e	p	iter	f^*	cc	ρ^*
bar-truss-3	35	25	28	6	64	10166.57280	1.63e-07	1541.8
bard1	5	8	1	3	14	17.00000	1.33e-07	10
bard1m	6	9	1	3	14	17.00000	1.33e-07	10
bard2*	12	24	5	3	26	6598.00000	5.48e-09	311.4
bard2m	12	24	5	3	36	-6598.00000	5.79e-09	361.1
bard3	6	8	3	1	15	-12.67871	8.09e-08	10
bard3m	6	10	1	3	20	-12.67871	1.44e-07	10
bem-milanc30-s	3436	4401	1968	1464	-	-	-	-
bilevel1	10	21	2	6	18	5.00000	1.04e-07	21.6
bilevel2	16	33	4	8	45	-6600.00000	1.33e-07	348.8
bilevel3	11	13	6	3	78	-12.67871	1.58e-07	10
bilin*	8	15	0	6	30	14.60002	1.98e-07	62.6
dempe	3	2	1	1	184	28.25301	1.00e-07	3968.7
design-cent-1*	12	9	6	3	15	1.86065	1.64e-07	10
design-cent-2*	13	13	6	3	22	3.48382	1.80e-07	10
design-cent-3*	15	9	6	3	30	3.72337	1.80e-07	10
design-cent-4*	22	23	10	8	20	3.07920	1.37e-07	10
desilva	6	8	2	2	12	-1.00000	3.63e-11	10
df1	2	6	0	1	19	5.86133e-08	5.86e-08	10
ex9.1.1	13	16	7	5	25	-13.00000	1.00e-07	10
ex9.1.10	11	15	5	3	18	-3.25000	1.00e-07	10
ex9.1.2	8	11	5	2	10	-6.25000	2.11e-07	10
ex9.1.3	23	26	15	6	39	-23.00000	1.93e-07	106.7
ex9.1.4	8	10	5	2	16	-37.00000	1.00e-07	13.7
ex9.1.5	13	18	7	5	18	-1.00000	3.23e-07	20
ex9.1.6	14	20	7	6	17	-21.00000	2.99e-07	10
ex9.1.7	17	23	9	6	51	-23.00000	1.85e-07	8506.2
ex9.1.8	11	15	5	3	18	-3.25000	1.00e-07	10
ex9.1.9	12	17	6	5	14	3.11111	1.45e-07	10
ex9.2.1	10	14	5	4	20	17.00000	1.52e-07	10
ex9.2.2	9	14	4	3	33	99.99742	2.00e-07	8392.3
ex9.2.3	14	21	8	4	14	5.00000	3.37e-07	10
ex9.2.4	8	9	5	2	13	0.50000	1.41e-07	10
ex9.2.5	8	11	4	3	23	9.00000	1.85e-08	20
ex9.2.6	16	22	6	6	14	-1.00000	1.24e-07	10
ex9.2.7	10	14	5	4	20	17.00000	1.52e-07	10
ex9.2.8	6	9	3	2	9	1.50000	1.23e-07	10
ex9.2.9	9	13	5	3	15	2.00000	1.32e-07	10
gauvin	3	6	0	2	19	20.00000	1.12e-07	10

$*$: The problem is to maximize the objective function

8. Chen, Y., Florian, M.: The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions, Optim., **32**, 193-209 (1995)
9. Clark, P.A., Westerberg, A.W.: A note on the optimality conditions for the bilevel programming problem, Naval Research Logistics Quarterly, **35**, 413-418 (1988)
10. DeMiguel, A.V., Friedlander, M.P., Nogales, F.J., Scholtes, S.: An interior point method for mpecs, Technical Report, London Business School, 2003
11. Dirkse, S.P., Ferris, M.C., Meeraus, A.: Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constraint optimization, Tech. Report NA-02/11, Oxford University Computing Lab, 2002
12. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles, Math. Program., Ser. A, **91**, 201-213 (2002)
13. Facchinei, F., Jiang, H.Y., Qi, L.: A smoothing method for mathematical programs with equilibrium constraints, Math. Program., **85**, 107-134 (1999)

name	n	m_i	m_e	p	iter	f^*	cc	ρ^*
gnash10	13	26	4	8	21	-230.82321	1.07e-07	270.3
gnash11	13	26	4	8	21	-129.91192	1.05e-07	218.0
gnash12	13	26	4	8	19	-36.93311	1.02e-07	229.6
gnash13	13	26	4	8	22	-7.06178	1.01e-07	234.3
gnash14	13	26	4	8	25	-0.17905	1.00e-07	214.0
gnash15	13	26	4	8	38	-354.69906	1.89e-07	90.2
gnash16	13	26	4	8	29	-241.44198	1.69e-07	80
gnash17	13	26	4	8	33	-90.74910	1.65e-07	80
gnash18	13	26	4	8	32	-25.69822	1.93e-07	40.0
gnash19	13	26	4	8	28	-6.11671	1.70e-07	28.5
hakonsen*	9	16	3	4	1000	0.77077	7.83e-02	1160.8
hs044-i	20	36	4	10	27	15.61777	1.84e-08	10
incid-set1-16	485	818	225	225	1000	3.36636e-05	7.34e-09	10.6
incid-set1-32	1989	3162	961	961	-	-	-	-
incid-set1-8	117	222	49	49	36	6.35692e-06	8.82e-09	10
incid-set1c-16	485	833	225	225	1000	4.14802e-05	9.68e-09	12.7
incid-set1c-32	1989	3193	961	961	-	-	-	-
incid-set1c-8	117	229	49	49	28	6.35310e-06	9.27e-09	10
incid-set2-16	485	593	225	225	1000	2.89786e-02	9.07e-06	10
incid-set2-32	1989	2201	961	961	-	-	-	-
incid-set2-8	117	173	49	49	1000	4.52492e-03	8.82e-09	10
incid-set2c-16	485	608	225	225	862	3.62824e-03	1.04e-08	10
incid-set2c-32	1989	2232	961	961	-	-	-	-
incid-set2c-8	117	180	49	49	760	5.47811e-03	8.23e-09	10
jr1	2	2	0	1	10	0.50000	3.82e-09	10
jr2	2	2	0	1	13	0.50000	1.62e-07	10
kth1	2	3	0	1	8	2.99978e-07	2.00e-14	10
kth2	2	3	0	1	10	5.85787e-08	5.86e-08	10
kth3	2	3	0	1	11	0.50000	1.41e-08	10
liswet1-050	152	151	52	52	101	1.40007e-02	8.90e-08	357.0
liswet1-100	302	301	102	100	239	1.37458e-02	9.38e-08	5120
liswet1-200#	602	601	202	200	424	1.72262e-02	9.60e-07	1.64e+05
nash1	6	8	2	2	19	1.82160e-13	1.00e-07	10
outrata31	5	10	0	4	18	3.20770	1.08e-07	10
outrata32	5	10	0	4	25	3.44940	1.08e-07	10
outrata33	5	10	0	4	17	4.60425	1.32e-07	10
outrata34	5	10	0	4	15	6.59268	1.63e-07	10
pack-comp1-16	467	561	225	225	82	0.61697	1.52e-07	2.39e+06
pack-comp1-32	1955	2205	961	961	-	-	-	-
pack-comp1-8	107	147	49	49	500	0.60000	2.46e-07	1.50e+05

∗: The problem is to maximize the objective function
#: Unable to reach the accuracy

14. Falk, J.E., Liu, J.: On bilevel programming, part I: general nonlinear cases, Math. Program., **70**, 47-72 (1995)
15. Fletcher, R., Leyffer, S.: Numerical experience with solving MPECs by nonlinear programming methods, Numerical Analysis Report NA/YYY, Dept. of Math., Univ. of Dundee, UK, 2001
16. Fletcher, R., Leyffer, S., Ralph, D., Scholtes, S.: Local convergence of SQP methods for mathematical programs with equilibrium constraints, Numerical Analysis Report NA/209, Dept. of Math., Univ. of Dundee, UK, 2002
17. Fourer, R., Gay, D.M., Kernighan, B.W.: AMPL: A Modeling Language for Mathematical Programming, Second Edition, THOMSON, Brooks/cole, 2003
18. Fukushima, M., Luo, Z.-Q., Pang, J.-S.: A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints, Comput. Optim. Appl., **10**, 5-34 (1998)

name	n	m_i	m_e	p	iter	f^*	cc	ρ^*
pack-comp1c-16	467	576	225	225		rank	deficient	
pack-comp1c-32	1955	2236	961	961	-	-	-	-
pack-comp1c-8	107	154	49	49	72	0.60000	1.02e-07	4.70e+6
pack-comp2-16‡	467	561	225	225	84	0.72715	5.30e-08	1.02e+6
pack-comp2-32	1955	2205	961	961	-	-	-	-
pack-comp2-8	107	147	49	49	36	0.67312	1.58e-08	2.27e+7
pack-comp2c-16	467	576	225	225	165	0.72749	5.05e-08	6.15e+5
pack-comp2c-32	1955	2236	961	961	-	-	-	-
pack-comp2c-8	107	154	49	49	27	0.67346	1.61e-08	1.25e+5
pack-rig1-16	380	483	204	158		rank	deficient	
pack-rig1-32	1622	1987	856	708	-	-	-	-
pack-rig1-8	87	123	46	32	46	0.78794	1.33e-07	10
pack-rig1c-16#	380	498	204	158	29	0.82796	1.35e-05	7.35e+5
pack-rig1c-32	1622	2018	856	708	-	-	-	-
pack-rig1c-8	87	130	46	32	37	0.78830	1.33e-07	10
pack-rig1p-16	445	592	225	203	114	0.82604	1.33e-07	1507.5
pack-rig1p-32	1855	2348	961	861	-	-	-	-
pack-rig1p-8	105	156	49	47	65	0.78794	1.33e-07	331.0
pack-rig2-16	375	473	204	149		rank	deficient	
pack-rig2-32	1580	1903	856	661	-	-	-	-
pack-rig2-8	85	119	46	30	39	0.78041	1.33e-07	16.0
pack-rig2c-16	375	488	204	149		rank	deficient	
pack-rig2c-32	2505	1825	1665	0	-	-	-	-
pack-rig2c-8	85	126	46	30	35	0.79931	1.33e-07	21.3
pack-rig2p-16	436	574	225	194		rank	deficient	
pack-rig2p-32	1808	2254	961	814	-	-	-	-
pack-rig2p-8	103	152	49	45	53	0.78041	1.33e-07	674.3
portfl1	87	98	13	12	29	1.76749e-5	1.40e-07	10
portfl2	87	98	13	12	32	1.48735e-5	1.42e-08	10
portfl3	87	98	13	12	29	9.04711e-6	1.40e-07	10
portfl4#	87	98	13	12	31	2.47929e-6	1.36e-08	10
portfl6	87	98	13	12	32	2.71395e-6	1.36e-08	10
qpec-100-1	105	202	0	100	40	9.90017e-2	1.91e-08	20
qpec-100-2	110	202	0	100	86	-6.44521	1.62e-08	10
qpec-100-3	110	204	0	100	62	-5.48167	1.91e-08	20
qpec-100-4	120	204	0	100	50	-4.05106	1.83e-08	10
qpec-200-1#	210	404	0	200	67	-1.93494	1.99e-07	182.2
qpec-200-2	220	404	0	200	89	-24.06934	1.83e-07	31.9
qpec-200-3+	220	408	0	200	81	-1.95346	1.97e-07	55.2
qpec-200-4	240	408	0	200	72	-6.21105	1.82e-07	20
qpec1	30	40	0	20	12	80.00000	2.19e-08	10
qpec2	30	40	0	20	37	44.98211	2.00e-07	2.24e+4

#: Unable to reach the accuracy
+: The initial penalty parameter $\rho_0 = 0.1$
‡: The initial penalty parameter $\rho_0 = 1000$

19. Fukushima, M., Pang, J.-S.: Some feasibility issues in mathematical programs with equilibrium constraints, SIAM J. Optim., **8**, 673-681 (1998)
20. Fukushima, M., Pang, J.-S.: Convergence of a smoothing continuation method for mathematical programs with complementarity constraints, in Ill-posed Variational Problems and Regularization Techniques, M. Theŕa and R. Tichatschke, eds., Springer-Verlag, New York, 1999, 99-110.
21. Fukushima, M., Tseng, P.: An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints, SIAM J. Optim., **12**, 724-739 (2002)

name	n	m_i	m_e	p	iter	f^*	cc	ρ^*
ralph1	2	3	0	1	29	-4.47014e-04	2.00e-07	1280
ralph2$^+$	2	3	0	1	20	-3.18371e-08	1.61e-08	4.1
ralphmod$^\#$	104	208	0	100	100	-683.03302	1.01e-06	953.3
scholtes1	3	3	0	1	16	2.00000	2.19e-08	10
scholtes2	3	3	0	1	13	15.00000	4.17e-16	15.3
scholtes3	2	3	0	1	13	0.50000	1.41e-08	10
scholtes4	3	5	0	1	27	-8.94127e-04	2.00e-07	2560
scholtes5	3	5	0	2	12	1.00000	1.17e-07	10
sl1$^\#$	8	12	2	3	96	1.00035e-04	7.90e-09	10
stackelberg1	3	5	1	1	30	-3266.66667	1.04e-07	60.2
tap-09	86	104	32	32	1000	1708.55984	1.58e+03	1195.7
tap-15	194	260	68	83		rank	deficient	
tollmpec*	2403	4078	628	1748	-	-	-	-
tollmpec1*	2403	4078	628	1748	-	-	-	-
water-FL	213	398	116	44		out of	domain	
water-net	66	124	36	14		out of	domain	

#: Unable to reach the accuracy
+: The initial penalty parameter $\rho_0 = 0.1$
$*$: The problem is to maximize the objective function

22. Gay, D.M.: Hooking your solver to AMPL, Preprint, Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, 2000
23. Gay, D.M., Overton, M.L., Wright, M.H.: A Primal-Dual interior method for nonconvex nonlinear programming, Advances in nonlinear programming: Proceedings of the 96 International conference on nonlinear programming, Y. Yuan, eds., Kluwer Academic Publishers, 1998
24. Gould, N.I.M., Orban, D., Sartenaer, A., Toint, Ph.L.: Superlinear convergence of primal-dual interior-point algorithms for nonlinear programming, SIAM J. Optim., **11**, 974-1002 (2001)
25. Harker, P.T., Pang, J.-S.: On the existence of optimal solutions to mathematical program with equilibrium constraints, Oper. Res. Lett., **7**, 61-64 (1988)
26. Harker, P.T., Pang, J.-S.: Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications, Math. Program., Ser. B, **48**, 161-220 (1990)
27. Huang, J., Pang, J.-S.: A mathematical programming with equilibrium constraints approach to the implied volatility surface of American options, Research Report, The Johns Hopkins Univ., 2001
28. Jiang, H.Y., Ralph, D.: Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, SIAM J. Optim., **10**, 779-808 (2000)
29. Leyffer, S.: The penalty interior point method fails to converge for mathematical programs with equilibrium constraints, Numerical analysis report NA/208, Department of Math., University of Dundee, 2002
30. Leyffer, S.: MacMPEC - www-unix.mcs.anl.gov/leyffer/MacMPEC/, 2002
31. Leyffer, S.: Complementarity constraint as nonlinear equation: Theory and numerical experience, Preprint, MCS Division, Argonne National Lab, 2003
32. Liu, X.-W., Sun, J.: A robust primal-dual interior point algorithm for nonlinear programs, accepted for publication in SIAM J. Optim.
33. Liu, X.-W., Sun, J.: Global convergence analysis of line search interior point methods for nonlinear programming without regularity assumptions, accepted for publication in J. Optim. Theory Appl.
34. Liu, X.-W., Yuan, Y.-X.: A robust algorithm for optimization with general equality and inequality constraints, SIAM J. Sci. Comput., **22**, 517-534 (2000)
35. Luo, Z.-Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints, Cambridge University Press, 1996
36. Marcotte, P.: Network design problem with congestion effects: A case of bilevel programming, Math. Program., **34**, 142-162 (1986)

37. Outrata, J.V.: On optimization problems with variational inequality constraints, SIAM J. Optim., **4**, 340-357 (1994)
38. Outrata, J.V., Zowe, J.: A numerical approach to optimization problems with variational inequality constraints, Math. Program., **68**, 105-130 (1995)
39. Pang, J.-S., Fukushima, M.: Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints, Comput. Optim. Appl., **13**, 111-136 (1999)
40. Raghunathan, A.U., Biegler, L.T.: An interior point method for mathematical programs with complementarity constraints, Technical Report, Department of Chemical Engineering, Carnegie Mellon University, 2003
41. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, SIAM J. Optim., **11**, 918-936 (2001)
42. Scholtes, S., Scheel, H.: Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity, Math. Oper. Res., **25**, 1-22 (2000)
43. Scholtes, S., Stöhr, M.: Exact penalization of mathematical programs with equilibrium constraints, SIAM J. Control Optim., **37**, 617-652 (1999)
44. Scholtes, S., Stöhr, M.: How stringent is the linear independence assumption for mathematical programs with complementarity constraints, Math. Oper. Res., **26**, 851-863 (2001)
45. Vicente, L.N., Calamai, P.: Bilevel and multilevel programming: A bibliography review, J. Global Optim., **5**, 291-306 (1994)