Generalized Stationary Points and an Interior Point Method for MPEC

Xinwei Liu and Jie Sun

Abstract— Mathematical program with equilibrium constraints (MPEC) has extensive applications in practical areas such as traffic control, engineering design, and economic modeling. Some generalized stationary points of MPEC are studied to better describe the limiting 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 without assuming strict complementarity or linear independence constraint qualification. Under very general assumptions, the algorithm can always find some point with strong or weak stationarity. In particular, it is shown that every limiting point of the generated sequence is a piecewise stationary point of MPEC if the penalty parameter of the merit function is bounded. Otherwise, a certain point with weak stationarity can be obtained. Preliminary numerical results are satisfactory, which include a case analyzed by Leyffer for which the penalty interior point algorithm failed to find a stationary solution.

Index Terms— Equilibrium constraints, global convergence, interior point methods, strict complementarity, variational inequality problems

I. INTRODUCTION

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):

$$y \in \mathcal{S}(x) \Longleftrightarrow \begin{cases} g(x,y) \le 0, \\ F(x,y)^\top (z-y) \ge 0, \forall z : g(x,z) \le 0, \end{cases} \tag{4}$$

$f : \Re^{n+m} \to \Re$, $c : \Re^{n+m} \to \Re^p$, $g : \Re^{n+m} \to \Re^\ell$, and $F : \Re^{n+m} \to \Re^m$. Throughout the paper, we suppose functions f, c and F are twice continuously differentiable and function g is triply continuously differentiable. We note that if $\ell = m$ and $g(x,y) = -y$, then the PVI is reduced to a parametric nonlinear complementarity problem and in this case MPEC is specifically called the mathematical program with complementarity constraints (MPCC).

MPEC includes the bilevel programming problem (e.g., [9], [34]) as its special case and has extensive applications in practical areas such as traffic control, engineering design, and economic modeling, see [2], [18], [25], [26]. 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 [25]). As such, MPEC is known to be a class of very difficult optimization problems and can not be solved directly by the methods for standard nonlinear programming [4], [6], [7].

There have been many papers dealing with MPEC in recent years. Some of them considered the existence and stationarity of the solution of MPEC, for example [12], [17], [19], [25], [27], [29], [33], while some other papers proposed algorithms for MPEC, see [8], [11], [13], [14], [20], [25], [26], [28], [30], [32]. Upon the success of interior point methods for linear programming, the interior point approach has been extended to solve nonlinear programming (NLP) problems and MPEC. The penalty interior point algorithm (PIPA) developed by Luo, Pang and Ralph [25] 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 very recently by Leyffer [21] 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 very simple MPCC. More recently, Benson, Shanno and Vanderbei [5] applied an interior-point method for NLP to MPEC, the possible difficulties in convergence were identified and some heuristics for implementation were suggested to overcome those difficulties.

In this paper we present a new interior point method for MPEC. The method, together with its convergence theory, is an extension of a method [22], [24] developed by the authors for inequality-constrained NLP. Our original motivation for that method was to overcome some convergence difficulties arising in applying interior point methods to NLP. In the context of MPEC, we first study the relations between MPEC and its NLP relaxation. For any given relaxation parameter $\theta > 0$, the method solves a corresponding barrier problem by an inner loop algorithm. The barrier parameter is a fixed fraction of θ, so it is decreased simultaneously with θ at every outer loop

Manuscript received October 28, 2002. This work was supported in part by Singapore-MIT Alliance and Grant RP314000-026/028/042-112 of National University of Singapore.

Xinwei Liu is with Singapore-MIT Alliance, National University of Singapore, Republic of Singapore (e-mail: smaliuxw@nus.edu.sg).

Jie Sun is with the School of Business and Singapore-MIT Alliance, National University of Singapore, Republic of Singapore (phone: 6874-6448; fax: 6779-2621; e-mail: bizsunj@nus.edu.sg).

iteration. The convergence properties of our method are as follows.

1) Global convergence results are derived without requiring the MPEC-LICQ and the strict complementarity conditions.
2) Under very general conditions, the algorithm can always find some point with strong or weak stationarity. In particular, it is shown that every limiting point of the generated sequence is a piecewise stationary point of the MPEC (which is also the B-stationary point if MPEC-LICQ holds at the point), provided that the penalty parameter of the merit function is bounded. Otherwise, one of the limiting points could be a singular stationary point, an infeasible stationary point, or a weak piecewise stationary point (All of the related definitions will be given later).
3) The numerical results are satisfactory, which include the solution of the example given by Leyffer [21] and an example for which the MPEC-LICQ does not hold at the optimal point.

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 some special cares are taken to avoid that the slack variables are reduced too fast.

The paper is organized as follows. In Section II, we define some weak stationary points of MPEC that will be used in the subsequent sections. In Section III, we describe a 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 a B-stationary point of the MPEC as the relaxation parameter tends to zero. In Section IV, we present a modified primal-dual interior point method and derive convergence results for the relaxed barrier problem. In Section V, we describe our algorithm for MPEC and give its global convergence results. In Section VI, we report our preliminary numerical results on a set of problems in the literature. We also present examples, in which the algorithm converges to weak stationary points.

Some notations ought to be clarified. 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 supscript 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 $\nabla g(x,y) = [\nabla g_1(x,y)\ \cdots\ \nabla g_\ell(x,y)]$, $\nabla g_{\mathcal{J}}(x,y) = [\nabla g_j(x,y) | j \in \mathcal{J}]$, where $\mathcal{J}$ is an index set. For functions involve 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, $\mathrm{diag}\,(v)$ stands for the diagonal matrix whose diagonal is the vector v.

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

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

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

II. GENERALIZED STATIONARY PROPERTIES

We make the following blanket assumption throughout this paper.

Assumption II.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 all $x \in \{x \in \Re^n | c(x,y) \le 0$ for some $y \in \Re^m\}$ and $j = 1, \cdots, \ell$, $g_j(x, \cdot)$ is convex.

It should be noted that Assumption II.1 always holds in the important special case of MPCC. Under Assumption II.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 II.1 holds and if (x,y) is bounded, then λ is also bounded and problem (1)-(3) is equivalent to

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

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

The following definition is well known.

Definition II.2: A point $(x,y) \in \mathcal{F}$ is a B-stationary point of MPEC if

$$
\nabla_x f(x,y)^\top d_x + \nabla_y f(x,y)^\top d_y \ge 0, \tag{10}
$$

for all $(d_x, d_y) \in \mathcal{T}(x,y;\mathcal{F})$, 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 piecewise stationary point of MPEC, defined below, is often used in algorithmic design.

Definition II.3: A point $(x,y) \in \mathcal{F}$ is a piecewise stationary point of MPEC, if for $\lambda \in M(x,y)$, and for each index set $\mathcal{J} \subseteq \mathcal{G}_{00}(x,y,\lambda)$, there exist multipliers $\zeta \in \Re^p$, $\eta \in \Re^\ell$ and $\pi \in \Re^m$ such that

$$\nabla f + \nabla c \zeta + \nabla g \eta + \nabla H \pi = 0, \tag{11}$$
$$\zeta^\top c = 0, \ \zeta \geq 0, \tag{12}$$
$$\pi^\top \nabla_y g_j \geq 0, \ \text{for } j \in \mathcal{J}, \tag{13}$$
$$\pi^\top \nabla_y g_j = 0, \ \text{for } j \in \mathcal{G}_{0+}, \tag{14}$$
$$\eta_j \geq 0, \ \text{for } j \in \mathcal{G}_{00} \backslash \mathcal{J}, \tag{15}$$
$$\eta_j = 0, \ \text{for } j \notin \mathcal{G}_0, \tag{16}$$

where we omit the variables (x,y) and (x,y,λ) for simplicity.

Definition II.4: For any $(x,y) \in \mathcal{F}$ and $\lambda \in M(x,y)$, the MPEC-LICQ holds at (x,y) if

$$\begin{pmatrix} \nabla H & \nabla c_{\mathcal{C}_0} & \nabla g_{\mathcal{G}_0} & 0 \\ \nabla_\lambda H & 0 & 0 & [e_j, j \in \mathcal{G}_0(\lambda)] \end{pmatrix} \tag{17}$$

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

Then we have the next result.

Proposition II.5: 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 piecewise stationary point of MPEC.

This proposition can be derived in a similar way to the derivation of Theorem 3.3.4 in [25], where the result has been proved under a more general setting. Similar results are reported in [29], [31].

To describe convergence results of our algorithm, we need various stationary properties in weaker sense.

Definition II.6:

(1) A point $(x,y) \in \mathcal{F}$ is called a weak piecewise stationary point of MPEC if there exist $\zeta \in \Re^p$, $\eta \in \Re^\ell$, and $\pi \in \Re^m$ such that (11)-(12), (14), and (16) hold.

(2) 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).

(3) 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

$$\min_{(x,y,\lambda)} \{\|c_+\|^2 + \|H\|^2 + \|g_+\|^2 + \|\lambda_-\|^2 + \|(\lambda \circ g + \theta e)_-\|^2\}, \tag{18}$$

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

$$\nabla c \ c_+ + \nabla H \ H + \nabla g \ g_+ + \nabla g \ \Lambda(\lambda \circ g + \theta e)_- = 0, \tag{19}$$
$$\nabla_y g^\top H + \lambda_- + \text{diag}\,(g)(\lambda \circ g + \theta e)_- = 0, \tag{20}$$

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

The optimal value of (18) is an ℓ_2 measure of the total infeasibility of problem (6)-(9). If (x,y,λ) is a feasible point, then for any $\theta \geq 0$, this measure is zero.

In general, a weak piecewise stationary point may not be a piecewise stationary point since (13) or (15) may not hold. However, it is easy to see that, under strict complementarity, the two concepts are identical since (13) and (15) are vacuous.

III. A RELAXATION SCHEME FOR MPEC

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

$$\min \ f(x,y) \tag{21}$$
$$\text{s.t.} \ c(x,y) \leq 0, \tag{22}$$
$$H(x,y,\lambda) = 0, \tag{23}$$
$$\lambda \geq 0, \ g(x,y) \leq 0, \ -\lambda \circ g(x,y) \leq \theta e, \tag{24}$$

where the complementarity constraints in the reformulated MPEC (6)-(9) are relaxed by inequalities.

It is obvious that if $\theta = 0$ then (21)-(24) reduces to (6)-(9). The following result shows the relationship between the MPEC-LICQ and the LICQ for the θ-relaxation in the usual sense of nonlinear programming (LICQ for NLP(θ) for short).

Proposition III.1: 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 every sufficiently small $\theta > 0$, the LICQ for NLP(θ) holds for every feasible point $(\bar{x},\bar{y},\bar{\lambda}) \in \mathcal{N}$.

To simplify the notation, we set

$$\tilde{G}(x,y,\lambda) = (c(x,y), \ g(x,y), \ -\lambda) \text{ and}$$
$$G_\theta(x,y,\lambda) = (\tilde{G}(x,y,\lambda), \ -\lambda \circ g(x,y) - \theta e). \tag{25}$$

Then

$$\nabla G_\theta(x,y,\lambda) = [\nabla c(x,y) \ \nabla g(x,y) \ 0 \ -[\nabla g(x,y)]\lambda], \tag{26}$$
$$\nabla_\lambda G_\theta(x,y,\lambda) = [0 \ 0 \ -I \ -\text{diag}\,(g(x,y))], \tag{27}$$

where I is the $\ell \times \ell$ identity matrix. The constraints of NLP(θ) can be written as $G_\theta(x,y,\lambda) \leq 0$, $H(x,y,\lambda) = 0$. The Lagrange function of program (21)-(24) is

$$L_\theta(x,y,\lambda,u,v) = f(x,y) + u^\top G_\theta(x,y,\lambda) + v^\top H(x,y,\lambda), \tag{28}$$

where $u \in \Re_+^{p+3\ell}$ and $v \in \Re^m$ are the multipliers. Let $\bar{u} = (u_1, \cdots, u_p)$, $\hat{u} = (u_{p+1}, \cdots, u_{p+\ell})$ and $\tilde{u} = (u_{p+2\ell+1}, \cdots, u_{p+3\ell})$. Now we show that any KKT point of NLP(θ) converges to a piecewise stationary point of MPEC if the primal and dual variables are bounded.

Proposition III.2: 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 limiting points, then (x^*,y^*) is a piecewise stationary point of the MPEC (1)-(3).

Before presenting our next result, we need the following definition:

Definition III.3: 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 [13], 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.

We have the following sufficient conditions for the dual boundedness required by Proposition III.2.

Proposition III.4: Suppose that $\{(\bar{x}, \bar{y}, \bar{\lambda})\}$ is bounded as $\theta \to 0$, Θ is an infinite set of θs 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 the second order necessary optimality condition of NLP(θ) holds at $(\bar{x}, \bar{y}, \bar{\lambda})$ for $\theta \in \Theta$, $(\bar{x}, \bar{y}, \bar{\lambda}) \mid \theta \in \Theta$ is asymptotically weakly nondegenerate, and the MPEC-LICQ holds at (x^*, y^*, λ^*).

IV. THE RELAXED BARRIER PROBLEM

We note that applying interior point approach to problem (6)-(9) directly will result in a conflict. Thus, we 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:

$$\begin{aligned} \min \quad & f(x,y) - \sum_{i=1}^{p} \mu \ln \bar{z}_i - \sum_{j=1}^{\ell} \mu \ln \hat{z}_j \\ & - \sum_{j=1}^{\ell} \mu \ln \hat{\lambda}_j - \sum_{j=1}^{\ell} \mu \ln \tilde{z}_j \qquad (29) \\ \text{s.t.} \quad & c(x,y) + \bar{z} = 0, \qquad (30) \\ & H(x,y,\lambda) = 0, \qquad (31) \\ & g(x,y) + \hat{z} = 0, \qquad (32) \\ & -\lambda + \hat{\lambda} = 0, \qquad (33) \\ & -\lambda \circ g(x,y) + \tilde{z} = \theta e. \qquad (34) \end{aligned}$$

By using (25), (29)-(34) can simply be written as

$$\begin{aligned} \min \quad & f(s) - \sum_{i=1}^{q} \mu \ln z_i \qquad (35) \\ \text{s.t.} \quad & G_\theta(s) + z = 0, \qquad (36) \\ & H(s) = 0, \qquad (37) \end{aligned}$$

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 (35)-(37) for fixed μ and derive global convergence results of the algorithm. The algorithm for MPEC is then presented in Section V, which takes the algorithm in this section as the inner loop and decreases μ in the outer loop.

A. THE ALGORITHM FOR PROBLEM (35)-(37)

Define the merit function with ℓ_2 penalty

$$\phi(s, z; \rho) = f(s) - \sum_{i=1}^{q} \mu \ln z_i + \rho \|(G_\theta(s) + z, H(s))\|, \quad (38)$$

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 (36) and (37), 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 $(\hat{d}^k_s, \hat{d}^k_z)$ is an approximate solution of program

$$\begin{aligned} \min \quad \psi_k(d_s, d_z) = {} & \tfrac{1}{2}(d_s^\top B_k d_s + d_z^\top Z_k^{-1} U_k d_z) + \rho_k \|(G^k_\theta + \\ & z^k + \nabla G^{k\top}_\theta d_s + d_z, H^k + \nabla H^{k\top} d_s)\| \qquad (39) \end{aligned}$$

such that some prescribed conditions (see the next subsection) hold. Then we compute the search direction $(d^k_s, d^k_z, d^k_u, d^k_v)$ by solving the modified primal-dual system of equations

$$\begin{aligned} B_k d_s + \nabla G^k_\theta d_u + \nabla H^k d_v &= \\ & -(\nabla f_k + \nabla G^k_\theta \, u^k + \nabla H^k v^k), \quad (40) \\ U_k d_z + Z_k d_u &= -(Z_k U_k e - \mu e), \quad (41) \\ \nabla G^{k\top}_\theta d_s + d_z &= \nabla G^{k\top}_\theta \hat{d}^k_s + \hat{d}^k_z, \quad (42) \\ \nabla H^{k\top} d_s &= \nabla H^{k\top} \hat{d}^k_s. \quad (43) \end{aligned}$$

Note that the right-hand-sides of (42) and (43) are different from the traditional interior point approach. For motivation of this modification the reader is referred to [23], [24].

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

Algorithm IV.1: (The algorithm for problem (35)-(37))

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

Step 2 Calculate the primal search direction (d^k_s, d^k_z) and the dual direction (d^k_u, d^k_v) by the primal-dual system of equations (40)-(43), where $(\hat{d}^k_s, \hat{d}^k_z)$ is derived by approximately minimizing (39);

Step 3 Let

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

where $d^k = (d^k_s, d^k_z)$ and

$$\begin{aligned} \delta(d^k_s, d^k_z) = {} & \|(G^k_\theta + z^k, H^k)\| - \|(G^k_\theta + z^k \\ & + \nabla G^{k\top}_\theta d^k_s + d^k_z, H^k + \nabla H^{k\top} d^k_s)\|. \end{aligned}$$

If

$$\pi_k(d^k; \rho_k) \le -\frac{1}{2} d^{k\top}_s B_k d^k_s - \frac{1}{2} d^{k\top}_z Z_k^{-1} U_k d^k_z, \quad (44)$$

let $\rho_{k+1} = \rho_k$; Otherwise, we replace ρ_k by a larger ρ_{k+1} (for example $\rho_{k+1} \geq 2\rho_k$) such that (44) holds;

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

$$\begin{aligned} &\phi(s^k + \sigma\hat{\alpha}_k d_s^k, z^k + \sigma\hat{\alpha}_k d_z^k; \rho_{k+1}) - \phi(s^k, z^k; \rho_{k+1}) \\ &\qquad \leq \sigma_0 \sigma \hat{\alpha}_k \pi_k(d^k; \rho_{k+1}), \end{aligned} \tag{45}$$

$$\begin{aligned} \beta_1 \mu e \leq (U_k + \gamma_k D_u^k) \max\{&z^k + \sigma\hat{\alpha}_k d_z^k, \\ &-G_\theta(s^k + \sigma\hat{\alpha}_k d_s^k)\} \leq \beta_2 \mu e, \end{aligned} \tag{46}$$

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

$$s^{k+1} = s^k + \alpha_k d_s^k, \tag{47}$$

$$z^{k+1} = \max\{z^k + \alpha_k d_z^k, -G_\theta^{k+1}\}, \tag{48}$$

and the new dual iterate is generated by

$$u^{k+1} = u^k + \gamma_k d_u^k, \ v^{k+1} = v^k + d_v^k; \tag{49}$$

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 IV.1 is only taken as an inner loop of our algorithm for MPEC, we will give the stopping criterion in the algorithm for MPEC.

B. CONVERGENCE OF ALGORITHM IV.1

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

Assumption IV.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 3.2, Proposition 3.3, Lemma 3.5 in [23] and Lemma 4.2 in [22].

Lemma IV.3: Under Assumption IV.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 IV.4: Under Assumption IV.2, if $(\hat{d}_s^k, \hat{d}_z^k)$ solves problem (39) 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 $\psi_k(\hat{d}_s^k, \hat{d}_z^k) \leq \psi_k(0,0)$, and there exist constants $\hat{\rho} > 0$ and $\varsigma > 0$ so that $\forall\ \rho_k \geq \hat{\rho}$,

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

(3) There exist $\nu \in (0,1)$, $\hat{\rho} > 0$ and $\varpi > 0$ so that $\forall\ \rho_k \geq \hat{\rho}$, $\|(\hat{d}_s^k, Z_k^{-1}\hat{d}_z^k)\| \leq \varpi \|(G_\theta^k + z^k, H^k)\|$ and $\psi_k(\hat{d}_s^k, \hat{d}_z^k) \leq \nu\psi_k(0,0)$ if one of the following conditions holds:

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

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

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

Remark. In practical implementations, we do not need the exact solution of problem (39). The approximate solutions which satisfy (1) – (4) can be computed very easily. We omit the details here and refer the interested reader to [22].

Lemma IV.5: Under Assumption IV.2, if $\{\rho_k\}$ is bounded, then $\{(d_s^k, d_z^k, d_u^k)\}$ and $\{v^k\}$ are bounded.

The following result shows that the algorithm converges to the KKT point of program (35)-(37) if $\{\rho_k\}$ is bounded.

Lemma IV.6: Under Assumption IV.2, if ρ_k is bounded, then

$$\lim_{k\to\infty} \|(d_s^k, d_z^k)\| = 0, \tag{50}$$

$$\lim_{k\to\infty} \|(G_\theta^{k+1} + z^{k+1}, H^{k+1})\| = 0, \tag{51}$$

$$\lim_{k\to\infty} \|Z_{k+1}U_{k+1}e - \mu e\| = 0, \tag{52}$$

$$\lim_{k\to\infty} \|\nabla f_{k+1} + \nabla G_\theta^{k+1} u^{k+1} + \nabla H^{k+1} v^{k+1}\| = 0. \tag{53}$$

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

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

Lemma IV.7: Under Assumption IV.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 $\nabla G_{\theta i}^*$, $i \in \mathcal{G}_0^*$, ∇H_j^*, $j = 1, \cdots, m$ 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

$$\lim_{k\in\mathcal{K}, k\to\infty} \left\| \begin{pmatrix} \nabla G_\theta^k & \nabla H^k \\ Z_k & 0 \end{pmatrix} \begin{pmatrix} G_\theta^k + z^k \\ H^k \end{pmatrix} \right\| = 0. \tag{54}$$

We summarize the results in the following theorem.

Theorem IV.8: Under Assumption IV.2, suppose $\{(s^k, z^k)\}$ is an infinite sequence generated by Algorithm IV.1, $\{\rho_k\}$ is the penalty parameter sequence. Then we have one of the following results:

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

$$\|(G_\theta^* + z^*, H^*)\| = 0,\ Z^*U^*e = \mu e, \tag{55}$$

$$\nabla f^* + \nabla G_\theta^* u^* + \nabla H^* v^* = 0, \tag{56}$$

namely, (s^*, z^*) is a KKT point of (35)-(37).
(2) The sequence $\{\rho_k\}$ is unbounded and there is a limiting point (s^*, z^*) which either satisfies that $\|((G^*_\theta)_+, H^*)\| = 0$ and that $\nabla H^*_j (j = 1, \cdots, m)$, $\nabla G^*_{\theta i} (i \in \mathcal{I} = \{i \in \{1, \cdots, q\} : G^*_{\theta i} = 0\})$ are linearly dependent, or satisfies that $\|((G^*_\theta)_+, H^*)\| \neq 0$ and that

$$\nabla G^*_\theta (G^*_\theta)_+ + \nabla H^* H^* = 0. \tag{57}$$

V. The Algorithm for MPEC and its Global Convergence

Based on the algorithm and analysis in last 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 [8], [30].

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 (35)-(37) 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 V.1: (The algorithm for the MPEC)

Step 1 Given 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 $\mu_0 > 0$ and penalty parameter $\rho_0 > 0$, scalar σ, constants $\tau > 0$, $\zeta > 0$, $\kappa \in (0, 1)$, the stopping tolerances $\epsilon > 0$, $\epsilon_1 > \epsilon_2 > 0$. Let $\theta_0 = \tau\mu_0$, $(\underline{x}^0, \underline{y}^0, \underline{\lambda}^0, \underline{z}^0, \underline{u}^0, \underline{v}^0) = (x^0, y^0, \lambda^0, z^0, u^0, v^0)$, $j := 0$;

Step 2 With using $(\underline{x}^j, \underline{y}^j, \underline{\lambda}^j, \underline{z}^j, \underline{u}^j, \underline{v}^j)$ as the starting point, solve the barrier problem (35)-(37) by Algorithm IV.1. The Algorithm IV.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:

$$(i) \quad \begin{cases} \|(G^{k_j}_{\theta_j} + z^{k_j}, H^{k_j})\| < \zeta\mu_j, \\ \|Z_{k_j} U_{k_j} e - \mu_j e\| < \zeta\mu_j, \\ \left\| \begin{pmatrix} \nabla_x f_{k_j} + \nabla_x G^{k_j}_{\theta_j} u^{k_j} + \nabla_x H^{k_j} v^{k_j} \\ \nabla_y f_{k_j} + \nabla_y G^{k_j}_{\theta_j} u^{k_j} + \nabla_y H^{k_j} v^{k_j} \\ \nabla_\lambda G^{k_j}_{\theta_j} u^{k_j} + \nabla_\lambda H^{k_j} v^{k_j} \end{pmatrix} \right\| \\ \qquad < \zeta\mu_j; \end{cases} \tag{58}$$

$$(ii) \quad \begin{cases} \|((G^{k_j}_0)_+, H^{k_j})\| \geq \epsilon_1, \\ \|(\nabla_E G^{k_j}_{\theta_j}(G^{k_j}_{\theta_j} + z^{k_j}) + \nabla_E H^{k_j} H^{k_j}, \\ \qquad Z_{k_j}(G^{k_j}_{\theta_j} + z^{k_j}))\| < \epsilon_2; \end{cases} \tag{59}$$

$$(iii) \quad \begin{cases} \|((G^{k_j}_0)_+, H^{k_j})\| < \epsilon_2, \\ \det(\left[\nabla_E (\tilde{G}^{k_j})_{\tilde{\mathcal{I}}_j} \ \nabla_E H^{k_j}\right]^\top \\ \qquad \left[\nabla_E (\tilde{G}^{k_j})_{\tilde{\mathcal{I}}_j} \ \nabla_E H^{k_j}\right]) < \epsilon_2, \end{cases} \tag{60}$$

where $Z_{k_j} = \text{diag}\,(z^{k_j})$ and $U_{k_j} = \text{diag}\,(u^{k_j})$, $G^{k_j}_0$ is the value of $G^{k_j}_\theta$ when $\theta = 0$, $\det(\cdot)$ is the determinant, $\tilde{G}^{k_j} = (c^{k_j}, g^{k_j}, -\lambda^{k_j})$, $\tilde{\mathcal{I}}_j = \{i | |(\tilde{G}^{k_j})_i| \leq \epsilon_2\}$, $\nabla_E(\tilde{G}^{k_j})_{\tilde{\mathcal{I}}_j}$ is the submatrix of $\nabla_E(\tilde{G}^{k_j})$ consisting of all columns indexed by $i \in \tilde{\mathcal{I}}_j$.

Set

$$(x^{j+1}, y^{j+1}, \lambda^{j+1}) = (x^{k_j}, y^{k_j}, \lambda^{k_j}), \tag{61}$$
$$(z^{j+1}, u^{j+1}, v^{j+1}) = (z^{k_j}, u^{k_j}, v^{k_j}) \tag{62}$$

and

$$\rho_{j+1} = \max\{\rho_{k_j}, \|(u^{j+1}, v^{j+1})\| + \sigma\}. \tag{63}$$

If Algorithm IV.1 terminates at (59) or (60), stop; If Algorithm IV.1 terminates at (58), go to the next step.

Step 3 If $\mu_j < \epsilon$, stop; Else calculate an approximate solution of (39) and then derive $(d^{k_j}_x, d^{k_j}_y, d^{k_j}_\lambda, d^{k_j}_z, d^{k_j}_u, d^{k_j}_v)$ by solving equations (40)-(43). Let

$$(\underline{x}^{j+1}, \underline{y}^{j+1}, \underline{\lambda}^{j+1}) = \begin{cases} (x^{k_j} + d^{k_j}_x, y^{k_j} + d^{k_j}_y, \lambda^{k_j} + d^{k_j}_\lambda), \\ \qquad \text{if } z^{k_j} + d^{k_j}_z \geq 0.2\xi z^{k_j} \\ (x^{k_j}, y^{k_j}, \lambda^{k_j}), \qquad \text{otherwise}, \end{cases} \tag{64}$$

$$(\underline{z}^{j+1}, \underline{u}^{j+1}, \underline{v}^{j+1}) = \begin{cases} (z^{k_j} + d^{k_j}_z, u^{k_j} + d^{k_j}_u, v^{k_j} + d^{k_j}_v), \\ \qquad \text{if } z^{k_j} + d^{k_j}_z \geq 0.2\xi z^{k_j} \\ (z^{k_j}, u^{k_j}, v^{k_j}), \qquad \text{otherwise}, \end{cases} \tag{65}$$

$\mu_{j+1} = \kappa\mu_j$, $\theta_j = \tau\mu_j$, $j := j + 1$ and go to Step 2.

Different from the algorithm for general nonlinear programming in [22], we update the penalty parameter ρ_j by the information on multipliers, see (63), where we do not need scalar σ to be positive.

It has been noted [15], [16] that the starting point for the new outer iteration should be selected carefully so that the unit steplength is accepted as the barrier is small. Based on our numerical experience, we take some strategy similar to [16] in Step 3 of the algorithm, which seems to have improved the performance.

The stopping conditions (58), (59) and (60) are based on the results of last section. Recall that these results require an assumption that $\nabla_E H_j(x^k, y^k, \lambda^k)$, $j = 1, \cdots, m$ are linearly independent, which is guaranteed if $F(x^k, \cdot)$ is strongly monotone and $g_j(x^k, \cdot)$, $j = 1, \cdots, \ell$ are convex for all $k \geq 0$. We have the following convergence results for the algorithm.

Theorem V.2: At termination, there are two possible results of Algorithm V.1.

(1) For some μ_j, Algorithm V.1 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.

(2) For each μ_j, Algorithm V.1 proceeds to Step 3, the algorithm terminates at an outer loop. Then it terminates at an approximate piecewise stationary point or an approximate weak piecewise stationary point of MPEC.

The following theorem further explains the case (2) of Theorem V.2, which does not require a proof.

Theorem V.3: Assume that Algorithm V.1 proceeds to Step 3 for each μ_j, $\epsilon = 0$ and an infinite sequence $\{(x^j, y^j, \lambda^j)\}$ is generated, moreover, $\{(x^j, y^j, \lambda^j)\}$ is uniformly bounded. The sequence $\{\rho_j\}$ is the penalty parameter sequence.

(1) If $\{\rho_j\}$ is bounded, then every limiting point of $\{(x^j, y^j)\}$ is a piecewise stationary point of MPEC (1)-(3). Moreover, if the MPEC-LICQ holds at this limiting point, then it is a B-stationary point of the MPEC.

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

VI. NUMERICAL RESULTS

Algorithm V.1 has been programmed in MATLAB 6.1 and implemented on a COMPAQ personal computer with a Pentium-III CPU and WINDOWS98 operating system. The computation of $(\hat{d}_x^k, \hat{d}_y^k)$ in Algorithm IV.1 is similar to Algorithm 6.1 in [22], where we select $\nu = 0.98$.

The initial parameters in Algorithm IV.1 are selected as $\mu_0 = 0.1$, $\rho_0 = 1$, $\sigma_0 = 0.1$, $\beta_1 = 0.01$ and $\beta_2 = 10$, $\xi = 0.005$. $B_0 = I$ is the identity matrix. For Algorithm V.1, we select $\sigma = -10$, $\tau = 2$, $\kappa = 0.01$, $\zeta = 100$ and $\epsilon = 10^{-6}$, $\epsilon_1 = 10\epsilon$ and $\epsilon_2 = 10^{-5}\epsilon$.

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

We first applied our algorithms to the set of test problems listed in the Appendix of [8]. All of them have been solved by [8]. Some of them were also used respectively by some other works [1], [3], [27], [28], [32] to test their new algorithms developed for MPEC.

For test problem 7, we let $w = \max\{0, x_1 + x_2 + y_1 - 2y_2 - 40\}$, then

$$f(x, y) = 2x_1 + 2x_2 - 3y_1 - 3y_2 + Rw^2 - 60, \qquad (66)$$

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

The initial x^0s are given by [8], but there is no information on selecting y^0 and λ^0. We set all components of y^0 and λ^0 as the same as the first component of x^0, that is,

$$y^0 = x_1^0 e_m, \ \lambda^0 = x_1^0 e_\ell \qquad (67)$$

where e_m and e_ℓ are respectively m-dimensional and ℓ-dimensional vectors of ones. Let $\omega = \max\{1, -0.5 \min(G_i^0 \mid i = 1, \cdots, m)\}$, the initial slack variables and the dual variables are given by

$$z^0 = \omega e_{(p+3\ell)}, \ u^0 = (\mu_0/\omega) e_{(p+3\ell)}, \ v^0 = 0. \qquad (68)$$

The numerical results are reported in Tables 1, 2 and 3, in which we label the problem as the same in [8], 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. Solutions and optimal values

Prob	x^*	f^*	ρ^*
1(a)	4.06041	3.207700	1
(b)	4.06041	3.207700	1
2(a)	5.15360	3.449404	1
(b)	5.15360	3.449404	1
3(a)	2.38942	4.604254	2
(b)	2.38942	4.604254	2
4(a)	1.37313	6.592684	2
(b)	1.37313	6.592684	2
5(a)	(0.50050,0.50050)	-0.999999	1
6(a)	93.33333	-3266.666667	61.5225
7(a)	(25.00125,30.00000)	4.999375	21.6345
8(1)	55.55129	-343.345260	191.9106
8(2)	42.53825	-203.155072	169.0610
8(3)	24.14506	-68.135650	212.2190
8(4)	12.37270	-19.154065	243.1791
8(5)	4.75356	-3.161181	253.8217
8(6)	50.00000	-346.893197	92.7425
8(7)	39.79144	-224.037202	74.3444
8(8)	24.25713	-80.785972	52.6498
8(9)	13.01965	-22.837119	41.3739
8(10)	6.00234	-5.349137	28.0686
9(a)	(5.00000,9.00000)	4.220791e-14	1
9(b)	(9.00416,5.99582)	3.127349e-10	1
9(c)	(9.89722,5.10278)	3.382036e-15	2
9(d)	(9.95627,5.04373)	9.460931e-15	2
9(e)	(5.00000,9.00000)	3.201918e-14	1
10(a)	(7.52468,3.78702, 11.47532,17.21298)	-6600.000000	310.9728
11(a)	(0.00045,2.00000)	-12.678711	2

In Table 1, we give the solutions and the optimal values obtained by our algorithm. Referring to Table 1 in [8], we notice that we have obtained the approximate optimal solutions for all test problems since they have the same optimal objective function values as given in [8]. However, the optimal solution are different for some test problems such as Problems 9 and 10. This difference is partially caused by the selection of the initial y^0 since the piecewise stationary point for some test problems may not be unique. For problems 9(b), 9(c), and 9(d), if we select $(y^0, \lambda^0) = -2x_1^0 e_{m+\ell}$, then $x^* = (5.00000, 9.00000)$ and the optimal objective values are $1.974146e$-14, $9.151050e$-17 and $2.495853e$-14, respectively. If $y^0 = 0$, $\lambda^0 = 5e_\ell$, then we have $x^* = (7.00001, 3.00001, 11.99999, 17.99999)$ for problem 10. We also report the optimal penalty parameters ρ^* in Table 1, which indicate the penalty parameters are bounded for all test problems.

In Table 2, we report the residuals of first-order conditions, constraint violations and complementarity, where RD= $\|\nabla f^* + \nabla G^* u^* + \nabla H^* v^*\|$, RP= $\|(\tilde{G}_+^*, H^*)\|$ ($\tilde{G}$ is defined by (25)), RC= ${z^*}^\top u^*$ and CC= $\|\lambda^* \circ g^*\|_\infty$. These data are not reported in [8]. We include them for future reference. The results in this table show that our algorithm obtained the approximate piecewise stationary points for all test problems including the problems without strict complementarity (for example Problem 1).

Table 2. Residuals on KKT conditions

Prob	RD	RP	RC	CC
1(a)	1.1842e-08	1.9222e-15	1.4000e-06	1.0814e-07
(b)	1.1726e-08	1.4856e-15	1.4000e-06	1.0814e-07
2(a)	4.0583e-06	3.7734e-13	1.4000e-06	1.0833e-07
(b)	3.8850e-06	3.5449e-13	1.4000e-06	1.0833e-07
3(a)	1.0552e-08	4.4187e-13	1.4000e-06	1.3204e-07
(b)	1.0624e-08	4.3863e-13	1.4000e-06	1.3204e-07
4(a)	1.1526e-09	7.3360e-12	1.4000e-06	1.6278e-07
(b)	1.4187e-09	6.5811e-12	1.4000e-06	1.6278e-07
5(a)	5.3756e-06	9.1420e-12	2.8191e-06	1.0063e-10
6(a)	1.9169e-07	0	5.0122e-07	6.2420e-08
7(a)	2.1770e-10	4.4848e-15	2.3964e-06	1.0416e-05
8(1)	6.0741e-07	2.1306e-14	2.4995e-06	1.0609e-07
8(2)	8.5715e-07	2.2295e-14	2.3634e-06	1.0421e-07
8(3)	1.1506e-08	2.2489e-15	2.6000e-06	1.0200e-07
8(4)	1.8626e-06	3.0479e-14	2.5939e-06	1.0093e-07
8(5)	2.5793e-06	5.8598e-14	2.6001e-06	1.0034e-07
8(6)	2.0148e-08	6.5607e-15	2.6000e-06	1.9226e-07
8(7)	5.3644e-07	3.4212e-11	2.6000e-06	1.9489e-07
8(8)	2.5263e-08	1.0001e-12	2.6000e-06	1.8815e-07
8(9)	5.7067e-06	3.9319e-11	2.6000e-06	1.9320e-07
8(10)	5.0826e-09	1.1041e-12	2.6000e-06	1.5840e-07
9(a)	8.4927e-06	3.0847e-15	1.0000e-06	1.0036e-05
9(b)	3.1536e-06	3.8350e-16	9.9999e-07	1.0000e-07
9(c)	5.8698e-07	2.4825e-15	1.0000e-06	1.0000e-07
9(d)	1.2207e-06	2.2205e-16	1.0000e-06	1.0000e-07
9(e)	8.3768e-06	2.7974e-15	1.0000e-06	1.0000e-07
10(a)	1.3571e-06	9.6152e-16	4.5000e-06	1.0000e-07
11(a)	3.6467e-06	4.0813e-07	1.5085e-06	1.4404e-07

We give the numbers of function evaluation (FN), gradient evaluation (GR), the number of all inner iterations (IT) in Table 3. 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. It should be noticed that calculations in Step 3 of Algorithm V.1 does not increase the number of evaluations of functions and gradients if the new iterate is not admitted since these values are known after Step 2. Otherwise, the numbers of iterations (IT) and function and gradient evaluations (FN and GR) are increased by 1. There are some differences from the calculations in [8], which calculate the numbers by summing up the evaluations of each component of the vector except the linear functions. For the approach in this paper we do not differentiate the linear and nonlinear functions. The function and gradient evaluations are mainly on functions $f(x,y)$, $c(x,y)$, $H(x,y,\lambda)$ and $g(x,y)$ since values on $\lambda \circ g(x,y)$ can be derived straightly. Since the barrier parameter is decreased by a factor 0.01, the number of outer iterations is 4 for all test problems.

In order to test the robustness of our algorithm, we resolve these problems by using the traditional primal-dual system of equations in the inner loop. The numerical results are similar for almost all problems except problem 7, for which the traditional algorithm failed to obtain the solution since for $\mu = 0.1$, $\alpha_k \to 0$, and the algorithm terminates at the point $x^* = (33.43867, 50.00911)$ after 185

Table 3. Some more results

Prob	IT	FN	GR	Prob	IT	FN	GR
1(a)	19	20	20	1(b)	25	26	26
2(a)	23	24	24	2(b)	26	28	27
3(a)	20	21	21	3(b)	24	25	25
4(a)	21	24	22	4(b)	25	30	26
5(a)	10	11	11	6(a)	22	30	23
7(a)	56	72	57	8(1)	22	23	23
8(2)	24	25	25	8(3)	27	28	28
8(4)	24	25	25	8(5)	24	26	25
8(6)	26	27	27	8(7)	29	30	30
8(8)	22	23	23	8(9)	25	28	26
8(10)	21	22	22	9(a)	20	21	21
9(b)	23	25	24	9(c)	17	18	18
9(d)	18	20	19	9(e)	17	19	18
10(a)	37	38	38	11(a)	26	35	27

iterations, $\rho^* = 2.1629\text{e+}07$, FN= 1161, GR= 185. This result shows that Algorithm IV.1 is probably more robust than the traditional interior method.

We then apply our algorithm to three special examples. The first example is presented by Leyffer in [21] to show a failure of PIPA.

$$\min \ x + y \tag{69}$$
$$\text{s.t.} \ x \in [-1, 1], \tag{70}$$
$$-1 + x + \lambda = 0, \tag{71}$$
$$y \geq 0, \ \lambda \geq 0, \ y\lambda = 0. \tag{72}$$

The standard starting point is $(0, 0.02, 1)$, and the optimal solution is $(-1, 0, 2)$. Our algorithm solves it successfully after 11 iterations. FN=GR=12, $\rho^* = 1$, RD=1.9281e-10, RP=2.2204e-16, RC=5.0020e-07 and CC=7.6393e-08.

The second example is

$$\min \ (x-2)^2 + y^2 \tag{73}$$
$$\text{s.t.} \ x \geq 0, \tag{74}$$
$$(1-x)^3 - \lambda = 0, \tag{75}$$
$$y \geq 0, \ \lambda \geq 0, \ y\lambda = 0, \tag{76}$$

of which the optimal point is $(1, 0, 0)$ and is also a singular stationary point of the problem. The initial point is $(-2, -2, -2)$. The algorithm stops at the approximate point $(0.9998, 0.0011, 0.0)$ with the multiplier vector is $(0.0, -7.9505e{+}06, \ 7.9505e{+}06, \ 0.0, 0.0)$ after 42 iterations, $\mu = 1.0e\text{-}05$, $\rho^* = 3.8724e{+}05$, FN= 43, GR= 42, RD= 1.0281, RP= 6.1176e-12, RC= 0.0 and CC= 2.3229e-15.

The third example is

$$\min \ x + (y-1) \tag{77}$$
$$\text{s.t.} \ x^2 + 1 \leq 0, \tag{78}$$
$$-x - \lambda = 0, \tag{79}$$
$$y \geq 0, \ \lambda \geq 0, \ y\lambda = 0, \tag{80}$$

which is obviously an infeasible MPEC. $(0, 0, 0)$ minimizes the ℓ_2-infeasibility of constraints. The initial point is

(1, 1, 1). Our algorithm stops at $(-0.0, 0.0010, 0.0007)$, an approximate infeasible stationary point after 87 iterations, FN= 125, GR= 87, $\mu = 1.0e$-03, $\rho^* = 2.8882e+11$. These results are interesting since they show that Algorithm V.1 may obtain certain weak stationary points when other methods fail to find meaningful solutions.

References

[1] E. Aiyoshi and K. Shimizu, Hierarchical decentralized systems and its new solution by a barrier method, IEEE Trans. Syst., Man, Cybern., SMC-11(1981), 444-449.

[2] G. Anandalingam and T.L. Friesz, eds., Hierarchical Optimization, Ann. Oper. Res., 1992.

[3] J.F. Bard, Convex two-level optimization, Math. Program., 40(1988), 15-27.

[4] O. Ben-Ayed and C.E. Blair, Computational difficulties of bilevel linear programming, Operations Research, 38(1990), 556-559.

[5] H.Y. Benson, D.F. Shanno and R.J. Vanderbei, Interior-point methods for nonconvex nonlinear programming: complementarity constraints, Research report ORFE-02-02, Operations Research and Financial Engineering, Princeton University.

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

[7] P.A. Clark and A.W. Westerberg, A note on the optimality conditions for the bilevel programming problem, Naval Research Logistics Quarterly, 35(1988), 413-418.

[8] F. Facchinei, H.Y. Jiang and L. Qi, A smoothing method for mathematical programs with equilibrium constraints, Math. Program., 85(1999), 107-134.

[9] J.E. Falk and J. Liu, On bilevel programming, part I: general nonlinear cases, Math. Program., 70(1995), 47-72.

[10] R. Fletcher and S. Leyffer, Numerical experience with solving MPECs by nonlinear programming methods, Numerical Analysis Report NA/YYY, Dept. of Math., Univ. of Dundee, UK, 2001.

[11] M. Fukushima, Z.-Q. Luo and J.-S. Pang, A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints, Comput. Optim. Appl., 10(1998), 5-34.

[12] M. Fukushima and J.S. Pang, Some feasibility issues in mathematical programs with equilibrium constraints, SIAM J. Optim., 8(1998), 673-681.

[13] M. Fukushima and J.S. Pang, 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.

[14] M. Fukushima and P. Tseng, An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints, SIAM J. Optim., 12(2002), 724-739.

[15] D.M. Gay, M.L. Overton and M.H. Wright, 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.

[16] N.I.M. Gould, D. Orban, A. Sartenaer and Ph.L. Toint, Superlinear convergence of primal-dual interior point algorithms for nonlinear programming, SIAM J. Optim., 11(2001), 974-1002.

[17] P.T. Harker and J.S. Pang, On the existence of optimal solutions to mathematical program with equilibrium constraints, Oper. Res. Lett., 7(1988), 61-64.

[18] P.T. Harker and J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications, Math. Program., Ser. B, 48(1990), 161-220.

[19] J. Huang and J.S. Pang, A mathematical programming with equilibrium constraints approach to the implied volatility surface of American options, Research Report, The Johns Hopkins Univ., 2001.

[20] H.Y. Jiang and D. Ralph, Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, SIAM J. Optim., 10(2000), 779-808.

[21] S. Leyffer, 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.

[22] X.W. Liu and J. Sun, A robust primal-dual interior point algorithm for nonlinear programs, http://www.fba.nus.edu.sg/depart/ds/sunjiehomepage/. Research Report, National Univ. of Singapore, 2001. Under second review of SIAM J. Optim.

[23] X.W. Liu and J. Sun, Conditions that ensure convergence of some interior point methods for nonlinear programs, http://www.fba.nus.edu.sg/depart/ds/sunjiehomepage/. Research report, National Univ. of Singapore, 2001.

[24] X.W. Liu and Y.X. Yuan, A robust algorithm for optimization with general equality and inequality constraints, SIAM J. Sci. Comput., 22(2000), 517-534.

[25] Z.Q. Luo, J.S. Pang and D. Ralph, Mathematical Programs with Equilibrium Constraints, Cambridge University Press, 1996.

[26] P. Marcotte, Network design problem with congestion effects: A case of bilevel programming, Math. Program., 34(1986), 142-162.

[27] J.V. Outrata, On optimization problems with variational inequality constraints, SIAM J. Optim., 4(1994), 340-357.

[28] J.V. Outrata and J. Zowe, A numerical approach to optimization problems with variational inequality constraints, Math. Program., 68(1995), 105-130.

[29] J.S. Pang and M. Fukushima, Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints, Comput. Optim. Appl., 13(1999), 111-136.

[30] S. Scholtes, Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, SIAM J. Optim., 11(2001), 918-936.

[31] S. Scholtes and H. Scheel, Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity, Math. Oper. Res., 25(2000), 1-22.

[32] S. Scholtes and M. Stöhr, Exact penalization of mathematical programs with equilibrium constraints, SIAM J. Control Optim., 37(1999), 617-652.

[33] S. Scholtes and M. Stöhr, How stringent is the linear independence assumption for mathematical programs with complementarity constraints, Math. Oper. Res., 26(2001), 851-863.

[34] L.N. Vicente and P. Calamai, Bilevel and multilevel programming: A bibliography review, J. Global Optim., 5(1994), 291-306.