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):  

\[
\begin{align*}
\min & \quad f(x, y) \\
\text{s.t.} & \quad c(x, y) \leq 0, \\
& \quad y \in S(x),
\end{align*}
\]

where \( S(x) \) is the solution set of a parametric variational inequality problem (PVI):  

\[
y \in S(x) \iff \begin{cases} g(x, y) \leq 0, \\
F(x, y)^\top (z - y) \geq 0, \forall z : g(x, z) \leq 0, \\
f : \mathbb{R}^{n+m} \to \mathbb{R}, c : \mathbb{R}^{n+m} \to \mathbb{R}^p, g : \mathbb{R}^{n+m} \to \mathbb{R}^l, \\
F : \mathbb{R}^{n+m} \to \mathbb{R}^m.
\end{cases}
\]

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 \( \theta \), so it is decreased simultaneously with \( \theta \) at every outer loop...
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 $\ell_2$-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 derive 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 subscript $k$ 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. $\nabla g_i(x, y) = (\nabla_x g_i(x, y), \nabla_y g_i(x, y))$, $i = 1, \ldots, \ell$, and $\nabla g(x, y) = [\nabla g_1(x, y), \ldots, \nabla g_\ell(x, y)]$, $\nabla g_J(x, y) = [\nabla g_j(x, y) | j \in J]$, where $J$ 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 $v$, $\text{diag}(v)$ stands for the diagonal matrix whose diagonal is the vector $v$.

Finally, we denote the feasible set of the MPEC by $\mathcal{F}$ and by strict complementarity we mean that $\mathcal{G}_0(x, y, \lambda) = \emptyset$.

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 $x \in X c(x, y) \leq 0$ for some $y \in \mathbb{R}^m$ and $j = 1, \ldots, \ell, g_j(x, y)$ is convex.

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

$$\left\{ \begin{array}{l}
F(x, y) + \sum_{j=1}^\ell \lambda j \nabla y g_j(x, y) = 0, \\
\lambda \geq 0, \ g(x, y) \leq 0, \ \lambda \circ g(x, y) = 0
\end{array} \right. \tag{5}$$

where $\circ$ denotes the Hadamard product. In general we designate the set of $\lambda$ 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 $\lambda$ is also bounded and problem (1)-(3) is equivalent to

$$\min f(x, y) \tag{6}$$

s.t. $c(x, y) \leq 0, \ H(x, y, \lambda) = 0, \ \lambda \geq 0, \ g(x, y) \leq 0, \ \lambda \circ g(x, y) = 0, \tag{9}$$

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

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

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

It is generally difficult to give an explicit expression of $T(x, y; \mathcal{F})$. Instead, the piecewise stationary point of MPEC, defined below, is often used in algorithmic design.
Definition II.3: A point \((x, y) \in F\) is a piecewise stationary point of MPEC if, for \(\lambda \in M(x, y)\), and for each index set \(J \subseteq G_{00}(x, y, \lambda)\), there exist multipliers \(\zeta \in \mathbb{R}^n\), \(\eta \in \mathbb{R}^p\) and \(\pi \in \mathbb{R}^{n_0}\) such that

\[
\nabla f + \nabla c\zeta + \nabla g\eta + \nabla H\pi = 0, \\
\zeta^\top c = 0, \quad \zeta \geq 0, \\
\pi^\top \nabla g_j^2 \geq 0, \quad \text{for } j \in J, \\
\pi_j \nabla g_{j^0} = 0, \quad \text{for } j \in G_{0+}, \\
\eta_j \geq 0, \quad \text{for } j \in G_{00}\backslash J, \\
\eta_j = 0, \quad \text{for } j \notin G_0.
\]

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

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

\[
\left(\begin{array}{ccc}
\nabla H & \nabla c & \nabla g_{00} \\
\nabla_{\lambda} H & 0 & 0 \\
0 & 0 & \left[c_j, j \in G_0(\lambda)\right]
\end{array}\right)
\]

has full column rank, where \(c_j\) is the \(j\)-th coordinate vector.

Then we have the next result.

Proposition II.5: If MPEC-LICQ holds at \((x^*, y^*) \in F\), 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 \((x, y) \in F\) is called a weak piecewise stationary point of MPEC if there exist \(\zeta \in \mathbb{R}^n\), \(\eta \in \mathbb{R}^p\), and \(\pi \in \mathbb{R}^{n_0}\) such that (11)-(12), (14), and (16) hold.

2. A point \((x, y) \in F\) is called a singular stationary point of MPEC if the MPEC-LICQ does not hold at \((x, y)\).

3. A point \((x, y) \in F\) is called an infeasible stationary point of MPEC if \((x, y) \notin F\), and for some \(\lambda \in \mathbb{R}^n\) and some \(\theta > 0\), \((x, y, \lambda)\) is a stationary point of the problem

\[
\min_{(x, y, \lambda)} \left\{ \|c - e\|^2 + \|H\|^2 + \|g\|^2 + \|\lambda - 0\|^2 + \|\lambda \circ g + \theta e\circ \lambda - 0\|^2 \right\},
\]

that is, \((x, y, \lambda)\) satisfies the following equations

\[
\nabla c e + \nabla H H + \nabla g g + \nabla g \lambda (\lambda \circ g + \theta e) = 0, \\
\nabla g \lambda + \lambda \circ g + \lambda \circ g = 0, \\
\Lambda = \text{diag}(\lambda), \\
H^\top = \Lambda + \Lambda + \text{diag}(\lambda).
\]

The optimal value of (18) is an \(\ell_2\) measure of the total infeasibility of problem (6)-(9). If \((x, y, \lambda)\) is a feasible point, then for any \(\theta > 0\), 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 \(\theta\)-relaxation of MPEC we mean the following nonlinear program (NLP(\(\theta\))

\[
\begin{align*}
\min \ f(x, y) \\
\text{s.t. } \ c(x, y) \leq 0, \quad H(x, y, \lambda) = 0, \quad \lambda \geq 0, \quad g(x, y) \leq 0, \quad -\lambda \circ g(x, y) \leq \theta e,
\end{align*}
\]

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 \(\theta\)-relaxation in the usual sense of nonlinear programming (LICQ for NLP(\(\theta\) for short).

Proposition III.1: For \((x^*, y^*) \in F\) and \(\lambda^* \in M(x^*, y^*)\), if the MPEC-LICQ holds at \((x^*, y^*)\), then there exists a neighborhood \(\mathcal{N}\) of \((x^*, y^*, \lambda^*)\) so that for every sufficiently small \(\theta > 0\), the LICQ for NLP(\(\theta\)) holds for every feasible point \((\bar{x}, \bar{y}, \bar{\lambda})\) in \(\mathcal{N}\).

To simplify the notation, we set

\[
\tilde{G}(x, y, \lambda) = (c(x, y), g(x, y), -\lambda) \quad \text{and} \quad G_\theta(x, y, \lambda) = (\tilde{G}(x, y, \lambda), -\lambda \circ g(x, y) - \theta e).
\]

Then

\[
\nabla G_\theta(x, y, \lambda) = [\nabla c(x, y) \ \nabla g(x, y) 0 - \nabla g(x, y)\lambda] (26)
\]

\[
\nabla \lambda G_\theta(x, y, \lambda) = |0 0 - I - \text{diag}(g(x, y))|, \quad (27)
\]

where \(I\) is the \(\ell \times \ell\) identity matrix. The constraints of NLP(\(\theta\)) can be written as \(G_\theta(x, y, \lambda) \leq 0, \quad H(x, y, \lambda) = 0\).

The Lagrange function of program (21)-(24) is

\[
L_\theta(x, y, \lambda, u, v) = \ f(x, y) + u^\top G_\theta(x, y, \lambda) + v^\top H(x, y, \lambda),
\]

where \(u \in \mathbb{R}^{n_0+3\ell}\) and \(v \in \mathbb{R}^{m_0}\) are the multipliers.

Let \(\bar{u} = (u_1, \cdots, u_p), \quad \bar{u} = (u_{p+1}, \cdots, u_{p+\ell})\) and \(\tilde{u} = (u_{p+2\ell+1}, \cdots, u_{p+3\ell})\). Now we show that any KKT point of NLP(\(\theta\)) 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(\(\theta\)), \((\tilde{u}, \bar{u}, \tilde{u})\) is the corresponding multiplier vector associated with constraint \((c, g, -\lambda \circ g - \theta e, H)\). If the sequence \(\{(\bar{x}, \bar{y}, \bar{\lambda}, \bar{u}, \tilde{u}, \bar{v})\}\) is uniformly bounded as \(\theta \to 0\) and \((x^*, y^*, \lambda^*, \tilde{u}^*, \bar{u}^*, \bar{v}^*)\) 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 \(\theta > 0\) such that for \(\theta = 0\) and all \(i \in G_{00}(x^*, y^*, \lambda^*) \cap I_\theta\), 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 \(I_\theta = \{i \mid \bar{\lambda}_i g_i(\bar{x}, \bar{y}) = 0\}\).

This definition is of similar nature to that given by Fukushima and Pang [13], which requires that \(\varsigma_i\) and
**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 second order necessary optimality condition of NLP(θ) as

\[ \ell \rightarrow 0 \]

Henceforth referred as the relaxed barrier problem:

\[ \min_{\theta} \psi_k(d_k, x_k) = \frac{1}{2} z_k^T B_k d_k + z_k^T Z_k^{-1} U_k d_k + \rho u_k \left\| G_k + z_k + H_k + \nabla H_k^\top d_k \right\| \]

such that some prescribed conditions (see the next subsection) hold. Then we compute the search direction \( d_k \) by solving the modified primal-dual system of equations

\[ B_k d_k + G_k^\top d_k + \nabla H_k d_k = -\nabla f_k + \nabla G_k^\top u_k + \nabla H_k v_k, \]

\[ U_k d_k + z_k d_k = -(z_k U_k - \mu e), \]

\[ \nabla G_k^\top d_k + d_k = \nabla G_k^\top d_k^k + d_k^k, \]

\[ \nabla H_k^\top d_k = \nabla H_k^\top d_k^k. \]

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 \( \theta \) and \( \mu \).

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

**Step 1** Given \( (x^0, y^0, u^0, v^0) \in R^{n+m+\ell} \times R^{n+\ell} \times R^{n+\ell} \times R^{n+m+\ell} \) and scalars \( \rho_0 > 0, \nu > 0, \) \( \eta = 0, 1, \xi \in (0, 1), \beta_1 < 1 < \beta_2, \sigma_0 \in (0, 1/\beta_2) \). Let \( k := 0 \).

**Step 2** Calculate the primal search direction \( d_k^p \) and the dual direction \( d_k^d \) by the primal-dual system of equations (40)-(43), where \( (d_k^p, d_k^d) \) is derived by approximately minimizing (39):

\[ \pi_k (d_k^p; \rho_k) \]

**Step 3** Let

\[ \pi_k (d_k^p; \rho_k) = -\frac{1}{2} d_k^p^T B_k d_k - \frac{1}{2} d_k^d^T Z_k^{-1} U_k d_k \]
let \( \rho_{k+1} = \rho_k \); Otherwise, we replace \( \rho_k \) by a larger \( \rho_{k+1} \) (for example \( \rho_{k+1} \geq 2\rho_k \)) such that (44) holds;
Step 4 Compute \( \hat{\alpha}_k \in [0, 1] \) such that \( z^k + \hat{\alpha}_k d_k^z \geq \zeta z^k \),
and select firstly \( \bar{\nu} \in (0, 1) \) and then \( \gamma_k \in [0, 1] \) as large as possible such that
\[
\phi(s^k + \alpha k d_k^z, z^k + \alpha k d_k^z; \rho_{k+1}) - \phi(s^k, z^k; \rho_{k+1}) \leq \sigma_0 \bar{\nu} \sigma_k \nu_k(d^k_k; \rho_k),
\]
(45)
\[\beta_1 \nu \mu \leq (U_k + \gamma_k D_k^p) \max \{z^k + \alpha k d_k^z, -G_\theta(s^k + \alpha k d_k^z)\} \leq \beta_2 \nu \mu,
\]
(46)
where \( D_k^p = \text{diag}(d_k^p) \). Let \( \alpha_k = \sigma \bar{\nu} \). The new primal iterate is generated by
\[
s^{k+1} = s^k + \alpha_k d_k^z, \quad z^{k+1} = \max\{z^k + \alpha_k d_k^z, -G_\theta^{k+1}\},
\]
(47)
(48)
and the new dual iterate is generated by
\[
u^{k+1} = u^k + \gamma k d_k u^k, \quad v^{k+1} = v^k + d_k^d;
\]
(49)
Step 5 If the stopping criterion holds, stop; else calculate values \( \nabla G_\theta^{k+1}, \nabla H^{k+1}, \nabla f_{k+1}, G_\theta^{k+1} \) 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 \( \Omega \subset \mathbb{R}^{n+m+1} \) such that \( s^k \in \Omega \) for all nonnegative integer \( k \).
(2) There exist constants \( \nu_1 \geq \nu_2 > 0 \) such that
\[
u_2 ||d||^2 \leq d^T B_k d \leq \nu_1 ||d||^2 \quad \text{for all} \ d \in \mathbb{R}^{n+m+1}.
\]
(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 \( \{d_k^k, d_k^d, u_k^k\} \) 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 \( (d_k^k, d_k^d) \) solves problem (39) exactly, then \( (d_k^k, d_k^d) \) 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 \( (d_k^k, d_k^d) \to 0 \).
(2) It holds that \( \psi_k(d_k^k, d_k^d) \leq \psi_k(0, 0) \), and there exist constants \( \overline{\rho} > 0 \) and \( \zeta > 0 \) so that \( \forall \rho_k \geq \overline{\rho} \),
\[
\psi_k(d_k^k, d_k^d) - \psi_k(0, 0) \leq -\zeta \rho_k \|
abla G_\theta^k (G_\theta^k + z^k) + \nabla H^k H^k, Z_k (G_\theta^k + z^k)\|_2^2.
\]
(3) There exist \( \nu \in (0, 1), \overline{\rho} > 0 \) and \( \zeta > 0 \) so that \( \forall \rho_k \geq \overline{\rho}, \|
abla \psi_k(d_k^k, Z_k^{-1} d_k^d)\| \leq \zeta \|
abla (G_\theta^k + z^k, H^k)\| \) and \( \psi_k(d_k^k, d_k^d) \leq \psi_k(0, 0) \) if one of the following conditions holds:
(i) \( \{z^k\} \) is componentwise bounded away from zero;
(ii) the vectors \( \nabla H_j^k \), \( j = 1, \ldots, m \), \( \nabla (G_\theta^k) \), \( i \in \mathcal{Q}_0 = \{i \mid z^k_i = 0, \ i = 1, \ldots, q\} \) are linearly independent.

(4) For all \( k \), \( (d_k^k, Z_k^{-1} d_k^d) / \sqrt{\rho_k} \) 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 \( \{d_k^k, d_k^d, u_k^k\} \) and \( \{v_k^k\} \) are bounded.

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

Lemma IV.6: Under Algorithm IV.2, if \( \rho_k \) is bounded, then
\[
\lim_{k \to \infty} \| (d_k^k, d_k^d) \| = 0,
\]
\[
\lim_{k \to \infty} \| (G_\theta^{k+1} + z^{k+1}, H^{k+1})\| = 0,
\]
\[
\lim_{k \to \infty} || Z_{k+1} U_{k+1} e - \mu e || = 0,
\]
\[
\lim_{k \to \infty} \| \nabla f_{k+1} + \nabla G_{\theta}^{k+1} u_{k+1} + \nabla H^{k+1} v_{k+1} \| = 0.
\]
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 \( \rho_k \) is unbounded, then
\[\{z^k\} \] 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 \( k \to \infty \) with \( \nabla G_\theta^k, i \in \mathcal{Q}_0, \nabla H_j^k, j = 1, \ldots, m \) being linearly dependent, where \( \mathcal{Q}_0^k = \{i \mid z^k_i = 0\} \).

(2) there is a subsequence \( \{s^k, z^k \} \in \mathcal{K} \) such that
\[
\lim_{k \to \infty} \left\| \begin{pmatrix} \nabla G_\theta^k & \nabla H^k \\ 0 & Z_k^{1/2} \end{pmatrix} \begin{pmatrix} G_\theta^k + z^k \\ H^k \end{pmatrix} \right\| = 0.
\]
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
\[
\| (G_\theta^* + z^*, H^*) \| = 0, \ Z^* U^* e = \mu e,
\]
\[
\nabla f^* + \nabla G_\theta^* u^* + \nabla H^* v^* = 0.
\]
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 \( \|((G_{\theta_0} + H^*^*)^+) \|= 0 \) and that \( \nabla H^i_j (j = 1, \ldots, m), \nabla G_{\theta_0} (i \in I = \{i \in \{1, \ldots, q \} : G_{\theta_0}^i = 0 \}) \) are linearly dependent, or satisfies that \( \|((G_{\theta_0} + H^*)^+) \| \neq 0 \) and that

\[
\nabla G_{\theta_0}^i (G_{\theta_0}^*)^+ + \nabla H^* H^* = 0.
\]

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 \( \theta \). 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 \( \mu \) and \( \theta \) simultaneously. In particular, the barrier parameter \( \mu \) is selected to be a fraction of \( \theta \) (so \( \theta \) is a multiple of \( \mu \)). 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 \( \mu \) is specified.

Algorithm V.1: (The algorithm for the MPEC)

Step 1 Given the initial point \( (x^0, y^0, \lambda^0, z^0, v^0, v^0) \) with \( (x^0, y^0, \lambda^0) \in \mathbb{R}^{n+m+\ell}, z^0 \in \mathbb{R}^{m+\ell}, u^0 \in \mathbb{R}^{m+\ell} \) and \( v^0 \in \mathbb{R}^m \), the initial barrier parameter \( \mu_0 > 0 \) and penalty parameter \( \rho_0 > 0 \), scalar \( \sigma \), constants \( \tau > 0, \zeta > 0, \kappa \in (0, 1) \), the stopping tolerances \( \epsilon > 0 \), \( \epsilon_1 > \epsilon_2 > 0 \). Let \( \theta_0 = \tau \mu_0, (x^0, y^0, \Lambda^0, z^0, v^0, v^0) = (x^0, y^0, \lambda^0, z^0, v^0, v^0), j := 0; \)

Step 2 With using \( (x^0, y^0, \lambda^0, z^0, v^0, v^0) \) as the starting point, solve the barrier problem (35)-(37) by Algorithm IV.1. The Algorithm IV.1 is terminated when the iterate \( (x^{(k)}, y^{(k)}, \lambda^{(k)}, z^{(k)}, u^{(k)}, v^{(k)}) \) satisfies one of the following groups of conditions:

\[
\begin{align*}
&\|((G_{\theta_j} + z^{(k)}H^{(k)})^+) \| < \zeta \mu_j, \\
&\|Z_{k_j} U_{k_j} e - \mu_j \kappa \| < \zeta \mu_j, \\
&\nabla x f_{k_j} + \nabla x G_{\theta_j}^i u^{(k)} + \nabla x H_{k_j}^i v^{(k)} < \zeta \mu_j, \\
&\nabla y f_{k_j} + \nabla G_{\theta_j}^i u^{(k)} + \nabla y H_{k_j}^i v^{(k)} < \zeta \mu_j, \\
&\nabla \lambda G_{\theta_j}^i u^{(k)} + \nabla \lambda H_{k_j}^i v^{(k)} < \zeta \mu_j; \\
&\|((G_{\theta_j}^i + z^{(k)}H^{(k)})^+) \| \geq \epsilon_1, \\
&\|\nabla E G_{\theta_j}^i (G_{\theta_j}^i + z^{(k)}H^{(k)}) + \nabla E H_{k_j}^i H_{k_j}^i, \\
&\|Z_{k_j} (G_{\theta_j}^i + z^{(k)}H^{(k)})^+ \| < \epsilon_2; \\
&\det(\nabla E (G_{\theta_j}^i + z^{(k)}H^{(k)})^+ \nabla E H_{k_j}^i) < \epsilon_2, \\
&\det(\nabla E (G_{\theta_j}^i + z^{(k)}H^{(k)})^+ \nabla E H_{k_j}^i) < \epsilon_2, \\
&\det(\nabla E (G_{\theta_j}^i + z^{(k)}H^{(k)})^+ \nabla E H_{k_j}^i) < \epsilon_2,
\end{align*}
\]

where \( Z_{k_j} = \text{diag}(z^{(k)}) \) and \( U_{k_j} = \text{diag}(u^{(k)}) \), \( G_{\theta_j}^i \) is the value of \( G_{\theta_j} \) when \( \theta = 0 \), \( \det() \) is the determinant, \( \tilde{\mathcal{I}}_j = \{i \mid \|\tilde{G}_{\theta_j}^i \| \leq \epsilon_2 \}, \)

\( \nabla E (G_{\theta_j}^i + z^{(k)}H^{(k)})^+ \tilde{\mathcal{I}}_j \) is the submatrix of \( \nabla E (G_{\theta_j}^i + z^{(k)}H^{(k)})^+ \) consisting of all columns indexed by \( i \in \tilde{\mathcal{I}}_j \).

Set

\[
(x^{(j+1)}, y^{(j+1)}, \lambda^{(j+1)}) = (x^{(k)}, y^{(k)}, \lambda^{(k)}),
\]

\[
(z^{(j+1)}, u^{(j+1)}, v^{(j+1)}) = (z^{(k)}, u^{(k)}, v^{(k)}),
\]

and

\[
\rho_{j+1} = \max \{\rho_k, ||(u^{(j+1)}, v^{(j+1)})|| + \sigma \}. \]

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 \( (d_{x_k}, d_{y_k}, d_{\lambda_k}, d_{z_k}, d_{u_k}, d_{v_k}) \) by solving equations (40)-(43). Let

\[
(x^{(j+1)}, y^{(j+1)}, \lambda^{(j+1)}) =
\begin{cases}
(x^{(k)}, d_{x_k} + \bar{d}_{x_k}, y^{(k)}, d_{y_k} + \bar{d}_{y_k}, \lambda^{(k)} + d_{\lambda_k} + \bar{d}_{\lambda_k}), \\
(x^{(k)}, d_{x_k} + \bar{d}_{x_k}, y^{(k)}, d_{y_k} + \bar{d}_{y_k}, \lambda^{(k)}), \\
(x^{(k)}, d_{x_k} + \bar{d}_{x_k}, y^{(k)}, d_{y_k} + \bar{d}_{y_k}, \lambda^{(k)} + d_{\lambda_k} + \bar{d}_{\lambda_k}),
\end{cases}
\]

\[
(z^{(j+1)}, u^{(j+1)}, v^{(j+1)}) =
\begin{cases}
(z^{(k)}, d_{z_k} + \bar{d}_{z_k}, u^{(k)} + d_{u_k} + \bar{d}_{u_k}, v^{(k)} + d_{v_k} + \bar{d}_{v_k}), \\
(z^{(k)}, d_{z_k} + \bar{d}_{z_k}, u^{(k)} + d_{u_k} + \bar{d}_{u_k}, v^{(k)}), \\
(z^{(k)}, d_{z_k} + \bar{d}_{z_k}, u^{(k)} + d_{u_k} + \bar{d}_{u_k}, v^{(k)} + d_{v_k} + \bar{d}_{v_k}),
\end{cases}
\]

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

Different from the algorithm for general nonlinear programming in [22], we update the penalty parameter \( \rho_j \) by the information on multipliers, see (63), where we do not need scalar \( \sigma \) 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 \( \nabla E H_{j}(y^{(k)} + \kappa \lambda^{(k)}, \lambda^{(k)}), j = 1, \ldots, m \) 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 \( k \geq 0 \). 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 \( \mu_j \), 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 \( \mu_j \), 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 \( \mu_j, \epsilon = 0 \) 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 \( (d^k_B, d^k_F) \) 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 \( \mu_0 = 0.1, \rho_0 = 0, \sigma_0 = 0.1, \beta_1 = 0.01 \) 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

\[
f(x, y) = 2x_1 + 2x_2 - 3y_1 - 3y_2 + Rw^2 - 60,
\]

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

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

\[
y^0 = x^0_1e_m, \quad \lambda^0 = x^0_1e_l
\]

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

\[
z^0 = \omega e_{(p+3l)}, \quad w^0 = (\mu_0/\omega)e_{(p+3l)}, \quad v^0 = 0.
\]

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.

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 \( (y^0, \lambda^0) = -2x^0, x^* = 5.00000, 0.00000) \) and the optimal objective values are 1.974146-14, 9.151050e-17 and 2.49553e-14, respectively. If \( y^0 = 0, \lambda^0 = 5\epsilon \), then we have \( x^* = (7.324085, 3.78702, 11.475962, 17.21298) \) for problem 10. We also report the optimal penalty parameters \( \rho^0 \) 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 \( RD = \|\nabla f^* + \nabla G^*u^* + \nabla H^*v^*\|, \quad RP = \|(\hat{G}^*_x, H^*)\| \) (\( \hat{G} \) is defined by (25)), \( RC = z^*\top u^* \) and \( CC = \|l^x \circ g^*\|_\infty \).

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).
We give the numbers of function evaluation (FN), gradient evaluation (GR), the number of all inner 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 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, ρ∗=1.9281e-10, RP=2.2204e-16, RC=5.0020e-07 and CC=7.6393e-08.

The second example is

\[
\begin{align*}
\min & \quad (x-2)^2 + y^2 \\
\text{s.t.} \quad & x \geq 0, \\
& (1-x)^3 - \lambda = 0, \\
& y \geq 0, \quad \lambda \geq 0, \quad y\lambda = 0,
\end{align*}
\]

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 (0.99998, 0.00110, 0) with the multiplier vector is (0.0, −7.9505e+06, 7.9505e+06, 0.0, 0.0) after 42 iterations, \(\mu = 1.0e-05\), \(\rho^* = 3.8724e+05\), FN= 43, GR= 42, RD= 1.0281, RP= 6.1176e-12, RC= 0.0 and CC= 2.3229e-15.

The third example is

\[
\begin{align*}
\min & \quad x + (y - 1) \\
\text{s.t.} \quad & x^2 + 1 \leq 0, \\
& -x - \lambda \geq 0, \\
& y \geq 0, \quad \lambda \geq 0, \quad y\lambda = 0,
\end{align*}
\]

which is obviously an infeasible MPEC. (0, 0, 0) minimizes the \(\ell_2\)-infeasibility of constraints. The initial point is
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