A SMOOTHING PENALTY METHOD FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

by

JIAPING ZHU
B.Sc., Nanjing University, 2002

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

Supervisor: Dr. Jane Ye and Co-Supervisor: Dr. Wu-Sheng Lu

ABSTRACT

In this thesis, a new smoothing penalty algorithm is introduced to solve a mathematical program with equilibrium constraints (MPEC). By smoothing the exact penalty function, an MPEC is reformulated as a series of subprograms which belong to a class of MPECs with simple linear complementarity constraints. To deal with the subproblems, a hybrid algorithm is proposed, which combines the active set algorithm, the $\delta-$active search algorithm and the PSQP algorithm. It is shown that the smoothing penalty algorithm converges globally to a M-stationary point of MPEC under weak conditions.

Supervisor: Dr. Jane Ye (Department of Mathematics and Statistics)
Co-Supervisor: Dr. Wu-Sheng Lu (Department of Electrical and Computer Engineering)

Contents

Acknowledgement

The author would like to take this opportunity to express the sincere appreciation to the University of Victoria, Department of Mathematics and Statistics for financial supporting and providing a great environment for learning and researching.

Many thanks to Dr. Jane Ye for her inspiration and patient supervision.

The author would also like to thank Dr. Guoshan Liu for his helpful suggestions.

Jiaping Zhu

Chapter 1

Introduction

In general, a *mathematical program with equilibrium constraints* (MPEC) is a mathematical programming problem whose constraints include variational inequalities. In economics and engineering, many problems can be formulated as MPECs. Among these applications, a special case draws much attention of the researchers, where the equilibrium constraints are in the form of complementarity constraints. Hence, an MPEC is usually referred to a *mathematical program with complementarity constraints.* The readers can refer to [2, 12, 17] for a survey about recent developments and applications of MPECs.

Consider an MPEC in the following form:

$$\text{(MPEC)} \qquad \min \quad f(x)$$
$$\text{s.t.} \quad g(x) \le 0, \quad h(x) = 0,$$
$$G(x) \ge 0, \quad H(x) \ge 0, \quad G(x)^\top H(x) = 0$$

where $f : R^n \to R$, $g : R^n \to R^p$, $h : R^n \to R^q$, $G : R^n \to R^m$, $H : R^n \to R^m$ are continuously differentiable and $a^\top$ denotes the transpose of a vector a.

From a nonlinear programming point of view, the presence of complementarity constraints makes this problem rather hard to deal with. There is no feasible point of the MPEC satisfying all inequality constraints strictly. This fact makes Mangasarian-Fromovitz constraint qualification (MFCQ) violated at any feasible point since MFCQ requires the existence of a strictly feasible point. Therefore, there may not exist Lagrange multipliers and hence well-developed algorithms in nonlinear programming are likely to have unstable numerical behavior. As a result, development of algorithm for MPEC is somewhat behind the analysis of optimality conditions.

Nevertheless, some progress has been announced recently. The difficulty with complementary constraints can be reduced by several different approaches to reformulate an MPEC. Consequently, various stationarity concepts arise (see e.g. [20, 22] for a detailed discussion). As observed in [22], the M-stationary condition is the most appropriate stationary condition for MPECs in the sense that it is the second strongest stationary condition (with the strongest one being the S-stationary condition) and it holds under almost all analogues of the constraint qualifications for nonlinear programming problems.

Moreover, some methods that aimed at solving MPECs numerically have been announced. A class of algorithms use smoothing nonlinear programming problems as approximating subproblems (e.g. [1, 3, 5, 7, 8, 15]), which guarantee to converge to a stationary point. However their convergence requires certain assumptions such as the nondegeneracy condition or the MPEC linear independent constraint qualification (MPEC LICQ)(cf. Chapter 2 in this thesis), which is somehow restrictive in practice. The piecewise sequential quadratic programming (PSQP) algorithm proposed in [18] and extended in [12, 13] uses quadratic programming problems as subproblems and it exhibits superlinear convergence under MPEC LICQ and piecewise second-order sufficient condition (PSOSC). But it is only locally convergent. For the class of MPECs with linear complementarity constraints, some algorithms have shown their promise from the numerical experiments [4, 6, 25]. Most recently, the PSQP algorithm for MPLCQ has been adjusted to provide a global convergence without MPEC LICQ in [26] by using a new technique called δ-active search. Moreover it has been extended to general MPECs by using a *partial* exact penalty function as a merit function in [9].

The main results of this thesis is taken from author's recent joint work [10] with G. S. Liu and J. Ye.

The thesis has two purposes. The primal one is to develop a smoothing exact penalty function algorithm for solving MPECs and show its global convergence to an M-stationary point. Note that the partial exact penalty function proposed in [9] leads to a *partial exact penalization*, which is, however, nonsmooth in general and hence no efficient algorithm is available. By smoothing the partial exact penalty function, we reformulate MPEC as a series of approximated mathematical programs with only simple linear complementarity constraints. Each of approximated mathematical programs is taken as a subproblem at each iteration. The second goal is to introduce a hybrid algorithm to solve the subproblems and establish its properties. The hybrid algorithm will take the advantage of the active set technique, the $\delta-$active search technique and the PSQP algorithm.

The thesis is organized as follows. In the next chapter, the preliminary knowledge is presented. In Chapter 3, we introduce a smoothing penalty method for computing an M-stationary point of MPEC and investigate its convergent properties. In Chapter 4, we propose a hybrid algorithm to solve the subproblems derived from smoothing penalty algorithm and its convergence analysis is provided in the same chapter. Some test problems and the preliminary numerical results with the MATLAB implementation are reported in Chapter 5. We conclude the thesis in Chapter 6 and attach the MATLAB codes in Appendix.

Chapter 2

Preliminaries

2.1 Stationary Points

As mentioned in Chapter 1, the MPEC has various stationarity points based on its different reformulations of complementarity constraints which make the problem intractable. The reader is referred to [22] for a detailed discussion on various stationarity concepts. In this section, we restate the definitions of piecewise stationary point, C-stationary point, M-stationary point and S-stationary point and indicate their interrelations. Given a feasible vector $\bar{x}$ of an MPEC, for convenience, we define the following index sets:

$$I_g := \{i : g_i(\bar{x}) = 0\},$$

$$\alpha := \alpha(\bar{x}) := \{i : G_i(\bar{x}) = 0, H_i(\bar{x}) > 0\},$$

$$\beta := \beta(\bar{x}) := \{i : G_i(\bar{x}) = 0, H_i(\bar{x}) = 0\},$$

$$\gamma := \gamma(\bar{x}) := \{i : G_i(\bar{x}) > 0, H_i(\bar{x}) = 0\}.$$

The set β is known as the *degenerate* set. If β is empty, the vector $\bar{x}$ is said to satisfy the *strict complementarity* condition or *nondegeneracy* condition.

Definition 2.1 *(1) A feasible point $\bar{x}$ of an MPEC is called Clarke stationary point (C-stationary point) if there exists a C-multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that*

$$0 = \nabla f(\bar{x}) + \sum_{i \in I_g} \lambda_i^g \nabla g_i(\bar{x}) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(\bar{x}) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(\bar{x}) + \lambda_i^H \nabla H_i(\bar{x})], \tag{2.1}$$

$$\lambda_i^g \geq 0, \quad \forall i \in I_g, \quad \lambda_i^G = 0, \quad \forall i \in \gamma, \quad \lambda_i^H = 0, \quad \forall i \in \alpha \tag{2.2}$$

and the following conditions hold:

$$\forall i \in \beta, \quad \lambda_i^G \lambda_i^H \geq 0.$$

For convenience, we will denote (2.2) as follows:

$$\lambda_{I_g}^g \geq 0, \quad \lambda_\gamma^G = 0, \quad \lambda_\alpha^H = 0$$

(2) A feasible point $\bar{x}$ of an MPEC is called a Mordukhovich stationary point (M-stationary point) if there exists a M-multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in$

R^{p+q+2m} *such that (2.1)-(2.2) hold and the following condition hold:*

$$\forall i \in \beta, \quad \textit{either } \lambda_i^G > 0,\ \lambda_i^H > 0 \quad \textit{or} \quad \lambda_i^G \lambda_i^H = 0.$$

(3) A feasible point $\bar{x}$ *of an MPEC is called a strong stationary point (S-stationary point) if there exists an S-multiplier* $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ *such that (2.1)-(2.2) hold and the following conditions hold:*

$$\forall i \in \beta,\ \lambda_i^G \geq 0\ \ \lambda_i^H \geq 0.$$

(4) A feasible point $\bar{x}$ *of an MPEC is called a piecewise stationary point if for each partition of the index set* β *into sets* β_1, β_2*, there exists a P-multiplier* $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ *such that (2.1)-(2.2) hold and the following conditions hold:*

$$\lambda_i^H \geq 0\ \ \forall i \in \beta_1 \textit{ and } \lambda_i^G \geq 0\ \forall i \in \beta_2.$$

From the definitions above, we can see the interrelations among the stationary concepts as illustrated in the following diagram.

$$\begin{array}{ccc} \text{S-Stationary Point} & \Rightarrow & \text{Piecewise Stationary Point} \\ \Downarrow & & \\ \text{M-Stationary Point} & & \\ \Downarrow & & \\ \text{C-Stationary Point} & & \end{array}$$

We now make use of the following example to show the various stationary concepts for MPEC defined above.

$$\begin{array}{lll} \text{(P)} & \min & -y \\ & \text{s.t.} & x - y = 0, \\ & & x \geq 0, \quad y \geq 0, \quad x^\top y = 0 \end{array}$$

where $x \in R$ and $y \in R$. This example is discussed in [11]. It is obvious that (0,0) is the only optimal solution of problem (P) since it is the only feasible solution of the problem. First, we will show, in the sense of the nonlinear programming, the KKT multiplier does not exist at (0,0). A KKT multiplier $(\lambda^x, \lambda^y, \lambda_1^h, \lambda_2^h)$ of (P) must satisfy

$$\begin{cases} 0 = (0, -1) + \lambda_1^h(1, -1) - \lambda^x(1, 0) - \lambda^y(0, 1) + \lambda_2^h(0, 0), \\ \lambda^x \geq 0, \quad \lambda^y \geq 0. \end{cases}$$

It is clear that the above system in terms of $(\lambda^x, \lambda^y, \lambda_1^h, \lambda_2^h)$ does not have any solution. Thus, the KKT multiplier in the sense of nonlinear programming does not exist.

To calculate the C-multipliers for problem (P), we consider the system in terms of $(\lambda^x, \lambda^y, \lambda^h)$:

$$\begin{cases} 0 = (0,-1) + \lambda^h(1,-1) - \lambda^x(1,0) - \lambda^y(0,1), \\ \lambda^x \lambda^y \geq 0. \end{cases}$$

The solution is $(\lambda^x, -1-\lambda^x, \lambda^x)$ for any $-1 \leq \lambda^x \leq 0$. Therefore, (0,0) is a C-stationary point.

The M-multipliers at (0,0) satisfy:

$$\begin{cases} 0 = (0,-1) + \lambda^h(1,-1) - \lambda^x(1,0) - \lambda^y(0,1), \\ \lambda^x > 0, \quad \lambda^y > 0, \quad or \quad \lambda^x \lambda^y = 0. \end{cases}$$

The above system implies that

$$(\lambda^x, \lambda^y, \lambda^h) = (0,-1,0)$$

or

$$(\lambda^x, \lambda^y, \lambda^h) = (-1,0,-1).$$

Since the optimal solution (0,0) for (P) is also the optimal solution for the subproblem:

$$\text{(P1)} \qquad \min \quad -y$$

$$\text{s.t.} \quad x - y = 0,$$

$$x \geq 0, \quad y = 0$$

and the subproblem

$$\text{(P2)} \qquad \min \quad -y$$

$$\text{s.t.} \quad x - y = 0,$$

$$x = 0, \quad y \geq 0.$$

The KKT multiplier for (P1) satisfy the following system:

$$\begin{cases} 0 = (0, -1) + \lambda^h(1, -1) - \lambda^x(1, 0) - \lambda^y(0, 1), \\ \lambda^x \geq 0, \end{cases}$$

and the KKT multiplier for (P2) satisfy the following system:

$$\begin{cases} 0 = (0, -1) + \lambda^h(1, -1) - \lambda^x(1, 0) - \lambda^y(0, 1), \\ \lambda^y \geq 0. \end{cases}$$

Therefore, the P multiplier set is

$$\{(\lambda^x, -1 - \lambda^x, \lambda^x) : \lambda^x \geq 0\} \bigcup \{(-1 - \lambda^y, \lambda^y, -1 - \lambda^y) : \lambda^y \geq 0\}.$$

The optimal solution (0,0) is not an S-stationary point since an S-multiplier

$(\lambda^x, \lambda^y, \lambda^h)$ of (P) must satisfy

$$\begin{cases} 0 = (0, -1) + \lambda^h(1, -1) - \lambda^x(1, 0) - \lambda^y(0, 1), \\ \lambda^x \geq 0, \quad \lambda^y \geq 0. \end{cases}$$

which has no solution.

2.2 Constraint Qualifications and Necessary Conditions

Notice that any feasible point of the MPEC does not satisfy the standard LICQ which guarantee the existence and uniqueness of Lagrange multipliers at a local minimizer of a nonlinear program. In order to gain the similar property, we make use of the following definition of LICQ for MPECs.

Definition 2.2 (MPEC LICQ) *Let $\bar{x}$ be a feasible point of an MPEC. We say that MPEC linear independence constraint qualification (MPEC LICQ) is satisfied at $\bar{x}$ if the gradient vectors*

$$\begin{aligned} &\nabla g_i(\bar{x}) \quad \forall i \in I_g, \\ &\nabla h_i(\bar{x}) \quad \forall i = 1, 2, \ldots, q, \\ &\nabla G_i(\bar{x}) \quad \forall i \in \alpha \cup \beta, \\ &\nabla H_i(\bar{x}) \quad \forall i \in \gamma \cup \beta \end{aligned}$$

are linearly independent.

MPEC LICQ discards the gradients of complementarity constrains, which make it different from the standard LICQ. Indeed, MPEC LICQ is the standard LICQ for the relaxed MPEC defined below and the S-stationary condition happens to be its KKT condition:

$$
\begin{array}{lll}
\text{(RMPEC)} & \min & f(x) \\
& \text{s.t.} & g(x) \leq 0, \quad h(x) = 0, \\
& & G_i(x) = 0, \quad H_i(x) > 0 \quad i \in \alpha, \\
& & G_i(x) > 0, \quad H_i(x) = 0 \quad i \in \gamma, \\
& & G_i(x) \geq 0, \quad H_i(x) \geq 0 \;\; i \in \beta.
\end{array}
$$

Before giving a necessary condition for an MPEC, we define the branch of MPEC in terms of each partition (β_1, β_2) of β:

$$
\begin{array}{lll}
\text{MPEC}_{(\beta_1,\beta_2)} & \min & f(x) \\
& \text{s.t.} & g(x) \leq 0, \quad h(x) = 0, \\
& & G_i(x) = 0, \quad H_i(x) > 0 \quad i \in \alpha, \\
& & G_i(x) > 0, \quad H_i(x) = 0 \quad i \in \gamma \\
& & G_i(x) \geq 0, \quad H_i(x) = 0 \;\; i \in \beta_1, \\
& & G_i(x) = 0, \quad H_i(x) \geq 0 \; i \in \beta_2.
\end{array}
$$

Theorem 2.1 *(see e.g. [12, proposition 4.3.7] or [21, Theorem 3.2]) Let $\bar{x}$ be a local optimal solution for an MPEC where all defining functions are continuously differentiable at $\bar{x}$. Suppose that MPEC LICQ is satisfied at $\bar{x}$, then it is an S-stationary point.*

Proof. Suppose that $\bar{x}$ is a local optimal solution for an MPEC. It is obvious that $\bar{x}$ is a local optimal solution of an MPEC if and only if it is a local optimal solution to MPEC(β_1, β_2) for all partition (β_1, β_2) of β. Since MPEC LICQ is satisfied at $\bar{x}$, the standard LICQ holds for all $MPEC_{(\beta_1,\beta_2)}$. Therefore, for any partition (β_1, β_2) of β, there exists a unique Lagrange multiplier $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that (2.1)-(2.2) and the following conditions hold:

$$\lambda_i^G \geq 0 \quad \forall i \in \beta_1, \quad \lambda_i^H \geq 0, \quad \forall i \in \beta_2. \tag{2.3}$$

Now we will prove that, under MPEC LICQ, there exists a common and unique Lagrange multiplier such that (2.3) is true for any partition of β. By contradiction, suppose that there are two different P-multipliers $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ associated with the partition (β_1, β_2) and $\lambda' = (\lambda^{g'}, \lambda^{h'}, \lambda^{G'}, \lambda^{H'}) \in R^{p+q+2m}$ associated with the partition $({\beta_1}', {\beta_2}') \neq (\beta_1, \beta_2)$

satisfying (2.1)-(2.2), respectively. Then

$$0 = \sum_{i\in I_g}(\lambda_i^g - \lambda_i^{g'})\nabla g_i(\bar{x}) + \sum_{i=1}^{q}(\lambda_i^h - \lambda_i^{h'})\nabla h_i(\bar{x})$$
$$-\sum_{i=1}^{m}[(\lambda_i^G - \lambda_i^{G'})\nabla G_i(\bar{x}) + (\lambda_i^H - \lambda_i^{H'})\nabla H_i(\bar{x})], \quad (2.4)$$
$$\lambda_\gamma^G - \lambda_\gamma^{G'} = 0, \quad \lambda_\alpha^H - \lambda_\alpha^{H'} = 0. \quad (2.5)$$

Since $(\lambda^g - \lambda^{g'}, \lambda^h - \lambda^{h'}, \lambda^G - \lambda^{G'}, \lambda^H - \lambda^{H'})$ is not a zero vector, there is a contradiction since MPEC LICQ is satisfied at the point $\bar{x}$. The S-stationary condition follows and $\bar{x}$ is an S-stationary point. ■

From the proof above, we can see that the piecewise stationary condition is equivalent to the S-stationary condition under the MPEC LICQ. Moreover, all concepts of stationary points coincide at any local optimal solution under this CQ.

To introduce other necessary conditions, we recall the definitions of two weaker constraint qualifications compared with MPEC LICQ .

Definition 2.3 (NNAMCQ) *Let $\bar{x}$ be a feasible point of MPEC where all defining functions are continuously differentiable at $\bar{x}$. We say that the No Nonzero Abnormal Multiplier Constraint Qualification (NNAMCQ) is satis-*

fied at $\bar{x}$ if there is no nonzero vector $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that

$$0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(\bar{x}) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(\bar{x}) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(\bar{x}) + \lambda_i^H \nabla H_i(\bar{x})],$$

$$\lambda_{I_g}^g \geq 0, \quad \lambda_\gamma^G = 0, \quad \lambda_\alpha^H = 0,$$

$$\text{either } \lambda_i^G > 0, \lambda_i^H > 0 \quad \text{or} \quad \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \beta.$$

Definition 2.4 (MPEC GMFCQ) *Let $\bar{x}$ be a feasible point of MPEC where all defining functions are continuously differentiable at $\bar{x}$. We say that MPEC generalized Mangasarian-Fromovitz constraint qualification (MPEC GMFCQ) is satisfied at z^* if*

(i) for every partition of β into sets P, Q, R with $R \neq \emptyset$, there exists an d such that

$$\nabla g_i(\bar{x})^\top d \leq 0 \quad \forall i \in I_g,$$

$$\nabla h_i(\bar{x})^\top d = 0 \quad \forall i = 1, 2, \ldots, q,$$

$$\nabla G_i(\bar{x})^\top d = 0 \quad \forall i \in \alpha \cup Q,$$

$$\nabla H_i(\bar{x})^\top d = 0 \quad \forall i \in \gamma \cup P,$$

$$\nabla G_i(\bar{x})^\top d \geq 0, \nabla H_i(\bar{x})^\top d \geq 0 \quad i \in R$$

and for some $i \in R$ either $\nabla G_i(\bar{x})^\top d > 0$ or $\nabla H_i(\bar{x})^\top d > 0$;

(ii) for every partition of β into sets P, Q, the gradient vectors

$$\nabla h_i(\bar{x}) \quad \forall i = 1, 2, \ldots, q,$$

$$\nabla G_i(\bar{x}) \quad \forall i \in \alpha \cup Q,$$

$$\nabla H_i(\bar{x}) \quad \forall i \in \gamma \cup P$$

are linearly independent and there exists an $d \in R^n$ such that

$$\nabla g_i(\bar{x})^\top d < 0 \quad \forall i \in I_g,$$

$$\nabla h_i(\bar{x})^\top d = 0 \quad \forall i = 1, 2, \ldots, q,$$

$$\nabla G_i(\bar{x})^\top d = 0 \quad \forall i \in \alpha \cup Q$$

$$\nabla H_i(\bar{x})^\top d = 0 \quad \forall i \in \gamma \cup P.$$

Note that in the sense of nonlinear programming, the MPEC GMPCQ is reduced to the standard MFCQ which requires that: the gradient vectors

$$\nabla h_i(\bar{x}) \quad \forall i = 1, 2, \ldots, q,$$

are linearly independent and there exists an $d \in R^n$ such that

$$\nabla g_i(\bar{x})^\top d < 0 \quad \forall i \in I_g,$$

$$\nabla h_i(\bar{x})^\top d = 0 \quad \forall i = 1, 2, \ldots, q.$$

In nonlinear programming, it is well-known that NNAMCQ is equivalent to MFCQ. For MPECs, the similar relation exists.

Proposition 2.2 *(See e.g. Ye [22, Proposition 2.1]) NNAMCQ is equivalent to MPEC GMFCQ.*

The following result is not new. It was essentially introduced by Ye and Ye in [24, Theorem 3.2] and further studied in Outrata [16] and Ye [23]. The reader is referred to Ye [22] for a simpler proof.

Theorem 2.3 (Kuhn-Tucker type necessary M-stationary condition) *(See e.g. Ye [22, Corollary 2.1]) Let $\bar{x}$ be a local optimal solution for an MPEC where all defining functions are continuously differentiable at $\bar{x}$. Suppose that either NNAMCQ or MPEC GMFCQ is satisfied at $\bar{x}$, then it is an M-stationary point.*

For a nonlinear programming problem, MFCQ leads to an exact penalization. Under NNAMCQ or MPEC GMFCQ, a similar property holds for MPECs.

Theorem 2.4 (Partial exact penalty) *(Liu, Ye and Zhu [10, Theorem 2.2]) Let $\bar{x}$ be a local solution of MPEC. If NNAMCQ or MPEC GMFCQ holds at $\bar{x}$, then $(\bar{x}, G(\bar{x}), H(\bar{x}))$ is also a local solution of the partial exact penalty problem $(FMPEC_\mu)$ for some $\mu > 0$:*

$$(FMPEC_\mu) \qquad \min \quad \phi_\mu(x,y,z)$$
$$s.t. \quad y \geq 0, \quad z \geq 0, \quad y^\top z = 0,$$

where $\phi_\mu(x,y,z) := f(x) + \mu(\|G(x)-y\|_1 + \|H(x)-z\|_1 + \|\max\{0, g(x)\}\|_1 + \|h(x)\|_1)$ *and* $\|\cdot\|_1$ *denotes the* L^1 *norm defined by*

$$\|x\|_1 = \sum_{i=1}^{m} |x_i| \tag{2.6}$$

for any $x \in R^m$.

Chapter 3

A Smoothing Penalty Method for MPEC

In this chapter, we introduce a smoothing penalty method in Section 3.1 and show its globally convergent properties in Section 3.2.

3.1 Smoothing Penalty Method

This new smoothing penalty method is motivated by Theorems 2.3 and 2.4. Since it makes the problem computationally difficult that the partial exact penalty function $\phi_\mu(x, y, z)$ is nondifferentiable, we smooth the partial exact penalty function $\phi_\mu(x,y,z)$ by introducing a smoothing parameter ρ. Consider the following smoothing problem:

$$(FMPEC_\mu^\rho) \qquad \min \quad \phi_\mu^\rho(x,y,z)$$
$$\text{s.t.} \quad y \geq 0,\ z \geq 0,\ y^\top z = 0,$$

where

$$\phi_\mu^\rho(x,y,z) := f(x) + \mu[\sum_{i=1}^{m} \sqrt{(G_i(x)-y_i)^2+\rho} + \sum_{i=1}^{m} \sqrt{(H_i(x)-z_i)^2+\rho}$$
$$+\frac{1}{2}\sum_{i=1}^{p}(\sqrt{g_i(x)^2+\rho} + g_i(x)) + \sum_{i=1}^{q} \sqrt{h_i(x)^2+\rho}]$$

is the smoothing function of partial exact penalty function $\phi_\mu(x,y,z)$ since the sequence $\{\phi_\mu^\rho(x,y,z)\}$ converges to $\phi_\mu(x,y,z)$ as $\rho \to 0$ and $\phi_\mu^\rho(x,y,z)$ is smooth.

The scheme of this smoothing algorithm is given below. At each iteration, we find an S-stationary point of $FMPEC_\mu^\rho$ for the given $\mu \geq 0$ and $\rho \geq 0$. The algorithm will stop at this point if prescribed termination rule is satisfied. Otherwise we go to the next iteration after we update the problem $FMPEC_\mu^\rho$ by adjusting μ dynamically and decreasing ρ according to certain rules. The convergence analysis in Section 3.2 shows that under the extended MPEC GMFCQ, any accumulated point of the sequence generated by this new algorithm is an M-stationary point of the MPEC.

In the following proposition, we can derive the form of S-multipliers for the problem $FMPEC_{\mu}^{\rho}$ due to its simple constraints.

Proposition 3.1 *For any positive integer k, any positive numbers ρ_k and μ_k, if (x_k, y_k, z_k) is an S-stationary point of $FMPEC_{\mu_k}^{\rho_k}$, then there exists unique S-multiplier $(\lambda^{x_k}, \lambda^{y_k})$ of $FMPEC_{\mu_k}^{\rho_k}$ in the following form:*

$$\lambda_i^{y_k} = \mu_k \frac{y_i^k - G_i(x_k)}{\sqrt{(G_i(x_k) - y_i^k)^2 + \rho_k}}, \quad i = 1, 2 \ldots, m, \tag{3.1}$$

$$\lambda_i^{z_k} = \mu_k \frac{z_i{}^k - H_i(x_k)}{\sqrt{(H_i(x_k) - z_i^k)^2 + \rho_k}}, \quad i = 1, 2 \ldots, m. \tag{3.2}$$

Proof. Let (x_k, y_k, z_k) be an S-stationary point of $FMPEC_{\mu_k}^{\rho_k}$. The uniqueness of the S-multiplier is ensured by MPEC LICQ which is satisfied at (x_k, y_k, z_k) naturally. Let $(\lambda^{y_k}, \lambda^{z_k})$ be the corresponding S-multiplier such that:

$$\begin{aligned} 0 = \nabla f(x_k) + \mu_k [&\sum_{i=1}^{m} \frac{G_i(x_k) - y_i^k}{\sqrt{(G_i(x_k) - y_i^k)^2 + \rho_k}} \nabla G_i(x_k) \\ &+ \sum_{i=1}^{m} \frac{H_i(x_k) - z_i^k}{\sqrt{(H_i(x_k) - z_i^k)^2 + \rho_k}} \nabla H_i(x_k) + \sum_{i=1}^{q} \frac{h_i(x_k)}{\sqrt{h_i(x_k)^2 + \rho_k}} \nabla h_i(x_k) \\ &+ \frac{1}{2} \sum_{i=1}^{p} (\frac{g_i(x_k)}{\sqrt{g_i(x_k)^2 + \rho_k}} + 1) \nabla g_i(x_k))], \end{aligned} \tag{3.3}$$

$$0 = \mu_k \frac{y_i^k - G_i(x_k)}{\sqrt{(G_i(x_k) - y_i^k)^2 + \rho_k}} - \lambda_i^{y_k} \qquad i = 1, 2 \ldots, m, \tag{3.4}$$

$$0 = \mu_k \frac{z_i^k - H_i(x_k)}{\sqrt{(H_i(x_k) - z_i^k)^2 + \rho_k}} - \lambda_i^{z_k} \qquad i = 1, 2 \ldots, m, \tag{3.5}$$

$$\lambda_i^{z_k} = 0, \quad i \in I_1^k, \quad \lambda_i^{y_k} = 0, \quad i \in I_3^k, \tag{3.6}$$

$$\lambda_i^{z_k} \geq 0, \quad \lambda_i^{y_k} \geq 0, \quad i \in I_2^k, \tag{3.7}$$

where

$$I_1^k = \{i : y_i^k = 0, z_i^k > 0\},$$

$$I_2^k = \{i : y_i^k = 0, z_i^k = 0\},$$

$$I_3^k = \{i : y_i^k > 0, z_i^k = 0\}.$$

From (3.4) and (3.5) we have (3.1) and (3.2), respectively. ■

Proposition 3.2 *For any positive integer k and any positive numbers ρ_k, μ_k, if (x_k, y_k, z_k) is an S-stationary point of $FMPEC_{\mu_k}^{\rho_k}$, then we have*

$$\begin{aligned} 0 = \quad & \nabla f(x_k) - \sum_{i=1}^{m} \lambda_i^{G_k} \nabla G_i(x_k) - \sum_{i=1}^{m} \lambda_i^{H_k} \nabla H_i(x_k) \\ & + \sum_{i=1}^{q} \lambda_i^{h_k} \nabla h_i(x_k) + \sum_{i=1}^{p} \lambda_i^{g_k} \nabla g_i(x_k) \end{aligned} \tag{3.8}$$

holds for multiplier $(\lambda^{G_k}, \lambda^{H_k}, \lambda^{h_k}, \lambda^{g_k})$ defined by

$$\lambda_i^{G_k} := \mu_k \frac{y_i^k - G_i(x_k)}{\sqrt{(G_i(x_k) - y_i^k)^2 + \rho_k}}, \quad i = 1, 2 \ldots, m, \tag{3.9}$$

$$\lambda_i^{H_k} := \mu_k \frac{z_i^k - H_i(x_k)}{\sqrt{(H_i(x_k) - z_i^k)^2 + \rho_k}}, \quad i = 1, 2 \ldots, m, \tag{3.10}$$

$$\lambda_i^{h_k} := \mu_k \frac{h_i(x_k)}{\sqrt{h_i(x_k)^2 + \rho_k}}, \quad i = 1, 2 \ldots, q \tag{3.11}$$

$$\lambda_i^{g_k} := \frac{1}{2} \mu_k \left(\frac{g_i(x_k)}{\sqrt{g_i(x_k)^2 + \rho_k}} + 1 \right), \quad i = 1, 2 \ldots, p. \tag{3.12}$$

Proof. Substituting (3.9)-(3.12) into (3.3), we have (3.8). ■

For convenience, we denote the feasible region of the problem $FMPEC_\mu^\rho$ by $\mathcal{WF}$. Since a feasible point of $FMPEC_\mu^\rho$ may not be feasible to MPEC, we call $\mathcal{WF}$ the weak feasible region of MPEC and a point in the weak feasible region a weak feasible point.

Now we are ready to formally state the smoothing penalty algorithm.

Algorithm 3.1

Step 1. Given a weak feasible point $(x_0, y_0, z_0) \in \mathcal{WF}$. Let $\mu_k > 0, \rho_k > 0$, $1 > \alpha_2 > \alpha_1 > 0$ and let $k = 1$.

Step 2. Terminate if a prescribed stopping rule is satisfied, otherwise find an S-stationary point (x_k, y_k, z_k) of $FMPEC_{\mu_k}^{\rho_k}$ with the the Lagrange multiplier $(\lambda^{G_k}, \lambda^{H_k}, \lambda^{h_k}, \lambda^{g_k})$ given in (3.9)-(3.12).

Step 3. Decrease ρ_k, let

$$\mu_{k+1} = \begin{cases} \mu_k, \quad if \ \sum_{i=1}^{m} |\lambda_i^{G_k}| + \sum_{i=1}^{m} |\lambda_i^{H_k}| + \sum_{i=1}^{q} |\lambda_i^{h_k}| + \sum_{i=1}^{p} |\lambda_i^{g_k}| \leq (1-\alpha_1)\mu_k \\ \sum_{i=1}^{m} |\lambda_i^{G_k}| + \sum_{i=1}^{m} |\lambda_i^{H_k}| + \sum_{i=1}^{q} |\lambda_i^{h_k}| + \sum_{i=1}^{p} |\lambda_i^{g_k}| + \alpha_2\mu_k, \quad otherwise. \end{cases}$$

and return to step 2 with k replaced by $k+1$.

In Algorithm 3.1, we did not specify the rules for termination and decreasing ρ_k. In fact, we only need to reduce ρ_k to zero step by step in order to obtain the convergence of the algorithm. A termination rule and a simple rule for decreasing ρ_k will be described in Chapter 5.

3.2 Global Convergence of Smoothing Penalty Method

Because of using the penalty function, we face the issue regarding the feasibility at an accumulated point of the generated sequence by this algorithm. In fact, this is a common issue of all the penalty function methods to optimization problems. We give the condition under which the feasibility can be ensured in the following proposition.

Proposition 3.3 *If the exact penalty parameter sequence* $\{\mu_k\}$ *is bounded and* $\lim_{k\to\infty} \rho_k = 0$*, then any accumulated point* $(\bar{x}, \bar{y}, \bar{z})$ *of the sequence generated by Algorithm 3.1 is feasible for FMPEC. Moreover* $\bar{x}$ *is feasible to MPEC.*

Proof. Suppose that $(\bar{x}, \bar{y}, \bar{z})$ is an accumulated point of the sequence generated by Algorithm 3.1. Then there is a subsequence $\{(x_k, y_k, z_k)\}$ of the sequence generated by the algorithm such that

$$\lim_{k\to\infty}(x_k, y_k, z_k) = (\bar{x}, \bar{y}, \bar{z}).$$

Let $(\lambda^{G^k}, \lambda^{H^k}, \lambda^{h_k}, \lambda^{g_k})$ be the Lagrange multiplier defined in (3.9)-(3.12). Since the sequence $\{\mu_k\}$ is bounded, Next, we are going to show that there is an integer K such that

$$\mu_k = \mu_K, \quad for \;\; all \;\; k \geq K.$$

If there is not such an integer K, then there exists a strictly increasing subsequence of $\{\mu_k\}$. Without loss of the generality, we assume that the sequence $\{\mu_k\}$ is strictly increasing. By the updating scheme, we have that

$$\begin{aligned} & \sum_{i=1}^{m}|\lambda_i^{G_k}| + \sum_{i=1}^{m}|\lambda_i^{H_k}| + \sum_{i=1}^{q}|\lambda_i^{h_k}| + \sum_{i=1}^{p}|\lambda_i^{g_k}| \\ > \;\; & \mu_k(1-\alpha_1) \end{aligned}$$

and

$$\mu_{k+1} = \sum_{i=1}^{m}|\lambda_i^{G_k}| + \sum_{i=1}^{m}|\lambda_i^{H_k}| + \sum_{i=1}^{q}|\lambda_i^{h_k}| + \sum_{i=1}^{p}|\lambda_i^{g_k}| + \alpha_2\mu_k.$$

Then we have the following inequality:

$$\mu_{k+1} > \mu_k(1+\alpha_2-\alpha_1) > \cdots > \mu_1(1+\alpha_2-\alpha_1)^k.$$

Since $\alpha_2 > \alpha_1$, $\mu_k \to +\infty$ as $k \to +\infty$ which contradicts to the boundedness of $\{\mu_k\}$.

Therefore, from the updating scheme of penalty parameter, we have following:

$$\begin{aligned}
&\sum_{i=1}^{m}|\lambda_i^{G_k}| + \sum_{i=1}^{m}|\lambda_i^{H_k}| + \sum_{i=1}^{q}|\lambda_i^{h_k}| + \sum_{i=1}^{p}|\lambda_i^{g_k}| \\
\leq\ & \mu_K(1-\alpha_1) \\
<\ & \mu_K,
\end{aligned}$$

for all $k \geq K$. Hence, without loss of generality, we can assume that

$$\lambda_i^G = \lim_{k\to+\infty} \lambda_i{}^{G_k},\ \lambda_i^H = \lim_{k\to+\infty} \lambda_i{}^{H_k},\ \lambda_i^h = \lim_{k\to+\infty} \lambda_i{}^{h_k},\ \lambda_i^g = \lim_{k\to+\infty} \lambda_i{}^{g_k}.$$

Then

$$\begin{aligned}
&\sum_{i=1}^{m}|\lambda_i{}^G| + \sum_{i=1}^{m}|\lambda_i{}^H| + \sum_{i=1}^{q}|\lambda_i{}^h| + \sum_{i=1}^{p}|\lambda_i{}^g| \\
\leq\ & \mu_K(1-\alpha_1) \\
<\ & \mu_K. \qquad (3.13)
\end{aligned}$$

From the definition (3.9), we have

$$y_i^k - G_i(x_k) = \frac{\lambda_i^{G_k}}{\mu_K}\sqrt{(G_i(x_k)-y_i^k)^2 + \rho_k}$$

for any i. Taking the limit on the both side of the above equation, we have

$$\bar{y}_i - G_i(\bar{x}) = \frac{\lambda_i^G}{\mu_K}|G_i(\bar{x}) - \bar{y}_i|.$$

If $\bar{y}_i - G_i(\bar{x}) \neq 0$, then $|\frac{\lambda_i^G}{\mu_K}| = 1$, which is contradicted with (3.13). Therefore, $\bar{y} - G(\bar{x}) = 0$. Similarly, we can prove that

$$\bar{z} - H(\bar{x}) = 0, \quad h(\bar{x}) = 0, \quad g(\bar{x}) \leq 0.$$

Moreover, $y_k \geq 0, z_k \geq 0, y_k^\top z_k = 0$ implies that $\bar{y} \geq 0, \bar{z} \geq 0, \bar{y}^\top \bar{z} = 0$. Therefore, $(\bar{x}, \bar{y}, \bar{z})$ is feasible to FMPEC. Obviously, $\bar{x}$ is feasible to MPEC.

■

The following proposition is of most importance for the new smoothing penalty algorithm. It proves that any accumulated point of the algorithm is an M-stationary point as long as the exact penalty parameter sequence $\{\mu_k\}$ is bounded.

Theorem 3.4 *Suppose that* $\lim_{k\to\infty} \rho_k = 0$ *and the exact penalty parameter sequence* $\{\mu_k\}$ *is bounded. Let* $(\bar{x}, \bar{y}, \bar{z})$ *be any accumulated point of the sequence generated by Algorithm 3.1. Then* $\bar{x}$ *is an M-stationary point of MPEC.*

Proof. In Proposition 3.3, we have proved that $(\bar{x}, \bar{y}, \bar{z})$ is feasible to FMPEC. The feasibility of $\bar{x}$ to MPEC and the fact that $\bar{y} = G(\bar{x}), \bar{z} = H(\bar{x})$ follows immediately. Without loss of generality, we may assume that

$$\lim_{k\to\infty} (x_k, y_k, z_k) = (\bar{x}, \bar{y}, \bar{z}),$$

where $\{(x_k, y_k, z_k)\}$ is the sequence generated by Algorithm 3.1. As in the proof of Proposition 3.3, we may assume that

$$\lambda_i^G = \lim_{k\to+\infty} \lambda_i^{G_k}, \quad \lambda_i^H = \lim_{k\to+\infty} \lambda_i^{H_k},$$

$$\lambda_i^h = \lim_{k\to+\infty} \lambda_i^{h_k}, \quad \lambda_i^g = \lim_{k\to+\infty} \lambda_i{}^{g_k},$$

where $\lambda_i^{G_k}, \lambda_i^{H_k}, \lambda_i^{h_k}$ and $\lambda_i^{g_k}$ are given by (3.9)-(3.12). Taking limit in (3.8), we have

$$0 = \nabla f(\bar{x}) - \sum_{i=1}^{m} \lambda_i^G \nabla G_i(\bar{x}) - \sum_{i=1}^{m} \lambda_i^H \nabla H_i(\bar{x}) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(\bar{x}) + \sum_{i=1}^{p} \lambda_i^g \nabla g_i(\bar{x}).$$

Clearly, $\lambda_{I_g}^g \geq 0$. Moreover $\lambda_i^g = 0$ for all $i \notin I_g$. Let $I_g^k = \{i : g_i(x_k) = 0\}$. By the continuity of g, for large k, $I_g^k \subseteq I_g$ which implies that for any $i \notin I_g$, $g_i(x_k) < 0$ and hence

$$\begin{aligned}
\lambda_i{}^g &= \lim_{k\to+\infty} \lambda_i{}^{g_k} \\
&= \lim_{k\to+\infty} \frac{1}{2}\mu_k\left(\frac{g_i(x_k)}{\sqrt{g_i(x_k)^2 + \rho_k}} + 1\right) \\
&= 0 \text{ by boundedness of } \mu_k.
\end{aligned}$$

We next show that $\lambda_i^H = 0$, for any $i \in \alpha$. For each $i \in \alpha$, we have $\bar{z}_i > 0$ which implies that for large enough k, $z_i^k > 0$. By the complementarity constraint $y_i^k z_i^k = 0$, we have $y_i^k = 0$, for large enough k. Thus $i \in I_1^k$ for k large enough. Note that $\lambda^{H_k} = \lambda^{z_k}$. By (3.6) and (3.7), we have $\lambda_i^{z_k} = 0$ which implies that $\lambda_i^H = \lim_{k\to+\infty} \lambda_i^{z_k} = 0$.

We prove that $\lambda_i^G = 0$, for any $i \in \gamma$, similarly.

Finally we prove that either $\lambda_i^G \lambda_i^H = 0$ or $\lambda_i^G > 0$, $\lambda_i^H > 0$, for any $i \in \beta$. If $\lambda_i^G \lambda_i^H = 0$ for any $i \in \beta$ then the proof is completed. Suppose that there is a $i \in \beta$, such that $\lambda_i^G \lambda_i^H \neq 0$. Note that $\lambda^{G_k} = \lambda^{y_k}$ and $\lambda^{H_k} = \lambda^{z_k}$. Then for k large enough,

$$\lambda_i^{y_k} \neq 0, \quad \lambda_i^{z_k} \neq 0$$

which implies that $i \in I_2^k$. Consequently, by (3.7),

$$\lambda_i^{y_k} > 0, \quad \lambda_i^{z_k} > 0.$$

Therefore,

$$\lambda_i^G = \lim_{k \to +\infty} \lambda_i^{y_k} > 0, \quad \lambda_i^H = \lim_{k \to +\infty} \lambda_i^{z_k} > 0.$$

Thus $\bar{x}$ is an M-stationary point of MPEC. ■

Notice that the boundedness of the penalty parameters set $\{\mu_k\}$ can guarantee an accumulated point to be an M-stationary point. The following result gives a sufficient condition for the boundedness of $\{\mu_k\}$. First we extend NNAMCQ and MPEC GMFCQ in the following definitions to accommodate a weak feasible point that may not be feasible to MPEC. Note that the definitions of the extended NNAMCQ and the extended MPEC GMFCQ coincide with NNAMCQ and MPEC GMFCQ at a feasible point, respectively.

Definition 3.1 (Extended NNAMCQ) *Let $(\bar{x}, \bar{y}, \bar{z})$ be a weak feasible solution. We say that the extended no nonzero abnormal multiplier constraint qualification (Extended NNAMCQ) is satisfied at $(\bar{x}, \bar{y}, \bar{z})$ if there is no nonzero vector $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that*

$$0 = \sum_{i \in I_{g+P}} \lambda_i^g \nabla g_i(\bar{x}) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(\bar{x}) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(\bar{x}) + \lambda_i^H \nabla H_i(\bar{x})], \tag{3.14}$$

$$\lambda_{I_{g+P}}^g \geq 0, \quad \lambda_{\gamma'}^G = 0, \quad \lambda_{\alpha'}^H = 0. \tag{3.15}$$

and the following condition hold:

$$\forall i \in \beta', \quad \text{either } \lambda_i^G > 0, \ \lambda_i^H > 0 \quad \text{or} \quad \lambda_i^G \lambda_i^H = 0,$$

where

$$I_{g+P} := \{i : g_i(\bar{x}) + P_i = 0\}, \quad P_i = -\max\{g_i(\bar{x}), 0\} \quad i = 1, \dots, p,$$

$$\alpha' := \alpha'(\bar{x}, \bar{y}, \bar{z}) := \{i : \bar{y}_i = 0, \bar{z}_i > 0\},$$

$$\beta' := \beta'(\bar{x}, \bar{y}, \bar{z}) := \{i : \bar{y}_i = 0, \bar{z}_i = 0\},$$

$$\gamma' := \gamma'(\bar{x}, \bar{y}, \bar{z}) := \{i : \bar{y}_i > 0, \bar{z}_i = 0\}.$$

Definition 3.2 (Extended MPEC GMFCQ) *Let $(\bar{x}, \bar{y}, \bar{z})$ be a weak feasible point. We say that the extended MPEC generalized Mangasarian-Fromovitz constraint qualification (Extended MPEC GMFCQ) is satisfied at $(\bar{x}, \bar{y}, \bar{z})$ if* *(i) for every partition of* β' *into sets* P, Q, R *with* $R \neq \emptyset$*, there exists an* d *such that*

$$\nabla g_i(\bar{x})^\top d \leq 0 \quad \forall i \in I_{g+P},$$

$$\nabla h_i(\bar{x})^\top d = 0 \quad \forall i = 1, 2, \ldots, q,$$

$$\nabla G_i(\bar{x})^\top d = 0 \quad \forall i \in \alpha' \cup Q,$$

$$\nabla H_i(\bar{x})^\top d = 0 \quad \forall i \in \gamma' \cup P,$$

$$\nabla G_i(\bar{x})^\top d \geq 0, \nabla H_i(\bar{x})^\top d \geq 0 \quad i \in R$$

and for some $i \in R$ *either* $\nabla G_i(\bar{x})^\top d > 0$ *or* $\nabla H_i(\bar{x})^\top d > 0$*;*

(ii) for every partition of β' *into sets* P, Q*, the gradient vectors*

$$\nabla h_i(\bar{x}) \quad \forall i = 1, 2, \ldots, q,$$

$$\nabla G_i(\bar{x}) \quad \forall i \in \alpha' \cup Q,$$

$$\nabla H_i(\bar{x}) \quad \forall i \in \gamma' \cup P$$

are linearly independent and there exists an $d \in R^n$ *such that*

$$\nabla g_i(\bar{x})^\top d < 0 \quad \forall i \in I_{g+P},$$

$$\nabla h_i(\bar{x})^\top d = 0 \quad \forall i = 1, 2, \ldots, q,$$

$$\nabla G_i(\bar{x})^\top d = 0 \quad \forall i \in \alpha' \cup Q$$

$$\nabla H_i(\bar{x})^\top d = 0 \quad \forall i \in \gamma' \cup P.$$

Proposition 3.5 *Assume that the extended MPEC GMFCQ (equivalently the extended NNAMCQ) is satisfied at any accumulated point $(\bar{x}, \bar{y}, \bar{z})$ of the sequence generated by Algorithm 3.1, then the exact penalty parameter sequence is bounded.*

Proof. The equivalence of the extended MPEC GMFCQ and the extended NNAMCQ follows similarly as in the case of MPEC GMFCQ and NNAMCQ (see [22, Proposition 2.1]). Suppose that the exact penalty parameter sequence μ_k is unbounded. Without loss of generality assume that μ_k is strictly increasing and $\lim_{k \to \infty} \mu_k = +\infty$. Let

$$\widetilde{\lambda}_i^{G_k} = \lambda_i^{G_k}/\mu_{k+1}, \quad \widetilde{\lambda}_i^{H_k} = \lambda_i^{H_k}/\mu_{k+1}, \quad \widetilde{\lambda}_i^{h_k} = \lambda_i^{h_k}/\mu_{k+1}, \quad \widetilde{\lambda}_i^{g_k} = \lambda_i^{g_k}/\mu_{k+1}. \tag{3.16}$$

It is clear that $\widetilde{\lambda}_i{}^{G^k}, \widetilde{\lambda}_i^{H_k}, \widetilde{\lambda}_i^{h_k}$ and $\widetilde{\lambda}_i^{g_k}$ are all bounded by 1. Dividing (3.8) by μ_{k+1} and substituting (3.16) in (3.8), we have

$$\begin{aligned} 0 = {} & \frac{\nabla f(x_k)}{\mu_{k+1}} - \sum_{i=1}^{m} \widetilde{\lambda}_i^{G_k} \nabla G_i(x_k) - \sum_{i=1}^{m} \widetilde{\lambda}_i^{H_k} \nabla H_i(x_k) \\ & + \sum_{i=1}^{q} \widetilde{\lambda}_i^{h_k} \nabla h_i(x_k) + \sum_{i=1}^{p} \widetilde{\lambda}_i^{g_k} \nabla g_i(x_k). \end{aligned} \tag{3.17}$$

Let

$$\lambda_i^G = \lim_{k \to +\infty} \widetilde{\lambda}_i^{G_k}, \ \lambda_i^H = \lim_{k \to +\infty} \widetilde{\lambda}_i^{H_k}, \ \lambda_i^h = \lim_{k \to +\infty} \widetilde{\lambda}_i^{h_k}, \ \lambda_i^g = \lim_{k \to +\infty} \widetilde{\lambda}_i^{g_k}$$

Taking limit in (3.17), we have

$$0 = -\sum_{i=1}^{m} \lambda_i^G \nabla G_i(\bar{x}) - \sum_{i=1}^{m} \lambda_i^H \nabla H_i(\bar{x}) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(\bar{x}) + \sum_{i=1}^{p} \lambda_i^g \nabla g_i(\bar{x}).$$

As in the proof of Theorem 3.4, we can prove that

$$\lambda_i^g = 0 \quad for \quad i \notin I_{g+P}, \quad \lambda_{I_{g+P}}^g \geq 0,$$

$$\lambda_{\alpha'}^H = 0, \quad \lambda_{\gamma'}^G = 0,$$

$$\text{either } \lambda_i^G \lambda_i^H = 0 \text{ or } \lambda_i^G > 0, \lambda_i^H > 0, \forall i \in \beta'.$$

By the extended NNAMCQ, we have that

$$\lambda^G = \lambda^H = \lambda^g = \lambda^h = 0. \tag{3.18}$$

However,

$$\begin{aligned}
&\sum_{i=1}^{m} |\lambda_i^G| + \sum_{i=1}^{m} |\lambda_i^H| + \sum_{i=1}^{q} |\lambda_i^h| + \sum_{i=1}^{p} |\lambda_i^g| \\
= &\sum_{i=1}^{m} |\lim_{k\to+\infty} \widetilde{\lambda}_i^{G_k}| + \sum_{i=1}^{m} |\lim_{k\to+\infty} \widetilde{\lambda}_i^{H_k}| + \sum_{i=1}^{q} |\lim_{k\to+\infty} \widetilde{\lambda}_i^{h_k}| + \sum_{i=1}^{p} |\lim_{k\to+\infty} \widetilde{\lambda}_i^{g_k}| \\
= &\lim_{k\to+\infty} [\ \sum_{i=1}^{m} |\widetilde{\lambda}_i^{G_k}| + \sum_{i=1}^{m} |\widetilde{\lambda}_i^{H_k}| + \sum_{i=1}^{q} |\widetilde{\lambda}_i^{h_k}| + \sum_{i=1}^{p} |\widetilde{\lambda}_i^{g_k}|\ ] \\
= &\lim_{k\to+\infty} \frac{1}{\mu_{k+1}} [\ \sum_{i=1}^{m} |\lambda_i^{G_k}| + \sum_{i=1}^{m} |\lambda_i^{H_k}| + \sum_{i=1}^{q} |\lambda_i^{h_k}| + \sum_{i=1}^{p} |\lambda_i^{g_k}|\ ] \\
= &\lim_{k\to+\infty} \frac{\mu_{k+1} - \alpha_2 \mu_k}{\mu_{k+1}} \\
\geq &\lim_{k\to+\infty} \frac{\mu_{k+1}(1-\alpha_2)}{\mu_{k+1}} \\
= &\ 1 - \alpha_2 > 0.
\end{aligned}$$

This contradicts to (3.18). Therefore, $\{\mu_k\}$ must be bounded. ■

Chapter 4

A Hybrid Algorithm for Solving the Subproblems

The purpose of this chapter is to introduce a hybrid algorithm for subproblems $FMPEC^{\rho}_{\mu}$ generated in Algorithm 3.1 and present its convergent properties.

4.1 A Hybrid Algorithm

In the Step 2 of Algorithm 3.1, we need to find an S-stationary point of subproblem $FMPEC^{\rho}_{\mu}$ for the given pair (μ, ρ). Some reported algorithms can be used to solve it. See [4, 6, 18, 26] for references. Alternatively, since the MPEC LICQ naturally holds for $FMPEC_\mu^\rho$, one can use the combination of the $\delta-$active search technique and the active set technique (see [9, 26] for the references) to obtain a piecewise stationary point which is equivalent to an S-stationary point under the MPEC LICQ . We now describe the hybrid algorithm.

Note that the feasible region $\mathcal{WF}$ of the problem $FMPEC_\mu^\rho$ is the union of *weak faces*

$$\mathcal{WF} = \bigcup_{(A,B)\in\mathcal{P}} \mathcal{WF}_{(A,B)},$$

where

$$\mathcal{WF}_{(A,B)} = \left\{ (x,y,z) : \begin{array}{lll} y_i = 0, & z_i \geq 0 & \forall i \in A \\ y_i \geq 0, & z_i = 0 & \forall i \in B \end{array} \right\}$$

and $\mathcal{P}$ is the set of all pairs (A, B) partitioning $\{1, 2, \cdots, m\}$. Let

$$I(y) := \{i : y_i = 0\}, \tag{4.1}$$

$$I(z) := \{i : z_i = 0\}. \tag{4.2}$$

Definition 4.1 *Let $(\tilde{x}, \tilde{y}, \tilde{z})$ be any weak feasible point. We call a pair $(A, B) \in \mathcal{P}$ an adjacent pair at $(\tilde{x}, \tilde{y}, \tilde{z})$ if*

$$A \subseteq I(\tilde{y}) \quad \text{and} \quad B \subseteq I(\tilde{z})$$

and denote by $\mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ *the set of all adjacent pairs at* $(\tilde{x},\tilde{y},\tilde{z})$.

The hybrid algorithm is also an iterative algorithm. We state the scheme below. Suppose that $(\tilde{x},\tilde{y},\tilde{z}) \in \mathcal{WF}$ is an iterative point but it is not an S-stationary point of $FMPEC^{\rho}_{\mu}$. Then we will use the active set technique to find a descent pair $(A,B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ such that the value of objective function $\phi^{\rho}_{\mu}(\tilde{x},\tilde{y},\tilde{z})$ decrease in the piece $\mathcal{WF}_{(\mathcal{A},\mathcal{B})}$. To find such a descent pair, first, we choose an adjacent pair $(A,B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ randomly and consider the following quadratic programming problem:

$$
\begin{aligned}
QMPEC_{(A,B)}(\tilde{x},\tilde{y},\tilde{z}) \qquad \min_{d} \quad & \nabla\phi^{\rho}_{\mu}(\tilde{x},\tilde{y},\tilde{z}) \cdot d + \frac{1}{2} d^{\top} D d \\
s.t. \quad & (d_y)_i + \tilde{y}_i = 0, \quad (d_z)_i + \tilde{z}_i \geq 0 \quad i \in A, \\
& (d_y)_i + \tilde{y}_i \geq 0, \quad (d_z)_i + \tilde{z}_i = 0 \quad i \in B,
\end{aligned}
$$

where D is a positive definite matrix and $d := (d_x, d_y, d_z) \in R^{n+2m}$. If the solution d of $QMPEC_{(A,B)}(\tilde{x},\tilde{y},\tilde{z})$ is nonzero, then d provides a descent direction of ϕ^{ρ}_{μ} to the weak face $\mathcal{WF}_{(A,B)}$ at $(\tilde{x},\tilde{y},\tilde{z})$ and hence (A,B) can be taken as a descent pair. Otherwise $d = 0$ is the optimal solution of $QMPEC_{(A,B)}(\tilde{x},\tilde{y},\tilde{z})$. Since all constraints are linear and the objective function is strictly convex, $d = 0$ is the optimal solution of $QMPEC_{(A,B)}(\tilde{x},\tilde{y},\tilde{z})$

if and only if there exist Lagrange multipliers λ^y and λ^z such that

$$0 = \nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) - (0, \lambda^y, 0)^\top - (0, 0, \lambda^z)^\top$$

$$\lambda_i{}^z \geq 0, \quad \tilde{z}_i \cdot \lambda_i{}^z = 0, \quad \forall i \in A$$

$$\lambda_i{}^y \geq 0, \quad \tilde{y}_i \cdot \lambda_i{}^y = 0, \quad \forall i \in B.$$

If $\lambda_i^y \geq 0$ and $\lambda_i^z \geq 0$ for any $i \in I_2 := \{i : \tilde{y}_i = 0,\ \tilde{z}_i = 0\}$, then $(\tilde{x}, \tilde{y}, \tilde{z})$ is an S-stationary point of $FMPEC_\mu^\rho$. Suppose that there is an $i \in I_2$ such that $\lambda_i^y < 0$ or $\lambda_i^z < 0$. Let

$$\begin{cases} \tilde{A} = A\backslash\{i\}, \quad \tilde{B} = B \cup \{i\}, \quad if \ \ \lambda_i^y < 0, \\ \tilde{A} = A \cup \{i\}, \quad \tilde{B} = B\backslash\{i\}, \quad otherwise, \end{cases}$$

then zero is not a stationary point of the problem $QMPEC_{(\tilde{A},\tilde{B})}(\tilde{x}, \tilde{y}, \tilde{z})$. Therefore, zero is not an optimal solution of it. Hence $(\tilde{A}, \tilde{B})$ can be taken as a descent pair. Now we have shown that if the current iterative point $(\tilde{x}, \tilde{y}, \tilde{z}) \in \mathcal{WF}$ is not a S-stationary point of $FMPEC_\mu^\rho$, then we can always find a descent pair $(\tilde{A}, \tilde{B})$ such that the nonzero solution $\tilde{d}$ to the quadratic program $QMPEC_{(\tilde{A},\tilde{B})}$ provides a descent direction on the weak face $\mathcal{WF}_{(\tilde{A},\tilde{B})}$. We then find the next iterative point along the descent direction $\tilde{d}$ by using Armijo's inexact line search method.

As explained in [9], if we only search a descent direction on one of the weak faces, the algorithm may find a stationary point in this weak face instead of a stationary point of $FMPEC_{\mu}^{\rho}$. In order to search in the other face, we use the δ-active search technique. We make use of a example in [9] as below to show the principle of the δ-active search technique:

$$\begin{array}{ll} \min\limits_{y,z} & f(y,z) := (y^2 - z + 1)^2 \\ s.t. & y \geq 0,\ z \geq 0,\ yz = 0. \end{array}$$

This problem is already in the form of FMPEC. The feasible region is the union of following two faces:

$$\{(y,z) : y \geq 0, z = 0\} \tag{4.3}$$

and

$$\{(y,z) : y = 0, z \geq 0\}. \tag{4.4}$$

Suppose the current iteration point (y_k, z_k) lies in the face (4.3), that is, $y_k \geq 0, z_k = 0$. To move to the other face from this face, we search the nearby point with more active constraints from the iterative point. Notice that $(0,0)$ is the only point which has more active constraints than (y_k, z_k) and lies in the both of faces. We find a descent direction $d = (0,1)$ of the function $f(y,z) = (y^2 - z + 1)^2$ on the face (4.4) by solving the quadratic subproblem associated with the face (4.4). After finding an appropriate step size p by Armijo-type inexact line search, we compare the values of objective function ϕ_μ^ρ at the current iteration point (y_k, z_k) and $(y_k + pd_k^y, z_k + pd_k^z)$ and set the next iterative point to be the one with the smaller value of ϕ_μ^ρ. In fact, the optimal solution of the problem happens to be $(0, 1)$.

The algorithm described above can not guarantee global convergence without further assumptions and is conceptual since the Lagrange multipliers can be computed only approximately. In order to make it implementable and globally convergent, we will make use of the ϵ-active set technique.

First we give the definition of an approximate S-stationary point for the subproblem

Definition 4.2 *Given $\epsilon > 0$, a feasible point $(\tilde{x}, \tilde{y}, \tilde{z})$ of $FMPEC_\mu^\rho$ is called an $\epsilon-$approximate S-stationary point if there are multipliers λ^y, λ^z such that*

$$\|\nabla\phi_\mu^\rho(\bar{x}, \bar{y}, \bar{z}) - (0, \lambda^y, 0) - (0, 0, \lambda^z)\|_1 \leq \frac{1}{2}\epsilon,$$

$$\lambda_i^y = 0,\ i \notin I^\epsilon(\tilde{y}),\quad \lambda_i^z = 0,\ i \notin I^\epsilon(\tilde{z}),$$

$$\lambda_i^y \geq 0,\ i \in I^\epsilon(\tilde{y}) \setminus I(\tilde{y}),$$

$$\lambda_i^z \geq 0,\ i \in I^\epsilon(\tilde{z}) \setminus I(\tilde{z}),$$

$$\lambda_i^y \geq -\epsilon,\ \lambda_i^y \geq -\epsilon,\ \textit{for any } i \in I(\tilde{y}) \cap I(\tilde{z}),$$

where

$$I^\epsilon(\tilde{y}) := \{i : \tilde{y} \leq \epsilon\},\ I^\epsilon(\tilde{z}) := \{i : \tilde{z} \leq \epsilon\},$$

and $I(\tilde{y})$, $I(\tilde{z})$ are defined in (4.1), (4.2), respectively.

We now describe the ϵ-active set technique for finding a decent pair as follows:

For an $\epsilon > 0$ and a randomly chosen pair $(A, B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$, assume that $(\tilde{x}, \tilde{y}, \tilde{z})$ is not an $\epsilon-$approximate S-stationary point of $FMPEC^{\rho}_{\mu}$. Then the following three cases may happen. The first case is that the optimal solution d of $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$ may satisfy $\|d\|_\infty \geq \frac{1}{2}\epsilon$. In this case, $(\tilde{A}, \tilde{B}) = (A, B)$ can be taken as a descent pair since $d \neq 0$. The second case is that $\|d\|_\infty < \frac{1}{2}\epsilon$ and $\|\bar{r}\|_1 \geq \frac{1}{2}\epsilon$, where

$$\bar{r} := \nabla\phi^{\rho}_{\mu}(\tilde{x}, \tilde{y}, \tilde{z}) - (0, \lambda^y, 0)^\top - (\lambda^z, 0, 0)^\top,$$

and (λ^y, λ^z) is the multipliers of the quadratic problem $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$. Then the optimal solution of the quadratic problem $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$ is still not zero. We still take $(\tilde{A}, \tilde{B}) = (A, B)$ as a descent pair in this case. The last case is that $\|\bar{r}\|_1 < \frac{1}{2}\epsilon$ and $\|d\|_\infty < \frac{1}{2}\epsilon$. Since $(\tilde{x}, \tilde{y}, \tilde{z})$ is not an $\epsilon-$approximate S-stationary point of $FMPEC^{\rho}_{\mu}$, there is an $\bar{i} \in I(\tilde{y}) \cap I(\tilde{z})$ such that

$$\lambda^y_{\bar{i}} < -\epsilon, \text{ or } \lambda^z_{\bar{i}} < -\epsilon.$$

Let

$$\begin{cases} \tilde{A} = A \backslash \{\bar{i}\}, \quad \tilde{B} = B \cup \{\bar{i}\}, \quad if \ \lambda_{\bar{i}}^{y} < -\epsilon, \\ \tilde{A} = A \cup \{\bar{i}\}, \quad \tilde{B} = B \backslash \{\bar{i}\}, \quad otherwise. \end{cases}$$

and take $(\tilde{A}, \tilde{B})$ as a descent pair. From the above process, we can either find a descent pair or conclude that the current iteration point $(\tilde{x}, \tilde{y}, \tilde{z})$ is an $\epsilon-$approximate S-stationary point of $FMPEC_{\mu}^{\rho}$. In Theorem 4.4, we prove that the descent pair that we choose must be an acceptable pair which guarantee the global convergence.

Now we are ready to state the algorithm for the subproblem.

Algorithm 4.1

Step 1. (Initialization) Given $(\bar{x}_0, \bar{y}_0, \bar{z}_0) \in \mathcal{WF}$, $0 < \sigma < 1$, $\delta_0 > 0$, $0 < \epsilon$, $0 < p < 1$, $0 < c_2 < 1$, and a symmetric positive definite matrix $D_1 \in R^{(n+2m)\times(n+2m)}$. Let $k := 0$.

Step 2. Let $\delta = \delta_0$ and $(x_{temp}, y_{temp}, z_{temp}) = (\bar{x}_k, \bar{y}_k, \bar{z}_k)$.

Step 3. Let

$$\begin{aligned} \tilde{x} &= \bar{x}_k \\ (\tilde{y})_i &= \begin{cases} 0, & (\bar{y}_k)_i \le \delta \\ (\bar{y}_k)_i, & otherwise \end{cases} \qquad i = 1, \cdots, m \end{aligned}$$

$$(\tilde{z})_i = \begin{cases} 0, & (\bar{z}_k)_i \leq \delta \\ (\bar{z}_k)_i, & otherwise \end{cases} \quad i = 1, \cdots, m$$

Step 4. Using the ϵ-active set technique, find a descent pair $(\tilde{A}, \tilde{B}) \in \mathcal{P}(\tilde{x}, \tilde{y}, \tilde{z})$ or conclude that $(\tilde{x}, \tilde{y}, \tilde{z})$ is an $\epsilon-$approximate S-stationary point of $FMPEC^{\rho}_{\mu}$.

Step 5. Find a solution $\tilde{d}$ to the problem $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$. Let m be the smallest nonnegative integer such that

$$\phi^{\rho}_{\mu}(\tilde{x}, \tilde{y}, \tilde{z}) - \phi^{\rho}_{\mu}((\tilde{x}, \tilde{y}, \tilde{z}) + p^m \tilde{d}) \geq \frac{1}{2} \sigma p^m (\tilde{d}^{\top} D_k \tilde{d}).$$

If

$$\phi^{\rho}_{\mu}(x_{temp}, y_{temp}, z_{temp}) > \phi^{\rho}_{\mu}((\tilde{x}, \tilde{y}, \tilde{z}) + p^m \tilde{d}),$$

set

$$(x_{temp}, y_{temp}, z_{temp}) = (\tilde{x}, \tilde{y}, \tilde{z}) + p^m \tilde{d}.$$

Step 6. If $\delta > 0$, let

$$\delta = c_2 \max\{(\bar{y}_k)_i, (\bar{y}_k)_i < \delta; \quad (\bar{z}_k)_i, (\bar{z}_k)_i < \delta\}$$

and go to Step 3.

Step 7. If $\delta = 0$, let

$$(\bar{x}_{k+1}, \bar{y}_{k+1}, \bar{z}_{k+1}) = (x_{temp}, y_{temp}, z_{temp}),$$

adjust D_k. Set $k = k + 1$ and go to Step 2.

Remark 4.1 *We may use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) rule or other rules to update* D_k.

4.2 Convergence of the Hybrid Algorithm

First we state the following blank assumptions in order to guarantee the global convergence.

(A1) There are constants $\beta_1 > 0$ and $\beta_2 > \beta_1$ such that $\beta_1 d^\top d \le d^\top D_k d \le \beta_2 d^\top d$ for any k and any $d \in R^{n+m+m}$.

(A2) The level set $\{(x, y, z) : \phi_\mu^\rho(x, y, z) \le \phi_\mu^\rho(\bar{x}_1, \bar{y}_1, \bar{z}_1)\}$ is bounded.

If the quadratic subproblem $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$ generates a solution d with the smallest value among all adjacent pairs $(A, B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$, then we call this descent direction d the steepest descent direction.

Definition 4.3 (steepest descent direction) *Let* $(\tilde{x}, \tilde{y}, \tilde{z})$ *be a weak feasible solution. If* $(\ddot{A}, \ddot{B}) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ *generates a solution* $\ddot{d}$ *to* $QMPEC_{(\ddot{A},\ddot{B})}(\tilde{x}, \tilde{y}, \tilde{z})$ *such that*

$$\nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) \cdot \ddot{d} + \frac{1}{2}\ddot{d}^\top D\ddot{d} \le \nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) \cdot d + \frac{1}{2}d^\top Dd$$

for all adjacent pairs $(A, B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ *and all solutions* d *of problem* $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$. *Then* $\ddot{d}$ *is called the steepest descent direction of* $FMPEC_\mu^\rho$ *at* $(\tilde{x}, \tilde{y}, \tilde{z})$ *and* $(\ddot{A}, \ddot{B})$ *is called a steepest descent pair of* $FMPEC_\mu^\rho$ *at* $(\tilde{x}, \tilde{y}, \tilde{z})$.

Ideally, the objective function $\phi_\mu^\rho(x, y, z)$ will descend the most from the point $(\tilde{x}, \tilde{y}, \tilde{z})$ along the steepest descent direction. However, finding the steepest descent direction requires checking all the adjacent pairs. This may be very time-consuming. If c is the cardinality of the degenerate set β, then 2^c quadratic programming problems may need to be checked to verify or disprove that steepest descendence. Indeed, we only need to obtain an adjacent pair $(\tilde{A}, \tilde{B}) \in \mathcal{P}_{(\bar{x},\bar{y},\bar{z})}$ with the solution $\tilde{d}$ to the problem $QMPEC_{(\tilde{A},\tilde{B})}(\tilde{x}, \tilde{y}, \tilde{z})$ which satisfies the following inequality

$$\nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) \cdot \tilde{d} + \frac{1}{2}\tilde{d}^\top D\tilde{d} \leq c_1(\nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) \cdot \ddot{d} + \frac{1}{2}\ddot{d}^\top D\ddot{d}), \tag{4.5}$$

for a given $0 < c_1 \leq 1$. We call such an adjacent pair $(\tilde{A}, \tilde{B})$ an *acceptable descent pair* and such a descent direction $\tilde{d}$ an *acceptable descent direction*, correspondingly. According to [9], we only need to verify that each descent pair obtained in step 4 is an *acceptable* descent pair to guarantee the global convergence of Algorithm 4.1.

The following result follows from the process of Algorithm 4.1.

Theorem 4.1 *If Algorithm 4.1 finitely stops, it must stop at an $\epsilon-$approximate S-stationary point of $FMPEC^{\rho}_{\mu}$.*

Since we use the δ-active set search technique and the Armijo's inexact line search rule in Algorithm 4.1, we can derive the following property similarly to [9, Algorithm 2.1]. We sketch the proof here which is slightly different with the one in [9, Theorem 3.1].

Theorem 4.2 *Assume that the assumptions $(A1)$ and $(A2)$ hold and each descent pair $(\tilde{A}, \tilde{B})$ obtained in step 4 of Algorithm 4.1 is an acceptable pair for some $c_1 > 0$. If the sequence of iterative points generated by Algorithm 4.1 is infinite, then each cluster point of the iterative sequence generated by Algorithm 4.1 is an S-stationary point of $FMPEC^{\rho}_{\mu}$.*

Proof. Suppose that $(\bar{x}, \bar{y}, \bar{z})$ is an accumulation point of the sequence $\{(x_k, y_k, z_k)\}$ generated by Algorithm 4.1. By assumption (A2), there must be a subsequence convergent to $(\bar{x}, \bar{y}, \bar{z})$. Without loss of the generality, we assume

$$\lim_{k \to \infty} \{(x_k, y_k, z_k)\} = (\bar{x}, \bar{y}, \bar{z}). \tag{4.6}$$

Similarly, by assumption (A1), there exits a positive definite matrix D such that

$$\lim_{k\to\infty} D_k = D \tag{4.7}$$

Let

$$I_1(y) = \{i|(y)_i = 0\},\ I_2(z) = \{i|(z)_i = 0\},$$

for given $y, z \in R^m$. Set

$$\begin{aligned} \tilde{x}_k &= x_k, \\ (\tilde{y}_k)_i &= \begin{cases} 0, & i \in I_1(\bar{y}), \\ (y_k)_i, & i \notin I_1(\bar{y}), \end{cases} \\ (\tilde{z}_k)_i &= \begin{cases} 0, & i \in I_2(\bar{z}), \\ (z_k)_i, & i \notin I_2(\bar{z}). \end{cases} \end{aligned}$$

Then it is clear that

$$\lim_{k\to\infty} (\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) = (\bar{x}, \bar{y}, \bar{z}). \tag{4.8}$$

Let $(\tilde{A}_k, \tilde{B}_k)$ be an acceptable descent pair found in the algorithm at $(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)$ and $\tilde{d}_k$ be the solution to the problem $QMPEC_{(\tilde{A}_k,\tilde{B}_k)}(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)$. We next show that

$$\lim_{k\to\infty} \tilde{d}_k = \tilde{d}.$$

Since there are only finite adjacent pairs in $\mathcal{P}_{(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k)}$ and (4.8) holds, without loss of the generality, there is a partition pair $(\tilde{A},\tilde{B})$ and an integer k_0 such that

$$\tilde{A}_k = \tilde{A}, \quad \tilde{B}_k = \tilde{B}.$$

for any $k > k_0$. By [9, Lemma 3.1], there is a δ^k used in the δ-active set searching of the kth iteration such that

$$I_1(\bar{y}) = I_1^k(\delta^k), \ I_2(\bar{z}) = I_2^k(\delta^k),$$

for any $k > k_0$, where

$$I_1^k(\delta^k) = \{i: \ (y_k)_i \leq \delta^k\}, \ \ I_2^k(\delta^k) = \{i: \ (z_k)_i \leq \delta^k\}.$$

Note that

$$\tilde{A}_k \subseteq I_1^k(\delta), \quad \tilde{B}_k \subseteq I_2^k(\delta).$$

for any $\delta \geq 0$ and $k > k_0$. Consequently,

$$\tilde{A} \subseteq I_1(\bar{y}), \quad \tilde{B} \subseteq I_2(\bar{z}),$$

i.e. $(\tilde{A},\tilde{B})$ is an adjacent pair at the point $(\bar{x},\bar{y},\bar{z})$.

Let $\tilde{d}$ be the solution to the problem $QMPEC_{(\tilde{A},\tilde{B})}(\bar{x},\bar{y},\bar{z})$. Notice that the LICQ is satisfied in the feasible region of the problem $QMPEC_{(\tilde{A},\tilde{B})}(\bar{x},\bar{y},\bar{z})$

and the quadratic subproblem is stable under perturbation by the sensitivity and stability theory (see e.g. [19]), we have

$$\lim_{k\to\infty} \tilde{d}_k = \tilde{d}. \tag{4.9}$$

We will finish the proof by contradiction. Assume that $(\bar{x}, \bar{y}, \bar{z})$ is not a piece-wise stationary point of FMPEC. Then we can find a adjacent pair $(\bar{A}, \bar{B}) \in \mathcal{P}_{(\bar{x},\bar{y},\bar{z})}$ such that $d = 0$ is not the optimal solution of $QMPEC_{(\bar{A},\bar{B})}(\bar{x}, \bar{y}, \bar{z})$. Since $d = 0$ is feasible for $QMPEC_{(\bar{A},\bar{B})}(\bar{x}, \bar{y}, \bar{z})$, the optimal solution $\bar{d}$ to $QMPEC_{(\bar{A},\bar{B})}(\bar{x}, \bar{y}, \bar{z})$ must satisfy

$$\nabla\phi_\mu^\rho((\bar{x}, \bar{y}, \bar{z})) \cdot \bar{d} + \frac{1}{2}\bar{d}^\top D\bar{d} < 0. \tag{4.10}$$

Let $\bar{d}_k$ be the solution to the problem $QMPEC_{(\bar{A},\bar{B})}(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)$. Then we must have

$$\lim_{k\to\infty} \bar{d}_k = \bar{d}. \tag{4.11}$$

By (4.5), we have

$$\begin{aligned} \nabla\phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) \cdot \tilde{d}_k + \frac{1}{2}\tilde{d}_k^\top D_k\tilde{d}_k &\le c_1(\nabla\phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) \cdot \ddot{d}_k + \frac{1}{2}\ddot{d}_k^\top D_k\ddot{d}_k) \\ &\le c_1(\nabla\phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) \cdot \bar{d} + \frac{1}{2}\bar{d}_k^\top D_k\bar{d}_k), \end{aligned} \tag{4.12}$$

for any large enough k. By (4.8)-(4.11), taking the limits on both sides of (4.12) we have

$$\begin{aligned}\nabla\phi_\mu^\rho(\bar{x},\bar{y},\bar{z})\cdot\tilde{d}+\tfrac{1}{2}\tilde{d}^\top D\tilde{d} &\le c_1(\nabla\phi_\mu^\rho(\bar{x},\bar{y},\bar{z})\cdot\bar{d}+\tfrac{1}{2}\bar{d}^\top D\bar{d})\\ &< 0\end{aligned}$$

which implies that $\tilde{d}$ is a descent direction of ϕ_μ^ρ at $(\bar{x},\bar{y},\bar{z})$.

In the following, we will show that there is a contradiction. On one hand, by (4.6), we have

$$\lim_{k\to\infty}\phi_\mu^\rho(x_k,y_k,z_k)=\phi_\mu^\rho(\bar{x},\bar{y},\bar{z}). \tag{4.13}$$

On the other hand, we will show that the next iterative point $(x_{k+1},y_{k+1},z_{k+1})$ provides a positive reduction on the penalty function ϕ_μ^ρ from (x_k,y_k,z_k). Let m be the smallest positive integer satisfying

$$\phi_\mu^\rho(\bar{x},\bar{y},\bar{z})-\phi_\mu^\rho((\bar{x},\bar{y},\bar{z})+p^m\tilde{d})>\frac{1}{2}\sigma p^m(\tilde{d}^\top D\tilde{d}).$$

Then there exists k_1 such that for all $k\ge k_1$,

$$\phi_\mu^\rho(\tilde{x}_k,\tilde{y}_k,\tilde{z}_k)-\phi_\mu^\rho((\tilde{x}_k,\tilde{y}_k,\tilde{z}_k)+p^m\tilde{d}_k)>\frac{1}{2}\sigma p^m(\tilde{d}_k^\top D\tilde{d}_k)$$

which implies that

$$m_k\le m \qquad \forall k\ge k_1, \tag{4.14}$$

where m_k is the smallest positive integer m satisfying

$$\phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) - \phi_\mu^\rho((\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) + p^{m_k}\tilde{d}_k) > \frac{1}{2}\sigma p^{m_k}(\tilde{d}_k^\top D_k \tilde{d}_k). \tag{4.15}$$

Therefore, there is a $k_2 \geq k_1$ such that

$$\begin{aligned}
&\phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) - \phi_\mu^\rho((\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) + p^{m_k}\tilde{d}_k) \\
&> \frac{1}{2}\sigma p^{m_k}(\tilde{d}_k^\top D_k \tilde{d}_k) \qquad \text{by (4.15)} \\
&\geq \frac{1}{2}\sigma p^{m}(\tilde{d}_k^\top D_k \tilde{d}_k) \qquad \text{by(4.14)} \\
&\geq \frac{1}{4}\sigma p^{m}(\tilde{d}^\top D \tilde{d}) \qquad (4.16)
\end{aligned}$$

for any $k \geq k_2$. By (4.6)–(4.8), there is a $k_3 \geq k_2$ such that for any $k > k_3$, we have

$$|\phi_\mu^\rho(x_k, y_k, z_k) - \phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)| < \frac{1}{8}\sigma p^m(\tilde{d}^\top D \tilde{d}). \tag{4.17}$$

By (4.16) and (4.17),

$$\begin{aligned}
&\phi_\mu^\rho(x_k, y_k, z_k) - \phi_\mu^\rho((\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) + p^{m_k}\tilde{d}_k) \\
&= \phi_\mu^\rho(x_k, y_k, z_k) - \phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) + \phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) - \phi_\mu^\rho((\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) + p^{m_k}\tilde{d}_k) \\
&> -\frac{1}{8}\sigma p^m(\tilde{d}^\top D \tilde{d}) + \frac{1}{4}\sigma p^m(\tilde{d}^\top D \tilde{d}) \\
&= \frac{1}{8}\sigma p^m(\tilde{d}^\top D \tilde{d})
\end{aligned}$$

for any $k \geq k_3$. Since we choose the next iterative point $(x_{k+1}, y_{k+1}, z_{k+1})$ which can provide the largest reduction, we have

$$\begin{aligned}
&\phi_\mu^\rho(x_k, y_k, z_k) - \phi_\mu^\rho(x_{k+1}, y_{k+1}, z_{k+1}) \\
&\geq \phi_\mu^\rho(x_k, y_k, z_k) - \phi_\mu^\rho((\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) + p^{m_k}\tilde{d}_k) \\
&\geq \frac{1}{8}\sigma p^m(\tilde{d}^\top D\tilde{d}) \\
&> 0 \qquad \forall k \geq k_4, \qquad (4.18)
\end{aligned}$$

which contradicts to (4.13) and the proof is completed. ■

Moreover we can verify that the descent pair obtained in Step 4 of Algorithm 4.1 is acceptable in Theorem 4.4. Then the premises of Theorem 4.2 satisfy.

First we need the following result.

Lemma 4.3 *Under Assumptions (A1) and (A2) there exists a constant* $M_1 > 0$ *such that*

$$\|\bar{r}\|_1 < M_1\bar{\epsilon}, \quad \|\hat{d}\|_\infty < M_1\bar{\epsilon},$$

for any $0 < \bar{\epsilon} \leq \epsilon$, *all iteration points* $(\tilde{x}, \tilde{y}, \tilde{z})$ *given in Algorithm 4.1 and all* $(A, B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ *with*

$$0 \geq \nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) \cdot \hat{d} + \frac{1}{2}\hat{d}^\top D\hat{d} \geq -\bar{\epsilon}, \qquad (4.19)$$

where $\hat{d}$ *is the optimal solution of* $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$.

Proof. We first prove that there is a constant $M_0 > 0$ and an $\tilde{\epsilon}$ with $0 < \tilde{\epsilon} \leq \epsilon$ such that

$$\|\hat{d}\|_\infty < M_0\bar{\epsilon},$$

for any $0 < \bar{\epsilon} \leq \tilde{\epsilon}$, all $(\tilde{x}, \tilde{y}, \tilde{z})$ and all $(A, B) \in \mathcal{P}_{(\tilde{x},\tilde{y},\tilde{z})}$ with (4.19). To the contrary, suppose that for any integer $k > 0$, there are $(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)$ and $\hat{d}_k$ such that

$$\|\hat{d}_k\|_\infty \geq k \times \frac{\epsilon}{2k},$$

and

$$0 \geq \nabla\phi_\mu^\rho(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k) \cdot \hat{d}_k + \frac{1}{2}\hat{d}_k^\top D_k \hat{d}_k \geq -\frac{\epsilon}{2k}.$$

By assumptions (A_1) and (A_2), there is a subsequence of $\{(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)\}$ and $\{\hat{d}_k\}$ which converge to $(\tilde{x}', \tilde{y}', \tilde{z}')$ and $\hat{d}'$ respectively. Moreover, D_k converges to $\bar{D}$ and $\hat{d}_k$ is the optimal solution of $QMPEC_{(\bar{A},\bar{B})}(\tilde{x}_k, \tilde{y}_k, \tilde{z}_k)$ for a fixed $(\bar{A}, \bar{B})$ because there is only finite number of adjacent pairs. By sensitivity analysis, we have that

$$\nabla\phi_\mu^\rho(\tilde{x}', \tilde{y}', \tilde{z}') \cdot \hat{d}' + \frac{1}{2}\hat{d}'^\top \bar{D}\hat{d}' = 0 \tag{4.20}$$

and

$$\|\tilde{d}'\|_\infty \geq \frac{1}{2}\epsilon \tag{4.21}$$

for the optimal solution $\hat{d}'$ of $QMPEC_{(\tilde{A},\tilde{B})}(\tilde{x}', \tilde{y}', \tilde{z}')$. We have $\hat{d}' = 0$ by (4.20). However, (4.21) implies $\hat{d}' \neq 0$. Hence there is a contradiction. Now by the assumption (A2), we know that the iteration points $(\tilde{x}, \tilde{y}, \tilde{z})$ given in the algorithm is bounded, so are $\hat{d}$. Suppose that the solutions $\hat{d}$ are bounded by a constant $M_b > 0$, then let $M_1 = \max\{M_0, \frac{M_b}{\tilde{\epsilon}}\}$, the conclusion follows.

■

Theorem 4.4 *If assumptions* (A1) *and* (A2) *hold, then for any given* $\epsilon > 0$, *each descent pair* $(\tilde{A}, \tilde{B})$ *obtained in Step 4 of Algorithm 4.1 is an acceptable pair for some* $c_1 > 0$.

Proof. Assume that a weak feasible point $(\tilde{x}, \tilde{y}, \tilde{z})$ is not an $\epsilon-$approximate S-stationary point of $FMPEC_\mu^\rho$. We want to prove that there is a $c_1 > 0$ such that (4.5) holds for any D, any iteration point $(\tilde{x}, \tilde{y}, \tilde{z})$ found in step 4 of Algorithm 4.1, any adjacent pair $(A, B) \in \mathcal{P}(\tilde{x}, \tilde{y}, \tilde{z})$ and any $\tilde{d}$, the solution of $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$.

Now we need to consider the three cases below according to the hybrid algorithm :

Case 1. If $\|\hat{d}\|_\infty \geq \frac{1}{2}\epsilon$, then we have

$$\nabla\phi_\mu^\rho(\tilde{x}, \tilde{y}, \tilde{z}) \cdot \hat{d} + \frac{1}{2}\hat{d}^\top D\hat{d} < -\frac{\epsilon}{2M_1} \tag{4.22}$$

by Lemma 4.3.

Case 2. If $\|\hat{d}\|_\infty < \frac{1}{2}\epsilon$ and $\|\bar{r}\|_1 \geq \frac{1}{2}\epsilon$, we still have (4.22) by Lemma 4.3.

Case 3. Assume that $\|\hat{d}\|_\infty < \frac{1}{2}\epsilon$ and $\|\bar{r}\|_1 < \frac{1}{2}\epsilon$ but $(\tilde{x}, \tilde{y}, \tilde{z})$ is not an ϵ-approximate S-stationary point. Since (λ^y, λ^z) is the multiplier for $QMPEC_{(A,B)}(\tilde{x}, \tilde{y}, \tilde{z})$,

$$0 = \nabla\phi^\rho_\mu(\tilde{x}, \tilde{y}, \tilde{z}) - (0, \lambda^y, 0)^\top - (0, 0, \lambda^z)^\top + D\hat{d}$$

$$\lambda_i^z \geq 0, \quad [(\hat{d}_z)_i + \tilde{z}_i] \cdot \lambda_i^z = 0, \quad \forall i \in A$$

$$\lambda_i^y \geq 0, \quad [(\hat{d}_y)_i + \tilde{y}_i] \cdot \lambda_i^y = 0, \quad \forall i \in B.$$

Consequently, (λ^y, λ^z) satisfies

$$\|\bar{r}\|_1 \leq \frac{1}{2}\epsilon$$

$$\lambda_i^y \geq 0, \ i \in I^\epsilon(\tilde{y}) \setminus I(\tilde{y})$$

$$\lambda_i^z \geq 0, \ i \in I^\epsilon(\tilde{z}) \setminus I(\tilde{z})$$

$$\lambda_i^y = 0, i \notin I^\epsilon(\tilde{y})$$

$$\lambda_i^z = 0, i \notin I^\epsilon(\tilde{z}).$$

Because $(\tilde{x}, \tilde{y}, \tilde{z})$ is not an $\epsilon-$ approximate stationary point of $FMPEC^\rho_\mu$, there is an $\bar{i} \in I(\tilde{y}) \cap I(\tilde{z})$ with

$$\lambda_{\bar{i}}^y < -\epsilon \tag{4.23}$$

or

$$\lambda_{\bar{i}}^{z} < -\epsilon. \tag{4.24}$$

In the cases (4.23) and (4.24), $(\tilde{A}, \tilde{B})$ is taken as

$$\tilde{A} = A \backslash \{\bar{i}\}, \quad \tilde{B} = B \cup \{\bar{i}\}$$

and

$$\tilde{A} = A \cup \{\bar{i}\}, \quad \tilde{B} = B \backslash \{\bar{i}\}$$

respectively.

For the case (4.23), consider the following problems:

$$\begin{cases} \nabla \phi_{\mu}^{\rho}(\tilde{x}, \tilde{y}, \tilde{z}) d \leq -\frac{1}{2}\epsilon \\ -e \leq d \leq e \\ d_{\bar{i}}^{\,y} \geq 0 \\ d_{i}^{y} \geq 0, i \notin I^{\epsilon}(\tilde{y}) \\ d_{i}^{\,y} = 0, i \in I^{\epsilon}(\tilde{y}) \setminus \{\bar{i}\}, \\ d_{i}^{\,z} = 0, \end{cases} \tag{4.25}$$

where $d = (d^x, d^y, d^z)$.

To show that the above linear system has a solution. Assume that it has no solution, then by Farkas's Lemma, the following dual linear system has a solution $(\bar{\lambda}_0, \bar{\lambda}^y, \bar{\lambda}^z, \pi^+, \pi^-)$:

$$0 > -\tfrac{1}{2}\bar{\lambda}_0\epsilon + e^\top\pi^+ + e^\top\pi^-$$

$$0 = \bar{\lambda}_0\nabla\phi^\rho_\mu(\tilde{x}, \tilde{y}, \tilde{z}) + \pi^+ - \pi^- - (0, \bar{\lambda}^y, 0)^\top - (0, 0, \bar{\lambda}^z)^\top$$

$$\bar{\lambda}_0 \geq 0,\ \pi^+ \geq 0,\ \pi^- \geq 0,$$

$$\bar{\lambda}^y_{\bar{i}} \geq 0,$$

$$\bar{\lambda}_i{}^y \geq 0, \quad i \notin I^\epsilon(\tilde{y}).$$

By the first inequality, we know that $\bar{\lambda}_0 > 0$ and $e^\top\pi^+ + e^\top\pi^- < \frac{1}{2}\bar{\lambda}_0\epsilon$. Then dividing the whole system by $\bar{\lambda}_0$ and let

$$\tilde{r} = -\frac{\pi^+ - \pi^-}{\bar{\lambda}_0}, \quad \tilde{\lambda}^y = \frac{\bar{\lambda}^y}{\bar{\lambda}_0}, \quad \tilde{\lambda}^z = \frac{\bar{\lambda}^z}{\bar{\lambda}_0},$$

then we have

$$\tilde{r} = \nabla\phi^\rho_\mu(\tilde{x}, \tilde{y}, \tilde{z}) - (0, \tilde{\lambda}^y, 0)^\top - (0, 0, \tilde{\lambda}^z)^\top,$$

$$\tilde{\lambda}^y_{\bar{i}} \geq 0,$$

$$\tilde{\lambda}^y_i \geq 0,\ i \notin I^\epsilon(\tilde{y}).$$

On one hand,

$$\|\tilde{r} - \bar{r}\|_1 \leq \|\tilde{r}\| + \|\bar{r}\|_1 < \epsilon$$

since $\|\bar{r}\|_1 < \frac{1}{2}\epsilon$ and

$$\|\tilde{r}\|_1 = \|\frac{\pi^+ - \pi^-}{\bar{\lambda}_0}\|_1 \leq \frac{\|\pi^+\|_1 + \|\pi^-\|_1}{\bar{\lambda}_0} < \frac{1}{2}\epsilon.$$

On the other hand,

$$\|\tilde{r}-\bar{r}\|_1 = \|(0,\tilde{\lambda}^y-\lambda^y,0)^\top + (0,0,\tilde{\lambda}^z-\lambda^z)^\top\|_1 \geq |\tilde{\lambda}_{\bar{i}}^y - \lambda_{\bar{i}}^y| > \epsilon$$

since $\tilde{\lambda}_{\bar{i}}^y \geq 0$ and $\lambda_{\bar{i}}^y < -\epsilon$. There is a contradiction. Thus, system (4.25) has a solution.

Lemma 4.5 *Let d be a solution to the system (4.25). Then*

$$\min_{0\leq t\leq\epsilon} \nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot(td) + \frac{1}{2}(td)^\top D(td) \leq -min\{\frac{1}{4}\epsilon^2, \frac{1}{8\beta}\epsilon^2\}.$$

Proof of Lemma 4.5. By assumption (A1), we have

$$\nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot(td) + \frac{1}{2}(td)^\top D(td) \leq -\frac{1}{2}\epsilon t + \frac{1}{2}\beta t^2,$$

where $\beta = (n+2m)\beta_2$. Let $\bar{t} = \arg\min_{0<t\leq\epsilon} -\frac{1}{2}\epsilon t + \frac{1}{2}\beta t^2$. If $\bar{t} = -\frac{-\epsilon}{2\beta} < \epsilon$ when $\beta > \frac{1}{2}$, then we have

$$\min_{0<t\leq\epsilon} -\frac{1}{2}\epsilon t + \frac{1}{2}\beta t^2 \quad = -\tfrac{\epsilon^2}{8\beta}.$$

Otherwise, $\bar{t} = \epsilon \leq -\frac{-\epsilon}{2\beta}$ when $\beta \leq \frac{1}{2}$, then we have

$$\begin{aligned}\min_{0<t\leq\epsilon} -\frac{1}{2}\epsilon t + \frac{1}{2}\beta t^2 \quad &= -\tfrac{1}{2}\epsilon^2 + \tfrac{1}{2}\beta\epsilon^2 \\ &\leq -\tfrac{\epsilon^2}{4}.\end{aligned}$$

Therefore

$$\min_{0<t\leq\epsilon} \frac{1}{2}\epsilon t + \frac{1}{2}\beta t^2 \leq -min\{\frac{1}{4}\epsilon^2, \frac{1}{8\beta}\epsilon^2\},$$

■

Let $\tilde{d}$ be the solution to the quadratic problem $QMPEC_{(\tilde{A},\tilde{B})}(\tilde{x},\tilde{y},\tilde{z})$. Then since td, where $0 \le t \le \epsilon$ and d is a solution of system (4.25), is feasible to the problem $QMPEC_{(\tilde{A},\tilde{B})}(\tilde{x},\tilde{y},\tilde{z})$ at $(\tilde{x},\tilde{y},\tilde{z})$, we have by Lemma 4.5 that

$$\begin{aligned}\nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot\tilde{d}+\frac{1}{2}\tilde{d}^\top D\tilde{d} &\le \min_{0\le t\le\epsilon}\nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot(td)+\frac{1}{2}(td)^\top D(td)\\ &\le -min\{\frac{1}{4}\epsilon^2,\frac{1}{8\beta}\epsilon^2\}.\end{aligned}$$

Hence we have proved that

$$\nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot\tilde{d}+\frac{1}{2}\tilde{d}^\top D\tilde{d}\le -min\{\frac{1}{4}\epsilon^2,\frac{1}{8\beta}\epsilon^2\}.$$

For the case (4.24), we can also get a similar result.

Combining Cases 1 and 2, we have

$$\nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot\tilde{d}+\frac{1}{2}\tilde{d}^\top D\tilde{d}\le -min\{\frac{1}{4}\epsilon^2,\frac{1}{8\beta}\epsilon^2,\frac{1}{2M_1}\epsilon\}. \tag{4.26}$$

By assumptions $(A1)$ and $(A2)$, there is an $M_2 > 0$ such that

$$\nabla\phi_\mu^\rho(\tilde{x},\tilde{y},\tilde{z})\cdot\ddot{d}+\frac{1}{2}\ddot{d}^\top D\ddot{d}\ge -M_2$$

for any (x,y,z) with $\phi_\mu^\rho(x,y,z)\le\phi_\mu^\rho(x_1,y_1,z_1)$, and the steepest descent pair $(\ddot{A},\ddot{B})\in\mathcal{P}_{(x,y,z)}$. Now, let

$$c_1 = min\{\frac{\epsilon^2}{4M_2},\frac{\epsilon^2}{8\beta M_2},\frac{\epsilon^2}{2M_1M_2}\},$$

then $(\tilde{A}, \tilde{B})$ is an acceptable descent pair and the proof of theorem is now complete. ■

Theorem 4.6 *If assumptions* $(A1)$ *and* $(A2)$ *hold, then Algorithm 4.1 finitely stops at an* $\epsilon-$*approximate stationary point of* $FMPEC_{\mu}^{\rho}$.

Proof. Let $\{(\bar{x}_k, \bar{y}_k, \bar{z}_k)\}$ be an iterative sequence generated by Algorithm 4.1. Then by virtue of assumption $(A2)$, without loss of generality, we may assume that this sequence converges to a cluster point $(\bar{x}, \bar{y}, \bar{z})$. By Theorems 4.2 and 4.4, if Algorithm 4.1 does not finitely stop at an $\epsilon-$approximate S-stationary point of $FMPEC_{\mu}^{\rho}$, then each cluster point of the iterative sequence generated by Algorithm 4.1 must be an S-stationary point of $FMPEC_{\mu}^{\rho}$. Hence $(\bar{x}, \bar{y}, \bar{z})$ must be a S-stationary point of $FMPEC_{\mu}^{\rho}$ if $\{(\bar{x}_k, \bar{y}_k, \bar{z}_k)\}$ is an infinite sequence.

Suppose that $(\bar{x}_k, \bar{y}_k, \bar{z}_k)$ is not an ϵ-approximate S-stationary of $FMPEC_{\mu}^{\rho}$ for all k. Then similarly as in the proof of Theorem 4.4 (see inequality (4.26)), in step 4 of Algorithm 4.1 we can find an acceptable descent pair $(A_k, B_k) \in \mathcal{P}_{(\bar{x}_k, \bar{y}_k, \bar{z}_k)}$ such that the solution $\bar{d}_k$ of $QMPEC_{(A_k,B_k)}(\bar{x}_k, \bar{y}_k, \bar{z}_k)$ satisfies

$$\nabla \phi_{\mu}^{\rho}(\bar{x}_k, \bar{y}_k, \bar{z}_k) \cdot \bar{d}_k + \frac{1}{2}\bar{d}_k^{\top} D \bar{d}_k \leq -\epsilon \tag{4.27}$$

for all sufficient large k. Since there is only a finite number of the index sets, without lose of generality, we may assume that the index set $\{I(\bar{y}_k)\}$ and the sequence $\{\bar{z}_k\}$ stay constant when k is large enough. Therefore, $\mathcal{P}_{(\bar{x}_k,\bar{y}_k,\bar{z}_k)}$ will remain the same as well when k is large enough. Since $\mathcal{P}_{(\bar{x},\bar{y},\bar{z})}$ is finite, we can find a pair $(\tilde{A},\tilde{B}) \in \mathcal{P}_{(\bar{x}_k,\bar{y}_k,\bar{z}_k)}$ such that the solution $\bar{d}_k$ of $QMPEC_{(\tilde{A},\tilde{B})}(\bar{x}_k,\bar{y}_k,\bar{z}_k)$ satisfies (4.27) for large enough k. By $(A2)$, without loss of generality we may assume that $\bar{d} := \lim_{k\to\infty} \bar{d}_k$. Taking limit as k goes to infinity in (4.27), we have

$$\nabla\phi_\mu^\rho(\bar{x},\bar{y},\bar{z}) \cdot \bar{d} + \frac{1}{2}\bar{d}^\top D\bar{d} \leq -\epsilon$$

and $\bar{d}$ is the solution of $QMPEC_{(\tilde{A},\tilde{B})}(\bar{x},\bar{y},\bar{z})$, by the standard result of sensitive analysis and the fact that the LICQ is satisfied in the feasible region of the problem $QMPEC_{(\tilde{A},\tilde{B})}(\bar{x},\bar{y},\bar{z})$. Consequently $\bar{d}$ is a descent direction of ϕ_μ^ρ at $(\bar{x},\bar{y},\bar{z})$ in the piece $\mathcal{WF}_{(\tilde{A},\tilde{B})}$. This contradicts the fact that $(\bar{x},\bar{y},\bar{z})$ is an S-stationary point of the problem $FMPEC_\mu^\rho$. Hence there is a k such that $(\bar{x}_k,\bar{y}_k,\bar{z}_k)$ is an ϵ-approximate S-stationary of $FMPEC_\mu^\rho$ and the proof of the theorem is completed. ∎

Chapter 5

Numerical Results

In order to measure the efficiency of the algorithm proposed in this paper, we have implemented the algorithms in MATLAB. The codes can be found in Appendix. We report the implementation details and some preliminary results for the problems solved by our algorithm.

5.1 Implementation Details

We first describe implementation details of Algorithm 3.1. We set $\rho = 10^{-5}$ and $\mu = 10$ initially. ρ is decreased by a factor of 2 at every iteration and μ is adjusted dynamically. In step 2, MPEC is reformulated to $FMPEC_{\mu}^{\rho}$ and then Algorithm 4.1 is called to solve $FMPEC_{\mu}^{\rho}$. Now, we give a termination rule to Algorithm 3.1. According to the convergent analysis of this algorithm, it should go to next iteration if

$$sum\_\lambda := \sum_{i=1}^{m} |\lambda_i{}^{G^k}| + \sum_{i=1}^{m} |\lambda_i{}^{H^k}| + \sum_{i=1}^{q} |\lambda_i{}^{h_k}| + \sum_{i=1}^{p} |\lambda_i{}^{g_k}| \geq (1-\alpha_1)\mu_k.$$

We simply choose $distance := \|(x_{k+1}, y_{k+1}, z_{k+1}) - (x_k, y_k, z_k)\| \leq tol$ to account for the roundoff errors, where the tolerance tol is set to 10^{-10}. Therefore, we stop when $sum\_\lambda \leq (1-\alpha_1)\mu_k$ and $distance \leq tol$.

Next we describe the implementation details for algorithm 4.1 which is used for solving the subproblems. Set $\sigma = p = c_2 = 0.5$ $\alpha_1 = 0.5 \times 10^{-5}$ and $\alpha_2 = 10^{-5}$. We also set $\delta_0 = max(y_0, z_0) + 1$ in order to make it large enough. In step 4, the solution d and multipliers λ_y, λ_z of the quadratic program $QMPEC_{A,B}(x, y, z)$ are found by calling the quadratic program solver QUADPROG from the MATLAB Optimization Toolbox. After an acceptable descent direction d is found, we implement Armijo inexact line search using the objective function $\phi_\mu^\rho(x, y, z)$ to determinate step size. [Notice that in Liu and Ye [9], the L_1 penalty function is used to perform Armijo inexact line search. As for the problem $FMPEC_\mu^\rho$, it only has simple linear complementarity constraints. Thus, the objective function $\phi_\mu^\rho(x, y, z)$ is exactly the same as its L_1 penalty function.] In step 7, we return to algorithm 3.1 when $\|(x_{k+1}, y_{k+1}, z_{k+1}) - (x_k, y_k, z_k)\| \leq tol$, otherwise we adjust the positive matrix D_k by BFGS rule as follows and go to the next iteration:

$$D_{k+1} = D_k + \frac{pp^\top}{p^\top q}(1 + \frac{q^\top D_k q}{p^\top q}) - \frac{D_k q p^\top + p q^\top D_k}{p^\top q},$$

where

$$p := (x_{k+1}, y_{k+1}, z_{k+1}) - (x_k, y_k, z_k)$$

and

$$q := \nabla\phi_\mu^\rho(x_{k+1}, y_{k+1}, z_{k+1}) - \nabla\phi_\mu^\rho(x_k, y_k, z_k).$$

5.2 Test Problems

We have performed some experiments on two sets of problems: small size problems taken from the literature and randomly generated problems. The first two problems have been solved by other algorithms. We refer the readers to the reference [13] for the first one and [4] for the second one.

Problem 1. This is an MPEC with quadratic objective function and linear complementarity constraints:

$$\begin{aligned} \min \quad & \frac{1}{2}(x^2 + y^2) \\ \text{s.t.} \quad & x \geq 1, \\ & F(x,y) := -x + y + 2 \geq 0, \quad y \geq 0 \\ & y^\top F(x,y) = 0. \end{aligned}$$

The feasible region of problem 1 is

$$\mathcal{F} = \{(x, 0) : 1 \leq x \leq 2\} \bigcup \{(x, x-2) : x \geq 2\}.$$

With initial points $(x_0, y_0) = (4, 2)$ and $(x_0, y_0) = (2, 0)$ respectively, the algorithm stopped at the unique optimal solution $(1, 0)$.

Problem 2. This problem is previously presented in [4] and reformed from the generalized Nash equilibria [3].

$$\begin{aligned} \min \quad & \frac{1}{2}[(x_1 + x_2 + y_1 - 15)^2 + (x_1 + x_2 + y_2 - 15)^2] \\ \text{s.t.} \quad & 0 \leq x_1 \leq 10, \quad 0 \leq x_2 \leq 10 \\ & z = Nx + My + q \geq 0, \quad y \geq 0, \\ & y^{\mathsf{T}} z = 0, \end{aligned}$$

where

$$N = \begin{bmatrix} \frac{8}{3} & 2 \\ 2 & \frac{5}{4} \end{bmatrix}, \quad M = \begin{bmatrix} 2 & \frac{8}{3} \\ \frac{5}{4} & 2 \end{bmatrix}, \quad q = \begin{bmatrix} -36 \\ -25 \end{bmatrix}.$$

With $(15/4, 4, 4, 15/4, 0, 0)$ as a initial point, the output $(7, 7.5, 0.5, 0.5, 0, 0)$ is one of the global optimal solution while with $(0, 0, 0, 0, -36, -25)$ as another initial point, it comes out with the other global optimal solution $(9, 6, 0, 0, 0, 0.5)$.

Problem 3. This is the problem we have presented in Chapter 4. With the initial points (1,0) in the face (4.3) and (0,2) in the face (4.4), we find the optimal solution (0,1).

Problem 4. This is a set of problems with following form:

$$\min_{x,y} \quad f(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix}^T (C + \mu_1 I)\begin{bmatrix} x \\ y \end{bmatrix} + c^T\begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{aligned} s.t. \quad & A_1x + B_1y = b_1, \\ & b_2 - (A_2x + B_2y) \geq 0, \\ & y^T[b_2 - (A_2x + B_2y)] = 0, \\ & y \geq 0, \end{aligned}$$

where $x \in R^{(n-m)}$, $n \geq m$, $y \in R^m$, $c \in R^n$, $C \in R^{n\times n}$, $b_1 \in R^q$, $b_2 \in R^m$, $A_1 \in R^{q\times(n-m)}$, $A_2 \in R^{m\times(n-m)}$, $B_1 \in R^{q\times m}$, $B_2 \in R^{m\times m}$, $\mu_1 > 0$ and I is the $n \times n$ identity matrix. We remark that the MPEC LICQ will not hold at any feasible point if $q > n$.

Now we describe the details in generating the test problems.

To generate a problem with a feasible point (x_0, y_0, z_0), we consider the reformulation of the problem as below:

$$\min_{x,y} \quad f(x,y) = \tfrac{1}{2} \begin{bmatrix} x \\ y \end{bmatrix}^T (C+\mu_1 I) \begin{bmatrix} x \\ y \end{bmatrix} + c^T \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{aligned} s.t. \quad & A_1 x + B_1 y = b_1, \\ & A_2 x + B_2 y + z = b_2, \\ & y^T z = 0, \\ & y \geq 0, \;\; z \geq 0, \end{aligned}$$

1) Generate Initial point

Set $0 < \sigma < 1$ to control the scale of the randomly generated variables. Let $x_0 \in R^{(n-m)}$ be a vector with random entries chosen from a uniform distribution on the interval (0,1). Then we reset $x_0 := tan(\sigma \pi x_0/2)$. So we have $x_0 \geq 0$. Similarly, we can generate $y_0 \geq 0$ and $z_0 \geq 0$.

For $i = 1, \cdots, (n-m)$, if $(x_0)_i < tol$, then reset $(x_0)_i = 0$.

For $i = 1, \cdots, m$,

if $|(y_0)_i - (z_0)_i| < tol$, then reset $(y_0)_i = (z_0)_i = 0$;

if $(y_0)_i < (z_0)_i$, then reset $(y_0)_i = 0$;

if $(y_0)_i > (z_0)_i$, then reset $(z_0)_i = 0$.

2) Generate Constraints

Set $0 < \sigma_{A1}, \sigma_{A2}, \sigma_{B1}, \sigma_{B2} < 1$. Generate a matrix $A_1 \in [-\frac{1}{2}, \frac{1}{2}]^{q \times (n-m)}$

with uniformly distributed random entries, then reset $A_1 := tan(\sigma_{A1}\pi A_1)$. Similarly we generate A_2, B_1 and B_2. Let

$$b_1 = \quad A_1x_0 + B_1y_0,$$

$$b_2 = \quad A_2x_0 + B_2y_0 + z_0.$$

3) Generate the Objective Function

Set $1 > \sigma_c > 0$, $1 > \sigma_C > 0$. Similar to the generation of x_0 and A_1, we can generate $\lambda > 0$, c and C.

Problem 5. This is a set of problems with quadratic objective function and quadratic complementarity constraints.

$$\begin{aligned} \min_x \quad & f(x) = c^\top x + \frac{1}{2}x^T(C + \mu_1 I)x \\ s.t. \quad & G(x) \geq 0,\ H(x) \geq 0,\ G(x)^\top H(x) = 0 \\ & x \geq 0, \end{aligned}$$

where

$$G_i(x) = A_{1i}x + \frac{1}{2}x^\top B_{1i}x - b_{1i}, \quad i = 1, \cdots, m,$$
$$H_i(x) = A_{2i}x + \frac{1}{2}x^\top B_{2i}x - b_{2i}, \quad i = 1, \cdots, m,$$

$x \in R^n$, $B_{1i}, B_{2i}(i = 1, \cdots, m)$ and $C \in R^{n\times n}$, A_{1i}, A_{2i} are the ith row of the matrices $A_1, A_2 \in R^{m\times n}$ respectively, b_{1i}, b_{2i} are the ith row of the matrices $b_1, b_2 \in R^m$ respectively.

We set the same value to parameters with that in Problem 4. First, we test a small size problem with

$$c = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & 3 & 5 \\ 3 & 8 & 5 \\ 5 & 5 & 4; \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 7 & 6 \\ 3 & 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 6 & 2 \end{bmatrix},$$

$$b_1 = b_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B_{11} = \begin{bmatrix} 6 & 7 & 8 \\ 1 & 7 & 9 \\ 4 & 5 & 3 \end{bmatrix},$$

$$B_{12} = \begin{bmatrix} 6 & 5 & 3 \\ 7 & 2 & 1 \\ 7 & 5 & 1 \end{bmatrix}, \quad B_{21} = \begin{bmatrix} 1 & 7 & 6 \\ 5 & 9 & 2 \\ 4 & 3 & 8 \end{bmatrix}, \quad B_{22} = \begin{bmatrix} 4 & 3 & 2 \\ 2 & 8 & 2 \\ 7 & 3 & 9 \end{bmatrix}.$$

The optimal solution of this problem is $(0, 0, 0)$. Note that MPEC LICQ is not satisfied at this point. The experiment produces the result with the magnitude 10e-08.

We randomly generate the problems and a feasible point of them in the same way with Problem 4.

5.3 Computational Results

We list the computational results in the following table.

No.	(n,m,p,q)	deg^1	$iter^2/qp^3$	μ_k	ρ_k	$f(x_0)$	$f(x_k)$
1	(2,1,1,0)	0	39/1400	10	1.819e-17	10	0.5
2	(4,2,4,0)	0	1/61	10	5e-06	11.406	8.8927e-12
3	(2,1,0,0)	0	1/67	10	5e-06	4	0
4	(10,5,0,8)	1	44/1236	14.233	5.6843e-19	160.52	30.332
4	(15,10,0,20)	0	45/4592	28.697	2.8422e-19	257.59	174.2
4	(40,20,0,20)	5	45/10184	44.107	2.8422e-19	572.6	440.01
4	(100,50,0,60)	8	20/20276	96.033	9.5367e-12	1417.3	1173.9
5	(3,2,3,0)	0	31/1233	10	4.6566e-15	27	-4.9235e-9
5	(10,6,10,0)	5	6/111	13.38	1.5625e-07	67.253	21.927
5	(20,10,20,0)	3	55/244	10	2.7756e-22	228.06	225.79
5	(40,20,40,0)	12	12/109	40.541	2.4414e-09	415.11	414.83

[1] The number of the degenerations at termination point.

[2] The number of the outer iterations.

[3] The number of the quadratic problem solved.

For all the tested problems, the algorithm terminated successfully. Both

Problems 4 and 5 do not satisfy the MPEC LICQ or the nondegereracy condition. Notice that the number of outer iterations does not increase along with the problem size. The size of the problem only affects the number of quadratic problems solved. While the value of penalty parameter roughly increases with the problem size.

We also notice that Problem 1 was solved with a large number of iterations. This problem has been previously solved in three iterations by the locally convergent PSQP algorithm under the MPEC LICQ in [13]. However, for this simple problem, one may not notice the time difference since it was solved in seconds. For a large size MPEC with simple linear complementarity constraints, the other methods, such as [4] and [13], may be good alternatives in terms of computing time. While the method proposed here has its advantages in solving MPECs with nonlinear complementarity constraints if the MPEC LICQ or the nondegereracy condition is not satisfied.

The algorithm is sensitive to α_1 and α_2. For example, changing α_2 from 10^{-5} to 10^{-10} for Problems 1 to 3 makes the output switched from the optimal solution to an M-stationary point. However, we only want to show the efficiency of the proposed algorithms. The optimal choices of parameters will not be investigated in the thesis.

Chapter 6

Conclusions

We have proposed a new smoothing penalty algorithm for solving mathematical programs with equilibrium constraints. In the new algorithm, the MPEC is reformulated to a series of approximate subproblems by smoothing the partial exact penalty function. The feasibility of an accumulated point is guaranteed under the extended MPEC GMFCQ which is a milder constraint qualification compared with strict complementarity condition or the MPEC LICQ. Moreover, the global convergence to an M-stationary point is established under the same constraint qualification.

An implementable hybrid algorithm is introduced to solve the subproblems. The subproblems belong to a class of mathematical programs with simple linear complementarity constraints and satisfy the MPEC LICQ naturally at any feasible point. By combining the active set technique, the $\delta-$active search technique and the PSQP algorithm, the hybrid algorithm globally converges to an S-stationary point without the strict complementarity condition.

The theoretical promises of the algorithms are supported by the reported numerical results. However, we realized that the effect of parameter ϵ has not been fully understood from the experiments we have done. In addition, the influence by other parameters is still under investigation. The other future work is to apply the algorithms to large-size realistic MPECs.

APPENDIX

```
%*********************************************************************
%This algorithm is used for finding a M-stationary point for MPEC.
%The outer loop is designed for updating parameters, mu and rho, of
%approximation partial exact penalty function. The inner loop can be any
%known algorithm to obtain a stationary point of MPLEC. Herein, we borrow the
%scheme of Algorithm 2.1 designed by G.S.Liu and J. J. Ye .
%*********************************************************************
clear all;
%initialization
rho=1e-05; %smoothing parameter
epsilon2=1e-01; %constant for updating mu

tol=1e-10;%termination tolerance
n=2;%n=input('dimention of x n='); %input dimention of x
m=1;%m=input('dimention of G and H m=') %input dimention of  G and H
x0=[1,0];%x0=input('initial feasible point x=')
y0=1;%y0=input('initial feasible point y=')
z0=0;%z0=input('initial feasible point z=')
mu=10; %penalty parameter
%end of initialization
iter=0; iter2=0; f0=f(x0);
% computing S-stationary point of FMPEC_\rho^\mu
[x,y,z,lambda_y,lambda_z,iter2]=mplcq(x0,y0,z0,rho,mu,n,m,iter2);
while norm([x,y,z]-[x0,y0,z0])>tol
    iter=iter+1
    %start to update mu
    blambda_h=lambda_h(x,rho);%computing vector lambda^h
```

```
    blambda_g=lambda_g(x,rho);%computing vector lambda^g
    aa=sum(abs(lambdaG(x,y,rho)));%sum(abs(lambda_y));
    bb=sum(abs(lambdaH(x,z,rho)));%sum(abs(lambda_z));
    cc=sum(abs(blambda_h));
    dd=sum(abs(blambda_g));
    sum_lambda=aa+bb+cc+dd;
    if sum_lambda>(1-epsilon2)*mu
        mu=sum_lambda+epsilon2*mu;
    end
    % end of updating mu
    rho=0.5*rho;    %decrease rho;
    x0=x;y0=y;z0=z;
    [x,y,z,lambda_y,lambda_z,iter2]=mplcq(x0,y0,z0,rho,mu,n,m,iter2);
    iter2
end


 function [x0,y0,z0,lambda_y,lambda_z,iter2]=mplcq(x0,y0,z0, rho, mu,n,m,iter2)
%************************************************************
% This algorithm is designed for finding a stationary point
% for MPLCQ
%*************************************************************
%initialization for fmpec_\rho^\mu
tol=1e-10%tolerant
sigma=0.5; %constant for armijo linear search
delta0=max(max(y0),max(z0))+1; %active set radius
```

```
delta=0;
p=0.5; %step size constant for armijo linear search
c1=0.5;%coeffection for reducing the active set radius
D=eye(n+2*m); %symetric positive matrix
epsilon=1e-5;%epsilon approximation
%end of initialization for fmpec_\rho^\mu
%tx --tilda x
%x0 -- x_k.
while delta==0

    delta=delta0;
    x_temp=x0; y_temp=y0; z_temp=z0;
    delta_temp=delta;
    while delta_temp>0
        tx=x0;
        for i=1:m
            if y0(i)<=delta
                ty(i)=0;
            else
                ty(i)=y0(i);
            end
        end
        for i=1:m
            if z0(i)<=delta
                tz(i)=0;
            else
                tz(i)=z0(i);
            end
        end
```

```
%start step 4: find a acceptable descent pair or conclude (tx,ty,tz) ia an
%approximate S-stationary point of FMPEC_\mu^\rho

%randomly give an adjacent pair (A,B)
        A=zeros(1,m);
        B=zeros(1,m);
        for i=1:m
            if ty(i)==0 & tz(i)>0
                A(i)=i;
            else if ty(i)>0 & tz(i)==0
                    B(i)=i;
                else if fix(i/2)*2==i
                          B(i)=i;
                    else
                     A(i)=i;
                    end
                end
            end
        end   % end of index set assignment to (A,B)


% find the optimal solution for the quadratic programming QMPEC.D is the
% positive definite matrix. A is the coefficient matrix with k equations
% first. b is the constant at right hand side.
% min 0.5d'Dd+numbdaphi(tx,ty,tz)'d
% s.t. C1d<=b1, C2d=b2

% give k
```

```
% k=n+m; % d_x=0
 C1=zeros(m,n+2*m); %Construct C1,C2,b1 and b2
 C2=zeros(m,n+2*m);
 for i=1:m
     if A(i)==i % i is in A,
         C1(i,n+m+i)=-1;% -d_z^i<=tz(i)
         b1(i)=tz(i);
         C2(i,n+i)=-1;% -d_y^i=ty(i)
         b2(i)=ty(i);
     else
         C1(i,n+i)=-1;%-d_y^i<=ty(i)
         b1(i)=ty(i);
         C2(i,n+m+i)=-1; % -d_z^i=tz(i)
         b2(i)=tz(i);
     end
 end % end of construction
 [tx,ty,tz];
 % find the solution for QMPEC
 [d,fval,exitflag,output,lambda]=quadprog(D,numbdaphi(tx,ty,tz,rho,mu),C1,b1,C2,b2);
 iter2=iter2+1; % compute objective function value
 %construct lambda_y, lambda_z
 for i=1:m
     if i==A(i)
         lambda_y(i)=lambda.eqlin(i);
         lambda_z(i)=lambda.ineqlin(i);
     else
         lambda_z(i)=lambda.eqlin(i);
         lambda_y(i)=lambda.ineqlin(i);
     end
```

```
end %end of construction
barr=D*d;
if sum(abs(barr))<0.5*epsilon & max(abs(d))<0.5*epsilon
    bari=[0,0];
    for i=1:m  % find smallest negative multiplier
        if ty(i)==0&tz(i)==0 % i in beta(ty,tz)
            if lambda.ineqlin(i)<-epsilon & bari(2)>lambda.ineqlin(i)
                bari(1)=i;
                bari(2)=lambda.ineqlin(i);
             end
             if lambda.eqlin(i)<-epsilon & bari(2)>lambda.eqlin(i)
                    bari(1)=i;
                    bari(2)=lambda.eqlin(i);
             end
        end
    end  % end of find smallest negative multiplier
    if bari(1)==0;
    %conclude (tx,ty,tz) is an apprx. S-stationry point and return (tx, ty,
       %tz) and lambda
           return;
    end
 % reconstruct C and b
     if bari(1)==A(bari(1))  % bari switch to B from A,
          C1(bari(1),bari(1)+n+m)=0; % minus bar i from A
          C2(bari(1),bari(1)+n)=0;
          C1(bari(1),bari(1)+n)=-1; % put bar i in B
          b1(bari(1))=ty(bari(1));
          C2(bari(1),bari(1)+m+n)=-1;% -d_z^i=tz(i)
          b2(bari(1))=tz(bari(1));
```

```
        else    %bar i is switched to A from B
              C1(bari(1), bari(1)+n)=0; % minus bar i from B
              C2(bari(1), bari(1)+m+n)=0;
              C1(bari(1),bari(1)+m+n)=-1;% put bar i in A
              b1(bari(1)+m)=tz(bari(1));
              C2(bari(1),bari(1)+n)=-1; %-d_y^i=ty(i)
              b2(bari(1))=ty(bari(1));
        end

 %find d again according to new pair (A,B)
       [d,fval,exitflag,output,lambda]=quadprog(D,numbdaphi(tx,ty,tz,rho,mu),C1,b1,C2,b2);
       iter2=iter2+1;
       %construct lambda_y, lambda_z
       %ff=fval
       for i=1:m
          if i==A(i)
              lambda_y(i)=lambda.eqlin(i);
              lambda_z(i)=lambda.ineqlin(i);
          else
              lambda_z(i)=lambda.eqlin(i);
              lambda_y(i)=lambda.ineqlin(i);
          end
       end %end of construction
end  %end of step 4

% Armijo's rule
step = 1;
temp=phi(tx+step*d(1:n)',ty+step*d(n+1:n+m)',tz+step*d(n+m+1:n+2*m)',rho,mu)-
phi(tx,ty,tz,rho,mu);
```

```
    while temp>-0.5*sigma*step*d'*D*d
         step = step*p;
         temp=phi(tx+step*d(1:n)',ty+step*d(n+1:n+m)',tz+step*d(n+m+1:n+2*m)',rho,mu)-
         phi(tx,ty,tz,rho,mu);
    end  % end of search
    if phi(x_temp,y_temp,z_temp,rho,mu)>
    phi(tx+step*d(1:n)',ty+step*d(n+1:n+m)',tz+step*d(n+m+1:n+2*m)',rho,mu)
         x_temp=tx+step*d(1:n)'
         y_temp=ty+step*d(n+1:n+m)'
         z_temp=tz+step*d(n+m+1:n+2*m)'
    end

    delta_temp=delta; %start to adjust delta
    ymax=0;
    zmax=0;
    for i=1:m
        if y0(i)<=delta
             ymax=max(ymax, y0(i));
        end
    end
    for i=1:m
        if z0(i)<=delta
             zmax=max(ymax,z0(i));
        end
    end
    delta=c1*max(ymax,zmax) %end of adjusting delta
end  % go to new delta active search
%if delta=0
p0 = [x_temp-x0,y_temp-y0,z_temp-z0];
```

```
    if norm(p0)<tol
        return
    end

    % Update D by BFGS update
    q = numbdaphi(x_temp, y_temp, z_temp,rho,mu)-numbdaphi(x0,y0,z0,rho,mu);
    D = D + q'*q/(p0*q') - D*p0'*(D*p0')'/(p0*D*p0');
    % Reset D to the identity if it fails to be positive definite
    if min(eig(D))<1e-5
        D=eye(n+2*m);
    end
    % the very end of BFGS algorithm

    x0=x_temp;
    y0=y_temp;
    z0=z_temp;
end  %go to next iteration
```

Bibliography

[1] X.J. CHEN AND F. FUKUSHIMA, *A smoothing method for a mathematical program with P-matrix linear complementarity constraints*, Computational Optimization and Applications, **27**(2004), pp. 223-246.

[2] S. DEMPE, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, 2002.

[3] F. FACCHINEI, H.Y. JIANG AND L. QI, *A smoothing method for mathematical programs with equilibrium constraints*, Mathematical Programming, **85**(1999), pp. 107-134.

[4] M. FUKUSHIMA, Z.Q. LUO and J.S. PANG, *A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints*, Computational Optimization and Application, **10**(1998), pp. 5-34.

[5] M. FUKUSHIMA, and J.S. PANG, *Convergence of a smoothing continuation method for mathematical programs with complementarity constraints*, Ill-Posed Variational Problems and Regularization Techniques, Edited by *M.Théra* and *R.Tichatschke*, springer, New York, NY, pp. 99-110, 2000.

[6] M. FUKUSHIMA and P. ZENG, *An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints*, SIAM Journal on Optimization, **12**(2002), pp. 724-739.

[7] H.Y. JIANG AND D. RALPH, *Extension of Quasi-Newton methods to mathematical programs with complementarity constraints*, Computational Optimization and Applications, **25**(2003), pp. 123-150.

[8] G. H. LIN and M. FUKUSHIMA, *New Relaxation Method for Mathematical Programs with Complementarity Constraints*, Journal of Optimization Theory and Applications, **118**(2003), pp. 81-116.

[9] G.S. LIU and J.J. YE , *A merit function piecewise SQP algorithm for solving mathematical programs with equilibrium constraints*, preprint.

[10] G.S. LIU, J.J. YE and J.P. ZHU , *A Smoothing Penalty Method for Solving Mathematical Programs with Equilibrium Constraints*, preprint.

[11] Y. LUCET and J. YE, *Sensitivity Analysis of the value function for optimization problems with variational inequality constraints*, SIAM Journal on Control and Optimization, **40**(2001), pp. 699-723.

[12] Z.Q. LUO, J.S. PANG and D. RALPH, *Mathematical Programs with Equilibrium Constraints,* Cambridge University Press, 1996.

[13] Z.Q. LUO, J.S. PANG and D. RALPH, *Piece-wise sequential quadratic programming for mathematical programs with nonlinear complementarity constraints*, in: M.C. Ferris and J.S. Pang, eds., Complementarity and Variational Problems: State of the Art, SIAM Publications, 1997.

[14] Z.Q. LUO, J.S. PANG, D. RALPH and S.Q. WU, *Exact penalization and stationary conditions of mathematical programs with equilibrium constraints*, Mathematical Programming **76**(1996), pp. 19-76.

[15] S. SCHOLTES, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal on Optimization, **11**(2001), pp. 918-936.

[16] J.V. OUTRATA, *Optimality conditions for a class of mathematical programs with equilibrium constraints,* Mathematics of Operations Research, **24**(1999), pp. 627-644.

[17] J.V. OUTRATA, M. KOČVARA AND J. ZOWE, *Nonsmooth Approach to Optimization Problem with Equilibrium Constraints: Theory, Application and Numerical Results,* Kluwer, Dordrecht, The Netherlands, 1998.

[18] D. RALPH, *Sequential quadratic programming for mathematical programs with linear complementarity constraints*, in: R.L. May and A.K. Easton eds., CTAC95 Computational Techniques and Applications, World Scientific, 1996.

[19] A.V. FIACCO, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.

[20] H. SCHEEL AND S. SCHOLTES, *Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity*, Mathematics of Operations Research, **25**(2000), pp. 1-22.

[21] J.J. YE *Optimality conditions for optimization problems with complementarity constraints* , SIAM Journal on Optimization, **2**(1999), pp. 374-387.

[22] J.J. YE, *Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints*, Journal of Mathematical Analysis and Applications, **307**(2005), pp. 350-369.

[23] J.J. YE, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM Journal on Optimizatio, **10**(2000), pp. 943-962.

[24] J.J. Ye and X.Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Mathematics of Operations Research, **22**(1997), pp. 977-977.

[25] J.Z. ZHANG AND G.S. LIU, *A new extreme point algorithm and its application in PSQP algorithms for solving mathematical programs with linear Complementarity constraints*, Journal of Global Optimization, **19**(2001), pp. 345-361.

[26] J.Z. ZHANG, S.Y. WANG AND G.S. LIU, *A globally convergent approximately active search algorithm for solving mathematical programs with linear complementarity constraints*, Numerische Mathematik, **98**(2004), pp. 539-558.