A Smoothing Newton Algorithm for Mathematical Programs with Complementarity Constraints*

Zheng-Hai Huang† and Jie Sun‡

Dedicated to Professor Jiye Han on the occasion of his 70 birthday

Abstract

We propose a smoothing Newton algorithm for solving mathematical programs with complementarity constraints (MPCCs). Under some reasonable conditions, the proposed algorithm is shown to be globally convergent and to generate a B-stationary point of the MPCC. Notable is that our results do not rely on the strict complementarity condition. Preliminary numerical results on some MacMPEC problems are reported.

Key words: Mathematical program with complementarity constraints, B-stationary point, smoothing algorithm, global convergence.

AMS subject classifications: 90C30, 90C33.

1 Introduction

The mathematical program with complementarity constraints (MPCC) has attracted much attention in recent research on optimization. Due to the complementarity constraint, the constraint system of MPCC fails to satisfy the standard constraint qualification condition. As such, its solution may become very difficult.

The MPCC considered in this paper takes the following form:

$$\begin{array}{rl} \text{Minimize} & f(x) \\ \text{Subject to} & g(x) \leq 0,\ h(x) = 0,\ \min\{G(x), H(x)\} = 0, \end{array} \tag{1.1}$$

where $f : \Re^n \to \Re$, $g : \Re^n \to \Re^p$, $h : \Re^n \to \Re^q$, $G : \Re^n \to \Re^m$, $H : \Re^n \to \Re^m$, and for any vectors $u, v \in \Re^m$, $\min\{u, v\}$ denotes an m-vector whose i-th component is $\min\{u_i, v_i\}$. Throughout this

*The research is partially supported by Grants RP312000-042/057-112 of National University of Singapore.

†Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, P.R. China. Email: zhenghaihuang@yahoo.com.cn. This author's work was also partially supported by the Scientific Research Foundation of Tianjin University for the Returned Overseas Chinese Scholars and the Scientific Research Foundation of Liu Hui Center for Applied Mathematics, Nankai University-Tianjin University.

‡Department of Decision Sciences, National University of Singapore, Singapore 119260, Republic of Singapore. Email: jsun@nus.edu.sg. Fax: (65)67792621.

paper, we assume that all functions involved in (1.1) are sufficiently smooth and the feasible set of (1.1), denoted by $\mathcal{F}$, is nonempty.

Various numerical methods for solving MPCCs have been proposed in the literature, such as, the sequential quadratic programming method [4, 5, 14], the regularization method [17, 23], the interior point method [1, 19], the penalty method [9, 10, 16, 24], the active-set method [7], and the smoothing method [2, 4, 6, 14]. The smoothing method can be roughly described as follows: by using some smoothing functions, the complementarity constraints are reformulated as a system of smoothing equations, hence the MPCC cab be solved by solving a sequence of nonlinear programs (NLPs).

The idea used in this paper is somewhat different from the existing smoothing methods. Instead of NLPs, we reformulate the MPCC as a system of smoothing equations and use a Newton-type method to solve the resulting system. Under some conditions, we show that the algorithm is globally convergent and generates a B-stationary point of (1.1). The analysis on global convergence does not rely on the strict complementarity condition.

The paper is organized as follows. In the next section, we recall some basic concepts and properties of the MPCC. In Section 3, the MPCC is reformulated as a system of smoothing equations. In Section 4, we propose a Newton-type algorithm to solve the smoothing equations and discuss convergence properties of the algorithm. We give some numerical results in the last section.

The following notations are used throughout the paper. All vectors are column vectors, the supscript T denotes the transpose. Suppose that l is an arbitrary integer, $\Re^l$ denotes the space of l-dimensional real column vectors, and $\Re^l_+$ (respectively, $\Re^l_{++}$) denotes the nonnegative (respectively, positive) orthant in $\Re^l$. For any vector $u \in \Re^l$, we denote by u_i the i-th component of u and, for any $\mathcal{K} \subset \mathcal{L} := \{1, 2, \dots, l\}$, by $u_{\mathcal{K}}$ the vector obtained after removing from u those u_i with $i \notin \mathcal{K}$. We sometime write u as $\mathrm{vec}\{u_i : i \in \mathcal{L}\}$. We denote by $\mathrm{diag}\{u_i : i \in \mathcal{L}\}$ the diagonal matrix whose i-th diagonal element is u_i. We denote by $\|u\|$ the 2-norm of u. For any vectors $u, v \in \Re^l$, we write $(u^T, v^T)^T$ as (u, v) for simplicity. We denote by $\Re^{l\times l}$ the space of $l \times l$ real matrices. For any continuously differentiable function $F = (F_1, F_2, \dots, F_l)^T : \Re^l \to \Re^l$ where l is a positive integer, we denote its Jacobian by $F' = \bigtriangledown F^T$ with $\bigtriangledown F = (\bigtriangledown F_1, \bigtriangledown F_2, \dots, \bigtriangledown F_l)$, where $\bigtriangledown F_i$ denotes the gradient of F_i for $i \in \mathcal{L}$. For any $\alpha, \beta \in \Re_{++}$, we write $\alpha = O(\beta)$ (respectively, $\alpha = o(\beta)$) to mean α/β is uniformly bounded (respectively, tends to zero) as $\beta \to 0$. Let $k \geq 0$ denote the iteration index. For any $\left(\mu, x, \lambda^\phi, \lambda^g, \lambda^h\right), \left(\mu_k, x^k, (\lambda^\phi)^k, (\lambda^g)^k, (\lambda^h)^k\right) \in \Re^{1+n+m+p+q}$, we always use the following notation throughout this paper:

$$z := \left(\mu, x, \lambda^\phi, \lambda^g, \lambda^h\right),\ y := \left(x, \lambda^\phi, \lambda^g, \lambda^h\right),\ u := (\mu, x),\ v := (\mu, \lambda^g),\ w := (\mu, \lambda^g, g(x)),$$
$$z^k := \left(\mu_k, x^k, (\lambda^\phi)^k, (\lambda^g)^k, (\lambda^h)^k\right),\ y^k := \left(x^k, (\lambda^\phi)^k, (\lambda^g)^k, (\lambda^h)^k\right),$$
$$u^k := (\mu_k, x^k),\ v^k := \left(\mu_k, (\lambda^g)^k\right),\ w^k := \left(\mu_k, (\lambda^g)^k, g(x^k)\right).$$

2 Preliminaries

Let $x^* \in \mathcal{F}$. We define the index sets:

$$\begin{array}{ll} \mathcal{I}^m := \{1, 2, \ldots, m\}, \\ \mathcal{I}^p := \{1, 2, \ldots, p\}, \\ \mathcal{I}^q := \{1, 2, \ldots, q\}, \\ \mathcal{I}^g := \{i \in \mathcal{I}^p : g_i(x^*) = 0\}, \\ \mathcal{I}^G := \{i \in \mathcal{I}^m : G_i(x^*) = 0 < H_i(x^*)\}, \\ \mathcal{I}^H := \{i \in \mathcal{I}^m : G_i(x^*) > 0 = H_i(x^*)\}, \\ \mathcal{I}^{GH} := \{i \in \mathcal{I}^m : G_i(x^*) = 0 = H_i(x^*)\}. \end{array} \tag{2.1}$$

If x^* is a solution of (1.1) and $\mathcal{I}^{GH} = \emptyset$, then x^* is defined as a *strict complementary solution* of (1.1). We say that the MPCC (1.1) satisfies the strict complementarity condition if an accumulation point of the sequence generated by an iterative algorithm is a strict complementary solution of (1.1). Such an assumption is often used in the convergence analysis of algorithms for solving the MPCC.

The following relaxed problem of (1.1) at x^* is frequently used in the convergence analysis of some algorithms for solving the MPCC (1.1):

$$\begin{array}{rll} \text{Minimize} & f(x) & \\ \text{Subject to} & g_i(x) \le 0, & \forall\, i \in \mathcal{I}^p, \\ & h_i(x) \le 0, & \forall\, i \in \mathcal{I}^q, \\ & G_i(x) = 0, & \forall\, i \in \mathcal{I}^G, \\ & H_i(x) = 0, & \forall\, i \in \mathcal{I}^H, \\ & G_i(x) \ge 0, & \forall\, i \in \mathcal{I}^{GH}, \\ & H_i(x) \ge 0, & \forall\, i \in \mathcal{I}^{GH}. \end{array} \tag{2.2}$$

We say that *the linear independence constraint qualification (LICQ) for the MPCC* (1.1) *(MPCC-LICQ for short) holds at* $x^* \in \mathcal{F}$ if

$$\begin{array}{c} \{\nabla G_i(x^*) : i \in \mathcal{I}^G \cup \mathcal{I}^{GH}\} \cup \{\nabla H_i(x^*) : i \in \mathcal{I}^H \cup \mathcal{I}^{GH}\} \\ \cup \{\nabla g_i(x^*) : i \in \mathcal{I}^g\} \cup \{\nabla h_i(x^*) : i \in \mathcal{I}^q\} \end{array}$$

are linearly independent.

Definition 2.1 *Let $x^* \in \mathcal{F}$. x^* is called a B-stationary point of the MPCC* (1.1) *if*

$$\nabla f(x^*)^T d \ge 0, \quad \forall\, d \in T_{\mathcal{F}}(x^*),$$

where $T_{\mathcal{F}}(x^*)$ is the contingent cone of $\mathcal{F}$ at x^* is defined as

$$T_{\mathcal{F}}(x^*) := \{d \in \Re^n : \exists\, t_k \downarrow 0 \text{ and } x^k \in \mathcal{F} \text{ such that } \lim_{k\to\infty} (x^k - x^*)/t_k = d\}.$$

We say that *the Karush-Kuhn-Tucker (KKT) conditions for the relaxed problem* (2.2) *at* $x^* \in \mathcal{F}$

is satisfied if there exists a vector $(\Lambda^*, \Gamma^*, \Omega^*, \Pi^*) \in \Re^{n+m+p+q}$ such that

$$\begin{aligned}
&\bigtriangledown f(x^*) - \textstyle\sum_{i\in\mathcal{I}^G\cup\mathcal{I}^{GH}} \Lambda_i^* \bigtriangledown G_i(x^*) - \sum_{i\in\mathcal{I}^H\cup\mathcal{I}^{GH}} \Gamma_i^* \bigtriangledown H_i(x^*) \\
&\quad + \textstyle\sum_{i\in\mathcal{I}^p} \Omega_i^* \bigtriangledown g_i(x^*) + \sum_{i\in\mathcal{I}^q} \Pi_i^* \bigtriangledown h_i(x^*) = 0, \\
&\Lambda^*_{\mathcal{I}^{GH}} \geq 0,\ G_{\mathcal{I}^{GH}}(x^*) \geq 0,\ (\Lambda^*_{\mathcal{I}^{GH}})^T G_{\mathcal{I}^{GH}}(x^*) = 0, \\
&\Gamma^*_{\mathcal{I}^{GH}} \geq 0,\ H_{\mathcal{I}^{GH}}(x^*) \geq 0,\ (\Gamma^*_{\mathcal{I}^{GH}})^T H_{\mathcal{I}^{GH}}(x^*) = 0, \\
&\Omega^* \geq 0,\ g(x^*) \leq 0,\ (\Omega^*)^T g(x^*) = 0, \\
&G_{\mathcal{I}^G}(x^*) = 0,\ H_{\mathcal{I}^H}(x^*) = 0,\ h(x^*) = 0.
\end{aligned} \tag{2.3}$$

Proposition 2.1 *[20, Proposition 4.3.7] If $x^* \in \mathcal{F}$ satisfies the KKT conditions* (2.3) *for the relaxed problem* (2.2)*, and if the MPCC-LICQ holds at x^*, then x^* is a B-stationary point of the problem* (1.1)*.*

Let

$$\phi(\mu, a, b) := a + b - \sqrt{a^2 + b^2 + 4\mu^2}, \quad \forall\ (\mu, a, b) \in \Re^3 \tag{2.4}$$

and

$$\Phi(u) := \text{vec}\,\{\phi(\mu, G_i(x), H_i(x)) : i \in \mathcal{I}^m\}, \quad \forall\, u := (\mu, x) \in \Re \times \Re^n. \tag{2.5}$$

Then, for any $u := (\mu, x) \in \Re_+ \times \Re^n$ with $\mu = 0$, the function Φ is not differentiable in general. Specifically, when $G_i(x^*) = H_i(x^*) = 0$ for some x^* and $i \in \mathcal{I}^{GH}$, the function Φ is not differentiable at $(0, x^*)$ and the Clarke generalized gradient (see, [3])

$$\partial_x \Phi_i(0, x^*) := \{r \in \Re^n : r = \lim_{k\to\infty} \bigtriangledown_x \Phi_i(0, x^k) \text{ with } x^k \to x^* \text{ and } \bigtriangledown_x \Phi_i(0, x^k) \text{ exists}\}$$

is contained in the set

$$(\partial_C)_x \Phi_i(0, x^*) := \{r \in \Re^n : r = \xi_i \bigtriangledown G_i(x^*) + \eta_i \bigtriangledown H_i(x^*) \text{ for some } (\xi, \eta) \in \mathcal{B}\}, \tag{2.6}$$

where $\mathcal{B} := \{(\xi, \eta) \in \Re^2 : (1 - \xi_i)^2 + (1 - \eta_i)^2 \leq 1\}$.

By using (2.5), the MPCC (1.1) can be approximated by a perturbed problem:

$$\begin{array}{rl}
\text{Minimize} & f(x) \\
\text{Subject to} & g(x) \leq 0,\ h(x) = 0,\ \Phi(u) = 0.
\end{array} \tag{2.7}$$

For any given $\mu > 0$, since the problem (2.7) is an ordinary NLP, under some constraint qualification, solving (2.7) is equivalent to solving the KKT system and there exist Lagrange multipliers $\lambda^\phi \in \Re^m$, $\lambda^g \in \Re^p$, and $\lambda^h \in \Re^q$ such that

$$\begin{aligned}
&\bigtriangledown f(x) - \textstyle\sum_{i\in\mathcal{I}^m} \lambda_i^\phi \bigtriangledown_x \Phi_i(u) + \sum_{i\in\mathcal{I}^p} \lambda_i^g \bigtriangledown g_i(x) + \sum_{i\in\mathcal{I}^q} \lambda_i^h \bigtriangledown h_i(x) = 0, \\
&\Phi(u) = 0, \\
&\lambda^g \geq 0, g(x) \leq 0, (\lambda^g)^T g(x) = 0, \\
&h(x) = 0,
\end{aligned} \tag{2.8}$$

where

$$\begin{aligned}\bigtriangledown_x \Phi_i(u) &= \left(1 - \frac{G_i(x)}{\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}}\right) \bigtriangledown G_i(x) \\ &\quad + \left(1 - \frac{H_i(x)}{\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}}\right) \bigtriangledown H_i(x), \quad \forall i \in \mathcal{I}^m.\end{aligned}$$

Most existing smoothing methods for the MPCC solve (2.7) by using some standard methods for the solution of the NLP (e.g., the sequential quadratic programming method), and make the smoothing parameter μ tend to zero (see, for example, [4, 14, 18]). We will use a different approach to solve the MPCC (1.1), which is discussed in the following two sections.

3 Smoothing Reformulation of the MPCC

In addition to the smoothing function ϕ defined by (2.4), we use two additional smoothing functions to construct a system of smoothing equations so that a Newton-type method can be applied to solve the this system to obtain a solution to (1.1). Let

$$\Theta(v) := \mathrm{vec}\{\theta(\mu, \lambda_i^g) : i \in \mathcal{I}^g\}, \quad \text{where } \theta(\mu, \lambda_i^g) := \left(\lambda_i^g + \sqrt{(\lambda_i^g)^2 + 4\mu^2}\right)/2, \quad \forall\, i \in \mathcal{I}^g \tag{3.1}$$

and

$$\begin{aligned}&\Psi(w) := \mathrm{vec}\{\psi(\mu, \lambda_i^g, -g_i(x)) : i \in \mathcal{I}^g\}, \quad \text{where} \\ &\psi(\mu, \lambda_i^g, -g_i(x)) := \lambda_i^g - g_i(x) - \sqrt{(\lambda_i^g + g_i(x))^2 + 4\mu^2}, \quad \forall\, i \in \mathcal{I}^g,\ \forall\, x \in \Re^n.\end{aligned} \tag{3.2}$$

Denote

$$L(z) := f(x) - \sum_{i \in \mathcal{I}^m} \lambda_i^\phi \Phi_i(u) + \sum_{i \in \mathcal{I}^p} \Theta_i(v) g_i(x) + \sum_{i \in \mathcal{I}^q} \lambda_i^h h_i(x). \tag{3.3}$$

Suppose that $c > 0$ is a given scalar. Let

$$H(z) := \begin{bmatrix} \mu \\ \bigtriangledown_x L(z) \\ \Phi(u) \\ \Psi(w) \\ -h(x) \end{bmatrix} + c\mu \begin{bmatrix} 0 \\ x \\ \lambda^\phi \\ \lambda^g \\ \lambda^h \end{bmatrix}, \tag{3.4}$$

where

$$\bigtriangledown_x L(z) := \bigtriangledown f(x) - \sum_{i \in \mathcal{I}^m} \lambda_i^\phi \bigtriangledown_x \Phi_i(u) + \sum_{i \in \mathcal{I}^p} \Theta_i(v) \bigtriangledown g_i(x) + \sum_{i \in \mathcal{I}^q} \lambda_i^h \bigtriangledown h_i(x).$$

The following proposition is the basis for our smoothing algorithm for solving the MPCC (1.1).

Proposition 3.1 *solving $H(z) = 0$ is equivalent to solving the KKT system* (2.8).

Proof. The result can be easily proved. We omit it here. $\square$

It is not difficult to see that the function H is continuously differentiable at any

$$z := (\mu, y) := (\mu, x, \lambda^\phi, \lambda^g, \lambda^h) \in \Re^{1+n+m+p+q}$$

with $\mu > 0$. Let H' denote the Jacobian matrix of H at $z \in \Re^{1+n+m+p+q}$ with $\mu > 0$. Then

$$H'(z) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ (\bigtriangledown_x L(z))'_\mu & \bigtriangledown_x^2 L(z)^T + cE^n & -\bigtriangledown \Phi(u)^T & (\bigtriangledown_{\lambda^g}\Theta(v) \bigtriangledown g(x))^T & \bigtriangledown h(x)^T \\ (\Phi(u))'_\mu & \bigtriangledown_x \Phi(u) & cE^m & 0 & 0 \\ (\Psi(w))'_\mu & -\bigtriangledown_g \Psi(w) \bigtriangledown g(x) & 0 & \bigtriangledown_{\lambda^g}\Psi(w) + cE^p & 0 \\ 0 & -\bigtriangledown h(x) & 0 & 0 & cE^q \end{bmatrix}, \tag{3.5}$$

where

$$\begin{aligned}
(\bigtriangledown_x L(z))'_\mu &= -\sum_{i\in\mathcal{I}^m} \lambda_i^\phi (\bigtriangledown_x \Phi_i(u))'_\mu + \sum_{i\in\mathcal{I}^p} (\Theta_i(v))'_\mu \bigtriangledown g_i(x), \\
(\bigtriangledown_x \Phi_i(u))'_\mu &= \frac{4\mu G_i(x)}{\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}} \bigtriangledown G_i(x) + \frac{4\mu H_i(x)}{\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}} \bigtriangledown H_i(x), \\
(\Theta_i(v))'_\mu &= \frac{4\mu}{\sqrt{(\lambda_i^g)^2 + 4\mu^2}}; \\
\bigtriangledown_x^2 L(z) &= \bigtriangledown^2 f(x) - \sum_{i\in\mathcal{I}^m} \lambda_i^\phi \bigtriangledown_x^2 \Phi_i(u) + \sum_{i\in\mathcal{I}^p} \Theta_i(v) \bigtriangledown^2 g_i(x) + \sum_{i\in\mathcal{I}^q} \lambda_i^h \bigtriangledown^2 h_i(x), \\
\bigtriangledown_x^2 \Phi_i(u) &= \frac{\partial_x \Phi_i(u)}{\partial G} \bigtriangledown^2 G_i(x) + \frac{\partial_x^2 \Phi_i(u)}{\partial G^2} \bigtriangledown G_i(x) \bigtriangledown G_i(x)^T + \frac{\partial_x^2 \Phi_i(u)}{\partial G \partial H} \bigtriangledown G_i(x) \bigtriangledown H_i(x)^T \\
&+ \frac{\partial_x^2 \Phi_i(u)}{\partial H \partial G} \bigtriangledown H_i(x) \bigtriangledown G_i(x)^T + \frac{\partial_x^2 \Phi_i(u)}{\partial H^2} \bigtriangledown H_i(x) \bigtriangledown H_i(x)^T + \frac{\partial_x \Phi_i(u)}{\partial H} \bigtriangledown^2 H_i(x), \\
\frac{\partial_x \Phi_i(u)}{\partial G} &= 1 - \frac{G_i(x)}{\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}}, \quad \frac{\partial_x \Phi_i(u)}{\partial H} = 1 - \frac{H_i(x)}{\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}}, \\
\frac{\partial_x^2 \Phi_i(u)}{\partial G^2} &= \frac{H_i(x)^2 + 4\mu^2}{(\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2})^3}, \quad \frac{\partial_x^2 \Phi_i(u)}{\partial H^2} = \frac{G_i(x)^2 + 4\mu^2}{(\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2})^3}, \\
\frac{\partial_x^2 \Phi_i(u)}{\partial G \partial H} &= \frac{G_i(x) H_i(x)}{(\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2})^3} = \frac{\partial_x^2 \Phi_i(u)}{\partial H \partial G};
\end{aligned}$$

$$\begin{aligned}
\bigtriangledown_{\lambda^g}\Theta(v) &= \operatorname{diag}\{(\theta(\mu, \lambda_i^g))'_\mu : i \in \mathcal{I}^g\}, \quad (\theta(\mu, \lambda_i^g))'_\mu = 2\mu/\sqrt{(\lambda_i^g)^2 + 4\mu^2}; \\
(\Phi(u))'_\mu &= \operatorname{vec}\{(\phi(\mu, G_i(x), H_i(x)))'_\mu : i \in \mathcal{I}^m\}, \quad (\phi(\mu, G_i(x), H_i(x)))'_\mu = -4\mu/\sqrt{G_i(x)^2 + H_i(x)^2 + 4\mu^2}; \\
(\Psi(w))'_\mu &= \operatorname{vec}\{(\Psi_i(w))'_\mu : i \in \mathcal{I}^m\}, \quad (\Psi_i(w))'_\mu = -4\mu/\sqrt{(\lambda_i^g + g_i(x))^2 + 4\mu^2}; \\
(\Psi(w))'_g &= \operatorname{diag}\{(\Psi_i(w))'_g : i \in \mathcal{I}^m\}, \quad (\Psi_i(w))'_g = -1 + \frac{\lambda_i^g + g_i(x)}{\sqrt{(\lambda_i^g + g_i(x))^2 + 4\mu^2}}; \\
(\Psi(w))'_{\lambda^g} &= \operatorname{diag}\{(\Psi_i(w))'_{\lambda^g} : i \in \mathcal{I}^m\}, \quad \text{and} \quad (\Psi_i(w))'_{\lambda^g} = 1 + \frac{\lambda_i^g + g_i(x)}{\sqrt{(\lambda_i^g + g_i(x))^2 + 4\mu^2}}.
\end{aligned}$$

Thus, Newton-type methods can be applied to solve the equation $H(z) = 0$. In the next section, we propose a smoothing Newton algorithm to solve the equation $H(z) = 0$ while making μ tend to zero.

4 Algorithm Description and Convergence Analysis

For any given $\beta, \gamma \in (0,1)$ and $c \in (0,\infty)$, we let $\alpha(z) := \beta\|H(z)\| \min\{1, \|H(z)\|^\gamma\}$.

Algorithm 4.1 *(A Smoothing Newton Algorithm)*

Step 0 *Choose* $\delta \in (0,1), \gamma \in (0,1]$, *and* $\mu_0, c \in (0,+\infty)$. *Let* $\left(x^0, (\lambda^\phi)^0, (\lambda^g)^0, (\lambda^h)^0\right) \in \Re^{n+m+p+q}$ *be an arbitrary vector. Set*

$$e^1 := (1,0,\dots,0) \in \Re^{1+n+m+p+q} \quad \text{and} \quad z^0 := \left(\mu_0, x^0, (\lambda^\phi)^0, (\lambda^g)^0, (\lambda^h)^0\right).$$

Choose $\beta \in (0,1)$ *such that* $\beta\|H(z^0)\| \le \mu_0$. *Set* $k := 0$.

Step 1 *If* $\|H(z^k)\| = 0$, *stop. Otherwise, set* $\Upsilon(z^k) := \alpha(z^k)e^1$.

Step 2 *Compute* $\Delta z^k := \left(\Delta\mu_k, \Delta x^k, \Delta(\lambda^\phi)^k, \Delta(\lambda^g)^k, \Delta(\lambda^h)^k\right) \in \Re^{1+n+m+p+q}$ *by*

$$H'(z^k)\Delta z^k = -H(z^k) + \Upsilon(z^k). \tag{4.1}$$

Step 3 *Let* ξ_k *be the maximum of the values* $1, \delta, \delta^2, \cdots$ *such that*

$$\|H(z^k + \xi_k \Delta z^k)\| \le \{1 - \sigma(1-\beta)\xi_k\} \|H(z^k)\|. \tag{4.2}$$

Step 4 *Set* $z^{k+1} := z^k + \xi_k \Delta z^k$ *and* $k := k+1$. *Go to Step 1.*

Similar framework of algorithms for solving variational inequalities and complementarity problems has been discussed extensively in the literature. See, for example, [13] and references therein.

Assumption 4.1 *Let* $\{z^k\}$ *be generated by Algorithm 4.1. Then* $H'(z^k)$ *is nonsingular for all* k.

Nonsingularity of Jacobian matrix of the function involved in Newton equation is an essential requirement in the Newton methods. The function H defined by (3.4) is a regularized reformulation function. Proper selection of the regularized parameter c is useful for improving nonsingularity of the Jacobian matrix (see, (3.5)). The following proposition give a sufficient condition of Assumption 4.1.

Proposition 4.1 *Let* $\{z^k\}$ *be generated by Algorithm 4.1. If* $(d^k)^T \bigtriangledown_x^2 L(z^k) d^k \ge 0$ *for any iteration index* k *and any* $d^k \in \Re^{1+n+m+p+q}$, *then Assumption 4.1 is satisfied.*

Proof. For any iteration index $k \ge 0$, let $dz := \left(d\mu, \Delta x, d\lambda^\phi, d\lambda^g, d\lambda^h\right) \in \Re^{1+n+m+p+q}$ satisfy $H'(z^k)dz = 0$. Then, it follows from (3.5) that $d\mu_k = 0$ and

$$\begin{aligned}
&(\bigtriangledown_x^2 L(z^k)^T + cE^n)dx - \bigtriangledown_x \Phi(u^k)^T d\lambda^\phi + (\bigtriangledown_{\lambda^g}\Theta(v^k) \bigtriangledown g(x^k))^T d\lambda^g + \bigtriangledown h(x^k)^T d\lambda^h = 0,\\
&\bigtriangledown_x \Phi(u^k)dx + cE^m d\lambda^\phi = 0,\\
&- \bigtriangledown_g \Psi(w^k) \bigtriangledown g(x^k)dx + (\bigtriangledown_{\lambda^g}\Psi(w^k) + cE^p)d\lambda^g = 0,\\
&- \bigtriangledown h(x^k)dx + cE^q d\lambda^h = 0.
\end{aligned} \tag{4.3}$$

By the last three equalities in (4.3), we have

$$\begin{array}{rcl} -(d\lambda^\phi)^T \bigtriangledown_x \Phi(u^k)dx &=& c(d\lambda^\phi)^T d\lambda^\phi, \\ (d\lambda^g)^T \bigtriangledown_{\lambda^g} \Theta(v^k) \bigtriangledown g(x^k)dx &=& (d\lambda^g)^T \bigtriangledown_{\lambda^g} \Theta(v^k)(\bigtriangledown_g \Psi(w^k))^{-1}(\bigtriangledown_{\lambda^g} \Psi(w^k) + cE^p)d\lambda^g, \\ (d\lambda^h)^T \bigtriangledown h(x^k)dx &=& c(d\lambda^h)^T d\lambda^h. \end{array}$$

From the first equality in (4.3) and the above equalities, we have

$$\begin{array}{l} (dx)^T(\bigtriangledown_x^2 L(z^k)^T + cE^n)dx + c(d\lambda^\phi)^T d\lambda^\phi \\ \quad +(d\lambda^g)^T \bigtriangledown_{\lambda^g} \Theta(v^k)(\bigtriangledown_g \Psi(w^k))^{-1}(\bigtriangledown_{\lambda^g} \Psi(w^k) + cE^p)d\lambda^g + c(d\lambda^h)^T d\lambda^h = 0. \end{array} \quad (4.4)$$

For any iteration index $k \geq 0$, since $\bigtriangledown_x^2 L(z^k)^T + cE^n$ and $\bigtriangledown_{\lambda^g}\Theta(v^k)(\bigtriangledown_g \Psi(w^k))^{-1}(\bigtriangledown_{\lambda^g}\Psi(w^k)+cE^p)$ are positive definite, it follows from (4.4) that

$$dx = 0, \quad d\lambda^\phi = 0, \quad d\lambda^g = 0, \quad \text{and} \quad d\lambda^h = 0.$$

Thus, the matrix $H'(z^k)$ is nonsingular for all $k \geq 0$. □

It is not difficult to see that the assumption given in Proposition 4.1 is trivially satisfied for the case that $G(x) = H(x) = 0$, f and g are convex functions, and h is a linear function (i.e., the case that the MPCC is a convex NLP).

Let $\mathcal{N}(\beta) := \{z \in \Re_{++} \times \Re^{n+m+p+q} : \alpha(z) \leq \mu\}$.

Theorem 4.1 *Let Assumption 4.1 be satisfied. Then,*

(i) Algorithm 4.1 is well-defined.

(ii) $\mu_k > 0$ *for any iteration index* k.

(iii) $\{\|H(z^k)\|\}$ *is monotonically decreasing.*

(iv) $z^k \in \mathcal{N}(\beta)$ *for any iteration index* k.

(v) $\{\mu_k\}$ *is monotonically decreasing.*

Proof. It is sufficient to show that results (i)-(v) hold for $l = k + 1$ under the assumption that results (i)-(v) hold for $l = k$ for $k > 0$. In the following, we assume that results (i)-(v) hold for $l = k$ for $k > 0$.

(i) To show that Algorithm 4.1 is well-defined for $l = k+1$, we need to show that equation (4.1) is solvable and line search (4.2) is finitely terminated for $l = k+1$. Since Proposition 4.1 implies that equation (4.1) is solvable, we only need to show that line search (4.2) is well-defined for $l = k + 1$. Let

$$R(z^k) := H(z^k + \xi\Delta z^k) - H(z^k) - \xi H'(z^k)\Delta z^k.$$

Thus, by (4.1),

$$\begin{array}{rcl} \|H(z^k + \xi\Delta z^k)\| &=& \|R(z^k) + H(z^k) + \xi H'(z^k)\Delta z^k\| \\ &\leq& \|R(z^k)\| + \{1 - \xi(1 - \beta)\}\|H(z^k)\|. \end{array} \quad (4.5)$$

Since H is continuously differentiable for $\mu_k > 0$, we have $\|R(z^k)\| = o(\xi)$. Thus,, it follows from (4.5) that line search (4.2) is well-defined for $l = k + 1$. Therefore, Algorithm 4.1 is well-defined.

(ii) From the first equation in (4.1), it follows that

$$\mu_{k+1} = \mu_k + \xi_k \Delta\mu_k = (1 - \xi_k)\mu_k + \xi_k \alpha(z^k) > 0, \tag{4.6}$$

which implies that the result (ii) holds for $l = k + 1$.

(iii) The result (iii) holds directly from (4.2).

(iv) By (iii), the sequence $\{\alpha(z^k)\}$ is monotonically decreasing. Thus,

$$\mu_{k+1} - \alpha(z^{k+1}) = (1 - \xi_k)\mu_k + \xi_k \alpha(z^k) - \alpha(z^{k+1}) \leq \alpha(z^k) - \alpha(z^{k+1})$$

implies that $\alpha(z^{k+1}) \leq \mu_{k+1}$. That is, the result (iv) holds for $l = k + 1$.

(v) From (4.6) and the result (iv), we have

$$\mu_{k+1} = (1 - \xi_k)\mu_k + \xi_k \alpha(z^k) \leq (1 - \xi_k)\mu_k + \xi_k \mu_k = \mu_k,$$

which implies that the result (v) holds for $l = k + 1$.

The proof is complete. □

Assumption 4.2 *The function H defined by* (3.4) *is coercive, i.e., for any given $\hat{\mu}$ and $\tilde{\mu}$ with $0 < \hat{\mu} \leq \tilde{\mu}$, and any sequence $\{z^k\}$ with $z^k := (\mu_k, y^k) \in \Re_{++} \times \Re^{n+m+p+q}$ satisfying $\mu_k \in [\hat{\mu}, \tilde{\mu}]$ for all k, it follows that* $\lim_{\|y^k\| \to +\infty} \|H(z^k)\| = +\infty$.

Assumption 4.3 $\mathcal{S} := \{y \in \Re^{n+m+p+q} : H(z) = 0\}$ *is nonempty and bounded.*

Remarks on Assumptions 4.2 and 4.3

1. For the regularized smoothing method to solve variational inequality and complementarity problems,
   - Assumption 4.2 is shown to be satisfied if the function involved in the problems is a P_0-function. See, for example, [21, Corollary 4.3] and [12, Lemma 2.5];
   - Similar conditions to Assumption 4.3 have been used extensively in the literature. See, for example, [21, the assumption in Lemma 4.4], [12, Assumption 1.1], and [13, Assumption 1]. Also see some discussions for the related conditions [11, 25].

2. For the smoothing method to solve the MPCC given in this paper, Assumptions 4.2 and 4.3 become relatively complex due to the complexity of the problem concerned. As an example, we give in the following some observations on Assumption 4.2. For any given $\hat{\mu}$ and $\tilde{\mu}$ with $0 < \hat{\mu} \leq \tilde{\mu}$, if for any sequence $\{z^k\}$ with $z^k := (\mu_k, y^k) \in \Re_{++} \times \Re^{n+m+p+q}$ satisfying $\mu_k \in [\hat{\mu}, \tilde{\mu}]$ for all k, one of the following conditions (A), (B), (C), and (D) is satisfied, then Assumption 4.2 holds.

   (A) If $\{\|(\nabla_x L(z^k), x^k)\|\} \to +\infty$ as $k \to \infty$, there exists an index $i_0 \in \{1, 2, \ldots, n\}$ such that

   (a) $|(\nabla_x L(z^k))_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(\nabla_x L(z^k))_{i_0}|}{|x^k_{i_0}|} = +\infty$; or

(b) $|x^k_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|x^k_{i_0}|}{|(\bigtriangledown_x L(z^k))_{i_0}|} = +\infty$; or

(c) $|(\bigtriangledown_x L(z^k))_{i_0}| \to +\infty$, $|x^k_{i_0}| \to +\infty$, and there exists a constant $c_{i_0} > 0$ such that $\lim_{k\to\infty} \frac{(\bigtriangledown_x L(z^k))_{i_0}}{x^k_{i_0}} = c_{i_0}$; or $\lim_{k\to\infty} \frac{(\bigtriangledown_x L(z^k))_{i_0}}{x^k_{i_0}} = -c_{i_0}$ and $\lim_{k\to\infty}(-c_{i_0} + c\mu_k) \not\to 0$.

(B) If $\{\|(\Phi(u^k), (\lambda^\phi)^k)\|\} \to +\infty$ as $k \to \infty$, there exists an index $i_0 \in \{1, 2, \ldots, n\}$ such that

(a) $|(\Phi(u^k))_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(\Phi(u^k))_{i_0}|}{|(\lambda^\phi)^k_{i_0}|} = +\infty$; or

(b) $|(\lambda^\phi)^k_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(\lambda^\phi)^k_{i_0}|}{|(\Phi(u^k))_{i_0}|} = +\infty$; or

(c) $|(\Phi(u^k))_{i_0}| \to +\infty$, $|(\lambda^\phi)^k_{i_0}| \to +\infty$, and there exists a constant $c_{i_0} > 0$ such that $\lim_{k\to\infty} \frac{(\Phi(u^k))_{i_0}}{(\lambda^\phi)^k_{i_0}} = c_{i_0}$; or $\lim_{k\to\infty} \frac{(\Phi(u^k))_{i_0}}{(\lambda^\phi)^k_{i_0}} = -c_{i_0}$ and $\lim_{k\to\infty}(-c_{i_0} + c\mu_k) \not\to 0$.

(C) If $\{\|(\Psi(w^k), (\lambda^g)^k)\|\} \to +\infty$ as $k \to \infty$, then there exists an index $i_0 \in \{1, 2, \ldots, n\}$ such that

(a) $|(\Psi(w^k))_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(\Psi(w^k))_{i_0}|}{|(\lambda^g)^k_{i_0}|} = +\infty$; or

(b) $|(\lambda^g)^k_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(\lambda^g)^k_{i_0}|}{|(\Psi(w^k))_{i_0}|} = +\infty$; or

(c) $|(\Psi(w^k))_{i_0}| \to +\infty$, $|(\lambda^g)^k_{i_0}| \to +\infty$, and there exists a constant $c_{i_0} > 0$ such that $\lim_{k\to\infty} \frac{(\Psi(w^k))_{i_0}}{(\lambda^g)^k_{i_0}} = c_{i_0}$; or $\lim_{k\to\infty} \frac{(\Psi(w^k))_{i_0}}{(\lambda^g)^k_{i_0}} = -c_{i_0}$ and $\lim_{k\to\infty}(-c_{i_0} + c\mu_k) \not\to 0$.

(D) If $\{\|(-h(x^k), (\lambda^h)^k)\|\} \to +\infty$ as $k \to \infty$, there exists an index $i_0 \in \{1, 2, \ldots, n\}$ such that

(a) $|(h(x^k))_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(h(x^k))_{i_0}|}{|(\lambda^h)^k_{i_0}|} = +\infty$; or

(b) $|(\lambda^h)^k_{i_0}| \to +\infty$ and $\lim_{k\to\infty} \frac{|(\lambda^h)^k_{i_0}|}{|(h(x^k))_{i_0}|} = +\infty$; or

(c) $|(h(x^k))_{i_0}| \to +\infty$, $|(\lambda^h)^k_{i_0}| \to +\infty$, and there exists a constant $c_{i_0} > 0$ such that $\lim_{k\to\infty} \frac{-(h(x^k))_{i_0}}{(\lambda^h)^k_{i_0}} = c_{i_0}$; or $\lim_{k\to\infty} \frac{-(h(x^k))_{i_0}}{(\lambda^h)^k_{i_0}} = -c_{i_0}$ and $\lim_{k\to\infty}(-c_{i_0} + c\mu_k) \not\to 0$.

These results can be easily proved by using (3.4).

We discuss in the following the global convergence of Algorithm 4.1.

Lemma 4.1 *Suppose that Assumptions 4.1-4.3 are satisfied, and that the sequence* $\{z^k := (\mu_k, y^k) := (\mu_k, x^k, (\lambda^\phi)^k, (\lambda^g)^k, (\lambda^h)^k)\}$ *is generated by Algorithm 4.1. Then,*

(i) $\lim_{k\to\infty} \|H(z^k)\| = 0$ *and* $\lim_{k\to\infty} \mu_k = 0$, *and*

(ii) the sequence $\{z^k\}$ *is bounded.*

Proof. (i) By Theorem 4.1(iii), there exist $H^* \geq 0$ and $\alpha^* \geq 0$ such that $\lim_{k\to\infty} \|H(z^k)\| = H^*$ and $\lim_{k\to\infty} \alpha(z^k) = \alpha^*$. We now prove that $H^* = 0$. Assume, on the contrary, that $H^* > 0$. Then, from Theorem 4.1(iv)(v) it follows that $0 < \alpha^* \leq \alpha(z^k) \leq \mu_k \leq \mu_0$. Thus, by using Assumption 4.2 we know that $\{z^k\}$ is bounded. Subsequencing if necessary, there exists $z^* := (\mu_*, y^*) \in$

$\Re \times \Re^{n+m+p+q}$ with $\mu_* = 0$ such that $\lim_{k\to\infty} z^k = z^*$. Then, $\lim_{k\to\infty} \|H(z^k)\| = \|H(z^*)\| = H^*$. By (4.1) we have

$$H'(z^*)\Delta z^* = -H(z^*) + \Upsilon(z^*). \tag{4.7}$$

Since $H^* > 0$, by Algorithm 4.1 it follows that $\xi_k \to \infty$ as $k \to \infty$. Hence, the stepsize $\hat{\xi}_k := \xi_k/\delta$ does not satisfy line search criterion (4.2) for all sufficiently large k, which yields

$$\left(\|H(z^k + \hat{\xi}_k \Delta z^k)\| - \|H(z^k)\|\right)/\hat{\xi}_k > -\sigma(1-\beta)\|H(z^k)\|,$$

and hence, $H(z^*)^T H'(z^*)\Delta z^*/\|H(z^*)\| \geq -\sigma(1-\beta)\|H(z^*)\|$. This, together with (4.7), implies that $-1+\beta+\sigma(1-\beta) \geq 0$, which contradicts the fact that $\sigma, \beta \in (0,1)$. Therefore, $H(z^*) = 0$. Moreover, by Theorem 4.1(v) and the definition of H, it follows that $\lim_{k\to\infty} \mu_k = 0$.

(ii) Define $H_0 : \Re^{n+m+p+q} \to \Re^{n+m+p+q}$ by $H_0(y) := H(0,y)$ for all $y \in \Re^{n+m+p+q}$, where the function H is defined by (3.4). Then, both H_0 and H are weakly univalent functions (see the related definition in [8]). Since Assumption 4.3 implies that the inverse image $H_0^{-1}(0)$ is nonempty and bounded, by using Theorem 2.5 in [22] we can obtain that the sequence $\{z^k\}$ is bounded. $\square$

Suppose that Assumptions 4.1-4.3 are satisfied, then by Lemma 4.1(ii), the iteration sequence $\{z^k\}$ has an accumulation point. In the remainder of this paper, we assume that

$$z^* := \left(\mu_*, x^*, (\lambda^\phi)^*, (\lambda^g)^*, (\lambda^h)^*\right) \in \Re^{1+n+m+p+q}$$

is an arbitrary accumulation point of $\{z^k\}$, and, by Lemma 4.1(i), $\mu_* = 0$.

Assumption 4.4 *The MPCC-LICQ is satisfied at x^*.*

It is easy to show that, for each $i \in \mathcal{I}^{GH}$, any accumulation point r^* of $\nabla_x \Phi_i(u^k)$ belongs to the set $(\partial_C)_x \Phi_i(0, x^*)$ defined by (2.6) and hence is represented as $r^* = \xi_i^* \nabla G_i(x^*) + \eta_i^* \nabla H_i(x^*)$ for some (ξ_i^*, η_i^*) such that $(1-\xi_i^*)^2 + (1-\eta_i^*)^2 \leq 1$.

Assumption 4.5 *The sequence $\{z^k\}$ is asymptotically weakly nondegenerate, i.e., neither ξ_i^* or η_i^* vanishes at any accumulation point r^* of $\nabla_x \Phi_i(u^k)$.*

Assumption 4.6 *$(d^k)^T \nabla_x^2 L(z^k) d^k \geq 0$ for sufficiently large k and any d^k.*

Now, we discuss the convergence of Algorithm 4.1.

Theorem 4.2 *Suppose that Assumptions 4.1-4.6 are satisfied, and that the sequence $\{z^k\}$ is generated by Algorithm 4.1. Then,*

(i) any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$, and

(ii) x^* is a B-stationary point of the problem (1.1).

Proof. The result (i) holds directly from Lemma 4.1(i) and a simple continuity argument. We only need to show the result (ii), which to some extent is motivated by the one in [6].

It is not difficult to show that

$$\partial_x \Phi_i(u^k) \to \nabla G_i(x^*),\ \forall\, i \in \mathcal{I}^G \quad \text{and} \quad \partial_x \Phi_i(u^k) \to \nabla H_i(x^*),\ \forall\, i \in \mathcal{I}^H$$

as $k \to \infty$. On the other hand, for each $i \in \mathcal{I}^{GH}$, let r^* be an arbitrary accumulation point of $\{\nabla_x \Phi(u^k)\}$, and hence, by Assumption 4.5, we have

$$r^* = \xi_i^* \nabla G_i(x^*) + \eta_i^* \nabla H_i(x^*), \quad \forall\, (\xi_i^*, \eta_i^*) > 0 \ \text{with}\ (1-\xi_i^*)^2 + (1-\eta_i^*)^2 \leq 1.$$

From the first two equations in $H(z^*) = 0$, it follows that

$$\begin{array}{c} \nabla f(x^*) - \sum_{i\in\mathcal{I}^G} (\lambda^\phi)_i^* \nabla G_i(x^*) - \sum_{i\in\mathcal{I}^{GH}} (\lambda^\phi)_i^* \xi_i^* \nabla G_i(x^*) - \sum_{i\in\mathcal{I}^{GH}} (\lambda^\phi)_i^* \eta_i^* \nabla H_i(x^*) \\ - \sum_{i\in\mathcal{I}^H} (\lambda^\phi)_i^* \nabla H_i(x^*) + \sum_{i\in\mathcal{I}^p} (\lambda^g)_i^* \nabla g_i(x^*) + \sum_{i\in\mathcal{I}^q} (\lambda^h)_i^* \nabla h_i(x^*) = 0. \end{array} \tag{4.8}$$

Since the MPCC-LICQ holds at x^*, by Proposition 2.1 we know that, to show that x^* is a B-stationary point of (1.1), we only need to show that (4.8) is the KKT condition for the relaxed problem (2.2). By comparing (4.8) with (2.3), we only need to show that

$$\Lambda_i^* := (\lambda^\phi)_i^* \xi_i^* \geq 0,\ \ \forall\, i \in \mathcal{I}^{GH} \quad \text{and} \quad \Gamma_i^* := (\lambda^\phi)_i^* \eta_i^* \geq 0,\ \ \forall\, i \in \mathcal{I}^{GH}. \tag{4.9}$$

We will derive a contradiction by assuming $\Lambda_{i_0}^* < 0$ for some $i_0 \in \mathcal{I}^{GH}$. For the case that $\Gamma_{i_0}^* < 0$ for some $i_0 \in \mathcal{I}^{GH}$ can be discussed similarly.

Since the MPCC-LICQ holds at x^*, it is not difficult to show that the vectors

$$\begin{array}{c} \{\nabla G_i(x^k) : i \in \mathcal{I}^G \cup \mathcal{I}^{GH}\} \cup \{\nabla H_i(x^k) : i \in \mathcal{I}^H \cup \mathcal{I}^{GH}\} \\ \cup \{\nabla g_i(x^k) : i \in \mathcal{I}^g\} \cup \{\nabla h_i(x^k) : i \in \mathcal{I}^q\} \end{array}$$

are linearly independent for all z^k sufficiently close to z^*. This implies that, for all z^k sufficiently close to z^*, there exists a sequence $\{d^k\}$ such that

$$\begin{array}{rcl} \nabla G_{i_0}(x^k)^T d^k &=& \frac{\partial \Phi_{i_0}(u^k)}{\partial H}, \quad i_0 \in \mathcal{I}^{GH}, \\ \nabla H_{i_0}(x^k)^T d^k &=& -\frac{\partial \Phi_{i_0}(u^k)}{\partial G}, \quad i_0 \in \mathcal{I}^{GH}, \\ \nabla G_i(x^k)^T d^k &=& 0, \quad \forall\, i \in \mathcal{I}^G \cup \mathcal{I}^{GH} \setminus \{i_0\}, \\ \nabla H_i(x^k)^T d^k &=& 0, \quad \forall\, i \in \mathcal{I}^H \cup \mathcal{I}^{GH} \setminus \{i_0\}, \\ \nabla g_i(x^k)^T d^k &=& 0, \quad \forall\, i \in \mathcal{I}^g, \\ \nabla h_i(x^k)^T d^k &=& 0, \quad \forall\, i \in \mathcal{I}^q. \end{array} \tag{4.10}$$

Now we consider the limit

$$\begin{array}{rcl} \lim_{k\to\infty} (d^k)^T \nabla_x^2 L(z^k) d^k &=& \lim_{k\to\infty} (d^k)^T \Bigg(\nabla^2 f(x^k) - \sum_{i\in\mathcal{I}^m} (\lambda^\phi)_i^k \nabla_x^2 \Phi_i(u^k) \\ && + \sum_{i\in\mathcal{I}^p} (\lambda^g)_i^k \nabla^2 g_i(x^k) + \sum_{i\in\mathcal{I}^q} (\lambda^h)_i^k \nabla^2 h_i(x^k) \Bigg) d^k. \end{array} \tag{4.11}$$

By using the fact that $z^k \to z^*$ as $k \to \infty$ (subsequencing if necessary), we can obtain the following results:

(a) For each $i \in \mathcal{I}^m$, sequences $\{\frac{\partial_x \Phi_i(u^k)}{\partial G}\}$ and $\{\frac{\partial_x \Phi_i(u^k)}{\partial H}\}$ are bounded.

(b) By Assumption 4.4 and the result (a), the sequence $\{d^k\}$ is bounded.

(c) Furthermore, it follows that sequences

$$\begin{aligned}
&\left\{(d^k)^T \bigtriangledown^2 f(x^k) d^k\right\},\\
&\left\{(\lambda^\phi)_i^k (d^k)^T \bigtriangledown_x^2 \Phi_i(u^k) d^k \ (i \in \mathcal{I}^m \setminus \{i_0\})\right\},\\
&\left\{(\lambda^g)_i^k (d^k)^T \bigtriangledown^2 g_i(x^k) d^k \ (i \in \mathcal{I}^p)\right\},\\
&\left\{(\lambda^h)_i^k (d^k)^T \bigtriangledown^2 h_i(x^k) d^k \ (i \in \mathcal{I}^q)\right\}
\end{aligned}$$

are bounded.

We consider the sequence $\left\{(\lambda^\phi)_{i_0}^k (d^k)^T \bigtriangledown_x^2 \Phi_{i_0}(u^k) d^k\right\}$. By using (4.10), we have

$$\begin{aligned}
&(d^k)^T \bigtriangledown_x^2 \Phi_{i_0}(u^k) d^k\\
&\quad = \frac{\partial_x \Phi_{i_0}(u^k)}{\partial G}(d^k)^T \bigtriangledown^2 G_{i_0}(x^k) d^k + \frac{\partial_x \Phi_{i_0}(u^k)}{\partial H}(d^k)^T \bigtriangledown^2 H_{i_0}(x^k) d^k - (d^k)^T \mathcal{R}^k d^k,
\end{aligned} \tag{4.12}$$

where

$$\begin{aligned}
\mathcal{R}^k \ := \ & -\frac{\partial_x^2 \Phi_{i_0}(u^k)}{\partial G^2}\left(\frac{\partial_x \Phi_{i_0}(u^k)}{\partial H}\right)^2 + 2\frac{\partial_x^2 \Phi_{i_0}(u^k)}{\partial G \partial H}\left(\frac{\partial_x \Phi_{i_0}(u^k)}{\partial G}\right)\left(\frac{\partial_x \Phi_{i_0}(u^k)}{\partial H}\right)\\
& -\frac{\partial_x^2 \Phi_{i_0}(u^k)}{\partial H^2}\left(\frac{\partial_x \Phi_{i_0}(u^k)}{\partial G}\right)^2.
\end{aligned}$$

For simplicity, we denote $G_{i_0 k} := G_{i_0}(x^k)$ and $H_{i_0 k} := H_{i_0}(x^k)$ for all iteration indices k. By using (4.10), we have that for all k,

$$\begin{aligned}
\mathcal{R}^k \ := \ & \frac{H_{i_0k}^2 + 4\mu_k^2}{\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}\right)^3}\left(\frac{\sqrt{G_{i_0k}^2 + H_{ik}^2 + 4\mu_k^2} - H_{i_0k}}{\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}}\right)^2\\
& + 2\frac{G_{i_0k}H_{i_0k}}{\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}\right)^3}\left(1 - \frac{G_{i_0k}}{\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}}\right)\left(1 - \frac{H_{i_0k}}{\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}}\right)\\
& + \frac{G_{i_0k}^2 + 4\mu_k^2}{\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}\right)^3}\left(\frac{\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} - G_{i_0k}}{\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}}\right)^2\\
= \ & \frac{(H_{i_0k}^2 + 4\mu_k^2)(G_{i_0k}^2 + 4\mu_k^2)^2}{\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}\right)^5\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + H_{i_0k}\right)^2}\\
& + \frac{2G_{i_0k}H_{i_0k}(H_{i_0k}^2 + 4\mu_k^2)(G_{i_0k}^2 + 4\mu_k^2)}{\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}\right)^5\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + G_{i_0k}\right)\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + H_{i_0k}\right)}\\
& + \frac{(G_{i_0k}^2 + 4\mu_k^2)(H_{i_0k}^2 + 4\mu_k^2)^2}{\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}\right)^5\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + G_{i_0k}\right)^2}.
\end{aligned} \tag{4.13}$$

Noticing that $G_{i_0k} \to 0$, $H_{i_0k} \to 0$, and $\mu_k \to 0$ as $k \to \infty$, we can show that sequences

$$\left\{ \frac{(H_{i_0k}^2 + 4\mu_k^2)(G_{i_0k}^2 + 4\mu_k^2)^2}{\left(G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2\right)^2 \left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + H_{i_0k}\right)^2} \right\},$$
$$\left\{ \frac{2G_{i_0k}H_{i_0k}(H_{i_0k}^2 + 4\mu_k^2)(G_{i_0k}^2 + 4\mu_k^2)}{\left(G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2\right)^2 \left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + G_{i_0k}\right)\left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + H_{i_0k}\right)} \right\},$$
$$\left\{ \frac{(G_{i_0k}^2 + 4\mu_k^2)(H_{i_0k}^2 + 4\mu_k^2)^2}{\left(G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2\right)^2 \left(\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2} + G_{i_0k}\right)^2} \right\}$$

are uniformly bounded away from zero. This, together with (4.13), implies that there exists a constant $\mathcal{C} > 0$ such that for all k,

$$\mathcal{R}^k \geq \frac{1}{\sqrt{G_{i_0k}^2 + H_{i_0k}^2 + 4\mu_k^2}}\mathcal{C},$$

which further implies that $\mathcal{R}^k \to +\infty$ as $k \to \infty$. Thus, by (4.12) it follows that

$$(d^k)^T \bigtriangledown_x^2 \Phi_{i_0}(u^k) d^k \to -\infty \quad \text{as } k \to \infty. \tag{4.14}$$

Since $\Lambda_{i_0} < 0$ and $\xi_{i_0}^* > 0$, it follows from (4.9) that $\lambda_{i_0}^\phi < 0$. Thus, by combining (4.11) with the result (c) and (4.14), we can further obtain that

$$(d^k)^T \bigtriangledown_x^2 L(z^k) d^k \to -\infty \quad \text{as } k \to \infty.$$

This contradicts Assumption 4.6, which indicates that (4.9) holds. Thus, (4.8) is the KKT condition for the relaxed problem (2.2). Furthermore, by Proposition 2.1 we obtain that x^* is a B-stationary point of (1.1).

The proof is complete. □

5 Numerical Results

In this section, we report some numerical experiments for Algorithm 4.1 implemented in Matlab. We tested the algorithm on some problems of small to medium size ($n \leq 140, m \leq 100$) taken from MacMPEC – a testing set of mathematical programs with equilibrium constraints (see, [15]). The numerical results are given in Table 1, where the names of problems are the same as those in MacMPEC. Throughout the computational experiments, the starting point was chosen as

$$z^0 := \left(\mu_0, x^0, (\lambda^\phi)^0, (\lambda^g)^0, (\lambda^h)^0\right) := (\mu_0, x^0, 0, \ldots, 0) \in \Re \times \Re^n \times \Re^{m+p+q},$$

where μ_0 was chosen as that in Table 1, and x^0 was chosen as same as that in MacMPEC; and the parameters used in the algorithm were chosen as

$$\sigma = 10^{-5},\ \delta = 0.5,\ \gamma = 0.2,\ \beta = 0.95 \min\{1.0, \mu_0/\|H(z^0)\|\}.$$

We used $\|H(z^k)\| \leq 10^{-6}$ as the stopping criterion. To improve the numerical implementation, the functions ϕ, θ, and ψ, defined by (2.4), (3.1), and (3.2), are replaced by

$$\begin{aligned}
\phi(\mu, a, b) &= a + b - \sqrt{a^2 + b^2 + 4(c_\phi \mu)^2}, \quad \forall\, (\mu, a, b) \in \Re^3, \\
\theta(\mu, a) &= \left(a + \sqrt{a^2 + 4(c_\theta \mu)^2}\right)/2, \quad \forall\, (\mu, a) \in \Re^2, \\
\psi(\mu, a, b) &= a + b - \sqrt{(a-b)^2 + 4(c_\psi \mu)^2}, \quad \forall\, (\mu, a, b) \in \Re^3,
\end{aligned}$$

respectively, where $c_\phi \geq 0, c_\theta \geq 0$, and $c_\psi \geq 0$ are three scalars. Obviously, such a modification does not destroy the convergence results in the last section. In our numerical implementation, we chose $c_\phi = c_\theta = c_\psi = 0.05$.

In Table 1, **Prob** denotes the problem to be tested; n denotes the dimension of the variable and m, p, and q denote the number of inequality constraints, equality constraints, and complementarity constraints, respectively. When the tested problems are reformulated as the form of the MPCC (1.1); c denotes the value of constant c given in (3.4); μ_0 denotes the value of μ_k when Algorithm 4.1 starts; **IT** denotes the number of iterations; **NF** denotes the number of function evaluations for the function H defined by (3.4); **ValH** denotes the value of $\|H(z^k)\|$ when the algorithm stops; and **Valf** denotes the value of the objective function $f(x)$ when the algorithm stops.

The quality of the solutions obtained in our experiment is comparable to that reported in the literature. For example, the obtained optimal values of problems dempe, ex9.1.1, ex9.1.3, ex9.1.7, and hakosen are the same as those in MacMPEC. For problems ex9.1.6, ex9.2.1, and ex9.2.7, the optimal values are the better than those in MacMPEC. For problem bilin the optimal value is better than that given by MacMPEC and is the same as the one in [19]. The optimal values of problem ex9.2.3, bard2, and bilevel1 are worse than the one in MacMPEC, but is the same as the one in [19].

Table 1 (Page 1): The numerical results of some problems in the MacMPEC

Prob	n	m	p	q	c	μ_0	IT	NF	ValH	Valf
bard2*	12	4	17	4	0.1	0.1	32	109	5.27e-7	6598
bard2m	12	8	17	0	0.5	0.1	185	809	8.71e-7	-6598
bard3	6	2	5	2	5	0.1	9	10	8.78e-9	-12.6787
bard3m	6	4	3	0	0.01	0.1	12	43	1.29e-8	-12.6787
bilevel1	10	6	5	2	0.01	0.1	28	56	9.73e-7	4.999999
bilevel2	20	12	9	4	0.01	0.1	38	113	2.34e-7	-6600
bilevel3	12	4	3	6	0.01	0.1	8	13	7.60e-9	-12.6787
bilin*	8	6	3	0	0.01	0.1	23	84	5.79e-7	14.6
dempe	3	1	0	1	0.01	0.1	8	9	3.05e-7	31.25
design-cent-1*	12	3	3	6	0.01	0.1	4	5	8.09e-7	1.860647
design-cent-2*	13	3	7	6	0.1	0.1	8	12	6.0e-9	3.483816
design-cent-4*	22	12	3	6	0.01	0.1	9	17	1.45e-7	3.079201
desilva	6	2	4	2	0.5	0.1	124	371	9.8e-7	-1.0
df1	2	1	4	0	0.01	5	0.1	8	3.94e-7	0.0
ex9.1.1	13	5	1	7	0.01	0.1	25	81	6.19e-8	-6.0
ex9.1.2	10	4	2	6	0.01	0.1	12	73	1.17e-7	-6.25
ex9.1.3	23	6	8	15	0.01	0.1	12	41	2.36e-7	-6.0
ex9.1.4	8	3	0	4	0.1	0.1	9	20	4.37e-8	-37.0
ex9.1.5	13	5	1	7	0.1	0.1	198	768	9.95e-7	-1.0
ex9.1.6	14	6	1	7	0.01	0.1	14	46	1.89e-10	-21.0
ex9.1.7	17	6	5	9	0.01	0.1	13	35	3.63e-7	-6.0
ex9.1.8	13	3	4	5	0.01	0.1	22	135	3.36e-7	-3.25
ex9.1.9	12	5	1	6	0.01	0.1	10	16	3.98e-8	3.11111
ex9.1.10	12	4	5	6	0.01	0.1	30	269	5.41e-8	-3.25
ex9.2.1	8	3	2	4	0.1	0.1	14	70	2.58e-8	17.313020
ex9.2.2	8	3	5	4	0.01	0.1	29	119	6.77e-7	100.0
ex9.2.3	16	6	7	8	0.1	0.1	43	168	6.30e-7	5.0
ex9.2.4	8	2	3	5	10	0.1	9	10	2.14e-7	0.499998
ex9.2.5	8	3	2	4	0.01	0.1	7	28	5.94e-8	9.0
ex9.2.6	12	4	4	6	0.01	0.1	69	220	9.19e-7	-1.0
ex9.2.7	10	4	2	5	0.01	0.1	13	76	4.03e-9	17.313020
ex9.2.9	9	3	4	5	1000	0.1	66	201	4.71e-7	1.999998
flp2	4	2	4	0	0.01	0.1	5	6	2.71e-9	0.0
flp4-1	80	30	30	0	0.01	0.1	6	7	2.14e-10	0.0
flp4-2	110	60	50	0	0.01	0.1	17	154	6.95e-7	0.0
flp4-3	140	70	100	0	0.01	0.1	43	416	3.66e-9	0.0
gauvin	3	2	2	0	0.01	0.1	5	6	4.53e-8	20.0
hakonsen*	7	4	4	1	0.5	0.1	15	32	2.53e-7	24.36682
jr1	2	1	0	0	0.01	0.1	7	9	1.54e-8	0.5
jr2	2	1	0	0	0.01	0.1	4	5	4.97e-8	0.5

*: The problem is to maximize the objective function

Table 1 (Page 2): The numerical results of some problems in the MacMPEC

Prob	n	m	p	q	c	μ_0	IT	NF	ValH	Valf
kth1	2	1	0	0	0.01	0.1	44	104	8.52e-7	1.37e-9
kth2	2	1	0	0	0.01	0.1	7	23	3.50e-7	-3.55e-9
kth3	2	1	0	0	0.01	0.1	7	12	4.36e-7	0.5
nash1	6	2	5	2	0.01	0.1	3	4	6.11e-8	1.30e-17
outrata31	5	4	2	0	2	0.1	194	871	9.53e-7	3.2077
outrata32	5	4	2	0	0.01	10	21	26	1.20e-7	3.449404
outrata33	5	4	2	0	0.1	10	19	21	5.76e-7	4.604253
outrata34	5	4	2	0	5	0.1	13	23	1.76e-9	6.592684
portf1-i-1	87	12	62	13	0.01	0.1	9	10	9.95e-8	1.502421e-5
portf1-i-2	87	12	62	13	1	0.1	7	8	1.88e-7	1.457245e-5
portf1-i-3	87	12	62	13	1	0.1	8	9	8.24e-10	6.264974e-6
portf1-i-4	87	12	62	13	1	0.1	7	8	1.55e-8	2.177332e-6
portf1-i-6	87	12	62	13	1	0.1	7	8	2.08e-8	2.361318e-6
qpec-100-1	105	100	2	0	0.1	20	31	40	3.24e-7	0.1103853
qpec1	30	20	0	0	1	0.1	108	314	9.77e-7	80.0
ralph1	2	1	1	0	0.1	0.1	76	217	8.88e-7	-1.47e-9
ralph2	2	1	0	0	0.01	0.1	4	5	4.28e-8	-1.83e-17
scholtes1	3	1	1	0	100	0.1	14	20	9.96e-7	1.999812
scholtes2	3	1	1	0	0.01	0.1	59	224	9.88e-7	15.0
scholtes3	2	1	0	0	0.01	0.1	18	23	8.12e-7	0.5
scholtes4	3	1	2	0	1	0.1	92	507	9.99e-7	-1.56e-7
scholtes5	3	2	0	0	0.01	0.1	6	7	1.66e-7	1.0
sl1	8	3	6	2	0.01	0.1	15	66	3.85e-10	1.0e-4
stackelberg1	3	1	4	1	0.01	0.1	10	21	1.05e-7	-3266.667

Acknowledgment. We are very grateful to the two referees for their valuable comments on the paper, which have considerably improved the paper.

References

[1] H. Y. Benson, D. F. Shanno, and R. J. Vanderbei, Interior-point methods for nonconvex nonlinear programming: complementarity constraints, *Preprint*, Operations Research and Financial Engineering, Princeton University, ORFE-02-02, July, 2002.

[2] X. Chen and M. Fukushima, A smoothing method for a mathematical program with P-matrix linear complementarity constraints, *Comput. Optim. Appl.*, 27 (2004), 223–246.

[3] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

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

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

[6] 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, Lecture Notes in Economics and Mathematical Systems*, Vol. 477, M. Théra and R. Tichatschke, eds., Springer-Verlag, Berlin/Heidelberg, 1999, pp. 105–116.

[7] M. Fukushima and P. Tseng, An implementable active-set algorithm for computing a B-stationary point of the mathematical program with linear complementarity constraints, *SIAM J Optim.*, 12 (2002), 24–739.

[8] M. S. Gowda and J. J. Sznajder, Weak univalence and connectedness of inverse images of continuous functions, *Math. Oper. Res.*, 24 (1999), 255–261.

[9] X. Hu and D. Ralph, Convergence of a penalty method for mathematical programming with equilibrium constraints, *Manuscript*, Department of Mathematics and Statistics, University of Melbourne, Australia, 2001.

[10] X. X. Huang, X. Q. Yang, and D. L. Zhu, A sequential smooth penalization approach to mathematical programs with complementarity constraints, *Manuscript*, Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, 2001.

[11] Z. H. Huang, Sufficient conditions on nonemptiness and boundedness of the solution set of the P_0 function nonlinear complementarity problem, *Oper. Res. Letters*, 30 (2002), 202–210.

[12] Z. H. Huang, J. Han, and Z. Chen, A predictor-corrector smoothing Newton algorithm, based on a new smoothing function, for solving the nonlinear complementarity problem with a P_0 function, *J. Optim. Theory Appl.*, 117 (2003), 39–68.

[13] Z. H. Huang, L. Qi, and D. Sun, Sub-quadratic convergence of a smoothing Newton algorithm for the P_0- and monotone LCP, *Math. Program.*, 99 (2004), 423–441.

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

[15] S. Leyffer, *MacMPEC – www-unix.mcs.anl.gov/leyffer/MacMPEC/.2002.*

[16] G. H. Lin and M. Fukushima, Some exact penalty result for nonlinear programs and mathematical programs with equilibrium constraints, *J. Optim. Theory Appl.*, 118 (2003), 67–80.

[17] G. H. Lin and M. Fukushima, New relaxation method for mathematical programs with complementarity constraints, *J. Optim. Theory Appl.*, 118 (2003), 81–116.

[18] X. Liu, G. Perakis, and J. Sun, A robust SQP method for mathematical programs with linear complementarity constraints, *Manuscript,* Singapore-MIT Alliance, National University of Singapore, 2003.

[19] X. Liu and J. Sun, Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints, *Math. Program.,* 101 (2004), 231-261.

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

[21] H. D. Qi, A regularized smoothing Newton method for box constrained variational inequality problems with P_0-functions, *SIAM J. Optim.*, 10 (2000), 315–330.

[22] G. Ravindran and M. S. Gowda, Regularization of P_0-functions in box variational inequality problems, *SIAM J. Optim.*, 11 (2001), 748–760.

[23] S. Scholtes, Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM J. Optim.*, 10 (2000), 918–936.

[24] S. Scholtes and M. Stöhr, Exact penalization of mathematical programs with complementarity constraints, *SIAM J. Control Optim.*, 37 (1999), 617–652.

[25] Y. B. Zhao and D. Li, Existence and limiting behavior of a non-interior-point trajectory for nonlinear complementarity problems without strict feasibility condition, *SIAM J. Control Optim.*, 40 (2001), 898–924.