A robust SQP method for mathematical programs with linear complementarity constraints *

Xinwei Liu[†], Georgia Perakis[‡], Jie Sun[§]

Abstract. The relationship between the mathematical program with linear complementarity constraints (MPCC) and its inequality relaxation is studied. A new sequential quadratic programming (SQP) method is presented for solving the MPCC based on this relationship. A certain SQP technique is introduced to deal with the possible infeasibility of quadratic programming subproblems. Global convergence results are derived without assuming the linear independence constraint qualification for MPEC and nondegeneracy of the complementarity constraints. Preliminary numerical results are reported.

Key words: mathematical programs with equilibrium constraints, sequential quadratic programming, complementarity, constraint qualification, nondegeneracy

*Research is partially supported by Singapore-MIT Alliance and School of Business, National University of Singapore.

†HPCES, Singapore-MIT Alliance, National University of Singapore. Email: smaliuxw@nus.edu.sg

‡Sloan School of Management, Massachusetts Institute of Technology, USA. Email: georgiap@mit.edu

§Department of Decision Sciences and Singapore-MIT Alliance, National University of Singapore. Fax: (65)6779-2621. Email: jsun@nus.edu.sg

1. Introduction

The mathematical program with equilibrium constraints (MPEC) has extensive applications in areas such as engineering design and economic modelling. It has been an active research topic in recent years. In this paper, we consider the following mathematical program with linear complementarity constraints (MPCC), which is a special case of the MPEC:

$$\min\ f(x,y) \tag{1.1}$$
$$\text{s.t. } Cx + Dy \leq c, \tag{1.2}$$
$$Ax + By = b, \tag{1.3}$$
$$Nx + My - w = q, \tag{1.4}$$
$$0 \leq w \perp y \geq 0, \tag{1.5}$$

where $f : \Re^{n+m} \to \Re$ is a twice continuously differentiable real-valued function, $C \in \Re^{\ell\times n}$, $D \in \Re^{\ell\times m}$, $A \in \Re^{p\times n}$, $B \in \Re^{p\times m}$, $N \in \Re^{m\times n}$, $M \in \Re^{m\times m}$are given matrices, $c,\ b,\ q$ are given $\ell,\ p,\ m$-dimensional vectors respectively.

Research work on the MPEC includes the monograph [19] that provides a comprehensive study on the MPEC, such as the exact penalization theory, optimality conditions, and some iterative algorithms. Based on different formulations including the piecewise smooth formulation and certain regularization scheme, in a series of papers [22, 23, 24] Scholtes et al. made an extensive study on stationarity and optimality of the MPEC, and presented some algorithms. Fukushima, Luo and Pang [8] presented a sequential quadratic programming (SQP) algorithm for the MPCC that is based on reformulating the complementarity constraints into a system of semismooth equations by using the Fischer-Burmeister function. The feasibility issues on the subproblems in SQP for MPEC were studied by Fukushima and Pang [9]. Ralph [21], Jiang and Ralph [12, 13] presented some SQP methods for MPCC and reported numercial results on implementations of their algorithms. Very recently, Fukushima and Tseng [11] proposed an active-set algorithm for MPCC.

It is noted that the global convergence results of all aforementioned algorithms are based on certain assumptions on the matrices in the problem and/or the nondegeneracy (strict complementarity) of the complementarity constraints. For example, Fukushima, Luo and Pang [8] assumed that M is a P_0-matrix with one of its principle submatrix being nondegenerate and that strict complementarity holds for the complementarity constraints. Fukushima and Tseng [11] assumed that a linearly independent constraint qualification for MPEC (MPEC-LICQ for short) holds. Similar assumptions are also needed in Jiang and Ralph [13].

There have been some algorithms with locally rapid convergence, e.g., [1, 5, 7, 16, 20]. Numerical results showed that they have good performances on some test problems such as MacMPEC (see [15]). However, it remains open to demonstrate their global convergences.

The following example indicates that the MPEC-LICQ and nondegeneracy may not hold for simple problems. Consider

$$\min\ x + y \tag{1.6}$$
$$\text{s.t. } -1 \le x \le 1, \tag{1.7}$$
$$1 + x - w = 0, \tag{1.8}$$
$$0 \le w \perp y \ge 0. \tag{1.9}$$

The optimal point is $(-1, 0, 0)$, at which both the MPEC-LICQ and nondegeneracy do not hold. Thus, it would make sense to develop a globally convergent algorithm that does not require the problem to satisfy the MPEC-LICQ or the nondegenracy assumption. Our target in this paper is to develop a globally convergent SQP method that does not require these assumptions. The algorithm is based on an inequality relaxation of the complementarity constraints. Some interesing features of the algorithm are as follows.

1. The feasibility of the problem is not required in advance, and the algorithm starts from an infeasible point which satisfies some linear constraints.

2. All SQP subproblems specified by the algorithm are feasible. Thus, a search direction always exists before termination.

3. Convergence results are proved without assuming the MPEC-LICQ or the nondegeneracy of complementarity constraints. The algorithm may find some point with certain strong or weak stationary properties.

The paper is organized as follows. In the next section, we describe a relaxation of MPCC and give some related results. We introduce a decomposed SQP technique and present a new SQP algorithm for MPCC in Section 3. We discuss stationary properties associated with the algorithm in Section 4 and prove the global convergence results of the algorithm in Section 5. Preliminary numerical results are reported in the last section.

For reader's convenience, we list some notations used in the paper. For any vectors $u \in \Re^s$ and $v \in \Re^t$, we have that $(u, v) = [u^\top\ v^\top]^\top \in \Re^{s+t}$, where $\top$ is the transpose. Suppose that $R \in \Re^{s\times t}$ is any $s \times t$ matrix, $\mathcal{S}$ is any subset of the indices $\{1, 2, \dots, t\}$, then $R_{\mathcal{S}}$ represents a submatrix of R consisting of its columns indexed by $\mathcal{S}$. A vector with supscript k and a matrix with subscript k corresponds to the iterate k, whereas a vector with subscript i represents its i-th component. The letter I stands for the identity matrix, whose size may be identified in the context. At last, "$\circ$" is the Hardamat product of vectors, that is, for $u \in \Re^s$ and $w \in \Re^s$, $u \circ w$ is a vector in $\Re^s$ with $(u \circ w)_i = u_i w_i$.

2. An inequality relaxation of MPCC

Let $\bar{z} = (\bar{x}, \bar{y}, \bar{w})$ be a feasible point of MPCC (1.1)-(1.5). A well-known relaxation of problem (1.1)-(1.5) associated with $\bar{z}$ is defined by

$$\min\ f(x, y) \tag{2.1}$$
$$\text{s.t. } Cx + Dy \le c, \tag{2.2}$$
$$Ax + By = b, \tag{2.3}$$

$$
\begin{aligned}
& Nx + My - w = q, && (2.4)\\
& y_j \geq 0,\ j \in \{j : \bar{w}_j = 0\}, && (2.5)\\
& y_j = 0,\ j \in \{j : \bar{w}_j \neq 0\}, && (2.6)\\
& w_j \geq 0,\ j \in \{j : \bar{y}_j = 0\}, && (2.7)\\
& w_j = 0,\ j \in \{j : \bar{y}_j \neq 0\}. && (2.8)
\end{aligned}
$$

We denote problem (2.1)-(2.8) by $\mathcal{R}(\bar{z})$. This kind of relaxed problems have played an important role in the development of theories and algorithms for MPECs, see [11, 19]. The following theorem is a crucial result (Theorem 2.1 of [11]).

Theorem 2.1 *Let $\bar{z}$ be a feasible solution of MPCC (1.1)-(1.5) such that the MPEC-LICQ holds at $\bar{z}$. Then $\bar{z}$ is a KKT point of the relaxed problem $\mathcal{R}(\bar{z})$ if and only if $\bar{z}$ is a Bouligand stationary point of the MPCC.*

In this paper we consider another relaxation, denoted by $\mathcal{N}(\tau)$, of the problem (1.1)-(1.5):

$$
\begin{aligned}
& \min\ f(x,y) && (2.9)\\
& \text{s.t. } Cx + Dy \leq c, && (2.10)\\
& \quad Ax + By = b, && (2.11)\\
& \quad Nx + My - w = q, && (2.12)\\
& \quad y \geq 0, && (2.13)\\
& \quad w \geq 0, && (2.14)\\
& \quad y \circ w \leq \tau e, && (2.15)
\end{aligned}
$$

where $\tau \geq 0$ is a scalar, $e = (1 \ \ldots \ 1)^\top$. The complementarity constraints in (1.1)-(1.5) are relaxed by inequalities (2.15). If $\tau = 0$, then problem (2.9)-(2.15) is equivalent to the MPCC. Here and below, to simplify the notation, we denote the primal variable by z with $z = (x, y, w)$ and denote the dual variables by $u = (\lambda, \mu, \nu, \xi, \zeta, \eta)$ where $\lambda, \mu, \nu, \xi, \zeta, \eta$ are the dual multiplier vector associated with the constraints (2.10)-(2.15), respectively.

This kind of relaxation has been proposed in [22]. Some newly developed locally convergent algorithms (e.g., see [5, 16, 20]) also used this relaxation. Suppose $\tau > 0$ and (x_τ, y_τ, w_τ) is a feasible point of $\mathcal{N}(\tau)$.

Since the proposed algorithm aims to find a point with certain stationary properties (discussed later) of problem (1.1)-(1.5) by relaxation problem $\mathcal{N}(\tau)$, we must study how the KKT point of $\mathcal{N}(\tau)$ is related to the original problem, which is the purpose of this section. The way we do it is to relate problem $\mathcal{N}(\tau)$ to $\mathcal{R}(\bar{z})$ and to apply Theorem 2.1.

It is well known that a KKT pair of $\mathcal{N}(\tau)$, denoted by (z_τ, u_τ), must satisfy

$$
\begin{aligned}
& \nabla_x f_\tau + C^\top \lambda_\tau + A^\top \mu_\tau + N^\top \nu_\tau = 0, && (2.16)\\
& \nabla_y f_\tau + D^\top \lambda_\tau + B^\top \mu_\tau + M^\top \nu_\tau - \xi_\tau + W_\tau \eta_\tau = 0, && (2.17)\\
& -\nu_\tau - \zeta_\tau + Y_\tau \eta_\tau = 0, && (2.18)\\
& {\lambda_\tau}^\top (Cx_\tau + Dy_\tau - c) = 0,\ {\xi_\tau}^\top y_\tau = 0,\ {\zeta_\tau}^\top w_\tau = 0,\ {\eta_\tau}^\top (y_\tau \circ w_\tau - \tau e) = 0. && (2.19)
\end{aligned}
$$

where $z_\tau = (x_\tau, y_\tau, w_\tau)$, $u_\tau = (\lambda_\tau, \mu_\tau, \nu_\tau, \xi_\tau, \zeta_\tau, \eta_\tau)$. Then we have the following result.

Theorem 2.2 *Suppose that $\{(z^k, u^k)\}$ is an infinite sequence, where z^k is a feasible point of $\mathcal{N}(\tau_k)$, and (z^k, u^k) satisfies the KKT conditions (2.16)-(2.19) of $\mathcal{N}(\tau_k)$. If $\{u^k\}$ is bounded, and $\tau_k \to 0$ as $k \to \infty$, then any limit point z^* of $\{z^k\}$ is a KKT point of the relaxed problem $\mathcal{R}(z^*)$.*

Proof. Since $\{u^k\}$ is bounded, there exists a subsequence $\{u^k : k \in \mathcal{K}\}$ and a vector u^* such that $u^k \to u^*$ for $k \in \mathcal{K}$ and $k \to \infty$. The limit point z^* is a feasible point of problem (1.1)-(1.5) because z^k is a feasible point of $\mathcal{N}(\tau_k)$. Let

$$\mathcal{I}_1 = \{i : y_i^* > w_i^* = 0\},\ \mathcal{I}_2 = \{i : y_i^* = w_i^* = 0\},\ \mathcal{I}_3 = \{i : 0 = y_i^* < w_i^*\}. \tag{2.20}$$

It follows from (2.16)-(2.19) that

$$\begin{aligned}
&\nabla_x f^* + C^\top \lambda^* + A^\top \mu^* + N^\top \nu^* = 0, && (2.21)\\
&\nabla_y f^* + D^\top \lambda^* + B^\top \mu^* + M^\top \nu^* = \pi^*, && (2.22)\\
&\lambda_i^*(C_i x^* + D_i y^* - c_i) = 0,\ i = 1, \ldots, \ell, && (2.23)\\
&\pi_i^* = 0, \qquad \nu_i^* = -\zeta_i^* + y_i^* \eta_i^*,\ i \in \mathcal{I}_1, && (2.24)\\
&\pi_i^* = \xi_i^*, \qquad \nu_i^* = -\zeta_i^*,\ i \in \mathcal{I}_2, && (2.25)\\
&\pi_i^* = \xi_i^* - w_i^* \eta_i^*, \qquad \nu_i^* = 0,\ i \in \mathcal{I}_3, && (2.26)
\end{aligned}$$

where $\xi_i^* \geq 0$, $\zeta_i^* \geq 0$, $\eta_i^* \geq 0$, and $\lambda_i^* \geq 0$.

A point $z^* = (x^*, y^*, w^*)$ is a KKT point of the relaxed problem $\mathcal{R}(z^*)$ if the point is feasible to the problem and there exist $\lambda^* \in \Re_+^\ell$, $\pi^* \in \Re^m$, $\mu^* \in \Re^p$ and $\nu^* \in \Re^m$ such that

$$\begin{aligned}
&\nabla_x f^* + C^\top \lambda^* + A^\top \mu^* + N^\top \nu^* = 0, && (2.27)\\
&\nabla_y f^* + D^\top \lambda^* + B^\top \mu^* + M^\top \nu^* = \pi^*, && (2.28)\\
&\lambda_i^*(C_i x^* + D_i y^* - c) = 0,\ i = 1, \ldots, \ell, && (2.29)\\
&\pi_i^* \geq 0,\ \pi_i^* y_i^* = 0,\ i \in \{i : w_i^* = 0\}, && (2.30)\\
&\nu_i^* \leq 0,\ \nu_i^* w_i^* = 0,\ i \in \{i : y_i^* = 0\}. && (2.31)
\end{aligned}$$

By comparing equations (2.21)-(2.26) with equations (2.27)-(2.31), we can see that z^* is also a KKT point of problem $\mathcal{R}(z^*)$. ∎

We now formally define the MPEC-LICQ for problem (1.1)-(1.5).

Definition 2.3 *The MPEC-LICQ holds at point z^* for problem (1.1)-(1.5) if the coefficient matrix*

$$\begin{pmatrix} (C^\top)_{\mathcal{S}_0^*} & A^\top & N^\top & & \\ (D^\top)_{\mathcal{S}_0^*} & B^\top & M^\top & I_{\mathcal{S}_y^*} & \\ & & I & & I_{\mathcal{S}_w^*} \end{pmatrix} \tag{2.32}$$

has full column rank, where $\mathcal{S}_0^* = \{i : C_i x^* + D_i y^* = c_i\}$, $\mathcal{S}_y^* = \{i : y_i^* = 0\}$ and $\mathcal{S}_w^* = \{i : w_i^* = 0\}$.

The nondegeneracy condition of problem (1.1)-(1.5) holds if and only if $y \circ w = 0$ and $y + w > 0$. Then we have the following result on the boundedness of $\{u^k\}$.

Lemma 2.4 *Suppose that $\{(z^k, u^k)\}$ is an infinite sequence, where z^k is a feasible point of $\mathcal{N}(\tau_k)$, and (z^k, u^k) satisfies the KKT conditions (2.16)-(2.19) of problem $\mathcal{N}(\tau_k)$. If $\tau_k \to 0$ and $z^k \to z^*$ as $k \to \infty$, both the MPEC-LICQ and the nondegeneracy hold at z^*, then $\{u^k\}$ is bounded.*

Proof. We prove it by the contrary. Assume $\|u^k\|_\infty \to \infty$ for $k \in \mathcal{K}$, where $\mathcal{K}$ is an infinite index set. By dividing $\|u^k\|_\infty$ on both sides of equations (2.16)-(2.18) and taking limit as $k \to \infty$, combining with the nondegeneracy at z^*, we derive the system of equations which shows that the column vectors of matrix (2.32) are linearly dependent, thus contradicts the MPEC-LICQ. The contradiction indicates that the result holds. ∎

The asymptotically weak nondegeneracy has been introduced by Fukushima and Pang [10] to substitute for the nondegeneracy condition in the convergence analysis of a smoothing continuation method for MPCC, which is described as follows.

Definition 2.5 *The infinite sequence $\{z^k\}$ has asymptotically weak nondegeneracy at z^* if there exist an infinite subsequence $\{z^k : k \in \mathcal{K}\}$ and two constants $\beta_1 > \beta_2 > 0$ such that $z^k \to z^*$ as $k \in \mathcal{K}$ and $k \to \infty$, and for sufficiently large $k \in \mathcal{K}$ and for any $i \in \{i : y_i^* = w_i^* = 0\} \bigcap \{i : y_i^k w_i^k = \tau_k\}$, there holds*

$$\beta_2 \leq y_i^k / w_i^k \leq \beta_1. \tag{2.33}$$

The condition on asymptotically weak nondegeneracy is weaker than the nondegeneracy (strictly complementarity) condition. If the nondegeneracy holds at z^*, then $\{i : y_i^* = w_i^* = 0\} = \emptyset$. Thus the condition on asymptotically weak nondegeneracy holds naturally.

Lemma 2.6 *Suppose that $\{(z^k, u^k)\}$ is an infinite sequence, where z^k is a feasible point of $\mathcal{N}(\tau_k)$, and (z^k, u^k) satisfies the KKT conditions (2.16)-(2.19) of problem $\mathcal{N}(\tau_k)$. Furthermore, we assume that $d^\top \nabla_z^2 L(z^k, u^k) d \geq 0$ for all k, where $\nabla_z^2 L(z^k, u^k)$ is the Hessian of Lagrangian of problem $\mathcal{N}(\tau_k)$ at point (z^k, u^k) and d is in the set*

$$\mathcal{S} = \left\{ d \in \Re^{n+2m} : \begin{array}{rl} C_i d_x + D_i d_y = 0, & i \in \mathcal{S}_0^k = \{i : C_i x^k + D_i y^k = c_i\}, \\ A d_x + B d_y = 0, & \\ N d_x + M d_y - d_w = 0, & \\ (d_y)_j = 0, & j \in \mathcal{S}_y^k = \{j : y_j^k = 0,\ w_j^k > 0\}, \\ (d_w)_j = 0, & j \in \mathcal{S}_w^k = \{j : w_j^k = 0,\ y_j^k > 0\}, \\ y_j^k (d_w)_j + w_j^k (d_y)_j = 0, & j \in \mathcal{S}_c^k = \{j : y_j^k w_j^k = \tau_k\}. \end{array} \right\}.$$

If $\tau_k \to 0$ and $z^k \to z^*$ as $k \to \infty$, both the MPEC-LICQ and the asymptotically weak nondegeneracy hold at z^*, then $\{u^k\}$ is bounded.

Proof. If $\{i : y_i^k w_i^k = \tau_k\} = \emptyset$, then $\eta^k = 0$. We suppose that $\{i : y_i^k w_i^k = \tau_k\} \neq \emptyset$ for all $k \geq 0$. Two cases need to be considered: (i) $\{i : y_i^* = w_i^* = 0\} \cap \{i : y_i^k w_i^k = \tau_k\} = \emptyset$ for all sufficiently large k; (ii) for sufficiently large k, $\{i : y_i^* = w_i^* = 0\} \cap \{i : y_i^k w_i^k = \tau_k\} \neq \emptyset$.

Case (i). In this case, for all sufficiently large k, if $i_0 \in \{i : y_i^k w_i^k = \tau_k\}$, then $y_{i_0}^* w_{i_0}^* = 0$ and $y_{i_0}^* + w_{i_0}^* > 0$. Thus, the result follows from the proof of Lemma 2.4.

Case (ii). We first prove that $\{\eta^k\}$ is bounded. If $\|\eta^k\|_\infty \to \infty$ as $k \in \mathcal{K}$ and $k \to \infty$, by the MPEC-LICQ, there exists $i_0 \in \{i : y_i^* = w_i^* = 0\} \cap \{i : y_i^k w_i^k = \tau_k\}$ such that $\eta_{i_0}^k \to \infty$ as $k \in \mathcal{K}$ and $k \to \infty$.

We select $d^k \in \mathcal{S}$ as follows: $(d_w^k)_i = 1$ and $(d_y^k)_i = -(y_i^k / w_i^k)$ for $i \in \{i : y_i^* = w_i^* = 0\} \cap \{i : y_i^k w_i^k = \tau_k\}$; $(d_y^k)_i = 0,\ i \in \{i : y_i^k = 0\}$; $(d_w^k)_i = 0,\ i \in \{i : w_i^k = 0\}$; $(d_y^k)_i = 0$ and $(d_w^k)_i = 0$ for $\{i : y_i^k w_i^k = \tau_k\} \backslash \{i : y_i^* = w_i^* = 0\}$. For sufficiently large k, this selection is guaranteed by the MPEC-LICQ. It is noted that $\|d_w^k\| \not\to 0$ as $k \to \infty$. Since

$$d^\top \nabla_z^2 L(z^k, u^k) d = d_x^\top \nabla_x^2 f_k d_x + 2 d_x^\top \nabla_{xy}^2 f_k d_y + d_y^\top \nabla_y^2 f_k d_y + 2 d_y^\top \mathrm{diag}(\eta^k) d_w, \tag{2.34}$$

by the asymptotically weak nondegeneracy, if $\|\eta^k\|_\infty \to \infty$ as $k \in \mathcal{K}$ and $k \to \infty$, then ${d^k}^\top \nabla_z^2 L(z^k, u^k) d^k \to -\infty$ as $k \in \mathcal{K}$ and $k \to \infty$, which contradicts that $d^\top \nabla_z^2 L(z^k, u^k) d \geq 0\ \forall d \in \mathcal{S}, \forall k$. This contradiction implies that $\{\eta^k\}$ is bounded. By equations (2.16)-(2.19) and the MPEC-LICQ, we have the desired result. ∎

The following result can be derived directly from Theorem 2.1, Theorem 2.2, Lemma 2.4 and Lemma 2.6.

Theorem 2.7 *Suppose that $\{(z^k, u^k)\}$ is an infinite sequence, where z^k is a feasible point of problem $\mathcal{N}(\tau_k)$, and (z^k, u^k) satisfies equations (2.16)-(2.19). Let $\tau_k \to 0$ as $k \to \infty$, z^* is any limit point of $\{z^k\}$. If either the MPEC-LICQ and the nondegeneracy, or the MPEC-LICQ, the asymptotically weak nondegeneracy hold at z^* and $d^\top \nabla_z^2 L(z^k, u^k) d \geq 0$ $\forall k \geq 0$ and $\forall d \in \mathcal{S}$, then z^* is a Bouligand stationary point of MPCC (1.1)-(1.5).*

Moreover, we have the following result under the MPEC-LICQ and the nondegeneracy.

Corollary 2.8 *Suppose that $\{(z^k, u^k)\}$ is an infinite sequence, where z^k is a feasible point of problem $\mathcal{N}(\tau_k)$, and (z^k, u^k) satisfies equations (2.16)-(2.19). Let $\tau_k \to 0$ as $k \to \infty$, z^* is any limit point of $\{z^k\}$. If both the MPEC-LICQ and the nondegeneracy hold at z^*, then there exist multipliers $(\lambda^*, \mu^*, \nu^*, \bar{\eta})$ such that $(z^*, \lambda^*, \mu^*, \nu^*, \bar{\eta})$ satisfies the following stationary conditions*

$$\nabla_x f(x^*, y^*) + C^\top \lambda^* + A^\top \mu^* + N^\top \nu^* = 0, \tag{2.35}$$

$$\nabla_y f(x^*, y^*) + D^\top \lambda^* + B^\top \mu^* + M^\top \nu^* + W^* \bar{\eta} = 0, \tag{2.36}$$

$$-\nu^* + Y^* \bar{\eta} = 0, \tag{2.37}$$

$$\lambda^* \geq 0,\ (\lambda^*)^\top (C x^* + D y^* - c) = 0. \tag{2.38}$$

Proof. By Lemma 2.4, for any z^*, there exists a $u^* = (\lambda^*, \mu^*, \nu^*, \xi^*, \zeta^*, \eta^*)$ which is a limit point of $\{u^k\}$ such that

$$\nabla_x f(x^*, y^*) + C^\top \lambda^* + A^\top \mu^* + N^\top \nu^* = 0, \tag{2.39}$$

$$\nabla_y f(x^*, y^*) + D^\top \lambda^* + B^\top \mu^* + M^\top \nu^* - \xi^* + W^* \eta^* = 0, \tag{2.40}$$

$$-\nu^* - \zeta^* + Y^* \eta^* = 0, \tag{2.41}$$

$$\lambda^* \geq 0,\ (\lambda^*)^\top (Cx^* + Dy^* - c) = 0. \tag{2.42}$$

Under the nondegeneracy condition, we can select $\bar{\xi}$, $\bar{\zeta}$ such that $\xi^* = (W^*)^2 \bar{\xi}$, $\zeta^* = (Y^*)^2 \bar{\zeta}$ as follows. If $\xi_i^* = 0$, let $\bar{\xi}_i^* = 0$, otheriwse if $\xi_i^* \neq 0$, we have $y_i^* = 0$ by (2.19), which implies that $w_i^* \neq 0$ by the nondegeneracy, thus $\bar{\xi}_i = (w_i^*)^{-2} \xi_i^*$, it is similar for the selection of $\bar{\zeta}$. Let $\bar{\eta} = \eta^* - W^* \bar{\xi} - Y^* \bar{\zeta}$, then, by (2.39)-(2.42), we have (2.35)-(2.38). ■

Under assumptions of MPEC-LICQ and nondegeneracy, Fukushima, Luo and Pang [8] proposed an algorithm converging to the point satisfying equations (2.35)-(2.38).

3. The algorithm

The SQP methods for MPECs presented in the literature such as [8, 12, 13, 21] replace the complementarity constraints with some smoothing equations, and then the MPECs are approximated by a new nonlinear programming with some constraint functions being asymptotically nonsmooth. Our algorithm is based on the above inequality relaxation (2.9)-(2.15). Some techniques originated from SQP for nonlinear programming, e.g., [3, 4, 17, 18, 25], are introduced to circumvent the possible inconsistency of subproblems.

Assumption 3.1

$$\begin{aligned} \mathcal{G} \equiv \{(x, y, w) \in \Re^n \times \Re^m_+ \times \Re^m_+ : Cx + Dy &\leq c, \\ Ax + By &= b, \\ Nx + My - w &= q\} \neq \emptyset. \end{aligned}$$

If $\mathcal{G} = \emptyset$, then the MPCC has no solution. However, the nonempty of $\mathcal{G}$ does not necessarily imply the existence of the solution of problem (1.1)-(1.5) since the complementarity constraints may not hold.

Suppose that $\mathcal{G} \neq \emptyset$. We do not require that the relaxation problem $\mathcal{N}(\tau)$ is feasible. For any $z \equiv (x, y, w) \in \mathcal{G}$, applying the SQP approach to the problem $\mathcal{N}(\tau)$, we derive the QP subproblem:

$$\min\ \psi(d) = \nabla f(x, y)^\top \begin{pmatrix} d_x \\ d_y \end{pmatrix} + \frac{1}{2}(d_x^\top\ d_y^\top\ d_w^\top) H \begin{pmatrix} d_x \\ d_y \\ d_w \end{pmatrix} \tag{3.1}$$

$$\text{s.t. } Cd_x + Dd_y \leq -(Cx + Dy - c), \tag{3.2}$$

$$Ad_x + Bd_y = 0, \tag{3.3}$$
$$Nd_x + Md_y - d_w = 0, \tag{3.4}$$
$$y + d_y \geq 0,\ w + d_w \geq 0, \tag{3.5}$$
$$Wd_y + Yd_w \leq -(WYe - \tau e), \tag{3.6}$$

where H is an approximate Lagrangian Hessian at z and is supposed to be positive definite. It is noted that z is a feasible point of problem (2.9)-(2.15) if and only if $d \equiv (d_x, d_y, d_w) = 0$ is a feasible point of the QP subproblem (3.1)-(3.6).

Problem (3.1)-(3.6) may have no feasible solution if z is not a feasible point for problem (2.9)-(2.15). In order to avoid this bad case, we introduce some decomposition technique in [3, 4, 17, 18, 25]. We firstly solve the problem $\mathcal{A}(z, \tau)$ as follows.

$$\min \|(Wd_y + Yd_w + WYe - \tau e)_+\|_1 \tag{3.7}$$
$$\text{s.t. } Cd_x + Dd_y \leq -(Cx + Dy - c), \tag{3.8}$$
$$Ad_x + Bd_y = 0, \tag{3.9}$$
$$Nd_x + Md_y - d_w = 0, \tag{3.10}$$
$$y + d_y \geq 0,\ w + d_w \geq 0, \tag{3.11}$$

where $\|\cdot\|_1$ is the so-called ℓ_1 norm. By introducing additional variables $v \in \Re^m_+$, this problem can be equivalently transformed to the following linear program $\mathcal{B}(z, \tau)$:

$$\min e^\top v \tag{3.12}$$
$$\text{s.t. } Cd_x + Dd_y \leq -(Cx + Dy - c), \tag{3.13}$$
$$Ad_x + Bd_y = 0, \tag{3.14}$$
$$Nd_x + Md_y - d_w = 0, \tag{3.15}$$
$$y + d_y \geq 0,\ w + d_w \geq 0, \tag{3.16}$$
$$Wd_y + Yd_w + (WYe - \tau e) - v \leq 0,\ v \geq 0. \tag{3.17}$$

The problem (3.12)-(3.17) is always feasible since $z \in \mathcal{G}$, and $d = 0$ together with $v = (WYe - \tau e)_+$ is its feasible solution. The problem is bounded because of its equivalence to problem (3.7)-(3.11).

Let $(\tilde{d}, \tilde{v})$ be the solution. Then $\|\tilde{v}\|_1 \leq \|(WYe - \tau e)_+\|_1$. The search direction is generated by solving the modified QP problem

$$\min \psi(d) \tag{3.18}$$
$$\text{s.t. } Cd_x + Dd_y \leq -(Cx + Dy - c), \tag{3.19}$$
$$Ad_x + Bd_y = 0, \tag{3.20}$$
$$Nd_x + Md_y - d_w = 0, \tag{3.21}$$
$$y + d_y \geq 0,\ w + d_w \geq 0, \tag{3.22}$$
$$Wd_y + Yd_w \leq \max\{W\tilde{d}_y + Y\tilde{d}_w, -(WYe - \tau e)\}, \tag{3.23}$$

where $d = (d_x, d_y, d_w)$ and $\psi(d)$ is defined by (3.1).

If $\tilde{v} = 0$ then problem (3.18)-(3.23) is precisely the same as problem (3.1)-(3.6). However, in general situation, there is an essential difference between them since problem (3.18)-(3.23) is always feasible.

Since problem (3.18)-(3.23) is a strictly convex quadratic program, it has the unique solution if and only if it has feasible solutions. The following result suggests a stopping criterion of the algotihm:

Proposition 3.2 *Assume that d is the unique solution of problem (3.18)-(3.23). If either $d = 0$ or $\nabla_x f^\top d_x + \nabla_y f^\top d_y \geq 0$, and $\|(WYe - \tau e)_+\|_1 = 0$, then z is a KKT point of problem $\mathcal{N}(\tau)$.*

Proof. $d \in \Re^{n+2m}$ is the solution of problem (3.18)-(3.23) if and only if d is feasible to the constraints (3.19)-(3.23) and there exist $\lambda \in \Re^\ell_+$, $\mu \in \Re^p$, $\nu \in \Re^m$, $\xi \in \Re^m_+$, $\zeta \in \Re^m_+$ and $\eta \in \Re^m_+$ such that

$$\begin{pmatrix} \nabla_x f \\ \nabla_y f \\ 0 \end{pmatrix} + H \begin{pmatrix} d_x \\ d_y \\ d_w \end{pmatrix} + \begin{pmatrix} C^\top \\ D^\top \\ 0 \end{pmatrix} \lambda + \begin{pmatrix} A^\top \\ B^\top \\ 0 \end{pmatrix} \mu + \begin{pmatrix} N^\top \\ M^\top \\ -I \end{pmatrix} \nu$$
$$- \begin{pmatrix} 0 \\ I \\ 0 \end{pmatrix} \xi - \begin{pmatrix} 0 \\ 0 \\ I \end{pmatrix} \zeta + \begin{pmatrix} 0 \\ W \\ Y \end{pmatrix} \eta = 0 \tag{3.24}$$

and

$$\lambda^\top (C(x + d_x) + D(y + d_y) - c) = 0, \ \xi^\top (y + d_y) = 0, \ \zeta^\top (w + d_w) = 0, \tag{3.25}$$

$$\eta^\top (W d_y + Y d_w - \max\{W\tilde{d}_y + Y\tilde{d}_w, \ -(WYe - \tau e)\}) = 0. \tag{3.26}$$

We prove that

$$\max\{W\tilde{d}_y + Y\tilde{d}_w, \ -(WYe - \tau e)\} = -(WYe - \tau e) \tag{3.27}$$

if $\|(WYe - \tau e)_+\|_1 = 0$. Suppose that $w_i(\tilde{d}_y)_i + y_i(\tilde{d}_w)_i > -(w_i y_i - \tau)$ for some i. Then

$$w_i(\tilde{d}_y)_i + y_i(\tilde{d}_w)_i + w_i y_i - \tau > 0. \tag{3.28}$$

Since $\|(WYe - \tau e)_+\|_1 = 0$, it follows that $\tilde{v} = 0$ and

$$W\tilde{d}_y + Y\tilde{d}_w + WYe - \tau e \leq \tilde{v}, \tag{3.29}$$

which contradicts (3.28). Thus, (3.27) holds. If $d = 0$, then the result follows immediately from (3.24)-(3.27).

By the fact that $\|(WYe - \tau e)_+\|_1 = 0$ and (3.23), $d = 0$ is feasible to problem (3.18)-(3.23), which implies that

$$\nabla_x f^\top d_x + \nabla_y f^\top d_y + \frac{1}{2} d^\top H d \leq 0. \tag{3.30}$$

Since $\nabla_x f^\top d_x + \nabla_y f^\top d_y \geq 0$, we have that $d^\top H d \leq 0$ and hence $d = 0$. The proof is completed. ∎

We define the merit function

$$\phi(z, \tau; \rho) = f(x, y) + \rho \|(y \circ w - \tau e)_+\|_1, \tag{3.31}$$

where $z = (x, y, w) \in \Re^{n+2m}$, $\tau > 0$ is the relaxation parameter, $\rho > 0$ is the penalty parameter, and $\|\cdot\|_1$ is the ℓ_1 norm. This function is the ℓ_1 penalty function, which plays a key role in globalizing the algorithm and helps to decide the stepsize in the derived search direction.

We are now ready to state the algorithm.

Algorithm 3.3 *(The algorithm for problem (1.1)-(1.5))*

Step 1. *Give* $\tau_0 > 0$, $\rho_0 > 0$, $\sigma \in (0, \frac{1}{2})$, $\delta \in (0, 1)$, *and* $H_0 \in \Re^{(n+2m)\times(n+2m)}$, *presd* $= 1$ *(termination parameter). Give* $z^0 \equiv (x^0, y^0, w^0) \in \mathcal{G}$, $\epsilon > 0$. *Let* $k := 0$*;*

Step 2. *Solve the LP subproblem* $\mathcal{B}(z^k, \tau_k)$. *Let* $(\tilde{d}_x^k, \tilde{d}_y^k, \tilde{d}_w^k, \tilde{v}^k)$ *be the solution. Then solve the modified QP subproblem (3.18)-(3.23) to obtain the search direction* $d^k \equiv (d_x^k, d_y^k, d_w^k)$, *and the multiplier vector* $u^k \equiv (\lambda^k, \mu^k, \nu^k, \xi^k, \zeta^k, \eta^k)$*;*

Step 3. *If either* $\|d^k\| \leq \epsilon$ *or* $\nabla f_k^\top d^k \geq -0.1\epsilon$, *and* $\|(W_k Y_k e - \tau_k e)_+\|_1 \leq \epsilon$, *then set* $z^{k+1} = z^k$, $\rho_{k+1} = \rho_k$, *presd* $= 0$ *and go to Step 5; Else if*

$$\psi_k(d^k) - \rho_k(\|(W_k Y_k e - \tau_k e)_+\|_1 - \|\tilde{v}^k\|_1) \leq 0, \tag{3.32}$$

then $\rho_{k+1} = \rho_k$, *else set*

$$\rho_{k+1} = \max\{2\rho_k, \frac{\psi_k(d^k)}{\|(W_k Y_k e - \tau_k e)_+\|_1 - \|\tilde{v}^k\|_1}\}; \tag{3.33}$$

Step 4. *Compute* $\Delta_k(d^k) = \nabla f_k^\top d^k + \rho_{k+1}(\|\tilde{v}^k\|_1 - \|(W_k Y_k e - \tau_k e)_+\|_1)$. *Select the stepsize* $\alpha_k \in (0, 1]$ *by backtracking such that*

$$\phi(z^{k+1}, \rho_{k+1}; \tau_k) \leq \phi(z^k, \rho_{k+1}; \tau_k) + \sigma \alpha_k \Delta_k(d^k), \tag{3.34}$$

where $z^{k+1} = z^k + \alpha_k d^k$*;*

Step 5. *If* $\tau_k \leq \epsilon$ *and either presd* $= 0$ *or* $\|(W_k Y_k e - \tau_k e)_+\|_1 - \|\tilde{v}^k\|_1 \leq 10^{-6}\epsilon$, *or* $\|y \circ w\|_\infty \leq \epsilon$ *and presd* $= 0$, *we terminate the algorithm; Otherwise if* $\tau_k > \epsilon$, *then* $\tau_{k+1} = \delta\tau_k$, *else* $\tau_{k+1} = \tau_k$. *Update* H_k *to* H_{k+1} *by some given procedure. Let* $k := k + 1$ *and go to Step 2.*

We make some remarks on the algorithm.

- The relaxation parameter τ_k is updated in each iteration, which is the same as that of Fukushima, Luo and Pang [8] and is different from the global convergent SQP methods for general MPECs such as those in [6, 13].

- The penalty parameter ρ_k is updated to ρ_{k+1} so that $\Delta_k(d^k) \le -\frac{1}{2}{d^k}^\top H_k d^k$ (see Proposition 3.4), in which case, the stepsize is selected so that the penalty function is decreased "sufficiently" along d^k for fixed parameters ρ_{k+1} and τ_k.

- It is noted that we do not incorporate the information on u^k in the update procedure of the penalty parameter, a general technique used in SQP methods for MPECs [6, 8, 13]. In this case, it is unnecessary that $\rho^* \ge \|u^k\|_\infty$ for all k.

- There are two stopping criteria in the algorithm. The first case is based on the result of Proposition 3.2 and that the relaxation parameter τ_k or the maximum of residues of complementarity constraints is small enough, whereas the other case is when τ_k is small enough and the equation $\|(W_kY_ke - \tau_k e)_+\|_1 = \|\tilde{v}^k\|_1$ holds approximately.

The next result is related to the penalty update of the algorithm.

Proposition 3.4 *If $\rho_{k+1} \ge \|u^k\|_\infty$, then*

$$\psi_k(d^k) - \rho_{k+1}(\|(W_kY_ke - \tau_k e)_+\|_1 - \|\tilde{v}^k\|_1) \le 0. \tag{3.35}$$

Consequently, $\Delta_k(d^k) \le -\frac{1}{2}{d^k}^\top H_k d^k$.

Proof. Since d^k solves problem (3.18)-(3.23), then d^k satisfies equations (3.24)-(3.26). With some further reductions,

$$\begin{aligned}\nabla f_k^\top d^k + {d^k}^\top H_k d^k &= {\lambda^k}^\top (Cx^k + Dy^k - c) - {\xi^k}^\top y^k - {\zeta^k}^\top w^k \\ &\quad - {\eta^k}^\top \max\{W_k\tilde{d}_y^k + Y_k\tilde{d}_w^k,\ -(W_kY_ke - \tau_k e)\} \\ &\le -{\eta^k}^\top \max\{W_k\tilde{d}_y^k + Y_k\tilde{d}_w^k,\ -(W_kY_ke - \tau_k e)\}.\end{aligned} \tag{3.36}$$

Let $\mathcal{I}$ be the index set such that $\max\{w_i^k(\tilde{d}_y^k)_i + y_i^k(\tilde{d}_w^k)_i,\ -(w_i^ky_i^k - \tau_k)\} \le 0$ for $i \in \mathcal{I}$. By problem (3.12)-(3.17), we have $\tilde{v}^k = (W_k\tilde{d}_y^k + Y_k\tilde{d}_w^k + W_kY_ke - \tau_k e)_+$. Thus, for $i \in \mathcal{I}$, if $w_i^k(\tilde{d}_y^k)_i + y_i^k(\tilde{d}_w^k)_i + w_i^ky_i^k - \tau_k \le 0$, then $w_i^ky_i^k - \tau_k \ge 0$, $\tilde{v}_i^k = 0$ and

$$\max\{w_i^k(\tilde{d}_y^k)_i + y_i^k(\tilde{d}_w^k)_i,\ -(w_i^ky_i^k - \tau_k)\} = \tilde{v}_i^k - (w_i^ky_i^k - \tau_k)_+; \tag{3.37}$$

otherwise, $w_i^ky_i^k - \tau_k \ge 0$, $\tilde{v}_i^k = w_i^k(\tilde{d}_y^k)_i + y_i^k(\tilde{d}_w^k)_i + w_i^ky_i^k - \tau_k$, which implies that (3.37) also holds. Since $\|\tilde{v}^k\|_1 \le \|(W_kY_ke - \tau_k e)_+\|_1$, by (3.36) we have

$$\psi_k(d^k) \le \|\eta^k\|_\infty(\|(W_kY_ke - \tau_k e)_+\|_1 - \|\tilde{v}^k\|_1). \tag{3.38}$$

Thus, the result follows from that $\|\eta^k\|_\infty \le \|u^k\|_\infty$. ∎

4. Stationary properties

Since we do not assume the MPEC-LICQ for the MPCC and the nondegeneracy of complementarity constraints, the algorithm may terminate at some points other than the Bouligand stationary point of problem (1.1)-(1.5). We present definitions on some points with generalized stationary properties in this section.

Definition 4.1

(1) A point $z^* \equiv (x^*, y^*, w^*)$ is called a strong stationary point *of problem (1.1)-(1.5) if the point is feasible to the problem and there exist $\lambda^* \in \Re^\ell_+$, $\mu^* \in \Re^p$, $\nu^* \in \Re^m$, and $\pi^* \in \Re^m$ such that equations (2.27)-(2.31) hold.*

(2) A point z^* is called a weak stationary point *of problem (1.1)-(1.5) if it is a feasible point to the problem, and there exist $\lambda^* \in \Re^\ell_+$, $\mu^* \in \Re^p$, $\nu^* \in \Re^m$, and $\pi^* \in \Re^m$ such that equations (2.27)-(2.29) hold.*

(3) A point z^* is called a singular stationary point *of problem (1.1)-(1.5) if it is a feasible point to the problem, and the MPEC-LICQ does not hold at z^*.*

(4) A point z^* is called an infeasible stationary point *of problem (1.1)-(1.5) if it is infeasible to the problem and is a KKT point of problem*

$$\min\ y^\top w \tag{4.1}$$

$$s.t.\ Cx + Dy \le c, \tag{4.2}$$

$$Ax + By = b, \tag{4.3}$$

$$Nx + My - w = q, \tag{4.4}$$

$$y \ge 0,\ \ w \ge 0. \tag{4.5}$$

The strong and some weak stationarities on MPEC have been defined in Scholtes [22]. Under the MPEC-LICQ, a strong stationary point of problem (1.1)-(1.5) is a Bouligand stationary point of the problem. A weak stationary point may not satisfy (2.30) and (2.31), which must hold for a strong stationary point.

The concepts on singular and infeasible stationary points are originated from nonlinear programming, e.g., see [3, 4, 17, 18, 25]. For completeness, we restate their definitions.

Definition 4.2 *Consider the standard nonlinear program*

$$\min\ \ f(x) \quad s.t.\ g(x) \le 0,\ \ h(x) = 0, \tag{4.6}$$

where $f : \Re^n \to \Re$, $g : \Re^n \to \Re^{m_1}$ and $h : \Re^n \to \Re^{m_2}$ are twice continuously differentiable functions.
(1) x^* is a singular stationary point of problem (4.6) if x^* is feasible to the problem and the LICQ does not hold at x^*.
(2) x^* is an infeasible stationary point of problem (4.6) if x^* is infeasible to the problem and is a stationary point of problem $\min_{x\in\Re^n} \|(g(x)_+,\ h(x))\|$, where $g(x)_+ = \max\{g(x), 0\}$ and $\|\cdot\|$ is the ℓ_2 norm.

The following results show the relations of singular and infeasible stationary points between the MPCC and its NLP relaxation problems.

Proposition 4.3

(i) *Suppose that* z^* *is feasible to the problem (1.1)-(1.5). Then for any* $\tau > 0$, z^* *is a singular stationary point of problem (1.1)-(1.5) if and only if* z^* *is a singular stationary point of the relaxation problem* $\mathcal{N}(\tau)$.

(ii) *If for every* k $z^k \in \mathcal{G}$ *is a singular stationary point of the relaxation problem* $\mathcal{N}(\tau_k)$, z^* *is a limit point of* $\{z^k\}$ *as* $\tau_k \to 0$, *then* z^* *is a singular stationary point of problem (1.1)-(1.5).*

(iii) *Suppose that for every* k $z^k \in \mathcal{G}$ *is an infeasible stationary point of the relaxation problem* $\mathcal{N}(\tau_k)$, z^* *is a limit point of* $\{z^k\}$ *as* $\tau_k \to 0$. *If* z^* *is an infeasible point of problem (1.1)-(1.5), then* z^* *is an infeasible stationary point of the problem.*

Proof. (i) Since z^* is feasible to the problem (1.1)-(1.5), constraints (2.15) are inactive. Then the result follows immediately from Definition 2.3 and Definition 4.1 (3).

(ii) For z^k sufficiently close to z^*, the linear dependence of gradients of active constraints of $\mathcal{N}(\tau_k)$ implies that the MPEC-LICQ does not hold at z^*. Then the result follows from Definition 2.3 and Definition 4.1 (3).

(iii) Since $z^k \in \mathcal{G}$, if z^k is an infeasible stationary point of $\mathcal{N}(\tau_k)$, then it is an infeasible point of $\mathcal{N}(\tau_k)$ and is a KKT point of minimizing $\|(y \circ w - \tau_k e)_+\|$ subject to constraints (4.2)-(4.5), that is, for some $i \in \{1, \dots, m\}$, $y_i^k w_i^k > \tau_k$, and there exist bounded multipliers $\lambda^k \in \Re^\ell_+$, $\mu^k \in \Re^p$, $\nu^k \in \Re^m$, $\xi^k \in \Re^m_+$, and $\zeta^k \in \Re^m_+$ such that

$$C^\top \lambda^k + A^\top \mu^k + N^\top \nu^k = 0, \tag{4.7}$$

$$W_k \eta^k + D^\top \lambda^k + B^\top \mu^k + M^\top \nu^k - \xi^k = 0, \tag{4.8}$$

$$Y_k \eta^k - \nu^k - \zeta^k = 0, \tag{4.9}$$

where $\eta^k = \partial \|v\||_{v=(y^k \circ w^k - \tau_k e)_+}$, $(\lambda^k)^\top (Cx^k + Dy^k - c) = 0$, $(\xi^k)^\top y^k = 0$, and $(\zeta^k)^\top w^k = 0$. Thus, if z^* is infeasible to the problem (1.1)-(1.5), then the result follows by taking limit on both sides of (4.7)-(4.9) as $k \in \mathcal{K}$ and $k \to \infty$, where is some infinite subset. ∎

In what follows, we will discuss some properties of the singular and infeasible stationary points of the original problem.

Proposition 4.4 *If the point* z^* *is a singular stationary point of problem (1.1)-(1.5), then*
(i) z^* *is a Fritz-John point of the relaxed problem* $\mathcal{R}(z^*)$, *where* $\mathcal{R}(z)$ *is defined by problem (2.1)-(2.8);*
(ii) $d^* = 0$ *is a solution of problem* $\mathcal{A}(z^*, 0)$, *where* $\mathcal{A}(z, \tau)$ *is defined by problem (3.7)-(3.11).*

Proof. (i) By the theory of nonlinear programming (e.g., see [2]), a point (x, y, w) is a Fritz-John point of the relaxed problem $\mathcal{R}(z^*)$, if it is a feasible point of the relaxed problem, and there exist multipliers $\lambda \in \Re^\ell_+$, $\mu \in \Re^p$, $\nu \in \Re^m$, $\xi \in \Re^m$ and $\zeta \in \Re^m$ such that

$$\begin{aligned}
C^\top\lambda + A^\top\mu + N^\top\nu \qquad &= 0, &(4.10)\\
D^\top\lambda + B^\top\mu + M^\top\nu - \xi \qquad &= 0, &(4.11)\\
-\nu \qquad - \zeta &= 0, &(4.12)
\end{aligned}$$

$$\lambda^\top(Cx + Dy - c) = 0;\ \xi_i y_i = 0,\ i \in \{i : w_i^* = 0\};\ \zeta_i w_i = 0,\ i \in \{i : y_i^* = 0\}. \quad (4.13)$$

Since z^* is a feasible point of problem (1.1)-(1.5), it is a local minimum point of problem (4.1)-(4.5). Thus, it is also a KKT of problem (4.1)-(4.5) because all constraints of this problem are linear constraints. Hence, there exist $\lambda^* \in \Re^\ell_+$, $\mu^* \in \Re^p$, $\nu^* \in \Re^m$, $\xi^* \in \Re^m_+$ and $\zeta^* \in \Re^m_+$ such that

$$\begin{aligned}
C^\top\lambda^* + A^\top\mu^* + N^\top\nu^* \qquad &= 0, &(4.14)\\
D^\top\lambda^* + B^\top\mu^* + M^\top\nu^* - \xi^* \qquad &= -w^*, &(4.15)\\
-\nu^* \qquad - \zeta^* &= -y^*, &(4.16)\\
{\lambda^*}^\top(Cx^* + Dy^* - c) = 0,\ {\xi^*}^\top y^* = 0,\ {\zeta^*}^\top w^* &= 0, &(4.17)
\end{aligned}$$

and ${y^*}^\top w^* = 0$. Moreover, that the MPEC-LICQ does not hold implies that at least one of the multipliers is non-zero vector. Then the result follows immediately from the above two system of equations.

(ii) z^* is feasible to the problem (4.1)-(4.5). Thus, $d^* = 0$ is a feasible solution of problem $\mathcal{A}(z^*, 0)$. Since $y^* \circ w^* = 0$, it is obvious that $d^* = 0$ is also the optimal solution of problem $\mathcal{A}(z^*, 0)$. ∎

Proposition 4.5 *If the point z^* is an infeasible stationary point of problem (1.1)-(1.5), then $d^* = 0$ is an optimal solution of problem $\mathcal{A}(z^*, 0)$.*

Proof. Since $\mathcal{A}(z, \tau)$ is a convex programming problem, a feasible point d of problem $\mathcal{A}(z, \tau)$ is its optimal solution if and only if there exist multipliers $\lambda \in \Re^\ell_+$, $\mu \in \Re^p$, $\nu \in \Re^m$, $\xi \in \Re^m_+$, and $\zeta \in \Re^m_+$ such that

$$\begin{aligned}
C^\top\lambda + A^\top\mu + N^\top\nu \qquad &= 0, &(4.18)\\
W\eta + D^\top\lambda + B^\top\mu + M^\top\nu - \xi \qquad &= 0, &(4.19)\\
Y\eta \qquad\qquad - \nu \qquad - \zeta &= 0, &(4.20)
\end{aligned}$$

where

$$\begin{aligned}
&\eta = \partial\|v\|_1|_{v=(Wd_y+Yd_w+WYe-\tau e)_+},\\
&\lambda^\top(C(x + d_x) + D(y + d_y) - c) = 0,\\
&\xi^\top(y + d_y) = 0,\ \text{and}\ \zeta^\top(w + d_w) = 0.
\end{aligned}$$

Since $W^*Y^*e \geq 0$, $\partial\|v\|_1|_{v=(W^*Y^*e)_+} = e$. Because z^* is a feasible point of problem (4.1)-(4.5), by comparing (4.18)-(4.20) with (4.14)-(4.17), if z^* is a KKT point of problem (4.1)-(4.5), then $d^* = 0$ is an optimal solution of problem $\mathcal{A}(z^*, 0)$. ∎

The following result is useful in the next section.

Proposition 4.6 *Let $z^* \in \mathcal{G}$ and let (d^*, v^*) be a solution of the linear programming problem $\mathcal{B}(z^*, 0)$, where $\mathcal{B}(z, \tau)$ is defined by problem (3.12)-(3.17). If $\|v^*\|_1 = (y^*)^\top w^*$, then z^* is a KKT point of problem (4.1)-(4.5).*

Proof. It is easy to note that $\bar{d} \equiv (\bar{d}_x, \bar{d}_y, \bar{d}_w) = 0$ and $\bar{v} = y^* \circ w^*$ satisfy the constraints of problem $\mathcal{B}(z^*, 0)$. Thus, $(\bar{d}, \bar{v})$ is an optimal solution of $\mathcal{B}(z^*, 0)$ since $\|\bar{v}\|_1 = (y^*)^\top w^*$. Hence, the KKT conditions for problem $\mathcal{B}(z^*, 0)$ hold at $(\bar{d}, \bar{v})$, that is, there exist multipliers $\bar{\lambda} \in \Re^p_+$, $\bar{\mu} \in \Re^m$, $\bar{\nu} \in \Re^m$, $\bar{\xi} \in \Re^m_+$, $\bar{\zeta} \in \Re^m_+$, $\bar{\eta} \in \Re^m_+$, and $\bar{\beta} \in \Re^m_+$ such that

$$
\begin{aligned}
C^\top\bar{\lambda} + A^\top\bar{\mu} + N^\top\bar{\nu} \qquad\qquad\qquad\qquad\qquad &= 0, &(4.21)\\
D^\top\bar{\lambda} + B^\top\bar{\mu} + M^\top\bar{\nu} - \bar{\xi} \qquad + W^*\bar{\eta} \qquad &= 0, &(4.22)\\
-\bar{\nu} \qquad - \bar{\zeta} + \ Y^*\bar{\eta} \qquad &= 0, &(4.23)\\
e - \qquad \bar{\eta} - \bar{\beta} &= 0, &(4.24)\\
(\bar{\lambda})^\top(Cx^* + Dy^* - c) = 0,\ (\bar{\xi})^\top y^* = 0,\ (\bar{\zeta})^\top w^* = 0,\ (\bar{\beta})^\top\bar{v} &= 0. &(4.25)
\end{aligned}
$$

By substituting $\bar{\eta} = e - \bar{\beta}$ into equations (4.22) and (4.23), and selecting $\lambda^* = \bar{\lambda}$, $\mu^* = \bar{\mu}$, $\nu^* = \bar{\nu}$, $\xi^* = \bar{\xi} + W^*\bar{\beta}$ and $\zeta^* = \bar{\zeta} + Y^*\bar{\beta}$, we have equations (4.14)-(4.17). Thus, the result is obtained. ∎

5. Global convergence

Suppose that $\epsilon = 0$ in Algorithm 3.3 and $\{z^k\}$ is an infinite sequence generated by the algorithm. Then $z^k \in \mathcal{G}$ for all $k \geq 0$. Moreover, $\tau_k \to 0$ as $k \to \infty$.

We need the following general assumption throughout this section.

Assumption 5.1

(1) $\mathcal{G} \neq \emptyset$;

(2) Sequence $\{(x^k, y^k)\}$ and sequence $\{\tilde{d}^k\}$ are bounded;

(3) There exist constants $\gamma_1 \geq \gamma_2 > 0$ such that, for all nonnegative integer k and $d \in \Re^{n+m+\ell}$, $\gamma_2\|d\|^2 \leq d^\top H_k d \leq \gamma_1\|d\|^2$.

Assumption 5.1 (1) can be checked easily by phase-I algorithm of linear programming, (2) implies that $\{w^k\}$ is bounded since the point z^k satisfies the equation (2.12). Under the assumption, $\tilde{v}^k$ is bounded.

In order to guarantee the boundedness of the solutions $\{\tilde{d}^k\}$ of problem $\mathcal{B}(z^k, \tau_k)$, we may either add some simple box constraints on d to the problem, or transform the linear program to the quadratic program by adding some coercive quadratic terms on d. These changes will not alter the convergence results developed for the algorithm. Thus, we make the boundedness of $\{\tilde{d}^k\}$ an assumption for flexibility in the algorithmic design.

Problem $\mathcal{B}(z^k, \tau_k)$ is a simple linear program, it has the following property.

Lemma 5.2 *Let $(\bar{d}^k, \bar{v}^k)$ be a feasible solution of problem $\mathcal{B}(z^k, \tau_k)$, where $\mathcal{B}(z, \tau)$ is defined by problem (3.12)-(3.17). Then for any $t \in [0, 1]$, $(t\bar{d}^k, t\bar{v}^k + (1-t)(W_k Y_k e - \tau_k e)_+)$ is also feasible to the problem.*

Proof. Since $z^k \in \mathcal{G}$, we have that $Cx^k + Dy^k - c \leq 0$. Thus, the result follows immediately from that $W_k Y_k e - \tau_k e \leq (W_k Y_k e - \tau_k e)_+$. ∎

Lemma 5.3 *The sequence $\{d^k\}$ is bounded.*

Proof. Since $\tilde{d}^k$ is feasible to the subproblem (3.18)-(3.23) at z^k for all k, we have that $\psi(d^k) \leq \psi(\tilde{d}^k)$. Thus,

$$\psi(d^k)/\|d^k\|^2 \leq \psi(\tilde{d}^k)/\|d^k\|^2, \tag{5.1}$$

If $\|d^k\| \to \infty$ as $k \in \mathcal{K}$ and $k \to \infty$, by taking limit on both sides of (5.1), we have $\gamma_2 \leq 0$ (where γ_2 is a constant in Assumption 5.1), which is a contradiction. ∎

In what follows, we prove that the stepsizes are bounded away from zero, thus the line search procedure is well-defined.

Lemma 5.4 *If $\rho_{k+1} \leq \rho$ ($\rho > 0$ is a constant), there exists a constant $\alpha_0 \in (0, 1]$ such that for all k,*

$$\phi(z^k + \alpha d^k, \rho_{k+1}; \tau_k) \leq \phi(z^k, \rho_{k+1}; \tau_k) + \sigma\alpha\Delta_k(d^k) \tag{5.2}$$

holds for all $\alpha \in [0, \alpha_0]$.

Proof. Note that

$$f(x^k + \alpha d_x^k, y^k + \alpha d_y^k) - f(x^k, y^k) - \alpha \nabla f_k^\top (d_x^k, d_y^k) \leq \frac{1}{2}\theta\alpha^2\|d^k\|^2 \tag{5.3}$$

for some positive constant θ.

By (3.17) and (3.23), we have

$$(W_k d_y^k + Y_k d_w^k + W_k Y_k e - \tau_k e)_+ \leq (W_k \tilde{d}_y^k + Y_k \tilde{d}_w^k + W_k Y_k e - \tau_k e)_+ \leq \tilde{v}^k. \tag{5.4}$$

Since

$$\begin{aligned}&\|((y^k + \alpha d_y^k) \circ (w^k + \alpha d_w^k) - \tau_k e)_+\|_1 - (1-\alpha)\|(y^k \circ w^k - \tau_k e)_+\|_1 \\ &\leq \alpha\|(y^k \circ w^k - \tau_k e + Y_k d_w^k + W_k d_y^k)_+\|_1 + \alpha^2\|d_y^k \circ d_w^k\|,\end{aligned} \tag{5.5}$$

combining with (5.4) and that $\|d_y^k \circ d_w^k\| \le \frac{1}{2}\|d^k\|^2$, we have that

$$\phi(z^k + \alpha d^k, \rho_{k+1}; \tau_k) - \phi(z^k, \rho_{k+1}; \tau_k) - \alpha\Delta_k(d^k) \le \frac{1}{2}(\rho_{k+1} + \theta)\alpha^2\|d^k\|^2. \tag{5.6}$$

By Proposition 3.4 and Assumption 5.1 (3),

$$\Delta_k(d^k) \le -\frac{1}{2}\gamma_2\|d^k\|^2. \tag{5.7}$$

Selecting $\alpha_0 = \min\{1,\ (1-\sigma)\gamma_2/(\rho+\theta)\}$, then for all $\alpha \le \alpha_0$, since $\rho_{k+1} \le \rho$,

$$\begin{aligned}&\phi(z^k + \alpha d^k, \rho_{k+1}; \tau_k) - \phi(z^k, \rho_{k+1}; \tau_k) - \sigma\alpha\Delta_k(d^k)\\ &\le \frac{1}{2}(\rho_{k+1} + \theta)\alpha^2\|d^k\|^2 - \frac{1}{2}(1-\sigma)\alpha\gamma_2\|d^k\|^2 \le 0,\end{aligned} \tag{5.8}$$

which is the desired result. ∎

Based on the above lemmas, we have the following result.

Theorem 5.5 *Suppose that the penalty parameter sequence $\{\rho_k\}$ is bounded. Let $\{z^k\}$ be an infinite sequence generated by Algorithm 3.3, z^* is any limit point of this sequence.*
(i) *If $(y^*)^\top w^* = 0$, then z^* is a strong stationary point of problem (1.1)-(1.5). Moreover, if MPEC-LICQ holds at z^*, then z^* is a Bouligand stationary point of problem (1.1)-(1.5);*
(ii) *If $(y^*)^\top w^* \ne 0$, then z^* is an infeasible stationary point of problem (1.1)-(1.5).*

Proof. (i) By the boundedness of $\{\rho_k\}$, without loss of generality, we may suppose that $\rho_k = \rho_0$ for all $k \ge 0$. It follows from (3.34) that $\phi(z^{k+1}, \rho_0; \tau_k) \le \phi(z^k, \rho_0; \tau_k)$. It is easy to derive that

$$\phi(z^k, \rho_0; \tau_k) - \phi(z^k, \rho_0; \tau_{k-1}) \le \rho_0 m(1-\delta)\tau_{k-1} \tag{5.9}$$

by the fact that

$$\|(w^k \circ y^k - \tau_k e)_+\|_1 \le \|(w^k \circ y^k - \tau_{k-1} e)_+\|_1 + m(1-\delta)\tau_{k-1}. \tag{5.10}$$

Since

$$\sum_{k=0}^{\infty} \tau_k = \frac{\tau_0}{1-\delta}, \tag{5.11}$$

we have

$$\limsup_{k\to\infty} \phi(z^k, \rho_0; \tau_k) < +\infty, \tag{5.12}$$

together with (5.9), we can deduce that $\{\phi(z^k, \rho_0; \tau_k)\}$ is convergent. Similarly, the sequence $\{\phi(z^k, \rho_0; \tau_{k-1})\}$ is also convergent.

Again by (3.34) and Lemma 5.4, we have

$$-\frac{1}{2}{d^k}^\top H_k d^k \ge \Delta_k(d^k) \to 0, \tag{5.13}$$

which implies that $\lim_{k\to\infty} d^k = 0$. Since d^k is the unique solution of problem (3.18)-(3.23), $d = 0$ is the solution of problem (3.18)-(3.23) at z^*. Thus, by Proposition 3.2, z^*

is a KKT point of problem $\mathcal{N}(0)$. Then, by Definition 4.1, z^* is a strong stationary point of problem (1.1)-(1.5).

If MPEC-LICQ holds at z^*, then the result follows from Definition 4.1 and Theorem 2.1.

(ii) Without loss of generality, we suppose that $\{z^k : k \in \mathcal{K}\} \to z^*$. In order to prove this result, we need only to prove that $\|v^*\|_1 = (y^*)^\top w^*$, where v^* is a limit point of $\{\tilde{v}^k : k \in \mathcal{K}\}$. Then the result follows from Proposition 4.6 and Definition 4.1 (4). We assume the contrary. Then there is some $i_0 \in \{1, \ldots, m\}$ such that $y^*_{i_0} w^*_{i_0} > v^*_{i_0} \geq 0$ since $v^*_i \leq y^*_i w^*_i$ for any $i \in \{1, \ldots, m\}$. By (i), $\lim_{k\to\infty} d^k = 0$, then by taking limit at both sides of (3.23) as $k \to \infty$, we have $\max\{W^* \tilde{d}^*_y + Y^* \tilde{d}^*_w, -y^* \circ w^*\} \geq 0$, which implies that

$$w^*_{i_0}(\tilde{d}^*_y)_{i_0} + y^*_{i_0}(\tilde{d}^*_w)_{i_0} \geq 0. \tag{5.14}$$

Taking limit on both sides of (3.17) as $k \to \infty$, we have

$$w^*_{i_0}(\tilde{d}^*_y)_{i_0} + y^*_{i_0}(\tilde{d}^*_w)_{i_0} \leq v^*_i - w^*_{i_0} y^*_{i_0} < 0, \tag{5.15}$$

which contradicts (5.14). The contradiction shows that $\|v^*\|_1 = (y^*)^\top w^*$. ∎

The following rsult is implied by the above theorem.

Corollary 5.6 *Suppose that $\{z^k\}$ is an infinite sequence generated by the algorithm, $\{\rho_k\}$ is the automatically generated penalty parameter sequence. If there is a limit point of $\{z^k\}$ which is a singular stationary point or a weak stationary point of problem (1.1)-(1.5), then $\rho_k \to \infty$.*

In the rest of this section, we consider the case where ρ_k is unbounded. Since the sequence $\{\rho_k\}$ is monotonically nondecreasing, $\rho_k \to \infty$.

Lemma 5.7 *If $\rho_k \to \infty$, then*
(i) $\lim_{k\to\infty} \|(y^k \circ w^k - \tau_k e)_+\|_1$ *exists;*
(ii) $\lim_{k\to\infty} \|(y^{k+1} \circ w^{k+1} - \tau_k e)_+\|_1$ *exists;*
(iii) $\lim_{k\to\infty}(\|\tilde{v}^k\|_1 - \|(y^k \circ w^k - \tau_k e)_+\|_1) = 0$.

Proof. (i) By (3.34), $\phi(z^{k+1}, \rho_{k+1}; \tau_k) \leq \phi(z^k, \rho_{k+1}; \tau_k)$. Thus,

$$\|(y^{k+1} \circ w^{k+1} - \tau_k e)_+\|_1 - \|(y^k \circ w^k - \tau_k e)_+\|_1 \leq \frac{1}{\rho_{k+1}}(f_k - f_{k+1}), \tag{5.16}$$

where $f_k = f(x^k, y^k)$ and $f_{k+1} = f(x^{k+1}, y^{k+1})$. Since

$$\|(y^{k+1} \circ w^{k+1} - \tau_k e)_+\|_1 \geq \|(y^{k+1} \circ w^{k+1} - \tau_{k+1} e)_+\|_1 - m(1-\delta)\tau_k, \tag{5.17}$$

we obtain

$$\|(y^{k+1} \circ w^{k+1} - \tau_{k+1} e)_+\|_1 - \|(y^k \circ w^k - \tau_k e)_+\|_1 \leq \frac{1}{\rho_{k+1}}(f_k - f_{k+1}) + m(1-\delta)\tau_k. \tag{5.18}$$

By Algorithm 3.3, we have either $\rho_{k+1} = \rho_k$ or $\rho_{k+1} \geq 2\rho_k$, $\forall k$. Without loss of generality, assume that there is a sequence of numbers $\{k_j\}$ with $k_0 = 0$ such that $\rho_{k_j} < \rho_{k_{j+1}}$ and $\rho_{i+1} = \rho_i$ if $k_j \leq i < k_{j+1} - 1$, $\forall j$. Moreover, by the assumption, there exists a constant $M > 0$ such that $|f_k| < M$ for all k. Thus,

$$\sum_{k=0}^{\infty} \frac{1}{\rho_{k+1}}(f_k - f_{k+1}) = \sum_{j=0}^{\infty} \frac{1}{\rho_{k_j}}(f_{k_j} - f_{k_{j+1}}) \leq 4M/\rho_0. \tag{5.19}$$

Together with (5.11), we have

$$\limsup_{k\to\infty} \|(y^{k+1} \circ w^{k+1} - \tau_{k+1}e)_+\|_1 < +\infty, \tag{5.20}$$

by (5.18), we can obtain the desired result.

(ii) The result follows from (5.16) and (5.17), since $\lim_{k\to\infty} \frac{1}{\rho_{k+1}}(f_k - f_{k+1}) = 0$ and $\lim_{k\to\infty} \tau_k = 0$.

(iii) By Lemma 5.3, if $\{\tilde{d^k}\}$ is bounded, then $\{d^k\}$ is bounded. Thus by Assumption 5.1, $\{\psi_k(d^k)\}$ is bounded. Dividing ρ_{k+1} on both sides of (3.35), then

$$\frac{1}{\rho_{k+1}}\psi_k(d^k) \leq \|\tilde{v}^k\|_1 - \|(W_kY_ke - \tau_k e)_+\|_1. \tag{5.21}$$

On the other hand, $\|\tilde{v}^k\|_1 \leq \|(W_kY_ke - \tau_k e)_+\|_1$. Thus, the result follows from that $\rho_{k+1} \to \infty$. ∎

Since $(y^k)^\top w^k - m\tau_k \leq \|(y^k \circ w^k - \tau_k e)_+\|_1 \leq (y^k)^\top w^k + m\tau_k$, then

$$\|(y^k \circ w^k - \tau_k e)_+\|_1 - m\tau_k \leq (y^k)^\top w^k \leq \|(y^k \circ w^k - \tau_k e)_+\|_1 + m\tau_k. \tag{5.22}$$

Thus, by Lemma 5.7 (i), $\lim_{k\to\infty}(y^k)^\top w^k$ exists. Since $z^k \in \mathcal{G}$ for all k, if $\lim_{k\to\infty}(y^k)^\top w^k = 0$, then any limit point of $\{z^k\}$ is a feasible point of problem (1.1)-(1.5); otherwise, any limit point of $\{z^k\}$ is an infeasible point of problem (1.1)-(1.5).

Theorem 5.8 *Suppose that $\rho_k \to \infty$ as $k \to \infty$. Assume that $\{z^k\}$ is an infinite sequence generated by Algorithm 3.3, and $\{u^k\}$ is the corresponding dual multiplier sequence.*
(i) *The sequence $\{u^k\}$ is unbounded;*
(ii) *The sequence $\{((y^k)^\top w^k)\}$ is convergent. Moreover, if $\lim_{k\to\infty}(y^k)^\top w^k \neq 0$, then all limit points of the sequence $\{z^k\}$ are infeasible stationary points of problem (1.1)-(1.5).*

Proof. (i) By Proposition 3.4, that $\rho_k \to \infty$ implies that there exists an infinite subsequence $\{u^k, k \in \mathcal{K}\}$ such that $u_k \to \infty$ as $k \to \infty$ and $k \in \mathcal{K}$. Thus, we have the result.

(ii) The existence of $\lim_{k\to\infty}(y^k)^\top w^k$ is derived by the former discussions. It follows from Lemma 5.7 that $\lim_{k\to\infty}\|\tilde{v}^k\|_1$ exists. Let v^* be any limit point of $\{\tilde{v}^k\}$. Then $\|v^*\|_1 = (y^*)^\top w^*$, where $z^* = (x^*, y^*, w^*)$ is any limit point of the sequence $\{z^k\}$. Thus, by Proposition 4.6, z^* is a KKT point of problem (4.1)-(4.5). Since $\lim_{k\to\infty}(y^k)^\top w^k \neq 0$,

z^* is infeasible to the problem (1.1)-(1.5). Thus, z^* is an infeasible stationary point of problem (1.1)-(1.5). ∎

6. Numerical results

We coded the algorithm in MATLAB and run the code under version 6.5 on a COMPAQ personal computer with pentium III processor 450 MHz and WINDOWS 98 system.

The initial paramters are selected as $\sigma = 0.01$, $\delta = 0.1$. The initial relaxation parameter is chosen as $\tau_0 = \max\{(y^0)^\top w^0/m, 1\} \geq 1$ so that we can observe how the algorithm with large relaxation parameter performs. The initial penalty parameter $\rho_0 = 1$. The tolerance for termination is $\epsilon = 5 \times 10^{-7}$. The initial approximate Hessian H_0 is the $(n + 2m)$-order identity matrix. We generate the initial iteration point by the quadratic programming:

$$\min \ \frac{1}{2} y^\top y + \frac{1}{2} w^\top w \tag{6.1}$$

$$\text{s.t. } Cx + Dy \leq c, \tag{6.2}$$

$$Ax + By = b, \tag{6.3}$$

$$Nx + My - w = q, \tag{6.4}$$

$$y \geq 0, \ \ w \geq 0, \tag{6.5}$$

which is solved by the M-function *quadprog.m* in MATLAB toolbox. Generally the derived point is not a feasible point of the original problem (1.1)-(1.5).

Note that the solution of problem (3.12)-(3.17) may not be feasible to problem (3.18)-(3.23) due to termination errors in solving problem (3.12)-(3.17) in a practical implementation of the algorithm. Hence in our implementation, instead of using problem (3.18)-(3.23) directly, at iterate k, we solve the subproblem

$$\min \ \psi_k(d) \tag{6.6}$$

$$\text{s.t. } Cd_x + Dd_y \leq \max\{C\tilde{d}_x^k + D\tilde{d}_y^k, \ -(Cx^k + Dy^k - c)\}, \tag{6.7}$$

$$Ad_x + Bd_y = A\tilde{d}_x^k + B\tilde{d}_y^k, \tag{6.8}$$

$$Nd_x + Md_y - d_w = N\tilde{d}_x^k + M\tilde{d}_y^k - \tilde{d}_w^k, \tag{6.9}$$

$$d_y \geq \min\{-y^k, \ \tilde{d}_y^k\}, \ d_w \geq \min\{-w^k, \ \tilde{d}_w^k\}, \tag{6.10}$$

$$W_k d_y + Y_k d_w \leq \max\{W_k \tilde{d}_y^k + Y_k \tilde{d}_w^k, \ -(W_k Y_k e - \tau_k e)\}. \tag{6.11}$$

The subproblems (3.12)-(3.17) and (6.6)-(6.11) are solved by the M-functions *linprog.m* and *quadprog.m* in MATLAB toolbox, respectively. We use the zero vector as the initial iterate in *linprog.m* and its solution is taken as the initial approximation of *quadprog.m* for each iteration. The approximate Hessian is updated by

$$H_{k+1} = H_k - \frac{H_k \Delta z_k \Delta z_k^\top H_k}{\Delta z_k^\top H_k \Delta z_k} + \frac{\Delta \bar{s}_k \Delta \bar{s}_k^\top}{\Delta \bar{s}_k^\top \Delta z_k}, \tag{6.12}$$

where $\Delta z_k = z_{k+1} - z_k$, $\Delta s_k = \nabla f_{k+1} - \nabla f_k$ and

$$\bar{\Delta s_k} = \begin{cases} \Delta s_k, & \Delta z_k^\top \Delta s_k \geq 0.2 \Delta z_k^\top H_k \Delta z_k, \\ \theta_k \Delta s_k + (1-\theta_k) H_k \Delta z_k, & \text{otherwise} \end{cases} \tag{6.13}$$

with $\theta_k = 0.8 \Delta z_k^\top H_k \Delta z_k / (\Delta z_k^\top H_k \Delta z_k - \Delta s_k^\top \Delta z_k)$.

Similar to [11], our test problems include mathematical programs with quadratic and non-quadratic objective functions. The MPCCs with quadratic objective functions are generated by a Generator QPECgen presented in [12]. The non-quadratic objective functions are derived by adding a cubic function $\frac{1}{3}\{\sum_{i=1}^n x_i^3 + \sum_{i=1}^m y_i^3\}$, which is an approach presented by Fukushima and Tseng [11]. It should be noted that the random number generator of MATLAB version 6.0 or 6.5 can be different from that of MATLAB version 5.3 although we use the same seed 0. Hence the solution (x_{gen}, y_{gen}) and its objective value f_{gen} for the randomly generated test problems may be different from those reported in [11, 12].

We report our results on the QPECgen problems in Table 1, where second_deg is the cardinality of the second-level degenerate index set, mono_M=1 indicates that M is monotone and mono_M=0 not necessarily monotone (both of them are defined in [12]), iter stands for the number of iterations, f_{gen}, f_0 and f^* are the values of objective functions at (x_{gen}, y_{gen}), the initial point and the solution point, respectively. It is noted that f^*s are almost the same as f_{gen} for 9 test problems, but they are different obviously for problems TP9, TP10, TP12. In Table 2, we report some results on these problems, including the penalty parameter at the solution ρ^*, the norms of differences and residues, respectively, defined by

$$\begin{aligned} Norm1 &= \|(x_0, y_0) - (x_{gen}, y_{gen})\|_\infty, \\ Norm2 &= \|(x^*, y^*) - (x_{gen}, y_{gen})\|_\infty, \\ Norm3 &= \|((Cx + Dy - c)_+, Ax + By - b, Nx + My - w - q, y_+, w_+)\|, \\ Norm4 &= \ell_2\text{-norm of residues of KKT conditions of } \mathcal{N}(\tau^*), \\ Norm5 &= \|y \circ w\|_\infty. \end{aligned}$$

In order to further observe the performance of the algorithm in solving TP9, TP10 and TP12, we select $x_0 = x_{gen}$, $y_0 = y_{gen}$ and $w_0 = Nx_0 + My_0 - q$. Thus, $f_0 = f_{gen}$, $Norm1 = 0$. The results are reported in Table 3. It is noted that for TP9 and TP10, the algorithm terminates at the solution which is much close to the given point (x_{gen}, y_{gen}), but for TP12, the algorithm finds an approximately feasible solution with less value of objective function.

We report our results on the MPCCs with non-quadratic objective (NP1-12) in Table 4. The initial points are the same as for problems TP1-12. For NP1-12, the values of $\|u^*\|_\infty$ are respectively 24.3691, 10.4147, 50.4726, 66.4546, 397.7876, 16.0829, 24.7346, 101.1405, 51.7503, $1.6443e+03$, 467.9976, 131.9828.

Since $Norm5 = 1.0000e-07$ for almost all problems except $Norm5 = 1.0047e-07$ for TP1 and $Norm5 = 1.0001e-07$ for NP12, we do not list them in all tables.

Table 1. Numerical results on QPECgen problems

Problem	(n, m, ℓ)	second_deg	mono_M	iter	f_{gen}	f_0	f^*
TP1	(8,50,4)	0	1	17	-176.644510	-165.505699	-176.644514
TP2	(8,100,4)	0	1	15	-663.293629	-660.049153	-663.293634
TP3	(8,150,4)	0	1	15	-581.974063	-579.649556	-581.974080
TP4	(8,200,4)	0	1	19	14.216617	14.412386	14.216597
TP5	(8,50,4)	4	1	19	-108.553062	-101.177468	-108.553066
TP6	(8,100,4)	4	1	15	-614.683018	-608.424829	-614.683024
TP7	(8,150,4)	4	1	12	-562.634084	-556.554199	-562.634095
TP8	(8,200,4)	4	1	17	126.844637	128.139324	126.844622
TP9*	(8,50,4)	4	0	23	-227.360868	-115.879956	-210.542637
TP10*	(8,100,4)	4	0	30	-262.606441	-156.170331	-253.296704
TP11	(8,150,4)	4	0	16	-454.065823	-432.013508	-454.065831
TP12*	(8,200,4)	4	0	13	-8.433461	-5.864383	-7.622670

(For problems with superscript $*$, f^* is not close to f_{gen} obviously)

Table 2. Penalty parameter and residues on QPECgen problems

Problem	ρ^*	Norm1	Norm2	Norm3	Norm4	$\|u^*\|_\infty$
TP1	1	0.4867	1.0157e-05	3.6566e-11	8.8891e-04	6.9936
TP2	1	0.1850	6.7750e-06	3.2570e-11	5.5434e-04	11.6470
TP3	2	0.2431	1.3548e-06	7.2246e-11	4.8193e-05	51.0818
TP4	2.9933	0.1433	1.1278e-04	4.7458e-11	2.3847e-04	39.0432
TP5	1	0.3999	1.3954e-05	1.6278e-11	0.0044	8.5664
TP6	1	0.3069	3.7609e-05	4.2932e-11	0.0011	15.3464
TP7	1	0.4152	2.8656e-05	3.3947e-11	0.0012	24.9688
TP8	4	0.2087	1.3471e-04	5.8273e-11	2.2272e-04	15.8659
TP9*	139.6098	0.7565	0.4114	2.7743e-11	9.8778e-04	927.1287
TP10*	100.3116	0.6991	0.3292	6.3196e-11	7.9732e-04	1.1609e+03
TP11	1	0.3179	9.3904e-06	6.2766e-11	6.2302e-04	14.9197
TP12*	48.7850	0.3234	0.2235	3.7794e-10	1.2252e-06	261.0823

Table 3. Using (x_{gen}, y_{gen}) as the initial point for TP9, TP10 and TP12

Problem	iter	f^*	ρ^*	Norm2	Norm3	Norm4
TP9	19	-227.360871	1	1.9085e-05	5.0641e-12	0.0015
TP10	18	-262.606445	1	1.4596e-05	3.7482e-11	0.0017
TP12	19	-8.524468	4.9981	0.0958	8.7796e-11	0.0176

Table 4. Results on MPCC with non-quadratic objective

Problem	iter	f_0	f^*	ρ^*	Norm3	Norm4
NP1	18	-162.753582	-174.182414	1	3.3770e-11	7.0585e-04
NP2	17	-653.592089	-656.880841	1	2.7578e-11	3.6231e-04
NP3	15	-569.324057	-571.632008	1	6.2597e-11	3.0102e-04
NP4	20	18.126318	17.919837	4	4.7013e-11	4.0734e-04
NP5	14	-98.927034	-106.408978	1	1.4906e-11	0.0015
NP6	15	-602.047401	-608.356983	1	3.1594e-11	0.0017
NP7	14	-546.259927	-552.370597	1	3.4027e-11	0.0015
NP8	16	131.709758	130.437858	4	4.6082e-11	3.6516e-04
NP9	27	-115.282412	-223.821337	149.4795	8.8084e-12	3.3461e-04
NP10	24	-155.096790	-249.908320	120.7078	4.8927e-11	0.0017
NP11	17	-424.833679	-442.946746	1	4.9029e-11	9.0709e-04
NP12	32	-5.148464	-7.069625	2.3161e+04	1.2333e-10	2.4291e-04

Tables 1-4 show that we have obtained the approximate strong stationary points of all test problems. They are also the KKT points of the relaxed problems (2.9)-(2.15) with the relaxation parameter $\tau \leq 5 \times 10^{-7}$. It is noted that the numbers of iterations of the algorithm are not very sensitive to the changes of the second-level degeneracy and the monotonity of a coefficient matrix, these parameters are proved to be important in global convergences of some existing algorithms. Moreover, comparing to the results for QPECgen, we also note that the numbers of iterations of the algorithm do not increase prominently for MPCC with non-quadratic objective functions. Another point of significance is that the algorithm just solves one quadratic programming subproblem at each iteration, which is different from the algorithm of Fukushima and Tseng [11], in which several quadratic programming subproblems may be solved in each iteration.

To observe the performance of the algorithm in comparing with the direct nonlinear programming approach, we also implement the algorithm by selecting $\tau_k = \tau^*$ for all k, where $\tau^* = Norm5$ which is derived together with the results in Tables 1 and 2. The algorithm is terminated when the ℓ_2 norm of the residues of KKT conditions of $\mathcal{N}(\tau^*) \leq Norm4 + \epsilon$ and the ℓ_∞ norm of the complementarity constraints$\leq Norm5 + \epsilon$, where $Norm4$ and $Norm5$ are derive from the tests in Tables 1 and 2. Some results are reported in Table 5, where the speed-up is calculated by 1−the ratio of the computational time of Algorithm 3.3 to that of the algorithm with fixing $\tau_k = \tau^*$.

Although the speed-up is not a very stable indicator in MATLAB enviroment and the errors may be changed with different tests, it may still give us the approximate understanding on the algorithm. It is noted that Algorithm 3.3 needs apparently less computational time and number of iterations as m is larger, whereas the the algorithm with fixing at a very small relaxation parameter seems to perform a little better as m is relatively small.

Table 5. Numerical results with $\tau_k = \tau^* \ \forall\ k \geq 0$.

Problem	iter	ρ^*	f^*	speed-up
TP1	16	1	-176.644514	3.07%
TP2	13	1	-663.293634	-13.81%
TP3	14	2	-581.974080	4.80%
TP4	19	2	14.216597	15.01%
TP5	14	1	-108.553066	-17.08%
TP6	13	2	-614.683024	-3.80%
TP7	15	4	-562.634095	30.75%
TP8	18	1	126.844622	10.69%
TP9*	18	1.2882e+03	-75.976175	-14.56%
TP10*	26	1.1597e+03	-257.182827	-6.27%
TP11*	22	212.5942	-441.110011	29.55%
TP12*	21	206.8722	-5.508203	40.51%

(For problems with superscript $*$, f^* is not close to f_{gen} obviously)

We also apply our algorithm to some special examples. The first example is problem (1.6)-(1.9) presented in the introduction. We select the initial point $(0, 1, 1)$. The algorithm terminates at $(-1.0000, 0, 0.0000)$ in 3 iterations. $\tau^* = 0.01$, $\rho^* = 2$, Norm3=$1.5904e-12$, Norm4=$3.9267e-13$, Norm5=0, $\|u^*\|_\infty = 1$. It is noted that the solution is a strong stationary point, although both the MPEC-LICQ and the nondegeneracy do not hold at this solution.

The second example is presented by Leyffer [14] to demonstrate a failure of PIPA:

$$\min\ x + y \tag{6.14}$$
$$\text{s.t. } x \in [-1, 1], \tag{6.15}$$
$$1 - x - w = 0, \tag{6.16}$$
$$0 \leq w \perp y \geq 0. \tag{6.17}$$

The standard starting point is $(0, 0.02, 1)$, the optimal solution is $(-1, 0, 2)$. The algorithm terminates at $(-1.0000, -0.0000, 2.0000)$ in 3 iterations. $\tau^* = 0.01$, $\rho^* = 2$. Norm3=$1.5543e-15$, Norm4=$8.6456e-15$, Norm5=$6.9389e-18$, $\|u^*\|_\infty = 1$.

The third example is an infeasible MPEC:

$$\min\ \frac{1}{2}(x^2 - y^2) + x + y \tag{6.18}$$
$$\text{s.t. } x \in [-1, 1], \tag{6.19}$$
$$2 \leq x + y \leq 3, \tag{6.20}$$
$$x + y + w = 4, \tag{6.21}$$
$$0 \leq w \perp y \geq 0. \tag{6.22}$$

The initial points are $(0.5, 2, 1.5)$ and $(0, 2.5, 1.5)$, respectively. It is easy to deduce that points $(1, 2, 1)$ and $(1, 1, 2)$ are minimizers of problem (4.1)-(4.5), which are also infeasible stationary points of the MPCC. The algorithm terminates at point $(1.0000, 2.0000, 1.0000)$ in 8 iterations, $\rho^* = 108.8951$ and 15.0923, Norm3=$1.4253e-13$ and $1.2091e-13$, Norm4=$3.5736e-14$ and $3.5230e-14$, respectively. $\|u^*\|_\infty = 10.9972$ and Norm5=2.0000 for both cases. These results show us that Algorithm 3.3 may obtain certain points with weak stationary properties when other methods may fail to find meaningful solutions.

References

[1] M. Anitescu, *On solving mathematical programs with complementarity constraints as nonlinear programs,* Preprint ANL/MCS-P864-1200, Argonne National Laboratory, Argonne, IL, 2000.

[2] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty, *Nonlinear Programming, Theory and Algorithms,* 2nd ed., John Wiley and Sons, New York, 1993.

[3] J.V. Burke, *A sequential quadratic programming method for potentially infeasible mathematical programs,* J. Math. Anal. Appl., 139(1989), 319-351.

[4] J.V. Burke and S.P. Han, *A robust sequential quadratic programming method,* Math. Prog., 43(1989), 277-303.

[5] A.V. de Miguel, M.P. Friedlander, F.J. Nogales and S. Scholtes, *A superlinearly convergent interior-point method for MPECs,* LBS working paper, 2003

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

[7] R. Fletcher, S. Leyffer, D. Ralph and S. Scholtes, *Local convergence of SQP methods for mathematical programs with equilibrium constraints,* University of Dundee Report NA\209, 2002.

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

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

[10] M. Fukushima and J.S. Pang, *Convergence of a smoothing continuation method for mathematical programs with complementarity constraints,* in Ill-posed Variational Problems and Regularization Techniques, M. Theŕa and R. Tichatschke, eds., Springer-Verlag, New York, 1999, 99-110.

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

[12] H.Y. Jiang and D. Ralph, *QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constraints,* Computational Optim. Appl., 13(1999), 25-59.

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

[14] S. Leyffer, *The penalty interior point method fails to converge for mathematical programs with equilibrium constraints,* Numerical analysis report NA/208, Department of Math., University of Dundee, 2002.

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

[16] S. Leyffer, *Complementarity constraints as nonlinear equations: theory and numerical experience,* Preprint, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439, USA.

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

[18] X.W. Liu and Y.X. Yuan, *A robust trust region algorithm for solving general nonlinear programming,* J. Comput. Math., 19(2001), 309-322.

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

[20] A. Raghunathan and L. Biegler, *MPEC formulations and algorithms in process engineering,* Technical report, Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, 2002.

[21] D. Ralph, *Sequential quadratic programming for mathematical programs with linear complementarity constraints,* in Computational Techniques and Applications: CTAC95, R.L. May and A.K. Easton (Eds.), World Scientific, 1996, 663-668.

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

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

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

[25] Y. Yuan, *On the convergence of a new trust region algorithm,* Numer. Math., 70(1995), 515-539.