Journal of Computational Mathematics, Vol.22, No.5, 2004, 735–752.

A SMOOTHING LEVENBERG-MARQUARDT TYPE METHOD FOR LCP *1)

Ju-liang Zhang   Jian Chen
(*Department of Management Science, School of Economics and Management, Tsinghua University, Beijing 100084, China.*)

Abstract

In this paper, we convert the linear complementarity problem to a system of semismooth nonlinear equations by using smoothing technique. Then we use Levenberg–Marquardt type method to solve this system. Taking advantage of the new results obtained by Dan, Yamashita and Fukushima [11, 33], the global and local superlinear convergence properties of the method are obtained under very mild conditions. Especially, the algorithm is locally superlinearly convergent under the assumption of either strict complementarity or certain nonsingularity. Preliminary numerical experiments are reported to show the efficiency of the algorithm.

Mathematics subject classification: 65K05, 90C30, 90C31, 90C33.
Key words: LCP, Levenberg–Marquardt method, Smoothing technique, P_0 matrix, Superlinear convergence.

1. Introduction

Consider the following linear complementarity problem (LCP)

$$\begin{array}{c} y = Mx + q, \\ x \geq 0,\ y \geq 0,\ x^T y = 0, \end{array} \tag{1}$$

where $M \in R^{n\times n}$, $x, y \in R^n$ and $x \geq 0$ means that $x_i \geq 0$ $(i = 1, \ldots, n)$. In this paper, we assume that the solution set of (1) is nonempty. Let X denote the solution set of (1).

LCP has many applications in economic and engineering, see [16] for a survey. A few experts have studied the problem and numerous algorithms were proposed for the problem, for examples, interior methods (see [37] and references therein), nonsmooth Newton methods (see [12, 15]) and smoothing methods (see [27] and references therein). Among these methods, some algorithms are superlinearly (quadratically) convergent under one of the following group of conditions:

(i). M is P matrix and one of the cluster points of the sequence generated by the algorithm is nondegenerate;

(ii). M is P_0 matrix and certain nonsingularity is satisfied at one of the cluster points of the sequence generated by the algorithm.

So we ask whether there exists a method which is superlinearly convergent without assumption of strict complementarity and nonsinguarity at the limit point. In this paper, we give a

* Received August 31, 2002; final revised March 24, 2004.
1) This work is supported in part by National Science Foundation of China (Grant No.70302003, 70071015, 70231010, 10171055) and by China Postdoctoral Science Fo-undation.

method which is superlinearly convergent under the assupmtion of that M is P_0 and either strict complementarity or nonsingularity.

As we known, only the method in [21] is convergent superlinearly/quadratically under the assumptions of either that M is P matrix and strict complementarity or that M is P_0 and certain nonsingularity. Our method is convergent superlinearly under each of the following group conditions simultaneously: (i). M is P_0 matrix and one of the cluster points of the sequence generated by the algorithm is strict complementarity; (ii). M is P_0 matrix and certain nonsingularity is satisfied at one of the cluster points of the sequence generated by the algorithm. So the results in this paper is stronger than the results in [21].

Levenberg–Marquardt method (LMM) is a classical algorithm for solving the following system of nonlinear equations

$$F(x) = 0,$$

where F is a mapping from R^n to R^m. At each iterative point x_k, the search direction is obtained by solving the following linear equation system

$$(F'(x_k)^T F'(x_k) + \mu_k I)d = -F'(x_k)^T F(x_k),$$

where $F'(x)$ denotes the Jacobian of $F(x)$ and $\mu_k > 0$ is a parameter. It is well known that the method is superlinearly convergent if μ_k is updated by an appropriate rule and certain nonsingularity of $F'(x)$ is satisfied at a limit point. Recently, Yamashita and Fukushima [33], Dan, Yamashita and Fukushima [11] proposed a new update rule for μ_k, i.e., $\mu_k = \|F(x_k)\|^\delta$ and proved that LMM is superlinearly(quadratically) convergent under local error bound condition without nonsingularity. Then they applied their method to the linear complementarity problem and obtained an algorithm which has superlinear (quadratical) convergence properties under the conditions that M is P_0 matrix and there exists a cluster point being nondegenerate. This result is very interesting.

In this paper, we convert LCP into a system of semismooth nonlinear equations $F(x, y, \tau) = 0$ by using smoothing technique and by viewing the smoothing parameter as an independent variable. Then LMM type algorithm is proposed to solve this semismooth system. It is similar to [11, 33] that $\mu^k = \|F(x^k, y^k, \tau^k)\|^\delta$. We can prove that our algorithm is globally convergent and the algorithm is superlinealy (quadratically) convergent under the assumption of either strict complementarity or certain nonsingularity. Note that there is an essential difference between our algorithm and the ones in [11, 33] since here we shall keep $\tau_k > 0$ at each iterative point $({x^k}^T, {y^k}^T, \tau^k)^T$. Therefore, our algorithm is not a simple application of Yamashita and Fukushima's algorithm.

Now we explain our notations. Throughout the paper, all of the vectors are column vector. R^n_+ denotes n-dimensional nonnegative orthant, i.e., $x \in R^n_+ \iff x_i \geq 0, i = 1, \ldots, n$, and R^n_{++} denotes the n-dimensional positive orthant, i.e., $x \in R^n_{++} \iff x_i > 0, i = 1, \ldots, n$. Sometimes we use (w, τ) for $(w^T, \tau)^T$ and $w = (x^T, y^T)^T$. $\|\cdot\|$ denotes 2-norm and $\|\cdot\|_\infty$ denotes ∞-norm.

The paper is organized as follows. In Section 2, we give some basic results on smoothing reformulation. The algorithm model and its global convergence are stated in Section 3. In Section 4, we show the local convergence properties of the algorithm. In Section 5, some preliminary numerical results are reported. In section 6, some discussions and conclusions are given.

2. Some Basic Results

Let $\psi : R^2 \to R$ be Fisher-Burmeister function

$$\psi(a, b) = a + b - \sqrt{a^2 + b^2}.$$

Define

$$\Phi(w) := \Phi(x,y) = \begin{pmatrix} Mx+q-y \\ \phi(x,y) \end{pmatrix}$$

where

$$\phi(x,y) = (\psi(x_1,y_1),\ldots,\psi(x_n,y_n))^T \in R^n.$$

Since ψ is a NCP-function, i.e.,

$$\psi(a,b) = 0 \iff a \geq 0, b \geq 0, ab = 0,$$

the following observation is a direct result of the definition of mapping Φ

$$w^* = (x^*, y^*) \text{ solves } (1) \quad \iff w^* \text{ solves } \Phi(w) = 0.$$

However, the system $\Phi(w) = 0$ is nonsmooth. Therefore, let $\psi_\tau : R^2 \to R$ denote the smoothed Fisher-Burmeister function

$$\psi_\tau(a,b) = a + b - \sqrt{a^2 + b^2 + 2\tau^2},$$

where $\tau > 0$ is a smoothing parameter. Then define

$$\Phi_\tau(w) = \Phi_\tau(x,y) = \begin{pmatrix} Mx+q-y \\ \phi_\tau(x,y) \end{pmatrix},$$

where $\phi_\tau(x,y) = (\psi_\tau(x_1,y_1),\ldots,\psi_\tau(x_n,y_n))^T \in R^n$.

So far we view τ as a parameter. In what follows, we will view τ as an independent variable. In order to make this clear, let us write

$$p(a,b,\tau) = \psi_\tau(a,b),$$

$$\theta(x,y,\tau) = \Phi_\tau(x,y).$$

Moreover, we use mapping $F : R^n \times R^n \times [0,+\infty) \to R^n \times R^n \times R$ defined by,

$$F(w,\tau) = F(x,y,\tau) = \begin{pmatrix} \theta(x,y,\tau) \\ \tau \end{pmatrix}.$$

Obviously, since $F(x,y,\tau) = 0$ can deduce $\tau = 0$ automatically, we obtain the following equivalent reformulation

$$w^* = (x^*, y^*) \text{ solves } (1) \quad \iff (w^*, 0) \text{ solves } F(w,\tau) = 0.$$

In this paper, we are interested in developing method for $F(w,\tau) = 0$. Let the solution set of $F(w,\tau) = 0$ be Ω.

Now we give some properties of $p(a,b,\tau)$, $\theta(x,y,\tau)$ and $F(w,\tau)$.

Lemma 2.1. *[14, 27, 26] Function $p(a,b,\tau)$ has the following properties:*

(i). $p(a,b,\tau)$ *is continuously differentiable on* $R^2 \times R_{++}$*;*

(ii). $p'(a,b,\tau) = \begin{pmatrix} 1 - \frac{a}{\sqrt{a^2+b^2+2\tau^2}} \\ 1 - \frac{b}{\sqrt{a^2+b^2+2\tau^2}} \\ -\frac{2\tau}{\sqrt{a^2+b^2+2\tau^2}} \end{pmatrix}$;

(iii). $|\psi(a,b) - p(a,b,\tau)| \leq \sqrt{2}\tau$ *for any* $(a,b,\tau) \in R^2 \times R_{++}$*;*

(iv). $p(a,b,\tau)$ *is strong semismooth at any* $(a,b,\tau) \in R^2 \times R_+$*, i.e.,*

$$p(a+\Delta a, b+\Delta b, \tau+\Delta\tau) - p(a,b,\tau) - V^T(\Delta a, \Delta b, \Delta\tau) = O(\|(\Delta a, \Delta b, \Delta\tau)\|^2)$$

for any $V \in \partial p(a+\Delta a, b+\Delta b, \tau+\Delta\tau)$ *and* $(\Delta a, \Delta b, \Delta\tau) \to 0$*, where* ∂p *is the generalized gradient of* p *in the sense of Clarke .*

Lemma 2.2. *[26] Function* $F(x,y,\tau)$ *has the following properties:*

(i). Function $F(x,y,\tau)$ *is continuously differentiable on* $R^n \times R^n \times R_{++}$ *and*

$$F'(x,y,\tau) = \begin{pmatrix} M & -I & 0 \\ D_{a,\tau} & D_{b,\tau} & d_\tau \\ 0 & 0 & 1 \end{pmatrix}$$

where

$$D_{a,\tau} = diag(\frac{\partial p}{\partial a}(x_1,y_1,\tau), \dots, \frac{\partial p}{\partial a}(x_n,y_n,\tau)) \in R^{n\times n},$$

$$D_{b,\tau} = diag(\frac{\partial p}{\partial b}(x_1,y_1,\tau), \dots, \frac{\partial p}{\partial b}(x_n,y_n,\tau)) \in R^{n\times n},$$

$$d_\tau = (\frac{\partial p}{\partial \tau}(x_1,y_1,\tau), \dots, \frac{\partial p}{\partial \tau}(x_n,y_n,\tau))^T \in R^n;$$

(ii). $F(x,y,\tau)$ *is local Lipschitz continuous and strong semismooth on* $R^n \times R^n \times R_+$*, i.e., for any* $(x,y,\tau) \in R^n \times R^n \times R_+$*, there exist* $L_1 > 0$*,* $L_2 > 0$ *and* $b_1 > 0$ *such that*

$$\|F(x+\Delta x, y+\Delta y, \tau+\Delta\tau) - F(x,y,\tau)\| \leq L_1\|(\Delta x, \Delta y, \Delta\tau)\| \tag{2}$$

$$\|F(x+\Delta x, y+\Delta y, \tau+\Delta\tau) - F(x,y,\tau) - T^T(\Delta x, \Delta y, \Delta\tau)\| \leq L_2\|(\Delta x, \Delta y, \Delta\tau)\|^2 \tag{3}$$

$\forall(\Delta x, \Delta y, \Delta\tau) \in N(0,b_1) = \{(\Delta x, \Delta y, \Delta\tau) | \|(\Delta x, \Delta y, \Delta\tau)\| \leq b_1, \tau+\Delta\tau \geq 0\}$, $T \in \partial F(x+\Delta x, y+\Delta y, \tau+\Delta\tau)$*, where* $\partial F(x,y,\tau)$ *is the generalized gradient in the sense of Clarke;*

(iii). For any solution $(w^*,0) \in \Omega$*, there exists a neighborhood* $N((w^*,0),b_2) = \{(w,\tau) | \|(w,\tau) - (w^*,0)\| \leq b_2, \tau \geq 0\}$ *of* $(w^*,0)$ *such that* $\|F(w,\tau)\|$ *provides an error bound for* $F(w,\tau) = 0$ *on* $N((w^*,0),b_2)$*, i.e., there exists a constant* $c_1 > 0$ *such that*

$$dist((w,\tau),\Omega) \leq c_1\|F(w,\tau)\|, \ \ \forall (w,\tau) \in N((w^*,0),b_2),$$

where $dist((w,\tau),\Omega) = \min_{(\bar{w},0)\in\Omega}\{\|(w,\tau) - (\bar{w},0)\|\}$.

Proof. (i). This is a direct result of Lemma 2.1.

(ii). It is similar to [22, 28] that we can prove that F is locally Lipschitz continuous. And it is similar to [26] that we can prove that F is strong semismooth on $R^n \times R^n \times R_+$.

(iii). From [29, 32, 17], we know that there exist $b_2 > 0$, $\bar{c} > 0$ such that when $w \in N(w^*,b_2) = \{w | \|w - w^*\| \leq b_2\}$

$$dist(w,X) \leq \bar{c}\|\Phi(w)\|. \tag{4}$$

Let

$$dist(w,X) = \|w - \bar{w}\| \text{ where } \bar{w} \in X \tag{5}$$

It follows from Lemma 2.1(iv) that

$$\begin{aligned} |\|\Phi(w)\| - \|\Phi_\tau(w)\|| &\leq \|\Phi(w) - \Phi_\tau(w)\| \\ &\leq \sqrt{2n}\tau. \end{aligned} \tag{6}$$

By (4), (5) and (6), we obtain

$$\begin{aligned} dist((w,\tau),\Omega) &\leq \|(w,\tau) - (\bar{w},0)\| \\ &\leq \|w - \bar{w}\| + \tau \\ &\leq \bar{c}\|\Phi(w)\| + \tau \\ &\leq \bar{c}\|\Phi_\tau(w)\| + \sqrt{2n}\bar{c}\tau + \tau \\ &\leq (\sqrt{2n}\bar{c} + 1)(\|\Phi_\tau(w)\| + \tau). \end{aligned} \tag{7}$$

On the other hand,

$$\|F(w,\tau)\| = \sqrt{\|\Phi_\tau(w)\|^2 + \tau^2} \geq \frac{\sqrt{2}}{2}(\|\Phi_\tau(w)\| + \tau). \tag{8}$$

Let $c_1 = \sqrt{2}(\sqrt{2n}\bar{c} + 1)$, then by (7) and (8), we obtain

$$dist((w,\tau),\Omega) \leq c_1\|F(w,\tau)\|, \ \forall (w,\tau) \in N((w^*,0),b_2).$$

3. Algorithm Model and Global Convergence

In what follows, we apply Levenberg-Marquardt type Method to solve $F(w,\tau) = 0$. The Levenberg-Marquardt Method(LMM) is classical but still one of the most popular methods for solving nonlinear equation system

$$F(w,\tau) = 0. \tag{9}$$

LMM is a kind of Newton-type method and at each iterative point (w^k,τ^k), the search direction is the solution $(\Delta w^k, \Delta\tau^k)$ of the following system of linear equations

$$(F'(w^k,\tau^k)^T F'(w^k,\tau^k) + \mu_k I) \begin{pmatrix} \Delta w \\ \Delta\tau \end{pmatrix} = -F'(w^k,\tau^k)^T F(w^k,\tau^k), \tag{10}$$

where μ_k is a positive parameter and $I \in R^{(2n+1)\times(2n+1)}$ is the identity matrix. Since the coefficient matrix is a positive definite matrix, (10) has the unique solution, which is a good property comparing with pure Newton method or Gauss-Newton method. Moreover, the search direction $(\Delta w^k, \Delta\tau^k)$ is a descent direction of the merit function $\Psi : R^{2n+1} \to R$

$$\Psi(w,\tau) = \frac{1}{2}\|F(w,\tau)\|^2.$$

Therefore, combining LMM with Armijio step based on the merit function Ψ, we may expect that LMM is globally convergent. In particular, we can show that any cluster (w^*,τ^*) is a stationary point of Ψ. In this case, if M is P_0 matrix, we can prove that $\tau^* = 0$ and w^* solves (1). Namely, we have the following theorem.

Theorem 3.1. *Suppose that (w^*,τ^*) is a stationary point of $\Psi(w,\tau)$. If M is P_0 matrix, then $\tau^* = 0$ and w^* solves (1).*

Proof. By Proposition 2.1 in [36] and Lemma 2.1 (ii), we can prove that if M is P_0 matrix and $\tau \neq 0$ then $F'(w,\tau)$ is nonsingularity. If (w^*,τ^*) is a stationary point of $\Psi(w,\tau)$, then

$F'(w^*,\tau^*)^T F(w^*,\tau^*) = 0$. Hence $F(w^*,\tau^*) = 0$. So $\tau^* = 0$. Then w^* is a stationary point of $\|\Phi(w)\|^2$. From Theorem 11 in [15] we know that w^* solves (1).

As for convergence rate of LMM, it is well known that LMM is superlinearly convergent if $F'(w^*,0)$ is nonsingular and μ_k is updated appropriately. This implies that (1) has the local unique solution. Recently, Dan, Yamashita and Fukushima [11] showed that LMM is superlinearly convergent if $\mu_k = \|F(w^k,\tau^k)\|^\delta$ $(0 < \delta \le 2)$ and $F(w,\tau)$ provides an error bound in a neighborhood of $(w^*,0)$, and LMM is quadratically convergent if $\mu_k = \|F(w^k,\tau^k)\|^2$. In this paper, μ_k is updated by $\mu_k = \|F(w^k,\tau^k)\|^\delta (0 < \delta \le 2)$.

However, we note that $\tau^k > 0$ must be satisfied if we propose an iterative method for solving $F(w,\tau) = 0$. Hence we can not apply general LMM to the system $F(w,\tau) = 0$ directly. We must modify the method to meet the requirement. Noting that the search direction $(\Delta w, \Delta\tau)$ given by (10) is the solution of the following optimization problem

$$\min_{(\Delta w,\Delta\tau)\in R^{2n+1}} \theta_k(\Delta w,\Delta\tau) = \frac{1}{2}\left\|F'(w^k,\tau^k)\begin{pmatrix}\Delta w\\ \Delta\tau\end{pmatrix} + F(w^k,\tau^k)\right\|^2 + \frac{1}{2}\mu_k\left\|\begin{pmatrix}\Delta w\\ \Delta\tau\end{pmatrix}\right\|^2,$$

in order to ensure that $\tau^k > 0\ \forall k$, we obtain search direction by solving the following optimization problem

$$\begin{array}{ll}\min_{(\Delta w,\Delta\tau)\in R^{2n+1}} & \theta_k(\Delta w,\Delta\tau)\\ \text{s.t.} & |\Delta\tau| \le \frac{1}{1+\mu_k}\tau^k.\end{array} \tag{11}$$

Clearly, (11) has the unique solution and $(\Delta w^k, \Delta\tau^k)$ is the solution of (11) if and only if there exists a real number $\lambda^k \ge 0$ such that

$$\left(F'(w^k,\tau^k)^T F'(w^k,\tau^k) + \mu_k I + \lambda^k\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right)\begin{pmatrix}\Delta w^k\\ \Delta\tau^k\end{pmatrix} = -F'(w^k,\tau^k)^T F(w^k,\tau^k), \tag{12}$$

$$|\Delta\tau^k| \le \frac{1}{1+\mu_k}\tau^k, \tag{13}$$

$$\lambda^k\left(|\Delta\tau^k| - \frac{1}{1+\mu_k}\tau^k\right) = 0, \tag{14}$$

$$\lambda^k \ge 0. \tag{15}$$

From (12), we know that (w^k,τ^k) is a stationary point of $\Psi(w,\tau)$ if $(\Delta w^k, \Delta\tau^k) = (0,0)$.

Note that there is only bound constraints on $\Delta\tau$ and there is no constraint on Δw, we can solve (11) as follows:

Set $\lambda^k = 0$ and solve (12). If $|\Delta\tau^k| \le \frac{1}{1+\mu_k}\tau^k$, we obtain the solution $(\Delta w^k, \Delta\tau^k)$ of (11). Otherwise, if $\Delta\tau^k > \frac{1}{1+\mu_k}\tau^k$, set $\Delta\tau^k = \frac{1}{1+\mu_k}\tau^k$. Solve (12) with respect to unknown variables $(\Delta w^k, \lambda^k)$. Then we obtain the solution $(\Delta w^k, \Delta\tau^k)$ of (11). If $\Delta\tau^k < -\frac{1}{1+\mu_k}\tau^k$, set $\Delta\tau^k = -\frac{1}{1+\mu_k}\tau^k$. Solve (12) with respect to unknown variables $(\Delta w^k, \lambda^k)$. Then we obtain the solution $(\Delta w^k, \Delta\tau^k)$ of (11).

Now we state our algorithm formally.

Algorithm 3.1.

Step 0. Choose parameters $\alpha \in (0,1)$, $\beta \in (0,1)$, $\gamma \in (0,1)$, $\delta \in (0,2]$, and initial point $(w^0,\tau^0) = (x^0,y^0,\tau^0) \in R^n \times R^n \times R_{++}$. Let $\mu_0 = \|F(w^0,\tau^0)\|^\delta$ and set $k := 0$;

Step 1. If (x^k,y^k,τ^k) satisfies termination condition, stop;

Step 2. Solve (11) to obtain $(\Delta w^k, \Delta\tau^k)$. If

$$\|F(w^k + \Delta w^k, \tau^k + \Delta\tau^k)\| \le \gamma \|F(w^k, \tau^k)\|, \tag{16}$$

then $w^{k+1} = w^k + \Delta w^k$, $\tau^{k+1} = \tau^k + \Delta\tau^k$, and go to Step 4. Otherwise, go to Step 3;

Step 3. Let m_k be the smallest nonnegative integer satisfying

$$\Psi(w^k + \beta^m \Delta w^k, \tau^k + \beta^m \Delta\tau^k) - \Psi(w^k, \tau^k) \le \alpha\beta^m \nabla\Psi(w^k, \tau^k)^T \begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix},$$

set $w^{k+1} = w^k + \beta^{m_k} w^k$, $\tau^{k+1} = \tau^k + \beta^{m_k}\Delta\tau^k$;

Step 4. Set $\mu_{k+1} = \|F(w^{k+1}, \tau^{k+1})\|^\delta$, $k := k+1$, go to Step 1.

Remark. (i). In Step 1, we do not give terminate rule clearly. In practice, we can use $\left\|\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix}\right\| \le \epsilon$ or $\|F'(w^k, \tau^k)F(w^k, \tau^k)\| \le \epsilon$ (ϵ is the precision prescribed) as the termination rule;

(ii). Since

$$\begin{aligned} & \nabla\Psi(w^k, \tau^k)^T \begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix} \\ = & -F(w^k, \tau^k)^T F'(w^k, \tau^k) \left(F'(w^k, \tau^k)^T F'(w^k, \tau^k) + \mu_k I + \lambda^k \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)^{-1} \\ & F'(w^k, \tau^k)^T F(w^k, \tau^k) \\ < & \ 0, \end{aligned}$$

the linesearch procedure in Step 3 can be carried out.

Now we analyze the convergence of the algorithm. In order to prove that the algorithm is globally convergent, we need the following assumption:

Assumption 3.1. The sequence $\{(w^k, \tau^k)\}$ generated by Algorithm 3.1 is bounded.

It is easy to prove that the sequence $\{\tau^k\}$ generated by the algorithm is bounded. For convenience, we assume that Assumption 3.1 holds. In fact, if we assume that X is bounded, then $\{w^k\}$ is bounded by Lemma 2.4, Lemma 3.2 in [17] and the fact

$$\|\Phi(w^k)\| \le \|\Phi_\tau(w^k)\| + 2\sqrt{n}\tau_k \le (2\sqrt{n}+1)\|F(w^k, \tau^k)\| \le (2\sqrt{n}+1)\|F(w^0, \tau^0)\|$$

for all k.

Theorem 3.2. *Suppose that the sequence $\{(w^k, \tau^k)\}$ is generated by Algorithm 3.1 and Assumption 3.1 holds. Then any cluster point of $\{(w^k, \tau^k)\}$ is a stationary point of $\Psi(w, \tau)$.*

Proof. Since $\nabla\Psi(w^k, \tau^k) \ne 0$ can imply that $(\Delta w^k, \Delta\tau^k) \ne 0$, we have

$$\begin{aligned} & \nabla\Psi(w^k, \tau^k)^T \begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix} \\ = & -({\Delta w^k}^T, \Delta\tau^k)^T \left(F'(w^k, \tau^k)^T F(w^k, \tau^k) + \mu_k I + \lambda^k \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix} . \\ < & \ 0. \end{aligned}$$

By Step 2, we know that $\{\Psi(w^k, \tau^k)\}$ is a monotonically decreasing sequence. Hence μ_k is monotonically decreasing and has a limit point. If $\mu_k \to 0$, then $F(w^k, \tau^k) \to 0$. Therefore, any cluster point of $\{(w^k,\tau^k)\}$ is a solution of $F(w,\tau)=0$, hence it is a stationary point of $\Psi(w,\tau)$. If $\lim_{k\to\infty}\mu_k=\bar{\mu}>0$, then we have

$$\begin{aligned}&\nabla\Psi(w^k,\tau^k)^T\begin{pmatrix}\Delta w^k\\ \Delta\tau^k\end{pmatrix}\\ =&-(\Delta {w^k}^T,\Delta\tau^k)^T\left(F'(w^k,\tau^k)^TF(w^k,\tau^k)+\mu_kI+\lambda^k\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)\begin{pmatrix}\Delta w^k\\ \Delta\tau^k\end{pmatrix}\\ <&-\bar{\mu}\left\|\begin{pmatrix}\Delta w^k\\ \Delta\tau^k\end{pmatrix}\right\|^2.\end{aligned}$$

It is similar to the standard arguments that we can prove that any cluster point of the sequence $\{(w^k,\tau^k)\}$ is a stationary point of $\Psi(w,\tau)$.

4. Local Convergence

In order to analyze the local convergence properties, we need the following assumption:

Assumption 4.1. (i). $\{(w^k,\tau^k)\}\to(w^*,0)$, where w^* is a solution of (1).

(ii). w^* is a strict complementarity solution, i.e., $x_i^*+y_i^*>0$, $\forall\ i=1,\ldots,n$.

First we give the following lemma, which says that $\{\|F'(w^k,\tau^k)\|\}$is a bounded sequence.And it is easy to prove the following lemma.

Lemma 4.1. *Suppose that $\{(w^k,\tau^k)\}$ is generated by Algorithm 3.1, then there exists $M_1>0$ such that*

$$\|F'(w^k,\tau^k)\|\le M_1,\ \forall k. \tag{17}$$

In what follows, let $\bar{w}^k$ denote a vector such that

$$\|(\bar{w}^k,0)-(w^k,\tau^k)\|=dist((w^k,\tau^k),\Omega),\ \bar{w}^k\in X. \tag{18}$$

Note that such $\bar{w}^k$ always exists even though the set Ω is nonconvex. Obviously, we have, by the definition of Ω,

$$\|\bar{w}^k-w^k\|=dist(w^k,X).$$

First we show that unit stepsize is accepted for all k sufficiently large. Now we give several lemmas.

Lemma 4.2. *Suppose that Assumption 4.1 holds and $\{(w^k,\tau^k)\}$ is generated by Algorithm 3.1. Let $b=\min\{b_1/2,b_2\}$ where b_1, b_2 are defined in Lemma 2.2. If $(w^k,\tau^k)\in N((w^*,0),b)$, then*

$$\left\|\begin{pmatrix}\Delta w^k\\ \Delta\tau^k\end{pmatrix}\right\|\le c_2 dist\left(\begin{pmatrix}w^k\\ \tau^k\end{pmatrix},\Omega\right),$$

$$\left\|F'(w^k,\tau^k)\begin{pmatrix}\Delta w^k\\ \Delta\tau^k\end{pmatrix}+F(w^k,\tau^k)\right\|\le c_3\left(dist\left(\begin{pmatrix}w^k\\ \tau^k\end{pmatrix},\Omega\right)^{1+\frac{\delta}{2}}\right),$$

where

$$c_2=\sqrt{\left(2L_2^2c_1^\delta b^{2-\delta}+(2M_1L_1^\delta b^\delta+1)\right)},$$

$$c_3=\sqrt{\left(2L_2^2b^{2-\delta}+(2M_1L_1^\delta b^\delta+1)L_1^\delta\right)},$$

and L_1, L_2 are defined as in Lemma 2.2.

Proof. Note that $(\bar{w}^k - w^k, -\frac{1}{1+\mu_k}\tau^k)$ is a feasible solution of (11), then

$$\theta_k(\Delta w^k, \Delta\tau^k) \le \theta_k(\bar{w}^k - w^k, -\frac{1}{1+\mu_k}\tau^k). \tag{19}$$

Since $(w^k, \tau^k) \in N((w^*, 0), b)$, then

$$\begin{aligned} \left\|\begin{pmatrix} \bar{w}^k - w^* \\ 0 \end{pmatrix}\right\| &\le \left\|\begin{pmatrix} \bar{w}^k \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\| + \left\|\begin{pmatrix} w^* \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\| \\ &\le \left\|\begin{pmatrix} w^* \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\| + \left\|\begin{pmatrix} w^* \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\| \le b_1. \end{aligned} \tag{20}$$

Hence $(\bar{w}^k, 0) \in N((w^*, 0), b_1)$. By Lemma 2.2 and (18), we know that

$$\mu_k = \left\|F(w^k, \tau^k)\right\|^\delta \ge \frac{1}{c_1^\delta}\left\|\begin{pmatrix} \bar{w}^k \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\|^\delta, \tag{21}$$

$$\mu_k = \left\|F(w^k, \tau^k)\right\|^\delta = \left\|F(w^k, \tau^k) - F(\bar{w}^k, 0)\right\|^\delta \le L_1^\delta \left\|\begin{pmatrix} \bar{w}^k \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\|^\delta. \tag{22}$$

By the definition of θ_k, Lemma 2.2, (20), (21), (22) and (17) we have

$$\begin{aligned} &\left\|\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix}\right\|^2 \\ \le\ & \frac{2}{\mu_k}\theta_k(\Delta w^k, \Delta\tau^k) \\ \le\ & \frac{2}{\mu_k}\theta_k(\bar{w}^k - w^k, -\frac{1}{1+\mu_k}\tau^k) \\ =\ & \frac{1}{\mu_k}\left(\left\|F(w^k, \tau^k) + F'(w^k, \tau^k)\begin{pmatrix} \bar{w}^k - w^k \\ -\frac{1}{1+\mu_k}\tau^k \end{pmatrix}\right\|^2 + \mu_k\left\|\begin{pmatrix} \bar{w}^k - w^k \\ -\frac{1}{1+\mu_k}\tau^k \end{pmatrix}\right\|^2\right) \\ \le\ & \frac{1}{\mu_k}\left(2\left\|F(w^k, \tau^k) + F'(w^k, \tau^k)\begin{pmatrix} \bar{w}^k - w^k \\ -\tau^k \end{pmatrix} - F(\bar{w}^k, 0)\right\|^2\right. \\ & \left. + 2M_1^2\mu_k^2{\tau^k}^2 + \mu_k\left\|\begin{pmatrix} \bar{w}^k - w^k \\ -\frac{1}{1+\mu_k}\tau^k \end{pmatrix}\right\|^2\right) \\ \le\ & \frac{1}{\mu_k}\left(2L_2^2\left\|\begin{pmatrix} \bar{w}^k - w^k \\ \tau^k \end{pmatrix}\right\|^4 + (2M_1^2L_1^\delta b^\delta + 1)\mu_k\left\|\begin{pmatrix} \bar{w}^k - w^k \\ -\tau^k \end{pmatrix}\right\|^2\right) \\ \le\ & \left(2L_2^2c_1^\delta b^{2-\delta} + (2M_1^2L_1^\delta b^\delta + 1)\right)\left\|\begin{pmatrix} \bar{w}^k \\ 0 \end{pmatrix} - \begin{pmatrix} w^k \\ \tau^k \end{pmatrix}\right\|^2. \end{aligned}$$

Hence let $c_2 = \sqrt{\left(2L_2^2c_1^\delta b^{2-\delta} + (2M_1^2L_1^\delta b^\delta + 1)\right)}$, we obtain the first inequality.

Now we prove that the second inequality holds. In a way similar to proving the first inequality, we can prove that

$$\begin{aligned} &\left\|F'(w^k, \tau^k)\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix} + F(w^k, \tau^k)\right\|^2 \\ \le\ & 2\theta_k(\Delta w^k, \Delta\tau^k) \\ \le\ & 2\theta_k\left(\bar{w}^k - w^k, -\frac{1}{1+\mu_k}\tau^k\right) \\ \le\ & 2L_2^2\left\|\begin{pmatrix} \bar{w}^k - w^k \\ \tau_k \end{pmatrix}\right\|^4 + (2M_1^2L_1^\delta b^\delta + 1)\mu_k\left\|\begin{pmatrix} \bar{w}^k - w^k \\ \tau_k \end{pmatrix}\right\|^2 \\ \le\ & (2L_2^2b^2 + (2M_1^2L_1^\delta b^\delta + 1)L_1^\delta)\left\|\begin{pmatrix} \bar{w}^k - w^k \\ \tau_k \end{pmatrix}\right\|^{(2+\delta)}. \end{aligned}$$

Let $c_3 = \sqrt{\left(2L_2^2 b^2 + (2M_1^2 L_1^\delta b^\delta + 1)L_1^\delta\right)}$, we obtain the result.

Lemma 4.3. *Suppose that Assumption 3.1 and Assumption 4.1 hold and* $\{(w^k, \tau^k)\}$ *is generated by Algorithm 3.1. If* $\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \begin{pmatrix} w^k+\Delta w^k \\ \tau^k+\Delta\tau^k \end{pmatrix} \in N\left(\begin{pmatrix} w^* \\ 0 \end{pmatrix}, b\right)$, *then*

$$dist\left(\begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta\tau^k \end{pmatrix}, \Omega\right) \leq c_4 \left(dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)\right)^{1+\frac{\delta}{2}},$$

where $c_4 = c_1(c_3 + L_2 c_2^{1+\frac{\delta}{2}}(2b)^{\frac{\delta}{2}})$. *Especially, there exists a positive constant* $b_3 > 0$ *such that*

$$dist\left(\begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta\tau^k \end{pmatrix}, \Omega\right) \leq b_3 \Longrightarrow dist\left(\begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta\tau^k \end{pmatrix}, \Omega\right) \leq \frac{1}{2} dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right).$$

Proof. By Lemma 4.2, we have

$$\left\|\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix}\right\| \leq c_2 dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right),$$

and

$$\left\|F'(w^k, \tau^k)\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix} + F(w^k, \tau^k)\right\| \leq c_3 \left(dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)\right)^{1+\frac{\delta}{2}}.$$

Then it follows from Lemma 2.2 (ii), (iii) that

$$\begin{aligned}
& \tfrac{1}{c_1} dist\left(\begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta\tau^k \end{pmatrix}, \Omega\right) \\
\leq\ & \|F(w^k + \Delta w^k, \tau^k + \Delta\tau^k)\| \\
\leq\ & \left\|F(w^k, \tau^k) + F'(w^k, \tau^k)\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix}\right\| + L_2 \left\|\begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix}\right\|^2 \\
\leq\ & c_3 dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)^{1+\frac{\delta}{2}} + L_2 c_2^{1+\frac{\delta}{2}}(2b)^{\frac{\delta}{2}} dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)^{1+\frac{\delta}{2}} \\
\leq\ & (c_3 + L_2 c_2^{1+\frac{\delta}{2}}(2b)^{\frac{\delta}{2}}) dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)^{1+\frac{\delta}{2}}.
\end{aligned}$$

Let $c_4 = c_1(c_3 + L_2 c_2^{1+\frac{\delta}{2}}(2b)^{\frac{\delta}{2}})$, then the conclusion holds.

Lemma 4.4. *Suppose that Assumption 3.1, 4.1 hold and* $\{(w^k, \tau^k)\}$ *are generated by Algorithm 3.1. Then there exists positive integer* $\bar{k}$ *such that* $m_k = 0$ *for all* $k \geq \bar{k}$, *i.e., the iteration formula is as follows*

$$\begin{pmatrix} w^{k+1} \\ \tau^{k+1} \end{pmatrix} = \begin{pmatrix} w^k \\ \tau^k \end{pmatrix} + \begin{pmatrix} \Delta w^k \\ \Delta\tau^k \end{pmatrix}, \ \forall k \geq \bar{k}.$$

Proof. Let $r = \min\left\{\frac{\hat{b}}{1+2c_2}, \frac{1}{2c_4}\right\}$, where $\hat{b} = \min\{b_1/2, b_2, b_3\}$ and b_1, b_2 are defined as Lemma 2.2 and b_3 is defined as Lemma 4.3. Since $\begin{pmatrix} w^* \\ 0 \end{pmatrix}$ is a solution of $F(w, \tau) = 0$, then there exists a positive integer $\bar{k} > 0$ by Assumption 4.1 such that

$$\|F(w^{\bar{k}}, \tau^{\bar{k}})\|^{\frac{\delta}{2}} \leq \frac{\gamma}{c_1^{1+\frac{\delta}{2}} c_4 L_1}, \gamma \text{ is chosen as in Algorithm 3.1} \tag{23}$$

and

$$\left\| \begin{pmatrix} w^{\bar{k}} \\ \tau^{\bar{k}} \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| \le r. \tag{24}$$

Now we prove that the following hold for all $k \ge \bar{k}$:

(i). (16) holds;

(ii). $\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix} \in N\left(\begin{pmatrix} w^* \\ 0 \end{pmatrix}, \hat{b}\right)$;

(iii). $\begin{pmatrix} w^{k+1} \\ \tau^{k+1} \end{pmatrix} = \begin{pmatrix} w^k \\ \tau^k \end{pmatrix} + \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix}$.

We prove these results by induction.

When $k = \bar{k}$, since

$$\begin{aligned} \left\| \begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| &\le \left\| \begin{pmatrix} w^k \\ \tau^k \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| + \left\| \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix} \right\| \\ &\le r + c_2 dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right) \\ &\le r + c_2 \left\| \begin{pmatrix} w^k \\ \tau^k \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| \\ &\le (1 + c_2) r \\ &\le \hat{b}. \end{aligned}$$

So (ii) holds.

Let $\begin{pmatrix} \hat{w}^k \\ 0 \end{pmatrix} \in \Omega$ be such that

$$\left\| \begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix} - \begin{pmatrix} \hat{w}^k \\ 0 \end{pmatrix} \right\| = dist\left(\begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix}, \Omega\right).$$

It is similar to (20) that we can prove that $(\hat{w}^k, 0) \in N((w^*, 0), b_1)$. Then by Lemma 4.3, Lemma 2.2 (ii) and (iii) and (23), we have

$$\begin{aligned} \|F(w^k + \Delta w^k, \tau^k + \Delta \tau^k)\| &= \|F(w^k + \Delta w^k, \tau^k + \Delta \tau^k) - F(\hat{w}^k, 0)\| \\ &\le L_1 dist\left(\begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix}, \Omega\right) \\ &\le L_1 c_4 dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)^{1+\frac{\delta}{2}} \\ &\le L_1 c_4 c_1^{1+\frac{\delta}{2}} \|F(w^k, \tau^k)\|^{1+\frac{\delta}{2}} \\ &\le \gamma \|F(w^k, \tau^k)\| \end{aligned} \tag{25}$$

So (i) holds. Hence (iii) holds by the definition of Algorithm 3.1.

Now we assume that (i), (ii) and (iii) hold for $k = \bar{k}, \bar{k}+1, \ldots, l$. We need to show that (i), (ii) and (iii) hold for $k = l + 1$.

Obviously, $\begin{pmatrix} w^{k+1} \\ \tau^{k+1} \end{pmatrix} \in N\left(\begin{pmatrix} w^* \\ 0 \end{pmatrix}, \hat{b}\right)$, $\forall k = \bar{k}, \bar{k}+1, \ldots, l$ by assumption.

It follows from (ii), (iii) and Lemma 4.3 that $\forall k = \bar{k}+1, \ldots, l, l+1$

$$\begin{array}{rl} & dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right) \\ \leq & \frac{1}{2} dist\left(\begin{pmatrix} w^{k-1} \\ \tau^{k-1} \end{pmatrix}, \Omega\right) \leq \cdots \leq \left(\frac{1}{2}\right)^{k-\bar{k}} \left\| \begin{pmatrix} w^{\bar{k}} \\ \tau^{\bar{k}} \end{pmatrix} - \begin{pmatrix} \bar{w}^{\bar{k}} \\ 0 \end{pmatrix} \right\| \\ \leq & \left(\frac{1}{2}\right)^{k-\bar{k}} \left\| \begin{pmatrix} w^{\bar{k}} \\ \tau^{\bar{k}} \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| \\ \leq & r\left(\frac{1}{2}\right)^{k-\bar{k}}. \end{array}$$

Therefore, by Lemma 4.2 we obtain that $\forall k = \bar{k}+1, \ldots, l, l+1$

$$\left\| \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix} \right\| \leq c_2 dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right) \leq c_2 r \left(\frac{1}{2}\right)^{k-\bar{k}}.$$

It follows that

$$\begin{array}{rl} & \left\| \begin{pmatrix} w^{l+1} + \Delta w^{l+1} \\ \tau^{l+1} + \Delta \tau^{l+1} \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| \\ \leq & \left\| \begin{pmatrix} w^{\bar{k}} \\ \tau^{\bar{k}} \end{pmatrix} - \begin{pmatrix} w^* \\ 0 \end{pmatrix} \right\| + \sum_{k=\bar{k}}^{l+1} \left\| \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix} \right\| \\ \leq & r + c_2 r \sum_{k=\bar{k}}^{l} \left(\frac{1}{2}\right)^{k-\bar{k}} \\ \leq & r + c_2 r \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^{k} \\ \leq & (1+c_2) r \\ \leq & \hat{b}. \end{array}$$

Then (ii) holds.

Note that $\{F(w^k, \tau^k)\}$ is a monotonically decreasing sequence, we have that

$$\|F(w^{l+1}, \tau^{l+1})\| \leq \cdots \leq \|F(w^{\bar{w}}, \tau^{\bar{k}})\| \leq \frac{\gamma}{c_4 L_1 c_1^{1+\frac{\delta}{2}}}.$$

Similar to the proof of (25), we can show that

$$\|F(w^{l+1} + \Delta w^{l+1}, \tau^{l+1} + \Delta \tau^{l+1})\| \leq \gamma \|F(w^{l+1}, \tau^{l+1})\|.$$

Hence (i) holds. Then (iii) holds by the definition of the algorithm.

Combining Lemma 4.3 and Lemma 4.4, we have the following theorem.

Theorem 4.1. *Suppose that Assumption 3.1 and Assumption 4.1 hold and $\{(w^k, \tau^k)\}$ is generated by Algorithm 3.1. Then $\left\{dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)\right\}$ converges to 0 superlinearly. If $\delta = 2$, then $\left\{dist\left(\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \Omega\right)\right\}$ converges to 0 quadratically.*

Proof. It follows from Lemma 4.4 that for all k sufficiently large, the iteration formula is as follows

$$\begin{pmatrix} w^{k+1} \\ \tau^{k+1} \end{pmatrix} = \begin{pmatrix} w^k \\ \tau^k \end{pmatrix} + \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix}$$

and

$$\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix} \in N\left(\begin{pmatrix} w^* \\ 0 \end{pmatrix}, b\right).$$

From Lemma 4.3, we know that the conclusion is true.
Remark. Here we assume that (w^k, τ^k) is convergent for simplicity. In fact, we need only assume that one cluster point of the sequence (w^k, τ^k) is a strict complementarity solution of (1). The we can use the technique in [33] to prove that that the whole sequence $\{(w^k, \tau^k)\}$ converges to the cluster point.

Now we study the superlinear convergence of the algorithm from the other hand.
Assumption 4.2. $\{\|(F'(w^k, \tau^k))^{-1}\|\}$is bounded above.

This assumption is used in many literatures to prove the superlinear convergence of their algorithms [2, 6]. It is noticed that this assumption implies that the solution set of (1) is singleton but not necessarily a strict complementarity solution.

From Assumption 4.2, we have that there exists a constant $c_1' > 0$ such that for sufficiently large k

$$\begin{aligned} \mu_k &= \|F(w^k, \tau^k)\| \\ &= \|F(w^k, \tau^k) - F(w^*, 0)\| \\ &= \left\| F'(w^k, \tau^k) \begin{pmatrix} w^* - w^k \\ -\tau^k \end{pmatrix} \right\|^{\delta} + O\left(\left\| \begin{pmatrix} w^* - w^k \\ -\tau^k \end{pmatrix} \right\|^{2\delta} \right) \\ &\geq \frac{1}{\|F'(w^k \tau^k)^{-1}\|} \left\| \begin{pmatrix} w^* - w^k \\ -\tau^k \end{pmatrix} \right\|^{\delta} + O\left(\left\| \begin{pmatrix} w^* - w^k \\ -\tau^k \end{pmatrix} \right\|^{2\delta} \right) \\ &\geq \frac{1}{c_1'} \left\| \begin{pmatrix} w^k - w^* \\ \tau^k \end{pmatrix} \right\|^{\delta}. \end{aligned} \tag{26}$$

From (26), it is simlar to Lemma 4.2, 4.3, 4.4 and Theorem 4.1 that we can obtain

Lemma 4.5. *Suppose that Assumption 4.2 holds and $\{(w^k, \tau^k)\}$ is generated by Algorithm 3.1. There exit $c_2' > 0, c_3' > 0$ such that if $(w^k, \tau^k) \in N((w^*, 0), b)$, then*

$$\left\| \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix} \right\| \leq c_2' \left\| \begin{pmatrix} w^k - w^* \\ \tau^k \end{pmatrix} \right\|,$$

$$\left\| F'(w^k, \tau^k) \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix} + F(w^k, \tau^k) \right\| \leq c_3' \left\| \begin{pmatrix} w^k - w^* \\ \tau^k \end{pmatrix} \right\|^{1+\frac{\delta}{2}}.$$

Lemma 4.6. *Suppose that Assumption 3.1 and Assumption 4.2 hold and $\{(w^k, \tau^k)\}$ is generated by Algorithm 3.1. There exists $c_4' > 0$ such that if $\begin{pmatrix} w^k \\ \tau^k \end{pmatrix}, \begin{pmatrix} w^k + \Delta w^k \\ \tau^k + \Delta \tau^k \end{pmatrix} \in N\left(\begin{pmatrix} w^* \\ 0 \end{pmatrix}, b \right)$, then*

$$\left\| \begin{pmatrix} w^k + \Delta w^k - w^* \\ \tau^k + \Delta \tau^k \end{pmatrix} \right\| \leq c_4' \left\| \begin{pmatrix} w^k - w^* \\ \tau^k \end{pmatrix} \right\|^{1+\frac{\delta}{2}}.$$

Especially, there exists a positive constant $b_3 > 0$ such that

$$\left\| \begin{pmatrix} w^k + \Delta w^k - w^* \\ \tau^k + \Delta \tau^k \end{pmatrix} \right\| \leq b_3 \Longrightarrow \left\| \begin{pmatrix} w^k + \Delta w^k - w^* \\ \tau^k + \Delta \tau^k \end{pmatrix} \right\| \leq \frac{1}{2} \left\| \begin{pmatrix} w^k - w^* \\ \tau^k \end{pmatrix} \right\|.$$

Lemma 4.7. *Suppose that Assumption 3.1, 4.2 hold and $\{(w^k, \tau^k)\}$ are generated by Algorithm 3.1. Then there exists positive integer $\bar{k}$ such that $m_k = 0$ for all $k \geq \bar{k}$, i.e., the iteration formula is as follows*

$$\begin{pmatrix} w^{k+1} \\ \tau^{k+1} \end{pmatrix} = \begin{pmatrix} w^k \\ \tau^k \end{pmatrix} + \begin{pmatrix} \Delta w^k \\ \Delta \tau^k \end{pmatrix}, \ \forall k \geq \bar{k}.$$

Theorem 4.2. *Suppose that Assumption 3.1 and Assumption 4.2 hold and* $\{(w^k, \tau^k)\}$ *is generated by Algorithm 3.1. Then*

$$\lim_{k\to\infty} \frac{\left\| \begin{pmatrix} w^{k+1} - w^* \\ \tau^{k+1} \end{pmatrix} \right\|}{\left\| \begin{pmatrix} w^k - w^* \\ \tau^k \end{pmatrix} \right\|^{1+\frac{\delta}{2}}} = 0.$$

5. Implementation and Numerical Experiments

In this section, we test our algorithm on some typical test problems. The program code was written in MATLAB and run in MATLAB 6.0 environment. The internal function QP in optimal toolbox is used to solve subproblem (11). The parameters are chosen as follows $\gamma = 0.9,\ \alpha = 0.1,\ \beta = 0.5,\ \delta = 1,\ \tau_0 = 0.1$. The stop criterion is $\|(\Delta w^k, \Delta\tau^k)\| \le 10^{-10}$. The numerical results are summarized in Table 1 and the test problems are introduced as follows.

LCP1: $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $q = (-1, -1)$. This problem is given in Cottle et [10], the initial point is $(0, \cdots, 0)$.

LCP2: $M = \begin{pmatrix} 0 & 0 & 10 & 20 \\ 0 & 0 & 30 & 15 \\ 10 & 20 & 0 & 0 \\ 30 & 15 & 0 & 0 \end{pmatrix}$, $q = (-1, -1, -1, -1)$.This problem is given in Cottle et [10], the initial point is $(0, \cdots, 0)$.

LCP3: $M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$, $q = -e$, $n = 16$.This linear complementarity problem is one for which Murty has shown that Lemke's complementary pivot algorithm is known to run in a number of pivots exponential in the number of variables in the problem (see [23]). The initial point is $(0, \cdots, 0)$.

LCP4: $M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 & 2 \\ 0 & 1 & 2 & \cdots & 2 & 2 \\ 0 & 0 & 1 & \cdots & 2 & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 2 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$, $q = -(1, \cdots, 1, 0)$. This problem is given in Chen and Ye [8], the initial point is $(0, \cdots, 0)$.

LCP5: $M = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}$, $q = (1, 0, -1)$. This problem is from Yamashita, Dan and Fukushima [35]. The initial point is $(0, \cdots, 0)$.

LCP6: $M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}$, $q = (0, -1, 0)$. This problem is from Yamashita, Dan and Fukushima [35]. The initial point is $(0, \cdots, 0)$.

LCP7: $M = \begin{pmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 0 & 1 \\ 2 & 0 & 2 & 2 \\ -1 & -1 & -2 & 0 \end{pmatrix}$, $q = (-8, -6, -4, 3)$. This problem is from Yamashita and Fukushima [36]. The initial point is $(0, \cdots, 0)$.

LCP8: $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$, $q = (0, 0, 1)$. This problem is from Chen and Ye [8]. The initial point is $(1, \cdots, 1)$.

LCP9: $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 1 \end{pmatrix}$, $q = (0, 0, 1)$. This problem is from Zhao and Li [38]. The initial point is $(1, \cdots, 1)$.

LCP10: $M = \begin{pmatrix} 4 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 4 & -2 & \cdots & 0 & 0 \\ 0 & 1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix}$, $q = -e$. This problem is from Ahn [1]. The initial point is $(0, \cdots, 0)$.

LCP11: $M = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 4 \end{pmatrix}$, $q = -e$. This problem is from Geiger and Kanzow [18]. The initial point is $(0, \cdots, 0)$.

LCP12: $M = diag(1/n, 2/n, \cdots, (n-1)/n, 1)$, $q = -e$. This problem is from Zhao and Li [38]. The initial point is $(0, \cdots, 0)$

Table 1. *Dim.* is the dimension n of the problem, *No. of the iter.* is the number of the iterations and *Residual* is $\|\Psi(x, y)\|$. Last 1(Last 2, Last 3) means the first (second, third) backward the terminate iteration and $\|F\| = \|F(x, y, \tau)\|$.

Problem	Dim.	No. of Iter.	Residual	τ	Last 3 $\|F\|$	Last 2 $\|F\|$	Last 1 $\|F\|$
LCP1	2	7	3.5e-34	1.2e-20	1.5e-5	2.3e-10	2.6e-17
LCP2	4	7	2.4e-22	9.2e-13	0.002	5.5e-6	2.2e-11
LCP3	16	23	0.02	7.7e-12	0.229	0.223	0.217
LCP4	100	21	1.9e-7	7.0e-15	0.002	0.001	6.2e-4
LCP4	300	28	7.8e-8	3.0e-14	0.001	6.0e-4	4.0e-4
LCP4	500	30	9.9e-8	1.8e-15	0.02	0.001	4.4e-4
LCP5	3	7	6.3e-33	2.8e-22	7.0e-6	2.3e-10	1.1e-16
LCP6	3	7	1.2e-25	2.5e-13	8.3e-5	1.5e-8	5.0e-13
LCP7	4	20	3.4e-25	1.1e-16	9.1e-4	8.7e-7	8.2e-13
LCP8	3	11	3.2e-26	4.0e-14	6.8e-4	4.8e-7	2.5e-13
LCP9	3	8	6.1e-24	1.1e-12	0.001	1.9e-6	3.5e-12
LCP10	300	18	2.0e-28	1.1e-15	3.7e-4	1.4e-7	2.0e-14
LCP10	500	21	3.7e-25	4.4e-14	9.7e-4	9.2e-7	8.6e-13
LCP11	300	20	6.3e-30	1.0e-17	1.1e-4	1.2e-8	3.5e-15
LCP11	500	24	7.0e-31	6.0e-21	1.8e-5	3.1e-10	1.2e-15
LCP12	20	56	0.82	3.0e-12	1.296	1.287	1.277

From Table 1, we see that the algorithm can solve these problems efficiently. From Column 6–Column 8, we know that $\|F\|$ tends to 0 rapidly at the end of the algorithm. This shows the superlinear convergence behavior of our method. However, there is no common knowledge on the choice for δ. For some problems, the lager δ is , the better the algorithm performs, whereas for other problems, the smaller δ is , the better the algorithm performs.

6. Conclusions

In this paper, we combine the smoothing technique and LMM to propose a new algorithm for $LCP(M, q)$. The global and local superlinear convergence of the algorithm are obtained under very mild conditions. Especially, the algorithm is locally superlinearly convergent under the assumption of that M is P_0 and either nonsingularity or strict complementarity. This property is interesting. Therefore the conditions are very weak comparing with the existing algorithms. Numerical experiments show that the algorithm is efficient. Furthermore, these experiments demonstrate the superlinear (even quadratical) convergence. However, we need to solve (11) exactly for the algorithm in this paper. This is expensive for large scale problems. How to improve the current version of the algorithm and to propose a more practicable algorithm is an interesting topic. Furthermore, to develop a method which is superlinearly convergent without assumption of nonsingularity and strict complementarity is a interesting topic. We are going to study the problems further.

Acknowledgement: The author gives indebted thanks to Professor Naihua Xiu and two anonymous referees for their valuable suggestions and comments.

References

[1] Ahn, B.H., Iterative methods for linear complementarity problem with upperbounds and lowerbounds, *Math. Prog.*, **26** (1983), 265-315.

[2] Burke, J., Xu, S., A non-interior predictor-corrector path following algorithm for the monotone linear complementarity problem, *Math. Prog.*, **87** (2000), 113-130.

[3] Chen, B., Chen, X., A global and local superlinear comtinuation-smoothing method for $P_0 + R_0$ and monotone NCP, *SIAM J. Optim.*, **9** (1999), 624-645.

[4] Chen, B., Harker, P.T., A non-interior-point comtinuation method for linear complementarity problems, *SIAM J. Matrix Anal. Appl.*, **14** (1993), 1168-1190.

[5] Chen, B., Xiu, N., A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing function, *SIAM J. Optim.*, **9** (1999), 605-623.

[6] Chen, B., Xiu, N., A superlinear noninterior one-step continuation method for monotone LCP in the absence of strict complementarity, **Journal of Optimization Theory and Application**, **108**:2 (2001), 317-332.

[7] Chen, C., Mangasarian, O.L., Smoothing methods for convex inequalities and linear complementarity problems, *Math. Prog.*, **71** (1995), 51-69.

[8] Chen, X., Ye, Y., On smoothing methods for the P_0 matrix linear complementarity problem, *SIAM J. Optim.*, **11** (2000), 341-363.

[9] Clark, F.H., Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[10] Cottle, R.W., Pang, J.-S., Stone, R.W., The Linear Complementarity Problem, Academic, New York, NY, 1992.

[11] Dan, H., Yamashita, N., Fukushima, M., Convergence Properties of the inexact Levenberg-Marquardt method under local error bound conditions, Technical Report 2001-003, Department of Applied Mathematics and Physics, Kyoto University, (Jannuary 2001).

[12] De Luca, T., Fancchinei, F., Kanzow, C., A semismooth equation approach to the solution of nonlinear complementarity problems, *Math. Prog.*, **75** (1996), 407-439.

[13] Engelke, S., and Konzow, C., Predictor-corrector smoothing methods for the solution of linear programs, Preprint 153, Institute of Applied Mathematics, University of Hamburg, Hamburg, March, 2000.

[14] Engelke, S., and Konzow, C., Improved smoothing-type methods for the solution of linear programs, Preprint, Institute of Applied Mathematics, University of Hamburg, Hamburg, March, 2000.

[15] Facchinei, F., Kanzow, C., A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, *Math. Prog.*, **76** (1997), 493-512.

[16] Ferris, M.C., Pang, J.-S, Engineering and economic applications of complementarity problems, *SIAM Review*, **39** (1997), 669-713.

[17] Fisher, A., An new NCP–function and its use for the solution of complementarity problems, In D.Z. Du, L. Qi, and R.S. Womersley, Recent Advances in Nonsmooth Optimization, (1995), 88-105.

[18] Geiger, C., Kanzow, K., On the resolution of monotone complementarity problems, *Computational Optimization and Applications*, **5** (1996), 155-173.

[19] Harker, P.T., Pang, J.-S, Finite-dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications, *Math. Prog.*, **48** (1990), 161-220.

[20] Hotta, K., Yoshise, A., Global convergence of a class of non-interior-point algorithm using Chen-Harker-Kanzow-Smale functions for nonlinear complementarity problems, *Math. Prog.*, **86** (1999), 105-133.

[21] Huang, Z., Qi, L., Sun, D., Sub-quadratic convergence of a smoothing Newton algorithm for the P_0- and monotone LCP, Mathematical Programming, to appear.

[22] Kanzow, C., Some noninterior continuation methods for linear complementarity problems, *SIAM J. Matrix Anal. Appl.*, **17** (1996), 851-868.

[23] Murty, K.G.,Linear Complementarity, Linear and Nonlinear Programming, Helderman-Verlag, Berlin, 1988.

[24] Qi, L., Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.*, **18** (1993), 227-244.

[25] Qi, L., Sun, J., A nonsmooth version of Newton's method, *Math., Prog.*, **58** (1993), 353-367.

[26] Qi, L., Sun, D., Smoothing functions and smoothing Newton method for complemenarity and variational inequality problems, *Journal of Optimization Theory and Application*, **113**:1 (2002), 121-147.

[27] Qi, L., Sun, D., A survey of some nonsmooth equations and smoothing Newton methods, in Andrew Eberhard, Barney Glover, Robin Hill and Daniel Ralph eds., Progress in optimization, 121–146, Appl. Optim., 30, Kluwer Acad. Publ., Dordrecht, 1999.

[28] Qi, L., Sun, D., Zhou, G., A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Math. Prog.*, **87** (2000), 1-35.

[29] Robinson, S.M., Some continuity properties of polyhedral multifunctions, *Math. Prog. Study*, **14** (1981), 206-214.

[30] Smale, S., Algorithms for solving equations, In Proceedings of the International Congress of Mathematicians, *AMS, Providence*, (1987), 172-195.

[31] Tseng, P., Error bounds and superlinear convergence analysis of some Newton-type methods to optimization, In: Di Pillo, G., Giannessi, F., eds. Nonlinear Optimization and Application 2, Kluwer Academic Publishers B.V., 1998.

[32] Tseng, P., Growth behaviour of aclass of merit functions for the nonlinear complementarity problem, *Jour of Optimization Theory and Applications*, **89** (1996), 17-37.

[33] Yamashita, N., Fukushima, M., On the rate of convergence of the Levenberg-Marquardt method, Computing [Suppl], **15** (2001), 239-249.

[34] Yamashita, N., Fukushima, M., The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem, *SIAM J. Optim.*, **11** (2001), 364-379.

[35] Yamashita, N., Dan, H., Fukushima, M., On the identification of degenerate indices in the nonlinear complementarity problem with proxmal point algorithm, Technical Report 2001-003, Department of Applied Mathematics and Physics, Kyoto University (February 2001).

[36] Yamashita, N., Fukushima, M., Modified Newton methods for solving a semismooth reformulation of monotone complementarity, *Math. Prog.*, **76** (1997), 469-491.

[37] Yoshise, A., Complementarity problems, In: Terlaky, T., eds. Interior point methods of mathematical programming, *Kluwer Academic Publishers*, (1996), 297-339.

[38] Zhao, Y., Li, D., A global and locally superlinearly convergent non-interior-point algorithm for P_0 LCPs, Tech. Report, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, 2001.