Reformulation: Nonsmooth, Piecewise Smooth,
Semismooth and Smoothing Methods, pp. 45–63
Edited by M. Fukushima and L. Qi

A Non-Interior Predictor-Corrector Path-Following Method for LCP

James V. Burke* and Song Xu*

Abstract In a previous work the authors introduced a non–interior predictor-corrector path following algorithm for the monotone linear complementarity problem. The method uses Chen–Harker–Kanzow–Smale smoothing techniques to track the central path and employs a refined notion for the neighborhood of the central path to obtain the boundedness of the iterates under the assumption of monotonicity and the existence of a feasible interior point. With these assumptions, the method is shown to be both globally linearly convergent and locally quadratically convergent. In this paper it is shown that this basic approach is still valid without the monotonicity assumption and regardless of the choice of norm in the definition of the neighborhood of the central path. Furthermore, it is shown that the method can be modified so that only one system of linear equations needs to be solved at each iteration without sacrificing either the global or local convergence behavior of the method. The local behavior of the method is further illuminated by showing that the solution set always lies in the interior of the neighborhood of the central path relative to the affine constraint. In this regard, the method is fundamentally different from interior point strategies where the solution set necessarily lies on the boundary of the neighborhood of the central path relative to the affine constraint. Finally, we show that the algorithm is globally convergent under a relatively mild condition.

Key Words linear complementarity, smoothing methods, path-following methods

1 INTRODUCTION

Consider the *linear complementarity problem*:

*Department of Mathematics, University of Washington, Seattle WA 98195. email: burke@math.washington.edu, songxu@math.washington.edu

LCP(q, M): Find $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$Mx - y + q = 0, \tag{1.1}$$

$$x \geq 0, y \geq 0, x^T y = 0, \tag{1.2}$$

where $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$.

In this paper, we study extensions, refinements, and properties of a non–interior predictor–corrector path following algorithm for this problem which was recently proposed by Burke and Xu [3]. Here the path to be followed is the *central path*

$$C = \{(x, y) : 0 < \mu,\ 0 < x,\ 0 < y,\ Mx - y + q = 0, \text{ and } Xy = \mu^2 e\} \tag{1.3}$$

where, following standard usage in the interior-point literature [16], we denote by $e \in \mathbb{R}^n$ the vector each of whose components is 1 and by X the diagonal matrix whose diagonal entries are given by the vector $x \in \mathbb{R}^n$. The algorithm is based on Chen-Harker-Kanzow-Smale smoothing techniques [6, 14, 19] and as such relies on the function

$$\phi(a, b, \mu) = a + b - \sqrt{(a - b)^2 + 4\mu^2}\ . \tag{1.4}$$

This function is a member of the Chen–Mangasarian class of smoothing functions for the problem LCP(q, M) [8]. It is easily verified that for $\mu > 0$

$$\phi(a, b, \mu) = 0 \text{ if and only if } 0 < a,\ 0 < b, \text{ and } ab = \mu^2. \tag{1.5}$$

Another function having this property is the smoothed Fischer-Burmeister function

$$\psi(a, b, \mu) = a + b - \sqrt{a^2 + b^2 + 2\mu^2}\ , \tag{1.6}$$

which is first studied by Kanzow [14]. For simplicity, in this paper, we will focus on the function ϕ. However, the same analysis can be easily carried out if the function ϕ is replaced by ψ.

Based on the functions ϕ, ψ and other smoothing functions, a number of non–interior path following algorithms have recently been proposed that are globally convergent or globally linearly convergent and possess rapid local convergence properties [2, 3, 5, 4, 6, 7, 9, 10, 12, 13, 14, 17, 18, 20, 21, 22]. Interested readers are referred to [3] for more references. In [3], Burke and Xu propose the first non–interior predictor–corrector algorithm for monotone LCP. The central idea is to apply Newton's method to equations of the form $F(x, y, \mu) = v$ for various choices of the right hand side v where the function $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+$ is given by

$$F(x, y, \mu) := \begin{bmatrix} Mx - y + q \\ \Phi(x, y, \mu) \\ \mu \end{bmatrix}, \tag{1.7}$$

with

$$\Phi(x, y, \mu) = \begin{bmatrix} \phi(x_1, y_1, \mu) \\ \cdots \\ \phi(x_n, y_n, \mu) \end{bmatrix}. \tag{1.8}$$

Note that

$$F(x, y, \mu) = 0 \tag{1.9}$$

if and only if (x, y) solves LCP(q, M), and

$$F(x, y, \mu) = \begin{bmatrix} 0 \\ 0 \\ \bar{\mu} \end{bmatrix} \quad \text{with} \quad \bar{\mu} \neq 0 \tag{1.10}$$

if and only if (x, y) is on the central path C with corresponding smoothing parameter $\bar{\mu}$. Burke and Xu [3] show that their algorithm is both globally linearly convergent and locally quadratically convergent under standard hypotheses and requires one or two matrix factorizations in each step. In this paper, we study this algorithm further and show that it has a number of very nice properties. (1) We observe that our convergence results are valid for any norm, in particular, they are valid for the ∞-norm, which is preferred in practical implementations. (2) We show that the solution set of the LCP is contained in the interior of every slice of the neighborhood relative to the affine constraints. Thus, this neighborhood is fundamentally different from those used in the interior-point literature, where the solution set necessarily lies on the boundary of the neighborhood relative to the affine constraints. This property explains why the non-interior method can be initiated from any point in the space and why the predictor steps are so efficient. (3) We show that the algorithm can be modified so that only one matrix factorization is needed in each iteration. (4) We show that essentially the same convergence theory can be obtained if the monotonicity and strict feasibility hypotheses are replaced by the hypotheses that the matrix M is both a P_0 and an R_0 matrix. (5) Finally, we establish the global convergence for the algorithm under relatively mild conditions.

The plan of the paper is as follows. In Section 2, we study some structural properties of the neighborhood of the central path. In Section 3, we state our predictor–corrector algorithm and show that it is well–defined. Section 4 contains the convergence analysis.

A few words about our notation are in order. All vectors are column vectors with the superscript T denoting transpose. The notation $\mathbb{R}^n$ is used for real n–dimensional space and $\mathbb{R}^{n\times n}$ is used to denote the set of all $n \times n$ real matrices. We denote the non–negative orthant in $\mathbb{R}^n$ by $\mathbb{R}^n_+$ and its interior by $\mathbb{R}^n_{++}$. Given $x, y \in \mathbb{R}^n$, we write $x \leq y$ to indicate that $y - x \in \mathbb{R}^n_+$. The notation $\|\cdot\|$ is used to denote a norm. Most of the results in this paper are established for an arbitrary norm. However, certain norms do play a special role. Given $x \in \mathbb{R}^n$, we denote by $\|x\|_1$, $\|x\|_2$, and $\|x\|_\infty$ the 1–norm, 2–norm, and ∞–norm of x, respectively. A matrix $M \in \mathbb{R}^{n\times n}$ is said to be a P_0 matrix if all of its principal minors are non–negative. The matrix M is said to be an R_0

matrix if the problem LCP$(0, M)$ has the unique solution $(x, y) = (0, 0)$. If the matrix M is positive semi–definite, then the problem LCP(q, M) is said to be a monotone linear complementarity problem.

2 A NEIGHBORHOOD OF THE CENTRAL PATH

Let $\|\cdot\|$ be a given norm on $\mathbb{R}^n$. By the equivalence of norms on $\mathbb{R}^n$, there exist positive constants n_l and n_u such that

$$n_l \|x\|_\infty \leq \|x\| \leq n_u \|x\|_\infty, \quad \text{for all} \quad x \in \mathbb{R}^n. \tag{2.1}$$

For example, when $\|.\|$ is the ∞-norm, $n_l = n_u = 1$, and when $\|.\|$ is the 2-norm, $n_l = 1$ and $n_u = \sqrt{n}$. We take the set

$$\mathcal{N}(\beta) := \left\{ (x, y) \;\middle|\; \begin{array}{l} Mx - y + q = 0, \;\Phi(x, y, \mu) \leq 0, \\ \|\Phi(x, y, \mu)\| \leq \beta\mu \text{ for some } \mu > 0 \end{array} \right\}, \tag{2.2}$$

as our neighborhood of the central path, where $\beta > 0$ is given. This neighborhood can be viewed as the union of the *slices*

$$\mathcal{N}(\beta, \mu) := \{ (x, y) \,:\, Mx - y + q = 0, \;\Phi(x, y, \mu) \leq 0, \;\|\Phi(x, y, \mu)\| \leq \beta\mu\} \tag{2.3}$$

for $\mu > 0$. When the norm is chosen to be the 2-norm, the neighborhood is reduced to the one studied by Burke and Xu [3].

The neighborhood (2.2) refines the neighborhood concept introduced in [2] by requiring that all points in the neighborhood satisfy the additional inequality $\Phi(x, y, \mu) \leq 0$. It will be shown that if the algorithm is initiated in this neighborhood, then the inequality $\Phi(x, y, \mu) \leq 0$ is automatically satisfied at subsequent iterates. Hence the addition of this inequality does not complicate the structure of the algorithm. In the monotone case, this inequality is key to establishing the boundedness of the iterates. However, the boundedness of the iterates can be assured in a number of ways. For example, the assumption that the matrix M is an R_0 matrix also suffices. Thus, in order to keep the discussion at a general level, we introduce the following boundedness hypothesis.

Hypothesis (A): For any $\beta > 0$ and $\mu_0 > 0$, the set

$$\bigcup_{0<\mu\leq\mu_0} \mathcal{N}(\beta, \mu)$$

is bounded

Lemma 2.1 *[2, Proposition 2.4] If M is an R_0 matrix, then for every $\beta > 0$ and $\mu_0 > 0$ the set*

$$\bigcup_{0<\mu\leq\mu_0} \{ (x, y) \,:\, Mx - y + q = 0, \;\|\Phi(x, y, \mu)\| \leq \beta\mu\}$$

is bounded.

Remark. Clearly, the boundedness of the set given in Lemma 2.1 implies the boundedness of the set given in Hypothesis (A).

Lemma 2.2 *Assume that the problem LCP(q, M) is monotone and has a feasible interior point. Then for any $\beta > 0$ and $\mu_0 > 0$, the set*

$$\bigcup_{0<\mu\leq\mu_0} \mathcal{N}(\beta,\mu)$$

is bounded. Indeed, for any $(x,y) \in \bigcup_{0<\mu\leq\mu_0} \mathcal{N}(\beta,\mu)$, we have for $i = 1, 2, \ldots, n$

$$-\frac{\beta\mu_0}{2n_l} \leq x_i \leq \frac{\bar{x}^T\bar{y} + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) + n\max\{\mu_0^2, \frac{\beta^2\mu_0^2}{4n_l^2}\}}{\bar{y}_i}$$

$$-\frac{\beta\mu_0}{2n_l} \leq y_i \leq \frac{\bar{x}^T\bar{y} + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) + n\max\{\mu_0^2, \frac{\beta^2\mu_0^2}{4n_l^2}\}}{\bar{x}_i},$$

where $(\bar{x}, \bar{y})$ is any feasible interior point, that is, a point satisfying

$$M\bar{x} - \bar{y} + q = 0, \quad \bar{x} > 0, \ \bar{y} > 0.$$

Proof. Let $\beta > 0$, $0 < \mu \leq \mu_0$, and $(x,y) \in \mathcal{N}(\beta,\mu)$ be given, and let $(\bar{x}, \bar{y})$ be a feasible interior point for LCP(q, M). First observe that if $-\delta \leq \phi(a,b,\mu)$, then $-\delta/2 < \min\{a,b\}$. To see this, note that the condition $-\delta \leq \phi(a,b,\mu)$ implies that

$$0 < \sqrt{[(a+\delta/2) - (b+\delta/2)]^2 + 4\mu^2} \leq (a+\delta/2) + (b+\delta/2). \qquad (2.4)$$

Squaring both sides and cleaning up yields $0 < \mu^2 \leq (a+\delta/2)(b+\delta/2)$. Thus, since at least one of $(a+\delta/2)$ and $(b+\delta/2)$ must be positive by (2.4), both must be positive yielding $-\delta/2 < \min\{a,b\}$. It follows from (2.1) that

$$\|\Phi(x,y,\mu)\|_\infty \leq \frac{1}{n_l}\|\Phi(x,y,\mu)\| \leq \frac{\beta\mu}{n_l}.$$

This observation implies that

$$x_i > -\frac{\beta\mu}{2n_l} \geq -\frac{\beta\mu_0}{2n_l}, \text{ and } y_i > -\frac{\beta\mu}{2n_l} \geq -\frac{\beta\mu_0}{2n_l}, \text{ for } i = 1, 2, \ldots, n \,. \qquad (2.5)$$

Next, note that if $0 \leq a$ and $0 \leq b$, then the inequality $\phi(a,b,\mu) \leq 0$ implies that $0 \leq a + b \leq \sqrt{(a-b)^2 + 4\mu^2}$. Again, by squaring and cleaning up, we see that this gives $ab \leq \mu^2$. This observation implies that

$$x_i y_i \leq \mu_0^2 \text{ for each } i \in \{1, \ldots, n\} \text{ with } 0 < x_i, \ 0 < y_i \,. \qquad (2.6)$$

We conclude the proof by noting that monotonicity yields $0 \le (\bar{x}-x)^T(\bar{y}-y)$, or equivalently $\bar{x}^T y + \bar{y}^T x \le \bar{x}^T \bar{y} + x^T y$. This inequality plus those in (2.5) and (2.6) yield the bound

$$\begin{aligned}
\sum_{y_i>0} \bar{x}_i y_i + \sum_{x_i>0} \bar{y}_i x_i &\le \bar{x}^T \bar{y} + x^T y - \left[\sum_{y_i<0} \bar{x}_i y_i + \sum_{x_i<0} \bar{y}_i x_i \right] \\
&\le \bar{x}^T \bar{y} + x^T y + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) \\
&\le \bar{x}^T \bar{y} + \sum_{\substack{x_i>0 \\ y_i>0}} x_i y_i + \sum_{\substack{x_i<0 \\ y_i<0}} x_i y_i + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) \\
&\le \bar{x}^T \bar{y} + n \max\{\mu_0^2, \frac{\beta^2\mu_0^2}{4n_l^2}\} + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) .
\end{aligned}$$

It follows that if $y_i > 0$, then

$$y_i \le \frac{\bar{x}^T \bar{y} + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) + n \max\{\mu_0^2, \frac{\beta^2\mu_0^2}{4n_l^2}\}}{\bar{x}_i},$$

and, if $x_i > 0$, then

$$x_i \le \frac{\bar{x}^T \bar{y} + \frac{\beta\mu_0}{2n_l}(\|\bar{x}\|_1 + \|\bar{y}\|_1) + n \max\{\mu_0^2, \frac{\beta^2\mu_0^2}{4n_l^2}\}}{\bar{y}_i}.$$

∎

An important property of the neighborhood $\mathcal{N}(\beta)$ that distinguishes it from its counter part in the interior point literature is that the solution set

$$\mathcal{S} = \{(x,y) : Mx - y + q = 0, x \ge 0, y \ge 0, x^T y = 0\}, \tag{2.7}$$

is contained in the interior of the slice $\mathcal{N}(\beta,\mu)$ relative to the affine set

$$\Lambda = \{(x,y) : Mx - y + q = 0\}, \tag{2.8}$$

for all $\mu > 0$ and $\beta > 2n_u$. This property partially explains why the non-interior path-following method can be initiated from any point in the space and why locally the Newton predictor steps are so efficient.

Theorem 2.1 *For any $\mu > 0$ and $\beta > 2n_u$, the solution set S is contained in the interior of the slice $\mathcal{N}(\beta,\mu)$ relative to the affine set Λ.*

Proof. Let

$$\mathcal{N}_\infty(\beta,\mu) := \{ (x,y) \,:\, Mx - y + q = 0,\ \Phi(x,y,\mu) \le 0,\ \|\Phi(x,y,\mu)\|_\infty \le \frac{\beta}{n_u}\mu\} . \tag{2.9}$$

It follows from (2.1) that

$$\mathcal{N}_\infty(\beta,\mu) \subseteq \mathcal{N}(\beta,\mu).$$

Thus, it suffices to prove that S is contained in the interior of the set $\mathcal{N}_\infty(\beta,\mu)$ relative to the affine set Λ.

First note that if $x_i y_i \le \mu^2$, then

$$\begin{aligned}\phi(x_i,y_i,\mu) &= x_i + y_i - \sqrt{(x_i-y_i)^2+4\mu^2}\\ &= x_i + y_i - \sqrt{(x_i+y_i)^2+4(\mu^2 - x_i y_i)}\\ &\le 0.\end{aligned}$$

Also, if $x_i \ge 0, y_i \ge 0$ and $x_i y_i = 0$, then either $x_i = 0$ or $y_i = 0$. If $x_i = 0$ and $y_i \ge 0$, then

$$\begin{aligned}\phi(x_i,y_i,\mu) &= x_i + y_i - \sqrt{(x_i-y_i)^2+4\mu^2}\\ &= y_i - \sqrt{y_i^2+4\mu^2}\\ &\ge y_i - (y_i + 2\mu)\\ &= -2\mu > -\frac{\beta}{n_u}\mu.\end{aligned}$$

Similarly, if $y_i = 0$ and $x_i \ge 0$, then again $\phi(x_i,y_i,\mu) > -\frac{\beta}{n_u}\mu$. Now if $(x^*,y^*) \in S$, then by the continuity of function ϕ, there is a $\delta > 0$ such that for all (x,y) in

$$\mathcal{O}(\delta) = \{(x,y) : \|x-x^*\|_\infty \le \delta,\ \ \|y-y^*\|_\infty \le \delta\}$$

we have $\phi(x_i,y_i,\mu) \ge -\frac{\beta}{n_u}\mu$ and $x_i y_i \le \mu^2$ for all $i = 1,\ldots,n$. Therefore for any $(x,y) \in \mathcal{O}(\delta)$, we have

$$\Phi(x,y,\mu) \le 0,\ \|\Phi(x,y,\mu)\|_\infty \le \frac{\beta}{n_u}\mu,$$

and so (x^*,y^*) is in the interior of the set $\mathcal{N}_\infty(\beta,\mu)$ relative to the affine set Λ.
∎

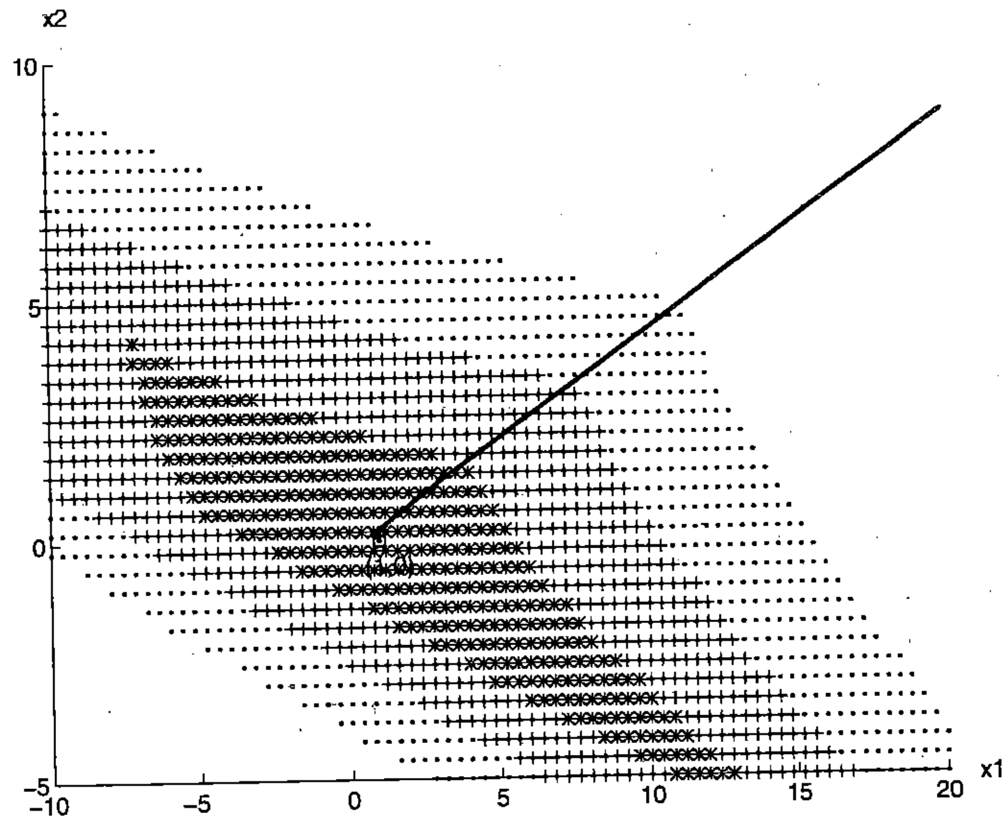
To illustrate this property, consider the problem LCP(q,M) given in [8, Example 5.1], where

$$M = \begin{bmatrix} 1 & 2\\ 2 & 5\end{bmatrix},\quad q = \begin{bmatrix} -1\\ -1\end{bmatrix}.$$

The unique solution of this problem is $(x_1,x_2) = (1,0)$ and $(y_1,y_2) = (0,1)$. For $\mu > 0$, let

$$\mathcal{N}_x(\beta,\mu) := \{\,(x_1,x_2) : \exists y \text{ such that } Mx - y + q = 0,\ \Phi(x,y,\mu) \le 0,$$
$$\|\Phi(x,y,\mu)\|_\infty \le \beta\mu\}$$

Figure 1.1 The nested slices of $\mathcal{N}_x(\beta, \mu)$ with $\beta = 4$ and $\mu = 15, 10, 5$.



be the projection of the slice $\mathcal{N}(\beta, \mu)$ onto the x-coordinate. In figure 1, several slices of $\mathcal{N}_x(\beta, \mu)$ are drawn with $\beta = 4$ and $\mu = 15, 10, 5$ respectively. Note that the solution $(x_1, x_2) = (1, 0)$ is in the interior of these slices.

We conclude this section by cataloging a few technical properties of the function $\phi(a, b, \mu)$ for later use.

Lemma 2.3 *The function ϕ defined in (1.4) has the following properties:*

(i) *[14] The function $\phi(a, b, \mu)$ is continuously differentiable on $\mathbb{R}^2 \times \mathbb{R}_{++}$.*

(ii) *[3, Lemma 2.2] The function $\phi(a, b, \mu)$ is concave on $\mathbb{R}^2 \times \mathbb{R}_{++}$,*

(iii) *[17, Lemma 2] For any $(a, b, \mu) \in \mathbb{R}^2 \times R_{++}$, we have*

$$\left\| \nabla^2 \phi(a, b, \mu) \right\|_2 \leq \frac{4}{\sqrt{(a-b)^2 + 4\mu^2}} \leq \frac{2}{\mu}.$$

3 A PREDICTOR-CORRECTOR ALGORITHM

The Algorithm: [3]

Step 0: (Initialization)

Choose $x^0 \in \mathbb{R}^n$, set $y^0 = Mx^0 + q$, and let $\mu_0 > 0$ be such that $\Phi(x^0, y^0, \mu_0) < 0$. Choose $\beta > 2n_u$ so that $\|\Phi(x^0, y^0, \mu_0)\| \le \beta\mu_0$. We now have $(x^0, y^0) \in \mathcal{N}(\beta, \mu_0)$. Choose $\bar{\sigma}$, α_1, and α_2 from $(0, 1)$.

Step 1: (The Predictor Step)

Let $(\Delta x^k, \Delta y^k, \Delta\mu_k)$ solve the equation

$$F(x^k, y^k, \mu_k) + \nabla F(x^k, y^k, \mu_k)^T \begin{bmatrix} \Delta x^k \\ \Delta y^k \\ \Delta\mu_k \end{bmatrix} = 0 . \tag{3.1}$$

If $\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, 0)\| = 0$, **STOP**, $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves $\mathrm{LCP}(q, M)$; else if

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \mu_k)\| > \beta\mu_k,$$

set

$$\hat{x}^k := x^k, \quad \hat{y}^k := y^k, \quad \hat{\mu}_k := \mu_k, \quad \text{and} \quad \eta_k = 1 ; \tag{3.2}$$

else let $\eta_k = \alpha_1^s$ where s is the positive integer such that

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \alpha_1^t \mu_k)\| \le \alpha_1^t \beta\mu_k, \tag{3.3}$$

for $t = 0, 1, \ldots, s$, and

$$\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \alpha_1^{s+1}\mu_k)\| > \alpha_1^{s+1}\beta\mu_k. \tag{3.4}$$

Set

$$\hat{x}^k := x^k + \Delta x^k, \quad \hat{y}^k := y^k + \Delta y^k, \quad \hat{\mu}_k := \eta_k \mu_k. \tag{3.5}$$

Step 2: (The Corrector Step)

Let $(\Delta\hat{x}^k, \Delta\hat{y}^k, \Delta\hat{\mu}_k)$ solve the equation

$$F(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \nabla F(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)^T \begin{bmatrix} \Delta\hat{x}^k \\ \Delta\hat{y}^k \\ \Delta\hat{\mu}_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ (1-\bar{\sigma})\hat{\mu}_k \end{bmatrix} \tag{3.6}$$

and let $\hat{\lambda}_k$ be the maximum of the value $1, \alpha_2, \alpha_2^2, \ldots$, such that

$$\left\|\Phi(\hat{x}^k + \hat{\lambda}_k \Delta\hat{x}^k, \hat{y}^k + \hat{\lambda}_k \Delta\hat{y}^k, (1 - \bar{\sigma}\hat{\lambda}_k)\hat{\mu}_k)\right\| \le (1 - \bar{\sigma}\hat{\lambda}_k)\beta\hat{\mu}_k. \tag{3.7}$$

Set

$$x^{k+1} = \hat{x}^k + \hat{\lambda}_k \Delta\hat{x}^k, \quad y^{k+1} = \hat{y}^k + \hat{\lambda}_k \Delta\hat{y}^k, \quad \mu_{k+1} = (1 - \bar{\sigma}\hat{\lambda}_k)\hat{\mu}_k, \tag{3.8}$$

and return to Step 1.

Remarks.

- In [3], the algorithm is stated using the 2–norm. It will be shown that the algorithm and its convergence analysis remain valid regardless of the choice of norm.

- Note that if the null step (3.2) is taken in Step 1, then the Newton equations (3.1) and (3.6) have the same coefficient matrix. Therefore only one matrix factorization is needed to implement both Steps 1 and 2 in this case. Otherwise, two matrix factorizations are needed.

- The algorithm can be modified so that only one matrix factorization is needed with the modified algorithm preserving the same nice convergence properties. The modification goes as follows. If $s = 0$ in the predictor step, then use the update

$$\hat{x}^k := x^k, \hat{y}^k := y^k, \hat{\mu}_k := \mu_k$$

instead of (3.5). On the other hand, if $s \geq 1$, use the update (3.5) and skip the corrector step. In other words, use the update

$$x^{k+1} = \hat{x}^k, y^{k+1} = \hat{y}^k, \mu_{k+1} = \hat{\mu}_k,$$

in the corrector step instead of (3.8).

- In the initialization step, setting

$$\mu_0 > \sqrt{\max_{\substack{i \in \{1,\dots,n\} \\ 0 < x_i^0,\ 0 < y_i^0}} x_i^0 y_i^0}$$

guarantees that the inequality $\Phi(x^0, y^0, \mu_0) < 0$ is satisfied. For example, one can choose $(x^0, y^0) = (0, q)$ in which case μ_0 can be taken to be any positive number.

- The condition that $\beta > 2n_u$ is only employed in the proof of local quadratic convergence. It is not required to verify the global linear convergence of the method. Theorem 2.1 motivates why this condition on β is required. Recall that when β is chosen in this way, the solution set is contained in the interior of the slices $\mathcal{N}(\beta, \mu)$ relative to the affine set Λ. Thus, eventually the Newton iterate associated with the predictor step remains in the interior of the current slice of the central path relative to the affine set Λ. Hence a full Newton step can be taken yielding the local quadratic convergence of the method.

- Observe that the function F has nonsingular Jacobian at a given point if and only if $\nabla_{(x,y)}F$ is nonsingular at that point. In [14, Theorem 3.5], it is shown that if $\mu > 0$ and the matrix M is a P_0 matrix, then $\nabla_{(x,y)}F(\bar{x}, \bar{y}, \mu)$ is nonsingular for all $(\bar{x}, \bar{y}) \in I\!R^{2n}$. Therefore, since the matrix M is assumed to be positive semi–definite and the algorithm is initiated with $\mu_0 > 0$, the Jacobians $\nabla F(x^k, y^k, \mu_k)$ and $\nabla F(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ are always nonsingular. Hence the Newton equations (3.1) and (3.6) yield unique solutions whenever (x^k, y^k, μ_k) and $(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ are well–defined. In addition, since $y^0 = Mx^0 + q$, we have $y^k = Mx^k + q$ and $\hat{y}^k = M\hat{x}^k + q$ for all well–defined iterates.

In analyzing the algorithm, it is helpful to take a closer look at the Newton equations (3.1) and (3.6). In (3.1) we have $\Delta\mu_k = -\mu_k$ and so (3.1) reduces to the system

$$\begin{aligned} M\Delta x^k - \Delta y^k &= 0 \\ \nabla_x \Phi(x^k, y^k, \mu_k)\Delta x^k + \nabla_y \Phi(x^k, y^k, \mu_k)\Delta y^k & \\ = -\Phi(x^k, y^k, \mu_k) + \mu_k \nabla_\mu \Phi(x^k, y^k, \mu_k). & \end{aligned} \tag{3.9}$$

Similarly, in (3.6), $\Delta\hat{\mu}_k = -\bar{\sigma}\hat{\mu}_k$ reducing (3.6) to the system

$$\begin{aligned} M\Delta \hat{x}^k - \Delta \hat{y}^k &= 0 \\ \nabla_x \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\Delta \hat{x}^k + \nabla_y \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\Delta \hat{y}^k & \\ = -\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \bar{\sigma}\hat{\mu}_k \nabla_\mu \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k). & \end{aligned} \tag{3.10}$$

Theorem 3.1 *Consider the algorithm described above and suppose that the matrix M is a P_0 matrix. If $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$ with $\mu_k > 0$, then either $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves LCP(q, M) or both $(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ and $(x^{k+1}, y^{k+1}, \mu_{k+1})$ are well–defined with the backtracking routines in Steps 1 and 2 finitely terminating. In the latter case, we have $(\hat{x}^k, \hat{y}^k) \in \mathcal{N}(\beta, \hat{\mu}_k)$ and $(x^{k+1}, y^{k+1}) \in \mathcal{N}(\beta, \mu_{k+1})$ with $0 < \mu_{k+1} < \hat{\mu}_k \le \mu_k$. Since $(x^0, y^0) \in \mathcal{N}(\beta, \mu_0)$ with $\mu_0 > 0$, this shows that the algorithm is well–defined.*

Proof. Let $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$ with $\mu_k > 0$. By the last remark given above, $(\Delta x^k, \Delta y^k, \Delta\mu_k)$ exists and is unique. Since $y^k + \Delta y^k = M(x^k + \Delta x^k) + q$, we have

$$\left\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, 0)\right\| = 0$$

if and only if $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves LCP(q, M). Therefore, if $(x^k + \Delta x^k, y^k + \Delta y^k)$ does not solve LCP(q, M), then by continuity, there exist $\epsilon > 0$ and $\bar{\mu} > 0$ such that $\left\|\Phi(x^k + \Delta x^k, y^k + \Delta y^k, \mu)\right\| > \epsilon$ for all $\mu \in [0, \bar{\mu}]$. In this case, the backtracking routine described in (3.3) and (3.4) of Step 1 is finitely terminating. Hence $(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$ is well–defined, with $0 < \hat{\mu}_k \le \mu_k$, and $(\Delta\hat{x}^k, \Delta\hat{y}^k, \Delta\hat{\mu}_k)$ is uniquely determined by (3.6). To see that the backtracking routine in Step 2 is finitely terminating, define $\theta(x, y, \mu) = \|\Phi(x, y, \mu)\|$. This is a convex composite function [1]. By (3.6)

$$\begin{aligned} &\theta'((\hat{x}^k, \hat{y}^k, \hat{\mu}_k); (\Delta\hat{x}^k, \Delta\hat{y}^k, \Delta\hat{\mu}_k)) \\ = \;& \inf_{\lambda > 0} \lambda^{-1} \left(\left\| \Phi(\hat{z}^k) + \lambda \nabla \Phi(\hat{z}^k)^T \begin{bmatrix} \Delta\hat{x}^k \\ \Delta\hat{y}^k \\ \Delta\hat{\mu}_k \end{bmatrix} \right\| - \left\|\Phi(\hat{z}^k)\right\| \right) \end{aligned}$$

$$\begin{aligned}
&\leq \left\| \Phi(\hat{z}^k) + \nabla\Phi(\hat{x}^k)^T \begin{bmatrix} \Delta\hat{x}^k \\ \Delta\hat{y}^k \\ \Delta\hat{\mu}_k \end{bmatrix} \right\| - \left\|\Phi(\hat{z}^k)\right\| \\
&= -\left\|\Phi(\hat{z}^k)\right\| \\
&< 0,
\end{aligned}$$

where $\hat{z}^k = (\hat{x}^k, \hat{y}^k, \hat{\mu}_k)$. Therefore, (3.7) can be viewed as an instance of a standard backtracking line search routine and as such is finitely terminating with $0 < \mu_{k+1} < \hat{\mu}_k \leq \mu_k$ (indeed, one can replace the value of $\bar{\sigma}$ on the right hand side of (3.7) by any number in the open interval $(0,1)$).

Since $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$, the argument given above implies that either $(x^k + \Delta x^k, y^k + \Delta y^k)$ solves LCP(q, M) or $\hat{y}^k = M\hat{x}^k + q$ with $\left\|\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)\right\| \leq \beta\hat{\mu}_k$ and $y^{k+1} = Mx^{k+1} + q$ with $\left\|\Phi(x^{k+1}, y^{k+1}, \mu_{k+1})\right\| \leq \beta\mu_{k+1}$. Thus, if $(x^k + \Delta x^k, y^k + \Delta y^k)$ does not solve LCP(q, M), we need only show that $\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0$ and $\Phi(x^{k+1}, y^{k+1}, \mu_{k+1}) \leq 0$ in order to have $(\hat{x}^k, \hat{y}^k) \in \mathcal{N}(\beta, \hat{\mu}_k)$ and $(x^{k+1}, y^{k+1}) \in \mathcal{N}(\beta, \mu_{k+1})$. First note that the component-wise concavity of Φ implies that for any $(x, y, \mu) \in I\!R^{2n+1}$ with $\mu > 0$, and $(\Delta x, \Delta y, \Delta\mu) \in I\!R^{2n+1}$ one has

$$\Phi(x + \Delta x, y + \Delta y, \mu + \Delta\mu) \leq \Phi(x, y, \mu) + \nabla\Phi(x, y, \mu)^T \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\mu \end{bmatrix} .$$

Hence, in the case of the predictor step, either (3.2) holds or

$$\begin{aligned}
&\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \\
&= \Phi(x^k + \Delta x^k, y^k + \Delta y^k, \eta_k\mu_k) \\
&\leq \Phi(x^k, y^k, \mu_k) + \nabla\Phi(x^k, y^k, \mu_k)^T \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ (\eta_k - 1)\mu_k \end{pmatrix} \\
&= \Phi(x^k, y^k, \mu_k) + \nabla\Phi(x^k, y^k, \mu_k)^T \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ -\mu_k \end{pmatrix} + \eta_k\mu_k\nabla_\mu\Phi(x^k, y^k, \mu_k) \\
&= \eta_k\mu_k\nabla_\mu\Phi(x^k, y^k, \mu_k) \leq 0.
\end{aligned}$$

In either case, $\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0$. For the corrector step, we have

$$\begin{aligned}
&\Phi(x^{k+1}, y^{k+1}, \mu_{k+1}) \\
&= \Phi(\hat{x}^k + \hat{\lambda}_k\Delta\hat{x}^k, \hat{y}^k + \hat{\lambda}_k\Delta\hat{y}^k, \hat{\mu}_k + \hat{\lambda}_k\Delta\hat{\mu}_k) \\
&\leq \Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) + \hat{\lambda}_k\nabla\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k)^T \begin{bmatrix} \Delta\hat{x}^k \\ \Delta\hat{y}^k \\ -\bar{\sigma}\hat{\mu}_k \end{bmatrix} \\
&= (1 - \hat{\lambda}_k)\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0,
\end{aligned}$$

since we have already shown that $\Phi(\hat{x}^k, \hat{y}^k, \hat{\mu}_k) \leq 0$. This completes the proof.

■

4 CONVERGENCE

In this section, we require the following hypothesis in order to establish the global linear convergence of the method.

Hypothesis (B): Given $\beta > 0$ and $\mu_0 > 0$, there exists a $C > 0$ such that

$$\left\|\nabla_{(x,y)} F(\bar{x}, \bar{y}, \mu)^{-1}\right\|_\infty \leq C, \tag{4.1}$$

for all $0 < \mu \leq \mu_0$ and $(\bar{x}, \bar{y}) \in \mathcal{N}(\beta, \mu)$.

In [2, Proposition 4.3], we show that a bound of this type exits under the assumption of a non–degeneracy condition due to Fukushima, Luo, and Pang [11, Assumption (A2)]. Similar results of this type have since been obtained by Chen and Xiu [7, Section 6], Tseng [20, Corollary 2], and Qi and Sun [17, Proposition 2].

In private discussions, Kanzow [15] points out that the Fukushima, Luo, and Pang non–degeneracy condition implies the uniqueness of the solution to LCP(q, M). Kanzow's proof easily extends to show that *any* condition which implies Assumption (B) also implies the uniqueness of the solution to LCP(q, M).

Proposition 4.1 *[3, Proposition 4.1] If Assumptions (A) and (B) hold, then LCP(q, M) has a unique solution.*

Theorem 4.1 *(Global Linear Convergence) Suppose that M is a P_0 matrix and hypotheses (A) and (B) hold. Let $\{(x^k, y^k, \mu_k)\}$ be the sequence generated by the algorithm. If the algorithm does not terminate finitely at the unique solution to LCP(q, M), then for $k = 0, 1, \ldots,$*

$$(x^k, y^k) \in \mathcal{N}(\beta, \mu_k) \tag{4.2}$$

$$(1 - \bar{\sigma}\hat{\lambda}_{k-1})\eta_{k-1} \cdots (1 - \bar{\sigma}\hat{\lambda}_0)\eta_0\mu_0 \;=\; \mu_k, \tag{4.3}$$

with

$$\hat{\lambda}_k \geq \bar{\lambda} := \min\left\{1, \frac{\alpha_2(1-\bar{\sigma})\beta}{n_u(2C^2(\frac{\beta}{n_l} + 2\bar{\sigma})^2 + \bar{\sigma}^2) + \bar{\sigma}(1-\bar{\sigma})\beta}\right\}, \tag{4.4}$$

where C is the constant defined in (4.1). Therefore μ_k converges to 0 at a global linear rate. In addition, the sequence $\{(x^k, y^k)\}$ converges to the unique solution of LCP(q, M).

Proof. The inclusion (4.2) has already been established in Theorem 3.1 and the relation (4.3) follows by construction.

For the sake of simplicity, set $(x, y, \mu) \doteq (\hat{x}_k, \hat{y}_k, \hat{\mu}_k)$ and $(\Delta x, \Delta y) = (\Delta\hat{x}^k, \Delta\hat{y}^k)$. Then for $i \in \{1, \ldots, n\}$ and $\lambda \in [0, 1]$, Lemma 2.3 and (3.6) imply that

$$|\phi(x_i + \lambda\Delta x_i, y_i + \lambda\Delta y_i, (1 - \bar{\sigma}\lambda)\mu)|$$

$$
\begin{aligned}
&= \Bigg| \phi(x_i, y_i, \mu) + \lambda \nabla \phi(x_i, y_i, \mu)^T \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix} + \\
&\quad \frac{\lambda^2}{2} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix}^T \nabla^2 \phi(x_i + \theta_i \lambda \Delta x_i, y_i + \theta_i \lambda \Delta y_i, (1 - \theta_i \bar{\sigma} \lambda)\mu) \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix} \Bigg| \\
&\le (1 - \lambda) |\phi(x_i^k, y_i^k, \mu)| + \\
&\quad \frac{\lambda^2}{2} \left\| \nabla^2 \phi(x_i + \theta_i \lambda \Delta x_i, y_i + \theta_i \lambda \Delta y_i, (1 - \theta_i \bar{\sigma} \lambda)\mu) \right\|_2 \left\| \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ -\bar{\sigma}\mu \end{pmatrix} \right\|_2^2 \\
&\le (1 - \lambda) |\phi(x_i, y_i, \mu)| + \frac{\lambda^2}{(1 - \bar{\sigma}\lambda)\mu} \|(\Delta x_i, \Delta y_i, -\bar{\sigma}\mu)\|_2^2
\end{aligned}
$$

for some $\theta_i \in [0, 1]$. Set $t_i := \|(\Delta x_i, \Delta y_i, -\bar{\sigma}\mu_k)\|_2^2$ for $i = 1, \ldots, n$, then

$$
\begin{aligned}
&\|\Phi(x + \lambda \Delta x, y + \lambda \Delta y, (1 - \bar{\sigma}\lambda)\mu)\| \\
\le\ & (1 - \lambda) \|\Phi(x, y, \mu)\| + \frac{\lambda^2}{(1 - \bar{\sigma}\lambda)\mu} \left\| \begin{pmatrix} t_1 \\ \cdots \\ t_n \end{pmatrix} \right\| \\
\le\ & (1 - \lambda) \|\Phi(x, y, \mu)\| + \frac{\lambda^2 n_u}{(1 - \bar{\sigma}\lambda)\mu} \left\| \begin{pmatrix} t_1 \\ \cdots \\ t_n \end{pmatrix} \right\|_\infty \\
\le\ & (1 - \lambda) \|\Phi(x, y, \mu)\| + \frac{\lambda^2 n_u}{(1 - \bar{\sigma}\lambda)\mu} \left(2 \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_\infty^2 + \bar{\sigma}^2 \mu^2 \right) \\
\le\ & (1 - \lambda)\beta\mu + \frac{\lambda^2 n_u}{1 - \bar{\sigma}\lambda} \left(2C^2 (\frac{\beta}{n_l} + 2\bar{\sigma})^2 + \bar{\sigma}^2 \right) \mu, \qquad (4.5)
\end{aligned}
$$

where the last inequality follows from (3.9) and the bound

$$
\begin{aligned}
&\left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_\infty \\
\le\ & \left\| \nabla_{(x,y)} F^{-1}(x, y, \mu) \right\|_\infty \left(\|\Phi(x, y, \mu)\|_\infty + \bar{\sigma}\mu \|\nabla_\mu \Phi(x, y, \mu)\|_\infty \right) \\
\le\ & C(\frac{\beta}{n_l} + 2\bar{\sigma})\mu. \qquad (4.6)
\end{aligned}
$$

It is easily verified that

$$
(1 - \lambda)\beta\mu + \frac{\lambda^2 n_u}{1 - \bar{\sigma}\lambda} \left(2C^2 (\frac{\beta}{n_l} + 2\bar{\sigma})^2 + \bar{\sigma}^2 \right) \mu \le (1 - \bar{\sigma}\lambda)\beta\mu,
$$

whenever

$$
\lambda \le \frac{(1 - \bar{\sigma})\beta}{n_u (2C^2 (\frac{\beta}{n_l} + 2\bar{\sigma})^2 + \bar{\sigma}^2) + \bar{\sigma}(1 - \bar{\sigma})\beta}.
$$

Therefore

$$\hat{\lambda}_k \geq \min\left\{1, \frac{\alpha_2(1-\bar{\sigma})\beta}{n_u(2C^2(\frac{\beta}{n_l}+2\bar{\sigma})^2+\bar{\sigma}^2)+\bar{\sigma}(1-\bar{\sigma})\beta}\right\}.$$

To conclude, note that the sequence $\{(x^k, y^k)\}$ is bounded by hypotheses (A) and Theorem 3.1. In addition, just as in (4.6), the relations (3.9) and (3.10) yield the bounds

$$\left\|\begin{pmatrix} \Delta x^k \\ \Delta y^k \end{pmatrix}\right\|_\infty \leq C(\frac{\beta}{n_l}+2)\mu_k \quad \text{and} \quad \left\|\begin{pmatrix} \Delta \hat{x}^k \\ \Delta \hat{y}^k \end{pmatrix}\right\| \leq C(\frac{\beta}{n_l}+2)\mu_k ,$$

since $0 < \bar{\sigma} < 1$ and $0 < \eta_k \leq 1$ for all k. Therefore, (4.3) and (4.4) imply that

$$\begin{aligned}\left\|(x^{k+1}, y^{k+1}) - (x^k, y^k)\right\|_\infty &\leq \left\|\begin{pmatrix} \Delta x^k \\ \Delta y^k \end{pmatrix}\right\|_\infty + \hat{\lambda}_k \left\|\begin{pmatrix} \Delta \hat{x}^k \\ \Delta \hat{y}^k \end{pmatrix}\right\|_\infty \\ &\leq 2C(\frac{\beta}{n_l}+2)\mu_k \leq 2C(\frac{\beta}{n_l}+2)(1-\bar{\sigma}\bar{\lambda})^k\mu_0 .\end{aligned}$$

Hence, $\{(x^k, y^k)\}$ is a Cauchy sequence and so must converge to the unique solution of LCP(q, M). ∎

Theorem 4.2 *(Local Quadratic Convergence) Suppose M is a P_0 matrix, assumption (B) holds and that the sequence $\{(x^k, y^k, \mu_k)\}$ generated by the algorithm converges to $\{(x^*, y^*, 0)\}$ where $\{(x^*, y^*)\}$ is the unique solution to LCP(q, M). If it is further assumed that the strict complementary slackness condition $0 < x^* + y^*$ is satisfied, then*

$$\mu_{k+1} = O(\mu_k^2), \tag{4.7}$$

that is, μ_k converges quadratically to zero.

Proof. First observe that due to the strict complementarity of $\{(x^*, y^*)\}$, Part (iii) of Lemma 2.3 indicates that there exist constants $\epsilon > 0$ and $L > 0$ such that

$$\left\|\nabla^2\phi(x, y, \mu)\right\|_2 \leq L, \quad \text{whenever} \quad \|(x, y, \mu) - (x^*, y^*, 0)\| \leq \epsilon. \tag{4.8}$$

Hence, for all k sufficient large and $\eta \in (0, 1]$, we have for each $i \in \{1, \ldots, n\}$ that

$$\begin{aligned}&|\phi(x_i^k + \Delta x_i^k, y_i^k + \Delta y_i^k, \eta\mu_k)| \\ = &\left|\phi(x_i^k, y_i^k, \mu_k) + \nabla^T\phi(x_i^k, y_i^k, \mu_k)\begin{pmatrix} \Delta x_i^k \\ \Delta y_i^k \\ (\eta-1)\mu_k \end{pmatrix}\right| +\end{aligned}$$

$$\begin{aligned}
&\frac{1}{2}\left|\begin{pmatrix}\Delta x_i^k\\ \Delta y_i^k\\ (\eta-1)\mu_k\end{pmatrix}^T \nabla^2\phi(z_i^k)\begin{pmatrix}\Delta x_i^k\\ \Delta y_i^k\\ (\eta-1)\mu_k\end{pmatrix}\right| \\
= &\left|\phi(x_i^k,y_i^k,\mu_k)+\nabla^T\phi(x_i^k,y_i^k,\mu_k)\begin{pmatrix}\Delta x_i^k\\ \Delta y_i^k\\ -\mu_k\end{pmatrix}+\eta\mu_k\nabla_\mu\phi(x^k,y^k,\mu_k)\right|+ \\
&\frac{1}{2}\left|\begin{pmatrix}\Delta x_i^k\\ \Delta y_i^k\\ (\eta-1)\mu_k\end{pmatrix}^T \nabla^2\phi(z_i^k)\begin{pmatrix}\Delta x_i^k\\ \Delta y_i^k\\ (\eta-1)\mu_k\end{pmatrix}\right| \\
\leq\ &\eta\mu_k|\nabla_\mu\phi(x^k,y^k,\mu_k)|+\frac{L}{2}\left\|(\Delta x_i^k,\Delta y_i^k,(\eta-1)\mu_k)\right\|_2^2 \\
=\ &2\eta\mu_k+\frac{L}{2}\left(\left\|(\Delta x_i^k,\Delta y_i^k)\right\|_2^2+(1-\eta)^2\mu_k^2\right)
\end{aligned}$$

where $z_i^k=(x_i^k+\theta_i\Delta x_i^k, y_i+\theta_i\Delta y_i^k,(1+\theta_i(\eta-1))\mu_k)$. Now using an argument similar to that used to obtain (4.5), we have

$$\begin{aligned}
&\left\|\Phi(x^k+\Delta x^k,y^k+\Delta y^k,\eta\mu_k)\right\| \\
\leq\ &n_u\left\|\Phi(x^k+\Delta x^k,y^k+\Delta y^k,\eta\mu_k)\right\|_\infty \\
\leq\ &n_u\left(2\eta\mu_k+\frac{L}{2}\left(2C^2(\frac{\beta}{n_l}+2)^2+1\right)\mu_k^2\right) \\
\leq\ &2n_u\eta\mu_k+n_uL\left(C^2(\frac{\beta}{n_l}+2)^2+\frac{1}{2}\right)\mu_k^2. \qquad (4.9)
\end{aligned}$$

Hence, since $\beta>2n_u$, the inequality (3.3) in Step 1 of the algorithm holds with $t=0$ for all k sufficiently large. It is easy to verify that

$$2n_u\eta\mu_k+n_uL\left(C^2(\frac{\beta}{n_l}+2)^2+\frac{1}{2}\right)\mu_k^2\leq\eta\beta\mu_k, \qquad (4.10)$$

whenever

$$\eta\geq\frac{n_uL\left(C^2(\frac{\beta}{n_l}+2)^2+\frac{1}{2}\right)\mu_k}{\beta-2n_u}.$$

Hence, by (3.4), we have

$$\alpha_1\eta_k\leq\frac{n_uL\left(C^2(\frac{\beta}{n_l}+2)^2+\frac{1}{2}\right)\mu_k}{\beta-2n_u},$$

and so

$$\eta_k\leq\frac{n_uL\left(C^2(\frac{\beta}{n_l}+2)^2+\frac{1}{2}\right)\mu_k}{\alpha_1(\beta-2n_u)} \qquad (4.11)$$

for all k sufficient large. Therefore, by (3.5),

$$\mu_{k+1}=O(\mu_k^2).$$

■

We now show that under relatively mild condition, the algorithm is globally convergent.

Theorem 4.3 *Suppose that M is a P_0 matrix and hypothesis (A) holds. Let $\{x^k, y^k, \mu_k\}$ be a sequence generated by the Algorithm. Then*

(i) *the sequence $\{\mu_k\}$ is monotonically decreasing and convergent to 0 as $k \to \infty r$,*

(ii) *the sequence $\{(x^k, y^k)\}$ is bounded and every accumulation point of $\{(x^k, y^k)\}$ is a solution to LCP(q, M).*

Proof. Since M is a P_0 matrix, the algorithm is well defined. By the construction of the algorithm, we can see that $\mu_{k+1} < \mu_k$ for $(k = 0, 1, \ldots)$. Hence the sequence $\{\mu_k\}$ is monotonically decreasing. Since $\mu_k \geq 0$ $(k = 0, 1, \ldots)$, there is a $\bar{\mu} \geq 0$ such that $\mu_k \to \bar{\mu}$. Since $(x^k, y^k) \in \mathcal{N}(\beta, \mu_k)$, the sequence (x^k, y^k) is bounded, by taking a subsequence if necessary, we may assume that $\{(x^k, y^k)\}$ converges to some point $(\bar{x}, \bar{y})$. If $\bar{\mu} = 0$, it follows from $\{(x^k, y^k)\} \in \mathcal{N}(\beta, \mu_k)$ that $(\bar{x}, \bar{y})$ is a solution of the LCP and we obtain the desired results. Suppose that $\bar{\mu} > 0$. Since $\nabla_{(x,y)} F(\bar{x}, \bar{y}, \bar{\mu})$ is nonsingular, there exist $\epsilon > 0$, $L > 0$ and $C > 0$ such that

$$\left\|\nabla_{(x,y)} F(x, y, \mu)^{-1}\right\|_\infty \leq C, \tag{4.12}$$

$$\left\|\nabla^2 \phi(x_i, y_i, \mu)\right\|_2 \leq L, \tag{4.13}$$

for all $(x, y, \mu) \in \mathcal{O}_{(\bar{x}, \bar{y}, \bar{\mu})} = \{(x, y, \mu) : \|(x, y, \mu) - (\bar{x}, \bar{y}, \bar{\mu})\| \leq \epsilon\}$. Similar to the proof of the global linear convergence result, we can show that there exists a $\bar{\lambda}$ such that $\hat{\lambda}_k \geq \bar{\lambda}$ for sufficient large k. Therefore, for sufficient large k, $\mu_{k+1} \leq c\mu_k$ for some constant $c \in (0, 1)$, which yields a contradiction. ■

References

[1] J. Burke, *Descent methods for composite nondifferentiable optimization problems*, Mathematical Programming, 33 (1987), pp. 260—279.

[2] J. Burke and S. Xu, *The global linear convergence of a non-interior path-following algorithm for linear complementarity problem.* To appear in Mathematics of Operations Research.

[3] ——, *A non–interior predictor–corrector path following algorithm for the monotone linear complementarity problem.* Preprint, Department of Mathematics, University of Washington , Seattle, WA 98195, December, 1997.

[4] B. Chen and X. Chen, *A global and local super–linear continuation method for $P_0 + R_0$ and monotone NCP.* To appear in SIAM Journal on Optimization.

[5] ——, *A global linear and local quadratic continuation method for variational inequalities with box constraints.* Preprint, Department of Management and Systems, Washington State University, Pullman, WA 99164-4736, March, 1997.

[6] B. CHEN AND P. HARKER, *A non–interior–point continuation method for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168—1190.

[7] B. CHEN AND N. XIU, *A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing functions.* To appear in SIAM Journal on Optimization.

[8] C. CHEN AND O. L. MANGASARIAN, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Comp. Optim. and Appl. , 5 (1996), pp. 97—138.

[9] X. CHEN, L. QI, AND D. SUN, *Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities*, Mathematics of Computation, 67 (1998), pp. 519-540.

[10] X. CHEN AND Y. YE, *On homotopy smooothing methods for variational inequalities.* To appear in SIAM J. Control Optim.

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

[12] K. HOTTA AND A. YOSHISE, *Global convergence of a class of non–interior–point algorithms using Chen–Harker–Kanzow functions for nonlinear complementarity problems.* Discussion Paper Series, No. 708, University of Tsukuba, Tsukuba, Ibaraki 305, Japan, December, 1996.

[13] H. JIANG, *Smoothed Fischer–Burmeister equation methods for the complementarity problem.* Report, Department of Mathematics, University of Melbourne, Parkville, Australia, June, 1997.

[14] C. KANZOW, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851—868.

[15] ——, *Private communications.* Seattle, Washington, May, 1997.

[16] M. KOJIMA, N. MEGGIDO, T. NOMA, AND A. YOSHISE, *A unified approach to interior point algorithms for linear complementarity problems*, Springer–Verlag, Berlin, 1991.

[17] L. QI AND D. SUN, *Improving the convergence of non–interior point algorithms for nonlinear complementarity problems.* To appear in Mathematics of Computation.

[18] L. QI, D. SUN, AND G. ZHOU, *A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational*

inequalities. Applied Mathematics Report, AMR 97/13, School of Mathematics University of New South Wales, Sydney 2052, Australia, June, 1997.

[19] S. SMALE, *Algorithms for solving equations.* Proceedings of the International Congress of Mathematicians, Berkeley, California, 1986.

[20] P. TSENG, *Analysis of a non–interior continuation method based on Chen–Mangasarian smoothing functions for complementarity problems*, Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi, eds., (Kluwer Academic Publisher, Nowell, MA. USA, 1998), 381-404.

[21] S. XU, *The global linear convergence of an infeasible non–interior path–folowing algorithm for complementarity problems with uniform P–functions.* Technique Report, Department of Mathematics, University of Washington, Seattle, WA 98195, December, 1996.

[22] ——, *The global linear convergence and complexity of a non-interior path-following algorithm for monotone LCP based on Chen-Harker-Kanzow-Smale smoothing function.* Technique Report, Department of Mathematics, University of Washington, Seattle, WA 98195, February, 1997.