The Global Linear Convergence of a Non-Interior Path-Following Algorithm for Linear Complementarity Problems

James V. Burke and Song Xu
Department of Mathematics
University of Washington
Seattle, WA 98195

Submitted to Mathematics of Operations Research
September, 1996

Revised, July 1997

Abstract

A non–interior path following algorithm is proposed for the linear complementarity problem. The method employs smoothing techniques introduced by Kanzow. If the LCP is P_0+R_0 and satisfies a non–degeneracy condition due to Fukushima, Luo, and Pang, then the algorithm is globally linearly convergent. As with interior point path following methods, the convergence theory relies on the notion of a neighborhood for the central path. However, the choice of neighborhood differs significantly from that which appears in the interior point literature. Numerical experiments are presented that illustrate the significance of the neighborhood concept for this class of methods.

1 Introduction

In this paper, we develop a non–interior path following method for the linear complementarity problem:

LCP(q, M): Find $(x^*, y^*) \in I\!R^n \times I\!R^n$ satisfying

$$Mx^* - y^* + q = 0, \tag{1.1}$$
$$x^* \geq 0, y^* \geq 0, (x^*)^T y^* = 0, \tag{1.2}$$

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

The global linear convergence of the method is established under a non–degeneracy assumption due to Fukushima, Luo, and Pang (see Definition 4.1 or [6, Assumption (A2)]) and the assumption that the matrix M is $P_0 + R_0$ (see Definition 2.2). The method is based on an algorithm proposed by Kanzow in [12]. It is also closely related to the algorithm proposed by Chen and Harker in [3]. The key difference with the Kanzow and Chen–Harker algorithms is our use of a new notion of neighborhood for the central path to update the continuation parameter.

Following standard usage in the interior point literature (for example see [13]), the *central path* is the set

$$\mathcal{C} := \{(x, y) : Mx - y + q = 0,\ Xy = \mu e \text{ with } 0 < x,\ 0 < y,\ \text{and } 0 < \mu\},$$

where e is the vectors of ones and $X = \text{diag}\,(x)$ the diagonal matrix with ith diagonal entry x_i for $i = 1, 2, \ldots, n$. For each $\beta \in (0, 1)$, the β neighborhood of $\mathcal{C}$ is defined to be the set

$$\mathcal{N}_s(\beta) := \left\{ (x, y) : Mx - y + q = 0,\ \frac{\|Xy - \mu e\|}{\mu} \leq \beta, \text{ with } 0 < x,\ 0 < y,\ \text{and } 0 < \mu \right\} .$$

Here, the subscript s is used to indicate that this is the *standard* neighborhood employed in the interior point literature for LCP(q, M). Very loosely speaking, standard interior point methods start with an initial point (x^0, y^0) lying in $\mathcal{N}_s(\beta)$ for some $\beta > 0$ and an initial value μ_0 satisfying $\|X^0 y^0 - \mu_0 e\| \leq \mu_0 \beta$. An update is obtained by choosing $\mu_1 < \mu_0$ and computing a Newton step $(\Delta x, \Delta y)$ at (x^0, y^0) based on the equations

$$Mx - y + q = 0,\ Xy = \mu_1 e. \tag{1.3}$$

The Newton step is then damped to ensure that the update (x^1, y^1) remains strictly positive and lies in $\mathcal{N}_s(\beta)$. This process is then iterated to termination. The trick is to implement the method so to ensure the existence of a sequence μ_k converging linearly to zero and satisfying $\left\|X^k y^k - \mu_k e\right\| \leq \mu_k \beta$. This yields the linear convergence of the vector $X^k y^k$ to zero which in turn provides the basis for complexity results. Numerous variations on this basic plan have been proposed. The most notable of which are the infeasible (i.e. $Mx^k - y^k + q$ may or may not be the zero vector) predictor–corrector strategies for which local super–linear convergence can also be established (e.g. see [17, 21]).

Kanzow's non–interior method for solving LCP(q, M) is based on the function

$$F_{\psi_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Psi_\mu(x, y) \end{bmatrix} , \tag{1.4}$$

where

$$\Psi_\mu(x, y) = \begin{bmatrix} \psi_\mu(x_1, y_1) \\ \ldots \\ \psi_\mu(x_n, y_n) \end{bmatrix}, \tag{1.5}$$

and

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

For $\mu > 0$, it is easy to show [12] that $\psi_\mu(a,b) = 0$ if and only if $0 < a,\ 0 < b$, and $ab = \mu$. Consequently, $F_{\psi_\mu}(x,y) = 0$ if and only if $(x,y) \in \mathcal{C}$. The key distinction between the system (1.3) and the system

$$F_{\psi_\mu}(x,y) = 0 \tag{1.7}$$

is that a solution to the system (1.3) may not be strictly positive and so may not lie on the central path. This partially explains why one must initiate interior point methods at strictly positive points and then damp the Newton steps to maintain this property. On the other hand, a solution to (1.7) must be strictly positive and so will lie on the central path. Thus, the non–negativity of any limit point is automatically assured without imposing additional non-negativity constraints. This is one of the reasons why the function F_{ψ_μ} is so effective in formulating non-interior path following methods.

In Kanzow's method an initial point (x^0, y^0) satisfying $Mx^0 - y^0 + q = 0$ and an initial value for μ_0 are chosen. It is not required that (x^0, y^0) be strictly positive or that it lie in some neighborhood of the central path. A global Newton strategy is then applied to the system $F_{\psi_{\mu_0}}(x,y) = 0$ until the value of the norm of $F_{\psi_{\mu_0}}(x,y)$ has been sufficiently reduced. The value of μ_0 is then updated to $\mu_1 < \mu_0$ and the process is iterated to termination. Kanzow shows in [12, Theorem 6.2] that if the matrix M is $P_0 + R_0$, then the sequence generated by his algorithm [12, Algorithm 6.1] has an accumulation point and that every such accumulation point is a solution to LCP(q, M). A similar algorithm is proposed by Chen and Harker in [3]. There the authors concentrate on establishing the existence of the central path and the continuity of the path at $\mu = 0$, however, no algorithmic convergence results are provided. The absence of any rate of convergence results for these algorithms is due to the somewhat ad hoc rules for updating the continuation parameter μ. This gap is bridged in interior point methods by requiring that the iterates remain in a pre–specified neighborhood of the central path.

In [20], Xu and Burke consider an interior point variation on the Chen–Harker–Kanzow path following techniques. Their algorithm does not require the feasibility of the affine constraint at each iteration, but it does require that the iterates remain strictly positive and stay in a given neighborhood $\mathcal{N}_s(\beta)$ of the central path. The positivity restriction allows the introduction of a rescaled Newton step producing iterates whose distance from the central path is easily controlled. The convergence behavior of the iterates can then be examined using standard interior point methodology. Xu and Burke establish the global linear convergence of their algorithm and use this result to establish the polynomial complexity of the method.

The complexity result in [20] is interesting since it provides some theoretical justification for the superb numerical performance demonstrated by the Kanzow and Chen–Harker algorithms. However, enforcing positivity in the system (1.7) is redundant since, unlike the system (1.3), this system automatically guarantees the positivity of its solutions. In this paper we describe a non–interior path following algorithm based on the function F_{ψ_μ} and establish its global linear convergence. The proof technique follows the pattern developed for interior point strategies. In this regard, the key is the introduction of a new notion of neighborhood for the central path that is better suited to the function F_{ψ_μ}. Just as with interior point methods, this neighborhood is used to adjust the value of the continuation parameter μ between iterations in a manner that insures the linear convergence of the values

$\left\|\Psi_{\mu_k}(x^k, y^k)\right\|$ to zero. Preliminary numerical experiments indicate that the algorithm is very promising.

The plan of the paper is as follows. In Section 2, we introduce a new notion of neighborhood for the central path and establish some of the properties of this neighborhood. The algorithm is stated and shown to be well defined in Section 3. The global linear convergence result is given in Section 4 and the numerical experiments are discussed in Section 5. A few concluding remarks are given in Section 6.

The notation we employ is standard. The notation borrowed from the interior–point literature has been discussed above. Vectors in $\mathbb{R}^n$ are assumed to be column vectors and unless otherwise stated the norm is the Euclidean norm. On one occasion we make use of the supremum norm. It is denoted by

$$\|x\|_\infty = \sup\{|x_i| \,:\, i = 1, 2, \ldots, n\} \,.$$

2 A Neighborhood of the Central Path

We take as our neighborhood of the central path the set

$$\mathcal{N}(\beta) := \left\{ (x, y) \,:\, Mx - y + q = 0,\ \frac{\|\Psi_\mu(x, y)\|^2}{\mu} \le \beta, \text{ with } 0 < \mu \right\} \,,$$

for $\beta > 0$. The square of the norm in this definition is used to ensure that the expression $\|\Psi_\mu(x, y)\|^2$ has linear growth with respect to μ. In the recent article [4], the authors suggest replacing μ by μ^2 in the definition of ψ_μ. With this substitution, the expression $\|\Psi_\mu(x, y)\|$ grows linearly with respect to μ and so one does not need to square the norm. On the other hand, by using the square of the norm, we can directly apply results due to Geiger and Kanzow ([11, Lemma 2.1] and [7, Lemmas 3.1 and 4.3]) showing that $\psi_0^2(a, b)$ is continuously differentiable on $\mathbb{R}^2$ and is twice continuously differentiable with uniformly bounded Hessian on $\mathbb{R}^2\backslash\{(0, 0)\}$. We extend these facts to ψ_μ with $\mu > 0$ and, in addition, show that $\psi_\mu^2(a, b)$ is globally Lipschitz in the continuation parameter $\mu > 0$. These properties allow us to mimic the interior point proof strategy in the non-interior setting.

Lemma 2.1 *The function ψ_μ defined in (1.6) has the following properties:*

1. *For every $\mu \ge 0$, the function ψ_μ^2 is continuously differentiable on $\mathbb{R}^2$.*
2. *One has*
$$\|\nabla^2(\psi_\mu^2(a, b))\| \le 4(5 + \sqrt{2}), \tag{2.8}$$
for all $(a, b) \in \mathbb{R}^2$ when $\mu > 0$ and for all $(a, b) \in \mathbb{R}^2\backslash\{(0, 0)\}$ when $\mu = 0$.
3. *For $\mu_1 \ge 0$, $\mu_2 \ge 0$ and $a, b \in \mathbb{R}$, we have*
$$|\psi_{\mu_1}^2(a, b) - \psi_{\mu_2}^2(a, b)| \le (2 + 2\sqrt{2})|\mu_1 - \mu_2|. \tag{2.9}$$

Proof The only troublesome case in Part 1 occurs at the points $(a,b) = (0,0)$ when $\mu = 0$. In this case one simply applies the inequality

$$|a| + |b| \leq \sqrt{2}\sqrt{a^2 + b^2} \tag{2.10}$$

to obtain a suitable bound on $\nabla\psi_\mu{}^2$ in the vicinity of the origin. This allows one to show that the limiting value of $\left\|\nabla\psi_\mu{}^2(a,b)\right\|$ is zero as a, b, and μ converge to zero. The proof of Part 2 follows directly from the pattern of proof provided in [7, Lemma 4.3] for the case $\mu = 0$. Part 3 also follows in a straightforward manner with the help of (2.10) and so we leave its proof as an exercise for the reader. □

We now recall some well established conditions that yield the existence of the central path. These conditions are also used to guarantee certain boundedness conditions for our neighborhood of the central path. We begin with a review of the necessary terminology.

Definition 2.2 *Let $M \in I\!R^{n\times n}$.*

(a) M is a P_0 matrix if each of its principal minors is non-negative.

(b) M is a P matrix if each of its principal minors is positive.

(c) M is an R_0 matrix if LCP$(0, M)$ has unique solution $(x, y) = (0, 0)$.

(d) M is said to be $P_0 + R_0$ if it is both a P_0 and an R_0 matrix.

(d) M is a non–degenerate matrix if each of its principal submatrices is nonsingular.

The set of P_0 matrices clearly contains the set of all positive semi–definite matrices. The positive semi–definite matrices give rise to the *monotone* linear complementarity problems of which both linear and convex quadratic programming are special cases. Every positive definite matrix is a P matrix, and a P matrix is a non–degenerate matrix that is both a P_0 and an R_0 matrix. Under the assumption that the matrix M is an R_0 matrix it is well–known that the solution set

$$S = \{(x,y) : 0 \leq x,\ 0 \leq y,\ Mx - y + q = 0, \text{ and } x^T y = 0\}. \tag{2.11}$$

to LCP(q, M) is bounded. The boundedness of S is key to the analysis of the limiting behavior of the central path as the continuation parameter μ tends to zero. This limiting behavior and the existence of the central path is addressed in the following theorem.

Theorem 2.3 *([12, Corollary 3.9]) If M is a P_0 and an R_0 matrix, then the equation $F_{\psi_\mu}(x, y) = 0$ has a unique solution $(x(\mu), y(\mu))$ for all $\mu > 0$. Moreover, the entire sequence $(x(\mu), y(\mu))$ converges to a solution of LCP(q, M) as μ tends to 0.*

Just as the assumptions in Theorem 2.3 can be used to establish the boundedness of the solution set S, these assumptions can also be used to assure the boundedness of certain *slices* of the neighborhood $\mathcal{N}(\beta)$.

Proposition 2.4 *Let $\beta > 0$, $\mu_0 > 0$ and consider the sets*

$$\mathcal{N}(\mu_0, \beta) := \left\{ (x, y) \,:\, Mx - y + q = 0,\ \frac{\|\Psi_\mu(x, y)\|^2}{\mu} \leq \beta,\ \textit{with } 0 < \mu \leq \mu_0 \right\} .$$

If M is an R_0 matrix, then the set $\mathcal{N}(\mu_0, \beta)$ is bounded.

Proof The pattern of proof is identical to that which is used to show the boundedness of S. Suppose to the contrary that there exists an unbounded sequence $\{(x^k, y^k)\} \in \mathcal{N}(\mu_0, \beta)$. Then there is also a sequence of scalars $\{\mu_k\}$ satisfying $\|\Psi_{\mu_k}(x, y)\|^2 \leq \beta\mu_k$ and $0 < \mu_k \leq \mu_0$. Since the sequence $\{((x^k, y^k)/\|(x^k, y^k)\|, \mu_k)\}$ is bounded, we may assume without loss of generality that this sequence converges to a point $((x^*, y^*), \mu_*) \in I\!R^{2n+1}$. By dividing the equation $Mx^k - y^k + q = 0$ through by $\|(x^k, y^k)\|$ and taking the limit as $k \to \infty$, we find that

$$Mx^* - y^* = 0. \tag{2.12}$$

In addition, for each $i = 1, \ldots, n$, we have

$$\frac{|\psi_{\mu_k}(x_i^k, y_i^k)|}{\|(x^k, y^k)\|} \leq \frac{\sqrt{\beta\mu_k}}{\|(x^k, y^k)\|} \leq \frac{\sqrt{\beta\mu_0}}{\|(x^k, y^k)\|}.$$

Again, taking the limit in k yields

$$\psi_0(x_i^*, y_i^*) = 0, \text{ for each } i = 1, \ldots, n. \tag{2.13}$$

But (2.13) and (2.12) taken together imply that $(x^*, y^*) \neq 0$ is a solution to LCP$(0, M)$. This contradiction yields the result. □

3 The Algorithm

Step 0 (Initialization)
Let $\mu_0 > 0$, $\beta > 0$, and $(x^0, y^0) \in I\!R^{2n}$ be given so that $(x^0, y^0) \in \mathcal{N}(\mu_0, \beta)$ with $\|\Psi_{\mu_0}(x^0, y^0)\|^2 \leq \mu_0\beta$, and choose $\sigma_i \in (0, 1)$ and $\alpha_i \in (0, 1)$ for $i = 1, 2$.

Step 1 (Computation of the Newton Direction)
Let $(\Delta x^k, \Delta y^k)$ solve the equation

$$F_{\psi_{\mu_k}}(x^k, y^k) + \nabla F_{\psi_{\mu_k}}(x^k, y^k)^T \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0. \tag{3.14}$$

Step 2 (Backtracking Line Search)
If $\Psi_{\mu_k}(x^k, y^k) = 0$, set $(x^{k+1}, y^{k+1}) = (x^k, y^k)$; otherwise, let λ_k be the maximum of the values $1,\ \alpha_1,\ \alpha_1^2,\ \ldots$ such that

$$\left\|\Psi_{\mu_k}(x^k + \lambda_k\Delta x^k, y^k + \lambda_k\Delta y^k)\right\|^2 \leq (1 - \sigma_1\lambda_k) \left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2, \tag{3.15}$$

and set $(x^{k+1}, y^{k+1}) = (x^k + \lambda_k\Delta x^k, y^k + \lambda_k\Delta y^k)$.

Step 3 (Update the Continuation Parameter)
Let γ_k be the maximum of the values $1,\ \alpha_2,\ \alpha_2^2,\ \ldots$ such that

$$||\Psi_{(1-\sigma_2\gamma_k)\mu_k}(x^{k+1}, y^{k+1})||^2 \leq \beta(1-\sigma_2\gamma_k)\mu_k, \tag{3.16}$$

and set $\mu_{k+1} = (1-\sigma_2\gamma_k)\mu_k,\ k = k+1$, and return to Step 1.

Remark An alternative to the backtracking procedure in Step 2 is to set

$$\lambda_k = \frac{1}{4(5+\sqrt{2})} \frac{\left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2}{||(\Delta x^k, \Delta y^k)||^2}$$

on every iteration. This choice of λ_k is justified by Part 2 of Theorem 3.1 and is sufficient for establishing the global linear convergence of the method. However, in our numerical experiments, the backtracking procedure yielded a significantly more efficient algorithm.

We now show that the algorithm is well–defined and implementable when it is assumed that M is a P_0 matrix.

Theorem 3.1 *Let $\beta > 0$. Suppose that M is a P_0 matrix and that $(x^k, y^k, \mu_k) \in I\!R^{2n+1}$ satisfies*

$$\mu_k > 0, \quad Mx^k - y^k + q = 0, \quad \textit{and} \quad \frac{||\Psi_{\mu_k}(x^k, y^k)||^2}{\mu_k} \leq \beta.$$

1. *[12, Theorem 3.5] The Jacobian $\nabla F_{\psi_{\mu_k}}(x^k, y^k)$ is non–singular. Hence, the Newton step in Step 1 of the algorithm exists and is unique.*

2. *If $\Psi_{\mu_k}(x^k, y^k) \neq 0$, then there is a*

$$\hat{\lambda_k} \geq \bar{\lambda_k} := \frac{1}{4(5+\sqrt{2})} \frac{\left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2}{||(\Delta x^k, \Delta y^k)||^2}$$

such that

$$\left\|\Psi_{\mu_k}(x^k + \lambda\Delta x^k, y^k + \lambda\Delta y^k)\right\|^2 \leq (1-\sigma_1\lambda)\left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2$$

for every $\lambda \in [0, \hat{\lambda_k}]$. Hence, the backtracking procedure for evaluating λ_k in Step 2 is finitely terminating. In addition, we have $\alpha_1\bar{\lambda_k} \leq \lambda_k$ with $\bar{\lambda_k} \leq 1$.

3. *There exists $\hat{\gamma_k} > 0$ such that*

$$||\Psi_{(1-\sigma_2\gamma)\mu_k}(x^{k+1}, y^{k+1})||^2 \leq \beta(1-\sigma_2\gamma)\mu_k,$$

for every $\gamma \in (0, \hat{\gamma_k})$. Hence, the backtracking procedure for evaluating γ_k in Step 3 is finitely terminating.

Proof 2. Since

$$\nabla(\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2)^T \begin{pmatrix} \Delta x^k \\ \Delta y^k \end{pmatrix} = -2\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2, \tag{3.17}$$

the Newton step is a direction of strict descent for the function $\|\Psi_{\mu_k}(\cdot,\cdot)\|^2$ at (x^k,y^k), and so the finite termination of the backtracking routine is straightforward.

We show that one can take $\hat{\lambda}_k \geq \bar{\lambda}_k$. For this, it clearly suffices to show that the inequality (3.15) holds for $\lambda_k = \sigma\bar{\lambda}_k$ for any $\sigma \in (0,1]$. Let $(\Delta x^k, \Delta y^k)$ be chosen to satisfy the Newton equation (3.14), and note that for $\lambda > 0$ and $i \in \{1,\ldots,n\}$, we have from Part 2 of Lemma 2.1 that

$$\begin{aligned}
& \psi^2_{\mu_k}(x_i^k + \lambda(\Delta x^k)_i, y_i^k + \lambda(\Delta y^k)_i) \\
= \;& \psi^2_{\mu_k}(x_i^k,y_i^k) + \lambda(\nabla(\psi^2_{\mu_k}(x_i^k,y_i^k)))^T \begin{bmatrix} (\Delta x^k)_i \\ (\Delta y^k)_i \end{bmatrix} + \\
& \lambda^2[(\Delta x^k)_i,(\Delta y^k)_i]\nabla^2(\psi^2_{\mu_k}(x_i^k+\theta_i\lambda(\Delta x^k)_i, y_i^k+\theta_i\lambda(\Delta y^k)_i)) \begin{bmatrix} (\Delta x^k)_i \\ (\Delta y^k)_i \end{bmatrix} \\
\leq \;& \psi^2_{\mu_k}(x_i^k,y_i^k) + 2\lambda\psi_{\mu_k}(x_i^k,y_i^k)(\nabla(\psi_{\mu_k}(x_i^k,y_i^k)))^T \begin{bmatrix} (\Delta x^k)_i \\ (\Delta y^k)_i \end{bmatrix} + \\
& \lambda^2 \left\|[(\Delta x^k)_i,(\Delta y^k)_i]\right\| \left\|\nabla^2(\psi^2_{\mu_k}(x_i^k+\theta_i\lambda(\Delta x^k)_i, y_i^k+\theta_i\lambda(\Delta y^k)_i))\right\| \left\| \begin{bmatrix} (\Delta x^k)_i \\ (\Delta y^k)_i \end{bmatrix} \right\| \\
\leq \;& \psi^2_{\mu_k}(x_i^k,y_i^k) - 2\lambda\psi^2_{\mu_k}(x_i^k,y_i^k) + \lambda^2 4(5+\sqrt{2}) \left\|[(\Delta x^k)_i,(\Delta y^k)_i]\right\|^2 \\
= \;& (1-2\lambda)\psi^2_{\mu_k}(x_i^k,y_i^k) + \lambda^2 4(5+\sqrt{2}) \left\|[(\Delta x^k)_i,(\Delta y^k)_i]\right\|^2,
\end{aligned}$$

for some $\theta_i \in [0,1]$. Hence

$$\left\|\Psi_{\mu_k}(x^k+\lambda\Delta x^k, y^k+\lambda\Delta y^k)\right\|^2 \leq (1-2\lambda)\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2 + \lambda^2 4(5+\sqrt{2}) \left\| \begin{bmatrix} \Delta x^k \\ \Delta y^k \end{bmatrix} \right\|^2. \tag{3.18}$$

The minimum value of the right–hand side of this inequality is attained at

$$\bar{\lambda}_k = \frac{1}{4(5+\sqrt{2})} \frac{\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2}{\|(\Delta x^k,\Delta y^k)\|^2},$$

with optimal value $0 \leq (1-\bar{\lambda}_k)\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2$. In particular, this implies that $\bar{\lambda}_k \leq 1$. Plugging $\sigma\bar{\lambda}_k$ in for λ in (3.18) yields

$$\begin{aligned}
\left\|\Psi_{\mu_k}(x^k+\sigma\bar{\lambda}_k\Delta x^k, y^k+\sigma\bar{\lambda}_k\Delta y^k)\right\|^2 &\leq \left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2 + \sigma(\sigma-2)\bar{\lambda}_k \left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2 \\
&\leq (1-\sigma\bar{\lambda}_k)\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2
\end{aligned}$$

which establishes the result.

3. If $||\Psi_{\mu_k}(x^k,y^k)|| = 0$, then

$$\frac{||\Psi_{\mu_k}(x^{k+1},y^{k+1})||^2}{\mu_k} = \frac{||\Psi_{\mu_k}(x^k,y^k)||^2}{\mu_k} = 0 < \beta;$$

and if $||\Psi_{\mu_k}(x^k, y^k)|| \neq 0$, then from Part 2

$$\frac{||\Psi_{\mu_k}(x^{k+1}, y^{k+1})||^2}{\mu_k} < \frac{||\Psi_{\mu_k}(x^k, y^k)||^2}{\mu_k} \leq \beta.$$

Therefore, in either case, there is an open interval $(0, \hat{\gamma}_k)$ over which the inequality

$$||\Psi_{(1-\sigma_2\gamma)\mu_k}(x^{k+1}, y^{k+1})||^2 < \beta(1 - \sigma_2\gamma)\mu_k$$

is valid for all $\gamma \in (0, \hat{\gamma}_k)$. □

4 Global Linear Convergence

To obtain the global linear convergence of the algorithm described in the previous section, we require the uniform boundedness of $\nabla F_{\psi_\mu}(x, y)^{-1}$ on the set $\mathcal{N}(\mu_0, \beta)$. We establish this uniform boundedness with the aid of a non–degeneracy condition on the problem LCP(q, M).

Definition 4.1 *Let S be the set of solutions to LCP(q, M) defined in (2.11) and set*

$$J := \{j | \text{ there exists } (x, y) \in S \text{ such that } y_j = 0\}.$$

We say that the problem LCP(q, M) satisfies the FLP condition if the principal submatrix M_{JJ} is non–degenerate.

Remarks 1. The FLP condition extends a similar notion due to Fukushima, Luo, and Pang [6, Assumption (A2)].

2. If M is a P matrix, then clearly LCP(q, M) satisfies the FLP condition.

3. An anonymous referee has observed that the FLP condition implies the uniqueness of the solution to LCP(q, M). In order to see this, first observe that our proof technique in Proposition 4.3 can be used to show that the FLP condition implies the non–singularity of every element of the so–called B-subdifferential [15, page 233] of F_{ψ_0} at every point in S. Thus, by [15, Proposition 2.5], S contains only isolated points. But then, by [10, Corollary 5], S must be a singleton.

The key step in establishing the uniform boundedness of $\nabla F_{\psi_\mu}(x, y)^{-1}$ is provided by the following technical lemma due to Fukushima, Luo, and Pang.

Lemma 4.2 *[6, Proposition 3.2] Let $\beta > 0$ and $\mu_0 > 0$ be given and assume that M is a P_0 matrix. Let (x^k, y^k, μ_k) be a sequence in $\mathbb{R}^{2n+1}$ satisfying $\left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2 \leq \beta\mu_k$ and $0 < \mu_k \leq \mu_0$ for all k. If the limit*

$$\lim_{k\to\infty} (\nabla_x \Psi_{\mu_k}(x^k, y^k), \nabla_y \Psi_{\mu_k}(x^k, y^k)) = (D_x, D_y)$$

exists and the principal submatrix M_{II} is non–degenerate, where

$$I = \{i : (D_x)_{ii} = 0\},$$

then the limiting matrix

$$\begin{bmatrix} M & -I \\ D_x & D_y \end{bmatrix}$$

is nonsingular.

Proposition 4.3 *Let $\mu_0 > 0, \beta > 0$ and assume that M is a P_0 and an R_0 matrix for which LCP(q, M) satisfies the FLP condition. Then for all $0 < \mu$ and $x, y \in I\!R^n$ satisfying*

$$0 < \mu \leq \mu^0, \text{ and } \|\Psi_\mu(x,y)\|^2 \leq \beta\mu$$

there exists a constant $L > 0$ such that

$$\|(\nabla F_{\psi_\mu}(x,y))^{-1}\| \leq L. \tag{4.1}$$

Proof Assume to the contrary that there is a sequence $\{(x^k, y^k, \mu_k)\}$ such that $0 < \mu_k \leq \mu_0$, $\left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2 \leq \beta\mu_k$, and $\|(\nabla F_{\psi_{\mu_k}}(x^k, y^k))^{-1}\| \geq k$. By Proposition 2.4, the sequence $\{(x^k, y^k, \mu_k)\}$ is bounded, hence we can assume that the sequence converges to some point (x^*, y^*, μ_*). If $\mu_* > 0$, then $\nabla F_{\psi_{\mu_*}}(x^*, y^*)$ is non–singular which implies the boundedness of the sequence $\|(\nabla F_{\psi_{\mu_k}}(x^k, y^k))^{-1}\|$. Hence it must be the case that $\mu_* = 0$. Therefore $(x^*, y^*) \in S$. In addition, from the definition of $\nabla F_{\psi_{\mu_k}}$, the sequence $(\nabla_x \Psi_{\mu_k}(x^k, y^k), \nabla_y \Psi_{\mu_k}(x^k, y^k))$ is also bounded, so with no loss in generality

$$\lim_{k\to\infty} (\nabla_x \Psi_{\mu_k}(x^k, y^k), \nabla_y \Psi_{\mu_k}(x^k, y^k)) = (D_x, D_y)$$

for some non–negative diagonal matrices D_x and D_y. Let $I = \{i | (D_x)_{ii} = 0\}$. It is easy to check that $I \subseteq J$. Therefore, M_{II} is non–degenerate since M_{JJ} is. Hence Lemma 4.2 implies that $\lim_k \nabla F_{\mu_k}(x^k, y^k)$ exists and is non–singular. But then again the sequence $\|(\nabla F_{\mu_k}(x^k, y^k))^{-1}\|$ must be bounded. This contradiction yields the result. □

We are now in position to state and prove the global linear convergence result for the algorithm described in the previous section.

Theorem 4.4 *Suppose that M is a P_0 and an R_0 matrix, and that the problem LCP(q, M) satisfies the FLP condition. Let (x^k, y^k, μ_k) be the sequence generated by the algorithm of Section 3. Then*

(i) *For $k = 0, 1, \ldots,$*

$$Mx^k - y^k + q = 0, \tag{4.2}$$

$$\frac{\|\Psi_{\mu_k}(x^k, y^k)\|^2}{\mu_k} \leq \beta, \text{ and} \tag{4.3}$$

$$(1 - \sigma_2\gamma_{k-1}) \ldots (1 - \sigma_2\gamma_0)\mu_0 = \mu_k. \tag{4.4}$$

(ii) *For all $k \geq 0$, we have*

$$\gamma_k \geq \bar{\gamma} := \min\{1, \frac{\beta\alpha_1\alpha_2\sigma_1\sigma_2^{-1}}{4(5+\sqrt{2})((2+2\sqrt{2})n+\beta)L^2}\}, \tag{4.5}$$

where L is the constant defined in (4.1). Therefore, μ_k converges to 0 at a global linear rate.

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

Proof (i) We establish (4.2)–(4.4) by induction on k. Clearly these relations hold for $k = 0$. Now assume that they hold for some $k > 0$. By Theorem 3.1, the algorithm is well defined and so (4.3) and (4.4) hold with k replaced by $k+1$. Since (3.14) is satisfied for all k with $Mx^0 - y^0 + q = 0$, we have that $Mx^k - y^k + q = 0$ for all k, and so, in particular, it is true when k replaced by $k+1$. Hence, by induction, (4.2)–(4.4) hold for all k.

(ii) We now show that for all k and $0 < \gamma < 1$, we have

$$\frac{||\Psi_{(1-\gamma)\mu_k}(x^{k+1}, y^{k+1})||^2}{(1-\gamma)\mu_k} \leq \frac{(1-\eta)\beta + (2+2\sqrt{2})n\gamma}{1-\gamma}, \tag{4.6}$$

for some $\eta \in (0, 1]$. This is done by separately establishing the cases where $\left\|\Psi_{\mu_k}(x^k, y^k)\right\|$ is equal to zero and not equal to zero.

We consider the case $\left\|\Psi_{\mu_k}(x^k, y^k)\right\| = 0$ first. In this case, $x^{k+1} = x^k$ and $y^{k+1} = y^k$. Thus, by Part 3 of Lemma 2.1,

$$\begin{aligned}
\frac{\left\|\Psi_{(1-\gamma)\mu_k}(x^{k+1}, y^{k+1})\right\|^2}{(1-\gamma)\mu_k} &= \frac{\left\|\Psi_{(1-\gamma)\mu_k}(x^k, y^k)\right\|^2}{(1-\gamma)\mu_k} \\
&\leq \frac{\left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2 + (2+2\sqrt{2})n\gamma\mu_k}{(1-\gamma)\mu_k} \\
&= \frac{(2+2\sqrt{2})n\gamma}{1-\gamma},
\end{aligned}$$

from which the inequality (4.6) follows for any choice of $\eta \in (0, 1]$.

Next consider the case $\left\|\Psi_{\mu_k}(x^k, y^k)\right\| \neq 0$. First observe that from Proposition 4.3, we have

$$\begin{aligned}
\left\|\begin{bmatrix} \Delta x^k \\ \Delta y^k \end{bmatrix}\right\|^2 &= \left\|(\nabla F_{\psi_{\mu_k}}(x^k, y^k)^T)^{-1}\begin{bmatrix} 0 \\ -\Psi_{\mu_k}(x^k, y^k) \end{bmatrix}\right\|^2 \\
&\leq \left\|(\nabla F_{\psi_{\mu_k}}(x^k, y^k))^{-1}\right\|^2 \left\|\begin{bmatrix} 0 \\ -\Psi_{\mu_k}(x^k, y^k) \end{bmatrix}\right\|^2 \\
&\leq L^2 \left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2.
\end{aligned}$$

Hence, if we define $\eta := (\alpha_1\sigma_1)/(4(5+\sqrt{2})L^2)$, then

$$\eta \leq \frac{\alpha_1\sigma_1}{4(5+\sqrt{2})} \frac{\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2}{\left\|\begin{bmatrix} \Delta x^k \\ \Delta y^k \end{bmatrix}\right\|^2},$$

with $\eta \in (0,1]$ by Part 2 of Theorem 3.1. From Step 2 of the algorithm and Part 2 of Theorem 3.1, we have

$$\left\|\Psi_{\mu_k}(x^{k+1},y^{k+1})\right\|^2 \leq (1-\eta)\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2.$$

Therefore, by Part 3 of Lemma 2.1,

$$\begin{aligned}
\frac{\left\|\Psi_{(1-\gamma)\mu_k}(x^{k+1},y^{k+1})\right\|^2}{(1-\gamma)\mu_k} &\leq \frac{\left\|\Psi_{\mu_k}(x^{k+1},y^{k+1})\right\|^2 + (2+2\sqrt{2})n\gamma\mu_k}{(1-\gamma)\mu_k} \\
&\leq \frac{(1-\eta)\left\|\Psi_{\mu_k}(x^k,y^k)\right\|^2 + (2+2\sqrt{2})n\gamma\mu_k}{(1-\gamma)\mu_k} \\
&\leq \frac{(1-\eta)\beta\mu_k + (2+2\sqrt{2})n\gamma\mu_k}{(1-\gamma)\mu_k} \\
&= \frac{(1-\eta)\beta + (2+2\sqrt{2})n\gamma}{1-\gamma},
\end{aligned}$$

whereby inequality (4.6) is confirmed.

It is easily verified that

$$\frac{(1-\eta)\beta + (2+2\sqrt{2})n\gamma}{1-\gamma} \leq \beta \quad \text{whenever} \quad \gamma \leq \frac{\eta\beta}{(2+2\sqrt{2})n+\beta}.$$

Therefore,

$$\sigma_2\gamma_k \geq \alpha_2 \frac{\eta\beta}{(2+2\sqrt{2})n+\beta},$$

or equivalently, $\gamma_k \geq \bar{\gamma}$.

(iii) The boundedness of the sequence $\{(x^k,y^k)\}$ follows from Proposition 2.4 . If $\{(x^{k_j},y^{k_j})\}$ is a convergent subsequence with limit (x^*,y^*), then it follows from (i) and (ii) that $||\Psi_0(x^*,y^*)|| = 0$ and $Mx^* - y^* + q = 0$, so $(x^*,y^*) \in S$. □

The algorithm of Section 3 requires a finite stopping criteria in order to be practical. To be useful, such a stopping criteria should say something about the quality of the approximate solution obtained at termination. The next lemma shows that such a finite stopping criteria can be based on the values μ_k.

Lemma 4.5 *Let $\epsilon > 0$. If $|\psi_\mu(a,b)| \leq \epsilon$, then*

$$-\epsilon \leq a, \quad -\epsilon \leq b, \quad \textit{and} \quad \frac{|ab-\mu|}{|a|+|b|+\sqrt{\mu}} \leq \epsilon.$$

Proof If $|\psi_\mu(a,b)| \leq \epsilon$, then

$$0 \leq \sqrt{a^2+b^2+2\mu} \leq \epsilon + a + b.$$

If $\epsilon + a < 0$, then

$$\epsilon + a + b < b \leq \sqrt{a^2+b^2+2\mu},$$

which is a contradiction. Hence $a \geq -\epsilon$. Similarly, $b \geq -\epsilon$.

Also

$$\begin{aligned} |\psi_\mu(a,b)| &= \frac{|(a+b)^2 - (a^2+b^2+2\mu)|}{|(a+b) + \sqrt{(a^2+b^2+2\mu)}|} \\ &\geq \frac{2|ab-\mu|}{(|a|+|b|) + (|a|+|b|+\sqrt{2\mu})}, \end{aligned}$$

which yields the result. □

Therefore, if $\mu_k \leq \epsilon$ and $\left\|\Psi_{\mu_k}(x^k,y^k)\right\| \leq \beta\mu_k$, then

$$-\sqrt{\beta\epsilon} \leq x_i^k, \quad -\sqrt{\beta\epsilon} \leq y_i^k, \quad \text{and} \quad \frac{|x_i^k y_i^k - \mu_k|}{|x_i^k| + |y_i^k| + \sqrt{\mu_k}} \leq \sqrt{\beta\epsilon} \quad \text{for every } i = 1,2,\ldots,n.$$

For high dimensional problems, a more practical stopping criteria is to require that $\mu_k \leq \epsilon$ and $\left\|\Psi_{\mu_k}(x^k,y^k)\right\|_\infty \leq \epsilon$. This criteria implies that

$$-\epsilon \leq x_i^k, \quad -\epsilon \leq y_i^k, \quad \text{and} \quad \frac{|x_i^k y_i^k - \mu_k|}{|x_i^k| + |y_i^k| + \sqrt{\mu_k}} \leq \epsilon \quad \text{for every } i = 1,2,\ldots,n.$$

Both of these stopping criteria have the rather nice feature that they induce termination when the *relative error* in the complementarity condition is small.

5 Numerical Experiments

5.1 Implementation Details

Although the algorithm of Section 3 is implementable as it stands, we have chosen to modify the implementation in order to make the comparison with Kanzow's algorithm [12] more straightforward. In particular, we are interested in observing how the updating strategy for the continuation parameter μ affects performance. For this reason, the implemented algorithm differs from Kanzow's algorithm only in the way that the continuation parameter μ is updated. This choice of implementation illustrates the benefits of an updating strategy based on the neighborhood $\mathcal{N}(\beta)$. The key difference between the algorithm of Section 3 and the one stated below is the use of a non–monotone line search strategy in Step 3.

The Implemented Algorithm

Step 0 (Initialization)
Let $\epsilon > 0$, $\sigma_1, \sigma_2 \in (0,1)$, $\alpha_1, \alpha_2 \in (0,1)$, and $p \in \{1,2,\ldots\}$. Choose $x^0 \in I\!R^n, y^0 \in I\!R^n, \mu_0 > 0$ such that

$$Mx^0 - y^0 + q = 0, \tag{5.1}$$

and let

$$\beta = \frac{||\Psi_{\mu_0}(x^0, y^0)||^2}{\mu_0}. \tag{5.2}$$

Step 1 (Check Termination Criterion)
If $||\min\{x^k, y^k\}|| \leq \epsilon$, stop; otherwise, continue to Step 2.

Step 2 (Computation of the Newton Direction)
Compute the Newton direction $(\Delta x^k, \Delta y^k)$ as the solution of the linear system:

$$F_{\psi_{\mu_k}}(x^k, y^k) + \nabla F_{\psi_{\mu_k}}(x^k, y^k)^T \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0. \tag{5.3}$$

Step 3 (Non–Monotone Backtracking Line Search)
Let $p_k := \min\{k, p\}$ and $\lambda_k = \alpha_1^{s_k}$, where s_k is the smallest nonnegative integer $s \in \{0, 1, \ldots, \}$ satisfying

$$\left\|\Psi_{\mu_k}(x^k + \alpha_1^s \Delta x^k, y^k + \alpha_1^s \Delta y^k)\right\|^2 \leq \max_{j=k-p_k+1,\ldots,k} \left\|\Psi_{\mu_j}(x^j, y^j)\right\|^2 - \sigma_1 \alpha_1^s \left\|\Psi_{\mu_k}(x^k, y^k)\right\|^2.$$

Let

$$x^{k+1} = x^k + \lambda_k \Delta x^k, \quad y^{k+1} = y^k + \lambda_k \Delta y^k. \tag{5.4}$$

Step 4 (Update the Continuation Parameter)
If $\left\|\Psi_{\mu_k}(x^{k+1}, y^{k+1})\right\| \geq \left\|\Psi_{\mu_k}(x^k, y^k)\right\|$, let $\mu_{k+1} = \mu_k$; else let $\gamma_k = \alpha_2^{t_k}$, where t_k is the smallest nonnegative integer $t \in \{0, 1, 2, \ldots\}$ satisfying

$$\frac{\left\|\Psi_{(1-\sigma_2\alpha_2^t)\mu_k}(x^{k+1}, y^{k+1})\right\|^2}{(1 - \sigma_2 \alpha_2^t)\mu_k} \leq \beta, \tag{5.5}$$

and let

$$\mu_{k+1} = (1 - \sigma_2 \gamma_k)\mu_k.$$

Set $k := k + 1$, and go to Step 1.

In our implementation, we choose $\epsilon = 10^{-6}$, $\sigma_1 = 10^{-4}$, $\alpha_1 = 0.75$, $\sigma_2 = 0.9999$, $\alpha_2 = 0.99$, $p = 5$. The initial $\mu_0 = \|q\|/n$.

The effects of an initial rescaling of the input data have been studied by Kanzow [12]. It was found that rescaling can have a dramatic impact on the numerical results. The rescaling suggested by Kanzow is to replace M and q by ΣM and ΣQ, respectively, where

$$\Sigma := \text{diag}\,(\sigma_1, \ldots, \sigma_n), \ \sigma_i := \begin{cases} 1/M(i,i) & \text{if} \quad M(i,i) \neq 0, \\ 1 & \text{if} \quad M(i,i) = 0. \end{cases} \tag{5.6}$$

	n=8	n=16	n=32	n=64	n=128	n=256
Example 5.1	5	5	5	5	5	5
Example 5.2	5	8	7	9	8	10

Table 1: The number of iterations the algorithm of Section 5.1 required to obtain a solution in Examples 5.1 and 5.2. These results are slightly better than those appearing in [12, Tables 1 and 2, Method 2].

One then solves the problem $LCP(\Sigma q, \Sigma M)$. Our numerical experiment is performed on both the original input data and the scaled input data.

5.2 Test Problems and Numerical Results

Example 5.1 *(Murty [14]): n variables,*

$$M = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 0 & 1 & 2 & \dots & 2 \\ 0 & 0 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{5.7}$$

$$q = (-1, \dots, -1)^T.$$

This is a standard test problem. The solution is $x^* = (0, \dots, 0, 1)^T$, $y^* = (1, \dots, 1, 0)^T$. The matrix in this example is a P-matrix. We take $x^0 = (1, \dots, 1)^T$ as our starting point. The numerical results for this test problem can be found in Table 1.

Example 5.2 *(Fathi [5]): n variables,*

$$M = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 5 & 6 & \dots & 6 \\ 2 & 6 & 9 & \dots & 10 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 6 & 9 & \dots & 4(n-1)+1 \end{pmatrix}, \tag{5.8}$$

$$q = (-1, \dots, -1)^T.$$

This is another standard test problem. The solution is $x^* = (1, 0, \dots, 0)^T, y^* = (0, 1, \dots, 1)^T$. The matrix M of this example is positive definite. We take $x^0 = (1, \dots, 1)^T$ as our starting point. The numerical results for this test problem can be found in Table 1.

n		Example 5.3		Example 5.4	
		NS	S	NS	S
n=50	Max.	12	9	15	10
	Avg.	10.1	7.4	12.1	8.6
	Min.	8	6	10	7
n=100	Max.	13	9	16	11
	Avg.	11	7.8	13.7	9
	Min.	8	6	12	8
n=150	Max.	15	9	16	10
	Avg.	12.1	7.9	13.8	9.3
	Min.	10	7	12	8
n=200	Max.	16	10	15	10
	Avg.	12.7	8.8	14.6	9.1
	Min.	11	8	14	8

Table 2: Number of iterations for examples 5.3 and 5.4. These results significantly improve on those appearing in [12, Tables 3 and 4, Method 2]. In addition, note that both the non–scaled and scaled versions of the algorithm perform quite well.

Example 5.3 *(Harker and Pang [8])*

The matrix M is computed as follows: Let $A, B \in I\!R^{n\times n}$ and $q, d \in I\!R^n$ be randomly generated such that $a_{ij}, b_{ij} \in (-5, 5), q_i \in (-500, 500), d_i \in (0.0, 0.3)$ and that B is skew-symmetric. Define $M = A^T A + B + \text{diag}\,(d)$. Then M is a $P-$matrix. In this example, the results using the problem input data as above is reported in the column NS. The result using Kanzow's rescaling technique described in (5.6) is reported in the column S. Ten problems are generated in this way for each of the dimensions $n = 50,\ 100,\ 150,\ 200$. The maximum, average, and minimum number of iterations needed by the algorithms are summarized in Table 2. In all runs, the starting point is chosen to be $x^0 = (0, \ldots, 0)^T$.

Example 5.4 *(Harker and Pang "hard examples" [8])*

In this example, M is computed in the same way as in the previous example, however, $q \in R^n$ is randomly generated with entries $q_i \in (-500, 0)$. Table 2 contains the numerical results. In this example, the results using the problem input data as above is reported in the column NS. The result using Kanzow's rescaling technique described in (5.6) is reported in the column S. In all the test runs, the starting point is chosen to be $x^0 = (0, \ldots, 0)^T$.

6 Concluding Remarks

In this paper we present the first global linear convergence result for non–interior path following smoothing methods for LCP. This result along with the work in [9] and [20] has initiated a flurry of activity on rate convergence analysis for non–interior point path following methods based on smoothing techniques [1, 2, 4, 16, 18, 19]. In all cases it is the notion of a neighborhood of the central path that provides the key to establishing the global linear convergence of the algorithms. However, the question of the complexity of these methods remains open. In [20], we show that an interior point implementation of a non–interior path following smoothing method has the same best polynomial–time complexity as is exhibited by the standard short–step interior point path following algorithm. This result along with the linear convergence results cited above hold forth the possibility of a polynomial complexity result for non–interior path following methods. In this regard, a deeper understanding of the dependence of the parameter L in Proposition 4.3 on the problem data will be crucial.

References

[1] B. Chen and X. Chen. 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.

[2] B. Chen and X. Chen. A global and local super–linear continuation method for P_0+R_0 and monotone NCP. Preprint, Department of Management and Systems, Washington State University, Pullman, WA 99164-4736, May, 1997.

[3] B. Chen and P.T. Harker. A non–interior–point continuation method for linear complementarity problems. *SIAM J. Matrix Anal. Appl.*, 14:1168—1190, 1993.

[4] 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. Preprint, Department of Management and Systems, Washington State University, Pullman, WA 99164-4736, February, 1997.

[5] Y. Fathi. Computational complexity of LCPs associated with positive definite matrices. *Math. Programming*, 17:335—344, 1979.

[6] M. Fukushima, Z.-Q. Luo, and J.-S. Pang. A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. Preprint, Department of Mathematical Sciences, Whiting Scool of Engineering, The Johns Hopkins University, Baltimore, Maryland, 21218–2692, USA, 1996.

[7] C. Geiger and C. Kanzow. On the resolution of monotone complementarity problems. *Comp. Optim. Appl.*, 5:155—173, 1996.

[8] P.T. Harker and J.-S. Pang. A damped Newton method for the linear complementarity problem. In E.L. Allgower and K. Georg, editors, *Computational Solution of Nonlinear Systems of Equations*, pages 265—284. Lectures in Applied Mathematics, Volume 26, AMS, Providence, Rhode Island, 1990.

[9] 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.

[10] C. Jones and M. S. Gowda. On the connectedness of solution sets in linear complementarity problems. Technique Report Number TR97–01, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21520, 1997.

[11] C. Kanzow. An unconstrained optimization technique for large–scale linearly constrained convex minimization problems. *Computing*, 53:101—117, 1994.

[12] C. Kanzow. Some noninterior continuation methods for linear complementarity problems. *SIAM J. Matrix Anal. Appl.*, 17:851—868, 1996.

[13] M. Kojima, N. Meggido, T. Noma, and A. Yoshise. *A unified approach to interior point algorithms for linear complementarity problems*. Springer–Verlag, Berlin, 1991.

[14] K.G. Murty. *Linear Complementarity, Linear and Nonlinear Programming*. Sigma Series in Allpied Mathematics 3, Heldermann, Berlin, 1988.

[15] L. Qi. Convergence analysis of some algorithms for solving nonsmooth equations. *Math. of Oper. Res.*, 18:227—244, 1993.

[16] L. Qi and D. Sun. Globally linearly, and globally and locally superlinearly convergent versions of the Hotta–Yoshise non–interior point algorithm for nonlinear complementarity problems. Applied Mathematics Report, School of Mathematics University of New South Wales, Sydney 2052, Australia, May, 1997.

[17] P. Tseng. Simplified analysis of an O(nL)–iteration infeasible predictor–corrector path–following method for monotone LCP. In R.P. Agarwal, editor, *Recent Trends in Optimization Theory and Applications*, pages 423—434. World Scientific Press, Singapore, 1994.

[18] 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.

[19] S. Xu. 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.

[20] S. Xu and J. V. Burke. A polynomial time interior point path–following algorithm for LCP based on Chen–Harker–Kanzow smoothing techniques. Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, September, 1996.

[21] Y. Zhang. On the convergence of a class of infeasible interior point algorithms for the horizontal linear complementarity problem. *SIAM J. Optimization*, 4:208—227, 1994.