Math. Program. 86: 91–103 (1999)

Digital Object Identifier (DOI) 10.1007/s10107990056a

Song Xu · James V. Burke

A polynomial time interior–point path–following algorithm for LCP based on Chen–Harker–Kanzow smoothing techniques

Received September 9, 1996 / Revised version received December 22, 1997
Published online March 16, 1999

Abstract. A polynomial complexity bound is established for an interior point path following algorithm for the monotone linear complementarity problem that is based on the Chen–Harker–Kanzow smoothing techniques. The fundamental difference with the Chen–Harker and Kanzow algorithms is the introduction of a rescaled Newton direction. The rescaling requires the iterates to remain in the interior of the positive orthant. To compensate for this restriction, the iterates are not required to remain feasible with respect to the affine constraints. If the method is initiated at an interior point that is also feasible with respect to the affine constraints, then the complexity bound is $O(\sqrt{n}L)$; otherwise, the complexity bound is $O(nL)$. The relations between our search direction and the one used in the standard interior-point algorithm are also discussed.

Key words. linear complementarity – polynomial complexity – path following – interior point method

1. Introduction

Consider the *monotone linear complementarity problem*:

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

$$Mx^* - y^* + q = 0, \tag{1}$$

$$x^* \geq 0, y^* \geq 0, (x^*)^T y^* = 0, \tag{2}$$

where $M \in \mathbb{R}^{n\times n}$ is positive semi–definite and $q \in \mathbb{R}^n$.

In this paper, we establish the polynomial complexity of an interior point path following algorithm for LCP. The proposed algorithm can be viewed as an interior point variation on the Chen and Harker [5] and the Kanzow [21] non–interior path following algorithms for LCP. The algorithm has the same best polynomial–time complexity as is exhibited by the standard short–step interior point path following algorithm. The results of this paper represent a first step toward understanding the relationship between interior and non–interior path following methods and provide a spring–board for discovering the complexity of the new non–interior path following algorithms for LCP.

S. Xu: Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, e-mail: s2xu@math.uwaterloo.ca

J.V. Burke: Department of Mathematics, Box # 354350, University of Washington, Seattle, WA 98195–4350, e-mail: burke@math.washington.edu.

This research is supported by the National Science Foundation Grant No. DMS-9303772

Mathematics Subject Classification (1991): 90C33

Path following (or continuation) methods for solving LCP are typically designed to follow the path in the positive orthant, $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$, determined by the equations $F_{\theta_\mu}(x, y) = 0$ for $\mu > 0$ where the function $F_{\theta_\mu} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ is given by

$$F_{\theta_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Theta_\mu(x, y) \end{bmatrix}, \tag{3}$$

with $\theta_\mu(a, b) = ab - \mu$ and

$$\Theta_\mu(x, y) = \begin{bmatrix} \theta_\mu(x_1, y_1) \\ \dots \\ \theta_\mu(x_n, y_n) \end{bmatrix}. \tag{4}$$

This path is called the *central path* [23]. Most path following methods attempt to follow the central path by applying Newton's method to the equations $F_{\theta_\mu}(x, y) = 0$ for decreasing values of μ. In this regard, predictor–corrector strategies are the most popular due to their rapid local convergence (for examples, see [29,31,34,35]). In a predictor–corrector strategy a predictor step ($\mu = 0$) is followed by a corrector step ($\mu > 0$) to return the iterates to a pre-specified neighborhood of the central path.

Interior point methods stay in the vicinity of the central path and remain in the positive orthant [23]. Each iterate of a *feasible* interior point method must satisfy the affine equation $0 = Mx - y + q$ while the iterates of an *infeasible* interior point method are not required to satisfy this equation.

Non–interior path following methods also follow the central path, but the iterates do not necessarily reside in the positive orthant. The first non–interior path following method for LCP was developed by Chen and Harker [5] and was based on a scaled version of the function

$$\upsilon_\mu(a, b) = \frac{a + b}{2} - \sqrt{\frac{(a - b)^2}{4} + \mu}.$$

Later Kanzow [21] developed non–interior path following methods based on the functions υ_μ and

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

It is easy to show that $\psi_\mu(a, b) = 0$ (or $\upsilon_\mu(a, b) = 0$) if and only if $0 \le a,\ 0 \le b$, and $ab = \mu$. Thus, the functions ψ_μ and υ_μ have a fundamental advantage over the function θ_μ which makes them well suited to non–interior path following methods. That is, the condition $\psi_\mu(a, b) = 0$ or ($\upsilon_\mu(a, b) = 0$) guarantees the non–negativity of the arguments a and b.

Using ψ_μ and υ_μ as building blocks, one defines the functions

$$F_{\psi_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Psi_\mu(x, y) \end{bmatrix}, \tag{5}$$

where

$$\Psi_\mu(x, y) = \begin{bmatrix} \psi_\mu(x_1, y_1) \\ \dots \\ \psi_\mu(x_n, y_n) \end{bmatrix}, \tag{6}$$

and

$$F_{\upsilon_\mu}(x, y) := \begin{bmatrix} Mx - y + q \\ \Upsilon_\mu(x, y) \end{bmatrix}, \tag{7}$$

where

$$\Upsilon_\mu(x, y) = \begin{bmatrix} \upsilon_\mu(x_1, y_1) \\ \dots \\ \upsilon_\mu(x_n, y_n) \end{bmatrix}. \tag{8}$$

Clearly, a point (x, y) is on the central path if and only if $F_{\psi_\mu}(x, y) = 0$ (or $F_{\upsilon_\mu}(x, y) = 0$).

Setting $\mu = 0$, we have $\upsilon_0(a, b) = \min\{a, b\}$. This instance of υ_μ has been studied extensively by Pang [26,27] and Harker and Pang [18]. Again taking $\mu = 0$, the function $\psi_0(a, b)$ was introduced by Fischer in [12] who attributes the function to Burmeister. In the growing literature associated with the function ψ_0 [8,10,11,13–15,20,22,30] it is often referred to as the Fischer–Burmeister function. Newton–like implementations based on these functions have proven to be quite successful. Extensions to solving nonlinear programming problems with equilibrium constraints are also being studied [9,16].

Two reasons for the growing interest in non–interior methods based on the functions υ_μ and ψ_μ are (1) these methods are ideally suited for application to the nonlinear complementarity problem where the interiority restriction on the iterates is quite severe, and (2) the numerical evidence on the efficiency of these methods is very impressive. We partially explain this numerical success by establishing the polynomial complexity of an interior point implementation. This is the first complexity result available for these methods and indicates that a similar complexity result may be possible for a non–interior implementation.

The functions υ_μ and ψ_μ are very closely related. By rewriting the expression under the square root, we see that

$$\upsilon_\mu(a, b) = \frac{a + b}{2} - \sqrt{\frac{(a + b)^2}{4} + (\mu - ab)}$$

and

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

The analysis of the algorithms based on these two functions are very similar differing only by a constant here and there. In what follows, we choose to focus on the function ψ_μ. However, whenever appropriate, we indicate how the analysis differs when the function υ_μ is used instead of ψ_μ.

The plan of the paper is as follows. In Section 2, we discuss the rescaled Newton direction for the functions F_{ψ_μ} and F_{υ_μ} and its relation to the direction used in the standard interior point methods. The algorithm and its complexity are presented in Section 3. We conclude in Section 4 with some remarks on the relationship between our algorithm and the algorithms studied by Mizuno [24,25].

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 with $\mathbb{R}^n_{++}$ being the positive orthant, i.e. the set of vectors in $\mathbb{R}^n$ that are componentwise positive. Following standard usage in the interior point literature, we denote by $e \in \mathbb{R}^n$ the vector each of whose components is 1, and, for the vectors $x,\ y$, and z in $\mathbb{R}^n$, we denote by $X,\ Y$, and Z the diagonal matrices whose diagonal entries are given by $x,\ y$, and z, respectively, e.g., $X_{ii} = x_i$ for $i = 1, 2, \ldots, n$. With this notation, the function $\Theta_\mu(x, y)$ defined in (4) can be written as $\Theta_\mu(x, y) = Xy - \mu e$. Given $x \in \mathbb{R}^n$, we denote by $\|x\|_1$, $\|x\|$, and $\|x\|_\infty$, the 1–norm, the 2–norm, and the ∞–norm of x, respectively, and by $x_{\min}$ the minimum component of the vector x.

2. The rescaled Newton directions

The first step in our analysis is to rescale the Newton step to yield iterates comparable to those of a standard interior point strategy. By analogy with the infeasible interior point strategies, at iteration k we compute a Newton step based on the equations

$$F_{\psi_{\mu^k}}(x, y) = \begin{bmatrix} s^k \\ 0 \end{bmatrix},$$

where $s^k := (1 - \gamma^k)(Mx^k - y^k + q)$ with $0 < \gamma^k < 1$. The equations for the Newton step $(\Delta x^k, \Delta y^k)$ take the form

$$M\Delta x - \Delta y = -\gamma^k(Mx^k - y^k + q) \tag{9}$$

$$D_{x^k}\Delta x + D_{y^k}\Delta y = -\Psi_{\mu^k}(x^k, y^k), \tag{10}$$

where

$$D_{x^k} := \operatorname{diag}\left(\frac{1}{\sqrt{2}} - \frac{x_i^k}{2\sqrt{\frac{(x_i^k)^2+(y_i^k)^2}{2} + \mu^k}}\right)$$

and

$$D_{y^k} := \operatorname{diag}\left(\frac{1}{\sqrt{2}} - \frac{y_i^k}{2\sqrt{\frac{(x_i^k)^2+(y_i^k)^2}{2} + \mu^k}}\right).$$

Observe that if (x, y) is on the central path, then

$$\sqrt{\frac{x_i^2 + y_i^2}{2} + \mu} = \frac{x_i + y_i}{\sqrt{2}}, \text{ for } i = 1, \ldots, n.$$

By replacing the expression $\sqrt{\frac{(x_i^k)^2+(y_i^k)^2}{2}+\mu^k}$ in the definitions of the diagonal matrices D_{x^k} and D_{y^k} by the expression $\frac{x_i^k+y_i^k}{\sqrt{2}}$ and then multiplying (10) through by the diagonal matrix diag $\left(\sqrt{2}(x_i^k+y_i^k)\right)$, we obtain the rescaled Newton equations

$$M\Delta x - \Delta y = -\gamma^k(Mx^k - y^k + q) \tag{11}$$
$$Y^k\Delta x + X^k\Delta y = -2\hat{\Psi}_{\mu^k}(x^k, y^k)\ . \tag{12}$$

where, for the sake of convenience, we define

$$\hat{\Psi}_\mu(x, y) := \text{diag}\left(\frac{x_i+y_i}{\sqrt{2}}\right)\Psi_\mu(x, y)\ .$$

The only difference between these rescaled Newton equations and the Newton equations used in a standard interior point path following strategy occurs in equation (12) where $2\hat{\Psi}_{\mu^k}(x^k, y^k)$ replaces the usual term $\Theta_{\mu^k}(x^k, y^k)$. The pattern of our development should now be clear. After a few identities and inequalities relating the functions $\hat{\Psi}_\mu(x, y)$ and $\Theta_\mu(x, y)$ have been established, a convergence theory and complexity analysis can be developed which is based on standard techniques from the theory of interior point path following methods. The necessary identities and inequalities are given in the next lemma.

Lemma 1. *For $a, b, \mu \in I\!R$ satisfying $a > 0, b > 0, \mu > 0$, we have*

$$\hat{\psi}_\mu(a, b) = \frac{(a+b)\theta_\mu(a, b)}{(a+b)+\sqrt{a^2+b^2+2\mu}}, \tag{13}$$
$${\psi_\mu}^2(a, b) = 2\hat{\psi}_\mu(a, b) - \theta_\mu(a, b),\ \text{and} \tag{14}$$
$$|\hat{\psi}_\mu(a, b)| \le |\theta_\mu(a, b)|, \tag{15}$$

where $\hat{\psi}_\mu(a, b) := \frac{a+b}{\sqrt{2}}\psi_\mu(a, b)$. In addition, given $0 < \beta < 1$, $0 < \mu$, and $(x, y) \in I\!R^n_{++} \times I\!R^n_{++}$ satisfying

$$\left\|\Theta_\mu(x, y)\right\| \le \beta\mu, \tag{16}$$

we have

$$\left\|2\hat{\Psi}_\mu(x, y) - \Theta_\mu(x, y)\right\| \le \frac{\beta^2\mu}{2(1-\beta)}. \tag{17}$$

Remark 1. The identity (13) is due to B. Chen [2].

Remark 2. Inequality (15) implies that for $\mu > 0$ and $\beta > 0$, the condition (16) (used in standard interior point methods to define the neighborhood of the central path) implies that the condition

$$\left\|\hat{\Psi}_\mu(x, y)\right\| \le \beta\mu$$

is also satisfied.

Remark 3. The identities (13) and (14), the inequalities (15) and (17), and the second remark remain valid with the expressions $\left[(a+b)+\sqrt{a^2+b^2+2\mu}\right]$, $\hat{\psi}_\mu$, and $\hat{\Psi}_\mu$ replaced by $\left[(a+b)+\sqrt{(a-b)^2+4\mu}\right]$, $\hat{\upsilon}_\mu$ and $\hat{\Upsilon}_\mu$, respectively, where $\hat{\upsilon}_\mu(a,b) := \frac{a+b}{2}\upsilon_\mu(a,b)$ and $\hat{\Upsilon}_\mu(x,y) := \mathrm{diag}\left(\frac{x_i+y_i}{2}\right)\Upsilon_\mu(x,y)$.

Remark 4. The bound (17) shows that the values $2\hat{\Psi}_{\mu^k}(x^k,y^k)$ approach the values $\Theta_{\mu^k}(x^k,y^k)$ used in the standard interior point methods as μ_k approaches 0. This partially explains why the interior point method based on the rescaled Newton direction studied in the next section has the same best polynomial-time complexity as the standard short step path-following interior point methods.

Proof. For $a,b,\mu \in I\!R$ satisfying $a>0, b>0, \mu>0$, the identity (13) is easily derived. The identity (14) and the inequality (15) follow readily from (13).

In order to see the bound (17), note that for any $(x,y) \in I\!R^n_{++} \times I\!R^n_{++}$ satisfying (16), we have

$$x_i y_i \geq (1-\beta)\mu \quad \text{for } i = 1,2,\ldots,n,$$

and so

$$\frac{(x_i+y_i)^2}{2} = \frac{(x_i-y_i)^2+4x_iy_i}{2} \geq 2x_iy_i \geq 2(1-\beta)\mu \quad \text{for } i=1,2,\ldots,n. \tag{18}$$

It now follows from the identity (14), the inequality (15), and (18) that

$$\begin{aligned}\left\|2\hat{\Psi}_\mu(x,y)-\Theta_\mu(x,y)\right\| &\leq \left\|\Psi_\mu(x,y)\right\|^2 \leq \frac{\left\|\hat{\Psi}_\mu(x,y)\right\|^2}{2(1-\beta)\mu} \leq \frac{\left\|\Theta_\mu(x,y)\right\|^2}{2(1-\beta)\mu}\\ &\leq \frac{\beta^2\mu^2}{2(1-\beta)\mu} = \frac{\beta^2\mu}{2(1-\beta)}.\end{aligned}$$

□

3. The algorithm

We present an algorithm based on the interior point algorithm proposed by Tseng [29]. The global linear convergence and complexity results are stated without proof since these proofs closely parallel those provided by Tseng [29].

Algorithm 1. Choose any $(\beta_1,\beta_2) \in I\!R^2$ satisfying

$$0<\beta_1<\beta_2<1,\ \frac{2\beta_1}{1-\beta_1}<\beta_2,\ \frac{\beta_1^2}{2(1-\beta_1)}+2\beta_1\beta_2+\beta_2^2(1-\beta_1)<\beta_1, \tag{19}$$

and any $(x^0,y^0,\mu^0) \in I\!R^{2n+1}_{++}$ satisfying $\left\|\Theta_{\mu^0}(x^0,y^0)\right\| \leq \beta_1\mu^0$. Let

$$\eta_1 = \frac{\beta_1 - \left[\frac{\beta_1^2}{2(1-\beta_1)}+2\beta_1\beta_2+\beta_2^2(1-\beta_1)\right]}{\sqrt{n}+\beta_1}. \tag{20}$$

For $k = 0, 1, \ldots$, compute $(x^{k+1}, y^{k+1}, \mu^{k+1})$ from (x^k, y^k, μ^k) according to

$$x^{k+1} = x^k + \Delta x^k, y^{k+1} = y^k + \Delta y^k, \mu^{k+1} = (1-\gamma^k)\mu^k, \tag{21}$$

where γ^k is the largest $\gamma \in (0, \eta_1]$ satisfying

$$\left\| 2\hat{\Psi}_{\mu^k}(x^k, y^k) + \gamma X^k (Mx^k - y^k + q) \right\| \leq \beta_2 \left(\mu^k - \left\| \Theta_{\mu^k}(x^k, y^k) \right\| \right), \tag{22}$$

and $(\Delta x^k, \Delta y^k)$ is the unique vector in $I\!R^{2n}$ satisfying

$$M\Delta x^k - \Delta y^k = -\gamma^k (Mx^k - y^k + q), \tag{23}$$

$$Y^k \Delta x^k + X^k \Delta y^k = -2\hat{\Psi}_{\mu^k}(x^k, y^k). \tag{24}$$

Remark 5. To implement the algorithm using the function υ_μ, begin by selecting the parameters β_1 and β_2 so that

$$0 < \beta_1 < \beta_2 < 1, \ \frac{2\beta_1}{1-\beta_1} < \beta_2, \ \frac{\beta_1^2}{(1-\beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1-\beta_1) < \beta_1. \tag{25}$$

Then set

$$\eta_1 = \frac{\beta_1 - \left[\frac{\beta_1^2}{(1-\beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1-\beta_1) \right]}{\sqrt{n} + \beta_1} \tag{26}$$

and replace the function $\hat{\Psi}_{\mu^k}$ in (22) and (24) by the function $\hat{\Upsilon}_{\mu^k}$.

Remark 6. The set of pairs (β_1, β_2) satisfying either (19) or (25) is non–empty. In both cases, it follows that $\eta_1 > 0$. For a choice of β_1 and β_2 satisfying both (19) and (25), take $\beta_1 = 0.09$, $\beta_2 = 0.2$.

The following Theorem shows that if the algorithm is initiated in the positive orthant, then it is well–defined and the iterates remain both in the positive orthant and the set $\{(x, y) \in I\!R^n \times I\!R^n : \|\Theta_\mu(x, y)\| \leq \beta_1\mu\}$ for decreasing values of μ.

Theorem 1. *Fix any $(\beta_1, \beta_2) \in I\!R^2$ satisfying (19). Let η_1 be given by (20). Suppose that $(x^k, y^k, \mu^k) \in I\!R^{2n+1}_{++}$ satisfies $\|\Theta_{\mu^k}(x^k, y^k)\| \leq \beta_1\mu^k$ and $(\Delta x^k, \Delta y^k)$ satisfies (23) and (24), with γ^k being the largest $\gamma \in (0, \eta_1]$ satisfying (22), then $\gamma^k > 0$ exists and*

$$(x^k + \Delta x^k, y^k + \Delta y^k) > 0, \tag{27}$$

$$\left\| \Theta_{(1-\gamma^k)\mu^k}(x^k + \Delta x^k, y^k + \Delta y^k) \right\| \leq \beta_1(1-\gamma^k)\mu^k. \tag{28}$$

Proof. For the sake of simplicity, denote $(x, y, \mu) = (x^k, y^k, \mu^k)$, $(\Delta x, \Delta y) = (\Delta x^k, \Delta y^k)$ and $\gamma = \gamma^k$ respectively. We first establish that $\gamma > 0$ exists. By Proposition 1, $\left\| 2\hat{\Psi}_\mu(x, y) \right\| \leq 2 \|\Theta_\mu(x, y)\| \leq 2\beta_1\mu$. By the choice of β_1 and β_2 in (19), we know that $2\beta_1 < \beta_2(1-\beta_1)$. Therefore,

$$\left\|2\hat{\Psi}_\mu(x,y)\right\| < \beta_2\mu(1-\beta_1) \le \beta_2\left(\mu - \left\|\Theta_\mu(x,y)\right\|\right),$$

which implies that $\gamma > 0$ exists since a strict inequality holds in (22) when $\gamma = 0$.

Next set $r = 2\hat{\Psi}_\mu(x,y)$, $s = Mx - y + q$, and $z = X^{-1}\Delta x$ and $\beta = \left\|\frac{\Theta_\mu(x,y)}{\mu}\right\|$. Then the system (23) and (24) can be rewritten as

$$\begin{aligned} MXz - \Delta y &= -\gamma s, \\ YXz + X\Delta y &= -r. \end{aligned}$$

It follows that

$$(YX + XMX)z = -r - \gamma Xs.$$

Since M is positive semidefinite, we have

$$z^T YXz \le z^T(YX + XMX)z = z^T(-r - \gamma Xs) \le \|z\|\,\|r + \gamma Xs\|, \tag{29}$$

which implies that

$$\|z\| \le \frac{\|r + \gamma Xs\|}{\min_i x_i y_i} \le \frac{\|r + \gamma Xs\|}{\mu(1-\beta)}, \tag{30}$$

where the second inequality follows from the inequality

$$Xy \ge (1-\beta)\mu e, \tag{31}$$

which is itself a consequence of the relation $\left\|\Theta_\mu(x,y)\right\| = \|Xy - \mu e\| \le \beta\mu$. By combining (30) with (22), we find that $\|z\| \le \beta_2 < 1$. Thus, in particular, $e + z > 0$. Let $x' = x + \Delta x$ and $y' = y + \Delta y$. Hence $x' = x + Xz = X(e+z) > 0$, since $x > 0$. From Lemma 1 and (24), for each $i = 1, \ldots, n$, we have

$$\begin{aligned} &|(x_i + (\Delta x)_i)(y_i + (\Delta y)_i) - \mu| \\ &= |x_i y_i - \mu + [x_i(\Delta y)_i + y_i(\Delta x)_i] + (\Delta x)_i(\Delta y)_i| \\ &= \left|2\hat{\psi}_\mu(x_i,y_i) - \frac{\hat{\psi}_\mu^2(x_i,y_i)}{(\frac{x_i+y_i}{\sqrt{2}})^2} - 2\hat{\psi}_\mu(x_i,y_i) + (\Delta x)_i(\Delta y)_i\right| \\ &\le \frac{\hat{\psi}_\mu^2(x_i,y_i)}{\frac{(x_i+y_i)^2}{2}} + |(\Delta x)_i(\Delta y)_i| \\ &\le \frac{\hat{\psi}_\mu^2(x_i,y_i)}{2(1-\beta)\mu} + |(\Delta x)_i(\Delta y)_i|, \end{aligned} \tag{32}$$

where (32) follows from the fact that $\frac{(x_i+y_i)^2}{2} \geq 2x_i y_i \geq 2(1-\beta)\mu$. Therefore,

$$\begin{aligned}
\left\|X'y' - \mu e\right\| &\leq \frac{1}{2(1-\beta)\mu} \left\| \begin{pmatrix} \hat{\psi}_\mu^2(x_1, y_1) \\ \cdots \\ \hat{\psi}_\mu^2(x_n, y_n) \end{pmatrix} \right\| + \|ZX\Delta y\| \quad \text{(by (32))} \\
&\leq \frac{1}{2(1-\beta)\mu} \left\| \begin{pmatrix} \hat{\psi}_\mu^2(x_1, y_1) \\ \cdots \\ \hat{\psi}_\mu^2(x_n, y_n) \end{pmatrix} \right\|_1 + \left\| Z(-2\hat{\Psi}_\mu(x,y) - Y\Delta x) \right\| \quad \text{(by (24))} \\
&\leq \frac{1}{2(1-\beta)\mu} \left\|\hat{\Psi}_\mu(x,y)\right\|^2 + 2\left\|Z\hat{\Psi}_\mu(x,y)\right\| + \|ZYXz\| \\
&\leq \frac{1}{2(1-\beta)\mu} \left\|\hat{\Psi}_\mu(x,y)\right\|^2 + 2\left\|Z\hat{\Psi}_\mu(x,y)\right\|_1 + \|ZYXz\|_1 \\
&\leq \frac{1}{2(1-\beta)\mu} \left\|\hat{\Psi}_\mu(x,y)\right\|^2 + 2\|z\|\left\|\hat{\Psi}_\mu(x,y)\right\| + z^T YXz \\
&\leq \frac{(\beta\mu)^2}{2(1-\beta)\mu} + 2\beta\beta_2\mu + \|z\|\,\|r + \gamma Xs\| \qquad \text{(Proposition 1 and (29))} \\
&\leq \frac{(\beta\mu)^2}{2(1-\beta)\mu} + 2\beta\beta_2\mu + \beta_2\beta_2\mu(1-\beta) \qquad \text{(by (22))} \\
&\leq \frac{\beta_1^2\mu}{2(1-\beta_1)} + 2\beta_1\beta_2\mu + \beta_2^2(1-\beta_1)\mu, \qquad (33)
\end{aligned}$$

where (33) follows from the fact that $\beta \leq \beta_1$ and

$$\begin{aligned}
2\beta\beta_2\mu + \beta_2^2(1-\beta)\mu &= 2\beta\beta_2\mu + \beta_2^2\mu - \beta_2^2\beta\mu \\
&= \beta(2\beta_2 - \beta_2^2)\mu + \beta_2^2\mu \\
&\leq \beta_1(2\beta_2 - \beta_2^2)\mu + \beta_2^2\mu \\
&= 2\beta_1\beta_2\mu + \beta_2^2(1-\beta_1)\mu.
\end{aligned}$$

Therefore, by (19) and (33), $||X'y' - \mu e|| \leq \beta_1\mu$. It follows from $x' > 0$ and $\beta_1 < 1$ that $y' > 0$. The triangle inequality, (33), and the inequality $\gamma < \eta_1$ now imply that

$$\begin{aligned}
\frac{||X'y' - (1-\gamma)\mu e||}{(1-\gamma)\mu} &\leq \frac{||X'y' - \mu e||}{(1-\gamma)\mu} + \frac{\gamma\sqrt{n}}{1-\gamma} \\
&\leq \frac{\frac{\beta_1^2}{2(1-\beta_1)} + 2\beta_1\beta_2 + \beta_2^2(1-\beta_1)}{1-\gamma} + \frac{\gamma\sqrt{n}}{1-\gamma} \\
&\leq \frac{\beta_1 - \eta_1(\sqrt{n} + \beta_1)}{1-\eta_1} + \frac{\eta_1\sqrt{n}}{1-\eta_1} = \beta_1 \,.
\end{aligned}$$

□

The following global linear convergence result is patterned on [29, Theorem 3.1].

Theorem 2. *Let S denote the set of solutions to LCP:*

$$S := \left\{(x, y) \,:\, 0 \le x,\ 0 \le y,\ y = Mx + q,\ \textit{and}\ x^T y = 0\right\},$$

and let β_1, β_2, η_1 and $\{(x^k, y^k, \mu^k, \gamma^k)\}_{k=0,1,\ldots,}$ be generated by the Algorithm of Section 3. Then

$$0 < (x^k, y^k), \tag{34}$$

$$\beta_1 \mu^k \ge \left\| \Theta_{\mu^k}(x^k, y^k) \right\|, \quad \textit{and} \tag{35}$$

$$\frac{\mu^k}{\mu^0}(Mx^0 - y^0 + q) = Mx^k - y^k + q, \tag{36}$$

for all k, where for $k > 0$

$$\mu^k = (1 - \gamma^{k-1}) \ldots (1 - \gamma^0)\mu^0. \tag{37}$$

Moreover, the sequence $\{(x^k, y^k)\}$ is bounded if and only if the solution set S is non–empty, in which case, for any $(x^*, y^*) \in S$, we have $\gamma^k \ge \min\{\eta_1, \eta_2\}$ for all k, where

$$\eta_2 = \begin{cases} \frac{[\beta_2(1-\beta_1)-2\beta_1]\mu^0 \min_i y_i^0}{[(1+\beta_1)n\mu^0+(x^0)^T y^0+(x^*)^T y^0+(x^0)^T y^*]\left\|Mx^0-y^0+q\right\|_\infty} & \textit{if}\ Mx^0 - y^0 + q \ne 0, \\ \infty & \textit{if}\ Mx^0 - y^0 + q = 0. \end{cases} \tag{38}$$

Thus, if S is nonempty, the Algorithm of Section 3 forces μ^k to zero at a global linear rate with the convergence ratio less than $1 - \min\{\eta_1, \eta_2\}$. Therefore, by standard results in the interior point literature (e.g., see [23]), one can find an element of S in $O((\min\{\eta_1, \eta_2\})^{-1}L)$ iterations, where L denotes the size of the binary encoding of the problem. It is easily seen that $\eta_1^{-1} = O(\sqrt{n})$, so it only remains to estimate η_2^{-1}. In the case where (x^0, y^0, μ^0) is chosen so that η_2^{-1} is $O(\sqrt{n})$ (such as when $Mx^0 - y^0 + q = 0$), the iteration count is $O(\sqrt{n}L)$. In the case where (x^0, y^0, μ^0) is the standard choice

$$x^0 = \rho_p e,\ y^0 = \rho_d e,\ \mu^0 = \rho_p \rho_d,$$

$$\rho_p \ge \frac{||x^*||_1}{n},\ \rho_d \ge \max\left\{\frac{||y^*||_1}{n}, ||\rho_p Me + q||_\infty\right\},$$

where (x^*, y^*) is any element of S, the formula (38) yields

$$\eta_2^{-1} \le \frac{3(4 + \beta_1)n}{\beta_2(1 - \beta_1) - 2\beta_1},$$

so the iteration count is $O(nL)$.

4. Concluding remarks

In Section 3, we present the first rate of convergence result and the first complexity result of any kind for a path following algorithm based on the Chen–Harker–Kanzow smoothing techniques. In the year following the announcement of this result there has been a flurry of activity on rate of convergence results for non-interior path following and smoothing methods for the complementarity problems and variational inequalities [1,3,4,6,28,32,33]. All of this work builds on new neighborhood concepts [1,19] for smoothing paths (e.g. the *central* path) that do not necessarily lie in the positive orthant. The first global linear convergence result for non-interior path following methods appears in [1]. The work in [3,4,6,32,33] builds on the ideas presented in [1] and [19]. In [32], Xu establishes the global linear convergence result for nonlinear complementarity problems. In [3,4,6] the authors extend the analysis to larger classes of smoothing functions [7,17] and, in addition, establish the local quadratic or super–linear convergence of their methods. In [28], the authors build on the approach developed in [19] and establish the global linear convergence or the local super–linear convergence of their method depending on the choice of parameters.

The interior point path following method studied in this paper is essentially a variation on standard interior point methods wherein the right hand side in the Newton equations is perturbed in a very special way. For this reason, it is possible to analyze the algorithm within the framework developed by Mizuno. In [24,25], Mizuno proposed a class of feasible interior point algorithms for monotone LCP which are based on the search direction $(\Delta x^k, \Delta y^k)$ satisfying the following equations

$$M\Delta x - \Delta y = 0, \tag{39}$$
$$Y^k\Delta x + X^k\Delta y = v^k - \sigma X^k y^k, \tag{40}$$

where $v^k \in I\!R^n_{++}$ and $\sigma > 0$. By adjusting the choice of the sequences $\{v^k\}$ with $\sigma = 1$, Mizuno is able to construct both path following and potential reduction methods and thereby provides a unifying framework within which a number of interior point methods can be studied. In order for this program to work, one must first show that the sequence $\{v^k\}$ satisfies the following three properties:

(A) $v^k > 0$ for $k = 0, 1, \ldots$.
(B) the sequence $\{v^k\}$ is an α–*sequence* for some $\alpha \geq 0$, that is, $v^{k+1} \in \mathcal{N}(v^k, \alpha)$ for all $k = 0, 1, 2, \ldots$, where $\mathcal{N}(v, \alpha) = \{u \in I\!R^n : \left\| V^{-0.5}(v - u)\right\| \leq \alpha\sqrt{v_{\min}}\}$, with $V = \mathrm{diag}\ (v)$, and
(C) there is an iteration index $m = O(\sqrt{n}L)$ such that $0 \leq v^m \leq 2^{-2L+1}e$.

A referee for this paper has observed that the algorithm of Section 3 can be cast within Mizuno's framework. To see this, define

$$v^k = 2X^k y^k - 2\hat{\Psi}_{\mu^k}(x^k, y^k). \tag{41}$$

With this definition, the Newton equations (39) and (40) are identical to the equations (11) and (12) when $\sigma = 2$. If one now assumes that $\beta \in (0, \frac{1}{2}]$ and $(\mu^k - \mu^{k+1})/\mu^k =$

$O(1/\sqrt{n})$, it can be shown that the sequence $\{v^k\}$ defined by (41) satisfies the conditions (A) and (B). The condition $\beta \in (0, \frac{1}{2}]$ can be enforced during the initialization phase of the algorithm. It is used to show that

$$\frac{1}{2}|\theta_\mu(x_i, y_i)| \leq |\hat{\psi}_\mu(x_i, y_i)| \leq |\theta_\mu(x_i, y_i)|, \text{ for } i = 1, 2, \ldots, n$$

whenever $(x, y) \in I\!R^n_{++} \times I\!R^n_{++}$ and

$$\left\| \Theta_\mu(x, y) \right\| \leq \beta\mu,$$

which in turn shows that condition (A) is satisfied. The bounds $\min\{\eta_1, \eta_2\} \leq \gamma^k \leq \eta_1$ (Theorem 2 and (20)) show that $O(1/\sqrt{n}) = \gamma^k$ if (x^0, y^0, μ^0) is chosen so that $\eta_2 = O(1/\sqrt{n})$ (for example, when $y^0 = Mx^0 + q$). This in turn implies that the condition $(\mu^k - \mu^{k+1})/\mu^k = O(1/\sqrt{n})$ is also satisfied. Finally, condition (C) can be verified using the complexity result established in this paper. This connection to Mizuno's work should provide a basis for developing a deeper understanding of the relationship between standard path following methods, potential reduction methods, and the path following method proposed in this paper.

Acknowledgements. We thank Bintong Chen for observing the identity (13) which greatly simplified the original proof of Lemma 1. In addition, we wish to express our gratitude to an anonymous referee for providing the details of the relationship between the rescaled Newton direction and Mizuno's convergence framework.

References

1. Burke, J., Xu, S. (1996): The global linear convergence of a non-interior path-following algorithm for linear complementarity problem. Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195. December 1996
2. Chen, B. (1997): Personal communication. January 1997
3. Chen, B., Chen, X. (1997): 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
4. Chen, B., Chen, X. (1997): 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. To appear in SIAM J. Optim.
5. Chen, B., Harker, P.T. (1993): A non–interior–point continuation method for linear complementarity problems. SIAM J. Matrix Anal. Appl. **14**, 1168–1190
6. Chen, B., Xiu, N. (1997): 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. To appear in SIAM J. Optim.
7. Chen, C., Mangasarian, O.L. (1996): A class of smoothing functions for nonlinear and mixed complementarity problems. Comput. Optim. Appl. **5**, 97–138
8. De Luca, T., Facchinei, F., Kanzow, C. (1996): A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 406–439
9. Facchinei, F., Jiang, H., Qi, L. (1996): A smoothing method for mathematical programs with equilibrium constraints. Preprint, School of Mathematis, University of New South Wales, Sydney, Australia. To appear in Math. Program.
10. Facchinei, F., Soares, J. (1995): Testing a new class of algorithms for nonlinear complementarity problems. In: Giannessi, F., Maugeri, A., eds., Variational Inequalities and Network Equilibrium Problems, pp. 69–83. Plenum Press, New York

11. Facchinei, F., Soares, J. (1997): A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. **7**, 225–247
12. Fischer, A. (1992): A special Newton–type optimization method. Optimization **24**, 269–284
13. Fischer, A. (1995): A Newton–type method for positive semi–definite linear complementarity problems. J. Optimization Theory Appl. **86**, 585–608
14. Fischer, A. (1995): An NCP–function and its use for the solution of complementarity problems. In: Du, D.-Z., Qi, L., Womersley, R.S., eds., Recent Advances in Nonsmooth Optimization, pp. 88–105. World Scientific Publishers, Singapore
15. Fischer, A. (1995): On the local super–linear convergence of a Newton–type method for LCP under weak conditions. Optimization Methods Software **6**, 83–107
16. Fukushima, M., Luo, Z.-Q., Pang, J.-S. (1998): A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. Comput. Optim. Appl. **10**, 5–34
17. Gabriel, S.A., Moré, J.J. (1997): Smoothing of mixed complementarity problems. In: Ferris, M.C., Pang, J.S., eds., Complementarity and Variational Problems: State of the Art, pp. 105–116. SIAM, Philadelphia, PA
18. Harker, P.T., Pang, J.-S. (1990): A damped Newton method for the linear complementarity problem. In: Allgower, E.L., Georg, K., eds., Computational Solution of Nonlinear Systems of Equations, pp. 265–284. Lectures in Applied Mathematics, Vol. 26, AMS, Providence, Rhode Island
19. Hotta, K., Yoshise, A. (1996): 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. To appear in Math. Program.
20. Kanzow, C. (1996): Global convergence properties of some iterative methods for linear complementarity problems. SIAM J. Optim. **6**, 326–341
21. Kanzow, C. (1996): Some noninterior continuation methods for linear complementarity problems. SIAM J. Matrix Anal. Appl. **17**, 851–868
22. Kanzow, C., Jiang, H. (1998): A continuation method for (strongly) monotone variational inequalities. Math. Program. **81**, 103–125
23. Kojima, M., Meggido, N., Noma, T., Yoshise, A. (1991): A unified approach to interior point algorithms for linear complementarity problems. Springer, Berlin
24. Mizuno, S. (1990): An O(n^3L) algorithm using a sequence for a linear complementarity problem. J. Oper. Res. Soc. Japan **33**, 66–75
25. Mizuno, S. (1992): A new polynomial time method for a linear complementarity problem. Math. Program. **56**, 31–43
26. Pang, J.-S. (1990): Newton's method for B–differentiable equations. Math. Oper. Res. **15**, 311–341
27. Pang, J.-S. (1991): A B–differentiable equation–based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity, and variational inequality problems. Math. Program. **51**, 101–131
28. Qi, L., Sun, D. (1997): 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. To appear in Math. Comp.
29. Tseng, P. (1994): Simplified analysis of an $O(nL)$–iteration infeasible predictor–corrector path–following method for monotone LCP. In: Agarwal, R.P., ed., Recent Trends in Optimization Theory and Applications, pp. 423–434. World Scientific Press, Singapore
30. Tseng, P. (1996): Growth behavior of a class of merit functions for the nonlinear complementarity problem. J. Optimization Theory Appl. **89**, 17–37
31. Wright, S.J. (1993): A path–following infeasible–interior–point algorithm for linear complementarity problems. Optimization Methods Software **2**, 79–106
32. Xu, S. (1996): 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
33. Xu, S. (1997): 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
34. Ye, Y., Anstreicher, K. (1993): On quadratic and $O(\sqrt{n}L)$–iteration convergence of a predictor–corrector algorithm for LCP. Math. Program. **62**, 537–551
35. Zhang, Y. (1994): On the convergence of a class of infeasible interior point algorithms for the horizontal linear complementarity problem. SIAM J. Optim., **4**, 208–227