Interior Point Methods for Nonlinear Complementarity Problems

Florian A. Potra[*] Yinyu Ye[†]

July 1991, revised July 1994

Abstract

We generalize several results for linear and nonlinear optimization to nonlinear complementarity problems (NCP) using interior-point algorithms. In particular, we discuss the global and local convergence rates, convergence toward a maximal complementarity solution, a criterion for using Newton's method, and the complexity of a class of NCPs.

Key words: Complementarity problem, interior-point method, convergence rate.
Abbreviated title: Interior-point methods for complementarity problems
AMS(MOS) classification: 90C33

[*]Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA

[†]Department of Management Sciences, The University of Iowa, Iowa City, IA, USA. Research supported in part by NSF Grant DDM-8922636.

1 Introduction

This paper attempts to generalize several new results that have been obtained in the ongoing research area of interior-point algorithms for linear and nonlinear optimization. The first type of results show that some convex nonlinear optimization problems (e.g., den Hertog et al. [3], Jarre [4], Monteiro and Adler [15], and Nesterov and Nemirovsky [16], satisfying certain Lipschitz conditions can be solved by polynomial-time interior-point methods with a global linear convergence rate. For details, see [16].

The second type of results combine the polynomial interior-point algorithm and Newton's method to achieve both global linear and local quadratic convergence rates for linear and nonlinear optimization. Such results are obtained for linear programming (Zhang and Tapia [20]) and entropy problems (Potra and Ye [17]). In the latter paper a checkable criterion is developed that makes it possible to switch to Newton's method at a certain stage of the interior-point algorithm, so that the quadratic convergence rate prevails. Such a hybrid algorithm is obviously desirable.

The third result is the convergence toward a maximal complementary solution when polynomial interior-point algorithms are applied to linear and nonlinear optimization problems (Güler and Ye [2]). A maximal complementary solution is a complementary solution in which the number of positive components is maximized. Such results were later used to develop a criterion to identify the optimal facet for linear programming and to prove the finite convergence of polynomial interior-point algorithms for linear programming with real data ([19]).

Interior-point algorithms for the complementarity problem (CP), a somehow more fundamental mathematical problem, are extensively studied by many researchers, starting with Megiddo [13] who developed the central path theory for the complementarity problem. In fact, part of the theory was established earlier by McLinden [12]. Then, Kojima, Megiddo, Mizuno, Noma and Yoshise developed several important algorithms for solving some CPs [5, 7, 8, 9, 10, 6]. A complete set of results can be found in the monograph [6].

We feel that there is a need to generalize the results from linear and nonlinear optimization problems to CPs. For example, although a global convergent algorithm is established by Kojima et al. for general monotone NCP, its convergence rate is problem dependent. Furthermore, since NCP is basically a system of (sign-restricted) nonlinear equations, an asymptotic superlinear or quadratic convergence rate is absolutely essential for good practical performance. The convergence towards a maximal complementary solution and the finite convergence property are also important for convex LCP with general data. Consider the monotone LCP: once the index set of the positive components at the maximal complementary solution is identified, the (exact) solution can be found by solving a system of linear equations or least squares. These facts motivated us to write the present paper. We mention that while some of these generalizations are straightforward, others are not trivial at all. More specifically, in Section 2 we will discuss a primal-dual interior-point algorithm for a class of monotone NCPs. Polynomial complexity is obtained in Section 3 for problems satisfying the so-called scaled Lipschitz condition. A symmetric version of this condition was proposed by Zhu [21], but our scaled condition assumes no symmetry on the Jacobian matrix. Then, in Section 3 we prove a global linear convergence rate of the algorithm for solving this class of problems. In Section 4 we analyze the property of convergence towards a maximal complementary solution of the primal-dual algorithm for monotone NCPs. In Section 5 we show that under some restrictive conditions it is possible to switch from the potential reduction method to Newton's method. The resulting hybrid algorithm has both polynomial complexity and local quadratic convergence. Because of the restrictive conditions such as the assumption that the Jacobians are invertible and smooth if scaled, the results of Section 5 do no longer apply to general LP, QP and LCP. However they apply to linearly constrained entrophy optimization problems having interior optimal solutions. We also discuss the quadratic convergence rate and the complexity of this hybrid algorithm in Section 6. For monotone LCPs, this approach implies finite convergence. Throughout this paper, for any vector x we denote by X the diagonal matrix having the same entries on the diagonal as the vector x. The symbol $\| \ \|$ denotes the Euclidian norm in $\mathbb{R}^n$ as well as the corresponding operator (matrix) norm, i.e.,

$$\|M\| = \sup\{\|Mx\| : \|x\| = 1\}.$$

2 A potential reduction algorithm for nonlinear complementarity problems with monotone operators

The nonlinear complementarity problem considered in this paper is defined as

$$\text{solve} \quad x^T \ F(x) = 0 \tag{2.1}$$

$$\text{subject to} \quad x \in \Omega := \{x \in \mathbb{R}^n : x \geq 0, \ F(x) \geq 0\}, \tag{2.2}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable nonlinear operator. We assume that F is monotone, i.e.:

$$(x - \overline{x})^T \ (F(x) - F(\overline{x})) \geq 0 \ \ \forall x, \overline{x} > 0 \in \mathbb{R}^n. \tag{2.3}$$

It is easily seen that this implies

$$h^T \ F'(x)h \geq 0 \qquad \forall \ x > 0, h \in \mathbb{R}^n, \tag{2.4}$$

where $F'(x)$ denotes the Jacobian matrix of $F(x)$. We also assume that the set of strictly feasible points,

$$\Omega^\circ := \{x \in I\!R^n : \; x > 0, \; F(x) > 0\} \tag{2.5}$$

is nonempty. Here $x > 0$ means that all components of the vector x are (strictly) positive. Under these assumptions it follows that the solution set

$$\mathcal{T} := \{(x, s) : x \in I\!R^n, \; x \geq 0, \; s = F(x) \geq 0, \; x^T s = 0\} \tag{2.6}$$

is a nonempty, compact, convex set (e.g., see Güler [1], McLinden [12], and monograph [6]). Starting with a point $x^\circ \in \Omega^\circ$ we would like to compute a sequence of points $\{x^k\}$ such that

$$x^k \in \Omega^\circ, \; k = 0, 1, \ldots, \quad \lim_{k \to \infty} (x^k)^T \; F(x^k) = 0. \tag{2.7}$$

In what follows we will describe an interior point primal-dual algorithm (Kojima et al. [9, 10], Kortanek et al. [11], and Monteiro and Adler [14, 15] that may be used to this effect. The natural "dual variable" is $s = F(x)$, and at each step of our primal-dual algorithm we will achieve a reduction of the following primal-dual potential function

$$\phi_\rho(x, s) = \rho \log(x^T s) - \sum_{i=1}^{n} \log(x_i s_i), \tag{2.8}$$

where $n \leq \rho \leq 2n$, and $x = [x_1, x_2, \ldots, x_n]^T > 0$, $s = [s_1, s_2, \ldots, s_n]^T > 0$. This potential function was introduced by Todd and Ye [18] and it satisfies the following inequality

$$(\rho - n) \log(x^T s) \leq \phi_\rho(x, s) - n \log n. \tag{2.9}$$

It follows that

$$x^T s \leq \epsilon \quad \text{whenever} \quad \phi_\rho(x, s) \leq n \log n - (\rho - n) |\log \epsilon|.$$

Therefore if the potential function is reduced by at least a positive constant at each step then we will have $x^T s \leq \epsilon$ after at most $O(\phi_\rho(x^o, s^o) - n \log n + (\rho - n) |\log \epsilon|)$ steps.

At step k of our potential reduction algorithm we start with an $x^k \in \Omega^\circ$ and consider the nonlinear equation

$$H_k(u) := X^k(F(x^k + u) - F(x^k)) + S^k u - \theta_k p^k = 0 \tag{2.10a}$$

where

$$s^k = F(x^k), \quad e = [1, 1, \ldots, 1]^T, \tag{2.10b}$$

$$p_k = \frac{(x^k)^T s^k}{\rho} \; e - X^k S^k e, \tag{2.10c}$$

$$\theta_k = \frac{\beta \min_{1 \leq i \leq n} \sqrt{x_i^k \; s_i^k}}{\|(X^k S^k)^{-1/2} \; p^k\|}. \tag{2.10d}$$

In general we cannot compute an exact solution of this nonlinear equation, so that in our potential reduction algorithm we only require that an approximate solution $u = \Delta x$ of (2.10) is found such that

$$\|H_k(\Delta x)\| \le C\beta^2 \min_{1\le i\le n} (x_i^k s_i^k). \tag{2.11}$$

We assume that the positive constants β and C are chosen such that

$$\beta(1 + C\beta) \le \frac{1}{2}, \quad \beta < 1/C, \tag{2.12a}$$

$$\gamma := \frac{\sqrt{3}}{2}\beta(1 - C\beta) - \beta^2(1 + C\beta)^2 > 0. \tag{2.12b}$$

With the above relation our potential reduction algorithm can be described as follows.

Algorithm 2.1
Initialization.

A1 Choose $x^0 \in \Omega^\circ$ *and constants* $\rho \in [n+\sqrt{n}, 2n]$, $\beta \in (0,1), C > 0$ *such that inequalities (2.12) are satisfied.*

Iteration. *For* $k = 0, 1, 2, \ldots,$

A2 find an approximate solution $u = \Delta x$ *of (2.10) such that (2.11) is satisfied;*

A3 set

$$x^{k+1} = x^k + \Delta x, \quad s^{k+1} = F(x^{k+1}), \quad \Delta s = s^{k+1} - s^k \tag{2.13}$$

Theorem 2.2 *Let* $F : I\!\!R^n \to I\!\!R^n$ *be a monotone operator. Then for any* $x^0 \in \Omega^\circ$, *Algorithm 2.1 is well defined and for each* $k = 0, 1, \ldots,$ *we have* $x^{k+1} \in \Omega^\circ$ *and*

$$\phi_\rho(x^{k+1}, s^{k+1}) \le \phi_\rho(x^k, s^k) - \gamma. \tag{2.14}$$

Proof. To simplify formulae we drop the index k, and let $x^+ = x + \Delta x$, $z = H_k(\Delta x)$ and $s^+ = s + \Delta s = F(x^+)$. Let $D = (XS)^{1/2}$. Note that

$$\|D^{-1}\| = 1/\min\sqrt{x_j s_j}, \quad \theta = \frac{\beta}{\|D^{-1}\|\|D^{-1}p\|}, \tag{2.15}$$

$$D^{-1}p = \frac{x^T s}{\rho} D^{-1}e - De \quad . \tag{2.16}$$

Moreover we have

$$\frac{\rho}{x^T s}\|D^{-1}p\| = \|D^{-1}e - \tfrac{\rho}{\|De\|^2}De\| \geq \frac{\sqrt{3}}{2}\|D^{-1}\| \quad \text{for } \rho \geq n + \sqrt{n} \tag{2.17}$$

(see Kojima et al. [10]). The following standard result is frequently used in interior-point algorithms (e.g., Kojima et al. [10]). If

$$\max(\|X^{-1}\ \Delta x\|_\infty, \|S^{-1}\ \Delta s\|_\infty) \leq 1/2 \quad , \tag{2.18}$$

then

$$\begin{aligned} & \phi_\rho(x^+, s^+) - \phi_\rho(x, s) \\ \leq\ & \frac{\rho}{x^T s}(x^T \Delta s + s^T \Delta x + \Delta x^T \Delta s) \quad - \quad e^T(X^{-1}\Delta x + S^{-1}\Delta s) + \delta^2 \ , \end{aligned} \tag{2.19a}$$

where

$$\delta^2 = \|X^{-1}\Delta x\|^2 \quad + \quad \|S^{-1}\Delta s\|^2 \quad . \tag{2.19b}$$

First we have

$$X\Delta s + S\Delta x = \theta p + z = \theta(p+q), \tag{2.20}$$

where $q = z/\theta$. From (2.11) and (2.16):

$$\begin{aligned} \|D^{-1}\| &\leq \|D^{-1}\|\|q\| = \frac{\|D^{-1}\|}{\theta}\|z\| \\ &\leq \frac{\|D^{-1}\|}{\theta}\frac{C\beta^2}{\|D^{-1}\|^2} = C\beta\|D^{-1}p\|. \end{aligned} \tag{2.21}$$

Thus,

$$\begin{aligned} \delta^2 &= \|D^{-2}X\Delta s\|^2 + \|D^{-2}S\Delta x\|^2 \\ &\leq \|D^{-1}\|^2(\|D^{-1}X\Delta s\|^2 + \|D^{-1}S\Delta x\|^2) \\ &= \|D^{-1}\|^2(\|D^{-1}(X\Delta s + S\Delta x)\|^2 - 2\Delta x^T\Delta s) \\ &= \|D^{-1}\|^2(\theta^2\|D^{-1}(p+q)\|^2 - 2\Delta x^T\Delta s) \quad \text{from (2.20)} \\ &\leq \|D^{-1}\|^2(\theta^2(\|D^{-1}p\| + \|D^{-1}q\|)^2 - 2\Delta x^T\Delta s) \\ &\leq \|D^{-1}\|^2(\theta^2\|D^{-1}p\|^2(1+C\beta)^2 - 2\Delta x^T\Delta s) \quad \text{from (2.21)} \\ &= \theta^2\|D^{-1}\|^2\|D^{-1}p\|^2(1+C\beta)^2 - 2\|D^{-1}\|^2\Delta x^T\Delta s \\ &= \beta^2(1+C\beta)^2 - 2\|D^{-1}\|^2\Delta x^T\Delta s \ , \quad \text{from (2.16)}. \end{aligned} \tag{2.22}$$

The monotony of $F(.)$ implies

$$\Delta x^T \Delta s \geq 0 \quad . \tag{2.23}$$

Therefore, we deduce that from (2.12) and (2.22)

$$\delta \leq \beta(1 + C\beta) \leq 1/2 \ . \tag{2.24}$$

This also implies that condition (2.18) holds, so that inequality (2.19a) will be satisfied. Moreover (2.18) shows that $x^+ \in \Omega^\circ$. Now we can write

$$\begin{aligned}
& \frac{\rho}{x^T s}(x^T \Delta s + s^T \Delta x) - e^T(X^{-1}\Delta x + S^{-1}\Delta s) \\
= \ & \frac{\rho}{x^T s} e^T((X\Delta s + S\Delta x) - \frac{x^T s}{\rho} D^{-2}(X\Delta s + S\Delta x)) \\
= \ & \frac{\rho}{x^T s} e^T(D - \frac{x^T s}{\rho} D^{-1})D^{-1}(X\Delta s + S\Delta x) \\
= \ & -\frac{\rho}{x^T s} p^T D^{-1} D^{-1}(X\Delta s + S\Delta x) \quad \text{from (2.17)} \\
= \ & -\frac{\rho\theta}{x^T s} p^T D^{-1} D^{-1}(p + q) \quad \text{from (2.20)} \\
= \ & -\frac{\rho\theta}{x^T s}(\|D^{-1}p\|^2 + p^T D^{-1} D^{-1} q) \\
\leq \ & -\frac{\rho\theta}{x^T s}(\|D^{-1}p\|^2 - \|D^{-1}p\|\|D^{-1}q\|) \\
\leq \ & -\frac{\rho\theta}{x^T s}(\|D^{-1}p\|^2 - C\beta\|D^{-1}p\|^2) \quad \text{from (2.21)} \\
= \ & -\frac{\rho\theta}{x^T s}(1 - C\beta)\|D^{-1}p\|^2 \\
\leq \ & -\frac{\sqrt{3}}{2}(1 - C\beta)\theta\|D^{-1}\|\|D^{-1}p\|, \quad \text{from (2.17)} \\
= \ & -\frac{\sqrt{3}}{2}(1 - C\beta)\beta, \quad \text{from (2.16).}
\end{aligned} \tag{2.25}$$

Finally, from (2.19), (2.22) and (2.25) we obtain:

$$\begin{aligned}
& \phi(x^+, s^+) - \phi(x, s) \\
\leq \ & -\frac{\sqrt{3}}{2}(1 - C\beta)\beta + \frac{\rho}{x^T s}\Delta x^T \Delta s + \delta^2
\end{aligned}$$

$$
\begin{aligned}
&\leq -\frac{\sqrt{3}}{2}\beta(1-C\beta)+\beta^2(1+C\beta)^2+(\frac{\rho}{x^Ts}-2\|D^{-1}\|^2)\Delta x^T\Delta s \\
&= -\frac{\sqrt{3}}{2}\beta(1-C\beta)+\beta^2(1+C\beta)^2+(\frac{\rho}{x^Ts}-\frac{2}{\min(x_js_j)})\Delta x^T\Delta s.
\end{aligned}
\tag{2.26}
$$

Since $x^Ts \geq n\min(x_js_j)$, we deduce that for any $\rho \leq 2n$ we have

$$
\frac{2}{\min(x_js_j)} \geq \frac{2n}{x^Ts} \geq \frac{\rho}{x^Ts} ,
$$

which indicates that the last term in (2.26) is nonpositive. Hence,

$$
\phi_\rho(x^+,s^+)-\phi_\rho(x,s) \leq -\frac{\sqrt{3}}{2}\beta(1-C\beta)+\beta^2(1+C\beta)^2 .
$$

□

The above theorem was developed for solving some nonlinear optimization problems (Potra and Ye [17]), and it represents a slight improvement of Theorem 3.2 of Kortanek et al. [11].

Finding a Δx that satisfies (2.11) means finding an approximate solution of equation (2.10). If F is linear then (2.10) can be solved exactly and in this case Theorem 2.1, with $C = 0$, reduces to a result of Kojima et al. [10] for the linear complementarity problem. In Section 3 we will see that if F satisfies a scaled Lipschitz condition then by applying one Newton step to (2.10) with starting point $u = 0$ we obtain a Δx that satisfies (2.11).

3 Complexity results for monotone operators satisfyingy a scaled Lipschitz condition

In what follows we assume that there is a monotone increasing function

$$
\psi : (0,1) \to (1,\infty) \tag{3.1a}
$$

such that the following scaled Lipschitz condition is satisfied:

$$
\|X(F(x+h)-F(x)-F'(x)h)\| \leq \psi(\alpha)h^TF'(x)h \tag{3.1b}
$$

whenever

$$
h \in \mathbb{R}^n, \quad x \in \mathbb{R}^n, \quad x > 0, \quad \|X^{-1}h\| \leq \alpha < 1 \tag{3.1c}
$$

We mention that such a condition was considered in the context of convex programming by Zhu [21]. A class of entrophy problems satisfying this condition is described in [17]. Let Δx be obtained by applying one Newton step to (2.10) with starting point $u = 0$, i.e.,

$$X^k F'(x^k)\Delta x + S^k \Delta x = \theta_k p^k. \tag{3.2}$$

The following lemma can be easily proved by using the techniques in Theorem 3.2 of [11].

Lemma 3.1 *Let F be a continuously differentiable monotone operator. Let $x^k \in \Omega^\circ$ and consider the notation in (2.10). If Δx is the solution of the linear system (3.2), then*

$$\|(X^k)^{-1}\Delta x\| \leq \beta.$$

In the next theorem we show that Δx defined as above satisfies condition (2.11) from Theorem 2.2.

Theorem 3.2 *Under the hypothesis of Lemma 3.1 assume that condition (3.1) is satisfied and denote $z^k = H_k(\Delta x)$. Then*

$$\|z^k\| \leq \frac{\psi(\beta)}{4}\beta^2 \min_{1\leq i\leq n} (x_i^k s_i^k) .$$

Proof. From (2.10) and (3.2) we have

$$z^k = X^k(F(x^k + \Delta x) - F(x^k) - F'(x^k)\Delta x),$$

so that by virtue of Lemma 3.1 and condition (3.1) we obtain

$$\|z^k\| \leq \psi(\beta)\Delta x^T F'(x^k)\Delta x. \tag{3.3}$$

From (3.2) and (2.10d) we have

$$\begin{aligned} \Delta x^T F'(x^k)\Delta x &= \Delta x^T S^k D^{-2} A\theta_k p^k - \|S^k D^{-1}\Delta x\|^2 \\ &\leq \|\theta_k D^{-1} p^k\|\|S^k D^{-1}\Delta x\| - \|S^k D^{-1}\Delta x\|^2 \\ &\leq \theta_k^2 \|D^{-1}p^k\|^2/4 = \beta^2 \min_{1\leq i\leq n} (x_i^k s_i^k)/4 \quad , \end{aligned}$$

and by using (3.3) we obtain the required inequality. □

Combining the results of Theorems 2.2 and 3.2 we can easily prove that the following algorithm is globally convergent and has polynomial complexity.

We are now ready to formally describe our algorithm. We will initialize the algorithm by taking $C = .25\psi(\beta)$ in (2.12). We note that finding a β such that (2.12) is satisfied requires eventually the analytic expression of the fuction $\psi(\beta)$. In case of entropy minimization problems the form of ψ is extremely simple and this poses no problems.

Algorithm 3.3
Initialization.

B1 Choose positive constants ρ, β such that

$$
\begin{aligned}
& n+\sqrt{n} \leq \rho \leq 2n, & \text{(3.4a)}\\
\gamma_1 : &= \beta(1+0.25\beta\psi(\beta)) \leq 0.5, & \text{(3.4b)}\\
\gamma_2 : &= 4-\beta\psi(\beta) > 0\ , & \text{(3.4c)}\\
\gamma : &= \frac{\sqrt{3}}{8}\beta\gamma_2-\gamma_1^2 > 0\ . & \text{(3.4d)}
\end{aligned}
$$

B2 Choose starting point $x^0 \in \Omega^\circ$.

Iteration. *For $k = 0, 1, 2, \ldots,$*

B3 solve linear system (3.2);

B4 set $x^{k+1} = x^k + \Delta x, \quad s^{k+1} = F(x^{k+1})$.

Corollary 3.4 *Algorithm 3.3 is well defined and*

$$x^k \in \Omega^\circ, \quad k = 0, 1, \ldots. \tag{3.5}$$

Moreover, the inequality $(x^k)^T s^k < \epsilon$ will be reached in at most

$$O(\phi_\rho(x^0, s^0) - n\log n + (\rho - n)|\log\epsilon|)$$

steps.

4 Maximal complementary solution

Let $\mathcal{X}, \mathcal{S}$ be the projections of the solution set $\mathcal{T}$ onto the primal and dual space respectively, i.e.,

$$\mathcal{X} = \{x \in \Omega : x^T F(x) = 0\}, \tag{4.1}$$

$$\mathcal{S} = \{s \in I\!R^n : s = F(x) \text{ for some } \quad x \in \mathcal{X}\}. \tag{4.2}$$

Recall that if F is monotone and Ω° is nonempty then $\mathcal{T}, \mathcal{X}, \mathcal{S}$ are nonempty, convex, and compact sets. For any $x \in \mathcal{X}$ let us define

$$\sigma(x) = \{i \in \{1, 2, \ldots, n\} : x_i > 0\}. \tag{4.3}$$

We say that $u \in \mathcal{X}$ is a maximal solution if $x \in \mathcal{X}$ and $\sigma(x) \supset \sigma(u)$ implies $\sigma(x) = \sigma(u)$ (see Güler and Ye [2]). $\mathcal{X}$ being nonempty, convex and compact clearly contains a maximal solution. It is easily shown that if $u, v \in \mathcal{X}$ are maximal solutions then $\sigma(u) = \sigma(v)$. Therefore we may define $\sigma(\mathcal{X}) = \sigma(u)$, where $u \in \mathcal{X}$ is a maximal solution. $\sigma(\mathcal{S})$ is defined in a similar manner. Now let (x^*, s^*) be maximal in $\mathcal{T}$. Then x^* is maximal in $\mathcal{X}$ and s^* is maximal in $\mathcal{S}$, i.e., $\sigma(x^*) = \sigma(\mathcal{X})$ and $\sigma(s^*) = \sigma(\mathcal{S})$. Indeed let $x \in \mathcal{X}$ be such that $\sigma(x) \supset \sigma(x^*)$, and denote $s = F(x)$. Because of the convexity of $\mathcal{T}$, $((x + x^*)/2,\ (s + s^*)/2) \in \mathcal{T}$, and we have obviously $\sigma((x + x^*)/2) = \sigma(x) \cup \sigma(x^*) = \sigma(x) \supset \sigma(x^*)$ and $\sigma((s + s^*)/2) \supset \sigma(s^*)$. From the maximality of (x^*, s^*) in $\mathcal{T}$ it follows that $\sigma(x^*) = \sigma((x + x^*)/2)$, $\sigma(s^*) = \sigma((s + s^*)/2)$. The first relation shows that $\sigma(x^*) = \sigma(x)$ so that x^* is maximal in $\mathcal{X}$. Similarly we show that s^* is maximal in $\mathcal{S}$.

If (x^*, s^*) is maximal in $\mathcal{T}$ we say that (x^*, s^*) is a maximal complementary solution. For any $(x, s) \in \mathcal{T}$ we have $\sigma(x) \cap \sigma(s) = \phi$, so that in particular $\sigma(\mathcal{X}) \cap \sigma(\mathcal{S}) = \phi$. A pair $(x, s) \in \mathcal{T}$ satisfying $\sigma(x) \cup \sigma(s) = \{1, 2, \ldots, n\}$ is called a strictly complementary solution. Every strictly complementary solution is a maximal complementary solution. Indeed let (x^*, s^*) be a strictly complementary solution and suppose there is $(x, s) \in \mathcal{T}$ such that $\sigma(x) \supset \sigma(x^*)$ and $\sigma(s) \supset \sigma(s^*)$ where at least one of the two inclusions is strict. In this case one derives immediately the contradiction $\sigma(x) \cap \sigma(s) \neq \phi$.

In the following theorem we prove that any limit point of a sequence $\{(x^k, s^k)\}$ satisfying (2.11) and (2.13) is a maximal complementary solution of (2.1).

Theorem 4.1 *Under the hypothesis of Theorem 2.2 suppose that (x^*, s^*) is a limit point of the sequence $\{(x^k, s^k)\}$. Then (x^*, s^*) is a maximal complementary solution of (2.1).*

Proof. According to Lemma 2 of Güler and Ye [2] it is sufficient to prove that

$$\frac{\min_{1 \le i \le n}\ (x_i^k s_i^k)}{(x^k)^T\ s^k} \ge \xi > 0, \tag{4.4}$$

with ξ independent of k. On the other hand Lemma 3 of the above cited paper shows that (4.4) is satisfied if there is a constant M independent of k such that

$$\phi_n(x^k, s^k) \le M, \tag{4.5}$$

where ϕ_n is defined as in (2.8) with $\rho = n$. As in the proof of Theorem 2.2, let us denote

$$x^k = x,\ s^k = s,\ x^{k+1} = x^+ = x + \Delta x,\ s^{k+1} = s^+ = s + \Delta s.$$

Let now $n + \sqrt{n} \le \rho \le 2n$. Then, by using Theorem 2.2 we can write:

$$\begin{aligned}
& \phi_n(x^+, s^+) - \phi_n(x,s) \\
= \; & \phi_\rho(x^+, s^+) - \phi_\rho(x,s) + \phi_n(x^+, s^+) - \phi_\rho(x^+, s^+) + \phi_\rho(x,s) - \phi_n(x,s) \\
\le \; & -\gamma - (\rho - n)\log\left(((x^+)^T s^+)/(x^T s)\right) \\
= \; & -\gamma - (\rho - n)\log\left(1 + \frac{x^T\Delta s + s^T \Delta x + \Delta x^T \Delta s}{x^T s}\right) \quad .
\end{aligned}$$

Since $\Delta x^T \Delta s \ge 0$, we have

$$\begin{aligned}
& \frac{x^T\Delta s + s^T \Delta x + \Delta x^T \Delta s}{x^T s} \ge \frac{x^T\Delta s + s^T \Delta x}{x^T s} \\
= \; & (\theta p^T e + z^T e)/(x^T s) \\
\ge \; & \frac{1}{x^T s}\left\{[\theta(n/\rho - 1)x^T s] - \frac{\sqrt{n}C\beta}{\|D^{-1}\|^2}\right\} \quad \text{(from (2.11))} \\
= \; & -\frac{\beta}{\|D^{-1}\|\|D^{-1}p\|}\frac{\rho - n}{\rho} - \frac{\sqrt{n}C\beta}{x^T s\|D^{-1}\|^2} \\
\ge \; & -\frac{2\beta(\rho - n)}{\sqrt{3}x^T s\|D^{-1}\|^2} - \frac{\sqrt{n}C\beta}{x^T s\|D^{-1}\|^2} \quad \text{(from (2.17))} \\
= \; & -\frac{\beta}{x^T s\|D^{-1}\|^2}[\frac{2(\rho - n)}{\sqrt{3}} + \beta C\sqrt{n}] \\
\ge \; & -\frac{\beta n}{x^T s\|D^{-1}\|^2}(\frac{2}{\sqrt{3}} + \beta C) \\
\ge \; & -\frac{2n}{\sqrt{3}x^T s\|D^{-1}\|^2}\beta(1 + \beta C) \ge \frac{-n}{\sqrt{3}x^T s\|D^{-1}\|^2} \quad \text{(from (2.12a))}.
\end{aligned}$$

Therefore

$$\phi_n(x^+, s^+) - \phi_n(x,s) \le -\gamma - (\rho - n)\log\left(1 - \frac{n}{\sqrt{3}x^T s\|D^{-1}\|^2}\right). \tag{4.6}$$

From the above inequality we immediately deduce that

$$\phi_n(x^+, s^+) - \phi_n(x,s) \le -\gamma - (\rho - n)\log\left(1 - \frac{1}{\sqrt{3}}\right). \tag{4.7}$$

Let $\omega > 0$ be sufficiently large so that

$$-\gamma - (\rho - n)\log\left(1 - \frac{1}{\sqrt{3}n^{\omega-1}}\right) < 0. \tag{4.8}$$

If $\phi_n(x, s) \geq \omega n \log(n)$, which implies

$$\log(x^T s \|D^{-1}\|^2) \geq \omega \log(n),$$

i.e.,

$$x^T s \|D^{-1}\|^2 \geq n^\omega,$$

then we have from (4.6) and (4.8) that

$$\phi_n(x^+, s^+) - \phi_n(x, s) \leq -\gamma - (\rho - n)\log\left(1 - \frac{1}{\sqrt{3}n^{\omega-1}}\right) < 0. \tag{4.9}$$

Consequently, from (4.7) and (4.9) it follows that (4.5) holds with

$$M = \max\{-\gamma - (\rho - n)\log(1 - 1/\sqrt{3}) + \omega n \log(n), \ \ \phi_n(x^0, s^0)\}. \tag{4.10}$$

The proof is complete. $\square$

Suppose now that problem (2.1) has a strictly complementary solution (x^*, s^*). Then

$$\sigma^* := \sigma(x^*) = \sigma(\mathcal{X}), \ \overline{\sigma}^* := \sigma(s^*) = \sigma(\mathcal{S}) = \{1, 2, \ldots, n\} \backslash \sigma^*. \tag{4.11}$$

Let $\{(x^k, s^k)\}$ be the sequence considered in Theorem 2.2, and denote

$$\sigma^k = \{i \in \{1, 2, \ldots, n\} : x_i^k \geq s_i^k\}. \tag{4.12}$$

In the following theorem we show that for k sufficiently large σ^k coincides with σ^*.

Theorem 4.2 *Under the hypothesis of Theorem 2.2 suppose that NCP has a strictly complementary solution* (x^*, s^*)*. Then there is* $K > 0$ *such that* $\sigma^k = \sigma^*$ *for any* $k \geq K$*.*

Proof. Let us denote

$$\eta := \min(\{x_i^* : i \in \sigma^*\} \cup \{s_i^* : i \in \overline{\sigma}^*\}), \ \epsilon = \eta^2 \xi, \tag{4.13}$$

where ξ is the constant from (4.4). Using Theorem 2.2 and inequality (2.9) it follows that there is $K = O(\phi_\rho(x^0, s^0) - n \log n + (\rho - n)|\log \epsilon|)$, such that

$$(x^k)^T s^k < \epsilon, \forall k \geq K. \tag{4.14}$$

From the monotony of the operator F we have

$$(x^k - x^*)^T(s^k - s^*) \geq 0,$$

so that for $k \geq K$,

$$\sum_{i\in\sigma^*} x_i^*\ s_i^k + \sum_{i\in\overline{\sigma}^*} s_i^*\ x_i^k \leq (x^k)^T\ s^k < \epsilon.$$

Then, by using (4.4) it is easily seen that for $i \in \sigma^*$ the following inequalities hold $\forall k \geq K$:

$$\begin{aligned} s_i^k &< \frac{\epsilon}{x_i^*} \leq \frac{\epsilon}{\eta} = \xi\eta, \\ x_i^k &= \frac{x_i^k s_i^k}{s_i^k} \geq \frac{\xi(x^k)^T s^k}{x_i^* s_i^k} x_i^* \geq \xi\eta. \end{aligned}$$

Similarly, for $i \in \overline{\sigma}^*$, we deduce

$$x_i^k < \xi\eta \leq s_i^k \qquad \forall k \geq K\ .$$

This completes the proof of our theorem. $\Box$

We note that in the context of linear programming the result of the above theorem is contained in Theorem 1 of Ye [19].

5 Local quadratic convergence towards a strictly complementary solution

In Section 4 we have seen that any limit point of the sequence $\{(x^k, s^k)\}$ is a maximal complementary solution. If the nonlinear complementarity problem has a strictly complementary solution then any such limit point (x^*, s^*) is a strictly complementary solution and, according to Theorem 4.2, for k sufficiently large,

$$\sigma^k := \{i \in \{1, 2, \ldots, n\} : x_i^k > s_i^k\} = \sigma^* := \sigma(x^*) = \sigma(\mathcal{X})\ .$$

Suppose for simplicity that

$$\sigma^k = \{1, 2, \ldots, m\} = \sigma^*\ . \tag{5.1}$$

Then the strictly complementary solution (x^*, s^*) is of the form

$$x^* = (y^*, 0),\ y^* \in I\!R^m,\ y^* > 0,\ s^* = (0, w^*),\ w^* \in I\!R^{n-m},\ w^* > 0 \tag{5.2}$$

The vectors y^* and w^* satisfy the condition

$$G(y^*) = 0, \quad w^* = \overline{G}(y^*), \tag{5.3}$$

where the mappings $G : I\!R^m \to I\!R^m$, $\overline{G} : I\!R^m \to I\!R^{n-m}$ are defined by

$$[G(y)]_i = [F(y,0)]_i, \quad i = 1, 2, \ldots, m, \tag{5.4a}$$

$$[\overline{G}(y)]_i = [F(y,0)]_{m+i}, \quad i = 1, 2, \ldots, n-m. \tag{5.4b}$$

Havind the strict complementary set σ^*, we solve the equation $G(y) = 0$ by Newton's method. In what follows we will construct a computational test that will insure that Newton's method

$$y^{j+1} = y^j - G'(y^j)^{-1} G(y^j) \quad j = 0, 1, \ldots, \tag{5.5}$$

with starting point

$$y^0 = [x_1^k, x_2^k, \ldots, x_m^k]^T. \tag{5.6}$$

is well defined and $\lim_{j\to\infty} y^j = y^*$. First we prove the following lemma.

Lemma 5.1 *Suppose that F satisfies the scaled Lipschitz condition (3.1). Then G satisfies*

$$\|Y(G(y+u) - G(y) - G'(y)u)\| \le \psi(\alpha) u^T G'(y) u, \tag{5.7a}$$

whenever

$$u \in I\!R^m, \; y \in I\!R^m, \; y > 0, \; \|Y^{-1}u\| \le \alpha < 1. \tag{5.7b}$$

Proof. Choose an arbitrary $\epsilon > 0$, take $x = (y, \epsilon, \epsilon, \ldots, \epsilon)$, $h = (u, 0, 0, \ldots, 0)$ in (3.1) and then let $\epsilon \to 0$. □

In what follows we assume, in addition to (3.1), that there is a monotone increasing function

$$\psi_1 : (0,1) \to (1, \infty) \tag{5.8a}$$

such that

$$\|(YG'(y+u)Y)^{-1}\| \le \psi_1(\alpha) \|(YG'(y)Y)^{-1}\|, \tag{5.8b}$$

whenever u and y satisfy (5.7b). We note that this condition is no longer satisfied for general LP and QP. However, it is satisfied for some entrophy optimization problems [17].

Because ψ and ψ_1 are monotone increasing and have values in $(1, \infty)$, so is the function

$$\psi_2(\alpha) = \frac{(1+\alpha)\psi(\alpha)\psi_1(\alpha)}{(1-\alpha)^2}. \tag{5.9}$$

It follows that the equation

$$\alpha\psi_2(\alpha) = \frac{1}{2} \tag{5.10}$$

has a unique solution $\alpha^* \in (0, 1/2)$ and that

$$\alpha\psi_2(\alpha) < \frac{1}{2} \quad \forall \in (0, \alpha^*). \tag{5.11}$$

With the above notation we can state the following important result.

Theorem 5.2 *Suppose that $G : I\!R^m \to I\!R^m$ is continuously differentiable and $G'(y)$ is invertible for all $y > 0$. Assume that conditions (5.7) and (5.8) hold. If the starting point $y^0 > 0$ satisfies*

$$\mu_0 := \|(Y^0 G'(y^0) Y^0)^{-1}\| \|Y^0 v^0\| \le \alpha^* \tag{5.12}$$

with $v^0 = G(y^0)$, then Newton's method (5.5) is well defined and produces a sequence of positive vectors $\{y^j\}$ such that

$$\mu_j := \|(Y^j G'(y^j) Y^j)^{-1}\| \|Y^j v^j\| \le c\tau^{2^j}, \tag{5.13}$$

$$\|(Y^j)^{-1}(y^{j+1} - y^j)\| \le c\tau^{2^j}, \tag{5.14}$$

where $v^j = G(y^j)$ and

$$c := \frac{1}{\psi_2(\mu_0)} < 1, \tau := \mu_0\psi_2(\mu_0) < \frac{1}{2}. \tag{5.15}$$

Proof. One step of Newton's method starting from some $y > 0$ can be written as

$$G'(y)\Delta y = -v = -G(y), \; y^+ = y + \Delta y. \tag{5.16}$$

We have clearly

$$Y^{-1}\Delta y = -(YG'(y)Y)^{-1}(Yv).$$

Therefore if

$$\|(YG'(y)Y)^{-1}\| \|Yv\| \le \mu < 1, \tag{5.17}$$

then

$$\|Y^{-1}\Delta y\| \le \mu < 1, \tag{5.18}$$

so that $y^+ > 0$. By denoting $v^+ = G(y^+)$ and by using (5.7) and (ref5.1) we get

$$\begin{aligned}
\|Y^+ v^+\| &\le \|Y^{-1}(Y + \Delta Y)\| \|Yv^+\| \le (1+\mu)\|Yv^+\| \\
&= (1+\mu)\|Y(G(y+\Delta y) - G(y) - G'(y)\Delta y)\| \\
&\le (1+\mu)\psi(\mu)\Delta y^T G'(y)\Delta y.
\end{aligned}$$

Equations (5.16)–(5.18) imply that

$$
\begin{aligned}
\Delta y^T G'(y)\Delta y &= -\Delta y^T v \leq \|Y^{-1}\Delta y\|\|Yv\| \\
&\leq \|(YG'(y)Y)^{-1}\|\|Yv\|^2 \\
&\leq \mu\|Yv\|,
\end{aligned}
$$

so that we obtain

$$
\begin{aligned}
\|Y^+v^+\| &\leq (1+\mu)\psi(\mu)\|(YG'(y)Y^{-1}\|\|Yv\|^2 \\
&\leq \mu(1+\mu)\psi(\mu)\|Yv\|. \qquad (5.19)
\end{aligned}
$$

On the other hand from (5.8) we have

$$
\begin{aligned}
\|(Y^+G'(y^+)Y^+)^{-1}\| &= \|(Y^+)^{-1}Y(YG'(y^+)Y)^{-1}Y(Y^+)^{-1}\| \\
&\leq \|Y(Y+\Delta Y)^{-1}\|^2\|(YG'(y^+)Y)^{-1}\| \\
&\leq \frac{1}{(1-\mu)^2}\|(YG'(y^+)Y)^{-1}\| \\
&\leq \frac{\psi_1(\mu)}{(1-\mu)^2}\|(YG'(y)Y)^{-1}\|. \qquad (5.20)
\end{aligned}
$$

Combining the last two inequalities we deduce that

$$
\begin{aligned}
&\|(Y^+G'(y^+)Y^+)^{-1}\|\|Y^+v^+\| \\
\leq\; &\mu\psi_2(\mu)\|(YG'(y)Y)^{-1}\|\|Yv\| \qquad (5.21)
\end{aligned}
$$

We prove now (5.13) by mathematical induction. It clearly holds for $j = 0$. Suppose it holds for some $j \geq 0$. Then (5.17) is satisfied with $y = y^j$ and $\mu = c\tau^{2^j}$. Consequently from (5.21) we deduce that

$$
\begin{aligned}
&\|(Y^{j+1}G(y^j)Y^{j+1})^{-1}\|\|Y^{j+1}\ v^{j+1}\| \\
\leq\; &\psi_2(c\tau^{2^j})c^2\tau^{2^{j+1}} \\
\leq\; &\psi_2(c\tau)c^2\tau^{2^{j+1}} = c\tau^{2^{j+1}} \quad .
\end{aligned}
$$

Hence (5.13) holds for all $j \geq 0$, and (5.14) follows from the fact that (5.17) implies (5.18). $\square$

Using (5.14) we can prove, as in Lemma 4 of Potra and Ye [17], the following result.

Lemma 5.3 *Under the hypothesis of Theorem 5.2, we have*

$$\|(Y^j)^{-1}Y^0\| < 3, \|(Y^0)^{-1}Y^j\| < 3, \quad j = 0, 1, \ldots$$

Now we have all the ingredients for proving that the sequences $\{y^j\}, \{v^j\}$ converge quadratically.

Theorem 5.4 *Under the hypothesis of Theorem 5.2 the sequence $\{y^k\}$ converges to a point $y^* > 0$ such that $v^* = G(y^*) = 0$ and the following estimates hold*

$$\|y^j - y^*\| \leq 3\|Y^0\|\tau^{2^j}, \tag{5.22}$$

$$\|v^j\| \leq 6c\mu_0\|Y^0v^0\|\tau^{(2^j-1)}, \tag{5.23}$$

$$\|Y^jv^j\| \leq \|Y^0v^0\|\tau^{(2^j-1)}. \tag{5.24}$$

Proof. From (5.12), (5.13) and (5.19) we deduce that

$$\begin{aligned}
\|Y^{j+1}v^{j+1}\| &\leq c\tau^{2^j}(1 + c\tau^{2^j})\psi(c\tau^{2^j})\|Y^jv^j\| \\
&\leq c\psi_2(c\tau^{2^j})\tau^{2^j}\|Y^jv^j\| \\
&\leq c\psi_2(\mu_0)\tau^{2^j}\|Y^jv^j\| = \tau^{2^j}\|Y^jv^j\|,
\end{aligned}$$

and (5.24) follows easily by induction. Then (5.23) is readily obtained by writing

$$\|v^j\| \leq \|(Y^j)^{-1}Y^0\|\|Y^jv^j\| \leq 3\|Y^jv^j\|.$$

On the other hand from (5.14) and Lemma 5.3,

$$\begin{aligned}
\|y^{j+1} - y^j\| &\leq \|Y^j\|\|(Y^j)^{-1}(y^{j+1} - y^j)\| \\
&\leq c\|Y^j\|\tau^{2^j} \\
&\leq c\|Y^0\|\|(Y^0)^{-1}Y^j\|\tau^{2^j} \\
&\leq 3c\|Y^0\|\tau^{2^j}.
\end{aligned}$$

It follows that $\{y^j\}$ is a Cauchy sequence and therefore converges to a point y^*. Because $y^j \geq 0\ \forall j$ we must have $y^* \geq 0$ and by invoking Lemma 5.3 we prove that $y^* > 0$. Finally the error estimate (5.22) is obtained from

$$\begin{aligned}
\|y^j - y^*\| &\le \sum_{i=0}^{\infty} \|y^{j+i} - y^{j+i+1}\| \\
&\le 3c\|Y^0\| \sum_{i=0}^{\infty} \tau^{2^j}\tau^{2^i} \\
&\le 3c\|Y^0\| \sum_{i=0}^{\infty} \tau^{2^j}\tau^{1+i} \\
&\le 3c\|Y^0\|\tau\tau^{2^j} \sum_{i=0}^{\infty} \tau^{i} \\
&\le 6c\mu_0\|Y^0\|\tau^{2^j}.
\end{aligned}$$

□

We have seen that under the hypothesis of Theorem 5.2 the sequence $\{y^k\}$ produced by Newton's method (5.5) converges to a point $y^* > 0$ such that $G(y^*) = 0$. If $w^* = \overline{G}(y^*) > 0$, then

$$x^* = (y^*, 0), \quad s^* = (0, \overline{G}(y^*)) \tag{5.25}$$

is a strictly complementary solution of (2.1). We know that in general the problem does not have a strictly complementary condition. In the following corollary we give a sufficient condition on y^0 for the existence of a strictly complementary solution.

Corollary 5.5 *Under the hypothesis of Theorem 5.2 suppose that there is a constant L such that*

$$\|\overline{G}(y^0 + h) - \overline{G}(y^0)\| \le L\|h\| \tag{5.26}$$

for all $h \in I\!R^m$ satisfying

$$\|h\| \le 6c\mu_0\|Y^0\|.$$

Assume also that

$$w^0 = \overline{G}(y^0) > 6Lc\|Y^0\|\mu_0 e, \tag{5.27}$$

where $e = [1, 1, \ldots, 1]^T \in I\!R^{n-m}$. Then (5.25) is a strictly complementary solution of (2.1).

Proof. The proof is immediate because according to (5.22) $\|y^0 - y^*\| \le 6c\mu_0\|Y^0\|$, and therefore

$$\begin{aligned}
w^* = \overline{G}(y^*) \quad &= \overline{G}(y^0) + (\overline{G}(y^*) - \overline{G}(y^0)) \\
&\ge \overline{G}(y^0) - 6Lc\mu_0\|Y^0\|e > 0.
\end{aligned}$$

We can always choose μ_0 small enough such that (5.27) is satisfied. □

6 The complexity of the hybrid potential-reduction Newton method

In this section we will show that if the nonlinear complementarity problem (2.1) has a strictly complementary solution of the form (5.25) then the starting point (5.6), with x^k defined as in Corollary 3.3 will satisfy condition (5.12) from the hypothesis of Theorem 5.2, provided k is large enough. We will give an explicit bound on the size of such a k and we will discuss the complexity of a hybrid algorithm that starts with algorithm (3.2), (3.4) and then switches to Newton's method (5.5) as soon as (5.12) and (5.27) are satisfied. First we consider the function

$$g(y) = \int_0^1 (y - y^*)^T G(y^* + t(y - y^*))dt. \tag{6.1}$$

By taking $x = (y, 0)$ in (2.3) we deduce that G is monotone, i.e.,

$$(y - \overline{y})^T (G(y) - G(\overline{y})) \geq 0 \quad \forall y, \overline{y}. \tag{6.2}$$

If G is differentiable this implies

$$u^T G'(y) u \geq 0 \quad \forall u, y. \tag{6.3}$$

If the Jacobian matrix $G'(y)$ is symmetric, then the function g given by (6.1) is convex and $G'(y)$ is its Hessian at y. However this is not the case in general. Nevertheless we can prove the following properties of g:

Lemma 6.1 *Let $G : \mathbb{R}^m \to \mathbb{R}^m$ be continuous and monotone, and let g be the function defined in (6.1) where $y^* > 0$ and $G(y^*) = 0$. Then*

(i) $\min_{y \in \mathbb{R}^m} g(y) = g(y^*) = 0$
(ii) *If $G(y) \geq 0$ then $g(y) - g(y^*) \leq y^T G(y)$*
(iii) $g(y) \geq g(y + \lambda(y^* - y)) \quad \forall y \in \mathbb{R}^m \ \forall \lambda \in (0, 1)$

Proof. $g(y^*) = 0$ by definition. Also from $G(y^*) = 0$ and (6.2) we obtain

$$\begin{aligned} g(y) &= \int_0^1 (y - y^*)^T [G(y^* + t(y - y^*)) - G(y^*)]dt \\ &= \int_0^1 \frac{1}{t}[(y^* + t(y - y^*)) - y^*]^T [G(y^* + t(y - y^*)) - G(y^*)]dt \geq 0 \ , \end{aligned}$$

which proves (i). Similarly, if $G(y)\geq 0$ then we can write

$$\begin{aligned} g(y) &= (y-y^*)^TG(y)+\int_0^1 (y-y^*)^T[G(y^*+t(y-y^*))-G(y)]dt \\ &= (y-y^*)^TG(y)-\int_0^1 \frac{1}{1-t}[(y^*+t(y-y^*))-y]^T[G(y^*+t(y-y^*))-G(y)]dt \\ &\leq (y-y^*)^TG(y)=y^TG(y)-(y^*)^TG(y)\leq y^TG(y). \end{aligned}$$

Let $y\in I\!R^m$, $0<\lambda<1$ and denote $\overline{y}=y+\lambda(y^*-y)$. Then it is easily checked that

$$\begin{aligned} & g(y)-g(\overline{y})=(1-\lambda)g(y)-g(\overline{y})+\lambda g(y) \\ =\ & (1-\lambda)\int_0^1 (y-y^*)^T[G(y^*+t(y-y^*))-G(y^*+t(1-\lambda)(y-y^*))]dt+\lambda g(y) \\ =\ & (1-\lambda)\int_0^1 \frac{1}{\lambda t}[(y^*+t(y-y*))-(y^*+t(1-\lambda)(y-y^*))]^T \\ \times\ & [G(y^*+t(y-y^*))-G(y^*+t(1-\lambda)(y-y^*))]dt+\lambda g(y) \\ \geq\ & \lambda g(y)\geq 0. \quad \square \end{aligned}$$

Property (6.3) can be written equivalently as

$$u^TG'(y+u)u\geq 0 \qquad \forall u,y.$$

In what follows we assume that there is a monotone decreasing function

$$\psi_3:(0,1)\to(0,1), \tag{6.4a}$$

such that

$$u^TG'(y+u)u\geq \psi_3(\alpha)\|Y^{-1}u\|^2/\|(YG'(y)Y)^{-1}\|, \tag{6.4b}$$

whenever

$$y>0,\ u\in I\!R^m,\ \|Y^{-1}u\|\leq\alpha<1. \tag{6.4c}$$

We note that if G' is symmetric then (5.8) implies that (6.4) is satisfied with $\psi_3(\alpha)=1/\psi_1(\alpha)$. Under the assumption (6.4) we can prove the following result.

Lemma 6.2 *Let $G:I\!R^m\to I\!R^m$ be monotone and continuously differentiable, and suppose that (6.4) is satisfied. Let $y\geq 0$, $v=G(y)\geq 0$ be such that*

$$y^Tv<\delta^2\leq\frac{(\alpha^2/2)\psi_3(\alpha)}{\|(Y^*G'(y^*)Y^*)^{-1}\|}, \tag{6.5}$$

for some $\alpha\in(0,1)$, where $y^*>0$, $G(y^*)=0$. Then

$$\|(Y^*)^{-1}(y-y^*)\|<\alpha. \tag{6.6}$$

Proof. From Lemma 6.1 (ii) we have

$$g(y) < \frac{(\alpha^2/2)\psi_3(\alpha)}{\|(Y^*G'(y^*)Y^*)^{-1}\|}. \tag{6.7}$$

By contradiction suppose that

$$\|(Y^*)^{-1}(y - y^*)\| \geq \alpha, \tag{6.8}$$

and consider the following minimization problem:

$$\begin{array}{c} \text{minimize} \quad g(y) \\ \text{subject to} \quad y > 0 \quad \text{and} \quad \|(Y^*)^{-1}(y - y^*)\| \geq \alpha. \end{array} \tag{6.9}$$

Because y^* is not feasible for this problem we deduce, by using Lemma 6.1 (iii), that the above minimization problem has a solution $\overline{y}$ on the boundary of the ellipsoid constraint, i.e.,

$$\|(Y^*)^{-1}(\overline{y} - y^*)\| = \alpha.$$

On the other hand according to (6.8) y is feasible for (6.9), so that by using (6.4) we obtain

$$\begin{aligned} g(y) &\geq g(\overline{y}) = \int_0^1 (\overline{y} - y^*)^T [G(y^* + t(\overline{y} - y^*)) - G(y^*)]dt \\ &= \int_0^1 \int_0^1 t(\overline{y} - y^*)^T \; G'(y^* + ts(\overline{y} - y^*))(\overline{y} - y^*)dsdt \\ &\geq \int_0^1 \int_0^1 \psi_3(\alpha)t(\|(Y^*)^{-1}(\overline{y} - y^*)\|^2/\|(Y^*G'(y^*)Y^*)^{-1}\|)ds\; dt \\ &\geq \frac{\alpha^2}{2}\psi_3(\alpha)/\|(Y^*G'(y^*)Y^*)^{-1}\|, \end{aligned}$$

which contradicts (6.7). The proof is complete. □

Now, suppose that the points obtained by Algorithm 3.3 satisfy

$$(x^k)^T s^k < \frac{\alpha^2}{2}\psi_3(\alpha)/\|(Y^*G'(y^*)Y^*)^{-1}\|. \tag{6.10}$$

Let y^0 be given by (5.6) and $v^0 = G(y^0)$. We have clearly $(y^0)^T v^0 \leq (x^k)^T s^k$ (??), so that by applying Lemma 6.2 we deduce that

$$\|(Y^0)^{-1}(y^0 - y^*)\| < \alpha.$$

In particular this implies

$$\|(Y^0)^{-1}Y^*\| < 1/(1-\alpha),$$

so that by using (5.8) we have

$$\|(Y^0G'(y^0)Y^0)^{-1}\| < \frac{\psi_1(\alpha)}{(1-\alpha)^2}\|(Y^*G(y^*)Y^*)^{-1}\|$$

On the other hand

$$\|Y^0v^0\| \le (y^0)^Tv^0 \le (x^k)^Ts^k \ (??)$$

and, by invoking (6.10) we obtain

$$\|(Y^0G'(y^0)Y^0)^{-1}\|\|Y^0v^0\| < \frac{\alpha^2\psi_1(\alpha)\psi_3(\alpha)}{2(1-\alpha)^2}. \tag{6.11}$$

Therefore (5.12) will be satisfied provided

$$\frac{\alpha^2\psi_1(\alpha)\psi_3(\alpha)}{(1-\alpha)^2} \le \alpha^*. \tag{6.12}$$

Because $\psi_3(\alpha) < 1$, the above inequality will be satisfied for all $0 < \alpha \le \alpha^{**}$, where α^{**} is the unique solution of the equation

$$\psi_4(\alpha) := \frac{\alpha^2\psi_1(\alpha)}{2(1-\alpha)^2} = \alpha^* \tag{6.13}$$

belonging to the interval (0,1).

The function $\psi_4 : (0,1) \to \mathbb{R}_+$ defined above is clearly monotone increasing and we have

$$\begin{aligned} \psi_4(\alpha^*) &\le \frac{(\alpha^*)^2(1+\alpha^*)\psi(\alpha^*)\psi_1(\alpha^*)}{2(1-\alpha^*)^2} \\ &= (\alpha^*)^2\psi_2(\alpha^*)/2 = \alpha^*/4 < \alpha^* \end{aligned}$$

and

$$\psi_4(1/2) = \frac{1}{2}\psi_1(\alpha) \ge \frac{1}{2} > \alpha^*.$$

Hence

$$\alpha^* < \alpha^{**} < 1/2. \tag{6.14}$$

It follows that (5.12) will be satisfied provided

$$(x^k)^T s^k < \frac{\alpha^2 \psi_3(\alpha)}{2\|(Y^* G'(y^*) Y^*)^{-1}\|} \tag{6.15}$$

for some $0 < \alpha \leq \alpha^{**}$. Moreover if (6.15) holds for some $0 < \alpha \leq \alpha^*$ then $\mu_0 \leq \frac{1}{2}\alpha$. For $\alpha \leq \min(\alpha^*, 1/(9\|Y^*\|))$ we obtain $6c\mu_0\|Y^0\| < 1$ (see (5.22)). Summing up we obtain the following theorem:

Theorem 6.3 *Under the hypothesis of Corollary 3.3 suppose that the problem (2.1) has a strictly complementary solution of the form (5.2). Assume that $G'(y)$ is invertible for all $y > 0$, and that (5.8) and (6.4) are satisfied. Then Algorithm 3.3 generates a sequence $\{x^k\}$ such that condition (5.12) is satisfied for y^0 given by (5.6) in at most*

$$O(\phi_\rho(x^0, s^0) + (\rho - n)|\log(\|Y^*\|^2 \|(Y^* G'(y^*) Y^*)^{-1}\|)|)$$

steps. From this starting point, Newton's method produces a sequence $\{y^j\}$ such that the inequality $\|y^j - y^*\| < \epsilon$ is attained in at most $\log_2|\log_2 \epsilon|$ steps. Note that if the problem is an LCP, then Newton's method produces the solution in one step.

Acknowledgment

The authors would like to thank Rongqin Sheng and three anonymous referees for their comments that lead to a much better presentation of our results.

References

[1] O. Güler. Existence of interior points and interior paths in nonlinear monotone complementarity problems. *Mathematics of Operations Research*, 18(1):128–147, 1993.

[2] O. Güler and Y. Ye. Convergence behavior of interior point algorithms. *Mathematical Programming*, 60:215–228, 1993.

[3] D. den Hertog, C. Roos, and T. Terlaky. On the classical logarithmic barrier function method for a class of smooth convex programming problems. *Journal of Optimization Theory and Applications*, 73:1–25, 1992.

[4] F. Jarre. *The method of analytic centers for smooth convex programs.* PhD thesis, Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Am Hubland, D–8700 Würzburg, West–Germany, 1989.

[5] M. Kojima, N. Megiddo, and T. Noma. Homotopy continuation methods for nonlinear complementarity problems. *Mathematics of Operations Research*, 16(4):754–774, 1991.

[6] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. A unified approach to interior point algorithms for linear complementarity problems. *Lecture Notes in Comput. Sci.*, 538, 1991.

[7] M. Kojima, S. Mizuno, and T. Noma. A new continuation method for complementarity problems with uniform P–functions. *Mathematical Programming*, 43:107–113, 1989.

[8] M. Kojima, S. Mizuno, and T. Noma. Limiting behavior of trajectories by a continuation method for monotone complementarity problems. *Mathematics of Operations Research*, 15(4):662–675, 1990.

[9] M. Kojima, S. Mizuno, and A. Yoshise. A polynomial–time algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44:1–26, 1989.

[10] M. Kojima, S. Mizuno, and A. Yoshise. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems. *Mathematical Programming*, 50:331–342, 1991.

[11] K. O. Kortanek, F. Potra, and Y. Ye. On some efficient interior point methods for nonlinear convex programming.

Linear Algebra and Its Applications, 152:169–189, 1991.

[12] L. McLinden. The analogue of Moreau's proximation theorem, with applications to the nonlinear complementarity problem. *Pacific Journal of Mathematics*, 88:101–161, 1980.

[13] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming : Interior Point and Related Methods*, pages 131–158. Springer Verlag, New York, 1989. Identical version in : *Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan*, pages 1–35, 1986.

[14] R. D. C. Monteiro and I. Adler. Interior path following primal–dual algorithms : Part I : Linear programming. *Mathematical Programming*, 44:27–41, 1989.

[15] R. D. C. Monteiro and I. Adler. An extension of Karmarkar–type algorithm to a class of convex separable programming problems with global linear rate of convergence. *Mathematics of Operations Research*, 15:408–422, 1990.

[16] Y. E. Nesterov and A. S. Nemirovsky. *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms.*

SIAM Publications. SIAM, Philadelphia, USA, 1994.

[17] F. A. Potra and Y. Ye. A quadratically convergent polynomial algorithm for solving entropy optimization problems. *SIAM Journal on Optimization*, 3:843–860, 1993.

[18] M. J. Todd and Y. Ye. A centered projective algorithm for linear programming.

Mathematics of Operations Research, 15:508–529, 1990.

[19] Y. Ye. On the finite convergence of interior–point algorithms for linear programming. *Mathematical Programming*, 57:325–335, 1992.

[20] Y. Zhang and R. A. Tapia. A superlinearly convergent polynomial primal–dual interior–point algorithm for linear programming. *SIAM Journal on Optimization*, 3:118–133, 1993.

[21] J. Zhu. A path following algorithm for a class of convex programming problems. *Zeitschrift für Operations Research—Methods and Models of Operations Research*, 36(4):539–377, 1992.