PREPRINT MCS-P357-0493, MATHEMATICS AND COMPUTER SCIENCE DIVISION, ARGONNE NATIONAL LABORATORY

Local Convergence of Interior-Point Algorithms for Degenerate Monotone LCP

Renato D.C. Monteiro* and Stephen Wright†

April 6, 1993

Abstract

Most asymptotic convergence analysis of interior-point algorithms for monotone linear complementarity problems assumes that the problem is nondegenerate, that is, the solution set contains a strictly complementary solution. We investigate the behavior of these algorithms when this assumption is removed.

1 Introduction

In the monotone linear complementarity problem (LCP), we seek a vector pair $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ that satisfies the conditions

$$y = Mx + q, \qquad x \geq 0, \qquad y \geq 0, \qquad x^T y = 0, \tag{1}$$

where $q \in \mathbb{R}^n$, and $M \in \mathbb{R}^{n \times n}$ is positive semidefinite. We use $\mathcal{S}$ to denote the solution set of (1).

An assumption that is frequently made in order to prove superlinear convergence of interior-point algorithms for (1) is the *nondegeneracy* assumption:

Assumption 1 *There is an $(x^*, y^*) \in \mathcal{S}$ such that $x_i^* + y_i^* > 0$ for all $i = 1, \cdots, n$.*

In general, we can define three subsets B, N, and J of the index set $\{1, \cdots, n\}$ by

$$\begin{aligned} B &= \{i = 1, \cdots, n \,|\, x_i^* > 0 \text{ for at least one } (x^*, y^*) \in \mathcal{S}\}, \\ N &= \{i = 1, \cdots, n \,|\, y_i^* > 0 \text{ for at least one } (x^*, y^*) \in \mathcal{S}\}, \\ J &= \{i = 1, \cdots, n \,|\, x_i^* = y_i^* = 0 \text{ for all } (x^*, y^*) \in \mathcal{S}\}. \end{aligned} \tag{2}$$

It is well known that B, N, and J form a partition of $\{1, \cdots, n\}$. Another useful result is the following.

*Systems and Industrial Engineering Department, University of Arizona, Tucson, AZ 85721. The work of this author was based on research supported by the National Science Foundation under grant DDM-9109404 and the Office of Naval Research under grant N00014-93-1-0234.

†Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439. The work of this author was based on research supported by the Office of Scientific Computing, U.S. Department of Energy, under Contract W-31-109-Eng-38.

Lemma 1.1 *There is an* $(x^*, y^*) \in \mathcal{S}$ *such that* $x_i^* > 0$ *for all* $i \in B$ *and* $y_i^* > 0$ *for all* $i \in N$.

Proof. Choose $|B| + |N|$ members (x^i, y^i) of $\mathcal{S}$ (where $|\cdot|$ denotes set cardinality) with the property that $x_i^i > 0$ for $i \in B$ and $y_i^i > 0$ for $i \in N$. Define

$$(x^*, y^*) = \frac{1}{|B| + |N|} \sum_{i \in B \cup N} (x^i, y^i).$$

Since $(x^i)^T y^j = (x^j)^T y^i = 0$ for any two solutions (x^i, y^i) and (x^j, y^j) of (1), it is easy to check that $y^* = Mx^* + q$ and $(x^*)^T y^* = 0$. Moreover, $x_i^* > 0$ for all $i \in B$ and $y_i^* > 0$ for all $i \in N$, giving the result. ∎

Assumption 1 can be restated simply as $J = \emptyset$.

An infeasible-interior-point algorithm solves (1) by generating a sequence of strictly positive iterates $\{(x^k, y^k)\}$, $k = 0, 1, 2, \cdots$, while aiming to satisfy the two equality relationships in (1) in the limit as $k \to \infty$. Feasible interior-point algorithms require all iterates to satisfy $y^k = Mx^k + q$ in addition to strict positivity $(x^k, y^k) > 0$.

This paper starts with general results about infeasible-interior-point algorithms. In our presentation, "Q-superlinear convergence" always means Q-superlinear convergence of the complementarity gap $(x^k)^T y^k$ to 0. In Section 2, we define the broad class of infeasible algorithms considered in this paper and show that no algorithm of this class can achieve Q-superlinear convergence when $J \neq \emptyset$. Ye and Anstreicher [10] presented a trivial LCP for which $J \neq \emptyset$ and observed that no feasible algorithm whose steps approach the primal-dual affine scaling directions can converge superlinearly for this example. Our result generalizes Ye and Anstreicher's observation to all instances of problem (1) for which $J \neq \emptyset$ and to a broader class of algorithms that includes many infeasible algorithms.

Section 3 proposes a scheme for estimating the index sets B, N, and J and shows that a finite termination scheme based on these estimates eventually yields an exact solution of (1). The results of this section generalizes the results obtained in Ye [9] for linear programs to the context of degenerate monotone LCPs.

Feasible algorithms in which the step vectors asymptotically converge to the primal-dual affine scaling direction are discussed in Section 4. In this case, we are able to prove stronger results about the asymptotic behavior of the steps $(\Delta x^k, \Delta y^k)$ and to derive formulae for the linear rate of convergence of the complementarity gap in terms of the current iterate and the current stepsize.

In Section 5, we analyze the two-step linear rate of convergence of the predictor-corrector algorithm. We show that this rate is linear with a constant of at most $1 - c/Q^{1/4}$, where $Q \leq \min\{|J|, n - |J|\}$ and c is not too small. This result shows that if $|J|$ contains only a few indices, or it contains all but a few indices, then we can expect a linear rate of convergence that is not too slow.

The following notation is used throughout the paper. Superscripts on matrices and vectors and subscripts on index sets and scalars denote iteration indices (usually k), while subscripts on matrices and vectors define components. The subvector x_B denotes $[x_i]_{i \in B}$, while the submatrix M_{BN} is $[M_{ij}]_{i\in B, j\in N}$. Subvectors and submatrices corresponding to other index sets are defined likewise. We use μ_k to denote the normalized complementarity gap $\mu_k = (x^k)^T y^k/n$. If $w \in \mathbb{R}^n$ then $\mathrm{diag}\,(w)$ denotes the diagonal matrix having the components of w as diagonal entries. The matrices X^k and Y^k are defined as $X^k = \mathrm{diag}\,(x^k)$ and $Y^k = \mathrm{diag}\,(y^k)$. If u and v are two vectors of the same length and $\nu \in \mathbb{R}$, then uv denotes the vector whose i-the component is $u_i v_i$ and u^ν denotes the vector whose i-th component is u_i^ν. The vector $(1, \cdots, 1)^T$, regardless of its dimension, is denoted by e. For any vector x, the notation x_+ is used for the vector whose i-th component is $\max(x_i, 0)$. Unless otherwise specified, $\|\cdot\|$ denotes $\|\cdot\|_2$. If $\{\delta_k\}$ and $\{\epsilon_k\}$ are two positive sequences, we say $\delta_k = \mathcal{O}(\epsilon_k)$ if $\limsup_k \delta_k/\epsilon_k < \infty$ and $\delta_k = o(\epsilon_k)$ if $\lim_k \delta_k/\epsilon_k = 0$. The distance function to the solution set $\mathcal{S}$ is defined as

$$\mathrm{dist}((x,y), \mathcal{S}) = \min_{(\bar{x},\bar{y})\in\mathcal{S}} \|(\bar{x},\bar{y}) - (x,y)\|.$$

2 Infeasible Algorithms: Local Convergence

In this section we largely restrict ourselves to discussing interior-point algorithms that fit the following framework, which we refer to as the *standard framework*.

(a) For all $k \geq 0$ we have $(x^k, y^k) > 0$ and

$$\|(x^k, y^k)\|_\infty \leq C_b, \tag{3}$$

for some constant $C_b > 0$.

(b) There is a $\bar{\gamma} \in (0,1)$ such that

$$x_i^k y_i^k \geq \bar{\gamma}\mu_k \qquad \text{for all } i = 1, \cdots, n \text{ and } k \geq 0. \tag{4}$$

(c) The step has the form $(x^{k+1}, y^{k+1}) = (x^k, y^k) + \alpha_k(\Delta x^k, \Delta y^k)$ with $\alpha_k \in (0,1]$. The search direction $(\Delta x^k, \Delta y^k)$ satisfies the equation

$$\begin{bmatrix} M & -I \\ Y^k & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \end{bmatrix} = \begin{bmatrix} r^k \\ -x^k y^k + \sigma_k \mu_k e \end{bmatrix}, \tag{5}$$

where $r^k = y^k - Mx^k - q$ denotes the residual vector and $\sigma_k \in [0,1]$ is the centering parameter.

(d) There are constants $\underline{\rho}$ and $\bar{\rho}$ such that $0 \leq \underline{\rho} \leq \bar{\rho}$ and

$$\underline{\rho}\frac{\|r^0\|}{\mu_0} \leq \frac{\|r^k\|}{\mu_k} \leq \bar{\rho}\frac{\|r^0\|}{\mu_0}. \tag{6}$$

For feasible interior-point algorithms, we have $r^k = 0$, $k = 0, 1, \cdots$, and so (d) is trivially satisfied.

Infeasible-interior-point algorithms that fit the standard framework include those of Zhang [11] and Wright [7]. This framework is broad enough to include most algorithms that have been proposed to date, including path-following and predictor-corrector algorithms (see, for example, Zhang [11] and Ji, Potra, and Huang [2]). Many algorithms use a central-path neighborhood different from (4); we use (4) partly because these alternative neighborhoods are usually subsets of our neighborhood when $\bar{\gamma}$ is chosen sufficiently small.

We note that the analysis of most algorithms does not require Assumption 1 to hold in order to prove global linear convergence or polynomial complexity. It is usually needed only to obtain a local convergence rate that is faster than the global rate (which is usually Q-linear or two-step Q-linear).

The following observation (see Zhang [11, Proposition 3.3]) will be needed in subsequent results. If we define

$$\nu_0 = 1, \qquad \nu_k \equiv \prod_{j=0}^{k-1} (1 - \alpha_j), \qquad \forall k > 0,$$

then,

$$r^k = \nu_k r^0, \qquad \forall k \geq 0. \tag{7}$$

The first few results of this section give upper and lower bounds on the components x_i^k and y_i^k and their ratios. In Lemma 2.1, we show that (3) can be a consequence of other frequently-made assumptions.

Lemma 2.1 *If either*

(a) $\underline{\rho} > 0$ *or*

(b) there is a strictly feasible point for (1), and the sequence $\{\mu_k\}$ *is bounded,*

then (3) holds.

Proof. For case (a), our proof uses techniques similar to those of Mizuno [4, Lemma 3.3], Potra [6, Lemma 4.1], and Wright [8, Lemma 3.2]. Given any point $(\bar{x}, \bar{y})$ with $(\bar{x}, \bar{y}) \geq 0$, $\bar{y} = M\bar{x} + q$, and the starting point (x^0, y^0), we have

$$\begin{aligned}
&M(\nu_k x^0 + (1 - \nu_k)\bar{x} - x^k) \\
&\quad = \nu_k M x^0 + (1 - \nu_k) M\bar{x} - Mx^k \\
&\quad = \nu_k (y^0 - q - r^0) + (1 - \nu_k)(\bar{y} - q) - (y^k - q - r^k) \\
&\quad = \nu_k y^0 + (1 - \nu_k)\bar{y} - y^k.
\end{aligned}$$

Hence, by positive semidefiniteness of M, we have

$$0 \leq (\nu_k x^0 + (1 - \nu_k)\bar{x} - x^k)^T (\nu_k y^0 + (1 - \nu_k)\bar{y} - y^k).$$

Rearranging this expression, we obtain

$$
\begin{aligned}
&\nu_k({x^0}^T y^k + {x^k}^T y^0) &(8)\\
&\leq \ \nu_k^2 {x^0}^T y^0 + {x^k}^T y^k + \nu_k(1-\nu_k)({x^0}^T \bar{y} + \bar{x}^T y^0) + (1-\nu_k)^2 \bar{x}^T \bar{y} - (1-\nu_k)(\bar{x}^T y^k + {x^k}^T \bar{y}).
\end{aligned}
$$

Since $(\bar{x}, \bar{y}) \geq 0$, $(x^k, y^k) \geq 0$, and $(1-\nu_k) \geq 0$, the last term on the right-hand side is nonnegative, so we can drop it without affecting the inequality. Also, we can choose $(\bar{x}, \bar{y}) \in \mathcal{S}$ to ensure $\bar{x}^T \bar{y} = 0$, and after some manipulation (8) becomes

$$
\begin{aligned}
{x^0}^T y^k + {x^k}^T y^0 &\leq \nu_k {x^0}^T y^0 + n\frac{\mu_k}{\nu_k} + ({x^0}^T \bar{y} + \bar{x}^T y^0) \\
&\leq {x^0}^T y^0 + n\frac{\mu_0}{\underline{\rho}} + ({x^0}^T \bar{y} + \bar{x}^T y^0),
\end{aligned}
$$

where the second inequality follows from (6). If we define

$$
C_b = \max_{i=1,\cdots,n} \left\{ \frac{1}{\min(x_i^0, y_i^0)} \left[n\mu_0 \left(1 + \frac{1}{\underline{\rho}} \right) + ({x^0}^T \bar{y} + \bar{x}^T y^0) \right] \right\},
$$

then (3) follows.

For (b), we choose $(\bar{x}, \bar{y})$ to be a strictly feasible point, so that $(\bar{x}, \bar{y}) > 0$. By rearranging (8), we obtain

$$
\begin{aligned}
&(1-\nu_k)(\bar{x}^T y^k + {x^k}^T \bar{y}) &(9)\\
&\leq \ \nu_k^2 {x^0}^T y^0 + {x^k}^T y^k + \nu_k(1-\nu_k)({x^0}^T \bar{y} + \bar{x}^T y^0) + (1-\nu_k)^2 \bar{x}^T \bar{y} - \nu_k({x^0}^T y^k + {x^k}^T y^0).
\end{aligned}
$$

Again, the final term is nonnegative, so we can drop it without altering the inequality. We also use $\nu_k \in [0,1]$ and

$$
1 - \nu_k \geq 1 - \nu_1 = \alpha_0 > 0, \qquad \forall k > 0 \tag{10}
$$

to write

$$
\bar{x}^T y^k + {x^k}^T \bar{y} \leq \frac{1}{\alpha_0}({x^0}^T y^0 + {x^k}^T y^k) + {x^0}^T \bar{y} + \bar{x}^T y^0 + \bar{x}^T \bar{y}.
$$

Boundedness of ${x^k}^T y^k$ can now be used to obtain (3). ∎

Our next result gives bounds on components of x_B^k and y_N^k.

Lemma 2.2 *There is a positive constant C_1 such that for all $k \geq 0$,*

$$
i \in B \ \Rightarrow \ y_i^k \leq C_1 \mu_k, \qquad x_i^k \geq \bar{\gamma}/C_1; \tag{11}
$$

$$
i \in N \ \Rightarrow \ x_i^k \leq C_1 \mu_k, \qquad y_i^k \geq \bar{\gamma}/C_1. \tag{12}
$$

Proof. Again, we choose $(\bar{x}, \bar{y})$ in (8) to be a solution of (1) for which $\bar{x}_i > 0$ for $i \in B$ and $\bar{y}_i > 0$ for $i \in N$. From (9) we have

$$\begin{aligned} &(1-\nu_k)(\bar{x}^T y^k + {x^k}^T \bar{y}) \\ &\leq \nu_k^2 {x^0}^T y^0 + {x^k}^T y^k + \nu_k(1-\nu_k)({x^0}^T \bar{y} + \bar{x}^T y^0) - \nu_k({x^0}^T y^k + {x^k}^T y^0). \end{aligned}$$

Dropping the (nonnegative) final term on the right-hand side and using $\nu_k \in [0,1]$, we have

$$(1-\nu_k)(\bar{x}^T y^k + {x^k}^T \bar{y}) \leq \nu_k n\mu_0 + n\mu_k + \nu_k(1-\nu_k)({x^0}^T \bar{y} + \bar{x}^T y^0). \tag{13}$$

Using (10), (13), and (6), we can write

$$\begin{aligned} \bar{x}^T y^k + {x^k}^T \bar{y} &\leq \frac{1}{\alpha_0}(\nu_k n\mu_0 + n\mu_k) + \nu_k({x^0}^T \bar{y} + \bar{x}^T y^0) \\ &\leq \frac{1}{\alpha_0}(n\bar{\rho}\mu_k + n\mu_k) + \frac{\bar{\rho}}{\mu_0}\mu_k({x^0}^T \bar{y} + \bar{x}^T y^0) \\ &\leq \bar{C}_1 \mu_k, \end{aligned}$$

where $\bar{C}_1$ is defined in an obvious way. The upper bounds on x_N^k and y_B^k follow when we define

$$C_1 = \bar{C}_1 \max\left(\max_{i \in B} \frac{1}{\bar{x}_i}, \max_{i \in N} \frac{1}{\bar{y}_i}\right).$$

To prove the remaining inequalities, we use (4). Taking $i \in B$, we have

$$x_i^k y_i^k \geq \bar{\gamma}\mu_k \;\Rightarrow\; x_i^k \geq \frac{\bar{\gamma}\mu_k}{y_i^k} \geq \frac{\bar{\gamma}\mu_k}{C_1 \mu_k} = \frac{\bar{\gamma}}{C_1}.$$

The lower bound on y_i^k, $i \in N$, is proved analogously. ∎

The following consequence of a result of Mangasarian and Shiau [3] bounds the distance to the solution set in terms of μ_k.

Lemma 2.3 *There is a positive constant C_2 such that for all $k \geq 0$,*

$$\operatorname{dist}((x^k, y^k), \mathcal{S}) \leq C_2 \mu_k^{1/2}. \tag{14}$$

Proof. From Mangasarian and Shiau [3, Theorem 2.7], there exist positive constants C_{2a} and C_{2b} such that

$$\begin{aligned} &\min_{\bar{x} \,|\, (\bar{x}, M\bar{x}+q) \in \mathcal{S}} \|\bar{x} - x^k\|_\infty \qquad (15) \\ &\leq C_{2a} \left\{ \left\| \begin{bmatrix} (x^k)^T(Mx^k+q) \\ -Mx^k - q \\ -x^k \end{bmatrix}_+ \right\|_2 + \left[(x^k)^T(Mx^k+q) + C_{2b} \left\| \begin{bmatrix} -Mx^k - q \\ -x^k \end{bmatrix}_+ \right\|_2 \right]^{1/2} \right\}. \end{aligned}$$

In our case, $Mx^k + q = y^k - r^k$ and $(x^k, y^k) > 0$, so by (6),

$$\{(x^k)^T(Mx^k+q)i\}_+ \le |(x^k)^T(Mx^k+q)| \le (x^k)^T y^k + \|x^k\|\|r^k\| \le \left(n + C_b\bar{\rho}\frac{\|r^0\|}{\mu_0}\right)\mu_k,$$

$$\|(-Mx^k-q)_+\| = \|(r^k-y^k)_+\| \le \|r_+^k\| \le \|r^k\| \le C_b\bar{\rho}\frac{\|r^0\|}{\mu_0}\mu_k,$$

$$\|(-x^k)_+\| = 0.$$

Substitution of these inequalities into (15) indicates that

$$\min_{\bar{x}\,|\,(\bar{x},M\bar{x}+q)\in\mathcal{S}} \|\bar{x}-x^k\|_\infty = O(\mu_k^{1/2}+\mu_k). \tag{16}$$

Now

$$\begin{aligned}
\operatorname{dist}((x^k,y^k),\mathcal{S}) &= \min_{(\bar{x},\bar{y})\in\mathcal{S}} \left\| \begin{bmatrix} \bar{x}-x^k \\ \bar{y}-y^k \end{bmatrix} \right\| \\
&= \min_{\bar{x}\,|\,(\bar{x},M\bar{x}+q)\in\mathcal{S}} \left\| \begin{bmatrix} \bar{x}-x^k \\ (M\bar{x}+q)-(Mx^k+q+r^k) \end{bmatrix} \right\| \\
&\le \min_{\bar{x}\,|\,(\bar{x},M\bar{x}+q)\in\mathcal{S}} (1+\|M\|)\|\bar{x}-x^k\| + \|r^k\| \\
&\le (1+\|M\|)\sqrt{n} \min_{\bar{x}\,|\,(\bar{x},M\bar{x}+q)\in\mathcal{S}} \|\bar{x}-x^k\|_\infty + \bar{\rho}\|r^0\|\mu_k/\mu_0 \\
&= \mathcal{O}(\mu_k^{1/2}+\mu_k),
\end{aligned}$$

giving the result. ∎

The next two results give bounds on the ratios of components of x^k to components of y^k.

Lemma 2.4 *If C_2 is the constant from Lemma 2.3, then for all $k \ge 0$,*

$$\frac{\bar{\gamma}}{C_2}\mu_k^{1/2} \le x_i^k \le C_2\mu_k^{1/2}, \qquad \frac{\bar{\gamma}}{C_2}\mu_k^{1/2} \le y_i^k \le C_2\mu_k^{1/2}, \qquad \forall i \in J, \tag{17}$$

and

$$\frac{\bar{\gamma}}{C_2^2} \le \frac{x_i^k}{y_i^k} \le \frac{C_2^2}{\bar{\gamma}} \qquad \forall i \in J. \tag{18}$$

Proof. We prove the bounds (17) only for x_i^k, $i \in J$, since the results for y_i^k are similar. Since $x_i^* = y_i^* = 0$ for all $i \in J$, we have

$$x_i^k = x_i^k - x_i^* \le \operatorname{dist}((x^k,y^k),\mathcal{S}) \le C_2\mu_k^{1/2},$$

which gives the upper bounds. For the lower bounds, we use (4) to write

$$x_i^k \ge \frac{\bar{\gamma}\mu_k}{y_i^k} \ge \frac{\bar{\gamma}\mu_k}{C_2\mu_k^{1/2}} = \frac{\bar{\gamma}}{C_2}\mu_k^{1/2}.$$

The bounds (18) follow immediately from (17). ∎

Lemma 2.5 *If C_1 is the constant from Lemma 2.2, then for all $k \geq 0$,*

$$\frac{x_i^k}{y_i^k} \leq \frac{C_1^2}{\bar{\gamma}} \mu_k, \qquad \forall i \in N, \tag{19}$$

$$\frac{y_i^k}{x_i^k} \leq \frac{C_1^2}{\bar{\gamma}} \mu_k, \qquad \forall i \in B. \tag{20}$$

$$\tag{21}$$

Proof. The proof follows immediately from (11) and (12). ∎

It is possible to prove global linear convergence and polynomial complexity for a number of algorithms that fit the standard framework without assuming nondegeneracy (see, for example, Zhang [11], Wright [7], and Ji, Potra, and Huang [2]). However, Assumption 1 is used to prove Q-superlinear convergence in Ye and Anstreicher [10], Ji, Potra, and Huang [2], and Wright [7]. We are therefore led to pose the question, Is nondegeneracy necessary for superlinear convergence? The following two results resolve this question in the affirmative.

Theorem 2.6 *Suppose that $J \neq \emptyset$. Then there is a constant $\epsilon > 0$ such that $\mu_{k+1}/\mu_k \geq \epsilon$ for all k sufficiently large.*

Proof. Suppose for contradiction that there is an infinite subsequence $\mathcal{K}$ such that $\mu_{k+1} = o(\mu_k)$ for $k \in \mathcal{K}$. Taking any index $i \in J$, we have from (17) that

$$\begin{aligned} 0 < x_i^{k+1} &\leq C_2 \mu_{k+1}^{1/2} = o(\mu_k^{1/2}), \\ 0 < y_i^{k+1} &\leq C_2 \mu_{k+1}^{1/2} = o(\mu_k^{1/2}). \end{aligned}$$

Now

$$x_i^{k+1} = x_i^k + \alpha_k \Delta x_i^k \;\Rightarrow\; \alpha_k \Delta x_i^k = -x_i^k + x_i^{k+1} = -x_i^k + o(\mu_k^{1/2})$$

and, similarly,

$$\alpha_k \Delta y_i^k = -y_i^k + o(\mu_k^{1/2}).$$

From (5), we have

$$\begin{aligned} & y_i^k \Delta x_i^k + x_i^k \Delta y_i^k = -x_i^k y_i^k + \sigma_k \mu_k \\ & \quad \Rightarrow \quad y_i^k(\alpha_k \Delta x_i^k) + x_i^k(\alpha_k \Delta y_i^k) = -\alpha_k x_i^k y_i^k + \sigma_k \alpha_k \mu_k \\ & \quad \Rightarrow \quad -2 x_i^k y_i^k + o(\mu_k^{1/2}) x_i^k + o(\mu_k^{1/2}) y_i^k = -\alpha_k x_i^k y_i^k + \sigma_k \alpha_k \mu_k. \end{aligned}$$

If we divide this last expression by $x_i^k y_i^k$, we obtain

$$-2 + \frac{o(\mu_k^{1/2})}{x_i^k} + \frac{o(\mu_k^{1/2})}{y_i^k} = -\alpha_k + \sigma_k \alpha_k \frac{\mu_k}{x_i^k y_i^k},$$

and using the lower bound in (17) for x_i^k and y_i^k, we conclude that

$$2 - \alpha_k = -\sigma_k \alpha_k \frac{\mu_k}{x_i^k y_i^k} + o(1) \leq -\sigma_k \alpha_k / n + o(1).$$

If we take the limit for $k \in \mathcal{K}$, $k \to \infty$, $2 - \alpha_k$ is bounded below by 1, while the right-hand side of the above inequality will eventually be less than any positive constant. This gives a contradiction, and therefore we cannot have $\mu_{k+1} = o(\mu_k)$. ■

Corollary 2.7 *When the assumptions of Theorem 2.6 hold, the sequence $\{\mu_k\}$ cannot exhibit ℓ-step superlinear convergence to zero, for any integer $\ell \geq 1$.*

Proof. The proof follows from $\mu_{k+\ell}/\mu_k \geq \epsilon^\ell$ for all sufficiently large k. ■

As well as precluding Q-superlinear convergence of infeasible-interior-point algorithms such as the one in Wright [7], Corollary 2.7 shows that 2-step superlinear convergence is not possible for predictor-corrector algorithms such as the one in Ji, Potra, and Huang [2] applied to degenerate problems.

3 Infeasible Algorithms: Finite Termination

In this section, we propose a technique for estimating the index sets B, N, and J and for performing a projection from the current iterate onto the solution set $\mathcal{S}$. We prove that finite termination can be achieved from any sufficiently advanced iterate (that is, when k is sufficiently large). The projection scheme is similar to that of Ye [9] for linear programming.

In the interests of generality, we make a number of assumptions on the iterates (x^k, y^k) themselves, rather than on the algorithm used to generate them. This strategy allows the analysis of this section to be applied to algorithms such as the one in [8], in addition to algorithms that fall within the standard framework of Section 2.

We require the following bounds to be satisfied for all sufficiently large k:

$$i \in J \quad \Rightarrow \quad \frac{1}{C_3} \leq \frac{x_i^k}{y_i^k} \leq C_3, \qquad 0 < x_i^k \leq C_2 \mu_k^{1/2}, \qquad 0 < y_i^k \leq C_2 \mu_k^{1/2}, \tag{22a}$$

$$i \in B \quad \Rightarrow \quad 0 < \frac{y_i^k}{x_i^k} \leq C_4 \mu_k, \qquad x_i^k \geq C_5, \tag{22b}$$

$$i \in N \quad \Rightarrow \quad 0 < \frac{x_i^k}{y_i^k} \leq C_4 \mu_k, \qquad y_i^k \geq C_5, \tag{22c}$$

$$\|(x^k, y^k)\| \leq C_b, \tag{22d}$$

$$\frac{\|r^k\|}{\mu_k} \leq \bar{\rho} \frac{\|r^0\|}{\mu_0}, \tag{22e}$$

$$\lim_k \mu_k = 0, \tag{22f}$$

where C_3, C_4, C_5, C_b, and $\bar{\rho}$ are positive constants with $C_3 > 1$ and $C_4 > 1$.

The inequalities (22a), (22b), and (22c) immediately suggest a scheme for estimating B, N, and J. We define estimates B_k, N_k, and J_k by

$$\begin{aligned} B_k &= \left\{ i \;\middle|\; \frac{y_i^k}{x_i^k} \le \min(1/2, \mu_k^{1/2}) \right\}, \\ N_k &= \left\{ i \;\middle|\; \frac{x_i^k}{y_i^k} \le \min(1/2, \mu_k^{1/2}) \right\}, \\ J_k &= \{1, 2, \cdots, n\} \backslash (B_k \cup N_k). \end{aligned}$$

Lemma 3.1 *For all k sufficiently large, we have*

$$B_k = B, \qquad N_k = N, \qquad J_k = J.$$

Proof. Choose a positive integer K_1 such that for $k \ge K_1$ the bounds (22) hold and, in addition,

$$\mu_k < \min(1/(2C_4^2), 1/C_3^2).$$

Then

$$i \in B \;\Rightarrow\; \frac{y_i^k}{x_i^k} \le C_4 \mu_k \le (C_4 \mu_k^{1/2}) \mu_k^{1/2} \le \mu_k^{1/2},$$

and, since $C_4 > 1$,

$$i \in B \;\Rightarrow\; \frac{y_i^k}{x_i^k} \le C_4 \mu_k \le 1/(2C_4) < 1/2.$$

Therefore $i \in B_k$. Similar logic shows that $i \in N \Rightarrow i \in N_k$ for $k \ge K_1$. For $i \in J$, we have

$$\frac{x_i^k}{y_i^k} \ge \frac{1}{C_3} > \mu_k^{1/2} \;\Rightarrow\; i \notin N_k,$$

while

$$\frac{y_i^k}{x_i^k} \ge \frac{1}{C_3} > \mu_k^{1/2} \;\Rightarrow\; i \notin B_k.$$

Therefore $i \in J_k$, and the proof is complete. ∎

On later iterations, when the index set estimates B_k, N_k, and J_k have stopped fluctuating from one iteration to the next, the following projection subproblem can be solved in an attempt to find an exact solution to (1). The problem is an equality-constrained quadratic program and hence is easier to solve than the original LCP.

$$\text{Problem } P(k): \qquad \min_{x_{B_k}, y_{N_k}} \frac{1}{2}\|x_{B_k} - x_{B_k}^k\|^2 + \frac{1}{2}\|y_{N_k} - y_{N_k}^k\|^2$$

subject to

$$\begin{aligned} 0 &= M_{B_k,B_k} x_{B_k} + q_{B_k}, \\ 0 &= M_{J_k,B_k} x_{B_k} + q_{J_k}, \\ y_{N_k} &= M_{N_k,B_k} x_{B_k} + q_{N_k}. \end{aligned}$$

Let $(\bar{x}^k_{B_k}, \bar{y}^k_{N_k})$ denote the optimal solution of this problem. A candidate solution $(\bar{x}^k, \bar{y}^k)$ of (1) is obtained by setting

$$\bar{x}^k_i = \begin{cases} (\bar{x}_{B_k})_i & i \in B_k \\ 0 & \text{otherwise} \end{cases}, \qquad \bar{y}^k_i = \begin{cases} (\bar{y}_{N_k})_i & i \in N_k \\ 0 & \text{otherwise} \end{cases}. \tag{23}$$

Theorem 3.2 *For all k sufficiently large, the solution of Problem $P(k)$ and the construction (23) yield a solution of (1).*

Proof. Assume that k is large enough that the bounds (22) and the result of Lemma 3.1 hold. (We can subsequently refer to B_k, N_k, and J_k as B, N, and J without confusion.) For any $(x^*, y^*) \in \mathcal{S}$, we have that (x^*_B, y^*_N) satisfies the equality constraints in Problem $P(k)$, so the feasible set for $P(k)$ is nontrivial. Using a change of variable

$$\tilde{x}^k_B = x_B - x^k_B, \qquad \tilde{y}^k_N = y_N - y^k_N,$$

we can reformulate $P(k)$ as

$$\min_{\tilde{x}^k_B, \tilde{y}^k_N} \frac{1}{2}\|\tilde{x}_B\|^2 + \frac{1}{2}\|\tilde{y}_N\|^2, \tag{24}$$

subject to

$$\begin{aligned} -M_{B,B}\tilde{x}^k_B &= q_B + M_{B,B}x^k_B, \\ -M_{J,B}\tilde{x}^k_B &= q_J + M_{J,B}x^k_B, \\ \tilde{y}^k_N - M_{N,B}\tilde{x}^k_B &= q_N + M_{N,B}x^k_B - y^k_N. \end{aligned} \tag{25}$$

Since the constraints (25) are consistent, Hoffman's lemma [1] implies that there is a sequence of vector pairs $\{(\tilde{x}^k_B, \tilde{y}^k_N)\}$ such that $(\tilde{x}^k_B, \tilde{y}^k_N)$ is feasible for (25) and

$$\left\| \begin{bmatrix} \tilde{x}^k_B \\ \tilde{y}^k_N \end{bmatrix} \right\| = \mathcal{O}\left(\left\| \begin{bmatrix} q_B + M_{B,B}x^k_B \\ q_J + M_{J,B}x^k_B \\ q_N + M_{N,B}x^k_B - y^k_N \end{bmatrix} \right\| \right). \tag{26}$$

Since $(\bar{x}^k_B - x^k_B, \bar{y}^k_N - y^k_N)$ is optimal for problem (24)–(25), we have

$$\left\| \begin{bmatrix} \bar{x}^k_B - x^k_B \\ \bar{y}^k_N - y^k_N \end{bmatrix} \right\| \le \left\| \begin{bmatrix} \tilde{x}^k_B \\ \tilde{y}^k_N \end{bmatrix} \right\|. \tag{27}$$

Now, using the bounds (22), we obtain

$$\begin{aligned} &\left\| \begin{bmatrix} q_B + M_{B,B}x^k_B \\ q_J + M_{J,B}x^k_J \\ q_N + M_{N,B}x^k_B - y^k_N \end{bmatrix} \right\| \\ \le\ & \|q + Mx^k - y^k\| + \left\| \begin{bmatrix} M_{B,N}x^k_N + M_{B,J}x^k_J - y^k_B \\ M_{J,N}x^k_N + M_{J,J}x^k_J - y^k_J \\ M_{N,N}x^k_N + M_{N,J}x^k_J \end{bmatrix} \right\| \\ =\ & \|r^k\| + \mathcal{O}(\|x^k_N\| + \|x^k_J\| + \|y^k_B\| + \|y^k_J\|) \\ =\ & \mathcal{O}(\mu_k^{1/2}). \end{aligned} \tag{28}$$

It follows from (26)–(28) that there is a constant $C_8 > 0$ such that

$$\|\bar{x}_B^k - x_B^k\| \le C_8 \mu_k^{1/2}, \qquad \|\bar{y}_N^k - y_N^k\| \le C_8 \mu_k^{1/2}.$$

Hence, we have

$$\begin{aligned} \bar{x}_B^k &\ge x_B^k - C_8 \mu_k^{1/2} e \ge C_5 e - C_8 \mu_k^{1/2} e, \\ \bar{y}_N^k &\ge y_N^k - C_8 \mu_k^{1/2} e \ge C_5 e - C_8 \mu_k^{1/2} e, \end{aligned}$$

and for k sufficiently large, we conclude that $(\bar{x}_B^k, \bar{y}_N^k) > 0$. The construction (23) then yields a vector pair $(\bar{x}^k, \bar{y}^k)$ that solves (1). ■

4 Feasible Algorithms

In this section we consider only *feasible* algorithms ($r^k \equiv 0$) in which all iterates lie in the relative interior of the feasible region. Condition (6) of the standard framework of Section 2 becomes redundant. Theorem 2.6 shows that no algorithm from this framework can converge superlinearly when $J \neq \emptyset$. Hence, the best we can hope to show is that the algorithm has a linear rate of convergence that is not too slow. Specifically, if we define

$$\Lambda \triangleq \limsup_{k \to \infty} \frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} = \limsup_{k \to \infty} \frac{\mu_{k+1}}{\mu_k}, \tag{29}$$

then we hope that the value $\Lambda \in (0, 1)$ is not too close to 1.

The main result of this section — Theorem 4.1 — assumes that σ_k converges to 0 and gives a formula for the linear rate of convergence in terms of just the current iterate and the current step size.

Theorem 4.1 *Assume that $r^k = 0$ for all $k \ge 0$ and $\lim_{k \to \infty} \sigma_k = 0$. Then Λ defined by (29) satisfies*

$$\Lambda = \limsup_{k \to \infty} \left[1 - \alpha_k + \alpha_k^2 \frac{{x_J^k}^T y_J^k}{4\, {x^k}^T y^k} \right]. \tag{30}$$

The proof of this theorem is postponed until we have proved several preliminary results.

Note that if α_k converges to 1, then the rate of convergence becomes simply

$$\Lambda = \limsup_{k \to \infty} \left[\frac{{x_J^k}^T y_J^k}{4\, {x^k}^T y^k} \right] \le \frac{1}{4}.$$

This expression indicates a reasonably fast linear convergence rate, but we have not been able to design an algorithm that achieves this rate since it is difficult to enforce the condition $\lim_{k \to \infty} \alpha_k = 1$.

We now derive lower and upper bounds on the linear rate of convergence. The following result is a consequence of Theorem 4.1.

Corollary 4.2 *Assume that $r^k = 0$ for all $k \geq 0$ and $\lim_{k\to\infty} \sigma_k = 0$. If $\liminf_{k\to\infty} \alpha_k = \bar{\alpha}$, then*

$$1 - \bar{\alpha} + \frac{1}{4}\bar{\alpha}^2 L \leq \limsup_{k\to\infty} \frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} \leq 1 - \bar{\alpha} + \frac{1}{4}\bar{\alpha}^2 U, \tag{31}$$

where

$$L = \bar{\gamma}\frac{|J|}{n}, \qquad U = \bar{\gamma}\frac{|J|}{n} + (1 - \bar{\gamma}) < 1.$$

Proof. Using relation (4), we can easily show that

$$L \leq \frac{{x_J^k}^T y_J^k}{{x^k}^T y^k} \leq U. \tag{32}$$

Hence,

$$\limsup_{k\to\infty} \frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} \leq \limsup_{k\to\infty} \left[1 - \alpha_k + \frac{1}{4}\alpha_k^2 U\right] = 1 - \bar{\alpha} + \frac{1}{4}\bar{\alpha}^2 U,$$

where the last equality follows because the quadratic $1 - \alpha + (U/4)\alpha^2$ is decreasing on $[0, 1]$. We have therefore proved the second inequality of (31); the first inequality follows by a similar argument. ∎

We now concentrate our efforts on the proof of Theorem 4.1.

Lemma 4.3 *Let $a \in \mathbb{R}^p_{++}$, $b \in \mathbb{R}^p_{++}$, and $z \in \mathbb{R}^p$ be given. Let $H \subseteq \mathbb{R}^{2p}$ be a subspace with the property that*

$$(u, v) \in H \Rightarrow u^T v \geq 0. \tag{33}$$

Then, the relations

$$au + bv = z, \qquad (u, v) \in H, \tag{34}$$

have at most one solution (u, v).

Proof. Assume for contradiction that $(u^1, v^1) \in \mathbb{R}^{2p}$ and $(u^2, v^2) \in \mathbb{R}^{2p}$ are two solutions of (34). We have $(u^1 - u^2, v^1 - v^2) \in H$ since H is a subspace. Therefore, in view of (33), we have

$$(u^1 - u^2)^T (v^1 - v^2) \geq 0. \tag{35}$$

On the other hand, we obtain

$$a\,(u^1 - u^2) + b\,(v^1 - v^2) = 0, \tag{36}$$

which in turn implies

$$v^1 - v^2 = -b^{-1}a\,(u^1 - u^2), \qquad u^1 - u^2 = -a^{-1}b\,(v^1 - v^2).$$

Combining these relations with relation (35), we obtain

$$-\|(b^{-1}a)^{1/2}(u^1 - u^2)\| \geq 0, \qquad -\|(a^{-1}b)^{1/2}(v^1 - v^2)\| \geq 0.$$

It follows from these last two inequalities that $u^1 = u^2$ and $v^1 = v^2$, so our result is proved. ∎

Lemma 4.4 *There hold*

$$\Delta x_N^k = \mathcal{O}(\mu_k), \qquad \Delta y_B^k = \mathcal{O}(\mu_k); \tag{37}$$

$$\Delta x_J^k = \mathcal{O}(\mu_k^{1/2}), \qquad \Delta y_J^k = \mathcal{O}(\mu_k^{1/2}); \tag{38}$$

$$\Delta x_B^k = \mathcal{O}(\mu_k^{1/2} + \sigma_k), \qquad \Delta y_N^k = \mathcal{O}(\mu_k^{1/2} + \sigma_k). \tag{39}$$

Proof. The proof is a modification of earlier results of Ye and Anstreicher [10] and Wright [7, 8]. For completeness, we include it in the appendix. ∎

Lemma 4.5 *Assume that $r^k = 0$ for all $k \geq 0$ and $\lim_{k\to\infty} \sigma_k = 0$. Then*

$$i \in J \;\Rightarrow\; \lim_{k\to\infty} \frac{\Delta x_i^k}{x_i^k} = -\frac{1}{2}, \qquad \lim_{k\to\infty} \frac{\Delta y_i^k}{y_i^k} = -\frac{1}{2}.$$

Proof. It follows from Lemmas 2.4 and 4.4 that the sequence $\{((x_J^k)^{-1}\Delta x_J^k, (y_J^k)^{-1}\Delta y_J^k)\}$ is bounded. Let (w^x, w^y) be an accumulation point of this sequence, so that there is an infinite subsequence $\mathcal{K}$ with

$$\lim_{k\in\mathcal{K}} ((x_J^k)^{-1}\Delta x_J^k, (y_J^k)^{-1}\Delta y_J^k) = (w^x, w^y). \tag{40}$$

By further restriction of $\mathcal{K}$ if necessary, we use Lemmas 2.4 and 4.4 again to deduce that there are vector pairs (l^x, l^y) and (δ^x, δ^y) such that

$$\lim_{k\in\mathcal{K}} \frac{1}{\mu_k^{1/2}} (x_J^k, y_J^k) = (l^x, l^y) > 0, \tag{41}$$

and

$$\lim_{k\in\mathcal{K}} \frac{1}{\mu_k^{1/2}} (\Delta x_J^k, \Delta y_J^k) = (\delta^x, \delta^y). \tag{42}$$

By using Lemma 4.4, the fact that $\lim_{k\to\infty} \sigma_k = 0$, and the relations

$$\begin{aligned} x_J^k \Delta y_J^k + y_J^k \Delta x_J^k &= \sigma_k \mu_k e - x_J^k y_J^k, \\ \Delta y^k - M \Delta x^k &= 0, \end{aligned}$$

we can easily verify that

$$l^x \delta^y + l^y \delta^x = -l^x l^y, \tag{43}$$

$$I_{.J}\delta^y - M_{.J}\delta^x \in \mathit{Range}[I_{.N}, -M_{.B}]. \tag{44}$$

Let $(\bar{x}, \bar{y})$ be an arbitrary solution of (1). Then, we have

$$\begin{aligned} 0 &= y^k - Mx^k - q \\ &= y^k - Mx^k - (\bar{y} - M\bar{x}) \\ &= (y^k - \bar{y}) - M(x^k - \bar{x}). \end{aligned}$$

This relation and the fact that $\bar{x}_{J\cup N} = 0$ and $\bar{y}_{B\cup J} = 0$ then imply

$$I_{.B}\frac{y_B^k}{\mu_k^{1/2}} + I_{.J}\frac{y_J^k}{\mu_k^{1/2}} - M_{.J}\frac{x_J^k}{\mu_k^{1/2}} - M_{.N}\frac{x_N^k}{\mu_k^{1/2}} \in \; Range[I_{.N}, -M_{.B}]. \tag{45}$$

It follows from Lemma 2.2 that the first and fourth terms in the left-hand side of relation (45) converge to 0, so we obtain

$$I_{.J}l^y - M_{.J}l^x \in Range[I_{.N}, -M_{.B}]. \tag{46}$$

It is easy to verify that the subspace H of the vectors $(u, v) \in \mathbb{R}^{2|J|}$ satisfying relation (46) (with (l^x, l^y) replaced by (u, v)) has the property expressed by relation (33). Hence, by Lemma 4.3, the system defined by (43) and (44) has at most one solution. From (46), it is clear that

$$(\delta^x, \delta^y) = -\tfrac{1}{2}(l^x, l^y)$$

solves (43) and (44) and is therefore the unique solution. Hence, from (40), (41), and (42), we obtain

$$(w^x, w^y) = ((l^x)^{-1}\delta^x, (l^y)^{-1}\delta^y) = -\tfrac{1}{2}(e, e).$$

Since every accumulation point of $\{((x_J^k)^{-1}\Delta x_J^k, (y_J^k)^{-1}\Delta y_J^k)\}$ is equal to $-\frac{1}{2}(e, e)$, the result follows. ∎

We are now in a position to give the proof of Theorem 4.1.

Proof of Theorem 4.1. From the standard framework of Section 2, it is straightforward to see that

$$\frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} = 1 - \alpha_k + \sigma_k\alpha_k + \alpha_k^2\frac{{\Delta x^k}^T \Delta y^k}{{x^k}^T y^k}. \tag{47}$$

Observe that relation (30), and therefore Theorem 4.1, follows immediately from (47) once we show that

$$\frac{{\Delta x^k}^T \Delta y^k}{{x^k}^T y^k} = \frac{{x_J^k}^T y_J^k}{4\,{x^k}^T y^k} + o(1). \tag{48}$$

To show (48), define

$$t_x^k = (x_J^k)^{-1}\Delta x_J^k - \tfrac{1}{2}e, \qquad t_y^k = (y_J^k)^{-1}\Delta y_J^k - \tfrac{1}{2}e. \tag{49}$$

By Lemma 4.5, we know that $t_x^k = o(1)$ and $t_y^k = o(1)$. Using (49), we have

$$\Delta x_i^k \Delta y_i^k = x_i^k y_i^k\,[\,\tfrac{1}{2} + (t_x^k)_i\,]\,[\,\tfrac{1}{2} + (t_y^k)_i\,] \;=\; \frac{x_i^k y_i^k}{4}\,[1 + q_i^k], \tag{50}$$

where

$$q_i^k \equiv 2\,[\,(t_x^k)_i + (t_y^k)_i + 2(t_x^k)_i(t_y^k)_i\,] = o(1). \tag{51}$$

From (50), we obtain

$$\frac{\Delta x_J^{k\,T} \Delta y_J^k}{x^{k^T} y^k} = \frac{x_J^{k\,T} y_J^k}{4\, x^{k^T} y^k} + \sum_{i \in J} q_i^k \frac{x_i^k y_i^k}{4 x^{k^T} y^k} = \frac{x_J^{k\,T} y_J^k}{4\, x^{k^T} y^k} + o(1). \tag{52}$$

By using Lemma 4.4, we can easily verify that

$$\frac{\Delta x_B^{k\,T} \Delta y_B^k}{x^{k^T} y^k} = o(1), \qquad \frac{\Delta x_N^{k\,T} \Delta y_N^k}{x^{k^T} y^k} = o(1). \tag{53}$$

Relation (48) now follows by combining (52) and (53). ∎

5 Convergence of the Predictor-Corrector Algorithm

In this section we use the results of the preceding section to analyze the two-step linear rate of convergence of the predictor corrector algorithm (see, for example, [2], [5], and [10]) when applied to degenerate LCPs. We show that this rate is less than or equal to $1 - c/Q^{1/4}$ where $Q \le \min\{|J|, n - |J|\}$ and c is a positive constant that is not too small. This result shows that if $|J|$ contains only a few indices, or all but a few indices, then we can attain a linear rate of convergence that is not too slow.

Although the predictor-corrector algorithm fits the standard framework of Section 2 (and can be described accordingly by letting predictor steps be taken at even values of k and corrector steps at odd values), it is more convenient to use the description below, which closely follows [2]. For a given constant $\beta \in (0, 1)$, define

$$\mathcal{N}(\beta) \equiv \{(x, y) \ge 0 \mid y = Mx + q, \;\; \|Xy - \mu e\| \le \beta\mu\},$$

where, as earlier, $\mu \equiv x^T y/n$. For $(x, y) > 0$ with $y = Mx + q$, the following system defines a direction $(\Delta x, \Delta y)$ which is used in the description of the predictor-corrector algorithm:

$$M\Delta x - \Delta y = 0, \tag{54a}$$

$$y\Delta x + x\Delta y = \sigma\mu e - xy, \tag{54b}$$

where $\sigma \in [0, 1]$.

Predictor-Corrector Algorithm: Let the constants $\beta \in (0, 1/4]$ and $\tau \in (0, \beta]$, and a strictly feasible solution $(x^0, y^0) \in \mathcal{N}(\beta)$ be given. Set $k = 0$, and go to step 1.

(1) Compute the predictor step $(\Delta x^k, \Delta y^k)$ by solving system (54) with $(x, y) = (x^k, y^k)$ and $\sigma_k = 0$.

(2) Compute the step size $\alpha_k > 0$ by

$$\alpha_k \equiv \max\{\alpha \in [0, 1] \mid (x^k + \tilde\alpha\Delta x^k, y^k + \tilde\alpha\Delta y^k) \in \mathcal{N}(\beta + \tau), \forall\tilde\alpha \in [0, \alpha]\},$$

and set $(\hat x^k, \hat y^k) \equiv (x^k, y^k) + \alpha_k(\Delta x^k, \Delta y^k)$.

(3) Compute the corrector step $(\Delta \hat{x}^k, \Delta \hat{y}^k)$ by solving system (54) with $(x, y) = (\hat{x}^k, \hat{y}^k)$ and $\sigma_k = 1$.

(4) Compute the new iterate as

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

Set $k = k + 1$, and go to step 1.

We now state several properties of the predictor-corrector algorithm. The main properties of the centering steps are given by the following result.

Lemma 5.1 *Let $\hat{\mu}_k \equiv (\hat{x}^k)^T \hat{y}^k / n$ for all $k \geq 0$. The following statements hold for every $k \geq 0$:*

(a) $\|\Delta \hat{x}^k \Delta \hat{y}^k\| \leq \beta \hat{\mu}^k$;

(b) $(\Delta \hat{x}^k)^T \Delta \hat{y}^k \leq \hat{\mu}^k / 8$;

(c) $(x^{k+1}, y^{k+1}) = (\hat{x}^k, \hat{y}^k) + (\Delta \hat{x}^k, \Delta \hat{y}^k) \in \mathcal{N}(\beta)$;

(d) $(x^{k+1})^T y^{k+1} = (\hat{x}^k + \Delta \hat{x}^k)^T (\hat{y}^k + \Delta \hat{y}^k) \leq [1 + 1/(8n)] \, (\hat{x}^k)^T \hat{y}^k$.

Proof. Statements (a) and (c) follows from arguments similar to the ones used in Lemma 2.3 of Ji, Potra, and Huang [2], while (b) and (d) are based on Lemma 3.1 of [2]. ∎

As a consequence of Lemma 5.1, we have the following result.

Corollary 5.2 *For all $k \geq 0$, we have*

$$\frac{(x^{k+1})^T y^{k+1}}{{x^k}^T y^k} \leq \left(1 + \frac{1}{8n}\right) \left(1 - \alpha_k + \alpha_k^2 \frac{(\Delta x^k)^T \Delta y^k}{{x^k}^T y^k}\right). \tag{55}$$

Proof. The result follows immediately from Lemma 5.1(d) and the fact that

$$(\hat{x}^k)^T \hat{y}^k = (x^k + \alpha_k \Delta x^k)^T (y^k + \alpha_k \Delta y^k) = (1 - \alpha_k) {x^k}^T y^k + \alpha_k^2 (\Delta x^k)^T \Delta y^k.$$

∎

For the predictor steps, we have the following result.

Lemma 5.3 *The following statements hold for every $k \geq 0$:*

(a) $\|\Delta x^k \Delta y^k\| \leq (x^k)^T y^k / 2$;

(b) $(\Delta x^k)^T \Delta y^k \leq (x^k)^T y^k / 4$;

(c) *The step size* α_k *satisfies*

$$\alpha_k \geq \left(\frac{\tau}{n}\right)^{1/2}$$

and

$$\alpha_k \geq \frac{2}{1+\sqrt{1+4\|z^k\|/\tau}}, \tag{56}$$

where

$$z^k \equiv \frac{1}{\mu_k}\left(\Delta x^k \Delta y^k - \frac{(\Delta x^k)^T \Delta y^k}{n} e\right); \tag{57}$$

(d) *If* $\tau \geq 1/(16n)$, *then*

$$\frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} \leq \left(1 - \frac{\sqrt{\tau}}{2\sqrt{n}}\right) = 1 - \mathcal{O}(1/\sqrt{n}). \tag{58}$$

For a proof of Lemma 5.3, we refer the reader to Lemma 3.1, Lemma 3.2, and Theorem 3.1 of Ji, Potra and Huang [2] and Lemma 2.2 of Ye and Anstreicher [10].

It is well known that Lemma 5.3(d) implies polynomial convergence of the predictor-corrector algorithm. Our final goal — stated in the next theorem — is to provide an upper bound on the ratio of successive elements of the sequence $\{(x^k)^T y^k\}$ which is sharper than the one implied by relation (58).

Theorem 5.4 *Assume that* $\tau \geq 1/(8n)$ *and* $0 < |J| < n$. *Then,*

$$\limsup_{k\to\infty} \frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} \leq 1 - \frac{(\tau/2)^{1/2}}{2Q^{1/4}}, \tag{59}$$

where

$$Q \equiv \frac{(n-|J|)|J|}{n} \leq \min\{|J|, n-|J|\}.$$

The proof of Theorem 5.4 is postponed until we have proved the following preliminary result.

Lemma 5.5 *Assume that* $n \geq 2$. *For all* k *sufficiently large, we have*

$$\|z^k\| \leq Q^{1/2}/2. \tag{60}$$

Proof. Define $w^k \in \mathbb{R}^n$ as

$$w^k = \begin{pmatrix} w_B^k \\ w_N^k \\ w_J^k \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \\ x_J^k s_J^k \end{pmatrix}.$$

Using Lemmas 4.4 and 4.5, we can easily verify that

$$\|\mu_k z^k\| = \left\|\Delta x^k \Delta y^k - \frac{(\Delta x^k)^T \Delta y^k}{n} e\right\| = \frac{1}{4}\left\|w^k - \frac{(x_J^k)^T y_J^k}{n} e\right\| + o(\mu_k). \tag{61}$$

Since $(x^k, y^k) \in \mathcal{N}(\beta)$, a simple argument shows that

$$\left\|x_J^k y_J^k - \frac{(x_J^k)^T y_J^k}{|J|} e\right\| \leq \|x_J^k y_J^k - \mu_k e\| \leq \beta\mu_k.$$

Also,

$$(x_J^k)^T y_J^k \leq (1+\beta)|J|\mu_k,$$

so we obtain

$$\begin{aligned}
&\left\|w^k - \frac{(x_J^k)^T y_J^k}{n} e\right\|^2 \\
&= \left\|x_J^k y_J^k - \frac{(x_J^k)^T y_J^k}{n} e\right\|^2 + (n-|J|)\left(\frac{(x_J^k)^T y_J^k}{n}\right)^2 \\
&= \left\|x_J^k y_J^k - \frac{(x_J^k)^T y_J^k}{|J|} e\right\|^2 + \left\|\frac{(x_J^k)^T y_J^k}{|J|} e - \frac{(x_J^k)^T y_J^k}{n} e\right\|^2 + (n-|J|)\left(\frac{(x_J^k)^T y_J^k}{n}\right)^2 \\
&\leq (\beta\mu_k)^2 + \frac{(n-|J|)^2}{n^2|J|}({x_J^k}^T y_J^k)^2 + (n-|J|)\left(\frac{(x_J^k)^T y_J^k}{n}\right)^2 \\
&= (\beta\mu_k)^2 + \frac{n-|J|}{n|J|}({x_J^k}^T y_J^k)^2 \\
&\leq (\beta\mu_k)^2 + \frac{n-|J|}{n|J|}|J|^2(1+\beta)^2\mu_k^2 \\
&\leq \mu_k^2(1+\beta)^2\left[\beta^2 + \frac{(n-|J|)|J|}{n}\right] \\
&\leq \mu_k^2(1+\beta)^2\left[\beta^2 + Q\right].
\end{aligned}$$

Hence, for all k sufficiently large, we have from (61) that

$$\|\mu_k z^k\| \leq \mu_k \frac{(1+\beta)}{4}\left[\beta^2 + Q\right]^{1/2} + o(\mu_k) \leq \mu_k \frac{(1+\beta)}{4}\left[\beta + Q^{1/2}\right] \leq \frac{\mu_k}{2} Q^{1/2},$$

where the last inequality follows from the fact that $\beta \leq 1/4$ and $Q^{1/2} \geq 1/\sqrt{2}$. ∎

We are now in a position to give the proof of Theorem 5.4.

Proof of Theorem 5.4. Using relations (56) and (60), we obtain

$$\alpha_k \geq \frac{2}{1 + \sqrt{1 + 2Q^{1/2}/\tau}} \geq \frac{2}{2\sqrt{2Q^{1/2}/\tau}} = \frac{\sqrt{\tau/2}}{Q^{1/4}}, \tag{62}$$

where the second inequality can easily be verified by using $\tau \le 1/4$ and $Q^{1/2} \ge 1/\sqrt{2}$. Using relations (55), (62), and Lemma (5.3)(b), we conclude that for all k sufficiently large, we have

$$\begin{aligned}
\frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} &\le \left(1+\frac{1}{8n}\right)\left(1-\frac{\alpha_k}{2}\right)^2 \\
&\le \left(1+\frac{1}{8n}-\frac{\alpha_k}{2}\right)\left(1-\frac{\alpha_k}{2}\right) \\
&\le \left(1-\frac{\alpha_k}{2}\right) \\
&\le \left(1-\frac{(\tau/2)^{1/2}}{2Q^{1/4}}\right)
\end{aligned}$$

where the third inequality follows from (62) and the fact that $\tau \ge 1/(8n)$ and $Q^{1/2} \le n$. ∎

A Appendix

We prove Lemma 4.4 for the infeasible case $r^k \ne 0$, although a proof for the feasible case would suffice for the purposes of Section 4. Before doing do, we prove some useful auxiliary results. First, we recall a result due to Ye and Anstreicher [10].

Lemma A.1 *Let M be a positive semi-definite matrix, and partition M as*

$$M = \begin{pmatrix} M_{\mathcal{P}\mathcal{P}} & M_{\mathcal{P}\mathcal{Q}} \\ M_{\mathcal{Q}\mathcal{P}} & M_{\mathcal{Q}\mathcal{Q}} \end{pmatrix},$$

where the pair of index sets $\mathcal{P} \subseteq \{1, \cdots, n\}$ and $\mathcal{Q} \subseteq \{1, \cdots, n\}$ forms a nontrivial partition of $\{1, \cdots, n\}$. Then

$$\mathit{Range}\begin{pmatrix} M_{\mathcal{P}\mathcal{P}} & M_{\mathcal{P}\mathcal{Q}} \\ 0 & I \end{pmatrix} = \mathit{Range}\begin{pmatrix} M_{\mathcal{P}\mathcal{P}}^T & M_{\mathcal{Q}\mathcal{P}}^T \\ 0 & -I \end{pmatrix}$$

As a consequence of Lemma A.1, we obtain the following result which will be explicitly used in the proof of Lemma 4.4.

Lemma A.2 *There holds*

$$\mathit{Range}\begin{pmatrix} M_{BB} & M_{BJ} & M_{BN} \\ 0 & 0 & I \end{pmatrix} = \mathit{Range}\begin{pmatrix} M_{BB}^T & M_{JB}^T & M_{NB}^T \\ 0 & 0 & -I \end{pmatrix}. \tag{63}$$

Proof. Applying Lemma A.1 with $\mathcal{P} = B \cup J$ and $\mathcal{Q} = N$, we obtain

$$\mathit{Range}\begin{pmatrix} M_{BB} & M_{BJ} & M_{BN} \\ M_{JB} & M_{JJ} & M_{JN} \\ 0 & 0 & I \end{pmatrix} = \mathit{Range}\begin{pmatrix} M_{BB}^T & M_{JB}^T & M_{NB}^T \\ M_{BJ}^T & M_{JJ}^T & M_{NJ}^T \\ 0 & 0 & -I \end{pmatrix},$$

which in turn immediately implies (63). ■

The following lemma is similar to Lemma 3.5 of Ye and Anstreicher [10] and Lemma 5.2 of Wright [7]. For this proof and the proof of Lemma 4.4, we drop the iteration index k on matrices and vectors for clarity, and define the diagonal matrix $D = X^{-1/2}Y^{1/2}$. (The principal submatrices D_N and D_B are defined in an obvious way.)

Lemma A.3 *The vector pair* $(\Delta x_B, \Delta y_N)$ *is the unique solution of the convex quadratic programming problem*

$$\min_{(w,z)} \tfrac{1}{2}\|D_B w\|^2 - \sigma_k \mu_k e^T X_B^{-1} w + \tfrac{1}{2}\|D_N^{-1} z\|^2 - \sigma_k \mu_k e^T Y_N^{-1} z, \tag{64}$$

subject to

$$\begin{array}{rcl} M_{BB} w & = & r_B - M_{BJ}\Delta x_J - M_{BN}\Delta x_N + \Delta y_B, \\ M_{JB} w & = & r_J - M_{JJ}\Delta x_J - M_{JN}\Delta x_N + \Delta y_J, \\ M_{NB} w - z & = & r_N - M_{NJ}\Delta x_J - M_{NN}\Delta x_N. \end{array} \tag{65}$$

Proof. By the Karush-Kuhn-Tucker conditions, a candidate solution (w,z) is optimal if it is feasible with respect to (65) and, in addition,

$$\left(\begin{array}{c} D_B^2 w - \sigma_k \mu_k X_B^{-1} e \\ D_N^{-2} z - \sigma_k \mu_k Y_N^{-1} e \end{array} \right) \in \mathit{Range} \left(\begin{array}{ccc} M_{BB}^T & M_{JB}^T & M_{NB}^T \\ 0 & 0 & -I \end{array} \right). \tag{66}$$

We prove the result by showing that $(\Delta x_B, \Delta y_N)$ satisfies this condition. Clearly, $(\Delta x_B, \Delta y_N)$ is feasible with respect to (65). Using (5), we have

$$\begin{array}{l} D_B^2 \Delta x_B - \sigma_k \mu_k X_B^{-1} e \\ \quad = \quad -y_B - \Delta y_B \\ \quad = \quad -y_B - (M_{BB}\Delta x_B - r_B + M_{BN}\Delta x_N + M_{BJ}\Delta x_J) \\ \quad = \quad -M_{BB}(x_B + \Delta x_B) - M_{BN}(x_N + \Delta x_N) - M_{BJ}(x_J + \Delta x_J), \end{array}$$

and

$$D_N^{-2}\Delta y_N - \sigma_k \mu_k Y_N^{-1} e = -(x_N + \Delta x_N).$$

Therefore,

$$\left(\begin{array}{c} D_B^2 w - \sigma_k \mu_k X_B^{-1} e \\ D_N^{-2} z - \sigma_k \mu_k Y_N^{-1} e \end{array} \right) \in \mathit{Range} \left(\begin{array}{ccc} M_{BB} & M_{BJ} & M_{BN} \\ 0 & 0 & I \end{array} \right). \tag{67}$$

From relations (63) and (67), it follows that $(\Delta x_B, \Delta y_N)$ satisfies (66). ■

We are now in a position to give the proof of Lemma 4.4.

Proof of Lemma 4.4. First, we show that

$$\|D\Delta x\| = \mathcal{O}(\mu_k^{1/2}), \qquad \|D^{-1}\Delta y\| = \mathcal{O}(\mu_k^{1/2}). \tag{68}$$

We introduce vector pairs $(\hat{u},\hat{v})$ and $(\bar{u},\bar{v})$ that satisfy

$$\begin{bmatrix} M & -I \\ Y & X \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}, \tag{69a}$$

$$\begin{bmatrix} M & -I \\ Y & X \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} 0 \\ -xy + \sigma_k \mu_k e \end{bmatrix}, \tag{69b}$$

so that

$$(\Delta x, \Delta y) = (\hat{u},\hat{v}) + (\bar{u},\bar{v}).$$

It follows exactly as in [8, Lemma 3.3] that

$$\|D\bar{u}\| = \mathcal{O}(\mu_k^{1/2}), \qquad \|D^{-1}\bar{v}\| = \mathcal{O}(\mu_k^{1/2}). \tag{70}$$

If we use (6) and the boundedness assumption (3), minor modifications to the proof of [8, Lemma 3.4] can be used to show

$$\|D\hat{u}\| = \mathcal{O}(\mu_k^{1/2}), \qquad \|D^{-1}\hat{v}\| = \mathcal{O}(\mu_k^{1/2}). \tag{71}$$

The relations (68) follow from (70) and (71).

We now prove (37) for Δx_N (the proof for Δy_B is similar). Taking $i \in N$, we have from (68) and (19) that

$$\begin{aligned} &\left| \left(\frac{y_i}{x_i}\right)^{1/2} \Delta x_i \right| \le \|D\Delta x\| = \mathcal{O}(\mu_k^{1/2}) \\ &\Rightarrow \quad |\Delta x_i| \le \left(\frac{x_i}{y_i}\right)^{1/2} \mathcal{O}(\mu_k^{1/2}) \le \left(\frac{C_1^2}{\bar{\gamma}}\mu_k\right)^{1/2} \mathcal{O}(\mu_k^{1/2}) = \mathcal{O}(\mu_k), \end{aligned}$$

as required.

To prove (38), we observe from (68) and (18) that for $i \in J$, we have

$$\begin{aligned} &\left| \left(\frac{y_i}{x_i}\right)^{1/2} \Delta x_i \right| = \mathcal{O}(\mu_k^{1/2}) \\ &\Rightarrow \quad |\Delta x_i| = \left(\frac{x_i}{y_i}\right)^{1/2} \mathcal{O}(\mu_k^{1/2}) = \mathcal{O}(\mu_k^{1/2}), \end{aligned}$$

as required. The proof for Δy_J is identical.

For the remaining inequality (39), we use the result of Lemma A.3. Since the feasible set for (65) is nonempty, there is a feasible vector pair $(\bar{w},\bar{z})$ such that

$$\begin{aligned} \|(\bar{w},\bar{z})\| &= \mathcal{O}(\|r\|) + \mathcal{O}(\|\Delta x_J\| + \|\Delta x_N\| + \|\Delta y_J\| + \|\Delta y_B\|) \\ &= \mathcal{O}(\mu_k^{1/2}), \end{aligned}$$

where the last equality follows from (6), (37), and (38). The remainder of the proof follows by using identical logic to that of Lemma 5.3 of Wright [7], so we omit the details. ∎

References

[1] A. J. Hoffman, *On approximate solutions of systems of linear inequalities*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 263–265.

[2] J. Ji, F. Potra, and S. Huang, *A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence*, Tech. Rep. 18, Department of Mathematics, University of Iowa, Iowa City, Iowa, August 1991.

[3] O. L. Mangasarian and T.-H. Shiau, *Error bounds for monotone linear complementarity problems*, Mathematical Programming, 36 (1986), pp. 81–89.

[4] S. Mizuno, *Polynomiality of Kojima-Meggido-Mizuno infeasible-interior-point algorithm for linear programming*, Tech. Rep. 1006, School of Operations Research and Industrial Engineering, Cornell University, 1992.

[5] S. Mizuno, M. Todd, and Y. Ye, *On adaptive-step primal-dual interior point algorithms for linear programming*, Tech. Rep. 944, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, N.Y., 1990. To appear in Mathematics of Operations Research.

[6] F. A. Potra, *An infeasible interior-point predictor-corrector algorithm for linear programming*, Tech. Rep. 26, Department of Mathematics, University of Iowa, Iowa City, Iowa, June 1992.

[7] S. J. Wright, *An infeasible-interior-point algorithm for linear complementarity problems*, Preprint MCS–P331–1092, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, October 1992.

[8] ——, *A path-following infeasible-interior-point algorithm for linear complementarity problems*, Preprint MCS–P334–1192, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, November 1992. To appear in Optimization Methods and Software.

[9] Y. Ye, *On the finite convergence of interior-point algorithms for linear programming*, Tech. Rep. 91–5, Department of Management Sciences, University of Iowa, Iowa City, February 1991.

[10] Y. Ye and K. Anstreicher, *On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP*, Tech. Rep. 91–20, Department of Management Sciences, University of Iowa, Iowa City, Iowa, November 1991.

[11] Y. Zhang, *On the convergence of an infeasible interior-point algorithm for linear programming and other problems*, Tech. Rep. 92–07, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Maryland, April 1992.