PREPRINT MCS-P418-0294, MCS DIVISION, ARGONNE NATIONAL LABORATORY

A Superquadratic Infeasible-Interior-Point Method for Linear Complementarity Problems

Stephen Wright*and Yin Zhang†

February 8, 1994

Abstract

We consider a modification of a path-following infeasible-interior-point algorithm described by Wright. In the new algorithm, we attempt to improve each new iterate by reusing the coefficient matrix factors from the latest step. We show that the modified algorithm has similar theoretical global convergence properties to the earlier algorithm, while its asymptotic convergence rate can be made superquadratic by an appropriate parameter choice.

1 Introduction

We describe an algorithm for solving the monotone linear complementarity problem (LCP), in which we aim to find a vector pair (x, y) with

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

where $q \in \mathbb{R}^n$ and M is an $n \times n$ positive semidefinite matrix. The solution set to (1) is denoted by $\mathcal{S}$, while the set $\mathcal{S}^c$ of strictly complementary solutions is defined as

$$\mathcal{S}^c = \{(x^*, y^*) \in \mathcal{S} \,|\, x^* + y^* > 0\}.$$

Our algorithm generates positive, not necessarily feasible iterates (x^k, y^k) and includes the infeasible-interior-point algorithm of Wright [10] (which is in turn based on earlier work of Zhang [12] and Wright [8]) as a special case. As in [10], the algorithm extends immediately to mixed monotone LCP with few complications.

*Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 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.

†Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland, 21228. The work of this author was based on research supported in part by the U.S. Department of Energy under Grant DE-FG02-93ER25171.

To motivate our method, we consider the following locally convergent algorithm for solving the system of nonlinear equations

$$F(z) = 0,$$

where $F : \mathbb{R}^N \to \mathbb{R}^N$ is continuously differentiable.

Choose $\tau \in (0,1)$, $I \geq 0$, $z^0 \in \mathbb{R}^N$; Set $k \leftarrow 0$;

loop:
  compute $d^k = -\nabla F(z^k)^{-1} F(z^k)$; $z \leftarrow z^k + d^k$;
  for $i = 0, 1, \cdots, I$ (improvement loop)
    compute $d = -\nabla F(z^k)^{-1} F(z)$;
    if $\|F(z+d)\| \leq \tau \|F(z)\|$
    then $z \leftarrow z + d$
    else $z^{k+1} \leftarrow z$; $k \leftarrow k+1$; go to loop;
  end for
  $z^{k+1} \leftarrow z$; $k \leftarrow k+1$; go to loop.

On each iteration, this method takes a single Newton step and follows it up with a number of Newton-like steps calculated with the old Jacobian $\nabla F(z^k)$. Simple analysis shows that if z^* is an isolated solution to the system $F(z) = 0$ with $\nabla F(z^*)$ nonsingular, and if $\|z^0 - z^*\|$ is small enough, then $\{z^k\}$ converges to z^*. Moreover, the inner loop (with iteration index i) eventually executes for all I iterations before control passes back to the main loop and, assuming that $\nabla F(z)$ is Lipschitz continuous at z^*, the convergence has Q-order $I+2$ (see, for example, [6]). Note that for each value of k, the Jacobian $\nabla F(z^k)$ is evaluated and factored only once and, in many contexts, the steps d calculated in the improvement loop are not expensive to compute.

Our algorithm, which we describe in Section 2, is identical to that of [10] in that it takes steps of two types — *safe* steps, which ensure global convergence, and *fast* steps, which ensure fast local convergence. As in our model algorithm above, each step is followed by an attempt to improve the new iterate without recomputing and refactoring the main coefficient matrix. The inner loop terminates when it fails to make significant progress, or after I iterations, whichever comes first.

The global convergence properties of the algorithm are at least as good as the algorithm of [10] in which no attempt at improvement is made. The global convergence and complexity analysis is identical to [10]. We state the relevant results, omitting most of the details, in Section 3. In Section 4, we prove some technical results about the steps computed within the inner improvement loop, and relate them to steps computed with an exact Jacobian. Our main local convergence result is proved in Section 5. We include some preliminary numerical results in Section 6.

In the remainder of the paper, we use $\mathbb{R}^n_+$ to denote the nonnegative orthant in $\mathbb{R}^n$. Subscripts on matrices and vectors indicate components, while superscripts on matrices and vectors and subscripts on scalars denote iteration numbers (usually k).

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

DISCLAIMER

2 The Algorithm

To describe the step between successive iterates, we define for any vector pair $(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ the following quantities:

$$\mu = x^T y/n, \qquad r = y - Mx - q, \qquad e = (1, 1, \cdots, 1)^T,$$

and, for any vector $x \in \mathbb{R}^n_+$,

$$X = \mathrm{diag}(x) = \mathrm{diag}(x_1, x_2, \cdots, x_n).$$

When $(x, y) = (x^k, y^k)$ (that is, the k-th iterate of the algorithm), we use r^k, μ_k, and X^k to denote r, μ, and X, respectively.

During the k-th iteration of the main loop, each search direction (u, v) and step length $\tilde{\alpha}$ is calculated as follows.

Given $(x, y) > 0$, $\tilde{\gamma} \in (0, 1)$, $\tilde{\beta} \in [0, 1)$, $\tilde{\sigma} \in [0, 1)$, solve

$$\begin{bmatrix} M & -I \\ Y^k & X^k \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r \\ -XYe + \tilde{\sigma}\mu e \end{bmatrix}. \tag{2}$$

Set

$$\tilde{\alpha} = \arg \min_{\alpha \in [0, \hat{\alpha}]} \mu(\alpha) \triangleq (x + \alpha u)^T (y + \alpha v)/n, \tag{3}$$

where $\hat{\alpha}$ is the largest number in $[0, 1]$ such that the following inequalities are satisfied for all $\alpha \in [0, \hat{\alpha}]$:

$$(x + \alpha u)^T (y + \alpha v) \geq (1 - \tilde{\beta})(1 - \alpha) x^T y, \qquad \text{if } r \neq 0 \tag{4a}$$

$$(x_j + \alpha u_j)(y_j + \alpha v_j) \geq (\tilde{\gamma}/n)(x + \alpha u)^T (y + \alpha v), \qquad j = 1, \cdots, n. \tag{4b}$$

The inequality (4b) ensures that the componentwise products $x_j y_j$ approach zero at approximately the same rate. They stay in a loosely defined neighborhood of the central path, where $x_j y_j = \mu$ for all $j = 1, \cdots, n$ — hence the term "path-following." The inequality (4a) ensures that when the current point is infeasible, the decrease in infeasibility $\|r\|$ on the current step is at least as great as the decrease in the complementarity gap μ, modulo a factor of $(1 - \tilde{\beta})$.

The basic form of our algorithm, given below, is the same as the one described in Wright [10], except for the addition of the **improve** procedure.

Given $\bar{\gamma} \in (0, \frac{1}{2})$, $\gamma_{\min}$ and $\gamma_{\max}$ with $0 < \gamma_{\min} < \gamma_{\max} \le \frac{1}{2}$, $\bar{\sigma} \in (0, \frac{1}{2})$,
$\rho \in (0, \bar{\gamma})$, and (x^0, y^0) with $x_j^0 y_j^0 \ge \gamma_{\max}\mu_0 > 0$;

$t_0 \leftarrow 1$, $\gamma_0 \leftarrow \gamma_{\max}$, $k \leftarrow 0$, $\nu_0 \leftarrow 1$;

while $\mu_k > 0$
  solve (2)–(4) with $(x, y) = (x^k, y^k)$, $\tilde{\sigma} = 0$, $\tilde{\beta} = \bar{\gamma}^{t_k}$, $\tilde{\gamma} = \gamma_{\min} + \bar{\gamma}^{t_k}(\gamma_{\max} - \gamma_{\min})$;
  if $(x^k + \tilde{\alpha}u)^T(y^k + \tilde{\alpha}v)/n \le \rho\mu_k$
  then $\beta_k \leftarrow \tilde{\beta}$, $t \leftarrow t_k + 1$, $\gamma \leftarrow \tilde{\gamma}$;
  else solve (2)–(4) with $(x, y) = (x^k, y^k)$, $\tilde{\sigma} \in [\bar{\sigma}, \frac{1}{2}]$, $\tilde{\beta} = 0$, $\tilde{\gamma} = \gamma_k$;
    $\beta_k \leftarrow 0$, $t \leftarrow t_k$, $\gamma \leftarrow \gamma_k$;
  end if
  $\alpha_k \leftarrow \tilde{\alpha}$, $\sigma_k \leftarrow \tilde{\sigma}$, $\nu \leftarrow \nu_k(1 - \alpha_k)$, $(x, y) \leftarrow (x^k, y^k) + \alpha_k(u, v)$;
  improve $((x, y), t, \nu, \gamma, (x^k, y^k))$;
  $t_{k+1} \leftarrow t$, $\nu_{k+1} \leftarrow \nu$, $\gamma_{k+1} \leftarrow \gamma$, $(x^{k+1}, y^{k+1}) \leftarrow (x, y)$, $k \leftarrow k + 1$;
end while.

We refer to the steps that are computed with $\tilde{\sigma} = 0$ as *fast* steps, because they lead to rapid local convergence, while the steps with $\tilde{\sigma} \in [\bar{\sigma}, \frac{1}{2}]$ are *safe* steps, because they ensure global convergence.

The **improve** procedure, which reuses the coefficient matrix in (2) to improve the new iterate, takes a combination of safe and fast steps, just like the main algorithm. The main difference is that the procedure is terminated if an improvement in μ of at least a factor of $\tau \in (\rho, 1)$ is not achieved. The user supplies the parameter τ and the nonnegative integer I, where I is the maximum number of steps that can be taken in **improve**.

improve $((x, y), t, \nu, \gamma, (x^k, y^k))$

Given $\tau \in (\rho, 1)$, $I \ge 0$,

for $i = 1, 2, \cdots, I$
  if $\mu = 0$ **then** return;
  solve (2)–(4) with $\tilde{\sigma} = 0$, $\tilde{\beta} = \bar{\gamma}^t$, $\tilde{\gamma} = \gamma_{\min} + \bar{\gamma}^t(\gamma_{\max} - \gamma_{\min})$;
  if $(x + \tilde{\alpha}u)^T(y + \tilde{\alpha}v)/n \le \rho\mu$
  then $t \leftarrow t + 1$, $\gamma \leftarrow \tilde{\gamma}$;
  else solve (2)–(4) with $\tilde{\sigma} \in [\bar{\sigma}, \frac{1}{2}]$, $\tilde{\beta} = 0$, $\tilde{\gamma} = \gamma$;
    if $(x + \tilde{\alpha}u)^T(y + \tilde{\alpha}v)/n > \tau\mu$ **then** return;
  end if
  $\nu \leftarrow \nu(1 - \tilde{\alpha})$, $(x, y) \leftarrow (x, y) + \tilde{\alpha}(u, v)$;
end for.

In the special case $I = 0$, **improve** is vacuous and the algorithm reduces to the method of [10]. We refer the reader to that paper for the intuitive motivation behind the use of safe and fast steps.

The inclusion of **improve** does not alter some of the fundamental properties of the iteration sequence (x^k, y^k). We still have

$$r^k = \nu_k r^0 \tag{5}$$

and also the following result, which is similar to Lemma 3.1 of [9].

Lemma 2.1 *Suppose that the initial point is infeasible, that is, $r^0 \neq 0$. Then the positive constant $\hat{\beta}$ defined by*

$$\hat{\beta} = \prod_{k=1}^{\infty} (1 - \bar{\gamma}^k)$$

is such that

$$\mu_k \geq \hat{\beta} \nu_k \mu_0 = \hat{\beta} \frac{\|r^k\|}{\|r^0\|} \mu_0, \qquad \forall k \geq 0.$$

We also have the following result, which shows that the algorithm either terminates finitely at a solution of (1) or else generates an infinite sequence $\{(x^k, y^k)\}$ of strictly positive iterates. The proof is a simple modification of [10, Lemma 3.2] and is omitted.

Lemma 2.2 *For all iterates generated by the algorithm, we have either $(x^k, y^k) > 0$ or else $\mu_k = 0$.*

We assume throughout the remainder of the paper that finite termination does not occur, that is, all iterates (x^k, y^k) and all the intermediate points (x, y) generated in the **improve** procedure are strictly positive.

3 Global Convergence

The analysis of global convergence and polynomial complexity is nearly identical to that of [10, Section 3]. We need only note that (5) still applies and that all iterates (x^k, y^k) satisfy $x_j^k y_j^k \geq \gamma_{\min} \mu_k$, $j = 1, \cdots, n$. The intermediate points generated by **improve** have the same properties. The technical results from [10, Section 3] can therefore be applied to show that nontrivial progress is made at each safe step. The presence of **improve** and the fast steps cannot hinder (and very often speed) the convergence.

In this section we summarize the main results from [10, Section 3] and state the sole assumption required for global convergence, which is as follows.

Assumption 1 $\mathcal{S} \neq \emptyset$.

Theorem 3.1 *If a safe step is taken at iteration k, then there is a constant $\omega > 0$ such that the step length α_k has*

$$\alpha_k \geq \frac{1}{\omega}.$$

If the initial point (x^0, y^0) *is chosen as*

$$(x^0, y^0) = (\xi_x e, \xi_y e), \tag{6}$$

where

$$\xi_x \geq \|x^*\|_\infty, \qquad \xi_y \geq \|y^*\|_\infty, \qquad \xi_y \geq \|q\|_\infty, \qquad \xi_y \geq \|Me\|_\infty \xi_x = \|Mx^0\|_\infty, \tag{7}$$

for some $(x^*, y^*) \in \mathcal{S}$, *then* $\omega = O(n^2)$.

Proof. See [10, Lemma 3.4, Theorem 3.5], where a different definition of ω is used. ■

The main global convergence result is as follows.

Theorem 3.2 *The complementarity gap* μ_k *converges geometrically to zero.*

Proof. As in Wright [10, Theorem 3.6], we can show that if a safe step is taken at iteration k, we have

$$(x^k + \alpha_k u)^T (y^k + \alpha_k v)/n \leq \left(1 - \frac{1}{4\omega}\right) \mu_k,$$

while if a fast step is taken, we have

$$(x^k + \alpha_k u)^T (y^k + \alpha_k v)/n \leq \rho \mu_k.$$

Since the complementarity gap may be decreased further by **improve**, we have $\mu_{k+1} \leq (x^k + \alpha_k u)^T (y^k + \alpha_k v)/n$ and therefore

$$\mu_{k+1} \leq \max\left(1 - \frac{1}{4\omega}, \rho\right) \mu_k,$$

from which the result follows. ■

Finally, we state the polynomial complexity result.

Theorem 3.3 [10, Corollary 3.7] *Let* $\epsilon > 0$ *be given. Suppose that the starting point is defined by (6), (7), where* $\mu_0 = \xi_x \xi_y \leq 1/\epsilon^\tau$ *for some constant* $\tau \geq 0$ *independent of* n. *Then there is an integer* K_ϵ *with*

$$K_\epsilon = O(n^2 \log(1/\epsilon))$$

such that $\mu_k \leq \epsilon$ *for all* $k \geq K_\epsilon$.

4 Technical Results

In the remainder of the paper, we turn our attention to the latter stages of the algorithm. We show that the algorithm eventually takes only fast steps (that is, the **then** branch of the main conditional statement is executed). Moreover, the **improve** routine eventually takes fast steps on all I of its iterations, so that a total of $I + 1$ fast steps are taken for each factorization of the coefficient matrix in (2).

In this section, we prove some results about the steps generated in this fast phase of the algorithm. In particular, we look at the effects of the inexact coefficient matrix in (2) on the steps calculated within **improve**.

We start by defining the two assumptions for the local convergence analysis, which will be implicitly assumed to hold throughout the remainder of the paper.

Assumption 2 $\mathcal{S}^c \neq \emptyset$.

Assumption 3 $\mathcal{S}$ *is bounded.*

For monotone LCP, a sufficient condition for Assumption 3 is the existence of a strictly feasible pair $(\bar{x}, \bar{y})$ such that $\bar{y} = M\bar{x} + q$, $(\bar{x}, \bar{y}) > 0$. This can be seen from the fact that for any $(x^*, y^*) \in \mathcal{S}$

$$(x^* - \bar{x})^T(y^* - \bar{y}) = (x^* - \bar{x})^T M(x^* - \bar{x}) \geq 0,$$

implying

$$\bar{x}^T y^* + \bar{y}^T x^* \leq \bar{x}^T \bar{y}.$$

By choosing any particular strictly complementary solution (x^*, y^*), we can define index sets B and N by

$$B = \{j \mid x_j^* > 0\}, \qquad N = \{j \mid y_j^* > 0\}.$$

It is well known that the global convergence of the algorithm guarantees that the iteration sequence $\{(x^k, y^k)\}$ approaches the solution set $\mathcal{S}$ (see the error bound result of Mangasarian [2], for example). Therefore, Assumption 3 implies the boundedness of the iteration sequence $\{(x^k, y^k)\}$, as given in the following lemma.

Lemma 4.1 *There is a positive constant* C_3 *such that* $\|(x^k, y^k)\| \leq C_3$ *for all* $k \geq 0$.

The next two results are simple modifications of results from Wright [10, Section 4]. Since we will apply these results to intermediate points generated by **improve** as well as to the main iterates (x^k, y^k), we state them in a more general form than in [10]. The proofs are, however, not affected. Boundedness of the iteration sequence is not necessary for either result, and neither is Assumption 3.

Lemma 4.2 ([10, Lemma 4.1]) *Let* $(x, y) \geq 0$ *be such that*

$$r = y - Mx - q = \nu r^0 \qquad \text{for some } \nu \in [0, \tfrac{1}{2}],$$

and $\mu = x^T y/n \geq \hat{\beta}\nu\mu_0$ *for this value of* ν. *Then for some constant* $C_4 > 0$ *we have*

$$\|x_N\| \leq C_4\mu, \qquad \|y_B\| \leq C_4\mu. \tag{8}$$

Lemma 4.3 *Let* (x, y) *be any point with the properties defined in Lemma 4.2, and suppose in addition that* $x_j y_j \geq \gamma_{\min}\mu$. *Let* $(\bar{u}, \bar{v})$ *be the search direction obtained by solving*

$$\begin{bmatrix} M & -I \\ Y & X \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} r \\ -XYe + \tilde{\sigma}\mu e \end{bmatrix}, \tag{9}$$

where $\tilde{\sigma} \in [0,1)$. Then there exists a positive constant C_5 such that

$$\|\bar{u}_N\| \leq C_5\mu, \qquad \|\bar{v}_B\| \leq C_5\mu. \tag{10}$$

If in addition $\tilde{\sigma} = 0$, there is a constant $C_6 > 0$ such that

$$\|\bar{u}_B\| \leq C_6\mu, \qquad \|\bar{v}_N\| \leq C_6\mu. \tag{11}$$

Proof. Follows from Lemma 4.2 and Theorem 4.5 of [10]. ∎

We now turn to the "approximate" fast steps computed by (2), where (x,y) is either the current iterate (x^k, y^k) or some intermediate point generated in the call to **improve** at iteration k. It is obvious from the algorithm definition that we have

$$\mu = x^T y/n \leq \mu_k. \tag{12}$$

We also assume that the point (x,y) is not too far from (x^k, y^k) in the sense that there is a constant $\chi \geq 1$ independent of k such that

$$\|(x^k - x, y^k - y)\| \leq \chi\mu_k. \tag{13}$$

The following result describes some characteristics of the actual search direction (u,v) calculated from (2), partly in terms of the *exact* search direction $(\bar{u}, \bar{v})$ that satisfies (9).

Lemma 4.4 *Let (x,y) be a vector pair satisfying the assumptions of Lemma 4.3 and, in addition, the properties (12) and (13). Then if $\tilde{\sigma} = 0$, there are positive constants C_7, C_8, and C_9 independent of k and χ such that the following bounds are satisfied:*

$$\|u - \bar{u}\| \leq C_7\chi\mu, \qquad \|v - \bar{v}\| \leq C_7\chi\mu, \tag{14}$$

$$\|(u,v)\| \leq C_8\chi\mu, \tag{15}$$

$$\|\bar{u}_N - u_N\| \leq C_9\chi\mu\mu_k, \qquad \|\bar{v}_B - v_B\| \leq C_9\chi\mu\mu_k. \tag{16}$$

Proof. From (2), we have that

$$\begin{bmatrix} M & -I \\ Y^k & X^k \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r \\ -XYe \end{bmatrix}, \tag{17}$$

while from (9), we have

$$\begin{bmatrix} M & -I \\ Y & X \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} r \\ -XYe \end{bmatrix},$$

and therefore

$$\begin{bmatrix} M & -I \\ Y^k & X^k \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} r \\ -XYe + (Y^k - Y)\bar{u} + (X^k - X)\bar{v} \end{bmatrix}. \tag{18}$$

From (17) and (18) we obtain

$$\begin{bmatrix} M & -I \\ Y^k & X^k \end{bmatrix} \begin{bmatrix} \bar{u} - u \\ \bar{v} - v \end{bmatrix} = \begin{bmatrix} 0 \\ (Y^k - Y)\bar{u} + (X^k - X)\bar{v} \end{bmatrix}. \tag{19}$$

Now from (13) and Lemma 4.3, there is a constant $\bar{C}_7$ independent of k and χ such that

$$\|(Y^k - Y)\bar{u} + (X^k - X)\bar{v}\| \leq \bar{C}_7 \chi \mu \mu_k. \tag{20}$$

Defining

$$D^k = (X^k)^{-1/2}(Y^k)^{1/2},$$

and multiplying the lower block in the system (19) by $(X^kY^k)^{-1/2}$, we obtain

$$D^k(\bar{u} - u) + (D^k)^{-1}(\bar{v} - v) = (X^kY^k)^{-1/2}[(Y^k - Y)\bar{u} + (X^k - X)\bar{v}]. \tag{21}$$

Using the upper block of (19), we have $(\bar{v} - v) = M(\bar{u} - u)$, and so it follows from positive semidefiniteness of M that

$$(\bar{u} - u)^T(\bar{v} - v) \geq 0. \tag{22}$$

By taking the Euclidean norm of both sides of (21), and using (22), we have

$$\|D^k(\bar{u} - u)\|^2 + \|(D^k)^{-1}(\bar{v} - v)\|^2 \leq \|(X^kY^k)^{-1/2}\|^2 \|(Y^k - Y)\bar{u} + (X^k - X)\bar{v}\|^2,$$

Therefore

$$\begin{aligned} \|D^k(\bar{u} - u)\| &\leq \|(X^kY^k)^{-1/2}\| \|(Y^k - Y)\bar{u} + (X^k - X)\bar{v}\|, \\ \|(D^k)^{-1}(\bar{v} - v)\| &\leq \|(X^kY^k)^{-1/2}\| \|(Y^k - Y)\bar{u} + (X^k - X)\bar{v}\|. \end{aligned}$$

Now since $x_j^k y_j^k \geq \gamma_{\min}\mu_k$, we have

$$\|(X^kY^k)^{-1/2}\| = \max_{j=1,\cdots,n} (x_j^k y_j^k)^{-1/2} \leq \gamma_{\min}^{-1/2} \mu_k^{-1/2}.$$

Therefore from (20) we have

$$\|D^k(\bar{u} - u)\| \leq \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2}.$$

Taking any $j = 1, \cdots, n$, we find that

$$\left| \frac{(y_j^k)^{1/2}}{(x_j^k)^{1/2}} (\bar{u}_j - u_j) \right| \leq \|D^k(\bar{u} - u)\| \leq \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2}.$$

Hence,

$$
\begin{aligned}
|\bar{u}_j - u_j| &\le \max_{j=1,\cdots,n} \frac{(x_j^k)^{1/2}}{(y_j^k)^{1/2}} \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2} \\
&\le \max_{j=1,\cdots,n} \frac{x_j^k}{(x_j^k y_j^k)^{1/2}} \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2} \\
&\le \frac{C_3}{\gamma_{\min}^{1/2} \mu_k{}^{1/2}} \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2} \\
&\le \frac{C_7}{\sqrt{n}} \chi \mu,
\end{aligned}
$$

for C_7 defined in an obvious way. We have proved the first inequality in (14); the proof of the second inequality is similar.

For (16), we repeat the logic above to obtain for $i \in N$ that

$$
\begin{aligned}
|\bar{u}_j - u_j| &\le \frac{(x_j^k)^{1/2}}{(y_j^k)^{1/2}} \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2} \\
&\le \frac{x_j^k}{(x_j^k y_j^k)^{1/2}} \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2} \\
&\le \frac{C_4 \mu_k}{\gamma_{\min}^{1/2} \mu_k^{1/2}} \bar{C}_7 \chi \gamma_{\min}^{-1/2} \mu \mu_k^{1/2} \\
&\le \frac{C_9}{\sqrt{n}} \chi \mu \mu_k,
\end{aligned}
$$

where C_9 is defined appropriately. The bound for $\|\bar{v}_B - v_B\|$ follows similarly.

To prove (15), we have from Lemma 4.3 and (14) that

$$
\|(u, v)\| \le \|(\bar{u}, \bar{v})\| + \|(u - \bar{u}, v - \bar{v})\| \le 2(C_5 + C_6)\mu + 2C_7 \chi \mu \le C_8 \chi \mu,
$$

where we have defined $C_8 = 2(C_5 + C_6 + C_7)$ and used the assumption that $\chi \ge 1$. ■

We now state the main result of this section, in which we obtain an estimate for the step length $\tilde{\alpha}$ along a (possibly approximate) fast step direction (u, v). The point (x, y) considered in this theorem represents either the main iterate (x^k, y^k) itself or one of the intermediate points generated by **improve** during iteration k. For the purpose of this result, we define the following positive constants, all of which are independent of k and χ:

$$
\begin{aligned}
\bar{C}_{10} &= C_9(C_3 + C_6) + C_7(C_4 + C_5) \\
C_{10} &= 2(C_5 C_6 + \bar{C}_{10} + C_7 C_9) \\
C_{12} &= \frac{2C_{10}}{(1 - \bar{\gamma})(\gamma_{\max} - \gamma_{\min})} \\
C_{13} &= 2(C_3 C_9 + C_4 C_7 + C_8^2), \\
C_{14} &= C_{12} + C_{13}/n.
\end{aligned}
$$

Theorem 4.5 *Let (x, y) be a point that satisfies the assumptions of Lemma 4.4, and in addition*

$$\mu \leq \min\left(1, \frac{n}{C_{13}\chi^2}\right). \tag{23}$$

Let t be a positive integer such that for γ defined by

$$\gamma = \gamma_{\min} + \bar{\gamma}^{t-1}(\gamma_{\max} - \gamma_{\min})$$

we have $x_j y_j \geq \gamma\mu$ for $j = 1, \cdots, n$, and suppose for this value of t that

$$C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^t} \leq \rho. \tag{24}$$

Then if a fast step is attempted from the point (x, y) with

$$\tilde{\sigma} = 0, \qquad \tilde{\beta} = \bar{\gamma}^t, \qquad \tilde{\gamma} = \gamma_{\min} + \bar{\gamma}^t(\gamma_{\max} - \gamma_{\min}),$$

and the search direction (u, v) is calculated from (2), the resulting step length $\tilde{\alpha}$ obtained from (2), (3), and (4) satisfies

$$\tilde{\alpha} \geq 1 - C_{12}\chi^2 \frac{\mu_k}{\bar{\gamma}^t}.$$

Moreover, the fast step is accepted with

$$(x + \tilde{\alpha}u)^T(y + \tilde{\alpha}v)/n \leq C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^t}\mu. \tag{25}$$

Proof. The proof is in three stages. First, we show that the tests (4) are satisfied for all α in the range

$$\left[0, 1 - C_{12}\chi^2 \frac{\mu_k}{\bar{\gamma}^t}\right]. \tag{26}$$

Second, we show that $\mu(\alpha)$ defined by (3) is decreasing on the interval $\alpha \in [0, 1]$. In the third stage, we show that

$$\mu(\tilde{\alpha}) \leq C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^t}\mu \leq \rho\mu, \tag{27}$$

which proves the result.

We first consider the condition (4a). From the left-hand side, we obtain

$$\begin{aligned}
&(x + \alpha u)^T(y + \alpha v) \\
&= (x + \alpha\bar{u} + \alpha(u - \bar{u}))^T(y + \alpha\bar{v} + \alpha(v - \bar{v})) \\
&= x^T y(1 - \alpha) + \alpha^2\bar{u}^T\bar{v} + \alpha(x + \alpha\bar{u})^T(v - \bar{v}) + \alpha(y + \alpha\bar{v})^T(u - \bar{u}) \\
&\qquad + \alpha^2(u - \bar{u})^T(v - \bar{v}).
\end{aligned} \tag{28}$$

Now, using Lemma 4.1 and the inequalities (8), (10), (11), (14), (15), and (16), we have

$$|\bar{u}^T\bar{v}| \leq 2C_5C_6\mu^2 \leq 2C_5C_6\mu\mu_k,$$

and

$$\begin{aligned}|(u-\bar{u})^T(y+\alpha\bar{v})| &\le \|u_N-\bar{u}_N\|(\|y_N\|+\|\bar{v}_N\|)+\|u_B-\bar{u}_B\|(\|y_B\|+\|\bar{v}_B\|)\\ &\le C_9\chi\mu\mu_k(C_3+C_6\mu)+C_7\chi\mu^2(C_4+C_5)\\ &\le \bar{C}_{10}\chi\mu\mu_k,\end{aligned}$$

where we have used $\mu \le 1$ and $\mu \le \mu_k$ to derive the last inequality. Similarly, we have

$$|(v-\bar{v})^T(x+\alpha\bar{u})| \le \bar{C}_{10}\chi\mu\mu_k,$$

while for the remaining term in (28) we have from Lemma 4.4 that

$$|(u-\bar{u})^T(v-\bar{v})| \le 2C_7C_9\chi^2\mu^2\mu_k \le 2C_7C_9\chi^2\mu\mu_k.$$

Hence, since $\chi \ge 1$, we have from the definition of C_{10} that

$$\left|(x+\alpha u)^T(y+\alpha v)-(1-\alpha)x^Ty\right| \le C_{10}\chi^2\mu\mu_k. \tag{29}$$

Since $\tilde{\beta}=\bar{\gamma}^t$, we have that (4a) is satisfied provided that

$$C_{10}\chi^2\mu\mu_k \le (1-\alpha)\bar{\gamma}^t n\mu,$$

which is certainly true provided that

$$1-\alpha \ge \frac{C_{10}\chi^2\mu_k}{n\bar{\gamma}^t}.$$

From the definition of C_{12}, since $1-\tilde{\gamma}$ and $\gamma_{\max}-\gamma_{\min}$ both lie in the range $(0,1)$, we have

$$\frac{C_{10}\chi^2\mu_k}{n\bar{\gamma}^t} \le \frac{C_{12}\chi^2\mu_k}{\bar{\gamma}^t},$$

so the inequality (4a) certainly holds for all α in the range (26).

Turning to the second inequality (4b), we have by an argument similar to the one above that

$$(x_j+\alpha u_j)(y_j+\alpha v_j) \ge x_jy_j(1-\alpha)-C_{10}\chi^2\mu\mu_k \ge \gamma\mu(1-\alpha)-C_{10}\chi^2\mu\mu_k,$$

while from (29), we have

$$(x+\alpha u)^T(y+\alpha v)/n \le (1-\alpha)\mu + C_{10}\chi^2\mu\mu_k/n.$$

Hence, the inequality (4b) holds provided that

$$\tilde{\gamma}\mu(1-\alpha)+C_{10}\chi^2\mu\mu_k(\tilde{\gamma}/n) \le \gamma\mu(1-\alpha)-C_{10}\chi^2\mu\mu_k,$$

which is certainly true whenever the inequality

$$(\gamma-\tilde{\gamma})(1-\alpha) \ge 2C_{10}\chi^2\mu_k \tag{30}$$

holds. Since

$$\gamma - \tilde{\gamma} = \left[\gamma_{\min} + \bar{\gamma}^{t-1}(\gamma_{\max} - \gamma_{\min})\right] - \left[\gamma_{\min} + \bar{\gamma}^{t}(\gamma_{\max} - \gamma_{\min})\right] = \bar{\gamma}^{t-1}(1-\bar{\gamma})(\gamma_{\max} - \gamma_{\min}),$$

we find that (30) holds whenever

$$1 - \alpha \geq \frac{2C_{10}\chi^2\mu_k}{\bar{\gamma}^{t-1}(1-\bar{\gamma})(\gamma_{\max} - \gamma_{\min})},$$

which, by definition of C_{12}, is true for α in the range (26).

For the second part of the proof, we show that $\mu(\alpha)$ defined by (3) is decreasing on the range $\alpha \in [0,1]$. Taking the derivative, we have

$$\begin{aligned} n\mu'(\alpha) &= (x^T v + y^T u) + 2\alpha u^T v \\ &= (x^T \bar{v} + y^T \bar{u}) + x^T(v - \bar{v}) + y^T(u - \bar{u}) + 2\alpha u^T v \\ &= -x^T y + x^T(v-\bar{v}) + y^T(u - \bar{u}) + 2\alpha u^T v. \end{aligned} \tag{31}$$

However, we can use Lemma 4.1 and relations (8), (14), and (16) to obtain

$$\begin{aligned} |x^T(v-\bar{v})| &\leq \|x_B\|\|v_B - \bar{v}_B\| + \|x_N\|\|v_N - \bar{v}_N\| \\ &\leq C_3C_9\chi\mu\mu_k + C_4C_7\chi\mu^2 \\ &\leq (C_3C_9 + C_4C_7)\chi\mu\mu_k, \end{aligned} \tag{32}$$

where we have used $\mu \leq \mu_k$ in the last inequality. A similar bound can be obtained for $|y^T(u-\bar{u})|$. For the final term in (31), we have

$$|u^T v| \leq C_8^2\chi^2\mu^2 \leq C_8^2\chi^2\mu\mu_k. \tag{33}$$

Substituting these relations in (31) and using the definition of C_{13}, we obtain

$$n\mu'(\alpha) \leq [-n + C_{13}\chi^2\mu_k]\mu.$$

It follows from (23) that the term in brackets is negative and hence $\mu'(\alpha) \leq 0$ for all $\alpha \in [0,1]$.

Finally, we observe that the step length $\tilde{\alpha}$ actually selected by the procedure will be at least as long as the upper bound of (26), so using (32), (33), and the definitions of C_{13} and C_{14}, we have

$$\begin{aligned} &(x + \tilde{\alpha}u)^T(y + \tilde{\alpha}v) \\ &\leq x^T y(1-\tilde{\alpha}) + |x^T(v-\bar{v})| + |y^T(u-\bar{u})| + |u^T v| \\ &\leq (x^T y)C_{12}\chi^2\frac{\mu_k}{\bar{\gamma}^t} + 2(C_3C_9 + C_4C_7)\chi\mu\mu_k + C_8^2\chi^2\mu\mu_k \\ &\leq (x^T y)\frac{\mu_k}{\bar{\gamma}^t}[C_{12} + C_{13}/n]\chi^2 \\ &= (x^T y)\frac{\mu_k}{\bar{\gamma}^t}C_{14}\chi^2. \end{aligned}$$

Therefore (25) holds. Acceptance of this step follows from (24), since we have $(x+\tilde{\alpha}u)^T(y+\tilde{\alpha}v)/n \le \rho\mu$. ∎

We close this section with a result that is important in defining the onset of the algorithm's fast phase.

Lemma 4.6 *There is a constant $\eta < 1$ such that*

$$\frac{\mu_{k+1}}{\bar{\gamma}^{t_{k+1}}} \le \eta \frac{\mu_k}{\bar{\gamma}^{t_k}}, \qquad \forall k \ge 0. \tag{34}$$

Proof. When the safe branch of the main algorithm is taken at iteration k, we have from the proof of Theorem 3.2 that

$$(x^k+\tilde{\alpha}u)^T(y^k+\tilde{\alpha}v)/n \le \left(1-\frac{1}{4\omega}\right)\mu_k,$$

while the value of t is unaltered. It is possible that in the subsequent call to **improve**, the value of t will be incremented. Whenever this happens, we are guaranteed that the complementarity gap μ decreases by a factor of at least ρ, so the ratio $\mu/\bar{\gamma}^t$ will also decrease by a factor of at least $\rho/\bar{\gamma} < 1$. Hence, when the safe branch is taken, we have

$$\frac{\mu_{k+1}}{\bar{\gamma}^{t_{k+1}}} \le \left(1-\frac{1}{4\omega}\right)\frac{\mu_k}{\bar{\gamma}^{t_k}}.$$

When the fast branch is taken, we have $t \leftarrow t+1$ and

$$(x^k+\tilde{\alpha}u)^T(y^k+\tilde{\alpha}v)/n \le \rho\mu_k,$$

so the ratio $\mu/\bar{\gamma}^t$ decreases by a factor of at least $\rho/\bar{\gamma}$. The comments above ensure that the subsequent call to **improve** can only accentuate this decrease, so in this case we have

$$\frac{\mu_{k+1}}{\bar{\gamma}^{t_{k+1}}} \le \frac{\rho}{\bar{\gamma}}\frac{\mu_k}{\bar{\gamma}^{t_k}}.$$

The result is obtained by defining

$$\eta = \max\left(1-\frac{1}{4\omega}, \frac{\rho}{\bar{\gamma}}\right).$$

∎

5 Local Convergence

In this section, we state and prove our two main local convergence results. First, we define a threshold value of $\mu_k/\bar{\gamma}^{t_k}$ below which both the main algorithm and the procedure **improve** take only fast steps. Second, we show that the resulting superlinear convergence has Q-order $I+2$.

Theorem 5.1 *Define*

$$\chi = 2(C_5 + C_6)\exp\left(\frac{C_8\rho}{1-\rho}\right), \tag{35}$$

and let K_1 be the smallest index such that $\nu_{K_1} \le 1/2$,

$$C_{14}\chi^2 \frac{\mu_{K_1}}{\bar{\gamma}^{t_{K_1}+I}} \le \rho, \tag{36}$$

and

$$\mu_{K_1} \le \min\left(1, \frac{n}{C_{13}\chi^2}\right). \tag{37}$$

Then the fast branch is taken in the main algorithm and, moreover, I fast steps are taken in the call to **improve**.

Proof. Existence of K_1 is guaranteed by Lemma 4.6. We choose any $k \ge K_1$. Our proof proceeds by showing first that the step taken from (x^k, y^k) in the main algorithm is a fast step. We then prove by induction that I fast steps are taken inside the procedure **improve**. Our main tool in both cases is Theorem 4.5.

For the first part of the proof, we apply Theorem 4.5 with

$$(x, y) = (x^k, y^k), \qquad t = t_k, \qquad \gamma = \gamma_k. \tag{38}$$

Note that the point (x, y) satisfies the assumptions of Lemmas 4.2 and 4.3 (by definition of K_1, r^k, ν_k, etc.) and the conditions (12) and (13) (trivially). Clearly also $x_j^k y_j^k \ge \gamma_k \mu_k$ for all $j = 1, \cdots, n$, and the condition (23) also holds. Because

$$C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^t} = C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^{t_k}} \le C_{14}\chi^2 \frac{\mu_{K_1}}{\bar{\gamma}^{t_{K_1}}} \le C_{14}\chi^2 \frac{\mu_{K_1}}{\bar{\gamma}^{t_{K_j}+I}} \le \rho,$$

the condition (24) also holds. Hence the conditions of Theorem 4.5 are satisfied by the choices (38), and therefore a fast step is taken by the main algorithm.

We turn now to the procedure **improve**. Our aim is to show inductively that if (x, y) is the current vector pair at the commencement of the i-th iteration of this procedure, then

$$\|(x^k, y^k) - (x, y)\| \le \left[2(C_5 + C_6)\prod_{l=1}^{i-1}(1 + C_8\rho^l)\right]\mu_k. \tag{39}$$

Moreover, we show that a fast step is taken from this vector (x, y) during the i-th iteration of **improve**. Note for future reference that

$$\log \prod_{l=1}^{i-1}(1 + C_8\rho^l) = \sum_{l=1}^{i-1}\log(1 + C_8\rho^l) \le \sum_{l=1}^{i-1} C_8\rho^l \le \frac{C_8\rho}{1-\rho},$$

and so

$$2(C_5 + C_6)\prod_{l=1}^{i-1}(1 + C_8\rho^l) \le \chi, \qquad i = 1, \cdots, I.$$

Consider the case $i = 1$, that is, the first iteration of **improve**. We aim to use Theorem 4.5 again, so we start by checking that the point just generated by the main algorithm satisfies the assumptions of this theorem. In other words, the choices

$$(x, y) = (x^k, y^k) + \alpha_k(u, v), \qquad t = t_k + 1,$$

must be shown to satisfy these assumptions. It is easy to see that the assumptions of Lemmas 4.2 and 4.3 and the condition (12) are satisfied. To see (13), note that the fast step just taken at iteration k of the main algorithm was computed with an *exact* coefficient matrix, that is, we have $(\bar{u}, \bar{v}) = (u, v)$. Hence we can apply Lemma 4.3 to deduce that

$$\|(x^k, y^k) - (x, y)\| \leq \|(u, v)\| = \|(\bar{u}, \bar{v})\| \leq 2(C_5 + C_6)\mu_k.$$

Thus the bound (39), and therefore also (13), holds for this point (x, y). The conditions (23) and $x_j y_j \geq \gamma\mu$ clearly hold, while (24) also holds because

$$C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^t} = C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^{t_k+1}} \leq C_{14}\chi^2 \frac{\mu_k}{\bar{\gamma}^{t_k+I}} \leq C_{14}\chi^2 \frac{\mu_{K_1}}{\bar{\gamma}^{t_{K_1}+I}} \leq \rho.$$

Hence Theorem 4.5 applies, and we have shown that a fast step is taken on the first iteration of **improve**.

We now consider the general iteration i of the internal loop of **improve**. We assume that our assertions hold for iterations 1 through $i-1$. Let (x^-, y^-) denote the value of (x, y) at the start of iteration $i-1$, and let (u^-, v^-) be the search direction calculated during this iteration, while as before (x, y) is the current point at the start of iteration i. To obtain an estimate of $\|(x^k, y^k) - (x, y)\|$, we note by our inductive hypothesis (39) that

$$\|(x^k, y^k) - (x^-, y^-)\| \leq \left[2(C_5 + C_6)\prod_{l=1}^{i-2}(1 + C_8\rho^l)\right]\mu_k.$$

We now apply Lemma 4.4 to the step (u^-, v^-) taken during iteration $i-1$, with χ replaced by $2(C_5 + C_6)\prod_{l=1}^{i-2}(1 + C_8\rho^l)$, to find that

$$\begin{aligned}
&\|(x^k, y^k) - (x, y)\| \\
&\leq \|(x^k, y^k) - (x^-, y^-)\| + \|(x^-, y^-) - (x, y)\| \\
&\leq 2(C_5 + C_6)\prod_{l=1}^{i-2}(1 + C_8\rho^l)\mu_k + \|(u^-, v^-)\| \\
&\leq 2(C_5 + C_6)\prod_{l=1}^{i-2}(1 + C_8\rho^l)\mu_k + C_8\left[2(C_5 + C_6)\prod_{l=1}^{i-2}(1 + C_8\rho^l)\right]\mu^- \\
&\leq \left[2(C_5 + C_6)\prod_{l=1}^{i-2}(1 + C_8\rho^l)\right](1 + C_8\rho^{i-1})\mu_k.
\end{aligned}$$

The final inequality follows from the fact that $\mu^- \leq \rho^{i-1}\mu_k$, since (x^-, y^-) is arrived at by taking $i-1$ fast steps (one step in the main algorithm, followed by $i-2$ iterations of the

improve loop), at each of which a reduction factor of at least ρ is achieved. We have now shown that the bound (39) continues to hold at iteration i. It is easy to check that the remaining conditions required by Theorem 4.5 hold. We mention only (24), which holds for $t = t_k + i$ because

$$C_{14}\chi^2\frac{\mu_k}{\bar{\gamma}^t} = C_{14}\chi^2\frac{\mu_k}{\bar{\gamma}^{t_k+i}} \leq C_{14}\chi^2\frac{\mu_k}{\bar{\gamma}^{t_k+I}} \leq C_{14}\chi^2\frac{\mu_{K_1}}{\bar{\gamma}^{t_{K_1}+I}} \leq \rho.$$

Hence, we can apply Theorem 4.5 again to deduce that a fast step is taken at iteration i, and our result is proved. ■

Our final result is to show high-order convergence of the sequence $\{\mu_k\}$ to zero. We show that this convergence has a Q-order of at least $I+2$, that is, for any $\epsilon > 0$

$$\limsup_{k\to\infty}\frac{\mu_{k+1}}{\mu_k^{I+2-\epsilon}} = 0.$$

An equivalent characterization of the Q-order $I+2$ convergence is the inequality (40) below (see Potra [7]).

Theorem 5.2 *The subsequence $\{\mu_k\}$, $k = 0, 1, \cdots$, converges to zero with Q-order $I+2$, that is,*

$$\liminf_{k\to\infty}\frac{\log\mu_{k+1}}{\log\mu_k} \geq I+2. \tag{40}$$

Proof. Consider $k \geq K_1$. Since a fast step is taken by the main algorithm and all I iterations of **improve**, and since Theorem 4.5 applies at all $I+1$ steps, we can apply the inequality (25) $I+1$ times to bound μ_{k+1} in terms of μ_k. The process yields

$$\mu_{k+1} \leq C_{14}^{I+1}\chi^{2(I+1)}\frac{\mu_k^{I+2}}{\bar{\gamma}^{(I+1)t_k+I(I+1)/2}}. \tag{41}$$

It follows from Lemma 4.6 and (41) that

$$\frac{\mu_{k+1}}{\mu_k} \leq \frac{C_{14}^{I+1}\chi^{2(I+1)}}{\bar{\gamma}^{I(I+1)/2}}\left(\frac{\mu_k}{\bar{\gamma}^{t_k}}\right)^{I+1} \to 0, \tag{42}$$

that is, $\{\mu_k\}$ converges to zero at least Q-superlinearly.

By taking logarithms, we obtain from (41) that

$$\log\mu_{k+1} \leq \log\left(\frac{C_{14}^{I+1}\chi^{2(I+1)}}{\bar{\gamma}^{I(I+1)/2}}\right) + (I+2)\log\mu_k - (I+1)t_k\log\bar{\gamma}.$$

We will assume that k is sufficiently large such that $\mu_k < 1$. From the above,

$$\frac{\log\mu_{k+1}}{\log\mu_k} \geq I+2+\log\left(\frac{C_{14}^{I+1}\chi^{2(I+1)}}{\bar{\gamma}^{I(I+1)/2}}\right)/\log\mu_k - (I+1)\log\bar{\gamma}\frac{t_k}{\log\mu_k}. \tag{43}$$

Obviously, as $k \to \infty$, the second term in the right-hand side vanishes. If we can show that the third term also goes to zero, then the conclusion (40) follows. Since $t_k \leq (I+1)k+1$, it suffices to prove

$$\lim_{k\to\infty} \frac{k}{\log \mu_k} = 0. \tag{44}$$

Suppose otherwise. Then there exist $\xi \in (0,1)$ and a subsequence $\{\mu_k\}_{\mathcal{K}} \subset \{\mu_k\}$ such that for all $k \in \mathcal{K}$

$$\frac{k}{\log \mu_k} \leq \frac{1}{\log \xi}, \quad \text{or equivalently} \quad \xi^k \leq \mu_k.$$

From (42), there exists a positive integer J such that for all $k \geq J$, $\mu_{k+1} \leq \frac{\xi}{2}\mu_k$. Hence, for all $k > J$ and $k \in \mathcal{K}$,

$$\xi^k \leq \mu_k \leq \left(\frac{\xi}{2}\right)^{k-J} \mu_J.$$

That is, for all $k > J$ and $k \in \mathcal{K}$, $2^k \leq \mu_J 2^J/\xi^J$. This is clearly a contradiction. ∎

6 Numerical Examples

We include some preliminary numerical results that compare the behavior of our algorithm with the method of [10], in which **improve** is vacuous ($I = 0$).

Our test problems have $M = A\Lambda A^T$, where $A \in \mathbb{R}^{n\times n}$ is dense with elements drawn from a uniform distribution in $[-1,1]$, and Λ is a diagonal matrix with diagonal elements $\Lambda_{ii} = 10^{4\zeta_i}$, where ζ_i is drawn from a uniform distribution in $[0,1]$. A solution (x^*, y^*) is generated so that even-numbered components of x^* and odd-numbered components of y^* are zero, and q is chosen so that the nonzero components of both vectors are uniformly distributed in $[0,1]$.

The algorithmic constants have the following values:

$$\gamma_{\min} = 10^{-6}, \qquad \gamma_{\max} = .002, \qquad \sigma_{\min} = 10^{-3}, \qquad \sigma_{\max} = .1,$$

$$\bar{\gamma} = .25, \qquad \rho = .99\bar{\gamma}.$$

We also modify the algorithms slightly so that only safe steps are attempted when the current value of μ is greater than 1 (that is, the fast step branch of the conditional statements in both the main algorithm and **improve** is bypassed). The value of $\tilde{\sigma}$ for the safe step at iteration k is chosen as

$$\sigma_k = \mathrm{mid}(\sigma_{\min}, \mu_k/\sqrt{n}, \sigma_{\max}),$$

where mid() denotes the median of its three arguments. Termination occurs when $\mu_k \leq 10^{-10}$.

Performance of the algorithm for $\tau = .8$ and $\tau = .9$ is shown in Tables 1 and 2, respectively. We tabulate the number of factorizations (which equals the number of iterations of the algorithm), together with the total number of linear system solutions performed, and the total number of corrector steps taken in **improve**. The behavior on these random problems is not too sensitive to the choices of the parameters I and τ; the choices $I = 3$ and $\tau = .8$ would probably be good choices in general.

Table 1: Performance of the algorithm for $\tau = .8$

		$n = 10$	$n = 50$	$n = 100$
	factorizations	24	38	38
$I = 0$	solves	36	56	54
	corrector steps	0	0	0
	factorizations	20	33	34
$I = 1$	solves	59	99	97
	corrector steps	10	11	10
	factorizations	18	31	32
$I = 2$	solves	65	110	105
	corrector steps	14	17	15
	factorizations	18	31	31
$I = 3$	solves	72	112	109
	corrector steps	17	20	19
	factorizations	17	30	31
$I = 5$	solves	74	114	113
	corrector steps	22	23	23

Table 2: Performance of the algorithm for $\tau = .9$

		$n = 10$	$n = 50$	$n = 100$
	factorizations	20	33	33
$I = 1$	solves	59	99	93
	corrector steps	10	11	12
	factorizations	18	31	32
$I = 2$	solves	65	110	105
	corrector steps	14	17	15
	factorizations	17	31	31
$I = 3$	solves	74	114	109
	corrector steps	20	21	19
	factorizations	17	30	31
$I = 5$	solves	76	114	115
	corrector steps	23	24	24

7 Final Comments

In this paper, we analyze an infeasible-interior-point algorithm that reuses matrix factors to accelerate convergence. In addition to the usual global convergence properties, the new algorithm possesses a local convergence rate of Q-order $I+2$.

The idea of reusing matrix factors was utilized in a number of works on interior-point methods. Among them, Mehrotra [3] and Zhang and Zhang [11] are most closely related to the current work. Mehrotra [3] also obtained a Q-order of $I+2$ convergence result, but it is for a feasible-interior-point algorithm. Moreover, his algorithm is in the Mizuno-Todd-Ye [5] predictor-corrector framework, thus always requiring two matrix factorizations per iteration. Zhang and Zhang [11] analyzed an infeasible-interior-point algorithm with $I=1$ that asymptotically requires only one matrix factorization per iteration. However, they only obtained Q-order 2 convergence instead of Q-order 3.

The higher-order convergence rates are probably of theoretic interest only. In practice, it is difficult to observe on computer a convergence rate higher than cubic. As can be seen from our preliminary numerical results, however, the approach of reusing matrix factors does have the tendency to reduce the number of factorizations required for solving LCP problems at a price of increasing the number of back solves. Since at each iteration matrix factorization is the dominant work in comparison to back solves, the potential reduction in computational work could be significant for large-scale LCP problems. For linear programming, the practical effectiveness of reusing matrix factors is already well documented [1, 4].

References

[1] I. J. Lustig, R. E. Marsten, and D. F. Shanno, *On implementing Mehrotra's predictor-corrector interior point method for linear programming*, SIAM Journal of Optimization, 2 (1992), pp. 435–449.

[2] O. L. Mangasarian, *Error bounds for nondegenerate monotone linear complementarity problems*, Mathematical Programming, 48 (1990), pp. 437–445.

[3] S. Mehrotra, *Asymptotic convergence in a generalized predictor-corrector method*, Technical Report, Dept. of Industrial Engineering and Management Science, Northwestern University, Evanston, Ill., October 1992.

[4] ——, *On the implementation of a primal-dual interior point method*, SIAM Journal of Optimization, 2 (1992), pp. 575–601.

[5] S. Mizuno, M. Todd, and Y. Ye, *On adaptive step primal-dual interior-point algorithms for linear programming*, Mathematics of Operations Research, 18 (1993), pp. 964–981.

[6] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York and London, 1970.

[7] F. A. POTRA, *On Q-order and R-order of convergence*, Journal of Optimization Theory and Applications, 63 (1989), pp. 415–431.

[8] 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, Ill., October 1992. Mathematical Programming, to appear.

[9] ——, *A path-following infeasible-interior-point algorithm for linear complementarity problems*, Optimization Methods and Software, 2 (1993), pp. 79–106.

[10] ——, *A path-following interior-point algorithm for linear and quadratic optimization problems*, Preprint MCS–P401–1293, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill., December 1993.

[11] D. ZHANG AND Y. ZHANG, *A Mehrotra-type predictor-corrector algorithm with polynomial complexity and Q-subquadratic convergence*, Tech. Rep. 93–13, Department of Mathematics and Statistics, University of Maryland Baltimore County, September 1993.

[12] Y. ZHANG, *On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problem*, Tech. Rep. 92–07, Department of Mathematics and Statistics, University of Maryland Baltimore County, April 1992. SIAM J. Optimization, to appear.