Simplified Analysis of An $O(nL)$-Iteration Infeasible Predictor-Corrector Path-Following Method for Monotone LCP*

P. Tseng†

June 25, 1994

ABSTRACT

We give a simplified analysis of an infeasible predictor-corrector path-following method for solving monotone linear complementarity problem. This method, like those studied by Mizuno et al. and by Potra and Sheng, (i) requires two factorizations and two backsolves per iteration, (ii) can find a solution in $O(\sqrt{n}L)$ or $O(nL)$ iterations, depending on the quality of the starting point, and (iii) has local quadratic convergence, provided a strictly complementary solution exists. The method decreases the centering parameter and the infeasibility at both predictor step and corrector step, and it is flexible in that either a primal-scaling or dual-scaling or primal-dual scaling can be used for the corrector step without affecting the global and local convergence properties of the method.

KEY WORDS. Monotone linear complementarity problem, infeasible-interior-point method, global linear convergence, local quadratic convergence.

---

* This research is supported by National Science Foundation Grant CCR-9311621.

† Department of Mathematics, University of Washington, Seattle, WA 98195, U.S.A. Email: *tseng@math.washington.edu*.

1 Introduction

Motivated by their practical performance, there recently have been numerous studies of the global/local convergence rate of infeasible-interior-point methods for linear programming [2], [4], [5], [8], [9], [10], [11], [13], [16], [17], [20], [30] and for solving monotone LCP [1], [12], [14], [15], [18], [19], [23], [24], [25], [26], [28], [29], [31]. The majority of these studies focus on primal-dual path-following methods using predictor-corrector steps, for which they establish global linear convergence and/or, under the assumption that a strictly complementary solution exists, local superlinear/quadratic convergence. However, the analyses tend to be complicated and hence not easily accessible to those outside the research area. (An exception is a method of Mizuno [10] which has a simple global convergence analysis. But this method uses feasible interior-point methods as subroutines and the analysis implicitly assumes an understanding of the latter methods.) In this paper, we present a simplified global and local convergence analysis of an infeasible predictor-corrector path-following method. This method shares the same features as those of Mizuno et al. [12] and of Potra and Sheng [19], namely, it (i) requires two factorizations and two backsolves per iteration, (ii) can start anywhere in the positive orthant and, assuming a solution exists, attain global linear convergence with convergence ratio that, depending on the quality of the starting point, ranges from $1 - O(\sqrt{n})^{-1}$ to $1 - O(n)^{-1}$ (also see [2], [9], [11], [18]), (iii) attains local quadratic convergence provided a strictly complementary solution exists. This method differs from existing methods in that the centering parameter and the infeasibility are decreased at both the predictor step and the corrector step. Also, this method is modularized in that either primal-scaling or dual-scaling or primal-dual-scaling can be used for the corrector step without affecting the global/local convergence properties of the method. The predictor step may also be removed without affecting the global convergence property of the method.

We will present our results in the context of the monotone LCP (LCP is short for "linear complementarity problem"), namely, that of finding an $(x^*, y^*) \in \Re^{2n}$ satisfying

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

where $M \in \Re^{n \times n}$ is positive semidefinite and $q \in \Re^n$ are given. However, our results also extend to the horizontal LCP studied in [29] and to the case where M is a P_*-matrix [6]. We denote by S the set of solutions (possibly empty) of the monotone LCP, i.e.,

$$S = \{\ (x^*, y^*) \in \Re^{2n} \mid (x^*, y^*) \text{ satisfies } (1.1)\ \}.$$

In our notation, all vectors are column vectors and superscript T denotes transpose. We denote by $\Re^n$ the n-dimensional real vector space and by $\Re^n_{++}$ the strictly positive orthant in

$\Re^n$. We denote by e the vector in $\Re^n$ whose components are all 1. For any $x \in \Re^n$, we denote by x_i the ith component of x, by X the $n \times n$ diagonal matrix whose ith diagonal entry is x_i for all i, and by $\|x\|_1$, $\|x\|$, $\|x\|_\infty$ the 1-norm, the 2-norm and the ∞-norm, respectively, of x. For any $I \subset \{1, ..., n\}$, we denote by x_I the vector with components x_i, $i \in I$.

2 An Infeasible Predictor-Corrector Algorithm

For any $w \in (0, \infty)$, let

$$\phi_w(x, y) = \|we - Xy\|/w \qquad \forall (x, y) \in \Re^{2n}.$$

The idea of the infeasible predictor-corrector path-following method is to begin each iteration with a $w > 0$ and an $(x, y) > 0$ that is approximately centred (in the sense that $\phi_w(x, y) \le \beta_0$ for some $\beta_0 \in (0, 1)$), take a damped Newton step (predictor step) for

$$Xy = 0, \qquad Mx + q - y = 0$$

(so to attain fast local convergence) and then followed by a Newton step (corrector step) for

$$Xy = we, \qquad Mx + q - y = -s$$

with s suitably chosen (so to make (x, y) approximately centred again). (See the introduction of [12] for a survey of the history of this method.) We will focus on the following version of the method which is similar to that in [12] and [19] but differs in that w^t and $Mx^t + q - y^t$ decrease in norm at *both* the predictor step and the corrector step.

Algorithm 2.1 *Choose any* $(\beta_0, \beta_1, \beta_2) \in \Re^3$ *satisfying* $0 < \beta_0 < \beta_1 < \beta_2 < \sqrt{\beta_0} < 1$ *and any* $(x^0, y^0, w^0) \in \Re^{2n+1}_{++}$ *satisfying* $\phi_{w^0}(x^0, y^0) \le \beta_0$. *Let*

$$\eta_1 = (\beta_0 - (\beta_2)^2))/(\beta_0 + \sqrt{n}). \tag{2.1}$$

For $t = 0, 1, ...,$ *compute* $(x^{t+1}, y^{t+1}, w^{t+1})$ *from* (x^t, y^t, w^t) *according to:*
Predictor Step:

$$\hat{x}^t = x^t + \theta^t \Delta x^t, \qquad \hat{y}^t = y^t + \theta^t \Delta y^t, \qquad \hat{w}^t = (1 - \theta^t) w^t, \tag{2.2}$$

where $(\Delta x^t, \Delta y^t)$ *is the unique vector in* $\Re^{2n}$ *satisfying*

$$Y^t \Delta x^t + X^t \Delta y^t = -X^t y^t, \qquad M\Delta x^t - \Delta y^t = -(Mx^t + q - y^t), \tag{2.3}$$

and θ^t is the largest $\theta \in (0,1]$ satisfying $(x^t + \theta\Delta x^t, y^t + \theta\Delta x^t) > 0$ and

$$\phi_{(1-\theta)w^t}(x^t + \theta\Delta x^t, y^t + \theta\Delta x^t) \leq \beta_1. \tag{2.4}$$

Corrector Step:

$$x^{t+1} = \hat{x}^t + \hat{\theta}^t\Delta\hat{x}^t, \qquad y^{t+1} = \hat{y}^t + \hat{\theta}^t\Delta\hat{y}^t, \qquad w^{t+1} = (1 - \gamma^t\hat{\theta}^t)\hat{w}^t, \tag{2.5}$$

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

$$\|\hat{w}^t e - \hat{X}^t\hat{y}^t - \gamma\hat{X}^t(M\hat{x}^t + q - \hat{y}^t)\|/\hat{w}^t \leq \beta_2, \tag{2.6}$$

$(\Delta\hat{x}^t, \Delta\hat{y}^t)$ is the unique vector in $\Re^{2n}$ satisfying

$$\hat{w}^t(\hat{X}^t)^{-1}\Delta\hat{x}^t + \hat{X}^t\Delta\hat{y}^t = \hat{w}^t e - \hat{X}^t\hat{y}^t, \qquad M\Delta\hat{x}^t - \Delta\hat{y}^t = -\gamma^t(M\hat{x}^t + q - \hat{y}^t), \tag{2.7}$$

and $\hat{\theta}^t$ is any $\theta \in [1, \infty)$ satisfying

$$(\hat{x}^t + \theta\Delta\hat{x}^t, \hat{y}^t + \theta\Delta\hat{y}^t) > 0, \qquad \phi_{(1-\gamma^t\theta)\hat{w}^t}(\hat{x}^t + \theta\Delta\hat{x}^t, \hat{y}^t + \theta\Delta\hat{y}^t) \leq \beta_0. \tag{2.8}$$

We use the primal-scaling direction in the corrector step so to simplify the global convergence analysis. This direction may be replaced by either a dual-scaling (corresponding to replacing "$\hat{w}^t(\hat{X}^t)^{-1}$" and the first "$\hat{X}^t$" in (2.7) with, respectively, $\hat{Y}^t$ and $\hat{w}^t(\hat{Y}^t)^{-1}$) or a primal-dual-scaling direction (corresponding to replacing "$\hat{w}^t(\hat{X}^t)^{-1}$" in (2.7) with $\hat{Y}^t$) without affecting the global and the local convergence properties of the method. In other words, Thms. 3.1 and 4.1 still hold for these modified methods (possibly with different constants). The predictor step, which is used to accelerate local convergence, can be removed (by setting $\theta^t = 0$) without affecting the global convergence property of the method. Note that (θ^t, γ^t), which will be shown to be well defined (see the proof of Thm. 3.1), can be computed very easily. The parameter $\hat{\theta}^t$ is introduced to accelerate the convergence of the method and can always be set to 1 (see the proof of Thm. 3.1).

3 Global Linear Convergence

Our first lemma shows that the direction $(\Delta\hat{x}^t, \Delta\hat{y}^t)$ is a centering direction at $(\hat{x}^t, \hat{y}^t)$. The proof of this lemma is patterned after that of [21, Lemma 1] for a primal-scaling interior-point method (also see [22, Lemma 1]).

Lemma 3.1 *Fix any* $(\beta_0,\beta_1,\beta_2)\in\Re^3$ *satisfying* $0<\beta_0<\beta_1<\beta_2\le\sqrt{\beta_0}<1$. *Let* η_1 *be given by* (2.1). *For any* $(\hat{x}^t,\hat{y}^t,\hat{w}^t)\in\Re^{2n+1}_{++}$ *satisfying* $\phi_{\hat{w}^t}(\hat{x}^t,\hat{y}^t)\le\beta_1$, *the quantity* $(\Delta\hat{x}^t,\Delta\hat{y}^t)$ *satisfying* (2.7), *with* γ^t *being the largest* $\gamma\in(0,\eta_1]$ *satisfying* (2.6), *also satisfies*

$$(\hat{x}^t+\Delta\hat{x}^t,\hat{y}^t+\Delta\hat{y}^t)>0,\qquad \phi_{(1-\gamma^t)\hat{w}^t}(\hat{x}^t+\Delta\hat{x}^t,\hat{y}^t+\Delta\hat{y}^t)\le\beta_0.$$

Proof. For simplicity, denote $(x,y,w)=(\hat{x}^t,\hat{y}^t,\hat{w}^t)$, $(\Delta x,\Delta y)=(\Delta\hat{x}^t,\Delta\hat{y}^t)$ and $\gamma=\gamma^t$ respectively. Let $r=we-Xy$, $s=Mx+q-y$ and $d=X^{-1}\Delta x$. Then (2.7) may be rewritten as

$$wd+X\Delta y=r,\qquad MXd-\Delta y=-\gamma s.$$

Since $XMX+wI$ is positive definite and hence invertible, we can solve for d to obtain

$$d=(XMX+wI)^{-1}(r-\gamma Xs).$$

This together with the positive semidefinite property of XMX yields

$$w\|d\|^2\le d^T(XMX+wI)d=d^T(r-\gamma Xs)\le\|d\|\|r-\gamma Xs\|, \tag{3.1}$$

and it readily follows that

$$\|d\|\le\|r-\gamma Xs\|/w. \tag{3.2}$$

Let $x'=x+\Delta x$ and $y'=y+\Delta y$. By (3.2) and (2.6), we have $\|d\|\le\beta_2<1$ and so $e+d>0$. Then $x>0$ yields $x'=x+Xd=X(e+d)>0$. Also, $x'=x+Xd$ implies $X'=X+DX$, which together with $y'=y+\Delta y$ yields

$$\begin{aligned} we-X'y' &= we-(I+D)X(y+\Delta y)\\ &= wd-DX(y+\Delta y)\\ &= wDd, \end{aligned}$$

where the last two equalities each follows from the first equation in (2.7). Thus

$$\begin{aligned} \|we-X'y'\| &= w\|Dd\|\\ &\le w\|Dd\|_1\\ &= w\|d\|^2\\ &\le \|r-\gamma Xs\|^2/w\\ &\le (\beta_2)^2 w, \end{aligned}$$

where the first inequality follows from properties of the 1-norm and the 2-norm; the second inequality follows from (3.1) and (3.2); the third inequality follows from (2.6). Since $\beta_2<1$

and $x' > 0$, the above relation implies that $y' > 0$. Also, using the triangle inequality and the above relation, we obtain

$$\begin{aligned}\phi_{(1-\gamma)w}(x', y') &= \|(1-\gamma)we - X'y'\|/((1-\gamma)w) \\ &\leq \|we - X'y'\|/((1-\gamma)w) + \gamma\sqrt{n}/(1-\gamma) \\ &\leq (\beta_2)^2/(1-\gamma) + \gamma\sqrt{n}/(1-\gamma).\end{aligned}$$

It follows from $\gamma \leq \eta_1$ that the righthand side of the above relation is below β_0. ■

The next lemma is needed to bound $\|(x^t, y^t)\|$ from above and certain components of (x^t, y^t) from below. This lemma, due originally to Mizuno [9, Lemma 3.3], is widely used in the convergence analysis of infeasible-interior-point methods. For completeness, we include a short proof.

Lemma 3.2 *Assume that $S \neq \emptyset$. For any $\mu \in [0, 1]$, any $(x^*, y^*) \in S$ and any $(x, y) \in \Re^{2n}$ satisfying $(x, y) \geq 0$ and $Mx + q - y = \mu(Mx^0 + q - y^0)$, we have*

$$\mu(x^T y^0 + y^T x^0) \leq x^T y + \mu((x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*), \tag{3.3}$$
$$(1-\mu)(x^T y^* + y^T x^*) \leq x^T y + \mu((x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*). \tag{3.4}$$

Proof. We have from $Mx + q - y = \mu(Mx^0 + q - y^0)$ and $Mx^* + q = y^*$ that

$$M(x - \mu x^0 - (1-\mu)x^*) - (y - \mu y^0 - (1-\mu)y^*) = 0.$$

Multiplying both sides by $x - \mu x^0 - (1-\mu)x^*$ and using the positive semidefinite property of M yields

$$(x - \mu x^0 - (1-\mu)x^*)^T(y - \mu y^0 - (1-\mu)y^*) \geq 0,$$

which can be rearranged to obtain

$$\begin{aligned}\mu(x^T y^0 + y^T x^0) + (1-\mu)(x^T y^* + y^T x^*) &\leq x^T y + (\mu x^0 + (1-\mu)x^*)^T(\mu y^0 + (1-\mu)y^*) \\ &= x^T y + \mu^2 (x^0)^T y^0 + \mu(1-\mu)((x^*)^T y^0 + (x^0)^T y^*),\end{aligned}$$

where the equality follows from $(x^*)^T y^* = 0$. Using the nonnegativity of (x, y), (x^0, y^0) and (x^*, y^*) and the fact $0 \leq \mu \leq 1$ completes the proof. ■

By using Lemmas 3.1 and 3.2, we can now establish the global linear convergence of Algorithm 2.1.

Theorem 3.1 *Let β_0, β_1, β_2, η_1, and $\{(x^t, y^t, w^t, \hat{x}^t, \hat{y}^t, \hat{w}^t, \theta^t, \gamma^t, \hat{\theta}^t)\}_{t=0,1,\ldots}$ be generated by Algorithm 2.1. Then,*

$$(x^t, y^t) > 0, \qquad \phi_{w^t}(x^t, y^t) \le \beta_0, \tag{3.5}$$

$$w^t = \mu^t w^0, \qquad Mx^t + q - y^t = \mu^t (Mx^0 + q - y^0), \tag{3.6}$$

for all t, where

$$\mu^t = (1 - \gamma^{t-1}\hat{\theta}^{t-1})(1 - \theta^{t-1}) \cdots (1 - \gamma^0\hat{\theta}^0)(1 - \theta^0). \tag{3.7}$$

If $\{(\hat{x}^t, \hat{y}^t)\}$ is bounded, then $S \ne \emptyset$. If $S \ne \emptyset$, then $\{(\hat{x}^t, \hat{y}^t)\}$ and $\{(x^t, y^t)\}$ are bounded and, for any $(x^*, y^*) \in S$, we have $\gamma^t \ge \min\{\eta_1, \eta_2\}$ for all t, where

$$\eta_2 \;=\; \frac{(\beta_2 - \beta_1) w^0 (\min_i y_i^0) / \|Mx^0 + q - y^0\|_\infty}{(1 + \beta_1) n w^0 + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*}. \tag{3.8}$$

Proof. It is easily seen that (3.5)-(3.6) hold for $t = 0$. Assume (3.5)-(3.6) hold for some $t \ge 0$. Then, the lefthand side of (2.4) is below β_0 when $\theta = 0$, so, by the choice of θ^t and (2.2), both θ^t and $(\hat{x}^t, \hat{y}^t, \hat{w}^t)$ are well defined and satisfy $(\hat{x}^t, \hat{y}^t, \hat{w}^t) > 0$ and $\phi_{\hat{w}^t}(\hat{x}^t, \hat{y}^t) \le \beta_1$. Then, γ^t is well defined (since when $\gamma = 0$, the lefthand side of (2.6) is equal to $\phi_{\hat{w}^t}(\hat{x}^t, \hat{y}^t)$ and hence below β_1) and, by Lemma 3.1, (2.8) holds for $\theta = 1$. Thus, $\hat{\theta}^t$ is well defined (and in particular can be set to 1) and, since $\theta = \hat{\theta}^t$ satisfies (2.8), (2.5) implies (3.5) holds when t is replaced by $t + 1$. Also, we have from (2.2) and (2.5) and the second equation in (2.3) and (2.7) that

$$w^{t+1} = (1 - \gamma^t\hat{\theta}^t)(1 - \theta^t) w^t, \qquad Mx^{t+1} + q - y^{t+1} = (1 - \gamma^t\hat{\theta}^t)(1 - \theta^t)(Mx^t + q - y^t)$$

so (3.6) holds when t is replaced by $t + 1$. By induction, (3.5)-(3.6) hold for all $t \ge 0$.

For any t, we have from $\phi_{\hat{w}^t}(\hat{x}^t, \hat{y}^t) \le \beta_1$ (shown earlier) that $\hat{X}^t \hat{y}^t \le (1 + \beta_1)\hat{w}^t e$, so

$$(\hat{x}^t)^T \hat{y}^t \le (1 + \beta_1) n \hat{w}^t = (1 + \beta_1) n (1 - \theta^t) \mu^t w^0, \tag{3.9}$$

where the equality follows from (2.2) and (3.6). Also, we have from the triangle inequality and $\phi_{\hat{w}^t}(\hat{x}^t, \hat{y}^t) \le \beta_1$ that, for any $\gamma \in [0, \infty)$,

$$\begin{aligned} \|\hat{w}^t e - \hat{X}^t \hat{y}^t - \gamma \hat{X}^t (M\hat{x}^t + q - \hat{y}^t)\| / \hat{w}^t &\le \phi_{\hat{w}^t}(\hat{x}^t, \hat{y}^t) + \gamma \|\hat{X}^t (M\hat{x}^t + q - \hat{y}^t)\| / \hat{w}^t \\ &\le \beta_1 + \gamma \|\hat{X}^t (M\hat{x}^t + q - \hat{y}^t)\| / \hat{w}^t. \end{aligned}$$

Let $\bar{\gamma}^t$ be the largest γ for which the lefthand side is below β_2 (so $\gamma^t = \min\{\eta_1, \bar{\gamma}^t\}$). Then $\bar{\gamma}^t$ must exceed the largest γ for which the righthand side is below β_2. This yields

$$\bar{\gamma}^t \ge \frac{(\beta_2 - \beta_1)\hat{w}^t}{\|\hat{X}^t (M\hat{x}^t + q - \hat{y}^t)\|} \ge \frac{(\beta_2 - \beta_1)\hat{w}^t}{\|\hat{x}^t\|_1 \|M\hat{x}^t + q - \hat{y}^t\|_\infty} = \frac{(\beta_2 - \beta_1) w^0}{\|\hat{x}^t\|_1 \|Mx^0 + q - y^0\|_\infty}, \tag{3.10}$$

where the equality follows from $\hat{w}^t = (1-\theta^t)\mu^t w^0$ and $M\hat{x}^t + q - \hat{y}^t = (1-\theta^t)\mu^t(Mx^0 + q - y^0)$ (see (2.2), (2.3) and (3.6)).

Assume $\{(\hat{x}^t, \hat{y}^t)\}$ is bounded. Then (3.10) yields that $\bar{\gamma}^t \geq \eta$ for some constant $\eta > 0$ and so $\gamma^t = \min\{\eta_1, \bar{\gamma}^t\} \geq \min\{\eta_1, \eta\}$ for all t. By (3.7) and $\hat{\theta}^t \geq 1$ for all t, $\{\mu^t\} \to 0$. Since $(\hat{x}^t, \hat{y}^t) > 0$, $M\hat{x}^t + q - \hat{y}^t = (1-\theta^t)\mu^t(Mx^0 + q - y^0)$ and (3.9) holds for all t, this implies any cluster point of $\{(\hat{x}^t, \hat{y}^t)\}$ is in S, so $S \neq \emptyset$.

Assume $S \neq \emptyset$ and fix any $(x^*, y^*) \in S$. By (3.5)-(3.6), we have (3.3) holds (with $(x, y) = (\hat{x}^t, \hat{y}^t)$, $\mu = (1-\theta^t)\mu^t$), so

$$\begin{aligned}
\|\hat{x}^t\|_1(\min_i y_i^0) + \|\hat{y}^t\|_1(\min_i x_i^0) &\leq (\hat{x}^t)^T y^0 + (\hat{y}^t)^T x^0 \\
&\leq (\hat{x}^t)^T \hat{y}^t/\mu^t + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^* \\
&\leq (1+\beta_1)nw^0 + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*,
\end{aligned}$$

where the first inequality follows from the nonnegativity of $(\hat{x}^t, \hat{y}^t)$ and (x^0, y^0) and the last inequality follows from (3.9). Since $(x^0, y^0) > 0$, this shows $\{(\hat{x}^t, \hat{y}^t)\}$ is bounded. A similar argument shows $\{(x^t, y^t)\}$ is bounded. Combining the above inequality with (3.10) yields

$$\bar{\gamma}^t \geq \frac{(\beta_2 - \beta_1)w^0(\min_i y_i^0)/\|Mx^0 + q - y^0\|_\infty}{(1+\beta_1)nw^0 + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*}$$

for every t. The righthand side is exactly η_2, so $\gamma^t = \min\{\eta_1, \bar{\gamma}^t\} \geq \min\{\eta_1, \eta_2\}$. ■

Thus, if S is nonempty, Algorithm 2.1 has global linear convergence of ratio at most $1 - \min\{\eta_1, \eta_2\}$ and, by using standard techniques, can find an element of S in $O(\min\{\eta_1, \eta_2\}^{-1}L)$ iterations, where L denotes the size of the problem encoding in binary. It is easily seen that $\eta_1^{-1} = O(\sqrt{n})$, so it only remains to estimate η_2^{-1}. In the case where (x^0, y^0, w^0) is chosen so that η_2^{-1} is $O(\sqrt{n})$ (such as when in addition $Mx^0 + q - y^0 = 0$), the iteration count is $O(\sqrt{n}L)$. In the case where (x^0, y^0, w^0) is chosen by the well known choice of (see [19, Cor. 3.6])

$$x^0 = \rho_p e, \quad y^0 = \rho_d e, \quad w^0 = \rho_p\rho_d, \quad \rho_p \geq \frac{\|x^*\|_1}{n}, \quad \rho_d \geq \max\left\{\frac{\|y^*\|_1}{n}, \|\rho_p Me + q\|_\infty\right\},$$

where (x^*, y^*) is any element of S, the formula (3.8) yields

$$\eta_2^{-1} \leq 3(4+\beta_1)n/(\beta_2 - \beta_1) = O(n),$$

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

To detect whether S is empty, we can adopt a technique discussed in [19, Thm. 3.7]: In Algorithm 2.1, for every t let ρ^t be the largest ρ such that

$$(\hat{x}^t)^T y^0 + (\hat{y}^t)^T x^0 > (\hat{x}^t)^T \hat{y}^t/\mu^t + (x^0)^T y^0 + \|(x^0, y^0)\|\rho.$$

Then, by (3.3), there is no $(x^*, y^*) \in S$ satisfying $\|(x^*, y^*)\| \leq \rho^t$.

4 Local Quadratic Convergence

The key to local quadratic convergence of $\{(x^t, y^t)\}$ is the following well known result (see, for example, [12, Lem. 12], [19, Lem. 4.4], [25, Thm. 4.5], [27, Thm. 3.6]) that shows $\|(\Delta x^t, \Delta y^t)\|$ is in the order of μ^t. However, our proof is simple as it uses only (3.3)-(3.4) and Hoffman's bound [3] and yields, as a byproduct, a new error bound on the distance $d((x, y), S) := \min_{(x^*,y^*)\in S} \|(x, y) - (x^*, y^*)\|$ (cf. [7]).

Lemma 4.1 *Assume there exists an $(x^*, y^*) \in S$ satisfying $(x^*_I, y^*_J) > 0$ for some $I \cup J = \{1, ..., n\}$ and $I \cap J = \emptyset$. Fix any $\beta_0 \in (0, 1)$, any $\bar{\mu} \in (0, 1)$ and any $(x^0, y^0, w^0) \in \Re^{2n+1}_{++}$. There exists a constant $C_1 > 0$ (depending on (x^*, y^*), (x^0, y^0, w^0), β_0, $\bar{\mu}$) such that, for any $(x, y) \in \Re^{2n}_{++}$ and any $\mu \in (0, \bar{\mu}]$ satisfying $\phi_{\mu w^0}(x, y) \leq \beta_0$ and $Mx+q-y = \mu(Mx^0+q-y^0)$, we have $d((x, y), S) \leq C_1\mu$ and, moreover, the vector $(\Delta x, \Delta y)$ satisfying*

$$Y\Delta x + X\Delta y = -Xy, \qquad M\Delta x - \Delta y = -(Mx + q - y) \tag{4.1}$$

satisfies $\|(\Delta x, \Delta y)\| \leq C_1\mu$.

Proof. For convenience, we will use "$a = O(b)$" to mean $a \leq Cb$ for some constant $C > 0$ depending on (x^*, y^*), (x^0, y^0, w^0), β_0 and $\bar{\mu}$. First, since $\phi_{\mu w^0}(x, y) \leq \beta_0$, we have $(1 - \beta_0)\mu w^0 \leq x_i y_i \leq (1 + \beta_0)\mu w^0$ for all i and so

$$(\min_i x_i y_i)^{-1} = O(\mu^{-1}), \qquad x^T y = O(\mu). \tag{4.2}$$

Second, by $\mu \leq \bar{\mu}$ and the nonnegativity of (x, y) and (x^*, y^*), we have

$$\begin{aligned}
(1 - \bar{\mu}) \min_{i\in I, j\in J}\{y^*_j, x^*_i\}\|(x_J, y_I)\|_1 &\leq (1 - \mu)((x_J)^T y^*_J + (y_I)^T x^*_I) \\
&\leq (1 - \mu)(x^T y^* + y^T x^*) \\
&\leq x^T y + \mu((x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*),
\end{aligned}$$

where the last inequality follows from (3.4). Since $(x_I^*, y_J^*) > 0$, this together with (4.2) shows

$$\|(x_J, y_I)\|_1 = O(x^T y + \mu) = O(\mu). \tag{4.3}$$

Third, by the nonnegativity of (x, y) and (x^0, y^0) and using (3.3), we have

$$\min_i \{y_i^0, x_i^0\}(\|x\|_1 + \|y\|_1) \le x^T y^0 + y^T x^0 \le x^T y/\mu + (x^0)^T y^0 + (x^*)^T y^0 + (x^0)^T y^*$$

so $(x^0, y^0) > 0$ and (4.2) imply $\|x\|_1 = O(1)$. Fourth, note the following two linear systems:

$$\begin{aligned} M\xi + q - \psi = \mu(Mx^0 + q - y^0), \quad \xi_J = x_J, \ \psi_I = y_I, \quad \xi \ge 0, \ \psi \ge 0, \\ M\xi + q - \psi = 0, \quad \xi_J = 0, \ \psi_I = 0, \quad \xi \ge 0, \ \psi \ge 0, \end{aligned} \tag{4.4}$$

both have solution ((x, y) and (x^*, y^*), respectively) so, by a result of Hoffman [3]), the system (4.4) has a solution $(\bar{x}, \bar{y})$ (which is easily seen to be in S) that satisfies

$$\|(x - \bar{x}, y - \bar{y})\| = O(\|(\mu(Mx^0 + q - y^0), x_J, y_I)\|). \tag{4.5}$$

Since $(\bar{x}, \bar{y}) \in S$ so $\bar{X}\bar{y} = 0$ and $M\bar{x} + q = \bar{y}$, we can rewrite (4.1) as

$$Y(x + \Delta x - \bar{x}) + X(y + \Delta y - \bar{y}) = (X - \bar{X})(y - \bar{y}), \quad M(x + \Delta x - \bar{x}) - (y + \Delta y - \bar{y}) = 0.$$

Substituting the second equation into the first gives

$$(YX + XMX)X^{-1}(x + \Delta x - \bar{x}) = (X - \bar{X})(y - \bar{y}),$$

and multiplying both sides by $X^{-1}(x + \Delta x - \bar{x})$ and using the positive semidefinite property of XMX yields

$$\begin{aligned} (\min_i x_i y_i)\|X^{-1}(x + \Delta x - \bar{x})\|^2 &\le (x + \Delta x - \bar{x})^T X^{-1}(YX + XMX)X^{-1}(x + \Delta x - \bar{x}) \\ &= (x + \Delta x - \bar{x})^T X^{-1}(X - \bar{X})(y - \bar{y}) \\ &\le \|X^{-1}(x + \Delta x - \bar{x})\|\|(X - \bar{X})(y - \bar{y})\|. \end{aligned}$$

Thus

$$\|X^{-1}(x + \Delta x - \bar{x})\| \le \|(X - \bar{X})(y - \bar{y})\| / \min_i x_i y_i$$

and since the lefthand side is greater than or equal to $\|x + \Delta x - \bar{x}\| / \|x\|_\infty$, we have

$$\begin{aligned} \|x + \Delta x - \bar{x}\| &\le \|x\|_\infty \|(X - \bar{X})(y - \bar{y})\| / \min_i x_i y_i \\ &= O(\|x - \bar{x}\|\|y - \bar{y}\|/\mu) = O(\mu), \end{aligned}$$

where the first equality follows from (4.2) and $\|x\| = O(1)$, and the second equality follows from $\|(x - \bar{x}, y - \bar{y})\| = O(\mu)$ (see (4.3) and (4.5)). Since $\|x - \bar{x}\| = O(\mu)$, the above relation shows $\|\Delta x\| = O(\mu)$. From the second equation in (4.1), we have

$$\|\Delta y\| = \|M\Delta x + Mx + q - y\| = \|M\Delta x + \mu(Mx^0 + q - y^0)\|,$$

and so $\|\Delta y\| = O(\mu)$. Finally, since $(\bar{x}, \bar{y}) \in S$, we have $d((x, y), S) \leq \|(x-\bar{x}, y-\bar{y})\| = O(\mu)$. ∎

By using Lemma 4.1, the local quadratic convergence of Algorithm 2.1 readily follows. The proof is fairly standard and is included for completeness.

Theorem 4.1 *Assume there exists an $(x^*, y^*) \in S$ satisfying $x^*_I > 0$, $y^*_J > 0$ for some $I \cup J = \{1, ..., n\}$ and $I \cap J = \emptyset$. Let β_0, β_1, β_2, η_1, and $\{(x^t, y^t, w^t, \hat{x}^t, \hat{y}^t, \hat{w}^t, \theta^t, \gamma^t, \hat{\theta}^t)\}_{t=0,1,...}$ be generated by Algorithm 2.1. Then, there exist constants $C_1 > 0$ and $C_2 > 0$ (depending on (x^*, y^*), (x^0, y^0, w^0), β_0, β_1, η_1, and η_2 given by (3.8)) such that*

$$d((x^t, y^t), S) \leq C_1\mu^t \quad \textit{and} \quad \mu^{t+1} \leq C_2(\mu^t)^2 \qquad \forall t \geq 1.$$

Proof. By Theorem 3.1, we have that the assumption of Lemma 4.1 is satisfied by $\bar{\mu} = 1 - \min\{\eta_1, \eta_2\}$ and $(x, y, \mu) = (x^t, y^t, \mu^t)$ for all $t \geq 1$. Hence, by Lemma 4.1, $d((x^t, y^t), S) \leq C_1\mu^t$ and $\|(\Delta x^t, \Delta y^t)\| \leq C_1\mu^t$ for some constant $C_1 > 0$. The latter implies, for any $\theta \in (0, 1]$,

$$\begin{aligned}
\phi_{(1-\theta)w^t}(x^t + \theta\Delta x^t, y^t + \theta\Delta y^t) &= \left\| w^t e - X^t y^t - \Delta X^t \Delta y^t \theta^2/(1-\theta) \right\| / w^t \\
&\leq \phi_{w^t}(x^t, y^t) + \|\Delta X^t \Delta y^t\| \theta^2/(1-\theta) w^t \\
&\leq \beta_0 + (C_1)^2 \mu^t/(1-\theta) w^0,
\end{aligned}$$

where the equality uses the first equation in (2.3) and the last inequality also uses $\theta^2 \leq 1$ and (3.5)-(3.6). It is easily checked that the righthand side is below β_1 for all $\theta \in [0, 1 - (C_1)^2\mu^t/((\beta_1 - \beta_0)w^0)]$, so we also have $(x^t + \theta\Delta x^t, y^t + \theta\Delta y^t) > 0$ (since if any component of $(x^t + \theta\Delta x^t, y^t + \theta\Delta y^t)$ reaches 0, the lefthand side would exceed 1). Since θ^t is the largest $\theta \in (0, 1)$ for which the lefthand side is below β_1 and $(x^t + \theta\Delta x^t, y^t + \theta\Delta y^t) > 0$, this yields

$$\theta^t \geq 1 - (C_1)^2\mu^t/((\beta_1 - \beta_0)w^0)$$

and it readily follows from (3.7) that $\mu^{t+1} \leq (1 - \theta^t)\mu^t \leq C_2(\mu^t)^2$ for some constant $C_2 > 0$. ∎

References

[1] Billups, S. C., and Ferris, M. C., Convergence of infeasible interior-point algorithms from arbitrary starting points, Technical Report, Computer Sciences Department, University of Wisconsin, Madison (October 1993).

[2] Freund, R., An infeasible-start algorithm for linear programming whose complexity depends on the distance from the starting point to the optimal solution, Working Paper 3559-93-MSA, Sloan School of Management, Massachusetts Institute of Technology, Cambridge (April 1993).

[3] Hoffman, A. J. On approximate solutions of systems of linear inequalities, *Journal of Research at National Bureau of Standards*, 49 (1952), 263-265.

[4] Kojima, M., Basic lemmas in polynomial-time infeasible-interior-point methods for linear programs, Research Report B-268, Department of Information Sciences, Tokyo Institute of Technology, Tokyo (1993).

[5] Kojima, M., Megiddo, N., and Mizuno S., A primal-dual exterior point algorithm for linear programming, *Mathematical Programming*, 61 (1993), 261-280.

[6] Kojima, M., Megiddo, N., Noma, T., and Yoshise, A., *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Lecture Notes in Computer Science 538, Springer-Verlag, Berlin, 1991.

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

[8] Miao, J., Two infeasible interior-point predictor-corrector algorithms for linear programming, Report RRR 20-93, RUTCOR, Rutgers University, New Brunswick (1993).

[9] Mizuno, S., Polynomiality of the Kojima-Megido-Mizuno infeasible interior point algorithm for linear programming, Technical Report 1006, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, (January 1993); to appear in *Mathematical Programming*.

[10] Mizuno, S., A simple predictor-corrector infeasible-interior-point algorithm for linear programming, Preprint, The Institute of Statistical Mathematics, Tokyo (October 1993).

[11] Mizuno, S. and Jarre, F., An infeasible-interior-point algorithm using projections onto a convex set, Preprint 209, Mathematische Institute der Universität Würzburg, Würzburg (August 1993); to appear in *Annals of Operations Research*.

[12] Mizuno, S., Jarre, F. and Stoer J., A unified approach to infeasible-interior-point algorithms via geometrical linear complementarity problems, Preprint Nr. 213, Mathematische Institute der Universität Würzburg, Würzburg (April 1994).

[13] Mizuno, S., Kojima, M. and Todd, M., Infeasible-interior-point primal-dual potential-reduction algorithms for linear programming, Technical Report 1023, School of Operations Research and Industrial Engineering, Cornell University, Ithaca (August 1992); to appear in *SIAM Journal on Optimization*.

[14] Monteiro, R. D. C., and Wright, S., Local convergence of interior-point algorithm for degenerate monotone LCP, Preprint MCS-P357-0493, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne (April 1993).

[15] Monteiro, R. D. C., and Wright, S., A superlinear infeasible-interior-point affine scaling algorithm for LCP, Preprint MCS-P361-0693, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne (June 1993).

[16] Potra, F., A quadratically convergent infeasible interior-point algorithm for linear programming, Technical Report 26, Department of Mathematics, University of Iowa, Iowa City (June 1992).

[17] Potra, F., An infeasible interior-point predictor-corrector algorithm for linear programming, Technical Report 28, Department of Mathematics, University of Iowa, Iowa City (July 1992).

[18] Potra, F., An $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence, Technical Report 50, Department of Mathematics, University of Iowa, Iowa City (January 1994).

[19] Potra, F. and Sheng, R., A modified $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence, Technical Report 54, Department of Mathematics, University of Iowa, Iowa City (April 1994).

[20] Stoer, J., The complexity of an infeasible-interior-point path-following method for the solution of linear programs, *Optimization Methods and Software*, 3 (1994), 1-12.

[21] Tseng, P., A simple complexity proof for a polynomial-time linear programming algorithm, *Operations Research Letters*, 8 (1989), 155-159.

[22] Tseng, P., Complexity analysis of a linear complementarity algorithm based on a Lyapunov function, *Mathematical Programming*, 53 (1992), 297-306.

[23] Wright, S. J., An infeasible-interior-point algorithm for linear complementarity problems, Preprint MCS-P331-1092, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne (October 1992); to appear in *Mathematical Programming.*

[24] Wright, S. J., A path-following infeasible-interior-point algorithm for linear complementarity problems, *Optimization Methods and Software*, 2 (1993), 79-106.

[25] Wright, S. J., A path-following interior-point algorithm for linear and quadratic problems, Preprint MCS-P401-1293, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne (December 1993).

[26] Wright, S. J. and Zhang, Y., A superquadratic infeasible-interior-point method for linear complementary problems, Technical Report YZ94-03, Department of Mathematics and Statistics, University of Maryland, Baltimore County; Preprint MCS-P418-0294, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne (February 1994).

[27] Ye, Y. and Anstreicher, K., On quadratic and $O(\sqrt{n}L)$-iteration convergence of predictor-corrector algorithm for LCP, *Mathematical Programming*, 62 (1993), 537-551.

[28] Zhang, D. and Zhang, Y., A Mehrotra-type predictor-corrector algorithm with polynomiality and Q-subquadratic convergence, Technical Report 93-13, Department of Mathematics and Statistics, University of Maryland, Baltimore County (September 1993).

[29] Zhang, Y., On the convergence of a class of infeasible interior-point algorithm for the horizontal linear complementarity problem, *SIAM Journal on Optimization*, 4 (1994), 208-227.

[30] Zhang, Y. and Zhang, D., Superlinear convergence of infeasible interior-point methods for linear programming, Technical Report 92-15, Department of Mathematics and Statistics, University of Maryland, Baltimore County (September 1992).

[31] Zhang, Y. and Zhang, D., On polynomiality of the Mehrotra-type predictor-corrector interior-point algorithms, Technical Report 93-12, Department of Mathematics and Statistics, University of Maryland, Baltimore County (June 1993).