A quadratically convergent O n L -iteration algorithm for linear programming

Yinyu Ye, Osman Güler, Richard A. Tapia, Yin Zhang

Institutions: University of Iowa, Delft University of Technology, Rice University, University of Maryland, Baltimore County

Published on: 14 Apr 1993 - Mathematical Programming (Springer-Verlag)

Topics: Compact convergence, Modes of convergence, Convergence tests, Rate of convergence and Quadratic growth

A Quadratically Convergent $O(\sqrt{n}L$-Iteration Algorithm for Linear Programming

Y. Ye
O. Güler
R.A. Tapia
Y. Zhang

August, 1991

TR91-26

A Quadratically Convergent $O(\sqrt{n}L)$-Iteration Algorithm for Linear Programming

Y. Ye, * *O. Güler,* ** *R. A. Tapia,* † *and Y. Zhang* ‡

August 1991

Abstract

Recently, Ye et al. [16] demonstrated that Mizuno-Todd-Ye's predictor-corrector interior-point algorithm for linear programming maintains the $O(\sqrt{n}L)$-iteration complexity while exhibiting superlinear convergence of the duality gap to zero under the assumption that the iteration sequence converges, and quadratic convergence of the duality gap to zero under the assumption of nondegeneracy. In this paper we establish the quadratic convergence result without any assumption concerning the convergence of the iteration sequence or nondegeneracy. This surprising result, to our knowledge, is the first instance of a demonstration of polynomiality and superlinear (or quadratic) convergence for an interior-point algorithm which does not assume the convergence of the iteration sequence or nondegeneracy.

Key words: Linear programming, primal and dual, superlinear and quadratic convergence, polynomiality

Abbreviated title: A quadratically convergent $O(\sqrt{n}L)$ algorithm for LP

---

* *Department of Management Sciences, The University of Iowa, Iowa City, Iowa, 52242. This author was supported in part by NSF Grant DDM-8922636, the Iowa Business School Summer Grant, and the Interdisciplinary Research Grant of the University of Iowa Center for Advanced Studies.*

** *Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft, The Netherlands.*

† *Department of Mathematical Sciences and Center for Research in Parallel Computation, Rice University, Houston, Texas, 77251-1892. This author was supported in part by NSF Coop. Agr. No. CCR-8809615, AFOSR 89-0363, DOE DEFG05-86ER25017 and ARO 9DAAL03-90-G-0093.*

‡ *Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21228. This author was supported in part by NSF Grant DMS-9102761 and DOE Grant DE-FG05-91ER25100.*

1. Introduction

Consider the primal linear program (LP):

$$\min \ c^T x$$
$$s.t. \quad Ax = b, \ x \geq 0,$$

and its dual (LD):

$$\max \ b^T y$$
$$s.t. \quad A^T y + s = c, \ s \geq 0,$$

where $A \in \mathbf{R}^{m \times n}$, $c \in \mathbf{R}^n$, and $b \in \mathbf{R}^m$. We say that s is feasible for (LD) if there exists y such that (y, s) is feasible for (LD). A feasible point is said to be strictly feasible if it is feasible and positive. We say that (x, s) is a (strictly) feasible pair for (LP) and (LD) if x is (strictly) feasible for (LP) and s is (strictly) feasible for (LD). It is well-known that for a feasible pair (x, s) the duality gap is given by $x^T s$. Hence a feasible pair (x^*, s^*) is optimal if and only if

$$x_j^* s_j^* = 0 \quad \text{for} \quad j = 1, 2, ..., n.$$

Moreover, consider a sequence of strictly feasible pairs $\{(x^k, s^k)\}$ such that the duality gap sequence $(x^k)^T s^k \to 0$. Then we say that this duality gap sequence converges Q-superlinearly to zero if

$$\lim_{k \to \infty} \frac{(x^{k+1})^T s^{k+1}}{(x^k)^T s^k} = 0,$$

and Q-quadratically to zero if

$$\limsup_{k \to \infty} \frac{(x^{k+1})^T s^{k+1}}{((x^k)^T s^k)^2} < +\infty.$$

In the context of the present work it is important to emphasize that the notions of convergence, superlinear convergence, or quadratic convergence of the duality gap sequence in no way require the convergence of the iteration sequence $\{(x^k, s^k)\}$. Of course, from Hoffman's lemma [4] it follows that in a particular sense the iteration sequence converges to the optimal solution set with the corresponding R-rate.

Recently, there has been an exciting outbreak of activity in the area of constructing primal-dual interior-point algorithms for either the linear programming problem (LP), or the (monotone) linear complementarity problem (LCP) with a strict complementarity solution, that are demonstrably superlinearly or quadratically convergent. For LP, these works include Zhang et al. [17], [18], Ye et al. [16] and McShane [9]. For LCP, these works include Kojima et al. [6][7], Zhang et al. [19], and Ji et al. [5].

For the moment assume that the iteration sequence $\{(x^k, s^k)\}$ has been generated by an interior-point algorithm. Consider the following assumptions:

A0 a strictly feasible pair (x^0, s^0) exists;

A1 the iteration sequence $\{(x^k, s^k)\}$ converges;

A2 the linear program is nondegenerate.

We intend A0 and A1 to apply to both LP and LCP. It is known that A2 implies A1 when the duality gap converges to zero. Note that A0 is assumed by all of the existing primal-dual interior-point algorithms. Concerning the results mentioned above, all of the superlinear convergence results assumed A1; while all of the quadratic convergence results assumed A2. Perhaps the most striking theoretical results obtained so far can be cataloged as follows:

- Quadratic convergence for the LCP that possesses a unique and strict complementarity solution (Kojima et al. [7]).
- $O(nL)$ iteration complexity and superlinear convergence assuming A1 (Zhang and Tapia [17] for LP and Ji et al. [5] for LCP) or quadratic convergence assuming A2 (Zhang and Tapia [17] for LP).
- $O(\sqrt{n}L)$ iteration complexity and superlinear convergence assuming A1 (Ye et al. [16] and McShane [9] for LP) or quadratic convergence assuming A2 (Ye et al. [16] for LP).

In these bounds L represents the data length of the problem being solved.

Certainly, the global property of polynomiality and the local property of superlinearity are desirable. However, the degree to which A1 is restrictive is an open question at this time. Moreover, A2 is not at all realistic, since most real-world LP problems are degenerate.

In what follows we consider the Mizuno-Todd-Ye predictor-corrector algorithm analyzed by Ye et al. [16]. We show that this $O(\sqrt{n}L)$ iteration complexity algorithm actually gives quadratic convergence of the duality gap to zero without assuming either A1 or A2. (Of course we must assume A0 as usual.)

In Section 2 we review the algorithm and collect several previously established estimates. Section 3 contains several technical results. Our main convergence result is given in Section 4, and a summary and concluding remarks are contained in Section 5.

2. The Predictor-Corrector Algorithm

In this section, we briefly describe the predictor-corrector LP algorithm of Mizuno et al. [11]. We employ the notation $X = \text{diag}(x)$, $S = \text{diag}(s)$, etc. and let Ω denote the collection of all feasible pairs (x, s). Consider the set

$$\mathcal{N}(\alpha) = \{(x,s) \in \Omega : \|Xs - \mu e\| \leq \alpha\mu \quad \text{and} \quad \mu = x^T s/n\},$$

where $\|.\|$ represents the l_2 norm, e is the vector of all ones, and α is a constant between 0 and 1.

To begin with choose a constant $0 < \beta \leq 1/4$ (a typical choice would be 1/4). All search directions d_x, d_s, and d_y will be defined as solutions of the following system of linear equations (Kojima et al. [8])

$$\begin{aligned} Xd_s + Sd_x &= \gamma\mu e - Xs \\ Ad_x &= 0 \\ A^T d_y + d_s &= 0 \end{aligned} \tag{1}$$

for some $(x, s) \in \Omega$, where $\mu = x^T s/n$. A typical iteration of the algorithm proceeds as follows. Given $(x^k, s^k) \in \mathcal{N}(\beta)$, we solve the system (1) with $(x, s) = (x^k, s^k)$ and $\gamma = 0$. Denote the resulting directions by d_x^k and d_s^k. For some step length $\theta > 0$ let

$$x(\theta) = x^k + \theta d_x^k, \quad s(\theta) = s^k + \theta d_s^k,$$

and $\mu(\theta) = x(\theta)^T s(\theta)/n$. This is the predictor step. The specific choice for θ will be stated after we consider the following lemma that is essentially due to Mizuno et al. [11].

Lemma 2.1. Let α be a constant between 0 and 1. If there exists a positive $\theta^k \leq 1$ such that

$$\|X(\theta)s(\theta) - \mu(\theta)e\| \leq \alpha\mu(\theta) \quad \text{for all} \quad 0 \leq \theta \leq \theta^k, \tag{2}$$

then $(x(\theta^k), s(\theta^k)) \in \mathcal{N}(\alpha)$.

The proof of Lemma 2.1 follows directly from a continuity argument. Lemma 2.1 basically says that the feasibility of $(x(\theta^k), s(\theta^k))$ is guaranteed as long as (2) is satisfied. Thus, we can choose the *largest* step length $\theta^k \leq 1$ such that (2) is satisfied for $\alpha = 2\beta$, and let

$$\hat{x}^k = x(\theta^k) \quad \text{and} \quad \hat{s}^k = s(\theta^k).$$

(Note that if $\theta^k = 1$, then we obtain an optimal pair in a finite number of iterations.)

Now we solve the system (1) with $(x, s) = (\hat{x}^k, \hat{s}^k) \in \mathcal{N}(2\beta)$, $\mu = (\hat{x}^k)^T \hat{s}^k/n$, and $\gamma = 1$. Let $x^{k+1} = \hat{x}^k + d_x$ and $s^{k+1} = \hat{s}^k + d_s$. It has been proved that $(x^{k+1}, s^{k+1}) \in \mathcal{N}(\beta)$ (Lemma 3 [11]). This is the corrector step.

We are now in a position to state the algorithm.

Large Step Predictor-Corrector Algorithm (Mizuno-Todd-Ye)

By the large step predictor-corrector algorithm we mean the algorithm defined above with the step length given by the *largest* θ^k satisfying the conditions of Lemma 2.1 with $0 < \beta \leq 1/4$ and $\alpha = 2\beta$.

The choice of θ^k in the algorithm requires one to find the roots of a quartic polynomial. From the proof of our main result we will see that the choice for θ^k need not be this involved and it suffices to choose θ^k as the lower bound given in Lemma 2.2 below, as was the case in Ye et al. [16]. These comments will be stated formally as a corollary to our main theorem.

Observe that the algorithm generates a sequence of feasible pairs satisfying

$$\|X^k s^k - \mu^k e\| \leq \beta\mu^k \tag{3.1}$$

and

$$(x^{k+1})^T s^{k+1} = (\hat{x}^k)^T \hat{s}^k = (1-\theta^k)(x^k)^T s^k. \qquad (3.2)$$

For convenience, in what follows let

$$\delta^k = D_x^k d_s^k / \mu^k.$$

¿From Mizuno et al. [11] (Lemmas 1, 2, and 4) we have that

$$\|\delta^k\| \le \sqrt{2}n/4, \qquad (4.1)$$

and for $0 < \beta \le 1/4$

$$\theta^k \ge \min\left\{\frac{1}{2}, (\frac{\beta}{2\|\delta^k\|})^{1/2}\right\}. \qquad (4.2)$$

Thus, these inequalities together with (3.2) imply that the iteration complexity of the large step predictor-corrector algorithm is $O(\sqrt{n}L)$. Note that the algorithm requires that the linear system (1) be solved twice at each iteration.

¿From relation (3.2), we see that if $(1-\theta^k) \to 0$ then the duality gap $(x^k)^T s^k$ converges to zero Q-superlinearly. Moreover, if $(1-\theta^k) = O((x^k)^T s^k)$ then the duality gap converges to zero Q-quadratically. In our convergence-rate analysis, as opposed to our complexity analysis, the big O notation represents a quantity that may or may not depend on n or L, the problem data, however this dependence will not be explicitly stated. The above lower bound in (4.2) for θ^k, due to Mizuno et al., is not sufficient to demonstrate superlinear convergence since it is at most 1/2. Thus, Ye et al. [16] derived the following lower bound for θ^k.

Lemma 2.2. If θ^k is the largest θ^k satisfying the conditions of Lemma 2.1 with $\alpha = 2\beta$, then

$$\theta^k \ge \frac{2}{\sqrt{1+4\|\delta^k\|/\beta}+1}.$$

Using the bound given in Lemma 2.2, Ye et al. [16] have proved that the large step predictor-corrector algorithm maintains the $O(\sqrt{n}L)$ iteration complexity, and also gives superlinear convergence under assumption A1 or quadratic convergence under assumption A2. In the next section we will show how to remove these assumptions and actually obtain quadratic convergence for general LP problems.

3. Technical Results

At the kth predictor step if θ^k is the largest θ^k satisfying the conditions of Lemma 2.1 with $\alpha = 2\beta$, then

$$\begin{aligned}
1-\theta^k &\le 1 - \frac{2}{\sqrt{1+4\|\delta^k\|/\beta}+1} \\
&= \frac{\sqrt{1+4\|\delta^k\|/\beta}-1}{\sqrt{1+4\|\delta^k\|/\beta}+1} \\
&= \frac{4\|\delta^k\|/\beta}{(\sqrt{1+4\|\delta^k\|/\beta}+1)^2} \\
&\le \|\delta^k\|/\beta. \qquad (5)
\end{aligned}$$

Our goal is to prove that $\|\delta^k\| = O((x^k)^T s^k)$ without using assumption A1 or A2.

We first introduce several technical lemmas. For simplicity, we drop the index k and recall the linear system during the predictor step

$$\begin{aligned}
Xd_s + Sd_x &= -Xs \\
Ad_x &= 0 \qquad (6) \\
A^T d_y + d_s &= 0.
\end{aligned}$$

Let $\mu = x^T s/n$ and $z = Xs$. Then from (3.1) we must have

$$(1-\beta)\mu \le z_j \le (1+\beta)\mu \quad \text{for} \quad j = 1, 2, ..., n. \qquad (7)$$

Define $D = X^{1/2}S^{-1/2}$ and denote by Π_L the orthogonal projection onto the linear subspace L of $\mathbf{R}^n$. Denote by $N(AD)$ and $R(DA^T)$ the null space of AD and the range of DA^T, respectively. We shall estimate $\|d_x\|$ and $\|d_s\|$. Our objective in this section is to demonstrate that $\|d_x\| = O(\mu)$ and $\|d_s\| = O(\mu)$.

We start by characterizing the solution to (6).

Lemma 3.1. If d_x and d_s are obtained from the linear system (6), then

$$\begin{aligned}
d_x &= -D\Pi_{N(AD)} r, \\
d_s &= -D^{-1}\Pi_{R(DA^T)} r,
\end{aligned}$$

where $r = Z^{1/2}e$.

Proof. The proof is straightforward, e.g., see Adler and Monteiro [1]. ∎

It is well known that for every linear program, a unique partition $A = (B,\ N)$ exists such that the primal optimal face is given by

$$\Omega_p = \{x = (x_B, x_N):\ Bx_B = b,\ x_B \geq 0,\ x_N = 0\}$$

and the dual optimal face is given by

$$\Omega_d = \{(y, s) = (y, (s_B, s_N)):\ s_B = c_B - B^T y = 0,\ s_N = c_N - N^T y \geq 0\}.$$

Strictly feasible solutions $x_B > 0$ and $s_N > 0$ exist on these optimal faces, respectively, and both faces are bounded under assumption A0. Here, we also use B and N to denote the partitioned column index sets. For all k, relation (3.1) implies that

$$\xi \leq x_j^k \leq 1/\xi \quad \text{for} \quad j \in B \tag{8.1}$$

and

$$\xi \leq s_j^k \leq 1/\xi \quad \text{for} \quad j \in N, \tag{8.2}$$

where $\xi < 1$ is a fixed positive number that is independent of k (Güler and Ye [3]).

Lemma 3.2. If d_x and d_s are obtained from the linear system (6) and $\mu = x^T s/n$, then

$$\|(d_x)_N\| = O(\mu) \quad \text{and} \quad \|(d_s)_B\| = O(\mu).$$

Proof. From Lemma 3.1, we obtain

$$\begin{aligned} \|D^{-1}d_x\| &\leq \|\Pi_{N(AD)}\|\|r\| \\ &\leq \|r\| = O(\sqrt{\mu}). \end{aligned}$$

We have from relations (7) and (8)

$$\begin{aligned} \|(d_x)_N\| &= \|D_N D_N^{-1}(d_x)_N\| \\ &\leq \|D_N\|\|D_N^{-1}(d_x)_N\| \\ &\leq \|D_N\| O(\sqrt{\mu}) \\ &= \|Z_N^{1/2} S_N^{-1}\| O(\sqrt{\mu}) \\ &= O(\sqrt{\mu}) O(\sqrt{\mu}) = O(\mu). \end{aligned}$$

This proves that $\|(d_x)_N\| = O(\mu)$. The proof that $\|(d_s)_B\| = O(\mu)$ is similar. ∎

The proofs of $\|(d_x)_B\| = O(\mu)$ and $\|(d_s)_N\| = O(\mu)$ are more involved. Towards this end, we first note

$$\begin{aligned} x + d_x &= D\Pi_{R(DA^T)}r, \\ s + d_s &= D^{-1}\Pi_{N(AD)}r. \end{aligned} \tag{9}$$

This is because from the first equation of (6) we have

$$\begin{aligned} S(x + d_x) &= -Xd_s \\ X(s + d_s) &= -Sd_x. \end{aligned}$$

Thus,

$$\begin{aligned} x + d_x &= -(XS^{-1})d_s = -D^2 d_s \\ s + d_s &= -(SX^{-1})d_x = -D^{-2}d_x, \end{aligned}$$

which gives relation (9).

The following lemma is essentially due to Adler and Monteiro [1] (also see Sonnevend et al. [12] and Witzgall et al. [14]).

Lemma 3.3. If d_x and d_s are obtained from the linear system (6), then $(d_x)_B$ is the solution to the (weighted) least-squares problem

$$\begin{aligned} \min_u \quad & (1/2)\|D_B^{-1}u\|^2 \\ s.t. \quad & Bu \;=\; -N(d_x)_N, \end{aligned} \tag{10.1}$$

and $(d_s)_N = -N^T d_y$ where d_y is a solution to the (weighted) least-squares problem

$$\begin{aligned} \min_v \quad & (1/2)\|D_N N^T v\|^2 \\ s.t. \quad & B^T v \;=\; -(d_s)_B. \end{aligned} \tag{10.2}$$

Proof. From (9), we see that

$$x_B + (d_x)_B \in R(D_B^2 B^T). \tag{11}$$

Since $s_B^* = 0$ for all optimal s^*, we must have $c_B \in R(B^T)$. Thus,

$$s_B = c_B - B^T y \in R(B^T),$$

which implies that

$$x_B = D_B^2 s_B \in R(D_B^2 B^T). \tag{12}$$

¿From (11) and (12) we have

$$(d_x)_B \in R(D_B^2 B^T).$$

Moreover, $(d_x)_B$ satisfies the equation

$$B(d_x)_B = -N(d_x)_N.$$

Thus, $(d_x)_B$ satisfies the Karush-Kuhn-Tucker conditions for the least squares problem (10.1).

Since $AD^2(s + d_s) = -Ad_x = 0$ and $AD^2 s = Ax = b$, it follows that

$$-b = AD^2 d_s = BD_B^2(d_s)_B + ND_N^2(d_s)_N. \tag{13}$$

Also, since $x_N^* = 0$ for all optimal x^*, we have $Bx_B^* = b$ implying $b \in R(B)$. Therefore, relation (13) implies

$$ND_N^2 N^T d_y = -ND_N^2(d_s)_N \ \in \ R(B).$$

Moreover, d_y satisfies the equation

$$B^T d_y = -(d_s)_B.$$

Thus, d_y satisfies the Karush-Kuhn-Tucker conditions for the least squares problem (10.2). ∎

Theorem 3.1. If d_x and d_s are obtained from the linear system (6) and $\mu = x^T s/n$, then

$$\|(d_x)_B\| = O(\mu) \quad \text{and} \quad \|(d_s)_N\| = O(\mu).$$

Proof. Since the least-squares problem (10.1) is always feasible, there must be a *feasible* $\bar{u}$ such that

$$\|\bar{u}\| = O(\|(d_x)_N\|),$$

which together with Lemma 3.2 implies

$$\|\bar{u}\| = O(\mu).$$

Furthermore, from Lemma 3.3 and relations (7) and (8)

$$\begin{aligned}
\|(d_x)_B\| &= \|D_B D_B^{-1}(d_x)_B\| \\
&\le \|D_B\|\|D_B^{-1}(d_x)_B\| \\
&\le \|D_B\|\|D_B^{-1}\bar{u}\| \\
&\le \|D_B\|\|D_B^{-1}\|\|\bar{u}\| \\
&= \|Z_B^{-1/2}X_B\|\|Z_B^{1/2}X_B^{-1}\|\|\bar{u}\| \\
&\le \|Z_B^{-1/2}\|\|X_B\|\|Z_B^{1/2}\|\|X_B^{-1}\|\|\bar{u}\| \\
&= O(\|\bar{u}\|) = O(\mu).
\end{aligned}$$

Similarly, we can prove the second statement of the theorem. ∎

4. Quadratic Convergence without Assumption A1 or A2

Lemma 3.2 and Theorem 3.1 indicate that at the kth predictor step, d_x^k and d_s^k satisfy

$$\|d_x^k\| = O(\mu^k) \quad \text{and} \quad \|d_s^k\| = O(\mu^k), \tag{14}$$

where $\mu^k = (x^k)^T s^k / n$, for all $k \ge 0$. We are now in a position to state our main result.

Theorem 4.1. Let $\{(x^k, s^k)\}$ be the sequence generated by the large step predictor-corrector algorithm. Then, with constants $0 < \beta \le 1/4$ and $\alpha = 2\beta$,

(i) the algorithm has iteration complexity $O(\sqrt{n}L)$;

(ii) $1 - \theta^k = O((x^k)^T s^k)$; and

(iii) $(x^k)^T s^k \to 0$ Q-quadratically.

Proof. The proof of (i), i.e., the $O(\sqrt{n}L)$-iteration complexity of the algorithm follows from inequalities (3.2), (4.1) and Lemma 2.2, which give

$$(x^{k+1})^T s^{k+1} \le \left(1 - \frac{2}{\sqrt{1 + \sqrt{2}n/\beta} + 1}\right)(x^k)^T s^k.$$

This also establishes

$$\lim_{k\to\infty} (x^k)^T s^k = 0.$$

¿From relation (14) we have

$$\|\delta^k\| = \|D_x^k d_s^k/\mu^k\| \le \|d_x^k\|\|d_s^k\|/\mu^k = O((\mu^k)^2)/\mu^k = O(\mu^k) = O((x^k)^T s^k),$$

which together with (5) establishes (ii). ¿From (3.2) we see that (ii) implies (iii). This proves the theorem. ∎

The asymptotic behavior of $(X^k)^{-1}d_x^k$ and $(S^k)^{-1}d_s^k$ at the predictor step seems interesting. Using relations (8) and (14), we have

$$|(d_x^k)_j/x_j^k| = O(\mu^k) \quad \text{for} \quad j \in B \tag{15.1}$$

and

$$|(d_s^k)_j/s_j^k| = O(\mu^k) \quad \text{for} \quad j \in N. \tag{15.2}$$

Moreover, at the predictor step (i.e., $\gamma = 0$), it follows from system (1) that

$$\frac{(d_x^k)_j}{x_j^k} + \frac{(d_s^k)_j}{s_j^k} = -1 \tag{16}$$

for every $j = 1, 2, ..., n$. Thus, relations (15) and (16) imply that

$$\lim_{k\to\infty} \frac{(d_x^k)_j}{x_j^k} = 0 \quad \text{and} \quad \lim_{k\to\infty} \frac{(d_s^k)_j}{s_j^k} = -1 \quad \text{for} \quad j \in B$$

and

$$\lim_{k\to\infty} \frac{(d_x^k)_j}{x_j^k} = -1 \quad \text{and} \quad \lim_{k\to\infty} \frac{(d_s^k)_j}{s_j^k} = 0 \quad \text{for} \quad j \in N.$$

Interestingly, the behavior of $(X^k)^{-1}d_x^k$ and $(S^k)^{-1}d_s^k$ was discussed by Tapia [13] in (1980) for complementarity problems where the strict complementarity solution is unique. He used these two vectors as the basis of an indicator theory for identifying variables which are zero at the solution. See also El-Bakry et al. [2]. These two vectors were also used by Mehrotra and Ye [10] as a criterion for identifying the partition B and N.

The following corollary formally states that we do not need to choose the largest step in the predictor step, but only a sufficiently large step. Thus, we are no longer required to find the zeros of a quartic polynomial.

Corollary 4.1. If the predictor-corrector algorithm is adapted with θ^k given by the lower bound in Lemma 2.2, then Theorem 4.1 also holds for the modified algorithm.

5. Summary and Concluding Remarks

Recently, Mizuno et al. [11] proposed a large step predictor-corrector interior-point algorithm for linear programming. They demonstrated that the algorithm possessed $O(\sqrt{n}L)$ iteration complexity. More recently, Ye et al. [16] proved that the algorithm, while maintaining $O(\sqrt{n}L)$ iteration complexity, exhibited superlinear convergence of the duality gap sequence to zero under the assumption that the iteration sequence converged, and exhibited quadratic convergence of the duality gap sequence to zero under the assumption of nondegeneracy. In this paper we have established the surprising result that the large-step predictor-corrector algorithm actually exhibits quadratic convergence of the duality gap to zero without the assumption of nondegeneracy or even the assumption that the iteration sequence converges. This result is the first instance of a demonstration of polynomiality and superlinear (or quadratic) convergence for an interior-point method which does not assume the convergence of the iteration sequence or nondegeneracy. We note that each iteration in the predictor-corrector algorithm requires the solutions of two linear systems—one in the predictor step and one in the corrector step.

Although the iteration sequence $\{(x^k, s^k)\}$ may not be convergent, it is a consequence of Hoffman's lemma [4] that the sequence $\{x^k\}$ converges R-quadratically to the primal optimal set Ω_p. The same is true for the sequence $\{s^k\}$ if we write the dual linear program and its optimal solution set in terms of s alone.

While seemingly quadratic convergence has often been observed in practice for primal-dual interior-point methods applied to degenerate problems, its effectiveness is compromised by the use of finite-precision arithmetic to solve the necessarily ill-conditioned linear systems. Hence our current result may have only theoretical value.

In this context the finite termination procedures of Ye [15] and Mehrotra and Ye [10] are of value in obtaining an optimal solution.

It has been observed in practice that the $O(\sqrt{n}L)$ algorithms that have been tested so far are generally less effective than are some of the $O(nL)$ algorithms (or other non-polynomial algorithms). Zhang, Tapia and Dennis [18] argued that several of these $O(\sqrt{n}L)$ algorithms possess particularly poor Q-convergence properties, i.e., they exhibit Q-linear convergence with convergence constants near 1 for large n. Therefore, some researchers may have embraced the belief that the $O(\sqrt{n}L)$ algorithms were less effective because none of them could achieve superlinear convergence. Now, we have demonstrated that a particular $O(\sqrt{n}L)$ algorithm actually has what we consider to be the optimal convergence rate for degenerate or nondegenerate problems. If, perchance, numerical experimentation still favors the $O(nL)$ algorithms, then we must conclude that their advantage is not due to their asymptotic behavior, but to some other, as yet unexplained, phenomenon.

References

[1] I. Adler and R. Monteiro, "Limiting behavior of the affine scaling continuous trajectories for linear programming problems," *Contemporary Mathematics* **114** (1990) 189-211.

[2] A. S. El-Bakry, R. A. Tapia and Y. Zhang, "A study of indicators for identifying zero variables in interior-point methods," TR91-15, Department of Mathematical Sciences, Rice University (Houston, TX, 1991).

[3] O. Güler and Y. Ye, "Convergence behavior of some interior-point algorithms," Working Paper 91-4, The College of Business Administration, The University of Iowa (Iowa City, IA, 1991), to appear in *Mathematical Programming.*

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

[5] J. Ji, F. Potra, R. A. Tapia and Y. Zhang, "An interior-point method with polynomial complexity and superlinear convergence for linear complementarity problems," Department of Mathematics, The University of Iowa (Iowa City, IA, 1991). Also TR91-23, Department of Mathematical Sciences, Rice University (Houston, TX, 1991).

[6] M. Kojima, Y. Kurita and S. Mizuno, "Large step interior-point algorithms for linear complementarity problems," Research Report B-243, Department of Information Sciences, Tokyo Institute of Technology (Tokyo, Japan, 1991), to appear in *SIAM J. on Optimization.*

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

[8] M. Kojima, S. Mizuno and A. Yoshise, "A polynomial-time algorithm for a class of linear complementarity problems," *Mathematical Programming* **44** (1989) 1-26.

[9] K. McShane, "A superlinearly convergent $O(\sqrt{n}L)$ iteration primal-dual linear programming algorithm," manuscript, 2537 Villanova Drive (Vienna, Virginia, 1991).

[10] S. Mehrotra and Y. Ye, "On finding the optimal face of linear programs," Technical Report, Department of IE and MS, Northwestern University (Evanston, IL, 1991).

[11] S. Mizuno, M. J. Todd, and Y. Ye, "On adaptive-step primal-dual interior-point algorithms for linear programming," Technical Report No. 944, School of ORIE, Cornell University (Ithaca, NY, 1990), to appear in *Mathematics of Operations Research.*

[12] G. Sonnevend, J. Stoer and G. Zhao, "On the complexity of following the central path of linear programs by linear extrapolation," *Methods of Operations Research* **63** (1989) 19-31.

[13] R. A. Tapia, "On the role of slack variables in quasi-Newton methods for constrained optimization," in L. C. W. Dixon and G. P. Szegö, eds., *Numerical Optimization of Dynamic Systems* (North-Holland, 1980) pp. 235-246.

[14] C. Witzgall, P. T. Boggs and P. D. Domich, "On the convergence behavior of trajectories for linear programming," *Contemporary Mathematics* **114** (1990) 161-187.

[15] Y. Ye, "On the finite convergence of interior-point algorithms for linear programming," Working Paper 91-5, The College of Business Administration, The University of Iowa (Iowa City, IA, 1991).

[16] Y. Ye, R. A. Tapia and Y. Zhang "A superlinearly convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming," TR91-22, Department of Mathematical Sciences, Rice University (Houston, TX, 1991).

[17] Y. Zhang and R. A. Tapia, "A quadratically convergent polynomial primal-dual interior-point algorithm for linear programming," TR90-40, Department of Mathematical Sciences, Rice University (Houston, TX, 1990), to appear in *SIAM J. on Optimization.*

[18] Y. Zhang, R. A. Tapia and J. E. Dennis, "On the superlinear and quadratic convergence of primal-dual interior-point linear programming algorithms," TR90-6, Department of Mathematical Sciences, Rice University (Houston, TX, 1990), to appear in *SIAM J. on Optimization.*

[19] Y. Zhang, R. A. Tapia and F. Potra, "On the superlinear convergence of interior-point algorithms for a general class of problems," TR90-7, Department of Mathematical Sciences, Rice University (Houston, TX, 1990), to appear in *SIAM J. on Optimization.*