Superlinear and Quadratic Convergence of Primal-Dual Interior-Point Methods for Linear Programming Revisited

Y. Zhang
R.A. Tapia

September, 1991

TR91-27

Superlinear and Quadratic Convergence of Primal-Dual Interior-Point Methods for Linear Programming Revisited *

Yin Zhang †and Richard A. Tapia ‡

September, 1991

Keywords: Linear programming, Primal-dual interior-point algorithms, Superlinear and quadratic convergence.

Abbreviated Title: Superlinear and Quadratic Convergence Revisited

*This research was supported in part by NSF Coop. Agr. No. CCR-8809615 and was initiated in June, 1991 while the first author was at Rice University as a visiting member of the Center for Research in Parallel Computation.

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

‡Department of Mathematical Sciences and the Center for Research in Parallel Computation, Rice University, Houston, Texas 77251-1892. This author was supported in part by AFOSR 89-0363, DOE DEFG05-86ER25017 and ARO 9DAAL03-90-G-0093.

Abstract

Recently, Zhang, Tapia and Dennis [19] produced a superlinear and quadratic convergence theory for the duality gap sequence in primal-dual interior-point methods for linear programming. In this theory, a basic assumption for superlinear convergence is the convergence of the iteration sequence; and a basic assumption for quadratic convergence is nondegeneracy. Several recent research projects have either used or built on this theory under one or both of the above mentioned assumptions. In this paper, we remove both assumptions from the Zhang-Tapia-Dennis theory.

1 Introduction

We consider linear programs in the standard form:

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b, \\ & x \geq 0, \end{array} \tag{1.1}$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n} (m < n)$ and A has full rank m.

It is known that the optimality conditions for (1.1) can be written as a $2n \times 2n$ nonlinear system with non-negative variables; namely,

$$F(x,y) = \begin{pmatrix} Ax - b \\ By - Bc \\ XYe \end{pmatrix} = 0, \quad (x,y) \geq 0, \tag{1.2}$$

where $B \in \mathbf{R}^{(n-m) \times n}$ is any matrix such that the columns of B^T form a basis for the null space of A, $X = diag(x)$, $Y = diag(y)$ and e is the n-vector of all ones.

The feasibility set of problem (1.2) is defined as

$$\mathcal{F} = \{(x,y) : x, y \in \mathbf{R}^n, Ax = b, By = Bc, (x,y) \geq 0\}.$$

A feasible pair $(x,y) \in \mathcal{F}$ is said to be strictly feasible if it is positive. In this work we assume that the relative interior of $\mathcal{F}$ is nonempty, i.e., strictly feasible points exist. We

denote the solution set of problem (1.2) by

$$\mathcal{S} = \{(x^*, y^*) : F(x^*, y^*) = 0, (x^*, y^*) \geq 0\}.$$

The primal-dual interior-point algorithms considered in this research can be motivated in several ways, e.g., path-following or potential reduction, but in essence they are all variants of Newton's method. The following generic algorithmic framework includes a majority of these existing primal-dual interior-point algorithms.

Algorithm 1 (Generic Primal-Dual Algorithm)

Given a strictly feasible pair (x^0, y^0). *For* $k = 0, 1, 2, \ldots,$ *do*

(1) *Choose* $\sigma^k \in [0, 1)$ *and set* $\mu^k = \sigma^k \frac{x^{k^T} y^k}{n}$.

(2) *Solve the following system for* $(\Delta x^k, \Delta y^k)$:

$$F'(x^k, y^k) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -F(x^k, y^k) + \mu^k \begin{pmatrix} 0 \\ e \end{pmatrix}.$$

(3) *Choose a step-length* $\alpha^k = \min(1, \tau^k \hat{\alpha}^k)$ *for* $\tau^k \in (0, 1)$ *and*

$$\hat{\alpha}^k = \frac{-1}{\min((X^k)^{-1} \Delta x^k, (Y^k)^{-1} \Delta y^k)}. \tag{1.3}$$

(4) *Form the new iterate*

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

This algorithmic framework was first suggested by Megiddo [9]. There are two fundamental parameters in Algorithm 1, the centering parameter σ^k and the step-length τ^k (or α^k). The choice of step-length α^k guarantees $(x^{k+1}, y^{k+1}) > 0$. It is easy to verify that the iterates satisfy the following useful relationships

$$x^{k+1^T} y^{k+1} = (1 - \alpha^k (1 - \sigma^k)) x^{k^T} y^k \tag{1.4}$$

and

$$(X^k)^{-1} \Delta x^k + (Y^k)^{-1} \Delta y^k = -e + \sigma_k \frac{x^{k^T} y^k}{n} (X^k Y^k)^{-1} e. \tag{1.5}$$

Practical implementations and numerical experiments performed by MacShane, Monma and Shanno [8] and by Lustig, Marsten and Shanno [6], among others, demonstrated that in general the primal-dual approach was computationally superior to both the primal and the dual approach.

The theoretical properties of algorithms in the framework of Algorithm 1 have been extensively analyzed. Most existing theoretical results for Algorithm 1 can be classified into two basic categories: worst-case complexity analysis and asymptotic convergence rate analysis. Early works focusing on the former include Kojima, Mizuno and Yoshise [5], Monteiro and Adler [12], and Todd and Ye [14].

The analysis of superlinear and quadratic convergence properties for Algorithm 1 started with Zhang, Tapia and Dennis [19] and has recently become an active research direction. For linear programming, works in this direction include Zhang and Tapia [18], Ye, Tapia and Zhang [17], McShane [7], and Ye, Güler, Tapia and Zhang [16]. For linear complementarity problems, works in this direction include Zhang, Tapia and Potra [20], Kojima, Kurita and Mizuno [4] and Ji, Potra, Tapia and Zhang [3].

In the context of the present work, it is important to observe that in establishing superlinear convergence of the duality gap (or complementarity) to zero, all of the above authors, except Ye *et al* [16], assumed the convergence of the iteration sequence for superlinear convergence and nondegeneracy for quadratic convergence. Ye *et al* [16] studied the Mizuno-Todd-Ye predictor-corrector algorithm, which takes $\sigma^k = 1$ and $\sigma^k = 0$ alternatively, and obtained the impressive result of quadratic convergence of the duality gap to zero without the assumption of nondegeneracy or the assumption of the convergence of the iteration sequence. This quadratic rate is based on counting the predictor and the corrector steps as a single step; otherwise, the rate becomes two-step quadratic. At this juncture, a question naturally arises: whether a convergence theory can be developed which guarantees superlinear convergence of Algorithm 1 without the assumption of the convergence of the iteration sequence, and quadratic convergence without the assumption of nondegeneracy. In this paper, we answer this question affirmatively by removing from the Zhang-Tapia-Dennis

convergence theory [19] the assumptions of the convergence of the iteration sequence and nondegeneracy. A question that still remains open is the existence of an algorithm in the form of Algorithm 1 which is one-step quadratically convergent for degenerate problems.

This paper is organized as follows. In Section 2, we introduce some known results that will be used in the development of our main result. In Section 3, we establish Theorem 3.2 which gives rather mild sufficient conditions for the superlinear and the quadratic convergence of Algorithm 1. In Section 4, we partially validate the assumptions of Theorem 3.2 by constructing simple choices of the parameters σ^k and α^k for all cases of the theorem except the case leading to one-step quadratic convergence. We make some concluding remarks in the final section.

2 Preliminaries

In this section, we introduce our notation and recall some known results that will be used in the development of our main results.

For brevity, we use the notation $z = (x, y) \in \mathbf{R}^{2n}$, i.e.,

$$z_i = \begin{cases} x_i, & 1 \le i \le n, \\ y_i, & n < i \le 2n. \end{cases} \tag{2.1}$$

The theorem below is Theorem 3.2 in [13]. It gives a bound on the steps generated by Algorithm 1.

Theorem 2.1 (Tapia-Zhang-Ye) *Let $\{z^k = (x^k, y^k)\}$ and $\{\Delta z^k\}$ be generated by Algorithm 1. Assume*

A1 $x^{kT} y^k \to 0$.

A2 $\min(X^k Y^k e)/x^{kT} y^k \ge \gamma/n$ *for all k and some $\gamma \in (0, 1)$.*

Then there exist constants $\beta' > 0$ and $\beta'' > 0$ such that

$$\|\Delta z^k\| \le \beta' x^{kT} y^k + \beta'' \sigma^k. \tag{2.2}$$

Lemma 2.1 below is Lemma 2 of Güler and Ye [2], tailored to fit the needs of this paper. Many interesting properties of interior-point algorithms follow from this simple fact.

Lemma 2.1 (Güler-Ye)

Let $\{z^k\} = \{(x^k, y^k)\}$ *be generated by Algorithm 1. Assume*

A1 ${x^k}^T y^k$ *converges to zero.*

A2 $\min(X^k Y^k e)/{x^k}^T y^k \geq \gamma/n$ *for all* k *and some* $\gamma \in (0,1)$.

Let z^* *be a limit point of* $\{z^k\}$. *Then*

(i) z^* *is a strictly complementary solution of Problem 1.2 and*

(ii) $\lim_{k\to\infty} \inf z_i^k > 0$ *for every* i *such that* $z_i^* > 0$.

El-Bakry, Tapia and Zhang [1] recently demonstrated that the relative interior of $\mathcal{S}$ and the set of solutions satisfying strict complementarity coincide. Furthermore, the zero-nonzero structure of solutions in the relative interior of $\mathcal{S}$ (equivalently solutions satisfying strict complementarity) is invariant.

For any $z^* = (x^*, y^*)$ in the relative interior of $\mathcal{S}$, define

$$I_x^+ = \{i : x_i^* > 0, 1 \leq i \leq n\} \quad \text{and} \quad I_y^+ = \{i : y_i^* > 0, 1 \leq i \leq n\}. \tag{2.3}$$

Since the zero-nonzero structure of the relative interior of $\mathcal{S}$ is invariant, the above two index sets are independent of the choice of z^*. By strict complementarity of z^*,

$$I_x^+ \bigcup I_y^+ = \{1, 2, ..., n\} \quad \text{and} \quad I_x^+ \bigcap I_y^+ = \emptyset. \tag{2.4}$$

3 Superlinear and Quadratic Convergence Revisited

In [19], Zhang, Tapia and Dennis established sufficient conditions for the superlinear and quadratic convergence of the duality gap sequence generated by Algorithm 1. Their results can be summarized as the following theorem.

Theorem 3.1 (Zhang-Tapia-Dennis)

Let $\{(x^k, y^k)\}$ *be generated by Algorithm 1. Assume*

A1 $\{(x^k, y^k)\}$ *converges to a solution* (x^*, y^*).

A2 $\min(X^k Y^k e)/x^{k^T} y^k \geq \gamma/n$ *for all* k *and some* $\gamma \in (0, 1)$.

A3 $\sigma^k \to 0$ *and* $\tau^k \to 1$.

Then the duality gap sequence $\{x^{k^T} y^k\}$ *converges to zero Q-superlinearly, i.e.,*

$$\frac{x^{k+1^T} y^{k+1}}{x^{k^T} y^k} \to 0.$$

Assume further

A4 $\sigma^k = O(x^{k^T} y^k)$ *and* $\tau^k = 1 - O(x^{k^T} y^k)$.

A5 x^* *is a nondegenerate vertex of (1.1).*

Then $\{(x^k, y^k)\}$ *converges to* (x^*, y^*) *Q-quadratically.*

Originally, Zhang, Tapia and Dennis also assumed strict complementarity for (x^*, y^*). Thanks to Lemma 2.1 that assumption has been effectively eliminated. Observe that under primal nondegeneracy and strict complementarity, the convergence of the duality gap sequence to zero implies the convergence of the iteration sequence because of the uniqueness of the solution.

Based on Theorem 3.1, Zhang and Tapia [18] constructed the first polynomial and superlinearly convergent primal-dual algorithm for linear programming. Its superlinear convergence was established under the assumption of the convergence of the iteration sequence. Furthermore, under the assumption of nondegeneracy, they also demonstrated quadratic convergence for the iteration sequence. More recently, Ye *et al* [16] gave the first proof of quadratic converegnce for a primal-dual algorithm for linear programming that did not assume the convergence of the iteration sequence or nondegeneracy. They demonstrated quadratic convergence of the duality gap to zero in the Mizuno-Todd-Ye predictor-corrector

algorithm [11]. As mentioned in Section 1, this rate is based on counting the predictor and the corrector steps as a single step.

Theorem 3.2 below improves Theorem 3.1 by replacing the assumption $(x^k, y^k) \to (x^*, y^*)$ with ${x^k}^T y^k \to 0$, and removing the nondegeneracy assumption.

Theorem 3.2 *Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1. Assume*

A1 ${x^k}^T y^k \to 0$;

A2 $\min(X^k Y^k e)/{x^k}^T y^k \geq \gamma/n$ *for all k and some $\gamma \in (0,1)$.*

A3 $\sigma^k \to 0$ *and* $\tau^k \to 1$.

(i) *Then the duality gap sequence $\{{x^k}^T y^k\}$ converges to zero Q-superlinearly, i.e.,*

$$\frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} \to 0.$$

Assume further

A4 $\sigma^k = O(({x^k}^T y^k)^\lambda)$ *and* $\tau^k = 1 - O(({x^k}^T y^k)^\lambda)$ *for some* $\lambda \in (0,1]$.

Then

(ii) *the duality gap sequence $\{{x^k}^T y^k\}$ converges to zero with Q-rate $1+\lambda$, i.e.,*

$$\limsup_{k\to\infty} \frac{{x^{k+1}}^T y^{k+1}}{({x^k}^T y^k)^{1+\lambda}} < \infty;$$

(iii) *the iteration sequence $\{(x^k, y^k)\}$ converges to a solution with R-rate $1+\lambda$.*

Proof: (i) From (1.4), Assumption A3 and the choice of $\alpha^k = \tau^k \hat{\alpha}^k$,

$$\limsup_{k\to\infty} \frac{{x^{k+1}}^T y^{k+1}}{{x^k}^T y^k} = \limsup_{k\to\infty}(1 - \tau^k \hat{\alpha}^k(1-\sigma^k)) = \limsup_{k\to\infty}(1-\hat{\alpha}^k).$$

Hence it suffices to prove $\hat{\alpha}^k \to 1$. Note that in Theorem 2.1 $\sigma^k \to 0$ implies $\|\Delta z^k\| \to 0$.

In view of (3.1) and (ii) of Lemma 2.1, we have that for any $i \in I_x^+$, $\Delta x_i^k \to 0$ imply $\Delta x_i^k / x_i^k \to 0$. Similarly, for $i \in I_y^+$ we have $\Delta y_i^k / y_i^k \to 0$.

It follows from (1.5), Assumptions A2 and A3 that

$$(X^k)^{-1}\Delta x^k + (Y^k)^{-1}\Delta y^k \to -e. \tag{3.1}$$

Therefore, for each $i \in \{1, 2, \ldots, n\}$ either

$$\frac{\Delta x_i^k}{x_i^k} \to 0, \quad \frac{\Delta y_i^k}{y_i^k} \to -1 \quad (i \in I_x^+) \tag{3.2}$$

or

$$\frac{\Delta x_i^k}{x_i^k} \to -1, \quad \frac{\Delta y_i^k}{y_i^k} \to 0 \quad (i \in I_x^+). \tag{3.3}$$

By the definition of $\hat{\alpha}^k$ (1.3), the relations (3.2) and (3.3) imply $\hat{\alpha}^k \to 1$.

(ii) We need to show

$$\frac{x^{k+1^T} y^{k+1}}{x^{k^T} y^k} = O((x^{k^T} y^k)^\lambda).$$

From (1.4) and Assumption A4

$$\frac{x^{k+1^T} y^{k+1}}{x^{k^T} y^k} = 1 - \tau^k \hat{\alpha}^k (1 - \sigma^k) = 1 - \hat{\alpha}^k + O((x^{k^T} y^k)^\lambda).$$

Hence it suffices to show $1 - \hat{\alpha}^k = O((x^{k^T} y^k)^\lambda)$. Furthermore, from (1.3), (3.2) and (3.3), $1 - \hat{\alpha}^k = O((x^{k^T} y^k)^\lambda)$ is equivalent to

$$1 + \frac{1}{\Delta z_i^k / z_i^k} = O((x^{k^T} y^k)^\lambda), \quad \text{for all } i \text{ such that } z_i^k \to 0.$$

Without loss of generality, consider $i \in I_x^+$ (i.e., $y_i^k \to 0$). It follows from (1.5) and Assumption A4 that

$$1 + \frac{1}{\Delta y_i^k / y_i^k} = 1 - \frac{1}{1 + \Delta x_i^k / x_i^k - O(\sigma^k)} = \frac{\Delta x_i^k / x_i^k - O((x^{k^T} y^k)^\lambda)}{1 + \Delta x_i^k / x_i^k - O((x^{k^T} y^k)^\lambda)}.$$

Therefore, it is sufficient to have

$$\Delta x_i^k = O((x^{k^T} y^k)^\lambda), \quad i \in I_x^+,$$
$$\Delta y_i^k = O((x^{k^T} y^k)^\lambda), \quad i \in I_y^+.$$

The above relations follow directly from Theorem 2.1 and Assumption A4.

(iii) The R-convergence rate of the iteration sequence follows essentially from (iv) of Theorem 4.1 of Tapia, Zhang and Ye [13]. This completes the proof. □

It is important to observe that the parameters σ^k and τ^k are under direct control of the algorithm designer. This is a point often missed by some readers.

4 Consistency Of Assumptions

In this section, we demonstrate the consistency of the assumptions made in Theorem 3.2 except for the case $\lambda = 1$.

While the current work was in progress, Ye [15] reported the following result. If the duality gap is small enough and if one takes pure Newton steps ($\sigma^k \equiv 0$) to the boundary of a gradually expanding neighborhood of the central path, then the duality gap sequence will converge to zero with Q-rate $1 + \lambda$ for $\lambda \in [\frac{1}{2}, 1)$. Here λ depends on the rate of the expansion. With additional effort, his result can be extended to include the case $\lambda \in (0, \frac{1}{2})$.

Although Ye's construction essentially can be used to demonstrate the consistency of the assumptions made in Theorem 3.2 except for the case $\lambda = 1$, for the sake of completeness we feel that it is expedient to include our own construction here. It is based on the same principle but is somewhat simpler and gives a slightly more general result.

Let

$$\sigma^k = \min(\sigma, \rho {x^k}^T y^k) \tag{4.1}$$

for some constants $\sigma \in (0, 1)$ and $\rho > 0$. For each k choose τ^k so that $\alpha^k = \tau^k \hat{\alpha}^k$ is the largest number in $(0, 1]$ such that for $k > 0$ and $\lambda \in (0, 1)$

$$\frac{\min(x_i^{k+1} y_i^{k+1})}{{x^{k+1}}^T y^{k+1}/n} \geq \gamma^k = \gamma^{k-1} - \theta^k ({x^k}^T y^k)^{1-\lambda} > 0, \tag{4.2}$$

where $\gamma^0 = 2\gamma$ for some constant $\gamma \in (0, 1/2)$ and $\theta^k \geq 0$ is chosen to ensure $\gamma^k > 0$.

We need the following two technical lemmas to prove our main result.

Lemma 4.1 *Let σ^k and α^k be chosen as in (4.1) and (4.2). If $\gamma^k \geq \gamma$, then there exists a constant $\eta \in (0, 1)$ such that*

$${x^{k+1}}^T y^{k+1} \leq \eta {x^k}^T y^k. \tag{4.3}$$

Proof: Since $1 - \sigma^k \geq 1 - \sigma > 0$, it follows from (1.4) that it is sufficient to establish a positive lower bound for α^k. It has been shown (see Lemma 3.4 of Zhang and Tapia [18]) that with the choice (4.2)

$$\alpha^k \geq \min\left(1, \frac{\sigma^k (1 - \gamma) {x^k}^T y^k / n}{\max |\Delta x_i^k \Delta y_i^k|}\right). \tag{4.4}$$

When $\sigma^k = \sigma$, it is known that α^k is bounded away from zero (see [18], for example). Otherwise, $\sigma^k = \rho {x^k}^T y^k$. By Theorem 2.1, $|\Delta x_i^k \Delta y_i^k| = O(({x^k}^T y^k)^2)$. Hence it follows from (4.4) that α^k is also bounded away from zero. □

Lemma 4.2 *Let $\{\sigma^k\}$ and $\{\alpha^k\}$ be chosen as in (4.1) and (4.2). There exist $\{\theta^k\}$ and a constant $\theta > 0$ such that $\theta^k \geq \theta$ for k sufficiently large and $\gamma^k \geq \gamma$.*

Proof: Let $\eta \in (0,1)$ be as in (4.3) and let

$$\theta^k = \theta \equiv \frac{\gamma(1-\eta^{1-\lambda})}{({x^0}^T y^0)^{1-\lambda}} > 0.$$

Since $\gamma^0 > \gamma$, the lemma holds for $k = 0$. Suppose that the lemma holds for $k = p-1$. Then (4.3) in Lemma 4.1 holds and

$$\gamma^k \geq \gamma^0 - \theta({x^0}^T y^0)^{1-\lambda} \sum_{k=0}^{p-1} \eta^{(1-\lambda)k} = 2\gamma - \gamma(1-\eta^{1-\lambda}) \sum_{k=0}^{p-1} \eta^{(1-\lambda)k} > \gamma.$$

So the lemma also holds for $k = p$. This completes the proof. □

The following theorem establishes, for general problems, the consistency of the assumptions made in Theorem 3.2 except for the case $\lambda = 1$ (i.e., except for quadratic convergence). Observe that the consistency of these assumptions for $\lambda = 1$ and for nondegenerate problems follows from Zhang and Tapia [18].

Theorem 4.1 *Let $\{(x^k, y^k)\}$ be generated by Algorithm 1 with the the choices (4.1) and (4.2), and let $\{\theta^k\}$ be such that Lemma 4.2 holds. Then for any $\lambda \in (0,1)$*

(i) *Assumptions A1-A4 of Theorem 3.2 are satisfied;*

(ii) *the duality gap sequence $\{{x^k}^T y^k\}$ converges to zero with Q-rate $1+\lambda$;*

(iii) *the iteration sequence $\{(x^k, y^k)\}$ converges to a solution with R-rate $1+\lambda$.*

Proof: (i) It suffices to demonstrate that Assumptions A1, A2 and A4 are simultaneously satisfied. Assumptions A1 and A2 follow from Lemmas 4.1 and 4.2. The first part of Assumptions A4 is enforced in (4.1). Hence, we only need to show $\tau^k = 1 - O(({x^k}^T y^k)^\lambda)$. From (1.4) it suffices to prove $\alpha^k = 1 - O(({x^k}^T y^k)^\lambda)$.

Let

$$f_i(\alpha) = (x_i^k + \alpha \Delta x_i^k)(y_i^k + \alpha \Delta y_i^k), \quad i = 1, 2, \ldots, n,$$

and

$$\bar{f}(\alpha) = \frac{1}{n} \sum_{i=1}^{n} f_i(\alpha).$$

Then condition (4.2) is equivalent to choosing α^k to be the largest number in $(0, 1]$ such that

$$h_i(\alpha) = f_i(\alpha) - \gamma^k \bar{f}(\alpha) \geq 0, \quad i = 1, 2, \ldots, n.$$

Denote $f_i = f_i(0)$, $i = 1, 2, \ldots, n$, and $\bar{f} = \bar{f}(0)$. It is clear that $f_i - \gamma^{k-1} \bar{f} \geq 0$. A straightforward calculation, together with Lemma 4.2, gives for k sufficiently large

$$\begin{aligned}
h_i(\alpha) &= \Delta x_i^k \Delta y_i^k \alpha^2 + \sigma^k (1 - \gamma^k) \bar{f} \alpha + (f_i - \gamma^k \bar{f})(1 - \alpha) \\
&= \Delta x_i^k \Delta y_i^k \alpha^2 + \sigma^k \theta^k (x^{k^T} y^k)^{1-\lambda} \bar{f} \alpha + \theta^k (x^{k^T} y^k)^{1-\lambda} \bar{f}(1 - \alpha) \\
&\quad + \sigma^k (1 - \gamma^{k-1}) \bar{f} \alpha + (f_i - \gamma^{k-1} \bar{f})(1 - \alpha) \\
&\geq \Delta x_i^k \Delta y_i^k \alpha^2 + \sigma^k \theta (x^{k^T} y^k)^{1-\lambda} \bar{f} \alpha + \theta (x^{k^T} y^k)^{1-\lambda} \bar{f}(1 - \alpha) \\
&= \bar{h}_i(\alpha).
\end{aligned}$$

Hence it suffices to show that all positive roots of $\bar{h}_i(\alpha)$, $i = 1, 2, \ldots, n$, are of the order $1 - O((x^{k^T} y^k)^\lambda)$. We only need to consider the case $\Delta x_i^k \Delta y_i^k < 0$; otherwise, the roots will be greater than one. Let $\xi = \theta (x^{k^T} y^k)^{1-\lambda} \bar{f}$, then

$$\bar{h}_i(\alpha) = \Delta x_i^k \Delta y_i^k \alpha^2 - (1 - \sigma^k) \xi \alpha + \xi.$$

The positive root of $\bar{h}_i(\alpha)$, when $\Delta x_i^k \Delta y_i^k < 0$, is

$$\bar{\alpha}_i = \frac{2}{\sqrt{(1 - \sigma^k)^2 - 4 \Delta x_i^k \Delta y_i^k / \xi} + 1 - \sigma^k}.$$

Note that $\sigma^k = O(x^{k^T} y^k)$ by (4.1), $|\Delta x_i^k \Delta y_i^k| = O((x^{k^T} y^k)^2)$ by Theorem 2.1 and $\xi = O((x^{k^T} y^k)^{2-\lambda})$. Therefore, $-4 \Delta x_i^k \Delta y_i^k / \xi = O((x^{k^T} y^k)^\lambda)$. Clearly, $\bar{\alpha}_i = 1 - O((x^{k^T} y^k)^\lambda)$ for all i with $\Delta x_i^k \Delta y_i^k < 0$. This implies $\alpha^k = 1 - O((x^{k^T} y^k)^\lambda)$ and proves (i).

The results in (ii) and (iii) follow directly from Theorem 3.2. □

5 Concluding Remarks

It is known that for degenerate problems the Jacobian matrix $F'(x, y)$ is necessarily singular at any solution (see [19], for example). Interestingly, in practice it had been observed that even for highly degenerate problems when the parameters were chosen as in Assumption A4, then primal-dual interior-point algorithms in the generic class of Algorithm 1 exhibited effectively quadratic convergence until numerical singularity was encountered. Fortunately, for most problems we tested, numerical singularity occurred at points extremely close to a solution; though there were exceptions. We believe that Theorem 3.2, together with Theorem 4.1, provides a satisfactory theoretical explanation for the above unexpected but pleasing phenomenon. Still, for degenerate problems, the existence of a one-step quadratically convergent algorithm remains an open question.

The theoretical and practical values of superlinear and quadratic convergence are well-accepted in the continuous numerical optimization community. In comparison to an average linearly convergent algorithm, a superlinearly or quadratically convergent algorithm generally exhibits significant advantages at two stages. At an intermediate stage the transition to necessarily fast local linear convergence accelerates the convergence process. Then at the final stage the constantly improving fast local convergence produces a highly accurate approximate solution very quickly. It should be clear that the onset of these stages is problem-dependent and can vary considerably from one problem to another.

Superlinear and quadratic convergence are relatively new concepts in linear programming. The area has been dominated by the simplex method for decades and since the simplex method has finite termination these notions, defined for infinite iterative procedures, do not apply. Recently, in trying to take advantage of the combinatorial information inherent in the linear program, finite termination techniques have been devised for interior-point methods. A procedure recently proposed by Mehrotra and Ye [10] has been shown to possess a solid theoretical foundation and good numerical performance. The existence of such finite termination procedures undoubtedly diminishes the significance of the final-stage advantage of superlinear and quadratic convergence. However, since at this juncture no finite termination

procedure has been shown to have the ability to terminate interior-point algorithms successfully at early stages, it is still safe to say that the final-stage advantage of superlinear and quadratic convergence still exists to some extent in linear programming. On the other hand, it is not at all clear if finite termination techniques have any effect on the intermediate-stage advantage of superlinear and quadratic convergence. These topics certainly merit further study.

Acknowledgment

The authors thank Yinyu Ye for constructive comments and discussions concerning this material.

References

[1] Amr El-Bakry, R. A. Tapia, and Yin Zhang. On the use of indicators in identifying zero variables for interior point methods. Technical Report TR91-15, Dept. Mathematical Sciences, Rice University, 1991.

[2] O. Güler and Y. Ye. Convergence behavior of some interior-point algorithms. Working Paper Series No. 91-4, Dept. of Management Sciences, The University of Iowa, 1991.

[3] J. Ji, F. Potra, R. A. Tapia, and Y. Zhang. An interior-point method for linear complementarity problems with polynomial complexity and superlinear convergence. Technical Report TR91-23, Dept. Mathematical Sciences, Rice University, 1991.

[4] M. Kojima, Y. Kurita and S. Mizuno. Large-step interior point algorithms for linear complementarity problems. Research Report on Information Sciences B-243, Dept. of Information Sciences, Tokyo Institute of Technology, 1991.

[5] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point method for linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 29–47. Springer-Verlag, New York, 1989.

[6] I.J. Lustig, R.E. Marsten, and D.F. Shanno. Computational experience with a primal-dual interior point method for linear programming. Technical Report SOR 89-17, Dept. Civil Eng. and O.R., Princeton University, 1989.

[7] K. McShane. A superlinearly convergent $O(\sqrt{n}L)$-iteration primal-dual linear programming algorithm. Manuscript, 2534 Villanova Drive, Vienna, Virginia, 1991.

[8] K.A. McShane, C.L. Monma, and D.F. Shanno. An implementation of a primal-dual interior point method for linear programming. *ORSA J. Computing*, 1:70–83, 1989.

[9] N. Megiddo. Pathways to the optimal set in linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 131–158. Springer-Verlag, New York, 1989.

[10] S. Mehrotra and Y. Ye. On finding the optimal facet 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. Tech. Report No. 944, School of ORIE, Cornell University, 1990, to appear in *Math. of O.R.*.

[12] R.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part I: linear programming. *Math. Prog.*, 44:27–41, 1989.

[13] R. A. Tapia, Y. Zhang and Y. Ye On the convergence of the iteration sequence in primal-dual interior-point methods. Technical Report No. 91-24, Dept. of Mathematical Sciences, Rice University, 1991.

[14] M.J. Todd and Y. Ye. A centered projective algorithm for linear programming. *Math. of O.R.*, 15:508–529,1990.

[15] Y. Ye. Improving the asymptotic convergence of interior-point algorithms for linear programming. Working paper Series No. 91-15, Dept. of Management Sciences, The University of Iowa, Sept., 1991.

[16] Y. Ye, O. Güler, R. A. Tapia and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Technical Report TR91-26, Dept. Mathematical Sciences, Rice University, 1991.

[17] Y. Ye, R. A. Tapia and Y. Zhang. A superlinearly convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Technical Report TR91-22, Dept. Mathematical Sciences, Rice University, 1991.

[18] Y. Zhang and R. A. Tapia. A quadratically convergent polynomial primal-dual interior-point algorithm for linear programming. Technical Report No. 90-40, Dept. of Mathematical Sciences, Rice University, 1990.

[19] Y. Zhang, R. A. Tapia and J. E. Dennis. On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms. Technical Report TR90-6, Dept. Mathematical Sciences, Rice University, 1990, to appear in *SIAM J. Optimization.*

[20] Y. Zhang, R. A. Tapia and F. Potra. On the superlinear convergence of interior point algorithms for a general class of problems. Technical Report TR90-9, Dept. Mathematical Sciences, Rice University, 1990, to appear in *SIAM J. Optimization.*