A Quadratically Convergent Polynomial Primal-Dual Interior-Point Algorithm for Linear Programming

Yin Zhang
R. A. Tapia

CRPC-TR90119
June, 1990

Center for Research on Parallel Computa
Rice University
P.O. Box 1892
Houston, TX 77251-1892

Revised April, 1991.

A Quadratically Convergent Polynomial Primal-Dual Interior-Point Algorithm for Linear Programming

Yin Zhang

and

R.A. Tapia

June, 1990
(Revised April, 1991)

TR90-40

A Quadratically Convergent Polynomial Primal-Dual Interior-Point Algorithm for Linear Programming *

Yin Zhang [†] and Richard Tapia [‡]

June, 1990
Revised April, 1991

*Research 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

[‡]Department of Mathematical Sciences, Rice University, Houston, Texas, 77251-1892

Abstract

The choice of the centering (or barrier) parameter and the step length parameter are the fundamental issues in primal-dual interior-point algorithms for linear programming. Various choices for these two parameters have been proposed that lead to polynomial algorithms. Recently, Zhang, Tapia and Dennis gave conditions that these choices must satisfy in order to achieve quadratic or superlinear convergence. However, it has not been shown that these conditions for fast convergence are compatible with the choices that lead to polynomiality. It is worth noting that none of the existing polynomial algorithms satisfies these fast convergence requirements. This paper gives an affirmative answer to the question: can an algorithm be both polynomial and superlinearly convergent? We construct and analyze a "large step" algorithm that possesses both polynomiality and Q-superlinear convergence. For nondegenerate problems, the convergence rate is actually Q-quadratic.

Key Words: Linear programming, Primal-dual interior-point algorithms, Polynomiality, Quadratic and superlinear convergence.

Abbreviated Title: A Quadratically Convergent Polynomial Algorithm

1 Introduction

We consider linear programs in the standard form:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq 0, \end{array} \tag{1.1}$$

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

The first-order optimality conditions for (1.1) can be written

$$\begin{pmatrix} Ax - b \\ A^T\lambda + y - c \\ XYe \end{pmatrix} = 0, \ \ (x, y) \geq 0, \tag{1.2}$$

where λ and y are dual variables, $X = \text{diag}(x)$, $Y = \text{diag}(y)$ and e has all components equal to one. To facilitate our presentation, we will eliminate the dual variable λ from the above system (though such an elimination may not be advisable from a practical point of view). Let $B \in \mathbf{R}^{(n-m) \times n}$ be any matrix such that the columns of B^T form a basis for the null space of A. Pre-multiply the second equation by the nonsingular matrix $[A^T \ B^T]^T$. Notice that $BA^T = 0$, so

$$0 = \begin{bmatrix} A \\ B \end{bmatrix} (A^T\lambda + y - c) = \begin{pmatrix} AA^T\lambda + A(y - c) \\ By - Bc \end{pmatrix}.$$

Since AA^T is nonsingular, λ is uniquely determined once y is known. Removing the equation for λ, we arrive at the following $2n$ by $2n$ nonlinear system with non-negativity constraints on the variables:

$$F(x, y) = \begin{pmatrix} Ax - b \\ By - Bc \\ XYe \end{pmatrix} = 0, \ \ (x, y) \geq 0. \tag{1.3}$$

for a vector $v \in \mathbf{R}^n$. The symbol $\|\cdot\|$ denotes the ℓ_2 norm unless otherwise stated.

The paper is organized as follows. In Section 2, we describe a general interior-point algorithmic framework for problem (1.1) based on the nonlinear system (1.3) and give a brief survey of existing results for algorithms that fall into this framework. In Sections 3 and 4, we specify our procedures for determining the step length and for choosing the centering parameter. We state our algorithm in Section 5. Polynomiality is established in Sections 6. Quadratic convergence for nondegenerate problems is established in Section 7, and superlinear convergence for all problems is established in Section 8. Concluding remarks are given in Section 9.

2 General Algorithm

Algorithm 1 (General Algorithm)

Given a strictly feasible pair (x_0, y_0). *For* $k = 0, 1, 2, \ldots,$ *do*

Step 1 *Compute the Newton step*

$$\begin{pmatrix} \Delta x_k^N \\ \Delta y_k^N \end{pmatrix} = -[F'(x_k, y_k)]^{-1} F(x_k, y_k)$$

and the centering step

$$\begin{pmatrix} \Delta x_k^C \\ \Delta y_k^C \end{pmatrix} = \tfrac{1}{n} x_k^T y_k [F'(x_k, y_k)]^{-1} \begin{pmatrix} 0 \\ e \end{pmatrix}.$$

Step 2 *Choose* $\sigma_k \in (0, 1)$ *and form the combined step*

$$\begin{pmatrix} \Delta x_k \\ \Delta y_k \end{pmatrix} = \begin{pmatrix} \Delta x_k^N \\ \Delta y_k^N \end{pmatrix} + \sigma_k \begin{pmatrix} \Delta x_k^C \\ \Delta y_k^C \end{pmatrix}.$$

Step 3 *Choose* $\alpha_k \in (0, \hat{\alpha}_k)$, *where*

$$\hat{\alpha}_k = \frac{-1}{\min(X_k^{-1} \Delta x_k, Y_k^{-1} \Delta y_k)}.$$

zero, since we must enforce the requirement $\alpha_k < \hat{\alpha}_k$ and $\hat{\alpha}_k$ is not directly under our control.

A number of existing primal-dual algorithms fit into the above general algorithmic framework with different choices for the parameters σ_k and α_k. For example, in the primal-dual algorithm of Kojima, Mizuno and Yoshise [2], σ_k is a constant and α_k is a particular function of σ_k. They showed that their algorithm requires at most $O(nL)$ iterations to reduce the duality gap by a factor of 2^{-L}. Other examples include the Todd and Ye [8] primal-dual potential reduction algorithm and the Monteiro and Adler [6] path-following primal-dual algorithm. Todd and Ye's algorithm uses the choice

$$\sigma_k = \frac{\sqrt{n}}{\sqrt{n} + \nu}$$

where ν is a constant. In Monteiro and Adler's algorithm,

$$\sigma_k = 1 - \frac{\delta}{\sqrt{n}}$$

where δ is a constant (Monteiro and Adler actually used $\delta = 0.35$ in their analysis). In both algorithms, a rather short step length α_k is required. Furthermore, both of these algorithms require at most $O(\sqrt{n}L)$ iterations to reduce the duality gap to 2^{-L}. This is the best complexity bound obtained for linear programming so far. Observe that all three algorithms use constant σ_k. In each of the three cases if σ denotes the constant value of σ_k, then Q-superlinear convergence is possible (see (2.6)) only if

$$\alpha_k \to \frac{1}{1-\sigma},$$

which seems extremely unlikely.

In analyzing the convergence of Algorithm 1, a central quantity is

$$\eta_k = \frac{x_k^T y_k / n}{\min(X_k Y_k e)}. \tag{2.7}$$

Since $\frac{1}{n} x_k^T y_k$ is the average value of the components of $X_k Y_k e$, it is clear that $\eta_k \geq 1$. In all the above mentioned polynomial algorithms, it is essential that the sequence $\{\eta_k\}$ be bounded.

exists in most real-world problems. For degenerate solutions, the best convergence that has been established is Q-superlinear, as stated in Theorem 2.2.

Although all the existing polynomial primal-dual interior-point algorithms satisfy assumption (ii) of Theorem 2.2, none of them satisfy assumption (iii), i.e., $\sigma_k \to 0$ and $\alpha_k \to 1$. In fact, in several polynomial algorithms, for example Todd and Ye's and Monteiro and Adler's, the values of σ_k are close to one. From Zhang, Tapia and Dennis [9] it follows that these algorithms will most likely have slow Q-linear convergence. Hence while their global behavior may be excellent, their local behavior can be improved.

Recently, in a number of performance-oriented primal-dual algorithms, for example the ones implemented by Choi *et. al.* [1], McShane *et. al.* [5] and Lustig *et. al.* [4], very small values of σ_k were used and also long steps were taken. Impressive numerical results were obtained from these implementations though polynomial complexity bounds are not known. Hence while their local behavior may be good, their global behavior is in question from a theoretical standpoint.

In this work, we develop a primal-dual interior-point polynomial algorithm that gives quadratic convergence for nondegenerate optimal solutions and superlinear convergence for degenerate solutions. Hence, from a mathematical point of view, both the global and the local behavior will be good. This new algorithm is still of a theoretical nature. However, the fact that polynomiality and quadratic or superlinear convergence can be achieved simultaneously by one algorithm provides motivation for practical implementations of the conditions $\sigma_k = O(x_k^T y_k)$ and $\alpha_k = 1 + O(x_k^T y_k)$ for fast convergence.

3 Determining the Step Length

In the previous section we mentioned that both polynomiality and superlinear convergence essentially require that the sequence $\{\eta_k\}$ be bounded. The most straightforward way of accomplishing this objective is to explicitly enforce the uniform bound on the

we have

$$f_k^i(\alpha) = f_k^i - (f_k^i - \sigma_k f_k^{\text{ave}})\alpha + s_k^i \alpha^2 \tag{3.5}$$

and

$$f_k^{\text{ave}}(\alpha) = f_k^{\text{ave}}[1 - (1 - \sigma_k)\alpha]. \tag{3.6}$$

Hence, $f_k^i(\alpha)$ is a quadratics (so $f_k^{\min}(\alpha)$ and $f_k^{\max}(\alpha)$ are piecewise quadratic) and $f_k^{\text{ave}}(\alpha)$ is linear.

For notational convenience, let us introduce the piecewise quadratic function

$$h(\alpha) \stackrel{\text{def}}{=} f_k^{\min}(\alpha) - \gamma_k f_k^{\text{ave}}(\alpha). \tag{3.7}$$

It follows that condition (3.3) is equivalent to

$$h(\alpha) \geq 0, \quad \alpha > 0. \tag{3.8}$$

In determining α_k we will use the following quantity:

$$\alpha_k^\gamma \stackrel{\text{def}}{=} \min\{\alpha > 0 : h(\alpha) = 0\}. \tag{3.9}$$

Recall that $\hat{\alpha}_k$ is defined in Step 3 of the general algorithm (see Section 2).

Lemma 3.1 *The quantity α_k^γ is well defined and $\alpha_k^\gamma \in (0, \hat{\alpha}_k)$. Moreover, condition (3.3) is satisfied for all $\alpha \in (0, \alpha_k^\gamma]$.*

Proof: Let us examine the function $h(\alpha)$. It follows from the definitions of γ_k and $\hat{\alpha}_k$ that

$$h(0) = f_k^{\min} - \gamma_k f_k^{\text{ave}} \geq 0$$

and

$$h(\hat{\alpha}_k) = f_k^{\min}(\hat{\alpha}_k) - \gamma_k f_k^{\text{ave}}(\hat{\alpha}_k) = -\gamma_k f_k^{\text{ave}}(\hat{\alpha}_k) < 0.$$

Hence it follows from the continuity of $h(\alpha)$ that $h(\alpha)$ has a root in $[0, \hat{\alpha}_k)$. When $h(0) > 0$, $h(\alpha)$ obviously has a root in $(0, \hat{\alpha}_k)$. When $h(0) = 0$, it can be verified that

We will also use the following quantity in determining α_k:

$$\alpha_k^{\Gamma} \stackrel{\text{def}}{=} \min\{\alpha > 0 : H(\alpha) = 0\}. \tag{3.15}$$

Analogous to Lemma 3.1 for condition (3.3), we have the following lemma for condition (3.11).

Lemma 3.2 *The quantity α_k^{Γ} is well-defined and $\alpha_k^{\Gamma} \in (0, \hat{\alpha}_k)$. Moreover, condition (3.11) is satisfied by all $\alpha \in (0, \alpha_k^{\Gamma}]$.*

Proof: The proof is similar to that for Lemma 3.1, so we omit it. □

Analogous to the expression (3.10) for condition (3.3), we have for condition (3.11)

$$\alpha_k^{\Gamma} = \min\{\alpha > 0 : f_k^i(\alpha) - \Gamma_k f_k^{\text{ave}}(\alpha) = 0,\ i = 1, 2, \ldots, n\}. \tag{3.16}$$

For the sake of simplicity, we will enforce the conditions

$$\gamma_k \leq 1/2 \ \text{ and } \ \Gamma_k \geq 2. \tag{3.17}$$

The specific values in (3.17) do not constitute a loss of generality because they will only affect expressions for some constants in our analysis. These values of γ_k and Γ_k will result in much simplified expressions for those constants.

From (2.2), we see that for fixed σ_k a larger step length α_k will produce a larger reduction in the duality gap. So it is always desirable to take the largest step length possible as long as other requirements are satisfied. Our procedure for determining the step length α_k is summarized as follows.

Procedure 1 (Step length Criterion)

Given

$$0 < \gamma \leq \min(1/2, f_0^{\text{min}}/f_0^{\text{ave}}), \ \ \max(2, f_0^{\text{max}}/f_0^{\text{ave}}) \leq \Gamma < n. \tag{3.18}$$

Step 1 *Choose $\gamma_k \in [\gamma, \min(1/2, f_k^{\text{min}}/f_k^{\text{ave}})]$ and $\Gamma_k \in [\max(2, f_k^{\text{max}}/f_k^{\text{ave}}), \Gamma]$.*

Lemma 3.4 *Let α_k be given by Procedure 1. Then*

$$\alpha_k \geq \min\left(1, \frac{(1-\gamma_k)\sigma_k f_k^{\text{ave}}}{-\min(s_k)}, \frac{(\Gamma_k - 1)\sigma_k f_k^{\text{ave}}}{\max(s_k)}\right). \tag{3.19}$$

Moreover,

$$\alpha_k \geq \min\left(1, \frac{\sigma_k f_k^{\text{ave}}}{2\|s_k\|_\infty}\right). \tag{3.20}$$

Proof: From (3.9), α_k^γ is a positive root of $f_k^i(\alpha) - \gamma_k f_k^{\text{ave}}(\alpha)$ for some index i. Noticing that for $\alpha \in [0,1]$, $f_k^{\text{ave}}(\alpha)$ is positive, and using Lemma 3.3, for all index i, $\alpha \in [0,1]$ and $\gamma_k \geq 0$, we have

$$\begin{aligned} f_k^i(\alpha) - \gamma_k f_k^{\text{ave}}(\alpha) &\geq f_k^{\min} - (f_k^{\min} - \sigma_k f_k^{\text{ave}})\alpha + \min(s_k)\alpha^2 - \gamma_k f_k^{\text{ave}}(\alpha) \\ &= (f_k^{\min} - \gamma_k f_k^{\text{ave}})(1-\alpha) + (1-\gamma_k)\sigma_k f_k^{\text{ave}}\alpha + \min(s_k)\alpha^2 \\ &\geq (1-\gamma_k)\sigma_k f_k^{\text{ave}}\alpha + \min(s_k)\alpha^2. \end{aligned} \tag{3.21}$$

If $\min(s_k) = 0$, then $h(\alpha) > 0$ for $\alpha \in (0,1]$. Therefore, we will have $\alpha_k^\gamma > 1$. Now assume $\min(s_k) < 0$. Then the quadratic in the right-hand side of the last inequality in (3.21) has a unique positive root

$$\bar{\alpha}_k = \frac{(1-\gamma_k)\sigma_k f_k^{\text{ave}}}{-\min(s_k)}.$$

Hence, if $\alpha_k^\gamma \leq 1$, from (3.21) we must have $\alpha_k^\gamma \geq \bar{\alpha}_k$. This proves that

$$\alpha_k^\gamma \geq \min\left(1, \frac{(1-\gamma_k)\sigma_k f_k^{\text{ave}}}{-\min(s_k)}\right). \tag{3.22}$$

Similarly, we can prove that

$$\alpha_k^\Gamma \geq \min\left(1, \frac{(\Gamma_k - 1)\sigma_k f_k^{\text{ave}}}{\max(s_k)}\right). \tag{3.23}$$

Combining (3.22) and (3.23), we obtain (3.19).

Finally, (3.20) follows from the facts that $\|s_k\|_\infty = \max\{-\min(s_k), \max(s_k)\}$ and

$$\frac{1}{2} \leq 1 - \gamma_k < 1 \leq \Gamma_k - 1.$$

This completes the proof. □

Procedure 2 (Centering Parameter Criterion)

Given

$$\sigma \in (0,1), \quad \rho^l = \frac{\gamma^2\sigma}{2n}, \quad \rho^u \geq \frac{\gamma^2\sigma}{n}. \tag{4.4}$$

Step 1 *Compute ω_k from (4.2).*

Step 2 *Compute $\rho_k^u = \min(\rho^u, \sigma/\omega_k)$.*

Step 3 *Choose $\rho_k \in [(\rho^l + \rho_k^u)/2, \rho_k^u]$.*

Step 4 *Let $\sigma_k = \rho_k\omega_k$.*

Since $\sigma_k = \rho_k\omega_k$ and $\rho_k \in [\rho^l, \rho_k^u]$, we have $\sigma_k \in [\rho^l\omega_k, \rho_k^u\omega_k]$. In addition, we require that σ_k be greater than the midpoint of the interval. This requirement is needed in our proof of superlinear convergence. It is evident that σ_k is bounded away from one because $\sigma_k \leq \sigma < 1$. The reasons why the centering parameter is so chosen will hopefully become clear as our discussion proceeds.

5 Algorithm Description

Now we formally state our primal-dual interior-point algorithm.

Algorithm 2 *Given a strictly feasible pair (x_0, y_0). Choose (recall (3.18))*

$$0 < \gamma \leq \min(1/2, f_0^{\text{min}}/f_0^{\text{ave}}), \quad \max(2, f_0^{\text{max}}/f_0^{\text{ave}}) \leq \Gamma < n,$$

and $\sigma \in (0,1)$. Set $\rho^l = \gamma^2\sigma/2n$ and $\rho^u = \gamma^2\sigma/n$ (i.e., (4.4)). For $k = 0, 1, 2, \ldots,$ do

Step 1 *Compute the Newton step and the centering step from Algorithm 1.*

Step 2 *Choose σ_k by Procedure 2 and form $(\Delta x_k, \Delta y_k)$ from Algorithm 1.*

Step 3 *Choose α_k by Procedure 1.*

Step 4 *Form (x_{k+1}, y_{k+1}) from Algorithm 1.*

Hence, it follows from (3.11), (3.20) and Procedure 2 that

$$\alpha_k \geq \min\left(1, \frac{\rho_k f_k^{\mathbf{ave}}}{8 f_k^{\mathbf{max}}}\right) \geq \min\left(1, \frac{\rho_k}{8\Gamma}\right) \geq \frac{\rho^l}{8\Gamma}. \tag{6.1}$$

Substituting ρ^l (see (4.4)) into the above expression, we obtain

$$\alpha_k \geq \frac{\sigma\gamma^2}{16\Gamma n}.$$

The proof is completed by substituting the above inequality into (2.6) and noticing that $\sigma_k \leq \sigma$. □

The following corollary follows immediately from Theorem 6.1. By a standard argument, it leads to polynomiality assuming integral data.

Corollary 6.1 *Assume $x_0^T y_0 \leq 2^{\nu L}$, where $L > 0$ and $\nu > 0$. Then in at most $O(nL)$ iterations, Algorithm 2 will produce (x_k, y_k) such that $x_k^T y_k \leq 2^{-L}$.*

Proof: From Theorem 6.1,

$$x_k^T y_k \leq (1 - \delta/n)^k x_0^T y_0 \leq (1 - \delta/n)^k 2^{\nu L}.$$

Let $(1 - \delta/n)^k 2^{\nu L} = 2^{-L}$ and take the natural logarithm of both sides. We have $k = -(\ln 2)(1 + \nu)L/\ln(1 - \delta/n)$. Observe that for $x \in (0, 1)$

$$\ln(1 - x) = -\sum_{k=1}^{\infty} \frac{x^k}{k} < -x.$$

Therefore,

$$k \leq (\ln 2)(1 + \nu)L/(\delta/n) = O(nL).$$

This completes the proof. □

7 Quadratic Convergence

In this section, we will apply Theorem 2.1 to establish that under strict complementarity and nondegeneracy assumptions our algorithm converges Q-quadratically. It can be

and

$$q_k^N = \begin{pmatrix} -1 \\ \vdots \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + O(x_k^T y_k), \qquad q_k^C = \begin{pmatrix} \frac{x_k^T y_k/n}{[X_k Y_k e]^1} \\ \vdots \\ \frac{x_k^T y_k/n}{[X_k Y_k e]^m} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + O(x_k^T y_k),$$

where the number of zero is m in p_k^N and p_k^C, and $n-m$ in q_k^N and q_k^C.

Now we are ready to state and prove our quadratic convergence theorem.

Theorem 7.1 (Quadratic convergence)

Let (x_*, y_*) be a solution of problem (1.3) and $\{(x_k, y_k)\}$ be generated by Algorithm 2. Assume

(i) *strict complementarity holds at (x_*, y_*),*

(ii) *x_* is a nondegenerate vertex of (1.1),*

(iii) *ρ^u is sufficiently large, e.g., $\rho^u \geq 16\Gamma$.*

Then $\{(x_k, y_k)\}$ converges to (x_*, y_*) Q-quadratically and $\{X_k Y_k e\}$, component-wise, converges to zero Q-superlinearly.

Proof: We first prove $\sigma_k = O(x_k^T y_k)$. Observe from Lemma 7.1 that for each index i either the "p" terms ($(p_k^N)^i$ and $(p_k^C)^i$) or the "q" terms ($(q_k^N)^i$ and $(q_k^C)^i$) are $O(x_k^T y_k)$. Thus, the quantity ω_k (see its definition (4.2)) is $O(x_k^T y_k)$. So is σ_k because $\sigma_k \leq \rho^u \omega_k$.

Since $\omega_k \to 0$, from the choice of ρ_k^u in Step 2 of we have for k sufficiently large

$$\rho_k^u = \rho^u \quad \text{and} \quad \rho_k \geq \frac{1}{2}(\rho^l + \rho^u). \tag{7.1}$$

We observe that if ρ^u is sufficiently large, e.g., $\rho^u \geq 16\Gamma$, (i.e., σ_k is not forced to approach zero too quickly), then the step length α_k will eventually be one, as can be seen from (6.1).

and define the distance from σ to the set Σ_k as

$$\text{dist}(\sigma, \Sigma_k) = \min\{|\sigma - \varsigma| : \varsigma \in \Sigma_k\}.$$

We choose σ_k according to Procedure 2 with the additional restriction that

$$\text{dist}(\sigma_k, \Sigma_k) \geq \pi_k \omega_k / (8n + 4). \tag{8.3}$$

In other words, we require not only

$$\sigma_k \in [0.5(\rho^l + \rho_k^u)\omega_k, \rho_k^u \omega_k] \tag{8.4}$$

but also that σ_k be bounded away from the set Σ_k by at least $\pi_k \omega_k / (8n + 4)$. Since $\{\pi_k\}$ is bounded away from zero, we see from (8.3) that $\{\text{dist}(\sigma_k, \Sigma_k)\}$ is bounded away from zero if $\{\omega_k\}$ is bounded away from zero.

The purpose of introducing condition (8.3) is to avoid the situation where $p_k^i = (p_k^N)^i + \sigma_k (p_k^C)^i$ (say) converges to zero but $(p_k^N)^i$ and $(p_k^C)^i$ do not. Although we believe that this situation is extremely unlikely to happen, we have not been able to rule it out.

Lemma 8.1 *The set of σ_k's satisfying (8.3) and (8.4) is nonempty.*

Proof: The length of the interval in (8.4) is $\pi_k \omega_k / 2$. Partition this interval into $2n + 1$ equal sub-intervals, each having length $\pi_k \omega_k / (4n + 2)$. If the interior of any one of the sub-intervals does not intersect Σ_k, then the midpoint of that sub-interval will satisfy (8.3) and (8.4). Since Σ_k has only $2n$ points, it cannot intersect the interiors of all the $2n + 1$ sub-intervals. This proves the lemma. □

Now we are well-equipped to prove our superlinear convergence theorem.

Theorem 8.1 (Superlinear convergence)

Let (x_*, y_*) be a solution of problem (1.3) and $\{(x_k, y_k)\}$ be generated by Algorithm 2 with the restriction (8.3) on the centering parameter σ_k. Assume

(i) *strict complementarity holds at (x_*, y_*),*

Now in view of (8.5) we also have $(p^N_{k_0})^i \to 0$. Similarly, we can prove that if $q^i_k \to 0$, then we have both $(q^N_{k_0})^j \to 0$ and $(q^C_{k_0})^j \to 0$. Therefore, for each index i, either $(p^V_{k_0})^i$ and $(p^C_{k_0})^i$, or $(q^N_{k_0})^i$ and $(q^C_{k_0})^i$ converge to zero. Since all these sequences are uniformly bounded (see the proof of Lemma 4.1), this leads to $\omega_{k_0} \to 0$ (see definition (4.2)), contradicting the hypothesis that $\{\omega_{k_0}\}$ is bounded away from zero. This proves that $\omega_k \to 0$. Consequently $\sigma_k \to 0$.

Now we prove $\alpha_k \to 1$. Note that (2.4) can be written as

$$p_k + q_k = -e + \sigma_k \tfrac{1}{n} x_k^T y_k (X_k Y_k)^{-1} e.$$

Since $\frac{1}{n} x_k^T y_k (X_k Y_k)^{-1} e$ is bounded above by $1/\gamma$, as $\sigma_k \to 0$, we have

$$p_k + q_k \to -e.$$

We have shown that for each i, either $p^i_k \to 0$ or $q^i_k \to 0$. Therefore, all p^i_k and q^i_k converge to either 0 or -1. This again implies that $\hat{\alpha}_k \to 1$ (see (7.2)). In view of (6.1) and (7.1), α_k will eventually be equal to one if ρ^u is sufficiently large, e.g., $\rho^u \geq 16\Gamma$. Hence,

$$1 \leq \alpha_k \leq \hat{\alpha}_k \to 1.$$

This completes the proof. □

9 Concluding Remarks

In this paper, we have shown that the two fundamental parameters in primal-dual interior-point algorithms for linear programming can be chosen in such a way that both polynomiality and superlinear convergence are achieved. If the solution is a nondegenerate vertex, then in addition to superlinear convergence we have quadratic convergence.

The current practices in some of the state-of-the-art implementations of primal-dual interior-point algorithms have the following common fundamental features. First, they allow iterates to be very close to the boundary of the positive orthant; second, they

[3] O. Güler and Yinyu Ye. Convergence behavior of some interior-point algorithms. Working paper series No. 91-4, the College of Business Administration, The University of Iowa, 1991.

[4] 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.

[5] 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.

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

[7] 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 Operations Research and Industrial Engineering, Cornell University, 1990.

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

[9] 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.