School of Operations Research and Industrial Engineering
Cornell University
Ithaca, New York 14853-3801

Technical Report No. 1003

May 1992

ON THE CONVERGENCE OF PRIMAL-DUAL INTERIOR POINT METHODS WITH WIDE NEIGHBORHOODS

Levent TUNÇEL[1]

[1]Research supported in part by Mathematical Sciences Institute of Cornell University, and also by NSF, AFOSR and ONR through NSF Grant DMS-8920550.

ON THE CONVERGENCE OF PRIMAL-DUAL INTERIOR POINT METHODS WITH WIDE NEIGHBORHOODS

*Levent TUNÇEL**

School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14853-3801

May, 1992

Abstract

We study primal-dual interior point methods for linear programs. After proposing a new primal-dual potential function we describe an algorithm that is a slightly modified version of existing primal-dual algorithms using wide neighborhoods. Assuming the optimal solution is non-degenerate, the algorithm is 1-step Q-quadratically convergent. We also study the degenerate case and show that the neighborhoods of the central path stay large as the iterates approach the optimal solutions.

Keywords: Linear programming, interior point methods, potential function, quadratic convergence.

*Research supported in part by Mathematical Sciences Institute of Cornell University, and also by NSF, AFOSR and ONR through NSF Grant DMS-8920550.

1 Introduction

There have been many recent developments on the asymptotic convergence rate of interior point methods. Assuming non-degeneracy and strict complementarity, Zhang, Tapia and Dennis [ZTD90] gave conditions on how to choose the centering parameter and the step size so that an interior point algorithm could achieve superlinear or quadratic convergence asymptotically. Zhang and Tapia [ZT90] provided an algorithm that has an $O(nt)$ bound on the number of iterations and is superlinearly convergent (t denotes the order of required improvement in precision). Under the non-degeneracy assumption, they showed that quadratic convergence can be achieved.

Very recently, superlinear and quadratic convergence results for $O(\sqrt{n}t)$ iteration algorithms were obtained:

McShane [Mc92] gives a primal-dual potential reduction algorithm that is superlinearly convergent under the assumption that the iterates converge, and quadratically convergent under the assumption that the optimum solution is non-degenerate. Mehrotra [Meh91] and Ye, Güler, Tapia and Zhang [YGTZ91] prove that the Mizuno-Todd-Ye predictor-corrector algorithm [MTY90] is quadratically convergent without assuming non-degeneracy or that the iterates converge.

In this paper, we study the primal-dual interior point algorithms with wide neighborhoods. Even though the best bound on the number of iterations for these algorithms is $O(nt)$, practically they are more promising than the algorithms that use smaller 2-norm neighborhoods. In the following section, we describe a generic primal-dual interior point algorithm which is due to Mizuno, Todd and Ye [MTY90] and very similar to the algorithms described by Kojima, Mizuno and Yoshise [KMY88], and Zhang and Tapia [ZT90]. In section 3, we propose a new primal-dual potential function for the primal-dual interior point methods with wide neighborhoods, and show that the new potential function fits in more nicely with the wide neighborhoods than the Todd-Ye primal-dual potential function [TY89]. We also show that as long as the iterates lie in a wide neighborhood of the central path the new potential function could be used to get an $O(nt)$ bound on the number of iterations. In section 4, we show that a slight modification of the algorithms suggested in Kojima et al. [KMY88], Mizuno et al. [MTY90], and Zhang and Tapia [ZT90], while keeping the $O(nt)$ bound on the number of iterations, asymptotically achieves 1-step Q-quadratic convergence under the assumption of non-degeneracy of the optimal solution. The difference between our result and Zhang and Tapia's is that we calculate only one projection per iteration. Section 5 includes the study of the degenerate case. We show that when the affine scaling direction stays small, the centering direction could be larger in norm. We provide a theorem showing that in this case the neighborhoods of the

central path do not converge to a point as they converge to the optimal solution set. This suggests that an algorithm keeping its iterates in a neighborhood of the central path (as defined in this paper) may not have convergent iterates.

2 Generic Primal-Dual Algorithm with Wide Neighborhoods

We consider linear programming problems in the following primal (P) and dual (D) forms:

(P)

$$\begin{array}{lrcl} \min & c^T x & & \\ & Ax & = & b \\ & x & \geq & 0, \end{array}$$

(D)

$$\begin{array}{lrcl} \max & b^T y & & \\ & A^T y + s & = & c \\ & s & \geq & 0, \end{array}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Without loss of generality, we will assume A has full row rank and that there exist interior solutions for both problems, i.e.,

$$\mathcal{F}_o := \{(x,s) > 0 : x \in F(P), s \in F(D)\} \neq \emptyset,$$

where $F(P)$ and $F(D)$ denote the set of feasible solutions for the primal and dual problems respectively. Most of the time we will deal only with s as a dual feasible solution. So, whenever we say $s \in F(D)$, we mean that $s \geq 0$ and there exists a $y \in \mathbb{R}^m$ such that $A^T y + s = c$. Given a vector x, X will denote the diagonal matrix whose entries are the components of x, and e will denote the

vector of ones. We will denote the components of a vector using subscripts and the iterate numbers using superscripts. Whenever we ignore superscripts it will be clear from the context what the iterate number should be. Now, we describe the central path and its neighborhoods.

The central path is given by the set of solutions to the following system of equalities (for $\mu > 0$):

$$\begin{aligned} Ax &= b, \quad x \geq 0 && (1)\\ A^T y + s &= c, \quad s \geq 0 && (2)\\ Xs &= \mu e && (3) \end{aligned}$$

Note that when $\mu = 0$, (1)-(3) give necessary and sufficient conditions for optimality. Analyses of the central path have been done by several authors (see, for instance, Sonnevend [So85], Megiddo [Meg88], and Bayer and Lagarias [BL89]). Our objective is to approximately follow this path to an optimal solution. We define some neighborhoods of the central path as given by Mizuno et al. [MTY90]. Let $\beta \in (0,1)$ be a constant; then a 2-norm neighborhood of the central path can be defined as

$$\mathcal{N}_2(\beta) := \{(x,s) \in \mathcal{F}_o : \|Xs - \mu e\|_2 \leq \beta \text{ for } \mu = \frac{x^T s}{n}\}.$$

Henceforth, μ always denotes $\frac{x^T s}{n}$. Kojima et al. [KMY89], Monteiro and Adler [MA89], and Mizuno et al. [MTY90] designed algorithms that use a 2-norm neighborhood. Using the ∞-norm or just one side of the ∞-norm, wider neighborhoods have been defined and used by Kojima et al. [KMY88], Zhang and Tapia [ZT90], and Mizuno et al. [MTY90]. Using the ∞-norm we have

$$\mathcal{N}_\infty(\beta) := \{(x,s) \in \mathcal{F}_o : \|Xs - \mu e\|_\infty \leq \beta\}$$

and using only one side of the ∞-norm,

$$\mathcal{N}_\infty^-(\beta) := \{(x,s) \in \mathcal{F}_o : \|Xs - \mu e\|_\infty^- \leq \beta\}.$$

Here, for $v \in \mathbb{R}^n$, $\|v\|_\infty^- := -\min\{0, \min\{v_j\}\}$. Clearly, for a given $\beta \in (0,1)$, $\mathcal{N}_2(\beta)$ is the smallest and $\mathcal{N}_\infty^-(\beta)$ is the largest of the three neighborhoods defined here.

Suppose we have an initial interior point solution $(x, s) \in \mathcal{F}_o$. Then a search direction $(\bar{dx}, \bar{ds})$ can be generated by solving the following set of equalities (see Kojima et al. [KMY88]):

$$\begin{aligned} A\bar{dx} &= 0 && (4)\\ A^T\bar{dy} + \bar{ds} &= 0 && (5)\\ S\bar{dx} + X\bar{ds} &= \gamma\mu e - Xs && (6) \end{aligned}$$

which is equivalent to solving

$$\begin{aligned} \bar{A}dx &= 0 && (7)\\ \bar{A}^T dy + ds &= 0 && (8)\\ dx + ds &= (-X^{1/2}S^{1/2} + \gamma\mu X^{-1/2}S^{-1/2})e && (9) \end{aligned}$$

where $\bar{A} := AX^{1/2}S^{-1/2}$. The equivalence of the above two systems can be easily checked by substituting $dx = X^{-1/2}S^{1/2}\bar{dx}$, $ds = X^{1/2}S^{-1/2}\bar{ds}$. The solution of (7)-(9) is

$$\begin{aligned} dx &= -P_{\bar{A}}(X^{1/2}S^{1/2} - \gamma\mu X^{-1/2}S^{-1/2})e\\ ds &= -(I - P_{\bar{A}})(X^{1/2}S^{1/2} - \gamma\mu X^{-1/2}S^{-1/2})e \end{aligned}$$

where $P_{\bar{A}} := I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}$, the projection matrix into the null space of $\bar{A}$. Now, a generic primal-dual algorithm can be easily described:

Algorithm:

Given $(x^o, s^o) \in \mathcal{N}$ with $(x^o)^T s^o \leq 2^t$, $k := 0$,

While (stopping criterion is not met) do

$(x, s) := (x^k, s^k)$

compute $(\bar{dx}, \bar{ds})$ from (4)-(6) using $\gamma = \gamma_k$

choose step size $\alpha_k \in (0, 1)$ such that $(x + \alpha_k \bar{dx}, s + \alpha_k \bar{ds}) \in \mathcal{N}$

$(x^{k+1}, s^{k+1}) := (x + \alpha_k \bar{dx}, s + \alpha_k \bar{ds})$

$k := k + 1$

end

The stopping criterion could be taken as $x^T s < 2^{-t}$ to get an $O(nt)$ complexity bound on the number of iterations. Since we are also interested in the asymptotic convergence rate of the algorithm, we leave the termination criterion flexible. Mizuno et al. [MTY90] study the above algorithm with $\mathcal{N} = \mathcal{N}_\infty(\beta)$ and $\mathcal{N} = \mathcal{N}_\infty^-(\beta)$. Very similar algorithms were also studied by Kojima et al. [KMY88] and Zhang and Tapia [ZT90]. The differences occur in how one chooses γ, the centering parameter, and α, the step size. In the next section, we propose a new potential function to have an alternative way of choosing the step size α.

3 A Primal-Dual Potential Function

In his seminal paper [Ka84], Karmarkar introduced a potential function which was very useful for proving an upper bound on the number of iterations for his algorithm. Since then many authors used variants of that potential function to prove bounds or to get search directions (see for instance Gonzaga [Go90] or Todd and Ye [TY90]). Under the assumption that the optimal objective function value is zero, Karmarkar's potential function for the primal problem is:

$$\phi^P(x) := n \log(c^T x) - \sum_{j=1}^{n} \log(x_j).$$

In dual form it corresponds to:

$$\phi^D(s) := n\log(-b^T y) - \sum_{j=1}^{n} \log(s_j).$$

Very interestingly, Powell [Po90] provides an example (in dual form) with infinitely many constraints on which Karmarkar's algorithm fails to converge to the optimum solution. Powell's result assumes that the step size is chosen to minimize the potential function for a given search direction. One problem with the potential function is that when one has infinitely many constraints $\frac{1}{n}\sum_{j=1}^{n}\log(s_j)$ becomes $\int \log(s)$ and the potential function loses its barrier property (i.e. we no longer have the property that $s^k \to \partial F(D) \Rightarrow \phi^D(s^k) \to \infty$). One way of getting around this problem is to try $\log(\inf(s))$ as the barrier. In the finite case for the primal problem this corresponds to

$$\psi_q^P(x) := (q+1)\log(c^T x) - \log(\min_j\{x_j\}),$$

where $q \geq 0$. ψ_q^P is not likely to serve as an effective potential function for it is not scale invariant. However, if we had a good scaling of the problem (for each iterate), it could work. We get the scaling from the dual problem and propose the following primal-dual potential function:

$$\psi_q(x,s) := (q+1)\log(\frac{x^T s}{n}) - \log(\min_j\{x_j s_j\})$$

where $q > 0$. We also consider ψ_q with $q = 0$, which can be viewed (see below) as a centering function. We compare $\psi_q(x,s)$ with the Todd-Ye primal-dual potential function [TY90]:

$$\phi_q(x,s) := (q+n)\log(x^T s) - \sum_{j=1}^{n}\log(x_j s_j)$$

where $q > 0$ (see also Tanabe [Ta87]). Again, we also consider ϕ_q with $q = 0$ which yields a different centering function.

There are two crucial properties needed from a primal-dual potential function. First, the potential function should be scale invariant with respect to primal-dual scaling. Secondly, it should measure the duality gap while keeping the objective function part ($\log(x^T s)$ or $\log(\frac{x^T s}{n})$) and the barrier part ($\sum_{j=1}^{n}\log(x_j s_j)$ or $\log(\min_j\{x_j s_j\})$) well balanced (this intuitively means that if we can decrease $x^T s$ "enough" the potential function should let the iterates move "closer" to the boundary and if we cannot then it should try to move us away from the boundary).

Proposition 3.1. *a) Let $\Lambda \in \mathbb{R}^{n\times n}$ be a diagonal positive definite matrix. Then*

$$\psi_q(\Lambda x, \Lambda^{-1}s) = \psi_q(x,s).$$

b) If for $(x,s) \in \mathcal{F}_o$ and $q > 0$, we have $\psi_q(x,s) \leq -qt - q\log(n)$, then

$$x^T s \leq 2^{-t}.$$

Proof: a) is clear. For b) note that

$$\psi_0(x,s) = \log(\frac{x^T s}{n}) - \log(\min_j\{x_j s_j\}) = \log(\frac{x^T s/n}{\min_j\{x_j s_j\}}) \geq 0. \tag{10}$$

So, $\psi_q(x,s) \leq -qt - q\log(n)$ implies $q\log(\frac{x^T s}{n}) \leq -qt - q\log(n)$, from which we deduce $x^T s \leq 2^{-t}$. □

We would like to point out that the wide neighborhoods $\mathcal{N}_\infty(\beta)$ and $\mathcal{N}_\infty^-(\beta)$ can also be written as

$$\mathcal{N}_\infty(\beta) = \{(x,s) \in \mathcal{F}_o : (1-\beta)\mu \leq x_j s_j \leq (1+\beta)\mu, \forall j\},$$

$$\mathcal{N}_\infty^-(\beta) = \{(x,s) \in \mathcal{F}_o : \min_j\{x_j s_j\} \geq (1-\beta)\mu\}.$$

Note that in the proof of Proposition 3.1 the way ψ_q balances $\log(\frac{x^T s}{n})$ and $\log(\min_j\{x_j s_j\})$ is closely related to the inequality defining the wide neighborhoods. Indeed, from (10) we have $\psi_0(x,s) \geq 0$ while it is easy to see that

$$\mathcal{N}_\infty(\beta) \subseteq \mathcal{N}_\infty^-(\beta) = \{(x,s) \in \mathcal{F}_o : \psi_0(x,s) \leq \log(\frac{1}{1-\beta})\}.$$

So, in a sense ψ_q seems a more natural potential function to use with these wide neighborhoods. The related algorithm described by Mizuno et al. [MTY90] uses the Todd-Ye potential function. As a result, in the corresponding convergence proof, one has to make a transition between the neighborhood information and the change in the value of the Todd-Ye potential function. Whereas here such a step is unnecessary due to the fact that the neighborhood $\mathcal{N}_\infty(\beta)$ defines a level

set of the centering function ψ_0 of the new potential function. Lastly, we would like to note that even though the ratio $\frac{\min_j\{x_j s_j\}}{x^T s/n}$ is a very relaxed measure of the centrality of (x, s), it has been widely and effectively used in many important results for interior point methods. It is interesting that simple bounds on this ratio are all one needs to establish nice results (see for instance [GY91], [Meh91], and [YGTZ91]).

4 Analysis and Convergence Results

First we show how one can use the new potential function to determine the step size such that an $O(nt)$ bound on the number of iterations can be achieved. Then we show how to phase out the centering parameter (as suggested by Zhang et al. [ZTD89] and Zhang and Tapia [ZT90]) to get quadratic convergence.

Let $\bar{\gamma} \in (0,1)$, $\beta \in (0,1)$ be constants. We define the centering parameter in a way that will keep the $O(nt)$ bound on the number of iterations and will make it fade away as $x^T s \to 0$ to ensure quadratic convergence:

$$\gamma_k := \min\{\bar{\gamma}, \rho \frac{(x^k)^T s^k}{n}\} \tag{11}$$

We will eventually let $\rho = Cn^2/\beta$ where C is a data dependent constant to be chosen later.

Let (x, s) be the current iterate with $(x, s) \in \mathcal{N}_\infty^-(\beta)$, and let (x^+, s^+) denote the next iterate. Then for a given step size $\alpha \in [0, 1]$,

$$x^+ = x - \alpha X^{1/2} S^{-1/2} P_{\bar{A}} v \tag{12}$$

$$s^+ = s - \alpha X^{-1/2} S^{1/2} (I - P_{\bar{A}}) v, \tag{13}$$

where $v := (X^{1/2}S^{1/2} - \gamma\mu X^{-1/2}S^{-1/2})e$. We will denote $v_p := P_{\bar{A}} v$. The analysis given here is very similar to the one given by Mizuno et al. [MTY90]. From (12)-(13) we obtain

$$(x^+)^T s^+ = x^T s - \alpha e^T X^{1/2} S^{1/2} [v_p + (v - v_p)] + \alpha^2 v_p^T (v - v_p) \tag{14}$$

$$= x^T s - \alpha(x^T s - \gamma\mu n) \tag{15}$$

$$= [1 - \alpha(1 - \gamma)] x^T s. \tag{16}$$

So, the duality gap decreases by a fraction dependent on α_k and γ_k. We also have

$$\begin{aligned} x_j^+ s_j^+ &= x_j s_j - \alpha\sqrt{x_j s_j}[(v_p)_j + (v - v_p)_j] + \alpha^2 (v_p)_j (v - v_p)_j & (17)\\ &= x_j s_j - \alpha(x_j s_j - \gamma\mu) + \alpha^2 (v_p)_j (v - v_p)_j & (18)\\ &= (1-\alpha) x_j s_j + \alpha\gamma\mu + \alpha^2 (v_p)_j (v - v_p)_j. & (19) \end{aligned}$$

So, to guarantee that $(x^+, s^+) \in \mathcal{N}_\infty^-(\beta)$ it suffices to have (for all j)

$$(1-\alpha)\frac{x_j s_j}{\mu} + \alpha\gamma - \frac{\alpha^2}{\mu}|(v_p)_j (v - v_p)_j| \geq [1 - \alpha(1-\gamma)](1-\beta).$$

We know $(x, s) \in \mathcal{N}_\infty^-(\beta)$, hence $\frac{x_j s_j}{\mu_k} \geq (1-\beta)$. So, it suffices to have

$$(1-\alpha)(1-\beta) + \alpha\gamma - \frac{\alpha^2}{\mu}|(v_p)_j (v - v_p)_j| \;\geq\; [1 - \alpha(1-\gamma)](1-\beta)$$

for all j. The above holds if for $\alpha_k \in [0, 1)$ we have

$$\alpha_k \leq \frac{\beta\gamma_k\mu_k}{|\min_j\{(v_p)_j (v - v_p)_j\}|}.$$

In order to ensure that $\alpha_k < 1$, we will choose

$$\alpha_k = \min\{\frac{\beta\gamma_k\mu_k}{|\min_j\{(v_p)_j (v - v_p)_j\}|}, \frac{1}{1+\gamma_k}\}. \qquad (20)$$

Note that in (20) the second term in the minimum is included to guarantee that $\alpha_k < 1$. As long as the minimum is given by the second term we would like to make sure that $(1 - \alpha_k)$ is $O(x^T s)$. There are many ways of doing this. Doing it our way shows a more explicit relation between the centering coefficient and the step size, and (11) shows that $(1 - \alpha_k)$ is of the correct order.

Now we turn to the analysis of the search direction. Clearly, we have

$$\|v\|_2 = (x^T s - 2\gamma_k \mu n + \gamma_k^2 \mu^2 \sum_{j=1}^{n} \frac{1}{x_j s_j})^{1/2}.$$

Since $(x, s) \in \mathcal{N}_\infty^-(\beta)$, we have

$$\sum_{j=1}^{n} \frac{\mu}{x_j s_j} \leq \frac{n}{1-\beta}.$$

This implies

$$\|v\|_2 \leq (1 - 2\gamma_k + \frac{\gamma_k^2}{1-\beta})^{1/2} (x^T s)^{1/2}. \tag{21}$$

For any $v \in \mathbb{R}^n$,

$$\|v_p\|_\infty \leq \|v_p\|_2 \leq \|v\|_2 \quad \text{and} \quad \|v - v_p\|_\infty \leq \|v - v_p\|_2 \leq \|v\|_2.$$

Hence, we have

$$|\min_j \{(v_p)_j (v - v_p)_j\}| \leq (1 - 2\gamma_k + \frac{\gamma_k^2}{1-\beta})(x^T s). \tag{22}$$

Taking the maximum allowable step size satisfying (20) defines Algorithm I. Here we will assume $\gamma_k = \bar{\gamma}$ to prove polynomiality, but at the end of this section it will be clear that once the quadratic convergence is obtained polynomiality does not get hurt. Now, define

$$C_o := \frac{\beta \bar{\gamma}}{(1 - 2\bar{\gamma} + \frac{\bar{\gamma}^2}{1-\beta})}.$$

By (19) and (20) we have that for all iterations the maximum allowable step size is at least

$$\min\{\frac{C_o}{n}, \frac{1}{1+\bar{\gamma}}\}.$$

So as in [MTY90], if we initially have $(x^o, s^o) \in \mathcal{N}_\infty^-(\beta)$ such that $(x^o)^T s^o \leq 2^t$, and if at each iteration we take a step of maximum allowable size, then, by (16), we get

$$(x^k)^T s^k \leq [1 - \frac{C_o}{n}(1 - \bar{\gamma})]^k \, 2^t.$$

Hence we conclude that in $O(nt)$ iterations we must have $(x, s) \in \mathcal{N}_\infty^-(\beta)$ such that $x^T s \leq 2^{-t}$. So, we proved:

Theorem 4.1. *If $(x^o)^T s^o \leq 2^t$, then in $O(nt)$ iterations Algorithm I will have a solution $(x, s) \in \mathcal{N}_\infty^-(\beta)$ such that $x^T s \leq 2^{-t}$.*

As long as the iterates lie in the neighborhood, we can also choose α_k such that ψ_q is minimized in the given search direction, while remaining in $\mathcal{N}_\infty^-(\beta)$. This way of choosing a step size yields what we will call Algorithm II. We have

Theorem 4.2. *Suppose $(x^o)^T s^o \leq 2^t$. Then for a large enough chosen q, the algorithm with line search on the potential function ψ_q (i.e. Algorithm II) will have a solution $(x, s) \in \mathcal{N}_\infty^-(\beta)$ such that $x^T s \leq 2^{-t}$ in $O(nt)$ iterations.*

Proof: We have

$$\psi_q(x^+, s^+) - \psi_q(x, s) = q \log(\frac{(x^+)^T s^+}{x^T s}) - \log(\frac{\min\{x_j^+ s_j^+\}}{(x^+)^T s^+/n}) + \log(\frac{\min\{x_j s_j\}}{x^T s/n})$$

Note that $\log(\frac{\min_j\{x_j s_j\}}{x^T s/n}) \leq 0$, and that $\log(\frac{\min_j\{x_j^+ s_j^+\}}{(x^+)^T s^+/n}) \geq \log(1 - \beta)$, which yields

$$\psi_q(x^+, s^+) - \psi_q(x, s) \leq q \log[1 - \alpha(1 - \gamma)] + \log(\frac{1}{1 - \beta}).$$

Now using the linear approximation for the log, we get:

$$\psi_q(x^+, s^+) - \psi_q(x, s) \leq -q\alpha(1 - \gamma) + \log(\frac{1}{1 - \beta}).$$

Since the maximum allowable step size (to stay in the neighborhood) is at least $\frac{C_o}{n}$, for a large enough q (e.g. one can choose $q \geq \frac{2n}{C_o(1-\bar{\gamma})} \log(\frac{1}{1-\beta})$ to guarantee a constant decrease in the potential function),

$$\psi_q(x^+, s^+) - \psi_q(x, s) \leq -\Omega(\frac{q}{n}).$$

By Proposition 3.1, $O(nt)$ iterations suffice. □

Remark: The same results also apply to the $\mathcal{N}_\infty$ neighborhoods.

Now, we turn to quadratic convergence. Our analysis is similar to the one given in Ye et al. [YGTZ91]. Here, we consider the search direction in a different scaling and that shortens some of the arguments slightly. Of course, they analyze only the affine scaling direction (for the "predictor" step) rather than a combination of the affine scaling and centering directions.

Let (x^*, s^*) be an optimal solution which is strictly complementary. i.e.

$$\{j : x_j^* > 0\} \cap \{j : s_j^* > 0\} = \emptyset$$

and

$$\{j : x_j^* > 0\} \cup \{j : s_j^* > 0\} = \{1, 2, \ldots, n\}$$

Note that such a solution always exists for linear programs (see Goldman and Tucker [GT56]) and since all the iterates lie in the neighborhood $\mathcal{N}_\infty^-(\beta)$ the algorithm will generate a sequence whose limit points are strictly complementary solutions (see Güler and Ye [GY91]). Without loss of generality, let $A = [B, N]$ be the corresponding partition of A; i.e., the components of x^* corresponding to the indices of the columns of B, forming the vector x_B^*, are strictly positive, and $x_N^* = 0$. Similarly we have $s_B^* = 0$, and $s_N^* > 0$.

Lemma 4.1. *Suppose (x^*, s^*) is non-degenerate. Then as $x^T s \to 0$, there exists a data dependent constant C such that*

(i) $\|(v_p)_N\|_2 \leq (x^T s)^{1/2}$

(ii) $\|(v - v_p)_B\|_2 \leq (x^T s)^{1/2}$

(iii) $\|(v_p)_B\|_2 \leq C(x^T s)^{3/2}$

(iv) $\|(v - v_p)_N\|_2 \leq C(x^T s)^{3/2}$.

Proof: Suppose $x^T s \leq n\bar{\gamma}/\rho$, so that $\gamma_k = \rho(x^T s/n)$. Then

(i) $\|(v_p)_N\|_2 \leq \|(v_p)\|_2 \leq \|v\|_2 \leq (x^T s)^{1/2}(1 - 2\rho\frac{(x^T s)}{n} + \frac{\rho^2(x^T s)^2}{n^2(1-\beta)})^{1/2}$.

The first two inequalities are clear, while the last one is given by (21). For $x^T s \leq 2n(1-\beta)/\rho$, we get $\|(v_p)_N\|_2 \leq (x^T s)^{1/2}$. Similarly, we have

(ii) $\|(v - v_p)_B\|_2 \leq \|(v - v_p)\|_2 \leq \|v\|_2 \leq (x^T s)^{1/2}$.

(iii) Note that v_p is the solution of the following least squares problem:

$$\min \{\tfrac{1}{2}\|v - v_p\|_2^2 : \bar{A}v_p = 0\}.$$

This implies that given $(v_p)_N$, $(v_p)_B$ is the solution to

$$\min \tfrac{1}{2}\|v_B - (v_p)_B\|_2^2$$

$$BD_B(v_p)_B = -ND_N(v_p)_N,$$

where $D := X^{1/2}S^{-1/2}$ and D_B, D_N give the corresponding partition of D. Note that $(x, s) \in \mathcal{N}_\infty^-(\beta)$ implies $\min\{x_j s_j\} \geq (1-\beta)\mu$, which implies

$$\max\{x_j s_j\} \leq [(n-1)\beta + 1]\mu.$$

Also note that by our assumption that $\mathcal{F}_o \neq \emptyset$, we have the optimal solution sets of both problems bounded. Since all iterates lie in the neighborhood $\mathcal{N}_\infty^-(\beta)$, using the result of Güler and Ye [GY91], we have that the limit points of the sequences generated by Algorithm I (or by Algorithm II) will be strictly complementary solutions. Now, $(1-\beta)\mu \leq x_j s_j \leq [(n-1)\beta + 1]\mu$. So, we have

$$\sqrt{(1-\beta)\mu} \leq x_j^{1/2} s_j^{1/2} \leq \sqrt{[(n-1)\beta + 1]\mu}. \tag{23}$$

This implies

$$\frac{\sqrt{(1-\beta)\mu}}{s_j} \leq x_j^{1/2} s_j^{-1/2} \leq \frac{\sqrt{[(n-1)\beta + 1]\mu}}{s_j}. \tag{24}$$

From (22) we also get

$$\frac{1}{\sqrt{[(n-1)\beta+1]\mu}} \leq x_j^{-1/2} s_j^{-1/2} \leq \frac{1}{\sqrt{(1-\beta)\mu}},$$

from which we deduce

$$\frac{x_j}{\sqrt{[(n-1)\beta+1]\mu}} \leq x_j^{1/2} s_j^{-1/2} \leq \frac{x_j}{\sqrt{(1-\beta)\mu}}, \tag{25}$$

Now using (24) for $j \in N$, and using (25) for $j \in B$, we conclude that there exist some data dependent constants C_B and C_N (see Güler and Ye [GY91]) such that for sufficiently large k

$(D_B)_j \in [\frac{1}{C_B}(x^T s)^{-1/2}, C_B(x^T s)^{-1/2}]$

and

$(D_N)_j \in [\frac{1}{C_N}(x^T s)^{1/2}, C_N(x^T s)^{1/2}]$.

Clearly, if the solution is non-degenerate then B is invertible and the solution to the least squares problem is given by

$$(v_p)_B = -D_B^{-1} B^{-1} N D_N (v_p)_N.$$

Assuming $(x^T s) \leq 2(1-\beta)n/\rho$, we get

$$\|(v_p)_B\|_2 \leq \|D_B^{-1}\|_2 \|B^{-1}N\|_2 \|D_N\|_2 \|(v_p)_N\|_2 \leq C(x^T s)^{3/2}.$$

(iv) Let $v_q := v - v_p$. Note that v_q is the solution of the following least squares problem:

$\min \{\frac{1}{2}\|v - v_q\|_2^2 : v_q = \bar{A}^T y \text{ for some } y \in \mathbb{R}^m\}$;

alternatively, $v_q = \bar{A}^T \bar{y}$ where $\bar{y}$ is the solution to $\min_y \frac{1}{2}\|v - \bar{A}^T y\|^2$. This implies that given $(v - v_p)_B$, $\bar{y}$ is the solution to

$\min \frac{1}{2}\|v_N - D_N N^T y\|_2^2$

$D_B B^T y = (v - v_p)_B \ (= (v_q)_B)$;

then to get $(v - v_p)_N$ we simply let $(v - v_p)_N := D_N N^T \bar{y}$.
Clearly, if the solution is non-degenerate then B^T is invertible and the solution to the least squares problem is given by

$$\bar{y} = B^{-T} D_B^{-1} (v - v_p)_B.$$

Hence

$$(v - v_p)_N = D_N N^T B^{-T} D_B^{-1} (v - v_p)_B.$$

Assuming $(x^T s) \leq 2(1 - \beta)n/\rho$, we get

$$\|(v - v_p)_N\|_2 \leq \|D_N\|_2 \|N^T B^{-T}\|_2 \|D_B^{-1}\|_2 \|(v - v_p)_B\|_2 \leq C(x^T s)^{3/2}.$$

□

Remark: Note that in the proof the non-degeneracy assumption implies the existence of the constants C_B and C_N, but we tried to avoid using the assumption at that point to keep the dependence of the proof on non-degeneracy assumption minimal (see section 5).

Now, showing the quadratic convergence of the algorithm is straight forward. In Algorihtm I we choose

$$\alpha_k = \min\{\frac{\beta\gamma_k\mu_k}{|\min_j\{(v_p)_j(v - v_p)_j\}|}, \frac{1}{1 + \gamma_k}\}.$$

Then by equation (16) we have

$$(x^{(k+1)})^T s^{(k+1)} = [1 - \alpha_k(1 - \gamma_k)](x^k)^T s^k \tag{26}$$

Note that by definition of γ_k, if $(x^k)^T s^k \leq \frac{\bar{\gamma} n}{\rho}$ then $\gamma_k = \rho \frac{(x^k)^T s^k}{n}$.
Also, if $(x^k)^T s^k \leq \frac{2(1-\beta)n}{\rho}$ we have

$$\begin{aligned} |\min_j\{(v_p)_j(v-v_p)_j\}| &\leq \max\{\|(v_p)_B\|_2\|(v-v_p)_B\|_2, \|(v_p)_N\|_2\|(v-v_p)_N\|_2\} \\ &\leq C(x^T s)^2. \end{aligned}$$

The first inequality is trivial, while the second follows from Lemma 4.1. Now we let $\rho = Cn^2/\beta$. Then the step size is

$$\alpha_k = \min\{\frac{\beta\gamma_k\mu_k}{|\min_j\{|(v_p)_j(v-v_p)_j\}|}, \frac{1}{1+\rho\mu_k}\} = \frac{1}{1+\rho\mu_k}.$$

So, using (26) we get

$$\begin{aligned} (x^{(k+1)})^T s^{(k+1)} &\leq [1 - \frac{1-\rho\mu_k}{1+\rho\mu_k}](x^k)^T s^k \\ &= \frac{2\rho}{n(1+\rho\mu_k)}[(x^k)^T s^k]^2 \\ &\leq \frac{2\rho}{n}[(x^k)^T s^k]^2 \\ &= \frac{2Cn}{\beta}[(x^k)^T s^k]^2. \end{aligned}$$

We proved:

Theorem 4.3. *Given $(x^o, s^o) \in \mathcal{N}_\infty^-(\beta)$ such that $(x^o)^T s^o \leq 2^t$, Algorithm I converges in $O(nt)$ iterations and assuming that (x^*, s^*) is non-degenerate, it asymptotically achieves 1-step Q-quadratic convergence.*

4.1 Implementation

Even though we proved asymptotic quadratic convergence, it is still unclear how one chooses C (and hence ρ) for a given instance of the problem. This clearly is related to the question of when one can see the quadratic convergence. If one is dealing with the integer LP model, i.e. all the data is integer and the size of the problem is L, then looking at the proof of Lemma 4.1. one can find a very rough over estimate of C a priori. Nevertheless, C would probably be $2^{O(L)}$. This is not a desirable answer. In this subsection we adjust C dynamically in a way that will preserve the polynomial time bound and eventually yield the quadratic convergence of the duality gap. We set

$$C(k) := \max\{C(k-1), \frac{|\min\{(v_p^{k-1})_j(v^{k-1} - v_p^{k-1})_j\}|}{[(x^{k-1})^T s^{k-1}]^2}\},$$

where $C(0) := \beta\bar{\gamma}/n(x^o)^T s^o$. So, $\rho_k = C(k)n^2/\beta$. We also change the definition of γ_k slightly:

$$\gamma_k := \begin{cases} \rho_k \frac{(x^k)^T s^k}{n}, & \text{if } \rho_k \frac{(x^k)^T s^k}{n} < \gamma_{k-1}; \\ \bar{\gamma}, & \text{otherwise.} \end{cases}$$

For simplicity we will assume $\bar{\gamma} \leq 1/4$ and we will set $\gamma_{-1} := \bar{\gamma}$. Note that the step size at the k^{th} iteration, α_k, is a function of γ_k which depends on ρ_k (and hence $C(k)$).

First we show that this way of updating γ_k and ρ_k preserves the polynomial time bound on the number of iterations. At iteration k, if $\gamma_k = \bar{\gamma}$ then we know that the duality gap decreases by at least a desired fraction. We will show that if $\gamma_k = \rho_k \frac{(x^k)^T s^k}{n}$ then at iteration $k-1$ we decrease the duality gap by a constant fraction, which is much more than what one needs for an $O(nt)$ bound on the number of iterations. Since $\gamma_k = \rho_k \frac{(x^k)^T s^k}{n}$, we must have $\rho_k \frac{(x^k)^T s^k}{n} < \gamma_{k-1}$. Using the definition of ρ_k we get

$$\frac{(x^k)^T s^k}{(x^{(k-1)})^T s^{(k-1)}} < \frac{\gamma_{k-1}\beta\mu_{k-1}}{|\min\{(v_p^{(k-1)})_j(v^{(k-1)} - v_p^{(k-1)})_j\}|}. \tag{27}$$

Now, there can be two cases: either $\alpha_{k-1} = \frac{1}{1+\gamma_{k-1}}$ or α_{k-1} is equal to the right hand side of (27). In the first case (using (16)), we have

$$\frac{(x^k)^T s^k}{(x^{(k-1)})^T s^{(k-1)}} = 1 - \frac{1-\gamma_{k-1}}{1+\gamma_{k-1}} = \frac{2\gamma_{k-1}}{1+\gamma_{k-1}}.$$

Since $\gamma_{k-1} \leq \bar{\gamma} \leq 1/4$, we have

$$\frac{(x^k)^T s^k}{(x^{(k-1)})^T s^{(k-1)}} \leq 1/2.$$

In the second case (27) implies

$$\frac{(x^k)^T s^k}{(x^{(k-1)})^T s^{(k-1)}} < \alpha_{k-1}.$$

Using equation (16) we have

$$1 - \alpha_{k-1}(1 - \gamma_{k-1}) < \alpha_{k-1},$$

from which we get $\alpha_{k-1} > \frac{1}{2-\gamma_{k-1}}$. So,

$$\frac{(x^k)^T s^k}{(x^{(k-1)})^T s^{(k-1)}} < 1 - \frac{1-\gamma_{k-1}}{2-\gamma_{k-1}} = \frac{1}{2-\gamma_{k-1}} \leq 4/7.$$

We showed that when we try to phase out the centering component, the previous iteration provides a very good reduction in the duality gap, hence the bound is preserved. Therefore, as long as $|\min\{(v_p)_j(v - v_p)_j\}|$ is the *right* order of the corresponding duality gap for a sequence of iterations, we will have a good estimate of the constant C. By Lemma 4.1, the sequence $C(k)$ is bounded above by a data dependent constant. Since the sequence is monotone non-decreasing, it converges. This shows that $\alpha_k \to 1$, however, to get quadratic convergence one still has to prove that $C(k-1)/C(k)$ goes to 1 at least as fast as $1 - O((x^k)^T s^k)$. We will use a cheap trick to avoid such analysis. After every r iterations we check if $C(k) = C(k-r)$. If the equality holds we do nothing; otherwise (if $C(k) > C(k-r)$), we let $C(k) := 2 * C(k)$ (note that this doubling does not affect the polynomial bound). Since we know that without the doubling of the $C(k)$s they converge to some number that is not greater than some data dependent constant C, we have that with the occasional doubling of the estimates of C the sequence $C(k)$ must eventually become constant. Therefore, there exists a large enough K such that for $k \geq K$ we have $\alpha_k = \frac{1}{1+\rho_k \mu_k}$ and quadratic convergence follows as in the arguments of Theorem 4.3.

5 Degeneracy

In case of degeneracy, having a centering component in the search direction may cause some trouble (note that Mehrotra [Meh91] and Ye et al. [YGTZ91] take the affine scaling direction (without any centering) as the search direction in the predictor step of Mizuno-Todd-Ye algorithm). Here, we will follow the proof of Lemma 4.1 without the non-degeneracy assumption and point out the difficulties arising from having a centering component. As we stated before the arguments given in Lemma 4.1 are similar to the arguments of Ye et al. [YGTZ91], but we deal with the solutions of the equations (7)-(9) rather than the solutions of (4)-(6). We will make our arguments in the primal space only, but the arguments for the dual space are essentially symmetric. We will start with an observation made by Ye et al. [YGTZ91] about the iterates s. They note that since $s^*_B = 0$ for all optimal $s^* \in F(D)$, we must have $c_B \in R(B^T)$ (range of B^T). But then for any feasible s, we have $s_B = c_B - B^T y$ which implies $s_B \in R(B^T)$. Hence there exists $\bar{y} \in \mathbb{R}^m$ such that $D_B B^T \bar{y} = (X^{1/2} S^{1/2} e)_B$.

So, if the search direction is the affine scaling direction; i.e., $v := (X^{1/2} S^{1/2} e)$ then we consider the folowing least squares problem

$$\text{(LS)} \quad \min \tfrac{1}{2} \|v_B - (v_p)_B\|_2^2$$

$$BD_B(v_p)_B = -ND_N(v_p)_N.$$

Karush-Kuhn-Tucker conditions for (LS) are given by

$$D_B B^T y + v_B - (v_p)_B = 0,$$

$$BD_B(v_p)_B = -ND_N(v_p)_N.$$

Since there exists a $\bar{y}$ such that $v_B = D_B B^T \bar{y}$, the necessary and sufficient conditions are equivalent to the following conditions:

$$D_B B^T \hat{y} + (v_p)_B = 0,$$

$$BD_B(v_p)_B = -ND_N(v_p)_N.$$

But these conditions are necessary and sufficient for the following least squares problem:

$$(LS') \quad \min \tfrac{1}{2}\|(v_p)_B\|_2^2$$

$$BD_B(v_p)_B = -ND_N(v_p)_N.$$

Now, suppose we are trying to find a feasible solution to the system of linear equalities $Br = -ND_N(v_p)_N$. Since we know that the system is feasible, it is clear that there exists a solution $\bar{r}$ and a data dependent constant C such that $\|\bar{r}\|_2 \leq C\|D_N\|_2\|(v_p)_N\|_2$. Then we let $w := D_B^{-1}\bar{r}$. It is clear that w is a feasible solution to (LS') and that

$$\|w\|_2 \leq \|D_B^{-1}\|_2\|\bar{r}\|_2 \leq C\|D_B^{-1}\|_2\|D_N\|_2\|(v_p)_N\|_2.$$

Using the same arguments as in Lemma 4.1 we have $\|w\|_2 \leq C(x^Ts)^{3/2}$. Since w is a feasible solution to (LS'), the optimal solution of (LS') (and hence (LS)) must have the same property; i.e., $\|(v_p)_B\|_2 \leq C(x^Ts)^{3/2}$.

Unfortunately, the same argument does not apply to the centering component, because it is clear that $(X^{-1/2}S^{-1/2}e)_B$ is not (in general) in the row space of BD_B. Since in the presence of degeneracy B does not necessarily have full column rank, in the same least squares problem (LS), one has to worry about the vectors which are close the the centering direction and are in the null space of BD_B. Since the centering direction does not lie in the row space of BD_B, the component of the centering direction that lies in the null space of BD_B becomes important in the analysis. If the norm of this component is large (it possibly could be as large as the norm of the centering direction), the analysis given here does not work. Even in this case we still have $\|(v_p)_B\| \leq O((x^Ts)^{3/2})$ (because the norm of the centering component is $O((x^Ts)^{3/2})$) but now the constant (hidden by "big Oh") is a multiple of ρ which does not allow us to pick ρ independently to ensure a large enough step size. From this argument we just gave, it seems that the centering component might make the norm of the projection larger (hence destroy the quadratic convergence property of the algorithm).

One might hope that degeneracy could keep the neighborhoods "larger" (in some sense) in some subspace so that the centering would not hurt the quadratic convergence property. This is one motivation for the Theorem 5.1 we will present. Another motivation is the fact that the recent quadratic convergence results of Mehrotra [Meh91] and Ye et al. [YGTZ91] do not rely on the assumption that the iterates are converging. So, one natural question is: "In case of degeneracy is it possible to have non-convergent iterates ?" Along these lines Roos [Ro92] and Todd [To92] provided examples in which the neighborhoods of the central path stay large as the central path converges to the analytic center of the optimal face. Theorem 5.1 provides a characterization of this behavior and proves that in case of degeneracy the neighborhoods $\mathcal{N}_2(\beta)$, $\mathcal{N}_\infty(\beta)$, and $\mathcal{N}_\infty^-(\beta)$ do not con-

verge to a point (the center of the optimal face) and stay "large" as the iterates approach the optimal face. So, for an algorithm to keep its iterates in one of these neighborhoods does not suffice to have convergent iterates.

We introduce some notation first. For each $\mu \geq 0$ let $(x_c(\mu), s_c(\mu))$ denote the solution of (1)-(3), i.e., $(x_c(\mu), s_c(\mu))$ is the central primal-dual solution corresponding to duality gap $n\mu$. Let $\{d^1, d^2, \ldots, d^r\}$ be an orthonormal basis for the null space of B. For each $i \in \{1, 2, \ldots, r\}$,

$$\eta_i(\beta, \mu) := \max\{\eta_i : x_c(\mu) \pm \eta_i \begin{pmatrix} d^i \\ 0 \end{pmatrix} \in \mathcal{N}_P(\beta, \mu)\}$$

where

$$\mathcal{N}_P(\beta, \mu) := \{x \in F(P) : \exists s, (x, s) \in \mathcal{N}_\infty^-(\beta) \text{ and } x^T s = n\mu\}.$$

x_c^* will denote the analytic center of the optimal face in the primal problem. Now, we can state the theorem.

Theorem 5.1. *Let* $\eta_i^* := \max\{\eta_i : x_c^* \pm \eta_i \begin{pmatrix} d^i \\ 0 \end{pmatrix} \in F(P)\} > 0$ *for all* $i \in \{1, 2, \ldots, r\}$. *Then there exist* $\bar{\mu}$ *and* $\epsilon \in (0, 1)$ *such that* $\mu \leq \bar{\mu}$ *implies* $\eta_i(\beta, \mu) \geq (1 - \epsilon)\eta_i^*$.

Proof: Since $x_c(\mu) \to x_c^*$, by the continuity of the central path, there exists $\bar{\mu}$ and $\epsilon_1 \in (0, 1)$ such that
$x_c(\mu) + \eta_i \begin{pmatrix} d^i \\ 0 \end{pmatrix} \in F(P)$ for $\mu \leq \bar{\mu}_1$ and $\eta_i \leq (1 - \epsilon_1)\eta_i^*$.

As we mentioned at the beginning of this section, for all $s \in F(D)$ we have $s_B \in R(B^T)$. Therefore,

$$\left[x_c(\mu) + \begin{pmatrix} d^i \\ 0 \end{pmatrix}\right]^T s = x_c(\mu)^T s, \forall s \in F(D).$$

i.e. $x^T s$ is constant for all feasible s as long as we move x in the null space of B.

Note that by definition we have $(x_c(\mu))_j (s_c(\mu))_j = \mu, \forall j$. So, if

$$\eta_i \leq \min_{d_j^i < 0} \{\frac{\beta\mu}{-d_j^i((s_c(\mu))_B)_j}\}, \quad \forall i \in \{1, 2, \ldots, r\} \tag{28}$$

then we get

$$((x_c(\mu))_B)_j((s_c(\mu))_B)_j + \eta_i d_j^i((s_c(\mu))_B)_j \geq (1-\beta)\mu$$

It follows from the fact that $x_c(\mu) \rightarrow x_c^*$ that there exist a $\bar{\mu}_2$ and a data dependent constant C_B such that if $\mu \leq \bar{\mu}_2$ then

$((x_c(\mu))_B)_j \in [\frac{1}{C_B}, C_B], \forall j.$

So, if $\mu \leq \bar{\mu}_2$ then

$((s_c(\mu))_B)_j \in [\frac{1}{C_B}\mu, C_B\mu].$

Therefore, the RHS of (28) is bounded away from zero by a data dependent constant (depending on the size of the optimal face). So, if $\mu \leq \min\{\bar{\mu}_1, \bar{\mu}_2\}$ then

$\eta_i(\beta, \mu) \geq \min\{\frac{\beta}{\eta_i^* C_B}, (1-\epsilon_1)\}\eta_i^*.$ □

Remark: Theorem 5.1 is also applicable to $\mathcal{N}_\infty$ and $\mathcal{N}_2$ neighborhoods. The proof for the $\mathcal{N}_\infty$ neighborhoods is essentially the same. The proof for the $\mathcal{N}_2$ neighborhoods follows from the fact that the theorem holds for $\mathcal{N}_\infty$ neighborhoods.

6 Conclusion

We introduced a new primal-dual potential function that fits in nicely with wide neighborhoods. The new potential function has some combinatorial flavor because of its barrier part. This might make it easier to get a candidate active set for determining the optimal solution. We also studied the asymptotic behavior of the primal-dual interior point algorithms in wide neighborhoods (proving quadratic convergence under the assumption of non-degeneracy) and showed how the centering component might make the norm of the search direction larger than desired in case of degeneracy. We showed that an algorithm keeping its iterates in a neighborhood of the central path does not necessarily have convergent iterates (in case of degeneracy).

In the light of Theorem 5.1, it is conceivable that perhaps the neighborhoods stay large in the nullspace of BD_B (as the iterates approach the optimal set) so that one does not have worry so much about the norm of the component of the search directions that is in the null space of BD_B to make sure that the

next iterate also lies in the same neighborhood. Possibly the analysis could then be restricted to the range of $D_B B^T$ and hence, in the degenerate case, even though the centering component is large in norm it may not hurt the quadratic convergence property.

Acknowledgment: The author wishes to thank Mike Todd for very valuable discussions and for suggestions to improve the presentation of this paper.

References:

[BL89] D.A. Bayer and J.C. Lagarias, The Nonlinear Geometry of Linear Programming: I.Affine and Projective Scaling Trajectories, II.Legendre Transform Coordinates and Central Trajectories, *Transactions of the American Mathematical Society 314*, pp. 499-581.

[GT56] A.J. Goldman and A.W. Tucker, Theory of Linear Programming, in: *Linear Inequalities and Related Systems*, H.W. Kuhn and A.W. Tucker, eds., Princeton University Press, Princeton, NJ, pp. 53-97.

[Go90] C.C. Gonzaga, Polynomial Affine Algorithms for Linear Programming, *Mathematical Programming* 49, pp. 7-21.

[GY91] 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.

[KMY88] M. Kojima, S. Mizuno, and A. Yoshise, A Primal-Dual Interior Point Algorithm for Linear Programming, in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods*, Springer-Verlag, New York, pp. 29-47.

[KMY89] M. Kojima, S. Mizuno, and A. Yoshise, A Polynomial Time Algorithm for a Class of Linear Complementarity Problems, *Mathematical Programming* 44, pp. 1-26.

[Mc92] K. McShane, *Primal-Dual Interior Point Algorithms for Linear Programming and the Linear Complementarity Problem*, Ph.D. Thesis, Cornell University, Ithaca, NY.

[Meg88] N. Megiddo, Pathways to the Optimal Set in Linear Programming, in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods*, Springer-Verlag, New York, pp.131-158.

[Meh91] S. Mehrotra, Quadratic Convergence in a Primal-Dual Method, *TR 91-15*, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.

[MTY90] S. Mizuno, M.J. Todd, and Y. Ye, On Adaptive-Step Primal-Dual Interior-Point Algorithms for Linear Programming, *TR-944 (to appear in Math. of OR*, School of OR and IE, Cornell University, Ithaca, NY.

[MA89] R.C. Monteiro and I. Adler, Interior Path Following Primal-Dual Algorithms. Part I: Linear Programming, *Mathematical Programming* 44, pp. 27-42.

[Po90] M.J.D. Powell, Karmarkar's Algorithm: A View from Nonlinear Programming, *IMA Bulletin*, Vol. 26, pp. 165-181.

[Ro92] K. Roos , *private communication.*

[Ta87] K. Tanabe, Centered Newton Method for Mathematical Programming, *in System Modelling and Optimization* (eds. M. Iri and K. Yajima), Springer-Verlag, The Proceedings of the 13th IFIP Conference, Tokyo,Aug. 31 - Sept. 4, pp. 197-206.

[To92] M.J. Todd , *private communication.*

[TY90] M.J. Todd and Y. Ye, A Centered Projective Algorithm for Linear Programming, *Mathematics of Operations Research* 15, pp. 508-529.

[YGTZ91] Y. Ye, O. Güler, R.A. Tapia, and Y. Zhang, A Quadratically Convergent $O(\sqrt{n}L)$-Iteration Algorithm for Linear Programming, *TR91-26*, Department of Mathematical Sciences, Rice University, Houston, TX.

[ZT90] 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.

[ZTD90] 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.