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

Technical Report No. 1018

August 1992

**CONSTANT POTENTIAL
PRIMAL-DUAL ALGORITHMS:
A FRAMEWORK**

*Levent TUNÇEL*¹

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

CONSTANT POTENTIAL PRIMAL-DUAL ALGORITHMS: A FRAMEWORK

*Levent TUNÇEL**

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

August 1992

Abstract

We study the primal-dual affine scaling algorithms for linear programs. Using an idea of Mizuno and Nagasawa and new potential functions we establish a framework for primal-dual algorithms that keep a potential function fixed. We show that if the neighborhood and the potential function used in the algorithm correspond to each other in a way we define then the convergence proofs simplify greatly. Our algorithms have the property that all the iterates can be kept in a neighborhood of the central path without using any centering in the search directions.

Keywords: Linear programming, interior point methods, primal-dual algorithms, potential function.

*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

In this paper, we study the primal-dual affine scaling algorithms for linear programs (see Monteiro, Adler and Resende [MAR90], Kojima, Megiddo, Noma and Yoshise [KMNY91]). We follow up on the recent algorithm by Mizuno and Nagasawa [MN92]. They suggest starting from an interior primal-dual solution that lies in a one-sided infinity neighborhood of the central path and determine the next iterate based on the Tanabe-Todd-Ye potential function (Tanabe [Ta87], Todd and Ye [TY90])

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

where ρ is a positive constant. Given the affine scaling direction at the current iterate, Mizuno and Nagasawa propose choosing the next iterate such that the value of this potential function does not increase (rather than finding the point which minimizes the potential function along the given search direction). Here we use the same idea with a potential function proposed by Tunçel [Tu92] for $\rho > 0$:

$$\psi(x, s; \rho) := (\rho + 1) \log\left(\frac{x^T s}{n}\right) - \log(\min_j \{x_j s_j\}).$$

We show that one can achieve the same complexity bounds achieved by Mizuno and Nagasawa [MN92] by keeping the potential function ψ constant. Since the potential function ψ has very nice properties related to the infinity-norm neighborhoods of the central path (see Tunçel [Tu92]), using the contours of ψ we can find a simple expression relating the decrease in the duality gap to the distance of the next iterate from the boundary of the feasible region. As a result our proofs are shorter and simpler. We then define compatibility of potential functions with neighborhoods and generalize the results we had for ψ and one-sided infinity neighborhoods. Ultimately, we specialize on the two-norm neighborhoods of the central path. Proposing a compatible potential function we show that the corresponding primal-dual affine scaling algorithm has the same worst case iteration complexity as the ones given in [MN92], but our algorithm is guaranteed to converge to a strictly complementary solution. We also show that in our framework one can easily guarantee that all the iterates lie in a neighborhood of the central path (again a property not shown for existing polynomial time affine scaling algorithms).

2 Primal-Dual Affine Scaling Algorithm

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

$$\begin{aligned} (P) \quad & \text{minimize} \quad c^T x \\ & Ax = b, \\ & x \geq 0, \\ (D) \quad & \text{maximize} \quad b^T y \\ & A^T y + s = c, \\ & s \geq 0, \end{aligned}$$

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 the iterate numbers it will be clear from the context what the iterate number is.

The central path is given by the set of solutions to the following system of equalities for $\mu > 0$ (see for instance Megiddo [Meg88]):

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

Given $\pi \in (0, 1)$ a one-sided infinity-norm neighborhood of the central path is defined (see for instance Mizuno, Todd and Ye [MTY90]) as:

$$\mathcal{N}_\infty^-(\pi) := \{(x, s) \in \mathcal{F}_o : \|Xs - \mu e\|_\infty^- \leq (1 - \pi)\mu\}.$$

Here, for $u \in \mathbb{R}^n$, $\|u\|_\infty^- := -\min\{0, \min\{u_j\}\}$. Note that $\mathcal{N}_\infty^-(\pi)$ is a level set of $\psi(x, s; 0)$:

$$\mathcal{N}_\infty^-(\pi) = \{(x, s) \in \mathcal{F}_o : \psi(x, s; 0) \leq \log(\frac{1}{\pi})\}.$$

Suppose we have an initial interior point solution $(x, s) \in \mathcal{F}_o$. Then the affine scaling direction $(\bar{d}x, \bar{d}s)$ can be generated by solving the following set of equalities (see Kojima, Mizuno and Yoshise [KMY88]):

$$A\bar{d}x = 0 \tag{4}$$

$$A^T\bar{d}y + \bar{d}s = 0 \tag{5}$$

$$S\bar{d}x + X\bar{d}s = -Xs \tag{6}$$

which is equivalent to solving

$$\bar{A}dx = 0 \tag{7}$$

$$\bar{A}^T dy + ds = 0 \tag{8}$$

$$dx + ds = -X^{1/2}S^{1/2}e \tag{9}$$

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

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

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, we describe a primal-dual affine scaling algorithm based on a potential function ψ and a neighborhood $\mathcal{N}$ of the central path:

Algorithm:

Given $(x^o, s^o) \in \mathcal{N}(\pi_0)$ with $(x^o)^T s^o \leq 2^t$, set $k := 0$.

While $((x^k)^T s^k > 2^{-t})$ do

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

compute $(\bar{d}x, \bar{d}s)$ from (4)-(6)

choose step size $\alpha_k \in (0, 1)$ such that $\psi(x + \alpha_k \bar{d}x, s + \alpha_k \bar{d}s; \rho) = \psi(x, s; \rho)$

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

$$k := k + 1$$

end

3 Analysis and Convergence Results for Wide Neighborhoods

We define

$$\begin{aligned} \mu_k &:= \frac{(x^k)^T s^k}{n}, \\ \pi_k &:= \frac{\min\{x_j^k s_j^k\}}{\mu_k}. \end{aligned}$$

We assume that $\pi_0 \in (0, 1)$ is a constant independent of n and t . Let (x, s) be the current iterate with $(x, s) \in \mathcal{N}(\pi)$, and let $(x(\alpha), s(\alpha))$ denote the next iterate for a given step size $\alpha \in [0, 1]$, so that

$$x(\alpha) = x - \alpha X^{1/2} S^{-1/2} P_{\bar{A}} v \tag{10}$$

$$s(\alpha) = s - \alpha X^{-1/2} S^{1/2} (I - P_{\bar{A}}) v, \tag{11}$$

where $v := X^{1/2}S^{1/2}e$. We will denote $v_p := P_{\bar{A}}v$ and $v_q := v - v_p$. From (10)-(11) we obtain

$$(x(\alpha))^T s(\alpha) = x^T s - \alpha e^T X^{1/2} S^{1/2} (v_p + v_q) + \alpha^2 v_p^T v_q \quad (12)$$

$$= x^T s - \alpha x^T s \quad (13)$$

$$= (1 - \alpha) x^T s. \quad (14)$$

We also have

$$x(\alpha)_j s(\alpha)_j = x_j s_j - \alpha \sqrt{x_j s_j} [(v_p)_j + (v_q)_j] + \alpha^2 (v_p)_j (v_q)_j \quad (15)$$

$$= x_j s_j - \alpha (x_j s_j) + \alpha^2 (v_p)_j (v_q)_j \quad (16)$$

$$= (1 - \alpha) x_j s_j + \alpha^2 (v_p)_j (v_q)_j \quad (17)$$

We define

$$\mu(\alpha) := \frac{x(\alpha)^T s(\alpha)}{n},$$

$$\pi(\alpha) := \frac{\min\{x_j(\alpha)s_j(\alpha)\}}{\mu(\alpha)}.$$

Then from (14) and (17) we have

$$\pi(\alpha) = \min\left\{\frac{x_j s_j}{\mu} + \frac{\alpha^2}{1 - \alpha} \frac{(v_p)_j (v_q)_j}{\mu}\right\} \quad (18)$$

Proposition 3.1. *Let (x, s) , $(x(\alpha), s(\alpha))$, π , $\pi(\alpha)$ be defined as above. Then*

$$\psi(x(\alpha), s(\alpha); \rho) = \psi(x, s; \rho) \iff (1 - \alpha)^\rho = \frac{\pi(\alpha)}{\pi} \quad (19)$$

Proof:

$$\begin{aligned} & \psi(x(\alpha), s(\alpha); \rho) = \psi(x, s; \rho) \\ \iff & \rho \log(\mu(\alpha)) - \log\left(\frac{\min\{x_j(\alpha)s_j(\alpha)\}}{\mu(\alpha)}\right) = \rho \log(\mu) - \log\left(\frac{\min\{x_j s_j\}}{\mu}\right) \end{aligned}$$

$$\begin{aligned}
&\iff \rho \log((1 - \alpha)\mu) - \log(\pi(\alpha)) = \rho \log(\mu) - \log(\pi) \\
&\iff \rho \log(1 - \alpha) = \log\left(\frac{\pi(\alpha)}{\pi}\right) \\
&\iff (1 - \alpha)^\rho = \frac{\pi(\alpha)}{\pi}.
\end{aligned}$$

□

Corollary 3.1. *After k iterations of the affine scaling algorithm described in section 2, we have,*

$$(x^k)^T s^k = \left(\frac{\pi_k}{\pi_o}\right)^{1/\rho} (x^o)^T s^o. \quad (20)$$

Proof: Follows from (14) and (19). □

Corollary 3.1 provides a very compact description of the contours of the potential function ψ and it relates the distance between the current iterate and the boundary of the feasible region to the duality gap directly. If π_k is very small relative to π_o (i.e. (x_k, s_k) is close to the boundary of the feasible region) then the duality gap must also be relatively small.

Lemma 3.1.

$$\frac{\pi(\alpha)}{\pi} \geq 1 - \left(\frac{\alpha^2}{1 - \alpha}\right) \frac{n}{\pi}.$$

Proof: Using (18) we get

$$\pi(\alpha) \geq \frac{\min\{x_j s_j\}}{\mu} - \left(\frac{\alpha^2}{1 - \alpha}\right) \frac{|\min\{(v_p)_j (v_q)_j\}|}{\mu} \quad (21)$$

$$= \pi - \left(\frac{\alpha^2}{1 - \alpha}\right) \frac{|\min\{(v_p)_j (v_q)_j\}|}{\mu} \quad (22)$$

Note that $\|v\|_2 = (x^T s)^{1/2}$, and for any $v \in \mathbb{R}^n$ we have $\|v_p\|_\infty \leq \|v_p\|_2 \leq \|v\|_2$ and $\|v_q\|_\infty \leq \|v_q\|_2 \leq \|v\|_2$. So, $|\min\{(v_p)_j (v_q)_j\}| \leq x^T s$. Hence, from (21)-(22) we get

$$\frac{\pi(\alpha)}{\pi} \geq 1 - \left(\frac{\alpha^2}{1-\alpha} \right) \frac{n}{\pi}.$$

□

Now, we settle a technical point showing that the contours of the potential function are well-defined.

Lemma 3.2. *Unless $(x(\alpha), s(\alpha))$ is feasible for $\alpha = 1$, (19) has exactly one root for $\alpha \in (0, 1)$.*

Proof: We define

$$\begin{aligned} g(\alpha) &:= (1 - \alpha)^{\rho+1}, \\ h_j(\alpha) &:= \frac{x_j s_j}{\pi \mu} (1 - \alpha) + \alpha^2 \frac{(v_p)_j (v_q)_j}{\pi \mu}, \\ f(\alpha) &:= g(\alpha) - \min \{h_j(\alpha)\}. \end{aligned}$$

Using (18) we have

$$(1 - \alpha)^\rho = \frac{\pi(\alpha)}{\pi} \iff f(\alpha) = 0.$$

We partition the indices as follows: $J^- := \{j : (v_p)_j (v_q)_j < 0\}$, $J^0 := \{j : (v_p)_j (v_q)_j = 0\}$, $J^+ := \{j : (v_p)_j (v_q)_j > 0\}$. Note that J^- and J^+ are non-empty, because if they both are empty then $(x(\alpha), s(\alpha))$ is feasible for $\alpha = 1$ and we have the optimal solution. Since $\sum_j (v_p)_j (v_q)_j = 0$, if one of them is non-empty then they both are.

Clearly, there exists a small enough $\epsilon > 0$ such that $f(\alpha) < 0$ for $\alpha \in (0, \epsilon)$ (e.g. $\epsilon = \frac{\pi \rho}{2n}$). It is also clear that if $(x(\alpha), s(\alpha))$ is not feasible for $\alpha = 1$, then $f(1) > 0$. By the continuity of f we conclude that there exists at least one root in $(0, 1)$. For $j \in J^0 \cup J^+$ we have

$$h_j(\alpha) \geq \frac{x_j s_j}{\pi \mu} (1 - \alpha) \geq (1 - \alpha) > (1 - \alpha)^{\rho+1}.$$

Hence, for $j \in J^0 \cup J^+$, $h_j(\alpha)$ does not intersect $g(\alpha)$ in $(0, 1)$. For $j \in J^-$ we have h_j as a concave function, so $\min_{j \in J^-} \{h_j(\alpha)\}$ is concave. Since $g(\alpha)$ is convex, the solution to (19) is unique. □

Lemma 3.3. *If α^* solves (19) for $\frac{1}{2t} \leq \rho \leq \sqrt{n}$ ($n \geq 4$) then $\alpha^* \geq \frac{\pi\rho}{2n}$.*

Proof: We want $\rho \log(1 - \alpha) = \log(\frac{\pi(\alpha)}{\pi})$. Note that if for $\bar{\alpha} \in (0, 1)$ we have

$$\rho \log(1 - \bar{\alpha}) \leq \log\left(\frac{\pi(\bar{\alpha})}{\pi}\right) \quad (23)$$

then $\alpha^* \geq \bar{\alpha}$ (from Lemma 3.2). We will show that $\bar{\alpha} = \frac{\pi\rho}{2n}$ satisfies (23) which will prove the lemma.

Using the linear approximation to the logarithm we have

$$\rho \log(1 - \bar{\alpha}) \leq -\rho\bar{\alpha} \quad \text{for } \bar{\alpha} \in (0, 1) \quad (24)$$

By Lemma 3.1 we have

$$\log\left(\frac{\pi(\bar{\alpha})}{\pi}\right) \geq \log\left[1 - \left(\frac{\bar{\alpha}^2}{1 - \bar{\alpha}}\right)\left(\frac{n}{\pi}\right)\right].$$

Note that for $\alpha = \frac{\pi\rho}{2n}$ and $\frac{1}{2t} \leq \rho \leq \sqrt{n}$, from which

$$\left(\frac{\bar{\alpha}^2}{1 - \bar{\alpha}}\right)\frac{n}{\pi} = \frac{\pi\rho^2/4n}{1 - \pi\rho/2n} \leq \frac{\pi\rho^2}{2n} \leq \frac{1}{2}.$$

Using the approximation $\log(1 + \lambda) \geq \lambda - \frac{\lambda^2}{2(1-|\lambda|)}$ for $|\lambda| < 1$, we get

$$\log\left(\frac{\pi(\bar{\alpha})}{\pi}\right) \geq -\left(\frac{1}{1 - \bar{\alpha}}\right)\left[\bar{\alpha}^2\frac{n}{\pi} + \frac{\bar{\alpha}^4 n^2 / \pi^2}{2(1 - \bar{\alpha} - \bar{\alpha}^2 n / \pi)}\right]. \quad (25)$$

From (24) and (25) we conclude that if

$$-\rho\bar{\alpha} + \rho\bar{\alpha}^2 \leq -\bar{\alpha}^2\frac{n}{\pi} - \frac{\bar{\alpha}^4 n^2 / \pi^2}{2(1 - \bar{\alpha} - \bar{\alpha}^2 n / \pi)} \quad (26)$$

then (23) holds. Letting $\bar{\alpha} = \frac{\pi\rho}{2n}$ and dividing both sides by $\rho\bar{\alpha}$ in (26) we get

$$-1 + \frac{\pi\rho}{2n} \leq -\frac{1}{2} - \frac{\pi\rho^2/8n}{2(1 - \pi\rho/2n - \pi\rho^2/4n)} \quad (27)$$

For given values of ρ and $n \geq 4$, the right hand side of (27) is at least $-\frac{1}{2} - \frac{\pi\rho^2}{8n} \geq -5/8$. The left hand side of (27) is at most $-1 + 1/4 = -6/8$. Therefore, (23) holds. $\square$

Now, we can state a theorem about the complexity of the algorithm.

Theorem 3.1. *The primal-dual affine scaling algorithm as stated in section 2 converges in $O(\frac{n2^{2t\rho t}}{\rho})$ iterations.*

Proof: From Corollary 3.1 we have $(x^k)^T s^k = \left(\frac{\pi_k}{\pi_o}\right)^{1/\rho} (x^o)^T s^o$. Since $(x^o)^T s^o \leq 2^t$, if at the K^{th} iteration $\frac{\pi_K}{\pi_o} < 2^{-2t\rho}$ then we must have $(x^K)^T s^K < 2^{-t}$ and the algorithm will stop. Otherwise, we get a uniform bound on the distances of the iterates from the boundary:

$$\pi_k \geq \pi_o 2^{-2t\rho}, \text{ for all } k \leq K.$$

By lemma 3.3, this uniform bound yields

$$\alpha^* > \frac{\pi\rho}{2n} \geq \frac{\pi_o\rho}{n2^{2t\rho+1}}. \quad (28)$$

From equality (14) and inequality (28) we conclude that we must have $(x^k)^T s^k < 2^{-t}$ in $O(\frac{n2^{2t\rho t}}{\rho})$ iterations. $\square$

Using the above theorem, for $\rho = 1/2t$, $\rho = 1$, and $\rho = \sqrt{n}$ we get the bounds $O(nt^2)$, $O(n2^{2t}t)$ and $O(\sqrt{n}2^{2t\sqrt{n}}t)$ respectively.

Remark 3.1. When $\rho = 1/2t$, we have $\pi_k \geq \pi_o/2$ for all iterates except the last one.

Remark 3.2. Suppose $(x, s) \in \mathcal{N}_{\infty}^-(\pi_o)$ is given. Then for some $\beta \in (0, \pi_o)$ (independent of the input) one can define

$$\pi := \frac{\min\{x_j s_j\}}{\mu} - \beta.$$

It is not hard to check that all the previous results hold with the new definition of π . With this definition, all the iterates have to lie in the neighborhood $\mathcal{N}_{\infty}^-(\beta)$. So, by a result of Güler and Ye [GY91] all the limit points of such a sequence of iterates are strictly complementary solutions. More importantly, one can attach Ye's scheme of making neighborhoods larger [Ye91] to the affine scaling algorithm provided here to obtain quadratic convergence. Hence creating a pure affine scaling algorithm that is polynomial and quadratically convergent.

4 A More General Framework : Compatibility

In this section we generalize some of the results from the previous section to certain pairs of neighborhoods and potential functions.

Given a constant $\pi > 0$ and a function $\delta(x, s)$ (which measures the centrality of (x, s)) a neighborhood of the central path will be defined as

$$\mathcal{N}(\delta, \pi) := \{(x, s) \in \mathcal{F}_o : \delta(x, s) \geq \pi\}.$$

Examples:

Ex.1.

$$\mathcal{N}_\infty^-(\pi) = \left\{ (x, s) \in \mathcal{F}_o : \left\| \frac{Xs}{\mu} - e \right\|_\infty^- \leq (1 - \pi) \right\}, \text{ so } \delta(x, s) = \frac{\min\{x_j s_j\}}{\mu}.$$

Ex.2.

$$\begin{aligned} \mathcal{N}_\infty(\pi) &= \left\{ (x, s) \in \mathcal{F}_o : \left\| \frac{Xs}{\mu} - e \right\|_\infty \leq (1 - \pi) \right\}, \text{ so} \\ \delta(x, s) &= \min \left\{ \frac{\min\{x_j s_j\}}{\mu}, 2 - \frac{\max\{x_j s_j\}}{\mu} \right\}. \end{aligned}$$

Ex.3.

$$\mathcal{N}_2(\pi) = \left\{ (x, s) \in \mathcal{F}_o : \left\| \frac{Xs}{\mu} - e \right\|_2 \leq (1 - \pi) \right\}, \text{ so } \delta(x, s) = 1 - \left\| \frac{Xs}{\mu} - e \right\|_2.$$

Ex.4. [To92]

For $u \in \mathbb{R}^n$, $u > 0$, define

$$\|u\|_R := \sqrt[n]{\prod_j u_j}.$$

Then

$$\mathcal{N}_R(\pi) = \left\{ (x, s) \in \mathcal{F}_o : \left\| \frac{Xs}{\mu} \right\|_R \geq \pi \right\}, \text{ so } \delta(x, s) = \left\| \frac{Xs}{\mu} \right\|_R.$$

□

Given $\mathcal{N}(\delta, \cdot)$ a primal-dual potential function that is *compatible* with $\mathcal{N}(\delta, \cdot)$ will be defined as

$$\psi(x, s; \rho) := \rho \log \left(\frac{x^T s}{n} \right) - \log(\delta(x, s)).$$

Note that ψ and $\mathcal{N}$ are compatible in a sense that

$$\mathcal{N}(\delta, \pi) = \left\{ (x, s) \in \mathcal{F}_o : \psi(x, s; 0) \leq \log \left(\frac{1}{\pi} \right) \right\}.$$

i.e. $\mathcal{N}(\delta, \pi)$ defines a level set of $\psi(x, s; 0)$.

Throughout this section we will use the following definitions:

$$\pi(\alpha) := \delta(x(\alpha), s(\alpha)),$$

$$\pi := \pi(0),$$

$$\pi_k := \delta(x^k, s^k).$$

In this setting Proposition 3.1 and Corollary 3.1 still hold. We easily have

Proposition 4.1. *Let ψ and $\mathcal{N}(\delta, \cdot)$ be compatible. Then*

$$\psi(x(\alpha), s(\alpha); \rho) = \psi(x, s; \rho) \iff (1 - \alpha)^\rho = \frac{\pi(\alpha)}{\pi} \tag{29}$$

Proof:

$$\begin{aligned} & \psi(x(\alpha), s(\alpha); \rho) = \psi(x, s; \rho) \\ \iff & \rho \log(\mu(\alpha)) - \log(\delta(x(\alpha), s(\alpha))) = \rho \log(\mu) - \log(\delta(x, s)) \\ \iff & \rho \log((1 - \alpha)\mu) - \log(\pi(\alpha)) = \rho \log(\mu) - \log(\pi) \\ \iff & \rho \log(1 - \alpha) = \log \left(\frac{\pi(\alpha)}{\pi} \right) \\ \iff & (1 - \alpha)^\rho = \frac{\pi(\alpha)}{\pi}. \end{aligned}$$

□

Corollary 4.1. *After k iterations of the affine scaling algorithm described in section 2, we have,*

$$(x^k)^T s^k = \left(\frac{\pi_k}{\pi_o} \right)^{1/\rho} (x^o)^T s^o. \quad (30)$$

Proof: Follows from (14) and (29). □

We can also generalize Lemma 3.3 with slight changes.

Lemma 4.1. *Suppose there exists a data dependent constant $M \geq 4$ such that*

$$\frac{\pi(\alpha)}{\pi} \geq 1 - \frac{\alpha^2}{1 - \alpha} \frac{M}{\pi}.$$

Also assume that there exists an $\epsilon > 0$ such that for $\alpha \in (1 - \epsilon, 1)$ we have

$$(1 - \alpha)^{\rho+1} > \frac{\delta(x(\alpha), s(\alpha))\mu(\alpha)}{\delta(x, s)\mu}.$$

Then there exists an $\alpha^ \in (0, 1)$ that solves (29) for $\frac{1}{2t} \leq \rho \leq \sqrt{M}$ such that $\alpha^* \geq \frac{\pi\rho}{2M}$.*

Proof: Because of our assumption that there exists an $\epsilon > 0$ such that for $\alpha \in (1 - \epsilon, 1)$ we have

$$(1 - \alpha)^{\rho+1} > \frac{\delta(x(\alpha), s(\alpha))\mu(\alpha)}{\delta(x, s)\mu},$$

and the fact that the functions in (19) are continuous, it suffices to show that for $\bar{\alpha} = \frac{\pi\rho}{2M}$ one has

$$(1 - \bar{\alpha})^\rho < \frac{\pi(\bar{\alpha})}{\pi}.$$

The proof of this fact can be obtained from the proof of Lemma 3.3 by simply replacing n by M in all relations. □

Note that in this general setting we don't provide an equivalent of Lemma 3.2; so, the equation (29) might have multiple solutions in $(\frac{\pi\rho}{2M}, 1)$. From a theoretical point of view it does not matter which of these solutions one uses as long as one finds a solution that lies in the interval $(\frac{\pi\rho}{2M}, 1)$ the following theorem holds. Now we can generalize Theorem 3.1:

Theorem 4.1. *Let ψ and $\mathcal{N}$ be compatible. Assume that the assumptions of Lemma 4.1 hold. Then the primal-dual affine scaling algorithm as stated in section 2 converges in $O(\frac{M 2^{2t\rho} t}{\rho})$ iterations.*

Proof: The proof is identical to the proof of Theorem 3.1 after replacing n by M . $\square$

Using the above theorem, for $\rho = 1/2t$, $\rho = 1$, and $\rho = \sqrt{M}$ we get the bounds $O(M t^2)$, $O(M 2^{2t} t)$ and $O(\sqrt{M} 2^{2t\sqrt{M}} t)$ respectively.

5 Analysis and Convergence Results for Two-Norm Neighborhoods

Now we specialize on the two-norm neighborhoods of the central path. We propose the following potential function that is compatible with $\mathcal{N}_2$ neighborhoods:

$$\psi_2(x, s; \rho) := \rho \log \left(\frac{x^T s}{n} \right) - \log \left(1 - \left\| \frac{Xs}{\mu} - e \right\|_2 \right).$$

So,

$$\delta(x, s) := 1 - \left\| \frac{Xs}{\mu} - e \right\|_2.$$

We define

$$\begin{aligned} \pi(\alpha) &:= 1 - \left\| \frac{X(\alpha)s(\alpha)}{\mu(\alpha)} - e \right\|_2, \\ \pi &:= \pi(0). \end{aligned}$$

Lemma 5.1.

$$\frac{\pi(\alpha)}{\pi} \geq 1 - \left(\frac{\alpha^2}{1 - \alpha} \right) \frac{n}{\pi}.$$

Proof: We have

$$\begin{aligned} \pi(\alpha) &= 1 - \left\| \frac{X(\alpha)s(\alpha)}{\mu(\alpha)} - e \right\|_2 \\ &= 1 - \left\| \frac{(1 - \alpha)Xs + \alpha^2 V_p(v_q)}{(1 - \alpha)\mu} - e \right\|_2 \end{aligned}$$

$$\begin{aligned}
&= 1 - \left\| \frac{Xs}{\mu} + \frac{\alpha^2}{1-\alpha} \frac{V_p(v_q)}{\mu} - e \right\|_2 \\
&\geq 1 - \left\| \frac{Xs}{\mu} - e \right\|_2 - \frac{\alpha^2}{(1-\alpha)\mu} \|V_p(v_q)\|_2.
\end{aligned}$$

The first equality is by definition, the second equality follows from (17). The inequality is simply a triangle inequality. Now, as in the proof of Lemma 3.1, $\|v_p\|_2 \leq \|v\|_2$ and $\|v_q\|_2 \leq \|v\|_2$ which implies

$$\|V_p v_q\|_2 \leq \|v\|_2^2 = n\mu.$$

Hence dividing both sides by π we get

$$\frac{\pi(\alpha)}{\pi} \geq 1 - \left(\frac{\alpha^2}{1-\alpha} \right) \frac{n}{\pi}$$

as desired. □

Lemma 5.2. *Unless $(x(1), s(1))$ is feasible (hence optimal) there exists an $\epsilon > 0$ such that for $\alpha \in (1 - \epsilon, 1)$ we have*

$$(1 - \alpha)^{\rho+1} > \frac{\delta(x(\alpha), s(\alpha))\mu(\alpha)}{\delta(x, s)\mu}.$$

Proof: For $\alpha \in (0, 1)$ we have

$$\frac{\delta(x(\alpha), s(\alpha))\mu(\alpha)}{\delta(x, s)\mu} = \frac{1 - \alpha - \left\| (1 - \alpha) \left(\frac{Xs}{\mu} - e \right) + \alpha^2 V_p v_q \right\|_2}{1 - \left\| \frac{Xs}{\mu} - e \right\|_2}$$

Clearly, if $\|V_p v_q\|_2 \neq 0$ then the right hand side above is strictly negative as $\alpha \rightarrow 1$. It is also clear that the right hand side is a continuous function of α in $(0, 1)$. Hence, there must exist an $\epsilon > 0$ such that for $\alpha \in (1 - \epsilon, 1)$ we have

$$(1 - \alpha)^{\rho+1} > \frac{\delta(x(\alpha), s(\alpha))\mu(\alpha)}{\delta(x, s)\mu}$$

as desired. □

Theorem 5.1. *Let $\psi := \psi_2$ and $\mathcal{N} := \mathcal{N}_2$ in the Algorithm described in section 2. Then using ψ_2 with $\rho = 1/2t$, $\rho = 1$, and $\rho = \sqrt{n}$ we get the iteration bounds $O(n t^2)$, $O(n 2^{2t} t)$ and $O(\sqrt{n} 2^{2t\sqrt{n}} t)$ respectively.*

Proof: Using lemmas 5.1 and 5.2 we see that the assumptions of Theorem 4.1 hold with $M = n$ then the conclusion of the theorem directly follows from Theorem 4.1. $\square$

Note that Remark 3.2 also applies here. Given $(x, s) \in \mathcal{N}_\infty^-(\pi_o)$ for some $\beta \in (0, \pi_o)$ (independent of the input), one can define

$$\pi := 1 - \beta - \left\| \frac{Xs}{\mu} - e \right\|_2.$$

It is not hard to check that all the previous results hold with the new definition of π . With this definition, all the iterates have to lie in the neighborhood $\mathcal{N}_\infty^-(\beta)$. Clearly this property applies to any pair of compatible potential function and neighborhood that fits into the framework given in section 4. So, using this framework one can guarantee that the all iterates lie in a certain prescribed neighborhood of the central path defined with respect to a certain function $\delta(x, s)$.

6 Conclusion

Using a different potential function we provided the same complexity bounds as those obtained by Mizuno and Nagasawa [MN92]. Since the potential function used by our algorithm fits in nicely with the one-sided infinity-norm neighborhoods, we could get a compact and useful description of the contours of the potential function for the affine scaling direction (one may also speculate that for the same reason potential function ψ can be more effective for the infinity neighborhoods than the Tanabe-Todd-Ye potential function). The contours of the potential function provide very useful information about the relation between the duality gap and the distance to the boundary which makes the convergence proof arguments simpler. Note that the lower bounds we provided on the step sizes (for $\rho = 1/2t$) are not better than the one given in [MAR90]; however, our algorithms are expected to take larger steps.

We also generalized these results by defining a relationship between potential functions and neighborhoods. By specializing on the two-norm neighborhoods we proposed a compatible potential function which allowed us to prove the same worst case complexity bound. Even though the potential function that is compatible with the infinity neighborhoods is unimodular in α , the same is not true for the two-norm potential function. On the other hand the two-norm potential function is differentiable in the interior of a corresponding neighborhood which might be useful in providing some information about good search directions.

We showed that using our framework one can guarantee that all the iterates in a desired neighborhood of the central path without any need for centering. As Mizuno [Mi92] pointed out, this framework is quite similar to the one given in [MTY90]. In that paper, all the iterates are kept in a neighborhood of the central path; here, the iterates are kept in a contour of a potential function. Because of the compatibility property a potential function is very closely related to its compatible neighborhood. One property that the contours possess is that at any point on a contour, one can take a large enough affine scaling step that will guarantee global convergence and will keep the next iterate on the same contour. Clearly the neighborhoods in [MTY90] do not possess this property. This approach brings more flexibility to interior point algorithms as one can use it for algorithms that have centering component or other components in their search directions.

Acknowledgment: The author wishes to thank Shinji Mizuno and Mike Todd for very valuable and inspiring discussions and suggestions.

References

- [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.
- [KMNY91] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise, *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems, Lecture Note in Computer Science*, Springer-Verlag, New York.
- [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.
- [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.
- [Mi92] S. Mizuno, *private communication*.
- [MN92] S. Mizuno and A. Nagasawa, A Primal-Dual Affine Scaling Potential Reduction Algorithm for Linear Programming, Research Memorandum No. 427, The Institute of Statistical Mathematics, Tokyo, Japan.
- [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.

- [MAR90] R.C. Monteiro, I. Adler and M.G.C. Resende, A Polynomial Time Primal-Dual Affine Scaling Algorithm for Linear and Convex Quadratic Programming and Its Power Series Extension, *Math. of OR* 15, pp. 191-214.
- [Ta87] K. Tanabe, Centered Newton Method for Mathematical Programming, *in System Modeling and Optimization* (eds. M. Iri and K. Yajima), Springer-Verlag, The Proceedings of the 13th IFIP Conference, Tokyo, Aug. 31 - Sept. 4 (1987), 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.
- [Tu92] L. Tunçel, On the Convergence of Primal-Dual Interior-Point Methods with Wide Neighborhoods, *TR-1003*, SORIE, Cornell University, Ithaca, NY.
- [Ye91] Y. Ye, On the Q-order Convergence of Interior Point Algorithms for Linear Programming, *Working Paper 91-17*, College of Business Administration, The University of Iowa, Iowa City, Iowa.