Superlinear Convergence of the Affine Scaling Algorithm

T. Tsuchiya
R. D. C. Monteiro

CRPC-TR92288
November 1992

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

Superlinear Convergence of the Affine Scaling Algorithm

T. Tsuchiya* R. D. C. Monteiro†

November 28, 1992

Abstract

In this paper we show that a variant of the long-step affine scaling algorithm analyzed by Tsuchiya and Muramatsu can have a two-step superlinear convergence property for general linear programming problems. Superlinear convergence of the dual estimate is also established. A plausible explanation is given for why step-size 2/3 is sharp for convergence of the dual estimates as long as we use fixed ratio step-size.

1 Introduction

The affine scaling algorithm, introduced by Dikin [6] in 1967, is one of the simplest and most efficient interior point algorithms for solving linear programming (LP) problems. Because of the theoretical and practical importance, there are a number of papers which study its global and local convergence [4, 6, 7, 8, 9, 12, 21, 22, 24, 25, 26, 27] as well as its continuous trajectory [3, 12, 28]. See [1, 2, 5, 14, 17, 18] for experiments and implementation issues of the algorithm.

Recently, Dikin [8] and Tsuchiya and Muramatsu [25] have succeeded in proving the global convergence for degenerate LP problems of the long step version of the affine scaling algorithm [27], that is the version in which the next iterate is determined by taking a fixed fraction $\lambda^k \in (0,1)$ of the whole step to the boundary of the feasible region. Dikin showed global convergence of the primal iterates and convergence of the dual estimate to the analytic center of the dual optimal face in the case of $\lambda^k = 1/2$. Subsequent to his work, independently, Tsuchiya and Muramatsu obtained an analogous result for $\lambda^k < 2/3$ (or $\lambda^k \leq 2/3$ in revision). They also demonstrated that the asymptotic reduction rate of the objective function value is exactly $1 - \lambda^k$ (if $\lambda^k \leq 2/3$). See the recent paper by Monteiro, Tsuchiya and Wang for a simplified and self-contained proof of these results [15].

*The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-Ku, Tokyo, 106, Japan. The work of this author was based on research supported by the Overseas Research Scholars of the Ministry of Education, Science and Culture of Japan

†Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona 85721. The work of this author was based on research supported by the National Science Foundation (NSF) under grant DDM-9109404 and the Office of Naval Research (ONR) under PR number 4116688—01.

In this paper we focus our attention on the local convergence property of the long-step affine-scaling algorithm. Specifically, we will show that a variant of the long-step affine scaling algorithm can enjoy a 2-step superlinear convergence property without sacrificing global convergence by choosing λ^k carefully. We define a trajectory in the space of always-active variables on the optimal face which plays a role similar to the central trajectory for other interior point algorithms, and take a long step as long as iterates stay close to this trajectory; detecting that iterates are leaving from the trajectory, we take $\lambda^k = 1/2$ as a corrector step.

This paper is organized as follows. In Section 2, we introduce basic assumptions, terminology and define the affine scaling algorithm. Section 3 and 4 are the main part of this paper. In Section 3, we state several basic results on the affine scaling direction near a dual degenerate face, which is defined to be a face on which objective function takes a constant value. This analysis gives a basis for obtaining the superlinear convergence result. Further, we take up the special case of homogeneous problems with unique solutions, and prove that the direction of approach to the optimal solution cannot converge to one point if we take any fraction greater than 2/3. Since dual estimate is a function of the direction of approach to the optimal solution, this result gives a plausible reasoning for why $\lambda^k = 2/3$ is sharp concerning convergence of dual estimates as observed in Tsuchiya and Muramatsu [25] and Hall and Vanderbei [10] (N. B. the latter result is stronger). In Section 4, we define a new variant of the affine scaling algorithm which takes either $\lambda^k = 1/2$ or $\lambda^k \sim 1$ alternatively, and show that this variant has 2-step superlinear convergence property with Q-order 1.3 with respect to the objective function value. Superlinear convergence of primal iterates to an interior point of the optimal face and dual estimates to the analytic center of the dual optimal face with the same R-order (1.3) is also shown. In Section 5 we prove some technical lemmas stated in Section 3. Finally, we give a concluding remark in Section 6.

The following notation is used throughout our paper. We denote the vector of all ones by e. Its dimension is always clear from the context. $\mathbb{R}^n$, $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ denote the n-dimensional Euclidean space, the nonnegative orthant of $\mathbb{R}^n$ and the positive orthant of $\mathbb{R}^n$, respectively. The set of all $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m\times n}$. Given an index set $J \subseteq \{1,\ldots,n\}$ and a vector $w \in \mathbb{R}^n$, we denote by w_J the subvector of w corresponding to J. Similarly, if E is an $m \times n$ matrix then E_J denotes the $m \times |J|$ submatrix of E corresponding to J. For a vector w, we let $\chi[w]$ denote the largest component of w. The Euclidean norm, the 1-norm and the ∞-norm are denoted by $\|\cdot\|$, $\|\cdot\|_1$ and $\|\cdot\|_\infty$, respectively. If J is a finite index set then $|J|$ denotes its cardinality, that is the number of elements of J. The superscript ^{T} denotes transpose.

2 Affine Scaling Algorithm

In this section, we state the main terminology and assumptions used throughout our paper and describe the affine scaling algorithm.

Consider the following linear programming problem

$$\begin{aligned} &\text{minimize}_x \quad c^T x \\ &\text{subject to} \quad Ax = b, \qquad x \geq 0, \end{aligned} \tag{1}$$

and its associated dual problem

$$\begin{array}{lll} \text{maximize}_{(y,s)} & b^T y & \\ \text{subject to} & A^T y + s = c, & s \geq 0, \end{array} \tag{2}$$

where $A \in \mathbb{R}^{m\times n}$, $c, x, s \in \mathbb{R}^n$ and $b, y \in \mathbb{R}^m$. We will denote by $\mathcal{P}^+$ and $\mathcal{P}^{++}$ the feasible region and the interior of the feasible region of (1), respectively.

We impose the following assumptions throughout this paper.

Assumption 1: $Rank(A) = m$;

Assumption 2: The objective function $c^T x$ is not constant over the feasible region of (1);

Assumption 3: Problem (1) has an interior feasible solution, that is $\mathcal{P}^{++} \neq \emptyset$;

Assumption 4: Problem (1) has an optimal solution, where the optimal value is c^*.

Given an index set $J \in \{1, ..., n\}$, the subset $\{x | Ax = b, x_J = 0, x \geq 0\}$ of $\mathcal{P}^+$ obtained by letting $x_J = 0$, is called a face of $\mathcal{P}^+$ if it is not empty. We specify the face with a partition (N, B) of the index set of the variables, where N consists of the indices for the variables whose values are always zero on that face and B being its complement. This index set N is referred to as the *always-active index set of the face.* In particular, the face is referred to as a *dual degenerate face* if the objective function takes a constant value on it. We regard a vertex also as a dual degenerate face. The always-active index set of a dual degenerate face is called a *dual degenerate index set.*

The whole set of the optimal solutions is referred to as the *optimal face.* We use (N_*, B_*) for the partition of the index set to determine the optimal face, where N_* is the dual degenerate index set for the optimal face and B_* is its complement.

The problem is called *homogeneous* when b is equal to zero. In the case, $x = 0$ is a dual degenerate face which is specified by the dual degenerate index set $N = \{1, ..., n\}$, and we have $B = \emptyset$. Conventionally, we define $Range(A_B) = \{0\}$ when $B = \emptyset$, and every quantity that contains subvectors and submatrices with the index set B is put to zero in interpretating lemmas, propositions and theorems. This will not cause any inconsistency.

We now introduce important functions which are used in the description and in the analysis of the affine scaling algorithm. For every $x \in \mathbb{R}^n_{++}$, let

$$y(x) \equiv (AX^2A^T)^{-1}AX^2c, \tag{3a}$$
$$s(x) \equiv c - A^T y(x), \tag{3b}$$
$$dx(x) \equiv X^2 s(x) = X[I - XA^T(AX^2A^T)^{-1}AX]Xc, \tag{3c}$$
$$d(x) \equiv Xs(x) = X^{-1}dx(x) \tag{3d}$$

where $X \equiv diag(x)$. We note that Assumption 1 implies that the inverse of AX^2A^T exists for every $x > 0$. The quantity $s(x)$ (or the pair $(y(x), s(x))$) is called the *dual estimate* associated with the point x. $dx(x)$ and $d(x)$ are referred to as the *affine scaling direction* and the *scaled affine scaling direction* associated with x, respectively.

The following lemma is well-known, and easily verified.

Lemma 2.1 *$dx(x)$ is given as the solution for the following optimization problem*

$$\begin{aligned} &\text{maximize}_{\,p} \quad \tilde{s}^T p - \tfrac{1}{2}\|X^{-1}p\|^2 \\ &\text{subject to} \quad Ap = 0, \end{aligned} \tag{4}$$

where $\tilde{s} \in c + Range(A^T)$. The search direction is not affected by choice of $\tilde{s}$.

We are ready to describe the affine scaling algorithm. For a good motivation of the method, we refer the reader to Dikin [6], Barnes [4], Vanderbei, Meketon and Freedman [27] and Vanderbei and Lagarias [26].

(Step 0) Assume $x^0 \in \mathcal{P}^{++}$ is available. Set $k := 0$.

(Step 1) Choose $\lambda^k \in (0,1)$, and

$$dx^k = dx(x^k); \tag{5a}$$

$$X^k = diag(x^k); \tag{5b}$$

$$x^{k+1} = x^k - \frac{\lambda^k}{\chi[(X^k)^{-1}dx^k]} dx^k. \tag{5c}$$

(Step 2) $k := k+1$ and return to (Step 1).

Note that if we choose $\lambda^k = 1$, the next iterate is just on the boundary of the feasible region. Thus the iteration is well-defined as long as we take $\lambda^k \in (0,1)$. In the sequel, we denote by x^k the kth iterate generated by the affine scaling algorithm. Given a function f of x^k, we abbreviate $f(x^k)$ as f^k. Further, we use X and X^k to denote $diag(x)$ and $diag(x^k)$, respectively. Similarly, given an index set J and k, we also use convention X_J and X_J^k to denote $diag(x_J)$ and $diag(x_J^k)$. In many of the existing implementation they fix λ^k to be a constant throughout the iterations. This type of step-size choice is referred to as a *fixed ratio step-size choice* in this paper.

3 Analysis of the Affine Scaling Search Direction near Dual Degenerate Faces

3.1 Preliminary Results

In this subsection we state preliminary results that are basis for deriving the main results in this paper. In order to make arguments clear, some of their proofs are put off to Section 5. Let us consider a dual degenerate face determined with the dual degenerate index set N and its complement B, where the objective function value is c'. Then there exists $(\bar{y}, \bar{s})$ such that

$$A_N^T \bar{y} + \bar{s}_N = c_N, \quad A_B^T \bar{y} = c_B, \tag{6}$$

with which we have

$$c^T x - c' = \bar{s}_N^T x_N \tag{7}$$

for all $x \in \mathcal{P}^+$. Let

$$\mathcal{Q}_N^+ \equiv \{x|\ x \in \mathcal{P}^+,\ c^T x - c' \geq 0\} \tag{8}$$

and

$$\mathcal{Q}_N^{++} \equiv \{x|\ x \in \mathcal{P}^{++},\ c^T x - c' > 0\}. \tag{9}$$

For $x \in \mathcal{Q}_N^{++}$, we define

$$u(x) \equiv \frac{d(x)}{c^T x - c'}. \tag{10}$$

The following estimates on $u(x)$ in the vicinity of the dual degenerate face are used in the consecutive analysis.

Lemma 3.1 *We have*

$$\frac{\|u_B(x)\|}{\|u_N(x)\|} = \frac{\|d_B(x)\|}{\|d_N(x)\|} = O(\|X_B^{-1}\|\|X_N\|). \tag{11}$$

in $\mathcal{Q}_N^{++}$.

Lemma 3.2 *We have*

$$e^T u_N(x) = 1 + O(\|X_B^{-1}\|^2\|X_N\|^2\|v_N(x)\|) \tag{12}$$

in $\mathcal{Q}_N^{++}$.

Now we consider the polyhedron

$$\mathcal{V}_N^+ \equiv \{v_N \geq 0 | A_N v_N \in Range(A_B), \bar{s}_N^T v_N = 1\} \tag{13}$$

and its interior

$$\mathcal{V}_N^{++} = \{v_N > 0 | A_N v_N \in Range(A_B), \bar{s}_N^T v_N = 1\}, \tag{14}$$

and define, for any $x \in \mathcal{Q}_N^{++}$,

$$v(x) \equiv \frac{x}{c^T x - c'} = \frac{x}{\bar{s}_N^T x_N}. \tag{15}$$

Obviously, $v_N(x)$ is a mapping from $x \in \mathcal{Q}_N^{++}$ to $\mathcal{V}_N^{++}$.

We define the analytic center of $\mathcal{V}_N^+$ as an optimal point for the problem

$$\text{minimize } -\sum_{i \in N} \log v_i, \quad \text{subject to } v_N \in \mathcal{V}_N^+, \tag{16}$$

which exists if and only if $\mathcal{V}_N^+$ is bounded.

We denote by $dv_N(v_N)$ the Newton direction at $v_N \in \mathcal{V}_N^{++}$ for the analytic center of $\mathcal{V}_N^+$. The direction $dv_N(v_N)$ is the optimal solution for the following optimization problem:

$$\begin{aligned} &\text{minimize } e^T V_N^{-1} p_N + \tfrac{1}{2}\|V_N^{-1} p_N\|^2, \\ &\text{subject to } A_N p_N \in Range(A_B), \quad \bar{s}_N^T p_N = 0. \end{aligned} \tag{17}$$

Here $V_N \equiv diag(v_N)$.

The Newton iteration is written as

$$v_N^+ = v_N - dv_N(v_N), \tag{18}$$

where v_N^+ is the next iterate. We define the scaled Newton direction as:

$$w_N(v_N) \equiv V_N^{-1} dv_N(v_N). \tag{19}$$

Note that the (scaled) Newton direction is well-defined even though the polyhedron $\mathcal{V}_N^+$ is not bounded.

Given an affine scaling sequence $\{x^k\}$ in $\mathcal{Q}_N^{++}$, we consider the sequence $\{v_N(x^k)\}$ in $\mathcal{V}_N^{++}$. The following results which show that $\{v_N(x^k)\}$ is approximately regarded as the sequence of the Newton iteration for the analytic center of $\mathcal{V}_N^+$ are the key observation in this paper. A proof of Theorem 3.3 will be given in Section 5. In the remaining part, we use the analogous notation to x and x^k concerning v and v^k, i.e., given an index set J and an iteration count k, we use V, V_J, V^k, V_J^k for $diag(v)$, $diag(v_J)$, $diag(v^k)$, $diag(v_J^k)$, etc.

Theorem 3.3 *We have*

$$\frac{u_N(x)}{\|u(x)\|^2} = w_N(v_N(x)) + e + r_N(x) \tag{20}$$

in $\mathcal{Q}_N^{++}$, *where* $\|r_N(x)\| = O(\|X_B^{-1}\|^2 \|X_N\|^2 \|v_N(x)\|)$.

Theorem 3.4 *Let* $x \in \mathcal{Q}_N^{++}$, *and let* $v_N^+(x,\lambda)$ *be* $v_N(x^+(\lambda))$, *where* $x^+(\lambda) \in \mathcal{Q}_N^{++}$ *is the point obtained with an affine scaling iteration with the step-size* λ *from* x. *Then we have*

$$v_N^+(x,\lambda) = v_N(x) - \frac{\lambda\delta(u(x))}{1-\lambda\delta(u(x))}(dv_N(v_N(x)) + V_N r_N(x)), \tag{21}$$

$$\delta(u) \equiv \frac{\|u\|^2}{\chi[u]}, \tag{22}$$

where $r_N(x)$ *is the same as in Theorem 3.3.*

Proof. We have

$$\begin{aligned} v_N^+(x,\lambda) - v_N(x) &= \frac{x_N^+(\lambda)}{c^T x^+(\lambda) - c'} - \frac{x_N}{c^T x - c'} \\ &= \frac{(c^T x - c')(x_N - \lambda dx_N(x)/\chi[d(x)]) - x_N(c^T x - c' - \lambda c^T dx(x)/\chi[d(x)])}{(c^T x^+(\lambda) - c')(c^T x - c')} \\ &= \frac{-\lambda(c^T x - c') dx_N(x)/\chi[d(x)] + x_N \lambda c^T dx(x)/\chi[d(x)]}{(c^T x - c' - \lambda c^T dx(x)/\chi[d(x)])(c^T x - c')} \\ &= \frac{\lambda c^T dx(x)/\chi[d(x)]}{c^T x - c' - \lambda c^T dx(x)/\chi[d(x)]}\left(\frac{x_N}{c^T x - c'} - \frac{dx_N(x)}{c^T dx(x)}\right) \\ &= \frac{\lambda\|u\|^2/\chi[u]}{1-\lambda\|u\|^2/\chi[u]}\left(v_N - V_N\frac{u_N}{\|u\|^2}\right) \\ &= \frac{\lambda\delta(u)}{1-\lambda\delta(u)}\left(v_N - V_N\frac{u_N}{\|u\|^2}\right). \end{aligned} \tag{23}$$

Substituting (20) into the last expression, the theorem is immediate. ■

The following lemma is easily verified, and will be referred frequently in the consecutive analysis.

Lemma 3.5 *Let $x \in \mathcal{Q}_N^{++}$, and let $x^+(\lambda) \in \mathcal{Q}_N^+$ is the point obtained with an affine scaling iteration with the step-size λ from x. Then we have*

$$\frac{c^T x^+(\lambda) - c'}{c^T x - c'} = 1 - \lambda\delta(u(x)) \geq 0. \tag{24}$$

3.2 Convergence Analysis of the Direction of Approach with Fixed Ratio Step-size Choices

Now we apply the results in the previous subsection to the case where the dual degenerate face is chosen to be the optimal face. We denote by N_* the dual degenerate index set for the optimal face of (1) and by B_* its complement. Let $\bar{s} = (\bar{s}_{N_*}, \bar{s}_{B_*}) = (\bar{s}_{N_*}, 0)$ be an interior point of the dual optimal face. Due to strictly complementarity, we have $\bar{s}_{N_*} > 0$ and

$$c^T x - c^* = \bar{s}_{N_*}^T x_{N_*} \tag{25}$$

for all $x \in \mathcal{P}^+$. Since $\bar{s}_{N_*} > 0$, we may regard $v_{N_*}(x)$ as the direction of x in the x_{N_*} space viewed from the optimal face of (1). The vector $v_{N_*}(x)$ is an element of the polyhedron

$$\mathcal{V}_{N_*}^+ = \{v_{N_*} \geq 0 | A_{N_*} v_{N_*} \in Range(A_{B_*}), \bar{s}_{N_*}^T v_{N_*} = 1\}. \tag{26}$$

Obviously this polyhedron is bounded. We denote by $v_{N_*}^*$ the analytic center of $\mathcal{V}_{N_*}^+$, and take

$$\psi(v_{N_*}) \equiv \|(V_{N_*}^*)^{-1}(v_{N_*} - v_{N_*}^*)\| \quad (V_{N_*}^* \equiv diag(v_{N_*}^*)) \tag{27}$$

as a measure for how $v_{N_*}^*$ is close to the analytic center. We easily verify the following lemmas:

Lemma 3.6 *If $v_{N_*} \to v_{N_*}^*$, we have*

$$\lim_{v \to v^*} \frac{\|w_{N_*}(v_{N_*})\|}{\psi(v_{N_*})} = \lim_{v \to v^*} \frac{\|(V_{N_*})^{-1} dv_{N_*}(v_{N_*})\|}{\|(V_{N_*}^*)^{-1}(v_{N_*} - v_{N_*}^*)\|} = 1. \tag{28}$$

Lemma 3.7 *There exist constants $M > 0$ and $\varepsilon > 0$ satisfying*

$$\psi(v_{N_*} - dv_{N_*}(v_{N_*})) \leq M\psi(v_{N_*})^2 \tag{29}$$

and

$$\|(V_{N_*}^*)^{-1} dv_{N_*}(v_{N_*})\| \leq 2\psi(v_{N_*}) \tag{30}$$

for all v_{N_*} satisfying $\psi(v_{N_*}) \leq \varepsilon$.

In order to obtain a geometrical image of the analysis, it is useful to consider the following set

$$\begin{aligned} \mathcal{M}_{N_*} &\equiv \{x \in \mathcal{P}^+ | \psi(v_{N_*}(x)) = 0\} \\ &= \{x \in \mathcal{P}^+ | x_{N_*} = \mu v^*_{N_*} \text{ for some } \mu > 0\}. \end{aligned} \tag{31}$$

This set is a cross section of $\mathcal{P}^+$ and the hyperplane

$$\{x \in \mathbb{R}^n | x_{N_*} = \mu v^*_{N_*}\}. \tag{32}$$

If the optimal face is a unique point, $\mathcal{M}_{N_*}$ is a line emanating from the optimal solution tangential to the limiting central trajectory. Generally, the dimension of $\mathcal{M}_{N_*}$ is given by "(the dimension of the optimal face) +1."

The role of $\mathcal{M}_{N_*}$ in the local analysis of affine scaling algorithm is similar to the role of central trajectory in the local analysis of many other interior point algorithms. Intuitively speaking, as long as we are approaching the optimal face from the direction of $\mathcal{M}_{N_*}$, the affine scaling algorithm can enjoy a nice local convergence property.

Lemma 3.8 *Let $\{x^l\}$ be a sequence on $\mathcal{Q}^{++}_{N_*}$. If*

$$\lim_{l\to\infty} u_{N_*}(x^l) = \frac{e}{|N_*|}, \tag{33}$$

every accumulation point of $\{v_{N_*}(x^l)\}$ is an interior point of $\mathcal{V}^+_{N_*}$.

Proof. By assumption, we have

$$\lim_{l\to\infty} u^l_{N_*} = \lim_{l\to\infty} V^l_{N_*} s_{N_*}(x^l) = \lim_{l\to\infty} S^l_{N_*} v^l_{N_*} = \frac{e}{|N_*|}, \tag{34}$$

where $S^l_{N_*} = diag(s_{N_*}(x^l))$. Let $\hat{v}_{N_*}$ be an accumulation point of $\{v^l_{N_*}\}$. Since the function $s(x)$ is bounded over the feasible region (cf. Lemma of [26](page 118) or Proposition 2.8 of [15]), the sequence $\{s^l\}$ is bounded as a whole sequence. Surpassing a subsequence $\{x^l\}_{l\in L}$ if necessary, we may assume that the sequence $\{v^l_{N_*}\}_{l\in L}$ is convergent to $\hat{v}_{N_*} \geq 0$, while $\{s^l_{N_*}\}_{l\in L}$ is convergent to an accumulation point $\hat{s}_{N_*}$ of $\{s^l_{N_*}\}$. Then we have

$$\lim_{l\in L} S^l_{N_*} v^l_{N_*} = \hat{S}_{N_*} \hat{v}_{N_*} = \frac{e}{|N_*|}, \tag{35}$$

where $\hat{S}_{N_*} \equiv diag(\hat{s}_{N_*})$. Since there exists a constant M_0 such that

$$\|s^l_{N_*}\| \leq M_0 \tag{36}$$

for all l, we have

$$\min_{i\in N_*} \hat{v}_i \geq \frac{1}{M_0|N_*|}. \tag{37}$$

Since the last inequality holds for any choice of $\hat{v}_{N_*}$, we see that every accumulation point of $\{v^l_{N_*}\}$ is in the interior of $\mathcal{V}^+_{N_*}$. This completes the proof.

■

Now we are ready to analyze the direction of approach to the optimal set under fixed ratio step-size choices.

3.2.1 Homogeneous Case

In [25], Tsuchiya and Muramatsu proved that for general LP problems, dual estimates converge to the analytic center of the dual optimal face if $\lambda^k \le 2/3$, and gave an example to show that $\lambda^k = 2/3$ is sharp on this property as long as we take fixed ratio step-size. More strongly, Hall and Vanderbei [10] demonstrated that $\lambda^k = 2/3$ is the largest fixed ratio step-size choice that can ensure convergence of dual estimates to one point. They both obtained these results by constructing small examples of homogeneous LP problems, and the basic observation that leads to their results is that the sequence cannot have the limiting direction of approach to the optimal solution if we take $\lambda^k = \lambda > 2/3$. Here, on the basis of the relationship between $\{x^k\}$ and $\{v_{N_*}(x^k)\}$ obtained in the last section, we give a plausible explanation for why 2/3 is sharp on convergence of dual estimates.

Let us consider a homogeneous problem with a unique optimal solution; i.e., $b = 0$ and the point $x^* = 0$ is the unique point where $c^T x^* = 0$ ($c^T x > 0$ for any other feasible solution x). This is an LP problem whose feasible region is a cone, where the origin is a unique optimal solution. We note that in the case we have $N_* = \{1, ..., n\}$ and $B_* = \emptyset$. Recall the rule stated in Section 2 that every quantity that contains subvectors and submatrices with B_* component, $\|X_{B_*}\|$, say, is put to zero in interpreting the lemmas and theorems in Section 3.1, if $B_* = \emptyset$. Then (12), (20) and (21) become

$$e^T u = e^T u_{N_*} = 1, \tag{38}$$

$$\frac{u_{N_*}(x)}{\|u(x)\|^2} = \frac{u(x)}{\|u(x)\|^2} = w_{N_*}(v_{N_*}(x)) + e, \tag{39}$$

and

$$v^+_{N_*}(x, \lambda) = v_{N_*}(x) - \frac{\lambda\delta(u(x))}{1 - \lambda\delta(u(x))} d v_{N_*}(v_{N_*}(x)), \tag{40}$$

respectively.

Below we demonstrate that the sequence $\{v^k_{N_*}\}$ converges quadratically to the analytic center if we take $\lambda^k = 1/2$, and $\{v^k_{N_*}\}$ has at least two different accumulation points if $\lambda^k = \lambda > 2/3$ (except for the special case where $v^k_{N_*}$ happens to be $v^*_{N_*}$ for some k). Since the dual estimate $s(x)$ is a function of the direction $v_{N_*}(x)$ of approach to x^*, it is likely that $s(x^k)$ do not converge if $v_{N_*}(x^k)$ has more than two different accumulation points. Thus the next theorem gives a nice geometrical explanation for why 2/3 appears in the bounds for convergence of dual estimates. We note that Dikin analyzed the same situation in [8] with $\lambda^k = 1/2$, and showed convergence of $v^k_{N_*}$ to $v^*_{N_*}$ by observing the reduction of the Karmarkar potential function [11] associated with this problem.

Theorem 3.9 *Let (1) be a homogeneous problem with a unique optimal solution $x^* = 0$, i.e., $N_* = \{1, ..., n\}$ and let $\{x^k\}$ be the sequence obtained by the affine scaling algorithm with fixed step-size $\lambda^k = \lambda > 0$, and assume also that $v^k_{N_*} \ne v^*_{N_*}$ for all k. Then, (I) if $\lambda = 1/2$, $\{v^k_{N_*}\}$ converges to $v^*_{N_*}$ quadratically; (II) if $\lambda > 2/3$, $\{v^k_{N_*}\}$ has at least two different accumulation points.*

Proof. Due to (38), we have

$$\delta(u^k) \ge \|u^k\| \ge \frac{1}{\sqrt{n}}. \tag{41}$$

Using Lemma 3.5, we have, for all k,

$$\frac{c^T(x^{k+1}-x^*)}{c^T(x^k-x^*)} = 1-\lambda\delta(u^k) \le 1-\frac{\lambda}{\sqrt{n}}, \tag{42}$$

which immediately implies that the sequence converges to the unique optimal point $x^*=0$.

Due to (40), the iterative formula for v_{N_*} is written as follows:

$$v_{N_*}^{k+1} \;=\; v_{N_*}^k - \frac{\lambda\delta^k}{1-\lambda\delta^k} dv_{N_*}(v_{N_*}^k). \tag{43}$$

First, we show that $\{v_{N_*}^k\}$ converges to the analytic center $v_{N_*}^*$ quadratically if $\lambda = 1/2$. It is shown in Lemma 4.1 of [15] that $\{u_{N_*}^k\}$ converges to $e/|N_*|$. Applying Lemma 3.8, we see every accumulation point of $\{v_{N_*}^k\}$ is in $\mathcal{V}_{N_*}^{++}$. From (39), we see $\lim_{k\to\infty} w_{N_*}^k = 0$. Since $v_{N_*}^*$ is a unique interior point in $\mathcal{V}_{N_*}^{+}$ such that $w_{N_*}(v_{N_*}) = 0$, $\{v_{N_*}^k\}$ converges to $v_{N_*}^*$.

To see quadratic convergence, it is enough to show that

$$\frac{\delta^k/2}{1-\delta^k/2} = 1 + O(\|v_{N_*}^k - v_{N_*}^*\|). \tag{44}$$

Indeed, we have

$$\frac{1}{\delta(u^k)} = \frac{\chi[u^k]}{\|u^k\|^2} = \frac{\chi[u_{N_*}^k]}{\|u^k\|^2} = \chi[\frac{u_{N_*}^k}{\|u^k\|^2}] = \chi[e + w_{N_*}(v_{N_*}^k)]. \tag{45}$$

Since $v_{N_*}^k \to v_{N_*}^*$, due to Lemma 3.6,

$$\|w_{N_*}^k\| \le 1.1\psi(v_{N_*}^k) \tag{46}$$

holds for sufficiently large k. The two relations above immediately implies (44).

Now we deal with the case of $\lambda^k = \lambda > 2/3$.

Assume by contradiction that $\{v_{N_*}^k\}$ converges to a point $\hat{v}_{N_*}$ of $\mathcal{V}_{N_*}^{+}$. Since $\{v_{N_*}^k\}$ converges to one point while we have

$$0 < \frac{\lambda/\sqrt{|N_*|}}{1-\lambda/\sqrt{|N_*|}} \le \frac{\lambda\delta^k}{1-\lambda\delta^k} \tag{47}$$

for all k, as is seen from (41), we see that $dv_{N_*}(v_{N_*}^k)$ tends to zero.

First, we assume that $\hat{v}_{N_*}$ is an interior point of $\mathcal{V}_{N_*}^{+}$ and derive a contradiction. Since $v_{N_*}^*$ is the unique interior point of $\mathcal{V}_{N_*}^{+}$ where $dv_{N_*}(v_{N_*}) = 0$, this implies $v_{N_*}^k \to v_{N_*}^*$. Then applying (39), we see $\delta^k \to 1$ as $k \to \infty$, and this implies that

$$\tilde{\lambda} = \lim_{k\to\infty} \frac{\lambda\delta^k}{1-\lambda\delta^k} > 2 \tag{48}$$

if $\lambda > 2/3$. Now (43) becomes exactly identical to the Newton iteration with step-size greater than 2, and the iterates cannot be convergent to $v_{N_*}^*$ except for the special case where $v_{N_*}^k = v_{N_*}^*$ takes place for some k, which is the contradiction.

Hence, $\hat{v}_{N_*}$ has to be on the boundary of $\mathcal{V}^+_{N_*}$. We observe that this also leads to a contradiction. To this end, we use the following property of the scaled Newton direction for the analytic center.

(Property of the Newton direction for the analytic center)

Let $\bar{v}_{N_*}$ be a point on the boundary of $\mathcal{V}^+_{N_*}$, and J be the index set corresponding to $\bar{v}_i$ whose value is 0 at $\bar{v}_{N_*}$. Then we have $w_J(v_{N_*}) \to -e$ as $v_{N_*} \to \bar{v}_{N_*}$.

This fact is implicitly used to analyze quadratic convergence of Iri and Imai's algorithm [23]. For the sake of completeness, we provide a proof in the appendix.

We apply this fact letting $\bar{v}_{N_*} = \hat{v}_{N_*}$. In view of the Newton iteration (18), the property means that the values of variables v_j $(j \in J)$ increase by the iteration of the Newton method at any interior point of $\mathcal{V}^+_{N_*}$ sufficiently close to $\bar{v}_{N_*}$; thus iterate cannot converge to $\hat{v}_{N_*}$, which is a contradiction, and completes the proof.

■

3.2.2 General Case

We can extend a similar argument for general cases. However, it turns out to be difficult to duplicate completely analogous simple results here, since we have $r^k_{N_*} \neq 0$ for general cases and effects of this term on $v^k_{N_*}$ is difficult to estimate. We restrict ourself to the case of $\lambda^k = 1/2$, and show how this analysis above is extended for general case; the result plays an important role in the next section.

Theorem 3.10 *If we take fixed ratio step-size $\lambda^k = 1/2$ for general problems, we have*

$$\psi(v^k_{N_*}) = O((c^T x^k - c^*)^2). \tag{49}$$

Proof. Due to Theorem 1.1 of [25], $\{x^k\}$ converges to an interior point of the optimal face. Let

$$\zeta^k = \frac{\psi(v^k_{N_*})}{(c^T x^k - c^*)^2} \tag{50}$$

for all k. By contradiction, we assume that there exists a subsequence $\{x^k\}_{k \in K}$ such that

$$\lim_{k \in K} \zeta^k = \infty \quad \text{and} \quad \zeta^k < \zeta^{k+1} \quad \text{for all } k \in K. \tag{51}$$

Due to Lemma 4.1 of [15] and (11), we have $u^k_{N_*} \to e/|N_*|$ and $u^k_{B_*} \to 0$. Applying Lemma 3.8, we see every accumulation point of $\{v^k_{N_*}\}$ is in $\mathcal{V}^{++}_{N_*}$.

Since $\mathcal{V}^+_{N_*}$ is bounded, $\|v^k_{N_*}\|$ is bounded and $\|x^k_{N_*}\| = O(c^T x^k - c^*)$. Hence we have $\|(X^k_{B_*})^{-1}\|^2 \|(X^k_{N_*})^2\| \|v^k_{N_*}\| = O((c^T x^k - c^*)^2)$. Together with this fact, Theorem 3.3 implies that

$$\|r^k_{N_*}\| \le M_0 (c^T x^k - c^*)^2 \tag{52}$$

for sufficiently large k, where M_0 is a positive constant, and hence that

$$\lim_{k \to \infty} w^k_{N_*} = 0. \tag{53}$$

Since $v_{N_*}^*$ is a unique interior point in $\mathcal{V}_{N_*}^+$ such that $w_{N_*}(v_{N_*}) = 0$, $\{v_{N_*}^k\}$ converges to $v_{N_*}^*$. Then we have $\psi(v_{N_*}^k) \le \varepsilon$ if k is sufficiently large, where ε is the constant in Lemma 3.7. Since $\lambda^k = 1/2$, we have, due to Theorem 3.4 and Lemma 3.7 that,

$$\begin{aligned}
\psi(v_{N_*}^{k+1}) &= \|(V_{N_*}^*)^{-1}(v_{N_*}^k - v_{N_*}^* - (1-\eta^k)(dv_{N_*}^k + V_{N_*}^k r_{N_*}^k))\| \\
&\quad \text{(use Theorem 3.4)} \\
&\le \|(V_{N_*}^*)^{-1}(v_{N_*}^k - v_{N_*}^* - dv_{N_*}^k)\| + \eta^k \|(V_{N_*}^*)^{-1} dv_{N_*}^k\| + \|(V_{N_*}^*)^{-1}\| \|V_{N_*}^k\| \|r_{N_*}^k\| \\
&\le M\psi(v_{N_*}^k)^2 + 2\eta^k \psi(v_N^k) + 2M_0 \|(V_{N_*}^*)^{-1}\| \|V_{N_*}^*\| (c^T x^k - c^*)^2, \\
&\quad \text{(use Lemma 3.7 and (52))}
\end{aligned} \tag{54}$$

where

$$\eta^k \equiv 1 - \frac{\delta^k/2}{1-\delta^k/2}. \tag{55}$$

Since $\delta^k \to 1$, we see $\eta^k \to 0$ as $k \to \infty$. On the other hand, we have, as shown in Theorem 1.1 of [25] (see also Theorem 4.2 of [15]),

$$\lim_{k\to\infty} \frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = \frac{1}{2}. \tag{56}$$

Then it follows that, for $k \in K$ sufficiently large,

$$\zeta^k \le \zeta^{k+1} \le 4.1M\zeta^k \psi(v_{N_*}^k) + 2 \cdot 4.1\eta^k \zeta^k + M_1, \tag{57}$$

where $M_1 \equiv 2M_0 \|(V_{N_*}^*)^{-1}\| \|V_{N_*}^*\|$. Dividing this inequality by ζ^k, we have

$$1 \le 4.1M\psi(v_{N_*}^k) + 8.2\eta^k + \frac{M_1}{\zeta^k} \tag{58}$$

for sufficiently large k, which, however, is a contradiction, because $\psi(v_{N_*}^k) \to 0$, $\eta^k \to 0$ and $\zeta^k \to \infty$ as $k \to \infty$ $(k \in K)$. This completes the proof. ■

4 A Superlinearly Convergent Affine Scaling Algorithm

In this section we will demonstrate global and superlinear convergence of the following variant of the long-step affine scaling algorithm.

Algorithm SLA

(Step 0) Assume $x^0 \in \mathcal{P}^{++}$ is available. Set $k := 0$.

(Step 1) Compute $dx^k \equiv dx(x^k)$ and $d^k \equiv d(x^k)$ according to (3c) and (3d), respectively, and let

$$N^k = \text{The index set consisting of } i \text{ such that } x_i^k \le \sqrt{|e^T d^k|}, \tag{59a}$$

$$g^k = \sqrt{\left| |N^k| - \frac{|e^T d_{N^k}^k|^2}{\|d^k\|^2} \right|}. \tag{59b}$$

(Step 2) If

$$g^k < |e^T d^k|^{0.95}, \tag{60}$$

then (Predictor step)

$$\lambda^k = \max(0.5, 1 - |e^T d^k|^{0.3}) \tag{61}$$

else (Corrector step)

$$\lambda^k = 0.5. \tag{62}$$

(Step 3)

$$x^{k+1} = x^k - \lambda^k \frac{dx^k}{\chi[d^k]}. \tag{63}$$

(Step 4) $k := k + 1$ and return to (Step 1).

Since the step-size is ensured to be greater than 0.5, convergence of the sequence is immediately seen from Lemma 1.2 of Tsuchiya and Muramatsu [25] (see also Theorem 2.6 of Monteiro et al. [15]). Specifically, we have the following lemma.

Lemma 4.1 *$\{x^k\}$ converges to an interior point of a dual degenerate face determined by a dual degenerate index set N. Further, if we denote by c^∞ the limiting objective function value, we have a constant $\widetilde{M} > 0$ such that*

$$v_N^k = \frac{x_N^k}{c^T x^k - c^\infty} \leq \widetilde{M} e \tag{64}$$

for all k.

With this lemma, the following lemma is obvious.

Lemma 4.2 *We have*

$$\|(X_B^k)^{-1}\| \|X_N^k\| \|v_N^k\| = O(c^T x^k - c^\infty), \quad \|(X_B^k)^{-2}\| \|X_N^k\|^2 \|v_N^k\| = O((c^T x^k - c^\infty)^2). \tag{65}$$

Lemma 4.3 *We have, for all k,*

$$\|u^k\| \leq 2. \tag{66}$$

Proof. Applying Lemma 3.5, we have

$$0 \leq 1 - \lambda^k \delta(u^k) = 1 - \lambda^k \frac{\|u^k\|^2}{\chi[u^k]} \leq 1 - \lambda^k \|u^k\| \leq 1 - \frac{1}{2}\|u^k\|, \tag{67}$$

from which the statement immediately follows. ∎

Lemma 4.4 *We have*

$$\lim_{k\to\infty} \frac{e^T d^k}{c^T x^k - c^\infty} = \lim_{k\to\infty} e^T u^k = 1. \tag{68}$$

Proof. From Lemma 3.1, Lemma 4.1 and Lemma 4.3, we have

$$\|u_B^k\| \le O(c^T x^k - c^*). \tag{69}$$

On the other hand, from Lemma 3.2 and Lemma 4.2, we have

$$e^T u_{N_*}^k = 1 + O((c^T x^k - c^*)^2). \tag{70}$$

From these relations, the lemma holds. ■

Lemma 4.5 *We have* $\lim_{k\to\infty} |e^T d^k| = 0$.

Proof. Immediate from the previous lemma. ■

Lemma 4.6 *We have* $N^k = N$ *for sufficiently large* k.

Proof. Due to Lemma 4.1 and Lemma 4.4, we have $x_i^k \le \widetilde{M}(c^T x^k - c^\infty) \le 1.1\widetilde{M} e^T d^k \le \sqrt{|e^T d^k|}$ for all $i \in N$ and k sufficiently large, while x_B^k is uniformly bounded away from zero. The lemma is immediate from these facts. ■

Now we show global convergence of the algorithm.

Theorem 4.7 $\{x^k\}$ *converges to an interior point of the optimal face of (1).*

Proof. We assume x^* is not in the interior of the optimal face and derive a contradiction. If the condition (60) is satisfied at most finite number of iterative steps, global convergence follows from Theorem 1.1 of Tsuchiya and Muramatsu [25]. Hence the condition (60) is satisfied infinitely many times. In the case, we have a subsequence $\{g^k\}_{k\in K}$ such that $\lim_{k\in K} g^k = 0$. We have

$$\lim_{k\in K} \frac{(e^T d_N^k)^2}{\|d^k\|^2} = \lim_{k\in K} \frac{(e^T u_N^k)^2}{\|u^k\|^2} = |N|. \tag{71}$$

We show that

$$\lim_{k\in K} u_N^k = \frac{e}{|N|}, \quad \lim_{k\in K} u_B^k = 0. \tag{72}$$

The second relation has been already obtained in the proof of Lemma 4.4. To show the first relation, let us take an accumulation point $\hat{u}_N$ of $\{u_N^k\}_{k\in K}$, which exists due to Lemma 4.3. Taking note of (70), (71) and the second relation of (72), we have

$$e^T \bar{u}_N = \sqrt{|N|}, \quad \|\bar{u}_N\| = 1, \tag{73}$$

where $\bar{u}_N \equiv \hat{u}_N/\|\hat{u}_N\|$. It immediately follows $\bar{u}_N = e/\sqrt{|N|}$. Since $e^T\hat{u}_N = 1$ holds by Lemma 4.4, we see the first relation of (72).

Surpassing a subset of K if necessary, we may assume that

$$\frac{1}{2|N|} e \le u_N^k = \frac{X_N^k s_N(x^k)}{c^T x^k - c^\infty}, \qquad \forall k \in K. \tag{74}$$

Clearly, this relation implies that $s_N^k > 0$ for all $k \in K$. Hence, it follows from Lemma 4.1 that

$$\frac{e}{2|N|} \le \frac{X_N^k s_N^k}{c^T x^k - c^\infty} \le \frac{\|X_N^k\| s_N^k}{c^T x^k - c^\infty} \le \widetilde{M} s_N^k, \qquad \forall k \in K. \tag{75}$$

Since the function $s(x)$ is bounded (cf. Lemma of [26](page 118) or Proposition 2.8 of [15]), the sequence $\{s^k\}$ has an accumulation point s^* such that $s_N^* > 0$ due to (75). It is known that $X^* s^* = 0$ (see, e.g., Proposition 2.3(c) of [15] for a proof; we change step-size in the iterations, but the proof there is substantially extended easily to this case) and $x_B^* > 0$. Hence, x^* and s^* satisfies the strictly complementarity condition. This implies that x^* is a point lying in the relative interior of the optimal face and contradicts our assumption. ∎

Thus, the limiting point exists in the interior of the optimal face. We provide a few more preliminary lemmas for the superlinear convergence result.

Lemma 4.8 *We have*

$$\|w_{N_*}^k\| = g^k + O(c^T x^k - c^*). \tag{76}$$

Proof. By using Theorem 3.3 and Lemma 4.2, we see

$$\|r_{N_*}^k\| = O((c^T x^k - c^*)^2). \tag{77}$$

Further, from Lemma 3.2 and Lemma 4.2, we have

$$\begin{aligned} e^T u_{N_*}^k &= 1 + O((c^T x^k - c^*)^2), && (78) \\ \frac{1}{\|u_{N_*}^k\|} &\le \sqrt{2|N_*|} && (79) \end{aligned}$$

for sufficiently large k. On the other hand, Lemma 3.1 and Lemma 4.1 imply that

$$\frac{\|u_{B_*}^k\|}{\|u^k\|} \le \frac{\|u_{B_*}^k\|}{\|u_{N_*}^k\|} = O(c^T x^k - c^*). \tag{80}$$

Below we use Theorem 3.3. With the help of the relations obtained above, we have

$$\begin{aligned} \|w_{N_*}^k + r_{N_*}^k\|^2 &= \left\|\frac{u_{N_*}^k}{\|u^k\|^2} - e\right\|^2 \\ &\quad \text{(use Theorem 3.3.)} \\ &= \frac{\|u_{N_*}\|^2}{\|u^k\|^4} - 2e^T \frac{u_{N_*}^k}{\|u^k\|^2} + |N_*| \\ &= \frac{1}{\|u^k\|^2} - \frac{\|u_{B_*}\|^2}{\|u^k\|^4} - 2e^T \frac{u_{N_*}^k}{\|u^k\|^2} + |N_*| \\ &= -\frac{1}{\|u^k\|^2} + |N_*| + \frac{1}{\|u^k\|^2}\left(-\frac{\|u_{B_*}\|^2}{\|u^k\|^2} - 2(e^T u_{N_*}^k - 1)\right) \\ &= -\frac{1}{\|u^k\|^2} + |N_*| + O((c^T x^k - c^*)^2) && (81) \\ &\quad \text{(use (79), (80) and (78).)} \end{aligned}$$

if k is sufficiently large. From this relation and (77), we have

$$\|w^k_{N_*}\|^2 = (|N_*| - \frac{1}{\|u^k\|^2}) + O((c^T x^k - c^*)^2). \tag{82}$$

From Lemma 4.6, we have

$$\begin{aligned}
\frac{(e^T d^k_{N^k})^2}{\|d^k\|^2} &= \frac{(e^T d^k_{N_*})^2}{\|d^k\|^2} \\
&= \frac{(e^T u^k_{N_*})^2}{\|u^k\|^2} \\
&= \frac{1}{\|u^k\|^2} + \frac{(e^T u^k_{N_*})^2 - 1}{\|u^k\|^2} \\
&= \frac{1}{\|u^k\|^2} + O((c^T x^k - c^*)^2) \qquad (83)\\
& \text{(use (78) and (79).)}
\end{aligned}$$

for sufficiently large k. From (82) and (83),

$$\|w^k_{N_*}\|^2 = |N_*| - \frac{(e^T d^k_{N_*})^2}{\|d^k\|^2} + O((c^T x^k - c^*)^2) = (g^k)^2 + O((c^T x^k - c^*)^2) \tag{84}$$

holds when k is sufficiently large. Then it easily follows that

$$\|w^k_{N_*}\| = g^k + O(c^T x^k - c^*). \tag{85}$$

■

Lemma 4.9 *Let $\{x^k\}_{k\in K}$ be a subsequence of $\{x^k\}$ such that $\lim_{k\in K} \|w^k_{N_*}\| \to 0$. Then we have $\lim_{k\in K} v^k_{N_*} = v^*_{N_*}$.*

Proof. Since $v^*_{N_*}$ is the unique interior point of $\mathcal{V}^+_{N_*}$ where $w_{N_*}(v_{N_*}) = 0$ holds, we are done if we can show that every accumulation point of $\{v^k_{N_*}\}_{k\in K}$ is an interior point of $\mathcal{V}^+_{N_*}$.

By using Lemma 3.2, Theorem 3.3 and Lemma 4.2, we have

$$\lim_{k\in K} \frac{(e^T u^k_{N_*})^2}{\|u^k\|^2} = |N_*|. \tag{86}$$

Then, making the same argument as the one when we obtained (72) from (71), we have

$$\lim_{k\in K} u^k_{N_*} = \frac{e}{|N_*|}. \tag{87}$$

Applying Lemma 3.8, we see that every limit point of $\{v^k_{N_*}\}$ is an interior point of $\mathcal{V}^+_{N_*}$. ■

Lemma 4.10 *We have*

$$0 \le \delta(u^k) \le 1. \tag{88}$$

Proof. It is enough to show the second equality. Let $\tilde{x}$ be the point obtained if we take $\lambda^k = 1$. Since $c^T\tilde{x} - c^* \geq 0$, we may apply Lemma 3.5, to have

$$0 \leq \frac{c^T\tilde{x} - c^*}{c^T x^k - c^*} = 1 - \delta(u^k), \tag{89}$$

then the result is immediate. ∎

Now, we are ready to prove the main theorem.

Theorem 4.11 *The algorithm SLA generates a sequence of the objective function values that converges 2-step superlinearly to the optimal value with Q-order 1.3. The primal iterates x^k and the dual estimates s^k converge 2-step superlinearly with the same R-order to an interior point of the optimal face and the analytic center of the dual optimal face, respectively.*

Proof. The basic idea of the algorithm is as follows; If the iterate is sufficiently close to $\mathcal{M}_{N_*}$, we can expect large reduction in the objective function value, hence take a long step (~ 1); This long step may sacrifice closeness to $\mathcal{M}_{N_*}$, but it can be recovered by taking $1/2$ in the next step which tends to bring the iterates close to the $\mathcal{M}_{N_*}$, as is expected from Theorem 3.9 and Theorem 3.10. The quantity g^k measures closeness to $\mathcal{M}_{N_*}$ in the algorithm, namely, we take a long step if g^k is sufficiently small; otherwise we use step-size $1/2$. The iteration with a long step is referred to as a predictor step, and with step-size $1/2$ is referred to as a corrector step. Roughly speaking, our goal is to show that we have a predictor step at least every one of two succeeding iterations asymptotically.

First we show that a predictor step is taken infinitely many times. Suppose that we have at most finite number of predictor steps. Let k_0 be the last iteration count where the predictor step occurs. Due to Theorem 3.10,

$$\psi(v^k) \leq O(c^T x^k - c^*)^2. \tag{90}$$

This implies that

$$\begin{aligned} g^k &\leq \|w^k_{N_*}\| + M_0(c^T x^k - c^*) \\ &\quad \text{(use Lemma 4.8.)} \\ &\leq 1.1\psi(v^k) + M_0(c^T x^k - c^*) \\ &\quad \text{(use Lemma 3.6.)} \\ &\leq 0.9(c^T x^k - c^*)^{0.95} \\ &\leq |e^T d^k|^{0.95} \qquad\qquad (91) \\ &\quad \text{(use Lemma 4.4.)} \end{aligned}$$

if k is sufficiently large, where $M_0 > 0$ is a constant, that is, the condition for another predictor step is satisfied again, which is a contradiction. Thus we have an infinite number of predictor steps.

Now we analyze what occurs at each predictor and corrector step asymptotically.

(I)Analysis of each predictor step

We surpass the subsequence $\{x^k\}_{K_1}$, where K_1 is the set consisting of all k where a predictor step is taken. The condition (60) is satisfied for all $k \in K_1$. Due to Lemma 4.8, we have

$$\|w^k_{N_*}\| - M_1(c^Tx^k - c^*) \leq g^k < |e^Td^k|^{0.95}, \tag{92}$$

where M_1 is a positive constant. Since we have

$$|e^Td^k|^{0.95} \leq 1.1(c^Tx^k - c^*)^{0.95} \tag{93}$$

for sufficiently large k because of Lemma 4.4, the relation (92) implies that

$$\begin{aligned} \|w^k_{N_*}\| &\leq M_1(c^Tx^k - c^*) + |e^Td^k|^{0.95} \\ &\leq M_1(c^Tx^k - c^*) + 1.1(c^Tx^k - c^*)^{0.95} \\ &\leq 1.2(c^Tx^k - c^*)^{0.95} \end{aligned} \tag{94}$$

holds for sufficiently large $k \in K_1$. As shown in Lemma 4.9, $\lim_{k\in K_1} w^k_{N_*} \to 0$ implies $\lim_{k\in K_1} v^k_{N_*} = v^*_{N_*}$. Then we have, from Lemma 3.6 and (94),

$$\psi(v^k_{N_*}) \leq 1.1\|w^k_{N_*}\| \leq 1.4(c^Tx^k - c^*)^{0.95} \tag{95}$$

for sufficiently large $k \in K_1$. Thus $\lim_{k\in K_1} \psi(v^k_{N_*}) = 0$ and

$$\psi(v^k_{N_*}) \leq \varepsilon \tag{96}$$

hold for sufficiently large $k \in K_1$, where ε is the constant appeared in Lemma 3.7.

Next we show

$$1 - M_2(c^Tx^k - c^*)^{0.95} \leq \delta(u^k) = \frac{\|u^k\|^2}{\chi[u^k]} \leq 1 + M_2(c^Tx^k - c^*)^{0.95} \tag{97}$$

when $k \in K_1$ is sufficiently large, where $M_2 > 0$ is a constant. From Lemma 3.1, Lemma 3.2 and Lemma 4.3, we have

$$\chi[u^k] = \chi[u^k_{N_*}] \tag{98}$$

as $k \to \infty$. Due to Lemma 4.2, we see that $r^k_{N_*}$ appearing in Theorem 3.3 is bounded by

$$\|r^k_{N_*}\| \leq M_3(c^tx^k - c^*)^2 \tag{99}$$

when k is sufficiently large, where $M_3 > 0$ is a constant. Applying Theorem 3.3 taking (95) and (99) into account, we see

$$\frac{1}{\delta^k} = \frac{\chi[u^k]}{\|u^k\|^2} = \frac{\chi[u^k_{N_*}]}{\|u^k\|^2} = \chi[\frac{u^k_{N_*}}{\|u^k\|^2}] = \chi[e + w^k_{N_*} + r^k_{N_*}] = 1 + O(c^tx^k - c^*)^{0.95}, \tag{100}$$

from which (97) immediately follows for all $k \in K_1$ sufficiently large.

With (97) and the relations

$$\lambda^k = 1 - |e^Td^k|^{0.3}, \tag{101}$$

$$\frac{c^Tx^{k+1} - c^*}{c^Tx^k - c^*} = 1 - \lambda^k\delta^k \quad \text{(cf. Lemma 3.5)}, \tag{102}$$

$$0.9(c^Tx^k - c^*)^{0.3} \leq |e^Td^k|^{0.3} \quad \text{(cf. Lemma 4.4)}, \tag{103}$$

$$0 < \delta(u^k) \leq 1 \quad \text{(cf. Lemma 4.10)} \tag{104}$$

which hold when k is sufficeintly large, we have

$$\begin{aligned}
0.9(c^Tx^k - c^*)^{0.3} &\leq |e^Td^k|^{0.3} \\
&= 1 - \lambda^k \\
&\leq 1 - \lambda^k\delta(u^k) \\
&= \frac{c^Tx^{k+1} - c^*}{c^Tx^k - c^*} \\
&\leq 1 - (1 - |e^Td^k|^{0.3})(1 - M_2(c^tx^k - c^*)^{0.95}) \\
&\quad \text{(use (97).)} \\
&\leq 1 - (1 - 0.9(c^tx^k - c^*)^{0.3})(1 - M_2(c^tx^k - c^*)^{0.95}) \\
&= 0.9(c^Tx^k - c^*)^{0.3}(1 - M_2(c^Tx^k - c^*)^{0.95} + 0.9^{-1}M_2(c^Tx^k - c^*)^{0.65}) \\
&\leq (c^Tx^k - c^*)^{0.3} \qquad (105)
\end{aligned}$$

for sufficiently large $k \in K_1$, namely, we obtain

$$0.9(c^Tx^k - c^*)^{1.3} \leq c^Tx^{k+1} - c^* \leq (c^Tx^k - c^*)^{1.3}. \qquad (106)$$

(II) Analysis of each corrector step

Now, we show that we cannot have two consecutive corrector steps after a sufficiently large number of iterations. In view of (106), this gives 2-step superlinear convergence of the objective function value. Surpass a subsequence $\{x^k\}_{k \in K_2}$ where kth iteration is predictor step but $(k+1)$st iteration is a corrector step. Obviously $K_2 \subset K_1$. We show that at $(k+2)$nd iteration the condition (60) is satisfied to take another predictor step again if $k \in K_2$ is sufficiently large. In the analysis below, we assume $k \in K_2$ is large enough so that the relations obtained in (I) and

$$(c^Tx^k - c^*)^{0.65} \leq \frac{\varepsilon}{2} \qquad (107)$$

holds.

(i) Change in the objective function value in the kth iteration

Since kth iteration is a predictor step, (106) is satisfied.

(ii) Change in the objective function value in the $(k+1)$st iteration

Since $(k+1)$st iteration is a corrector step where $\lambda^{k+1} = 0.5$, we have, from (104) and (105), that

$$0.45(c^Tx^k - c^*)^{0.3} \leq \frac{c^Tx^{k+2} - c^*}{c^Tx^k - c^*} = (1 - \frac{\delta^{k+1}}{2})\frac{c^Tx^{k+1} - c^*}{c^Tx^k - c^*} \leq (c^Tx^k - c^*)^{0.3}, \qquad (108)$$

or equivalently,

$$0.45(c^Tx^k - c^*)^{1.3} \leq (c^Tx^{k+2} - c^*) \leq (c^Tx^k - c^*)^{1.3}. \qquad (109)$$

(iii) Change in the measure of centrality ψ in the kth iteration

We apply Theorem 3.4. Due to (96), we can use Lemma 3.7. By using (101), (102), (103), (104) and (99), $\psi(v^{k+1})$ is bounded from above as follows for sufficiently large $k \in K_2$:

$$
\begin{aligned}
\psi(v_{N_*}^{k+1}) &= \|(V_{N_*}^*)^{-1}(v_{N_*}^k - v_{N_*}^* - \frac{\lambda^k\delta^k}{1-\lambda^k\delta^k}(dv_{N_*}^k + V_{N_*}^k r_{N_*}^k))\| \\
&\quad \text{(use Theorem 3.4.)} \\
&= \|(V_{N_*}^*)^{-1}(v_{N_*}^k - v_{N_*}^* - \frac{\lambda^k\delta^k}{1-\lambda^k\delta^k}(v_{N_*}^k - v_{N_*}^* + \Delta v_{N_*}^k + V_{N_*}^k r_{N_*}^k))\| \\
&\quad (\Delta v_{N_*}^k \equiv dv_{N_*}^k - (v_{N_*}^k - v_{N_*}^*).) \\
&= \|(V_{N_*}^*)^{-1}(\frac{1-2\lambda^k\delta^k}{1-\lambda^k\delta^k}(v_{N_*}^k - v_{N_*}^*) - \frac{\lambda^k\delta^k}{1-\lambda^k\delta^k}(\Delta v_{N_*}^k + r_{N_*}^k))\| \\
&\le \frac{1}{1-\lambda^k}\|(V_{N_*}^*)^{-1}(v_{N_*}^k - v_{N_*}^*)\| + \frac{1}{1-\lambda^k}\|(V_{N_*}^*)^{-1}(\Delta v_{N_*}^k + V_{N_*}^k r_{N_*}^k)\| \\
&\quad \text{(use } 0 \le \delta^k \le 1 \text{ and } 0 \le \lambda^k \le 1.) \\
&\le \frac{1}{|e^T d^k|^{0.3}}(\psi(v_{N_*}^k) + \|(V_{N_*}^*)^{-1}\Delta v_{N_*}^k\| + \|(V_{N_*}^*)^{-1}\|\|V_{N_*}^k\|\|r_{N_*}^k\|) \\
&\quad \text{(use } 1-\lambda^k = |e^t d^k|^{0.3}.) \\
&\le \frac{1}{0.9(c^Tx^k - c^*)^{0.3}}(\psi(v_{N_*}^k) + M\psi(v_{N_*}^k)^2 + 2M_3\|(V_{N_*}^*)^{-1}\|\|V_{N_*}^*\|(c^Tx^k - c^*)^2) \\
&\quad \text{(use (103), Lemma 3.7 and (99).)} \\
&\le \frac{1}{0.9(c^Tx^k - c^*)^{0.3}} \times \\
&\qquad (1.4(c^Tx^k - c^*)^{0.95} + 1.4^2(c^Tx^k - c^*)^2 M + 2M_3\|(V_{N_*}^*)^{-1}\|\|V_{N_*}^*\|(c^Tx^k - c^*)^2) \\
&\quad \text{(use (95).)} \\
&= \frac{(c^Tx^k - c^*)^{0.65}}{0.9}(1.4 + 1.4^2(c^Tx^k - c^*)^{1.05}M + 2M_3\|(V_{N_*}^*)^{-1}\|\|V_{N_*}^*\|(c^Tx^k - c^*)^{1.05}) \\
&\le 2(c^Tx^k - c^*)^{0.65}. \qquad (110) \\
&\quad \text{(holds when } k \text{ is sufficiently large.)}
\end{aligned}
$$

(iv) Recovery in the measure of centrality ψ in the $(k+1)$st iteration

Since (110) and (107) holds, we have

$$\psi(v_{N_*}^{k+1}) \le \varepsilon. \qquad (111)$$

In the similar manner as we obtained in (97), we have

$$1 - M_4(c^Tx^k - c^*)^{0.65} \le \delta^{k+1} = \frac{\|u^{k+1}\|^2}{\chi[u^{k+1}]} \le 1 + M_4(c^Tx^k - c^*)^{0.65} \qquad (112)$$

for sufficiently large $k \in K_2$, where $M_4 > 0$ is a constant. Then, letting

$$\eta^{k+1} \equiv 1 - \frac{\delta^{k+1}/2}{1-\delta^{k+1}/2}, \qquad (113)$$

we have

$$\eta^{k+1} \leq M_5(c^T x^k - c^*)^{0.65}, \tag{114}$$

where $M_5 > 0$ is a constant. Again we apply Theorem 3.4 and Lemma 3.7, and obtain

$$\begin{aligned}
\psi(v_{N_*}^{k+2}) &= \|(V_{N_*}^*)^{-1}(v_{N_*}^{k+2} - v_{N_*}^*)\| \\
&= \|(V_{N_*}^*)^{-1}(v_{N_*}^{k+1} - v_{N_*}^* - (1-\eta^{k+1})(dv_{N_*}^{k+1} + V_{N_*}^{k+1} r_{N_*}^{k+1}))\| \\
&\leq \|(V_{N_*}^*)^{-1}(v_{N_*}^{k+1} - v_{N_*}^* - dv_{N_*}^{k+1})\| + \eta^{k+1}\|(V_{N_*}^*)^{-1} dv_{N_*}^{k+1}\| + \|(V_{N_*}^*)^{-1}\|\|V_{N_*}^{k+1}\|\|r_{N_*}^{k+1}\|. \\
&\leq M\psi(v_{N_*}^{k+1})^2 + 2\eta^{k+1}\psi(v^{k+1}) + \|(V_{N_*}^*)^{-1}\|\|V_{N_*}^{k+1}\|\|r_{N_*}^{k+1}\|, \\
&\quad \text{(apply Lemma 3.7.)} \\
&\leq (4M + 4M_5)(c^T x^k - c^*)^{1.3} + 2\|(V_{N_*}^*)^{-1}\|\|V_{N_*}^*\| M_3 (c^T x^k - c^*)^2, \\
&\quad \text{(use (110), (114) and (99).)} \\
&\leq 5(M + M_5)(c^T x^k - c^*)^{1.3} \\
&\quad \text{(holds if } k \text{ is sufficiently large.)} \\
&\leq \frac{5(M + M_5)}{0.45}(c^T x^{k+2} - c^*) \\
&\quad \text{(use (109).)} \\
&= M_6(c^T x^{k+2} - c^*),
\end{aligned} \tag{115}$$

where $M_6 \equiv 5(M + M_5)/0.45$.

(v) Relationship between g^{k+2} and $|e^T d^{k+2}|$

From (115), we have $\psi(v_{N_*}^{k+2}) \to 0$ as $k \in K_2$ tends to infinity. Using Lemma 3.6 and (115),

$$\|w_{N_*}^{k+2}\| \leq 1.1\psi(v_{N_*}^{k+2}) \leq 1.1 M_6(c^T x^{k+2} - c^*) \tag{116}$$

for sufficiently large $k \in K_2$. By using Lemma 4.8, we have

$$g^{k+2} \leq M_7(c^T x^{k+2} - c^*), \tag{117}$$

where M_7 is a constant. On the other hand, due to Lemma 4.4,

$$0.9(c^T x^{k+2} - c^*)^{0.95} < |e^T d^{k+2}|^{0.95} \tag{118}$$

for sufficiently large k. Comparing (117) and (118), we see

$$g^{k+2} \leq M_7(c^T x^{k+2} - c^*) \leq 0.9(c^T x^{k+2} - c^*)^{0.95} < |e^T d^{k+2}|^{0.95} \tag{119}$$

is satisfied if $c^T x^{k+2} - c^*$ is sufficiently small, i.e., $k \in K_2$ is sufficiently large. This means that (60) is satisfied at $(k+2)$nd iteration to take another predictor step again if $k \in K_2$ is sufficiently large, and we are done.

(III) Superlinear convergence property of the sequence

Now, we know that every one of two consecutive steps is a predictor step asymptotically. Then 2-step superlinear convergence of the objective function value with Q-order 1.3 is seen from (106). R-superlinear convergence of x^k with R-order 1.3 follows easily from Theorem 2.6 of [15]. Finally, we show that R-superlinear convergence of s^k to the analytic center of the dual optimal face. Since we have (95) and (110), it is not difficult to see that $v^k_{N_*}$ converges 2-step superlinearly to $v^*_{N_*}$ with R-order 1.3. By using the relation $s^k = (V^k)^{-1}u^k$ and Theorem 3.3, we see the limiting point s^* of s^k is,

$$s^* = (s^*_{N_*}, s^*_{B_*}) = ((V^*_{N_*})^{-1}e/|N_*|, 0). \tag{120}$$

It is not difficult to check that this point is the analytic center of the dual optimal face (cf. Section 4 of [15]). Due to Theorem 3.3, Lemma 4.2, Lemma 3.1, we have

$$\begin{aligned} \|s^k - s^*\| &= \|((V^k_{N_*})^{-1}u^k_{N_*} - s^*_{N_*}, (V^k_B)^{-1}u^k_{B_*} - 0)\| \\ &= \|((V^k_{N_*})^{-1}u^k_{N_*} - (V^*_{N_*})^{-1}e/|N_*|, (V^k_B)^{-1}u^k_{B_*})\| \\ &\le \|((V^k_{N_*})^{-1}u^k_{N_*} - (V^k_{N_*})^{-1}e/|N_*|, (V^k_B)^{-1}u^k_{B_*})\| \\ &\quad + \|(V^k_{N_*})^{-1}e/|N_*| - (V^*_{N_*})^{-1}e/|N_*|\| \\ &\le M_8(c^Tx^k - c^*)^2 + M_9\psi(v^k_{N_*}), \end{aligned} \tag{121}$$

where M_8 and M_9 are positive constants, from which R-superlinear convergence of $\|s^k - s^*\|$ with R-order 1.3 follows. ∎

5 Technical Lemmas

In this section we prove lemmas and theorems mentioned at Section 3.1. We use the same notations as in Section 3.1.

We start with the proof of Lemma 3.1.

Proof of Lemma 3.1 Since $\bar{s} = (\bar{s}_N, 0) \in c + Range(A^T)$, we know from Lemma 2.1 that $dx(x)$ solves problem

$$\begin{aligned} &\text{maximize}_{\,p} \quad \tfrac{1}{2}\bar{s}_N^T p_N - \|X^{-1}p\|^2 \\ &\text{subject to} \quad Ap = 0. \end{aligned} \tag{122}$$

It follows that $dx_B(x)$ solves the problem

$$\begin{aligned} &\text{minimize}_{\,p_B} \quad \tfrac{1}{2}\|X_B^{-1}p_B\|^2 \\ &\text{subject to} \quad A_Bp_B = -A_Ndx_N(x). \end{aligned} \tag{123}$$

Now, since $A_Bp_B = -A_Ndx_N(x)$ has a solution $dx_B(x)$, it should have a solution $\bar{p}_B \in \mathbb{R}^{|B|}$ such that

$$\|\bar{p}_B\| \le C\|dx_N(x)\| \tag{124}$$

where C is a constant independent of x. Hence, we have

$$\begin{aligned} \|d_B(x)\| &= \|X_B^{-1}dx_B(x)\| \\ &\le \|X_B^{-1}\bar{p}_B\| \\ &\le \|X_B^{-1}\|\|\bar{p}_B\| \\ &\le \|X_B^{-1}\|C\|dx_N(x)\| \\ &\le C\|X_B^{-1}\|\|X_N\|\|X_N^{-1}dx_N(x)\| \\ &= C\|X_B^{-1}\|\|X_N\|\|d_N(x)\|. \end{aligned} \tag{125}$$

The lemma immediately follows from this result. ∎

Now we prove Lemma 3.2 and Theorem 3.3. Let us consider the LP problem

$$\begin{array}{ll} \text{minimize}_x & c^T x \\ \text{subject to} & A_N x_N + A_B x_B = b, \qquad x_N \geq 0, \end{array} \tag{126}$$

obtained by removing the constraint $x_B \geq 0$ from (1). Due to the relation (7) and $A_B \tilde{x}_B = b$ for some $\tilde{x}$ on the dual degenerate face, this problem can be written in the following form:

$$\begin{array}{ll} \text{minimize}_x & \bar{s}_N^T x_N \\ \text{subject to} & A_N x_N \in Range(A_B), \qquad x_N \geq 0. \end{array} \tag{127}$$

An auxiliary search direction, called the *homogeneous affine scaling direction*, plays an important role in the proofs. The homogeneous affine scaling direction $\widetilde{dx}_N(x_N)$ is defined as the affine scaling search direction for the problem (127) as follows:

$$\begin{array}{ll} \text{maximize}_{p_N} & \bar{s}_N^T p_N - \frac{1}{2}\|X_N^{-1} p_N\|^2 \\ \text{subject to} & A_N p_N \in Range(A_B). \end{array} \tag{128}$$

We define

$$\tilde{u}_N(x) = \frac{d_N(x)}{\bar{s}_N^T x_N}. \tag{129}$$

In the remaining part of this section, we concentrate our efforts to prove the following three lemmas.

Lemma 5.1 *We have*

$$\frac{\tilde{u}_N(x)}{\|\tilde{u}_N(x)\|^2} = w_N(v_N(x)) + e \tag{130}$$

in $\mathcal{Q}_N^{++}$.

Lemma 5.2 *We have*

$$e^T \tilde{u}_N(x_N) = 1 \tag{131}$$

in $\mathcal{Q}_N^{++}$.

Lemma 5.3 *We have*

$$\tilde{u}_N(x) = u_N(x) + \tilde{r}_N(x), \tag{132}$$

where $\tilde{r}_N(x) = O(\|X_B^{-1}\|^2 \|X_N\|^2 \|v_N(x)\|)$ *in* $\mathcal{Q}_N^{++}$.

Befor going into their proofs, we give the proofs of Lemma 3.2 and Theorem 3.3. Lemma 3.2 is immediate from Lemma 5.2 and Lemma 5.3. We prove Theorem 3.3 below.

Proof of Theorem 3.3 In view of Lemma 5.1, it is enough to check that

$$\frac{u_N(x)}{\|u(x)\|^2} = \frac{\tilde{u}_N(x)}{\|\tilde{u}_N(x)\|^2} + O(\|(X_B)^{-1}\|^2 \|X_N\|^2 \|v_{N*}\|). \tag{133}$$

Using Lemma 5.3, we have

$$\begin{aligned}
\frac{u_N(x)}{\|u(x)\|^2} &= \frac{\tilde{u}_N(x)+\tilde{r}_N(x)}{\|u(x)\|^2} \\
&= \frac{\tilde{u}_N(x)}{\|u(x)\|^2}+\frac{\tilde{r}_N(x)}{\|u(x)\|^2} \\
&= \frac{\tilde{u}_N(x)}{\|\tilde{u}_N(x)\|^2}\frac{\|\tilde{u}_N(x)\|^2}{\|u(x)\|^2}+\frac{\tilde{r}_N(x)}{\|u(x)\|^2} \\
&= \frac{\tilde{u}_N(x)}{\|\tilde{u}_N(x)\|^2}+\frac{\tilde{u}_N(x)}{\|\tilde{u}_N(x)\|^2}(\alpha(x)\beta(x)-1)+\frac{\tilde{r}_N(x)}{\|u(x)\|^2}.
\end{aligned} \tag{134}$$

where

$$\alpha(x)\equiv\frac{\|\tilde{u}_N(x)\|^2}{\|u_N(x)\|^2}=\frac{\|\tilde{u}_N(x)\|^2}{\|\tilde{u}_N(x)+\tilde{r}_N(x)\|^2},\quad \beta(x)\equiv\frac{\|u_N(x)\|^2}{\|u(x)\|^2}=\frac{\|u_N(x)\|^2}{\|u_N(x)\|^2+\|u_B(x)\|^2}. \tag{135}$$

Due to Lemma 5.2 and Lemma 5.3, we have

$$\frac{1}{\|\tilde{u}_N(x)\|^2}\le 2|N|,\quad \frac{1}{\|u_N(x)\|^2}\le 2|N|+O(\|(X_B)^{-1}\|^2\|X_N\|^2\|v_N(x)\|). \tag{136}$$

From these relations, we see the last term $\tilde{r}_N/\|u\|^2$ in (134) is a quantity of $O(\|(X_B)^{-1}\|^2$ $\|X_N\|^2\|v_N(x)\|)$.

It is remaining to show that the term

$$\frac{\tilde{u}_N}{\|\tilde{u}_N(x)\|}(\alpha(x)\beta(x)-1) \tag{137}$$

is a quantity of $O(\|(X_B)^{-1}\|^2\|X_N\|^2\|v_N(x)\|)$.

Because of (136), it is enough to show

$$\alpha(x)\beta(x)=1+O(\|(X_B)^{-1}\|^2\|X_N\|^2\|v_N(x)\|). \tag{138}$$

Taking the relation

$$\frac{1}{\|\bar{s}_N\|}\le\frac{\|x_N\|}{\bar{s}_N^T x_N}=\|v_N(x)\| \tag{139}$$

into account, we have

$$\alpha(x)=\frac{\|\tilde{u}_N(x)\|^2}{\|\tilde{u}_N(x)+\tilde{r}_N(x)\|^2}\le\frac{1}{(1-\|\tilde{r}_N(x)\|/\|\tilde{u}_N(x)\|)^2}=1+O(\|X_N\|^2\|(X_B)^{-1}\|^2\|v_N(x)\|) \tag{140}$$

and

$$\beta(x)=\frac{\|u_N(x)\|^2}{\|u_N(x)\|^2+\|u_B(x)\|^2}=1+O(\|X_N\|^2\|(X_B)^{-1}\|^2)=1+O(\|X_N\|^2\|(X_B)^{-1}\|^2\|v_N(x)\|). \tag{141}$$

This completes the proof. ■

Now we are going to prove the three lemmas stated above.

Proof of Lemma 5.1 The linear space

$$\{x|A_Nx_N \in \ Range(A_B)\} \tag{142}$$

is rewritten as

$$\{x|\tilde{A}_Nx_N = 0\} \tag{143}$$

with an appropriate row full rank matrix $\tilde{A}_N$.

Then $\widetilde{dx}_N(x)$ is written as the optimizer of

$$\begin{array}{ll} \text{maximize}_{\,p_N} & \bar{s}_N^T p_N - \frac{1}{2}\|X_N^{-1}p_N\|^2 \\ \text{subject to} & \tilde{A}_N p_N = 0, \end{array} \tag{144}$$

whereas $dv_N(v_N(x))$ is written as the optimizer of

$$\begin{array}{ll} \text{minimize}_{\,q_N} & e^T V_N^{-1} q_N + \frac{1}{2}\|V_N^{-1}q_N\|^2 \\ \text{subject to} & \tilde{A}_N p_N = 0, \bar{s}_N^T q_N = 0, \end{array} \tag{145}$$

where $V_N \equiv diag(v_N(x))$.

It is not difficult to see that the explicit representation of $\widetilde{dx}_N(x)$ is given by:

$$\widetilde{dx}_N(x) = X_N(I - X_N\tilde{A}_N^T(\tilde{A}_N X_N^2 \tilde{A}_N^T)^{-1}\tilde{A}_N X_N)X_N\bar{s}_N, \tag{146}$$

and hence

$$\tilde{u}_N(x) = (I - V_N\tilde{A}_N^T(\tilde{A}_N V_N^2 \tilde{A}_N^T)^{-1}\tilde{A}_N V_N)V_N\bar{s}_N. \tag{147}$$

As to $dv_N(v_N)$, by taking note of $\bar{s}_N \notin Range(\tilde{A}_N)$, $\tilde{A}_N v_N = 0$ and $\bar{s}_N^T v_N = 1$, we have

$$\begin{aligned} dv_N(v_N) &= V_N\left(I - V_N\begin{pmatrix} \tilde{A}_N^T & \bar{s}_N \end{pmatrix}\begin{pmatrix} \tilde{A}_N V_N^2 \tilde{A}_N^T & \tilde{A}_N V_N^2 \bar{s}_N \\ \bar{s}_N^T V_N^2 \tilde{A}_N^T & \bar{s}_N^T V_N^2 \bar{s}_N \end{pmatrix}^{-1}\begin{pmatrix} \tilde{A}_N \\ \bar{s}_N^T \end{pmatrix} V_N\right)e \\ &= v_N - V_N\begin{pmatrix} \tilde{A}_N^T & \bar{s}_N \end{pmatrix}\begin{pmatrix} \tilde{A}_N V_N^2 \tilde{A}_N^T & \tilde{A}_N V_N^2 \bar{s}_N \\ \bar{s}_N^T V_N^2 \tilde{A}_N^T & \bar{s}_N^T V_N^2 \bar{s}_N \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= v_N - V_N\begin{pmatrix} \tilde{A}_N^T & \bar{s}_N \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \end{aligned} \tag{148}$$

where α, β are the solutions of the equation

$$\begin{pmatrix} \tilde{A}_N V_N^2 \tilde{A}_N^T & \tilde{A}_N V_N^2 \bar{s}_N \\ \bar{s}_N^T V_N^2 \tilde{A}_N^T & \bar{s}_N^T V_N^2 \bar{s}_N \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{149}$$

By block manipulation of the matrix, we see that

$$\alpha = -(\tilde{A}_N V_N^2 \tilde{A}_N^T)^{-1}\tilde{A}_N V_N^2 \bar{s}_N \beta, \tag{150}$$

$$\beta = \frac{1}{\bar{s}_N^T V_N \left(I - V_N \tilde{A}_N^T(\tilde{A}_N V_N^2 \tilde{A}_N^T)^{-1}\tilde{A}_N V_N\right) V_N \bar{s}_N}. \tag{151}$$

From this result, it is easy to see that

$$dv_N(v_N) = v_N - \frac{V_N\left(I - V_N\tilde{A}^T(\tilde{A}_N V_N^2 \tilde{A}_N^T)^{-1}\tilde{A}_N V_N\right)V_N\bar{s}_N}{\bar{s}_N^T V_N\left(I - V_N\tilde{A}_N^T(\tilde{A}_N V_N^2 \tilde{A}_N^T)^{-1}\tilde{A}_N V_N\right)V_N\bar{s}_N}. \tag{152}$$

Now substituting the definition of $\tilde{u}_N(x)$, we see that

$$dv_N(v_N(x)) = v_N(x) - V_N(x)\frac{\tilde{u}_N(x)}{\|\tilde{u}_N(x)\|^2}, \tag{153}$$

from which the desired result immediately follows. ∎

Next we prove Lemma 5.2 and Lemma 5.3.

Proof of Lemma 5.2 Since $\widetilde{dx}_N(x)$ solves problem (128), it must satisfy the optimality conditions for (128). That is, there exist $\tilde{\eta} = \tilde{\eta}(x) \in \mathbb{R}^m$ and $\widetilde{dx}_B = \widetilde{dx}_B(x) \in \mathbb{R}^{|B|}$ such that

$$\begin{aligned} \bar{s}_N - X_N^{-2}\widetilde{dx}_N(x) - A_N^T\tilde{\eta} &= 0, && (154)\\ A_N\widetilde{dx}_N(x) &= -A_B\widetilde{dx}_B, && (155)\\ A_B^T\tilde{\eta} &= 0. && (156)\end{aligned}$$

Let $\tilde{x}$ be an arbitrary point of the dual degenerate face. Then, $\tilde{x}_N = 0$ and hence $A_B\tilde{x}_B = b$. Thus, if $x \in \mathcal{P}^+$ we obtain

$$A_N x_N = b - A_B x_B = A_B\tilde{x}_B - A_B x_B = A_B(\tilde{x}_B - x_B). \tag{157}$$

Using this relation and relation (154), we obtain

$$e^T(X_N)^{-1}\widetilde{dx}_N(x) = \bar{s}_N^T x_N - x_N^T A_N^T\tilde{\eta} = \bar{s}_N^T x_N - (\tilde{x}_B - x_B)^T A_B^T\tilde{\eta} = \bar{s}_N^T x_N \tag{158}$$

where the last equality follows from relation (156). Using relations (158), we obtain

$$e_N^T\tilde{u}(x) = \frac{e^T(X_N)^{-1}\widetilde{dx}_N(x)}{\bar{s}_N^T x_N} = 1. \tag{159}$$

∎

The following result provides a preliminary relation between the affine scaling direction $dx(x)$ and the homogeneous affine scaling direction $\widetilde{dx}_N(x)$.

Lemma 5.4 *The vector $(p_B, p_N) \equiv (dx_B(x), dx_N(x) - \widetilde{dx}_N(x))$ is the unique solution of the following QP problem:*

$$\begin{aligned} &\textit{minimize}_{(p_B,p_N)} && \tfrac{1}{2}\|X_B^{-1}p_B\|^2 + \tfrac{1}{2}\|X_N^{-1}p_N\|^2\\ &\textit{subject to} && A_B p_B + A_N p_N = -A_N\widetilde{dx}_N(x). \end{aligned} \tag{160}$$

Proof. The vector $(dx_B(x),\, dx_N(x) - \widetilde{dx}_N(x))$ is clearly feasible for problem (160). To prove that $(dx_B(x),\, dx_N(x) - \widetilde{dx}_N(x))$ is optimal for (160), it is sufficient to show that there exists $y \in \mathbb{R}^m$ such that

$$\begin{aligned} X_B^{-2} dx_B(x) - A_B^T y &= 0, & (161)\\ X_N^{-2}(dx_N(x) - \widetilde{dx}_N(x)) - A_N^T y &= 0. & (162) \end{aligned}$$

We already know that $\widetilde{dx}_N(x)$ and some $\tilde{\eta} \in \mathbb{R}^m$ satisfy (154) and (156). Since $dx(x)$ solves problem (4) (with $\tilde{s} = \bar{s}$ we know that there exists $\eta = \eta(x) \in \mathbb{R}^m$ such that

$$\begin{aligned} X_B^{-2} dx_B(x) - A_B^T \eta &= 0, & (163)\\ \bar{s}_N + X_N^{-2} dx_N(x) - A_N^T \eta &= 0. & (164) \end{aligned}$$

Combining relations (154), (156), (163) and (164), we obtain

$$\begin{aligned} X_B^{-2} dx_B(x) - A_B^T(\eta - \tilde{\eta}) &= 0 & (165)\\ X_N^{-2}(dx_N(x) - \widetilde{dx}_N(x)) - A_N^T(\eta - \tilde{\eta}) &= 0 & (166) \end{aligned}$$

which gives relations (161) if we define $y = \eta - \tilde{\eta}$. ■

Now, we are ready to prove Lemma 5.3. For the purpose of simplifying notation, we define

$$\tilde{d}_N(x) \equiv X_N^{-1} \widetilde{dx}_N(x). \tag{167}$$

Proof of Lemma 5.3 Fix $x > 0$ and define $\tau_N \equiv dx_N(x) - \widetilde{dx}_N(x)$. Clearly, $A_N \tau_N \in Range(A_B)$. The system $A_B p_B = -A_N \tau_N$ has a solution $p_B = \tau_B$ such that

$$\|\tau_B\| \le C \|\tau_N\|. \tag{168}$$

Define $\widetilde{dx}_B(x) \equiv dx_B(x) - \tau_B$. Then, it is easy to see that

$$A_B \widetilde{dx}_B(x) = -A_N \widetilde{dx}_N(x). \tag{169}$$

This relation implies that $(p_B, p_N) = (\widetilde{dx}_B(x), 0)$ is feasible to problem (160). Using Lemma 5.4, we obtain

$$\|X_B^{-1} dx_B(x)\|^2 + \|X_N^{-1}(dx_N(x) - \widetilde{dx}_N(x))\|^2 \le \|X_B^{-1} \widetilde{dx}_B(x)\|^2. \tag{170}$$

Hence, we obtain

$$\begin{aligned} \|d_N(x) - \tilde{d}_N(x)\|^2 &= \|X_N^{-1}(dx_N(x) - \widetilde{dx}_N(x))\|^2 \\ &\le \|X_B^{-1} \widetilde{dx}_B(x)\|^2 - \|X_B^{-1} dx_B(x)\|^2 \\ &= [X_B^{-1}(\widetilde{dx}_B(x) + dx_B(x))]^T [X_B^{-1}(\widetilde{dx}_B(x) - dx_B(x))] \\ &\le \|X_B^{-1}(dx_B(x) + \widetilde{dx}_B(x))\| \|X_B^{-1}(\widetilde{dx}_B(x) - dx_B(x))\|. \qquad (171) \end{aligned}$$

We now bound each of the terms of the last expression in (171).

In view of Lemma 2.1, we may replace c with $\bar{s}$ in the definition (3c) of $d(x)$. Taking note of $\bar{s}_B = 0$, we have

$$\|d(x)\| \leq \|X_N \bar{s}_N\|. \tag{172}$$

Similarly, we obtain

$$\|\tilde{d}(x)\| \leq \|X_N \bar{s}_N\|. \tag{173}$$

Using relations Lemma 3.1, (172), (173) and (168), we obtain

$$\begin{aligned} \|X_B^{-1}(dx_B(x) + \widetilde{dx}_B(x))\| &\leq \|2X_B^{-1} dx_B(x) - X_B^{-1}\tau_B\| \\ &\leq 2\|d_B(x)\| + \|X_B^{-1}\tau_B\| \\ &\leq 2C\|X_B^{-1}\|\|X_N\|\|d_N(x)\| + C\|\tau_N\|\|X_B^{-1}\| \\ &\leq C\|X_B^{-1}\|\|X_N\|[2\|d_N(x)\| + \|d_N(x) - \tilde{d}_N(x)\|] \\ &\leq 4C\|X_B^{-1}\|\|X_N\|\|X_N\bar{s}_N\|. \end{aligned} \tag{174}$$

We also have that

$$\begin{aligned} \|X_B^{-1}(dx_B(x) - \widetilde{dx}_B(x))\| &\leq \|X_B^{-1}\| \, \|dx_B(x) - \widetilde{dx}_B(x)\| \\ &\leq \|X_B^{-1}\| \, \|\tau_B\| \\ &\leq C\|X_N^{-1}\| \, \|\tau_N\| \\ &\leq C\|X_B^{-1}\|\|dx_N(x) - \widetilde{dx}_N(x)\| \\ &\leq C\|X_B^{-1}\|\|X_N\| \, \|d_N(x) - \tilde{d}_N(x)\|. \end{aligned} \tag{175}$$

Combining relations (171), (174) and (175), we obtain

$$\|d_N(x) - \tilde{d}_N(x)\|^2 \leq 4C^2\|X_B^{-1}\|^2\|X_N\|^2\|X_N\bar{s}_N\|\|d_N(x) - \tilde{d}_N(x)\|. \tag{176}$$

Since $c^T x - c' = \bar{s}_N^T x_N$, we obtain the desired relation (132) by dividing both sides by $(c^T x - c')\|d_N(x) - \tilde{d}_N(x)\|$. ∎

6 Concluding Remarks

So far, we demonstrated that the affine scaling algorithm can have superlinear convergence property with an $Q(R)$-order 1.3. Practical efficiency of this algorithm is not clear at this moment, but we believe that this result has some interest in its own right, since this is an unexpected result in view of the nature of the affine scaling algorithm as a steepest descent method. Now we conclude this paper with some remarks for further research.

One interesting problem is to improve the order of convergence. The order 1.3 is not tight in this paper and we can obtain a better bound even without changing major part of this algorithm. However, if we want to come up with any convergence order less than or equal two as many primal-dual algorithms (e.g. [29] and [13]) and Iri and Imai's algorithm [23] enjoy, it would be necessary to make some more devices, even if it is possible.

This analysis can be directly applied to the long-step variant of Karmarkar's algorithm [11] analyzed in [16]. It should be possible to show that Karmarkar's algorithm can enjoy superlinear convergence as well without sacrificing its polynomial complexity by adopting the step-size choice proposed here (with a slight modification).

Another algorithm that is related to this analysis is Todd's low complexity algorithm [19]. Todd's low complexity algorithm uses the affine scaling direction with step-size shorter than 1/2 in its predictor step (it uses the step-size 1/5 in terms of the ellipsoid in the scaled space). As we showed in the paper, the affine scaling direction can work as a kind of corrector step towards the central trajectory when $\lambda^k \leq 1/2$, and this may give a good explanation for why the algorithm does not need corrector step asymptotically (cf. [20]). It looks possible to apply our analysis to modify this algorithm so that it can have superlinear convergence property without sacrificing good polynomial complexity. This is an interesting topic for further research.

Appendix

The notations in this appendix has local meaning. Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and consider a bounded polyhedron $\mathcal{T} \equiv \{x \in \mathbb{R}^n | Ax = b,\ x \geq 0\}$ whose interior is nonempty. The analytic center is the optimal solution for the following problem:

$$\begin{array}{l} \text{minimize } -\sum_{i=1,n} \log x_i. \\ \text{subject to } Ax = b, \quad x \geq 0. \end{array} \tag{177}$$

The Newton direction $dx(x)$ for the analytic center of $\mathcal{T}$ at an interior point x is given by the optimal solution for the following problem:

$$\begin{array}{l} \text{minimize}_{\ p}\ f_x(p) \\ \text{subject to } Ap = 0, \end{array} \tag{178}$$

where

$$f_x(p) \equiv e^T X^{-1} p + \frac{1}{2}\|X^{-1}p\|^2, \tag{179}$$

$X \equiv diag(x)$. The Newton iteration is written as

$$x^+ = x - dx(x), \tag{180}$$

where x^+ is the next iterate. We define the scaled Newton direction as:

$$d(x) \equiv X^{-1} dx(x). \tag{181}$$

Under these settings, we prove the following proposition. We think that this proposition is enough to understand "(Property of the Newton direction for the analytic center)" in page 11.

Proposition *Let $\bar{x}$ be a point on the boundary of $\mathcal{T}$, and let N, B be the index sets where $\bar{x}_N = 0$ and $\bar{x}_B > 0$. Then we have*

$$\lim_{x \to \bar{x}} d_N(x) = -e. \tag{182}$$

Proof. It is well-known that $\|d(x)\| \leq \sqrt{n}$ for any interior point of $\mathcal{T}$. By contradiction, assume that (182) does not hold. Then we can take a sequence $\{x^l\}$ of interior points of $\mathcal{T}$, where $\lim_{l\to\infty} x^l = \bar{x}, \lim_{l\to\infty} d(x^l) = -\tilde{d}$ and $\tilde{d}_N \neq e$. We simply denote $d^l \equiv d(x^l)$ and $dx^l \equiv dx(x^l)$. Let

$$f_x^N(p_N) \equiv e^T X_N^{-1} p_N + \frac{1}{2}\|X_N^{-1} p_N\|^2, \tag{183}$$

$$f_x^B(p_B) \equiv e^T X_B^{-1} p_B + \frac{1}{2}\|X_B^{-1} p_B\|^2. \tag{184}$$

Then, we have

$$f_x(p) = f_x^N(p_N) + f_x^B(p_B). \tag{185}$$

Since d^l converges to $-\tilde{d}$, each function has its limiting value as $l \to \infty$. Let

$$f_*^N \equiv \lim_{l\to\infty} f_{x^l}^N(dx_N^l), \quad f_*^B \equiv \lim_{l\to\infty} f_{x^l}^B(dx_B^l). \tag{186}$$

Since $\tilde{d}_N^l \neq e$, we see

$$f_*^N > -\frac{1}{2}|N|. \tag{187}$$

Now, it is easy to check $(x_N^l, (x_B^l - \bar{x}_B))$ is a feasible solution for (178). Hence, $dx_B^l - (x_B^l - \bar{x}_B)$ sasisfy the following equation with respect to τ_B:

$$A_B \tau_B = -A_N dx_N^l + A_N x_N^l. \tag{188}$$

Obviously, this equation has a solution $\hat{\tau}_B^l$ such that

$$\|\hat{\tau}_B^l\| \leq C(\|dx_N^l\| + \|x_N^l\|) \leq C(\sqrt{n}+1)\|x_N^l\|, \tag{189}$$

where C is a constant determined only from A (and the partition (N, B)) and we used the inequality

$$dx_N^l \leq \|X_N^l\|\|d^l\| \leq \sqrt{n}\|x_N^l\|. \tag{190}$$

Now, letting

$$\widetilde{dx}_N^l \equiv -x_N^l, \quad \widetilde{dx}_B^l \equiv \hat{\tau}_B^l + dx_B^l, \tag{191}$$

we easily see that $\widetilde{dx}^l$ is a feasible solution for (178) where

$$f_{x^l}^N(\widetilde{dx}_N^l) = -\frac{1}{2}|N|, \quad \lim_{l\to\infty} f_{x^l}^B(\widetilde{dx}_B^l) = f_*^B. \tag{192}$$

The limit in the second relation holds since $\bar{x}_B > 0$, $\lim_{l\to\infty} \hat{\tau}_B^l = 0$ due to (189). In view of (187), this means that for sufficiently large l, we have

$$f_{x^l}(\widetilde{dx}^l) < f_{x^l}(dx^l), \tag{193}$$

which, however, is a contradiction to the definition of dx^l. This completes the proof. ∎

Acknowledgment

This research was carried out while the first author was visiting the Center for Research on Parallel Computation and the Department of Computational and Applied Mathematics of Rice University, Houston, USA. He thanks his host Prof. J. E. Dennis and the colleagues there for providing the excellent research environment during his stay at Rice University.

Part of this research was supported by the National Science Foundation (NSF) under Grant No. DDM-9109404 and the Office of Naval Research (ONR) under contract 92-SIE-316. The second author thanks NSF for the financial support received during the completion of this work.

The first author also thanks the second author for having arranged and supported a trip with his NSF grant to visit the University of Arizona where this research was initiated. He also thanks the University of Arizona for the congenial scientific atmosphere that it provided.

References

[1] I. Adler, N. K. Karmarkar, M. G. C. Resende, and G. Veiga. Data structures and programming techniques for the implementation of Karmarkar's algorithm. *ORSA Journal on Computing*, 1:84–106, 1989.

[2] I. Adler, N. K. Karmarkar, M. G. C. Resende, and G. Veiga. An implementation of Karmarkar's algorithm for linear programming. *Mathematical Programming*, 44:297–335, 1989. Errata in *Mathematical Programming*, 50:415, 1991.

[3] I. Adler and R. D. C. Monteiro. Limiting behavior of the affine scaling continuous trajectories for linear programming problems. *Mathematical Programming*, 50:29–51, 1991.

[4] E. R. Barnes. A variation on Karmarkar's algorithm for solving linear programming problems. *Mathematical Programming*, 36:174–182, 1986.

[5] Yun-Chan Cheng, David J. Houck, Jun-Min Liu, Marc S. Meketon, Lev Slutsman, Robert J. Vanderbei, and Pyng Wang. The AT&T KORBX System. *AT&T Technical Journal*, 68(3):7–19, 1989.

[6] I. I. Dikin. Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk SSSR*, 174:747–748, 1967. Translated in : *Soviet Mathematics Doklady*, 8:674–675, 1967.

[7] I. I. Dikin. On the convergence of an iterative process. *Upravlyaemye Sistemi*, 12:54–60, 1974. (In Russian).

[8] I. I. Dikin. The convergence of dual variables. Technical Report, Siberian Energy Institute, Irkutsk, Russia, December 1991.

[9] C. C. Gonzaga. Convergence of the large step primal affine–scaling algorithm for primal nondegenerate linear programs. Technical Report ES–230/90, Dept. of Systems

Engineering and Computer Science, COPPE Federal University of Rio de Janeiro, 21941 Rio de Janeiro, RJ, Brazil, September 1990.

[10] L. A. Hall and R. J. Vanderbei. Two-thirds is sharp for affine scaling. Technical Report SOR–92/9, Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey, 08544, September 1992.

[11] N. K. Karmarkar. A new polynomial–time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[12] N. Megiddo and M. Shub. Boundary behavior of interior point algorithms in linear programming. *Mathematics of Operations Research*, 14:97–114, 1989.

[13] S. Mehrotra. Quadratic convergence in a primal-dual method. Technical Report 91–15, Dept. of Industrial Engineering and Management Science, Northwestern University, Evanston, IL 60208, USA, September 1991.

[14] C. L. Monma and A. J. Morton. Computational experience with the dual affine variant of Karmarkar's method for linear programming. *Operations Research Letters*, 6:261–267, 1987.

[15] R. D. C. Monteiro, T. Tsuchiya, and Y. Wang. A simplified global convergence proof of the affine scaling algorithm. Technical Report, Dept. of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA, October 1992.

[16] M. Muramatsu and T. Tsuchiya. A convergence analysis of a long-step variant of the projective scaling algorithm. Research Memorandum 454, The Institute of Statistical Mathematics, 4–6–7 Minami–Azabu, Minato–ku, Tokyo 106, Japan, Nobemver 1992.

[17] M. G. C. Resende and G. Veiga. An efficient implementation of a network interior point method. Working Paper, AT&T Bell Laboratory, Murray Hill, NJ, USA, Feburary 1992.

[18] L. Sinha, B. Freedman, N. Karmarkar, A. Putcha, and K. Ramakrishnan. Overseas network planning. *Proceedings of "the Third International Network Planning Sysmposium – Networks' 86" (IEEE Communications Society, held on June 1-6, 1986, Tarpon Springs, Florida, USA)*, pages 121–124, 1986.

[19] M. Todd. A low complexity interior-point algorithm for linear programming. *SIAM Journal on Optimization*, 2:198–209, 1992.

[20] M. Todd and J.-P. Vial. Todd's complexity interior-point algorithm is a predictor-corrector path-following algorithm. *Operations Research Letter*, 11:199–207, 1992.

[21] P. Tseng and Z. Q. Luo. On the convergence of the affine–scaling algorithm. Technical Report CICS–P–169, Center for Intelligent Control Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, 1989.

[22] T. Tsuchiya. Global convergence of the affine–scaling methods for degenerate linear programming problems. *Mathematical Programming*, 52:377–404, 1991.

[23] T. Tsuchiya. Quadratic convergence of Iri and Imai's method for degenerate linear programming problems. Research Memorandum 412, The Institute of Statistical Mathematics, 4–6–7 Minami–Azabu, Minato–ku, Tokyo 106, Japan, 1991.

[24] T. Tsuchiya. Global convergence property of the affine scaling method for primal degenerate linear programming problems. *Mathematics of Operations Research*, 17(3):527–557, 1992.

[25] T. Tsuchiya and M. Muramatsu. Global convergence of a long–step affine scaling algorithm for degenerate linear programming problems. Research Memorandum 423, The Institute of Statistical Mathematics, 4–6–7 Minami–Azabu, Minato–ku, Tokyo 106, Japan, January 1992 (revised, September, 1992).

[26] R. J. Vanderbei and J. C. Lagarias. I. I. Dikin's convergence result for the affine–scaling algorithm. In J. C. Lagarias and M. J. Todd, editors, *Mathematical Developments Arising from Linear Programming : Proceedings of a Joint Summer Research Conference held at Bowdoin College, Brunswick, Maine, USA, June/July 1988*, volume 114 of *Contemporary Mathematics*, pages 109–119. American Mathematical Society, Providence, Rhode Island, USA, 1990.

[27] R. J. Vanderbei, M. S. Meketon, and B. A. Freedman. A modification of Karmarkar's linear programming algorithm. *Algorithmica*, 1(4):395–407, 1986.

[28] C. Witzgall, P. T. Boggs, and P. D. Domich. On the convergence behavior of trajectories for linear programming. In J. C. Lagarias and M. J. Todd, editors, *Mathematical Developments Arising from Linear Programming : Proceedings of a Joint Summer Research Conference held at Bowdoin College, Brunswick, Maine, USA, June/July 1988*, volume 114 of *Contemporary Mathematics*, pages 161–187. American Mathematical Society, Providence, Rhode Island, USA, 1990.

[29] Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$–iteration algorithm for linear programming. Technical Report TR–91–26, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, August 1991.