



FLORIAN POTRA

**A path-following method for linear complementarity
problems based on the affine invariant Kantorovich
Theorem**

A path-following method for linear complementarity problems based on the affine invariant Kantorovich Theorem ^{*}

Florian A. Potra[†]

August, 2000

Abstract A path following algorithm for linear complementarity problems is presented. Given a point z that approximates a point $z(\tau)$ on the central path with complementarity gap τ , one determines a parameter $\theta \in (0, 1)$ so that this point satisfies the hypothesis of the affine invariant Kantorovich Theorem for the equation defining $z((1 - \theta)\tau)$. It is shown that θ is bounded below by a multiple of $n^{-1/2}$, where n is the dimension of the problem. Since the hypothesis of the Kantorovich Theorem is satisfied the sequence generated by Newton's method, or by the simplified Newton method, will converge to $z((1 - \theta)\tau)$. We show that the number of steps required to obtain an acceptable approximation of $z((1 - \theta)\tau)$ is bounded above by a number independent of n . Therefore the algorithm has $O(\sqrt{n}L)$ -iteration complexity. The parameters of the algorithm can be determined in such a way that only one Newton step is needed each time the complementarity gap is decreased.

Keywords: Linear complementarity problem, interior-point algorithm, path-following, Kantorovich Theorem

Abbreviated title: Path following with Kantorovich

AMS (MOS) subject classification: 49M15, 65H10, 65K05, 90C33.

^{*}This work was supported in part by NSF, Grant DMS-9996154 .

[†]Department of Mathematics and Statistics, University of Maryland Baltimore County. Email: potra@math.umbc.edu.

1 Introduction

Theoretical and experimental work done over the past decade has demonstrated that primal-dual path following algorithms are among the best solution methods for linear programming (LP), quadratic programming (QP), and Linear Complementarity Problems (LCP) (see for example the excellent monograph of Steve Wright [16] and the recent survey paper [10]). On one hand primal-dual path following methods form the basis of the best general purpose practical algorithms, and on the other hand they have interesting theoretical properties. For example, it has been proved that a large class of primal-dual path following algorithms can solve linear programming problems with rational data in $O(\sqrt{n}L)$ iterations, where L is the length of a binary coding of the input data and n is the dimension of the problem. This is the best computational complexity result obtained so far in the literature. Although the same computational complexity result can be obtained by using other classes of interior-point methods (for example potential reduction methods), path following algorithms have a number of other attractive properties that distinguishes them among interior-point methods (see the recent monographs [14, 16, 17]).

In the present paper we show that the affine invariant Kantorovich Theorem [2] can be used for constructing a class of path following algorithm for LCP that have $O(\sqrt{n}L)$ iteration complexity. Given a point z that approximates a point $z(\tau)$ on the central path of the LCP with complementarity gap τ , the algorithms compute a parameter $\theta \in (0, 1)$ so that z satisfies the hypothesis of the affine invariant form of the Kantorovich Theorem for the equation defining $z((1-\theta)\tau)$. It is shown that θ is bounded below by a multiple of $n^{-1/2}$, where n is the dimension of the problem. Since the hypothesis of the Kantorovich Theorem is satisfied the sequence generated by Newton's method, or the simplified Newton method, with starting point z will converge to $z((1-\theta)\tau)$. We show that the number of Newton (or simplified Newton) steps required to obtain an acceptable approximation of $z((1-\theta)\tau)$ is bounded above by a number independent of n . Therefore a point with complementarity gap less than ϵ can be obtained in at most $O(\sqrt{n} \log(\epsilon_0/\epsilon))$ Newton and simplified Newton steps, where ϵ_0 is the complementarity gap of the starting point. For problems with rational input data of bitlength L this implies that an exact solution can be obtained in at most $O(\sqrt{n}L)$ iterations plus a rounding procedure involving $O(n^3)$ arithmetic operations (see [16]).

The use of Newton method techniques in the study of interior point methods goes back to the celebrated paper of Renegar [12]. The affine invariant Kantorovich theorem was used by Renegar and Shub [13] to prove $O(\sqrt{n}L)$ -iteration complexity for some Newton based interior point methods for LP and QP. We note that Renegar and Shub used the affine invariant Kantorovich Theorem to prove a variant of Smale's theorem [15], which was then applied to analyze the respective interior point methods. The approach of the present paper is different in that it uses the Kantorovich Theorem directly. More importantly, we compute

the **optimal** decrease of the duality gap that satisfies the hypothesis of the Kantorovich Theorem instead of using a fixed decrease of the duality gap.

The paper is organized as follows. In Section 2 we state the affine invariant Kantorovich Theorem for general nonlinear equations in \mathbb{R}^n . In Section 3 we show how the Kantorovich Theorem can be used to construct a path-following algorithm for linear complementarity problems. Section 4 contains some technical lemmas that are used in the proof of our main result. In the last section we prove that our path following algorithm has $O(\sqrt{n}L)$ -iteration complexity. We show that the parameters of the algorithm can be determined in such a way that only one Newton step is needed each time the complementarity gap is decreased. Cases where a given number of Newton steps or simplified Newton steps are needed are also discussed.

Conventions. We denote by \mathbb{N} the set of all nonnegative integers. \mathbb{R} , \mathbb{R}_+ , \mathbb{R}_{++} denote the set of real, nonnegative real, and positive real numbers respectively. For any real number γ , $\lceil \gamma \rceil$ denotes the smallest integer greater or equal to γ . Given a vector x , the corresponding upper case symbol denotes as usual the diagonal matrix X defined by the vector. The symbol e represents the vector of all ones, with dimension given by the context.

We denote component-wise operations on vectors by the usual notations for real numbers. Thus, given two vectors u, v of the same dimension, uv , u/v , etc. will denote the vectors with components $u_i v_i$, u_i/v_i , etc. This notation is consistent as long as component-wise operations always have precedence in relation to matrix operations. Note that $uv \equiv Uv$ and if A is a matrix, then $Auv \equiv AUv$, but in general $Auv \neq (Au)v$. Also if f is a scalar function and v is a vector, then $f(v)$ denotes the vector with components $f(v_i)$. For example if $v \in \mathbb{R}_+^n$, then \sqrt{v} denotes the vector with components $\sqrt{v_i}$, and $1-v$ denotes the vector with components $1 - v_i$. Traditionally the vector $1 - v$ is written as $e - v$, where e is the vector of all ones. Inequalities are to be understood in a similar fashion. For example if $v \in \mathbb{R}^n$ then $v \geq 3$ means that $v_i \geq 3$, $i = 1, \dots, n$. Traditionally this is written as $v \geq 3e$. If $\|\cdot\|$ is a vector norm on \mathbb{R}^n and A is a matrix then the operator norm induced by $\|\cdot\|$ is defined by $\|A\| = \max\{\|Ax\|; \|x\| = 1\}$. As a particular case we note that if U is the diagonal matrix defined by the vector u , then $\|U\|_2 = \|u\|_\infty$. Finally we want to introduce a less standard notation. If $x, s \in \mathbb{R}^n$ then the vector $z \in \mathbb{R}^{2n}$ obtained by concatenating x and s will be denoted by $\lceil x, s \rceil$, i.e.,

$$z = \lceil x, s \rceil = \begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} x^T & s^T \end{bmatrix}^T. \quad (1.1)$$

2 The Kantorovich Theorem

The celebrated Kantorovich Theorem gives sufficient conditions for the existence and uniqueness of a solution of nonlinear operator equation in a Banach space and shows that under that condition the sequences generated by Newton's method and the simplified Newton method converge to the solution. In what follows we will present this theorem only in a finite dimensional setting. For our application it is essential that we use the "affine-invariant" version of the Kantorovich Theorem whose importance was stressed in the paper of Deuffhard and Heindl [2]. As it will become apparent from Section 3, the traditional form of the Kantorovich Theorem, as presented for example in the classical monograph of Ortega and Rheinboldt [7], is inadequate for our application.

Let $F : D \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$ be a nonlinear operator defined on a domain D of the p -dimensional linear space \mathbb{R}^p with values in \mathbb{R}^p . Let $\|\cdot\|$ be a given norm on \mathbb{R}^p , and let z be a point of D such that the closed ball of radius ρ centered at z ,

$$\overline{B}(z, \rho) := \{y \in \mathbb{R}^p : \|y - z\| \leq \rho\}, \quad (2.1)$$

is included in D , i.e.,

$$\overline{B}(z, \rho) \subset D. \quad (2.2)$$

We assume that the Jacobian $F'(z)$ is nonsingular and that the following affine-invariant Lipschitz condition is satisfied

$$\|F'(z)^{-1}(F'(y) - F'(\bar{y}))\| \leq \omega \|y - \bar{y}\|, \forall y, \bar{y} \in \overline{B}(z, \rho). \quad (2.3)$$

The Kantorovich Theorem essentially states that if the quantity

$$\alpha := \|F'(z)^{-1}F(z)\| \quad (2.4)$$

is small enough, in the sense that

$$\kappa := \alpha\omega \leq \frac{1}{2} \quad (2.5)$$

then there is a z^* with $F(z^*) = 0$ and the sequences produced by Newton's method

$$z_{\mathcal{N}}^0 = z, z_{\mathcal{N}}^{k+1} = z_{\mathcal{N}}^k - F'(z_{\mathcal{N}}^k)^{-1}F(z_{\mathcal{N}}^k), k = 0, 1, \dots \quad (2.6)$$

and by the simplified Newton method

$$z_{\mathcal{S}}^0 = z, z_{\mathcal{S}}^{k+1} = z_{\mathcal{S}}^k - F'(z)^{-1}F(z_{\mathcal{S}}^k), k = 0, 1, \dots \quad (2.7)$$

are well defined and converge to z^* . More precisely we have

Theorem 2.1 (*The Kantorovich Theorem*) Let $F : D \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$ be a differentiable mapping and let $z \in D$ be such that $F'(z)$ is nonsingular. Assume that conditions (2.3) – (2.5) are satisfied and that the radius ρ is large enough in the sense that

$$\hat{\rho} := \frac{1 - \sqrt{1 - 2\kappa}}{\omega} \leq \rho. \quad (2.8)$$

Then:

1. F has a zero z^* in the closed ball $\overline{B}(z, \hat{\rho})$;
2. The open ball $B(z, \check{\rho})$ with radius

$$\check{\rho} := \frac{1 + \sqrt{1 - 2\kappa}}{\omega} \quad (2.9)$$

does not contain any zero of F different from z^* ;

3. The iterative procedures (2.6) and (2.7) produce sequences belonging to the open ball $B(z, \hat{\rho})$ that converge to z^* ;
4. If $\kappa < 1/2$ then for Newton's method we have the following estimates

$$\|z_{\mathcal{N}}^k - z^*\| \leq \frac{2\beta \lambda^{2^k}}{1 - \lambda^{2^k}}, \quad (2.10)$$

with

$$\beta = \frac{\sqrt{1 - 2\kappa}}{\omega}, \quad \lambda = \frac{1 - \sqrt{1 - 2\kappa} - \kappa}{\kappa}, \quad k = 1, 2, \dots, \quad (2.11)$$

and for the simplified Newton methods we have

$$\|z_{\mathcal{S}}^k - z^*\| \leq \frac{2\beta \lambda^2}{1 - \lambda^2} \xi^{k-1}, \quad k = 1, 2, \dots \quad (2.12)$$

where

$$\xi = 1 - \sqrt{1 - 2\kappa}. \quad (2.13)$$

Proof. The above theorem is essentially the theorem presented in [2] with the error bounds (2.10) from [9]. The only thing that remains to be proved is (2.12). In order to prove these error bounds we write

$$z_{\mathcal{S}}^{k+1} - z_{\mathcal{S}}^* = -F'(z)^{-1} \left[F(z_{\mathcal{S}}^k) - F(z_{\mathcal{S}}^*) - F'(z)(z_{\mathcal{S}}^k - z_{\mathcal{S}}^*) \right].$$

Using (2.3) we deduce that

$$\|z_{\mathcal{S}}^{k+1} - z_{\mathcal{S}}^*\| \leq \frac{1}{2}\omega(\|z_{\mathcal{S}}^k - z\| + \|z_{\mathcal{S}}^* - z\|) \|z_{\mathcal{S}}^k - z_{\mathcal{S}}^*\|.$$

Since $z_{\mathcal{S}}^k, z_{\mathcal{S}}^{k-1} \in B(z, \hat{\rho})$ it follows that

$$\|z_{\mathcal{S}}^{k+1} - z_{\mathcal{S}}^*\| \leq \omega \hat{\rho} \|z_{\mathcal{S}}^k - z_{\mathcal{S}}^*\| = \xi \|z_{\mathcal{S}}^k - z_{\mathcal{S}}^*\|, \quad k = 1, 2, \dots$$

Using (2.10) for the first Newton step we obtain

$$\|z_{\mathcal{S}}^k - z^*\| \leq \|z_{\mathcal{S}}^1 - z^*\| \xi^{k-1} \leq \frac{2\beta\lambda^2}{1-\lambda^2} \xi^{k-1}, \quad k = 1, 2, \dots$$

■

It is known (see [9]) that the estimates (2.10) are sharp in the sense that for any α and ω with $\alpha\omega < .5$ there is a function that satisfies the hypothesis of the Kantorovich Theorem for which the estimates are verified with equality for all k . For the traditional form of the Kantorovich sharp apriori bounds were obtained in [8] and [3]. The equivalence between the bounds of [8] and [3] is proved in [9].

The estimates (2.12) are not sharp in this sense described above, but they are sufficiently good for our purpose. The sharp estimates for the simplified Newton method given in [9] are not in explicit form and therefore they are more difficult to use in complexity analysis. In most textbooks (e.g. [7]) the Lipschitz condition (2.3) is replaced by the following two conditions:

$$\begin{aligned} \|F'(z)^{-1}\| &\leq \zeta, \\ \|(F'(y) - F'(w))\| &\leq \eta \|y - w\|, \quad \forall y, w \in \overline{B}(z, \rho). \end{aligned}$$

If these conditions holds then (2.3) is clearly satisfied with $\omega = \zeta\eta$. However, in our application estimating ζ and η would be much more difficult and the product $\zeta\eta$ would severely overestimate ω which would make it impossible to obtain the desired complexity result.

3 The Linear Complementarity Problem

Given two matrices $Q, R \in \mathbb{R}^{n \times n}$ and a vector $b \in \mathbb{R}^n$, the horizontal linear complementarity problem (HLCP) consists in finding a pair of vectors (x, s) such that

$$\begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\geq 0. \end{aligned} \tag{3.1}$$

The standard (monotone) linear complementarity problem (LCP) is obtained by taking $R = -I$ and Q positive semidefinite. The linear programming problem (LP) and the quadratic programming problem (QP) can be formulated as an HLCP. Therefore HLCP provides a convenient and general framework for studying interior point methods.

Throughout this paper we assume that the HLCP (3.1) is monotone in the sense that:

$$Qu + Rv = 0 \text{ implies } u^T v \geq 0, \text{ for any } u, v \in \mathbb{R}^n. \quad (3.2)$$

This condition is satisfied if the HLCP is a reformulation of a QP and for many other interesting classes of problems (see the excellent monograph [1]). If the HLCP is a reformulation of a LP then the following stronger condition holds

$$Qu + Rv = 0 \text{ implies } u^T v = 0, \text{ for any } u, v \in \mathbb{R}^n. \quad (3.3)$$

In this case we say that the HLCP is skew-symmetric. Since in the skew-symmetric case we can often obtain sharper estimates we are going to consider this particular case as well in this paper.

Since HLCP is uninteresting for $n = 1$ we will assume throughout this paper that $n \geq 2$.

It is known (see [6]) that if the HLCP has an interior point (i.e., there is $(x, s) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ satisfying $Qx + Rs = b$), then for any parameter $\tau > 0$ the nonlinear system:

$$\begin{aligned} xs &= \tau e \\ Qx + Rs &= b \\ x, s &\geq 0, \end{aligned} \quad (3.4)$$

where $e = (1, 1, \dots, 1)^T$ is the vector of all ones, has a unique positive solution $z(\tau) = [x(\tau), s(\tau)]$. The set of all such solutions defines the central path \mathcal{C} of the HLCP. It can be proved that $(x(\tau), s(\tau))$ converges to a solution of the HLCP as $\tau \rightarrow 0$. Therefore one could numerically solve the HLCP by following the central path as $\tau \rightarrow 0$. An implementation of this approach for solving the HLCP is called a path following algorithm.

At a basic step of a path following algorithm an approximation (x, s) of $(x(\tau), s(\tau))$ has already been computed for some $\tau > 0$. Then the algorithm determines a smaller value of the central path parameter

$$\tau_+ = (1 - \theta)\tau, \quad (3.5)$$

where the value $\theta \in (0, 1)$ is computed by some procedure. Finally one computes an approximation (x^+, s^+) of $(x(\tau_+), s(\tau_+))$ and the procedure is repeated with (x^+, s^+, τ_+) instead of (x, s, τ) .

In what follows we will show how the above basic step of a path following algorithm can be implemented by using the Kantorovich theorem. It is convenient to introduce the

notations

$$z = \begin{bmatrix} x \\ s \end{bmatrix}, z(\tau) = \begin{bmatrix} x(\tau) \\ s(\tau) \end{bmatrix}, z^+ = \begin{bmatrix} x^+ \\ s^+ \end{bmatrix}, z(\tau_+) = \begin{bmatrix} x(\tau_+) \\ s(\tau_+) \end{bmatrix}, \text{ etc.} \quad (3.6)$$

For any $\sigma > 0$ we define the nonlinear operator

$$F_\sigma(z) = \begin{bmatrix} xs - \sigma e \\ Qx + Rs - b \end{bmatrix}. \quad (3.7)$$

The equation (3.4) defining $z(\tau)$ becomes

$$F_\tau(z) = 0, \quad (3.8)$$

while the equation defining $z(\tau_+)$ can be written as

$$F_{(1-\theta)\tau}(z) = 0. \quad (3.9)$$

We assume that the starting point z belongs to the interior of the feasible set of the HLCP

$$\mathcal{F}^0 := \{z = [x, s] \in \mathbb{R}_{++}^{2n}; Qx + Rs = b\}. \quad (3.10)$$

Since both Newton's method and the simplified Newton method solve linear equations exactly, all the points generated by those methods will satisfy the linear equation from (3.10). Moreover the iterates will remain strictly positive, and hence will belong to \mathcal{F}^0 .

Suppose that the starting point $z \in \mathcal{F}^0$ approximates in a certain sense a point $z(\tau)$ on the central path for some given $\tau > 0$. We want to determine the largest possible parameter $\theta \in (0, 1)$ such that z satisfies the Kantorovich Theorem for (3.9). In order to do that we have to estimate the quantities α and ω that appear in Theorem 2.1. Since the Jacobian of F_σ does not depend on σ we will denote

$$F'(z) := F'_\sigma(z) = \begin{bmatrix} S & X \\ Q & R \end{bmatrix}. \quad (3.11)$$

In order to estimate the Lipschitz constant ω from the Kantorovich Theorem we notice that

$$L := F'(z)^{-1}(F'(y) - F'(\bar{y})) = F'(z)^{-1} \begin{bmatrix} V & U \\ 0 & 0 \end{bmatrix}, \quad (3.12)$$

where U, V are the diagonal matrices corresponding to the n -dimensional vectors u, v defined by

$$[u, v] := y - \bar{y}.$$

For any $\bar{z} \in \mathbb{R}^{2n}$ we denote

$$t = L\bar{z}, \quad t = [t_x, t_s], \quad \bar{z} = [\bar{z}_x, \bar{z}_s]. \quad (3.13)$$

It follows that the n -dimensional vectors t_x, t_s are solutions of the linear system

$$\begin{aligned} st_x + xt_s &= u\bar{z}_s + v\bar{z}_x \\ Qt_x + Rt_s &= 0. \end{aligned}$$

From Lemma 4.1 (to be proved in the following section) it follows that

$$\|Dt_x\|_2^2 + \|D^{-1}t_s\|_2^2 \leq (\min xs)^{-1} \|u\bar{z}_s + v\bar{z}_x\|_2^2, \quad (3.14)$$

where D is the diagonal matrix given by

$$D = X^{-1/2}S^{1/2}. \quad (3.15)$$

Therefore it is convenient to introduce the following norm on \mathbb{R}^{2n} :

$$\|q\|_z := \sqrt{\|Dq_x\|_2^2 + \|D^{-1}q_s\|_2^2}, \quad \forall q = [q_x, q_s] \in \mathbb{R}^{2n}. \quad (3.16)$$

This norm depends on the current point z and therefore it is called a local norm. The significance of local norms to interior point methods theory is beautifully illustrated in the very recent monograph of Renegar [11].

We majorize the right-hand side of (3.14) noticing that by using a two-dimensional Cauchy inequality and the above definition of the norm we can write

$$\begin{aligned} \|u\bar{z}_s + v\bar{z}_x\|_2 &\leq \|u\bar{z}_s\|_2 + \|v\bar{z}_x\|_2 \\ &\leq \|Du\|_2 \|D^{-1}\bar{z}_s\|_2 + \|D^{-1}v\|_2 \|D\bar{z}_x\|_2 \\ &\leq \sqrt{\|Du\|_2^2 + \|D^{-1}v\|_2^2} \sqrt{\|D\bar{z}_x\|_2^2 + \|D^{-1}\bar{z}_s\|_2^2} \\ &= \|y - w\|_z \|\bar{z}\|_z. \end{aligned}$$

Now from (3.12),(3.13),(3.14) and (3.17) we deduce that

$$\|L\bar{z}\|_z \leq (\min xs)^{-1/2} \|y - w\|_z \|\bar{z}\|_z.$$

Hence the operator norm of L induced by our vector norm satisfies

$$\|L\|_z \leq (\min xs)^{-1/2} \|y - w\|_z.$$

Thus we have obtained the following result:

Lemma 3.1 *If $z \in \mathcal{F}^0$ and $F'(z)$ is defined by (3.11) then*

$$\|F'(z)^{-1}(F'(y) - F'(w))\|_z \leq \omega(z)\|y - w\|_z, \forall y, w \in \mathbb{R}^{2n}$$

where the Lipschitz constant $\omega(z)$ is given by

$$\omega(z) = \frac{1}{\sqrt{\min xs}}. \quad (3.17)$$

In order to verify the hypothesis of the Kantorovich Theorem for the equation $F_\tau(z) = 0$ we also need to consider the quantity

$$\alpha(z, \tau) := \|F'(z)^{-1}F_\tau(z)\|_z. \quad (3.18)$$

It is convenient to introduce the Kantorovich measure of proximity

$$\kappa(z, \tau) := \alpha(z, \tau)\omega(z) \quad (3.19)$$

as well as the optimality measure (the normalized primal-dual gap)

$$\mu := \mu(z) := \frac{x^T s}{n}. \quad (3.20)$$

If for a given interior point z and a given positive parameter τ we have $\kappa(z, \tau) \leq .5$ then the Kantorovich Theorem is satisfied and the sequences generated by Newton's method and the simplified Newton method with starting point z will converge to the point $z(\tau)$ on the central path. We are ready to describe now our algorithm.

ALGORITHM 1

Given $0 < \kappa_1 < \kappa_2 < .5$, $\epsilon > 0$, and $z^0 \in \mathcal{F}^0$ satisfying $\kappa(z^0, \mu(z^0)) \leq \kappa_1$;

Set $k \leftarrow 0$ and $\tau_0 \leftarrow \mu(z^0)$;

repeat(outer iteration)

Set $(z, \tau) \leftarrow (z^k, \tau_k)$, $\bar{z} \leftarrow z^k$;

Determine the largest $\theta \in (0, 1)$ such that $\kappa(z, (1 - \theta)\tau) \leq \kappa_2$;

Set $\tau \leftarrow (1 - \theta)\tau$, $\bar{z} \leftarrow z$;

repeat(inner iteration)

Set $z \leftarrow z - F'(z)^{-1}F_\tau(z)$, for Newton's method

(or $z \leftarrow z - F'(\bar{z})^{-1}F_\tau(z)$ for the simplified Newton);

until $\kappa(z, \tau) \leq \kappa_1$;

Set $(z^{k+1}, \tau_{k+1}) \leftarrow (z, \tau)$;

until $(x^{k+1})^T s^{k+1} \leq \epsilon$.

In Section 5 we will prove that the parameter θ computed in the outer iteration is bounded below by a quantity of the form χ/\sqrt{n} , where χ is a positive quantity depending only on κ_1 and κ_2 . Since the hypothesis of the Kantorovich Theorem is satisfied the inner iteration will terminate in finite number of steps. Therefore the above algorithm is globally convergent. More precisely we will prove that the inner iteration will terminate in at most m steps where the number m depends only on κ_1 and κ_2 . Therefore the total number of Newton and simplified Newton steps required for Algorithm 1 is at most $O(\sqrt{n} \log((x^0)^T s^0)/\epsilon)$.

4 Some Technical Results

In this section we present some simple inequalities that will be used in the proof of our main result. First we bound the size of the solution of the following linear system.

$$\begin{aligned} su + xv &= a \\ Qu + Rv &= 0 \end{aligned} \quad (4.1)$$

where $z = [x, s] \in \mathbb{R}_{++}^{2n}$ and $a \in \mathbb{R}^n$ are given vectors. In order to treat together the monotone case and the skew-symmetric case it is convenient to introduce the following notation.

$$\varsigma := \begin{cases} \sqrt{2}, & \text{if HLCP is monotone} \\ 1 & \text{if HLCP is skew-symmetric} \end{cases} \quad (4.2)$$

Lemma 4.1 *The solution of (4.1) satisfies*

$$\frac{1}{\varsigma} \left\| \frac{a}{\sqrt{xs}} \right\|_2 \leq \left\| \begin{bmatrix} u \\ v \end{bmatrix} \right\|_z \leq \left\| \frac{a}{\sqrt{xs}} \right\|_2.$$

Proof. By dividing the first equation in (4.1) with \sqrt{xs} we obtain:

$$Du + D^{-1}v = a/\sqrt{xs},$$

where the diagonal matrix D is given by (3.15). It follows that

$$\|Du\|_2^2 + \|D^{-1}v\|_2^2 + 2u^T v = \|a/\sqrt{xs}\|_2^2.$$

The desired inequalities are obtained by noticing that in the skew-symmetric case we have $u^T v = 0$ while in the monotone case we have

$$0 \leq u^T v \leq 2 \|Du\|_2 \|D^{-1}v\|_2 \leq \|Du\|_2^2 + \|D^{-1}v\|_2^2.$$

■

The above lemma allows us to establish a relationship in between $\alpha(z, \tau)$ defined in (3.18) and the proximity measure

$$\delta(z, \tau) := \left\| \sqrt{\frac{\tau}{xs}} - \sqrt{\frac{xs}{\tau}} \right\|_2 \quad (4.3)$$

considered in [4] (see also [14]).

Corollary 4.2 For any $z \in \mathbb{R}_{++}^{2n}$ and any $\tau > 0$ we have

$$\frac{\sqrt{\tau}}{\varsigma} \delta(z, \tau) \leq \alpha(z, \tau) \leq \sqrt{\tau} \delta(z, \tau),$$

where ς is given by (4.2).

Now we establish a relation between $\delta(z, \tau)$ and $\delta(z, \mu(z))$.

Lemma 4.3 The following identity holds for any $z \in \mathbb{R}_{++}^{2n}$ and any $\tau > 0$

$$\delta^2(z, \tau) = \frac{\tau}{\mu} \delta^2(z, \mu(z)) + \frac{(\tau - \mu)^2}{\tau\mu} n.$$

Proof. The above identity can be easily verified by simplifying the right hand side. ■

The following obvious inequality will be often used in the sequel

$$\omega(z) \geq (\mu(z))^{-1/2}. \quad (4.4)$$

Corollary 4.4 For any $z \in \mathbb{R}_{++}^{2n}$ and any $\tau > 0$ with $\varsigma\kappa(z, \tau) < \sqrt{n}$ there holds

$$\left(1 - \frac{\varsigma\kappa(z, \tau)}{\sqrt{n}}\right)\mu(z) \leq \tau \leq \left(1 + \frac{\varsigma\kappa(z, \tau)}{\sqrt{n}}\right)\mu(z), \quad (4.5)$$

and

$$\delta^2(z, \mu(z)) \leq (\varsigma\kappa(z, \tau))^2 \left[1 + \frac{(\varsigma\kappa(z, \tau))^2}{n - (\varsigma\kappa(z, \tau))^2}\right], \quad (4.6)$$

where ς is given by (4.2).

If HLCP (3.1) is skew-symmetric we also have

$$\delta^2(z, \mu(z)) \leq \kappa^2(z, \mu(z)). \quad (4.7)$$

Proof. By taking $a = \tau e - xs$ in Lemma 4.1 and using Lemma 4.3 together with (4.4) we deduce that

$$\begin{aligned}\zeta^2 \kappa^2(z, \tau) &= \zeta^2 \omega^2(z) \alpha^2(z, \tau) \geq \tau \omega^2(z) \delta^2(z, \tau) \geq \frac{\tau}{\mu} \delta^2(z, \tau) \\ &= \frac{\tau^2}{\mu^2} \delta^2(z, \mu(z)) + \frac{(\tau - \mu)^2}{\mu^2} n.\end{aligned}\quad (4.8)$$

The above inequality implies

$$\zeta^2 \kappa^2(z, \tau) \geq \frac{(\tau - \mu)^2}{\mu^2} n \quad (4.9)$$

and

$$\delta^2(z, \mu(z)) \leq (\zeta \kappa(z, \tau))^2 + \frac{\mu^2 - \tau^2}{\tau^2} (\zeta \kappa(z, \tau))^2 - \frac{(\tau - \mu)^2}{\tau^2} n. \quad (4.10)$$

Inequality (4.9) clearly implies (4.5). By denoting

$$t = (\zeta \kappa(z, \tau))^2, \quad \tau = (1 - \gamma)\mu, \quad \text{where } \gamma < 1$$

we have

$$\frac{\mu^2 - \tau^2}{\tau^2} (\zeta \kappa(z, \tau))^2 - \frac{(\tau - \mu)^2}{\tau^2} n = \frac{(1 - (1 - \gamma)^2)t - \gamma^2 n}{(1 - \gamma)^2} =: f(\gamma).$$

The derivative of the function above is

$$f'(\gamma) = \frac{2(n\gamma - t)}{(\gamma - 1)^3}.$$

We deduce that the maximum of $f(\gamma)$ for $\gamma \in (-\infty, 1)$ is attained at $\gamma = t/n$ and we have $f(t/n) = t^2/(n - t)$. Inequality (4.6) follows from this observation and (4.10). Finally, (4.7) is obtained by setting $\zeta = 1$ and $\tau = \mu$ in (4.6). \blacksquare

In the following lemma we obtain bounds on the elements of xs in terms of the proximity $\delta(z, \mu(z))$. The result follows from Lemma 3.2 in [5], but we give a simple proof for the sake of completeness.

Lemma 4.5 *If*

$$\delta^2(z, \mu(z)) \leq 2\eta$$

then

$$\frac{1}{1 + \eta + \sqrt{2\eta + \eta^2}} \leq \frac{xs}{\mu} \leq 1 + \eta + \sqrt{2\eta + \eta^2}.$$

Proof. We first note that

$$\delta^2(z, \mu(z)) = \sum_{i=1}^n \frac{\mu}{x_i s_i} - n.$$

We denote by ν an arbitrary entry of $\mu/(xs)$. Without loss of generality we assume $\nu = \mu/(x_1 s_1)$. Then by using the inequality between the arithmetic and the harmonic mean we can write

$$\begin{aligned} 2\eta &\geq \nu - n + \sum_{i=2}^n \frac{\mu}{x_i s_i} \geq \nu - n + \frac{(n-1)^2 \mu}{\sum_{i=2}^n x_i s_i} \\ &= \nu - n + \frac{(n-1)^2 \mu}{n\mu - \frac{\mu}{\nu}} = \nu - n + \frac{(n-1)^2 \nu}{n\nu - 1}. \end{aligned}$$

It follows that ν satisfies the following quadratic inequality:

$$n\nu^2 - 2(1 + \eta)n\nu + 2\eta + n \leq 0.$$

Hence ν must lie between the roots of the corresponding quadratic equation, i.e.,

$$1 + \eta - \sqrt{2\eta + \eta^2 - 2\eta/n} \leq \nu \leq 1 + \eta + \sqrt{2\eta + \eta^2 - 2\eta/n}.$$

This implies the following bounds that are independent of n

$$1 + \eta - \sqrt{2\eta + \eta^2} \leq \nu \leq 1 + \eta + \sqrt{2\eta + \eta^2},$$

and the proof is complete by noticing that the lower bound above is equal to the inverse of the upper bound. \blacksquare

As an immediate consequence of Corollary 4.4 and Lemma 4.5 we obtain the following result.

Corollary 4.6 *For any $z = [x, s] \in \mathbb{R}_{++}^{2n}$ and any $\tau > 0$ with $\varsigma\kappa(z, \tau) < \sqrt{n}$ we have*

$$\psi^{-1} \leq \frac{xs}{\mu(z)} \leq \psi,$$

where

$$\psi = \psi(z, \tau) = 1 + \vartheta + \sqrt{2\vartheta + \vartheta^2}, \quad (4.11)$$

$$\vartheta = \vartheta(z, \tau) = \frac{1}{2}(\varsigma\kappa(z, \tau))^2 \left[1 + \frac{(\varsigma\kappa(z, \tau))^2}{n - (\varsigma\kappa(z, \tau))^2} \right], \quad (4.12)$$

and ς is defined by (4.2).

If HLCP (3.1) is skew-symmetric we also have

$$\varrho^{-1} \leq \frac{xs}{\mu(z)} \leq \varrho,$$

where

$$\varrho = \varrho(z) = 1 + .5 \kappa^2(z, \mu(z)) + \sqrt{\kappa^2(z, \mu(z)) + .25 \kappa^4(z, \mu(z))}. \quad (4.13)$$

We end this section by finding an upper bound for $\kappa(z, \tau)$ in terms of the proximity measure

$$\delta_K(z, \tau) := \frac{\|xs - \tau e\|_2}{\tau} \quad (4.14)$$

considered in [6].

Lemma 4.7 For any $(z, \tau) \in \mathbb{R}_{++}^{2n} \times \mathbb{R}_{++}$ with $\delta_K(z, \tau) < 1$ we have

$$\kappa(z, \tau) \leq \frac{\delta_K(z, \tau)}{1 - \delta_K(z, \tau)}.$$

Proof. From the obvious relation

$$xs = \tau - (\tau - xs) \geq \tau - \|\tau - xs\|_2 = \tau(1 - \delta_K(z, \tau))$$

we deduce that

$$\omega(z) \leq \frac{1}{\sqrt{\tau(1 - \delta_K(z, \tau))}}.$$

Also, by using (4.3) and Corollary 4.2 we obtain

$$\alpha(z, \tau) \leq \sqrt{\tau} \delta(z, \tau) \leq \omega(z) \|\tau e - xs\|_2 = \omega(z) \delta_K(z, \tau) \tau.$$

Since $\kappa(z, \tau) = \omega(z) \alpha(z, \tau)$ we obtain the desired bound. ■

5 Polynomial Complexity

In this section we return to Algorithm 1 and prove that it is globally convergent for any choice of parameters $0 < \kappa_1 < \kappa_2 < .5$. More precisely we will show that at each outer iteration the duality gap is reduced by a factor of $1 - \chi/\sqrt{n}$ where χ is a constant depending only on κ_1 and κ_2 . Also we will prove that each inner iteration will terminate in at most m steps, where m depends only on κ_1 and κ_2 . This implies that Algorithm 1 has $O(\sqrt{n})L$ iteration complexity.

First we prove that all the points z produced by Algorithm 1 are strictly feasible .

Proposition 5.1 *All the points z generated by Algorithm 1 belong to \mathcal{F}^0 . Moreover if HLCP (3.1) is monotone then*

$$\left(1 - \frac{\sqrt{2}\kappa_1}{\sqrt{n}}\right)\mu(z^k) \leq \tau_k \leq \left(1 + \frac{\sqrt{2}\kappa_1}{\sqrt{n}}\right)\mu(z^k), \quad (5.1)$$

and if HLCP (3.1) is skew-symmetric then $\tau_k = \mu(z^k)$.

Proof. First we prove that all points generated by Algorithm 1 are strictly positive. Since $z^0 \in \mathcal{F}^0$ we have $\bar{z} > 0$ when $k = 0$. We also have $\kappa(\bar{z}) \leq \kappa_1$. According to the Kantorovich Theorem all the points z generated in the inner iteration satisfy

$$\|z - \bar{z}\|_{\bar{z}} < \frac{1 - \sqrt{1 - 2\kappa_1}}{\omega(\bar{z})}. \quad (5.2)$$

We will prove that

$$\|z - \bar{z}\|_{\bar{z}} < (\omega(\bar{z}))^{-1} \text{ implies } z > 0.$$

If $z = [x, s]$ and $\bar{z} = [\bar{x}, \bar{s}]$ we can write

$$x = \bar{x} - (\bar{x} - x) = \bar{x} \left(e - \bar{x}^{-1}(\bar{x} - x) \right) \geq \left(1 - \|\bar{x}^{-1}(\bar{x} - x)\|_{\infty} \right) \bar{x}$$

Since

$$\begin{aligned} \|\bar{x}^{-1}(\bar{x} - x)\|_{\infty} &= \|(\bar{x}\bar{s})^{-1/2}\bar{D}(\bar{x} - x)\|_{\infty} \\ &\leq \omega(\bar{z}) \|\bar{D}(\bar{x} - x)\|_{\infty} \leq \omega(\bar{z}) \|z - \bar{z}\|_{\bar{z}} < 1 \end{aligned}$$

we deduce that $x > 1$. The positivity of s is proved similarly.

We will prove now that any $z = [x, s]$ produced during the inner iteration satisfies

$$Qx + Rs = b.$$

Since $\bar{z} \in \mathcal{F}$ the above equality is certainly true for $z = \bar{z}$. This property is preserved by the Newton or the simplified Newton steps. For example a Newton step $z^+ = z - F'(z)^{-1}F_{\tau}(z)$ can be written as

$$\begin{aligned} su + xv &= \tau e - xs \\ Qu + Rv &= 0 \end{aligned} \quad z^+ = z + [u, v], \text{ where } z = [x, s], \quad z^+ = [x^+, s^+] \quad (5.3)$$

and we have

$$Qx^+ + Rs^+ = Qx + Rs + Qu + Rv = Qx + Rs = b.$$

A similar argument applies if z^+ is produced by a simplified Newton step $z^+ = z - F'(\bar{z})^{-1}F_\tau(z)$:

$$\begin{aligned}\bar{s}u + \bar{x}v &= \tau e - xs \\ Qu + Rv &= 0\end{aligned}$$

$$z^+ = z + [u, v], \text{ where } z = [x, s], \bar{z} = [\bar{x}, \bar{s}], z^+ = [x^+, s^+]. \quad (5.4)$$

An induction argument completes our claim that all points generated by Algorithm 1 are strictly feasible.

We prove now the second part of our proposition. Since (5.1) is an immediate consequence of Corollary 4.5 we only have to prove that if HLCP is skew-symmetric, then

$$\tau_k = \mu(z^k), \quad k = 0, 1, \dots \quad (5.5)$$

For $k = 0$ this is true because Algorithm 1 starts with $\tau_0 = \mu(z^0)$. Since τ is updated in the outer iteration, at the beginning of the inner iteration we no longer have $\tau = \mu(z)$. However we will prove that this property is restored after the first Newton step and then it is preserved throughout the inner iteration by subsequent Newton or simplified Newton steps. After a Newton step of the form (5.3) we have

$$x^+s^+ = xs + (su + xv) + uv = \tau e + uv.$$

Since $u^Tv = 0$ it follows that $\mu(z^+) = \tau$.

After a simplified Newton step of the form (5.3) there holds

$$x^+s^+ = xs + (su + xv) + uv = \tau e + (s - \bar{s})u + (x - \bar{x}). \quad (5.6)$$

Let us denote $\bar{u} = x - \bar{x}$ and $\bar{v} = s - \bar{s}$. Since $z, \bar{z} \in \mathcal{F}$ we have $Q\bar{u} + R\bar{v} = 0$. Therefore $\bar{u}^T\bar{v} = 0$. Also $Q(u - \bar{u}) + R(v - \bar{v}) = 0$ and we can write

$$0 = (u - \bar{u})^T(v - \bar{v}) = u^Tv - (u^T\bar{v} + v^T\bar{u}) + \bar{u}^T\bar{v} = -(u^T\bar{v} + v^T\bar{u}).$$

The above inequality and (5.6) imply $\mu(z^+) = \tau$. ■

At each step of the outer iteration we are given $(z, \tau) \in \mathbb{R}_{++}^{2n} \times \mathbb{R}_{++}$ with $\kappa(z, \tau) \leq \kappa_1$ and we want to determine the largest $\theta \in (0, 1)$ such that $\kappa(z, (1 - \theta)\tau) \leq \kappa_2$. The following Lemma will enable us to find a lower bound for this θ .

Lemma 5.2 *For any $(z, \tau, \theta) \in \mathbb{R}_{++}^{2n} \times \mathbb{R}_{++} \times (0, 1)$ with $\kappa(z, \tau) < \sqrt{n}$ we have*

$$\kappa(z, (1 - \theta)\tau) \leq \kappa(z, \tau) + \theta\sqrt{\psi(z, \tau)}\sqrt{(\zeta\kappa(z, \tau))^2 + n},$$

where ς and ψ are defined by (4.2) and (4.11) respectively.

If HLCP (3.1) is skew-symmetric we also have

$$\kappa(z, (1 - \theta)\tau) \leq \kappa(z, \tau) + \theta \sqrt{\varrho(z)} \sqrt{(\varsigma \kappa(z, \tau))^2 + n},$$

where ϱ is defined by (4.13)

Proof. For any $\theta \in (0, 1)$ we have

$$\alpha(z, (1 - \theta)\tau) = \| [u(\theta), v(\theta)] \|_z,$$

where $[u(\theta), v(\theta)]$ is the solution of the linear system

$$\begin{aligned} su(\theta) + xv(\theta) &= (1 - \theta)\tau - xs \\ Qu + Rv &= 0. \end{aligned}$$

It follows that

$$[u(\theta), v(\theta)] = [u, v] + \theta[\tilde{u}, \tilde{v}],$$

where

$$\begin{aligned} su + xv &= \tau - xs \\ Qu + Rv &= 0 \end{aligned}$$

and

$$\begin{aligned} s\tilde{u} + x\tilde{v} &= -\tau \\ Q\tilde{u} + R\tilde{v} &= 0. \end{aligned}$$

Using (4.8) we obtain

$$\begin{aligned} \| [\tilde{u}, \tilde{v}] \|_z^2 &\leq \sum_{i=1}^n \frac{\tau^2}{x_i s_i} = \frac{\tau^2}{\mu} \sum_{i=1}^n \frac{\mu}{x_i s_i} = \frac{\tau^2}{\mu} [\delta^2(z, \mu) + n] \\ &= \mu \left[\frac{\tau^2}{\mu^2} \delta^2(z, \mu) + \frac{\tau^2}{\mu^2} n \right] = \mu \left[(\varsigma \kappa(z, \tau))^2 + \frac{\tau^2 - (\tau - \mu)^2}{\mu^2} \right] \\ &= \mu \left[(\varsigma \kappa(z, \tau))^2 + \left(1 - (1 - \tau/\mu)^2\right) n \right] \leq \mu \left[(\varsigma \kappa(z, \tau))^2 + n \right]. \end{aligned}$$

Since $\| [u, v] \|_z = \alpha(z, \tau)$, we can write

$$\alpha(z, (1 - \theta)\tau) \leq \| [u, v] \|_z + \theta \| [\tilde{u}, \tilde{v}] \|_z \leq \alpha(z, \tau) + \sqrt{\mu} \theta \sqrt{(\varsigma \kappa(z, \tau))^2 + n}.$$

Our claim is obtained by multiplying with $\omega(z)$ and using Corollary 4.6. ■

Now we are in the position to find an explicit lower bound for θ in terms of κ_1 , κ_2 , and \sqrt{n} . In order to treat the monotone case and the skew-symmetric case together it is convenient to introduce the following notation for $i = 1, 2$

$$\psi_i := \begin{cases} 1 + \vartheta_i + \sqrt{2\vartheta_i + \vartheta_i^2}, & \text{if HLCP is monotone} \\ 1 + \eta_i + \sqrt{2\eta_i + \eta_i^2}, & \text{if HLCP is skew-symmetric} \end{cases} \quad (5.7)$$

where,

$$\vartheta_i = \kappa_i^2 \left(1 + \frac{\kappa_i^2}{1 - \kappa_i^2} \right), \quad \eta_i = \frac{\kappa_i^2}{2}, \quad i = 1, 2. \quad (5.8)$$

Corollary 5.3 *The parameter θ determined at each outer iteration of Algorithm 1 satisfies the inequality*

$$\theta \geq \frac{\kappa_2 - \kappa_1}{\sqrt{\psi_1} \sqrt{\zeta^2 \kappa_1^2 + n}} \geq \frac{\chi}{\sqrt{n}},$$

where

$$\chi = \chi(\kappa_1, \kappa_2) = \frac{\sqrt{2} (\kappa_2 - \kappa_1)}{\sqrt{2 + \zeta^2 \kappa_1^2} \sqrt{\psi_1}}, \quad (5.9)$$

with ζ and ψ_1 defined by (4.2) and (5.7) respectively.

Proof. The corollary is an immediate consequence of Lemma 5.2. ■

In next theorem we will obtain a bound on the number of steps of the inner iteration that depends only on κ_1 and κ_2 .

Theorem 5.4 *If Newton's method is used in Algorithm 1 then each inner iteration terminates in at most $\mathcal{N}(\kappa_1, \kappa_2)$ steps, where*

$$\mathcal{N}(\kappa_1, \kappa_2) = \lceil \log_2 \left(\frac{\log_2(\zeta_{\mathcal{N}})}{\log_2 \left((1 - \sqrt{1 - 2\kappa_2} - \kappa_2) / \kappa_2 \right)} \right) \rceil, \quad (5.10)$$

and

$$\zeta_{\mathcal{N}} = \frac{(1 - \zeta \kappa_2 / \sqrt{2})(\kappa_2^2 - (1 - \sqrt{1 - 2\kappa_2} - \kappa_2)^2) \kappa_1}{2\sqrt{2} \kappa_2^2 \sqrt{1 - 2\kappa_2} (\sqrt{\psi_2} + 1 - \sqrt{1 - 2\kappa_2}) (1 + \kappa_1)}.$$

If the simplified Newton method is used in Algorithm 1 then each inner iteration terminates in at most $\mathcal{S}(\kappa_1, \kappa_2)$ steps

$$\mathcal{S}(\kappa_1, \kappa_2) = \lceil \frac{\log_2(\zeta_{\mathcal{S}})}{\log_2(1 - \sqrt{1 - 2\kappa_2})} + 1 \rceil, \quad (5.11)$$

and

$$\zeta_S = \frac{(1 - \varsigma \kappa_2 / \sqrt{2})(\kappa_2^2 - (1 - \sqrt{1 - 2\kappa_2} - \kappa_2)^2) \kappa_1}{2\sqrt{2} \sqrt{1 - 2\kappa_2} (1 - \sqrt{1 - 2\kappa_2} - \kappa_2)^2 (\sqrt{\psi_2} + 1 - \sqrt{1 - 2\kappa_2}) (1 + \kappa_1)}.$$

Proof. The inner iteration starts with a vector $\bar{z} \in \mathcal{F}^0$ and a parameter $\tau \in \mathbb{R}_{++}$ such that

$$\kappa(\bar{z}, \tau) \leq \kappa_2 \quad (5.12)$$

and it terminates when a point z is obtained such that

$$\kappa(z, \tau) \leq \kappa_1. \quad (5.13)$$

Note that \bar{z} and τ are constant during the inner iteration. Since we suppose that $\kappa_2 < .5$, it follows that the sequences produced by either Newton's method or the simplified Newton method with starting point \bar{z} converge to the solution $z(\tau)$ of the equation $F_\tau(z) = 0$. In what follows we will prove that (5.13) is satisfied provided $\|z - z(\tau)\|_{\bar{z}}$ is small enough. From Lemma 4.7 it follows that (5.13) is satisfied if

$$\delta_K(z, \tau) \leq \frac{\kappa_1}{1 + \kappa_1}. \quad (5.14)$$

With

$$\bar{z} = [\bar{x}, \bar{s}], \quad z = [x, s], \quad \bar{D} = \bar{X}^{-1/2} \bar{S}^{1/2}$$

we can write

$$\begin{aligned} \|xs - \tau e\|_2 &= \|xs - x(\tau)s(\tau)\|_2 = \|xs - xs(\tau) + xs(\tau) - x(\tau)s(\tau)\|_2 \\ &= \left\| \bar{D}x \bar{D}^{-1}s + \bar{D}x \bar{D}^{-1}s(\tau) - \bar{D}x(\tau) \bar{D}^{-1}s(\tau) \right\|_2 \\ &= \left\| \bar{D}x \bar{D}^{-1}(s - s(\tau)) + \bar{D}^{-1}s(\tau) \bar{D}(x - x(\tau)) \right\|_2 \\ &\leq \left\| \bar{D}x \right\|_\infty \left\| \bar{D}^{-1}(s - s(\tau)) \right\|_2 + \left\| \bar{D}^{-1}s(\tau) \right\|_\infty \left\| \bar{D}(x - x(\tau)) \right\|_2 \\ &\leq \sqrt{\left\| \bar{D}x \right\|_\infty^2 + \left\| \bar{D}^{-1}s(\tau) \right\|_\infty^2} \|z - z(\tau)\|_{\bar{z}}. \end{aligned}$$

In order to bound $\left\| \bar{D}x \right\|_\infty$ we use Corollary 4.6 and the identity $\bar{D}\bar{x} = (\bar{x}\bar{s})^{1/2}$ to obtain

$$\left\| \bar{D}x \right\|_\infty \leq \left\| \bar{D}\bar{x} \right\|_\infty + \left\| \bar{D}(x - \bar{x}) \right\|_\infty \leq \sqrt{\psi_2 \mu(\bar{z})} + \|z - \bar{z}\|_{\bar{z}}.$$

Similarly we get

$$\left\| \bar{D}^{-1}s(\tau) \right\|_\infty \leq \sqrt{\psi_2 \mu(\bar{z})} + \|z(\tau) - \bar{z}\|_{\bar{z}}.$$

Because $z(\tau)$ is the solution of the equation $F_\tau(z) = 0$ and z is produced by Newton's method (or the simplified Newton method) with starting point \bar{z} the Kantorovich Theorem ensures that

$$\|z(\tau) - \bar{z}\|_{\bar{z}}, \|z - \bar{z}\|_{\bar{z}} \leq \frac{1 - \sqrt{1 - 2\kappa_2}}{\omega(\bar{z})}.$$

Taking into account (4.4) we deduce that

$$\sqrt{\|\bar{D}x\|_\infty^2 + \|\bar{D}^{-1}s(\tau)\|_\infty^2} \leq \sqrt{2\mu(\bar{z})} (\sqrt{\psi_2} + 1 - \sqrt{1 - 2\kappa_2}).$$

Therefore,

$$\delta_K(z, \tau) \leq \frac{\sqrt{\mu}}{\tau} \sqrt{2} (\sqrt{\psi_2} + 1 - \sqrt{1 - 2\kappa_2}) \|z - z(\tau)\|_{\bar{z}}.$$

It follows that (5.14) is satisfied if

$$\|z - z(\tau)\|_{\bar{z}} \leq \frac{\tau}{\sqrt{\mu}} v, \quad (5.15)$$

where,

$$v = \frac{\kappa_1}{\sqrt{2}(1 + \kappa_1)(\sqrt{\psi_2} + 1 - \sqrt{1 - 2\kappa_2})}.$$

If Newton's method is used in the inner iteration then from the Kantorovich Theorem it follows that (5.15) is satisfied after m Newton steps provided that

$$\frac{2\beta_2 \lambda_2^{2m}}{1 - \lambda_2^{2m}} \leq \frac{\tau}{\sqrt{\mu}} v, \quad \text{where } \beta_2 = \frac{\sqrt{1 - 2\kappa_2}}{\omega(\bar{z})}, \quad \lambda_2 = \frac{1 - \kappa_2 - \sqrt{1 - 2\kappa_2}}{\kappa_2}.$$

From Corollary 4.4 we obtain the following bound

$$\frac{\tau}{\sqrt{\mu}} = \frac{\tau\sqrt{\mu}}{\mu} \geq (1 - \varsigma\kappa_2/\sqrt{2}) \sqrt{\mu}. \quad (5.16)$$

Using (5.16) together with the fact that

$$1 - \lambda_2^{2m} \geq 1 - \lambda_2^2, \quad \text{for } m \geq 1$$

and (4.4) we deduce that (5.15) is satisfied if

$$\lambda_2^{2m} \leq \frac{(1 - \varsigma\kappa_2/\sqrt{2})(1 - \lambda_2^2)v}{2\sqrt{1 - 2\kappa_2}}.$$

It is easily seen that the right hand side coincides with $\zeta_{\mathcal{N}}$. By taking twice the logarithm of both sides we obtain the claim of our Theorem in case the inner iteration uses Newton steps.

Let us analyze now the case where the inner iteration uses simplified Newton steps. Using (2.12) we deduce in a similar fashion that the Kantorovich theorem it follows that (5.15) is satisfied after m steps if

$$\xi_2^{m-1} \leq \frac{\zeta_{\mathcal{N}}}{\lambda_2^2} = \zeta_{\mathcal{S}}.$$

By taking the logarithm of both sides we obtain the claim of our theorem for the simplified Newton method. ■

Corollary 5.5 *Algorithm 1 terminates using at most $O(\sqrt{n} \log(\frac{x^0 \tau_{s^0}}{\epsilon}))$ Newton and simplified Newton steps.*

Proof. The argument is standard. From Corollary 5.3 it follows that

$$\tau_k \leq \left(1 - \frac{\chi}{\sqrt{n}}\right)^k \tau_0,$$

and the assertion follows by using the fact that $\tau_0 = \mu(z^0)$ together with Proposition 5.1 and Theorem 5.4. ■

It is interesting to see how the complexity depends on κ_1 and κ_2 . Roughly speaking the upper bounds on the number of steps of the inner iteration \mathcal{N} and \mathcal{S} are increasing in κ_2 and decreasing in $\kappa_2 - \kappa_1$, while the coefficient χ that characterizes the guaranteed decrease of the primal dual gap at each outer iteration is increasing in $\kappa_2 - \kappa_1$ and decreasing in κ_1 . One could obtain different scenarios for different ratios between κ_1 and κ_2 . In the following two corollaries we give some numerical values in case $\kappa_2 = 2 \kappa_1$. We note that the numerical values given in those corollaries are only upper bounds. They are obtained by plugging numerical values in the formulae (5.9), (5.10) and (5.11). We also note that our estimates are true for all values of $n \geq 2$. For larger values of n the estimates can be slightly improved (see e.g. (4.5)). Also when comparing \mathcal{N} and \mathcal{S} one has to consider that \mathcal{N} Newton steps require \mathcal{N} matrix factorizations and \mathcal{N} backsolves, while \mathcal{S} simplified Newton steps require one matrix factorization and \mathcal{S} backsolves. In case the matrices Q and R are dense one matrix factorization requires $O(n^3)$ arithmetic operations, while one backsolve requires only $O(n^2)$ arithmetic operations.

Corollary 5.6 *If the HLCP is monotone then:*

1. $\mathcal{N}(.12, .24) = 1, \quad \chi(.12, .24) > .1 ;$

2. $\mathcal{N}(.21, .42) = 2$, $\mathcal{S}(.21, .42) = 5$, $\chi(.21, .42) > .17$;
3. $\mathcal{N}(.24, .48) = 3$, $\mathcal{S}(.24, .48) = 12$, $\chi(.24, .48) > .196$.
4. $\mathcal{N}(.245, .49) = 4$, $\mathcal{S}(.245, .49) = 18$, $\chi(.245, .49) > .199$.

Corollary 5.7 *If the HLCP is skew-symmetric then:*

1. $\mathcal{N}(.125, .25) = 1$, $\chi(.12, .24) > .11$;
2. $\mathcal{N}(.215, .43) = 2$, $\mathcal{S}(.215, .43) = 5$, $\chi(.215, .43) > .19$;
3. $\mathcal{N}(.24, .48) = 3$, $\mathcal{S}(.24, .48) = 10$, $\chi(.24, .48) > .2$.
4. $\mathcal{N}(.245, .49) = 4$, $\mathcal{S}(.245, .49) = 16$, $\chi(.245, .49) > .21$.

Acknowledgments

This paper was written when the author was visiting the Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB). The author would like to thank Peter Deuffhard, President of ZIB, for the excellent working conditions at ZIB, and for very interesting discussions on various mathematical problems related to the subject of the present paper.

References

- [1] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Academic Press, Boston, MA, 1992.
- [2] P. Deuffhard and G. Heindl. Affine invariant convergence theorems for Newton's method and extensions to related methods. *SIAM J. Numer. Anal.*, 16:1–10, 1979.
- [3] W. B. Gragg and R. A. Tapia. Optimal error bounds for the Newton-Kantorovich theorem. *SIAM J. Numer. Anal.*, 11:10–13, 1974.
- [4] B. Jansen, C. Roos, T. Terlaky, and J.-P. Vial. Primal-dual algorithms for linear programming based on the logarithmic barrier method. *J. Optim. Theory Appl.*, 83(1):1–26, 1994.
- [5] B. Jansen, C. Roos, T. Terlaky, and J.-Ph. Vial. Primal-dual target-following algorithms for linear programming. *Ann. Oper. Res.*, 62:197–231, 1996. Interior point methods in mathematical programming.

- [6] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. A unified approach to interior point algorithms for linear complementarity problems. *Lecture Notes in Comput. Sci.*, 538, 1991.
- [7] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York, 1970.
- [8] A. M. Ostrowski. *Solution of Equations in Euclidean and Banach Spaces*. Academic Press, New York, 1973.
- [9] F. A. Potra. On the a posteriori error estimates for Newton's method. *Beitraege Numer. Math.*, 12:125–138, 1984.
- [10] F. A. Potra and S. J. Wright. Interior-point methods. Technical report, Department of Mathematics and Statistics, UMBC, Baltimore, MD 21250, USA, December 1999.
- [11] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. SIAM Publications, to appear.
- [12] James Renegar. A polynomial-time algorithm, based on Newton's method, for linear programming. *Math. Programming*, 40(1 (Ser. A)):59–93, 1988.
- [13] James Renegar and Michael Shub. Unified complexity analysis for Newton LP methods. *Math. Programming*, 53(1, Ser. A):1–16, 1992.
- [14] C. Roos, J.-Ph. Vial, and T. Terlaky. *Theory and Algorithms for Linear Optimization : An Interior Point Approach*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.
- [15] Steve Smale. Newton's method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.
- [16] S. J. Wright. *Primal–Dual Interior–Point Methods*. SIAM Publications, Philadelphia, 1997.
- [17] Y. Ye. *Interior Point Algorithms : Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.