Math. Program., Ser. A 102: 47–70 (2005)

Digital Object Identifier (DOI) 10.1007/s10107-003-0501-8

Florian A. Potra

The Kantorovich Theorem and interior point methods*

Received: February 11, 2000 / Accepted: December 15, 2003
Published online: February 6, 2004 – © Springer-Verlag 2004

Abstract. The Kantorovich Theorem is a fundamental tool in nonlinear analysis which has been extensively used in classical numerical analysis. In this paper we show that it can also be used in analyzing interior point methods. We obtain optimal bounds for Newton's method when relied upon in a path following algorithm for linear complementarity problems.

Given a point z that approximates a point $z(\tau)$ on the central path with complementarity gap τ, a parameter $\theta \in (0, 1)$ is determined for which the point z satisfies the hypothesis of the affine invariant form of the Kantorovich Theorem, with regards to the equation defining $z((1-\theta)\tau)$. The resulting iterative algorithm produces a point with complementarity gap less than ϵ in at most $O(\sqrt{n}\log(\epsilon_0/\epsilon))$ Newton steps, or simplified Newton steps, where ϵ_0 is the complementarity gap of the starting point and n is the dimension of the problem. Thus we recover the best complexity results known to date and, in addition, we obtain the best bounds for Newton's method in this context.

Key words. Linear complementarity problem – Interior point algorithm – Path-following – Kantorovich Theorem

1. Introduction

The Kantorovich Theorem [16] is one of the most important tools in nonlinear analysis and in classical numerical analysis. However, with a few notable exceptions, it has not been utilized in the mathematical programming community. The purpose of this paper is to provide a bridge between the two research communities by showing that the Kantorovich Theorem can be used in analyzing interior point methods, and can be used for obtaining *optimal* bounds along the way.

The Kantorovich Theorem has been successfully used by numerical analysts for obtaining optimal bounds for many iterative procedures. The original paper of Kantorovich [16] contains optimal a-priori bounds for Newton's method, albeit not in explicit form. Explicit forms of those a-priori bounds were found independently by Ostrowski [34] and Gragg & Tapia [12] (see [40] for the equivalence of the bounds). The paper of Gragg and Tapia [12] also contains sharp a-posteriori bounds for Newton's method. By using different techniques and/or different a-posteriori information, the a-posteriori error bounds for Newton's method were refined in a series of papers including [22–24, 36, 39, 53–55]. Various extensions of the Kantorovich Theorem have been used to obtain error bounds for Newton-like methods [6, 7, 24, 25, 29, 37, 51], inexact Newton methods

F.A. Potra: University of Maryland Baltimore County, e-mail: potra@math.umbc.edu

Mathematics Subject Classification (2000): 49M15, 65H10, 65K05, 90C33

* Work supported by the National Science Foundation under Grants No. 9996154, 0139701

[26], the secant method [35], Halley's method [9], etc. A survey of the results obtained before 2000 is contained in [56].

The Kantorovich Theorem has also been used for proving existence and uniqueness of solutions of nonlinear equations arising in various fields. In order to give the reader an idea of the spectrum of applications of the Kantorovich Theorem, we mention applications to such diverse problems as the chaotic behavior of maps [48], optimal shape design [19], nonlinear Riemann-Hilbert problems [3], optimal control [13], nonlinear finite element [50], integral equations [2, 5, 10], flow problems [20], nonlinear elliptic problems [28], semidefinite programming [30], and multiparameter singularly perturbed systems [27] - to cite only a few papers published in the past five years. We also mention that the Kantorovich Theorem has recently been extended for analyzing Newton's method on Riemannian manifolds [11]. The relation between the Kantorovich Theorem and other existence theorems for solutions of nonlinear equations is investigated in [1, 21].

In the present paper we show that the Kantorovich Theorem can be very useful in constructing and analyzing an optimal-complexity path following algorithm for linear complementarity problems. We have chosen to apply the Kantorovich Theorem to linear complementarity problems because such problems provide a convenient framework for analyzing primal-dual interior point methods. 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 monograph of Steve Wright [52] and the recent survey paper [41]). 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 [46, 52, 57]).

We will show that the affine invariant Kantorovich Theorem [8] can be used for constructing a class of path following algorithms 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 proven that θ is bounded below by a multiple of $n^{-1/2}$. Since the hypothesis of the Kantorovich Theorem is satisfied, the sequence generated by Newton's method, or by 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 (or simplified Newton) steps, where ϵ_0 is the complementarity gap of the starting point. For linear programming problems with rational input data of bit length 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 [52]).

The use of Newton method techniques in the study of interior point methods goes back to the celebrated paper of Renegar [43]. The affine invariant Kantorovich Theorem was used by Renegar and Shub [45] 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 [47], 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. A beautiful analysis of the damped Newton method for minimizing self-concordant functions, and for solving variational inequalities with strongly self-concordant monotone operators, can be found in the monograph of Nesterov and Nemirovskii [31]. Their analysis is also affine invariant and provides the basis for constructing interior point methods for a much larger class of problems than that considered in the present paper.

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.

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 $v_i \geq 3,\ i = 1, \ldots, n$. Traditionally, this is written as $v \geq 3e$. If $\| . \|$ is a vector norm on $\mathbb{R}^n$ and A is a matrix then the operator norm induced by $\| . \|$ 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 $\| U \|_2 = \| u \|_\infty$.

2. The Kantorovich Theorem

The celebrated Kantorovich Theorem gives sufficient conditions for the existence and uniqueness of solutions of nonlinear operator equations in Banach spaces, and shows that under those conditions the sequences generated by Newton's method and the simplified Newton method converge to the solution. In what follows we 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 Deuflhard and Heindl [8]. As will become apparent in Section 3, the "traditional form" of the Kantorovich Theorem as presented, for example, in the classical monograph of Ortega and Rheinboldt [33, Th. 12.6.2] cannot be directly applied to our problem. We mention that the original paper of Kantorovich [16], as well as the book of Kantorovich and Akilov [17], contain an affine invariant form of the theorem (see also [6]). However, it is the "traditional form" that appears in most books and papers published in the western literature, including [32–34, 42, 49].

Let $F : D \subset \mathbb{R}^p \to \mathbb{R}^p$ 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 $\| \, . \, \|$ 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,

$$\overline{B}(z, \rho) := \{y \in \mathbb{R}^p \, : \, \|y - z\| \leq \rho\}, \tag{1}$$

is included in D, i.e.,

$$\overline{B}(z, \rho) \subset D \, . \tag{2}$$

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

$$\|F'(z)^{-1}(F'(y) - F'(\overline{y}))\| \leq \omega \|y - \overline{y}\|, \; \forall y, \overline{y} \in \overline{B}(z, \rho). \tag{3}$$

The Kantorovich Theorem essentially states that if the quantity

$$\alpha := \|F'(z)^{-1}F(z)\| \tag{4}$$

is small enough, in the sense that

$$\kappa := \alpha\omega \leq \frac{1}{2} \tag{5}$$

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

$$z_{\mathcal{N}}^0 = z, \, z_{\mathcal{N}}^{k+1} = z_{\mathcal{N}}^k - F'(z_{\mathcal{N}}^k)^{-1}F(z_{\mathcal{N}}^k), \; k = 0, 1, \ldots \tag{6}$$

and by the simplified Newton method

$$z_{\mathcal{S}}^0 = z, \, z_{\mathcal{S}}^{k+1} = z_{\mathcal{S}}^k - F'(z)^{-1}F(z_{\mathcal{S}}^k), \; k = 0, 1, \ldots \tag{7}$$

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

Theorem 1. *(The Kantorovich Theorem) Let $F : D \subset \mathbb{R}^p \to \mathbb{R}^p$ be a differentiable mapping and let $z \in D$ be such that $F'(z)$ is nonsingular. Assume that conditions (2) – (5) are satisfied and that the radius ρ is large enough in the sense that*

$$\hat{\rho} := \frac{1-\sqrt{1-2\kappa}}{\omega} \leq \rho. \tag{8}$$

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*

$$\check{\rho} := \frac{1+\sqrt{1-2\kappa}}{\omega} \tag{9}$$

   does not contain any zero of F different from z^*;
3. *The iterative procedures (6) and (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*

$$\left\| z_{\mathcal{N}}^k - z^* \right\| \leq \frac{2\beta\,\lambda^{2^k}}{1-\lambda^{2^k}}, \tag{10}$$

   with

$$\beta = \frac{\sqrt{1-2\kappa}}{\omega}, \quad \lambda = \frac{1-\sqrt{1-2\kappa}\; - \kappa}{\kappa}, \;\; k = 1, 2, \ldots \quad , \tag{11}$$

   and for the simplified Newton methods we have

$$\left\| z_{\mathcal{S}}^k - z^* \right\| \leq \frac{2\,\beta\,\lambda^2}{1-\lambda^2}\,\xi^{k-1}, \;\; k = 1, 2, \ldots \quad . \tag{12}$$

   where

$$\xi = 1 - \sqrt{1-2\kappa}\,. \tag{13}$$

Proof. The above theorem is essentially the affine invariant version of the Kantorovich Theorem presented in [8] with the error bounds (10) from [40]. The only thing that remains to be proved is (12). In order to do that we first write

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

Using (3) and the fact that the function difference can be expressed as the integral of its Jacobian we deduce successively that

$$\begin{aligned}
\left\| z_{\mathcal{S}}^{k+1} - z^* \right\| &= \left\| \int_0^1 F'(z)^{-1}\left[F'\left(z^* + t(z_{\mathcal{S}}^k - z^*)\right) - F'(z)\right](z_{\mathcal{S}}^k - z^*)dt \right\| \\
&\leq \left\| z_{\mathcal{S}}^k - z^* \right\| \int_0^1 \left\| F'(z)^{-1}\left[F'\left(z^* + t(z_{\mathcal{S}}^k - z^*)\right) - F'(z)\right] \right\| dt \\
&\leq \omega \left\| z_{\mathcal{S}}^k - z^* \right\| \int_0^1 \left(t \left\| z_{\mathcal{S}}^k - z \right\| + (1-t)\left\| z^* - z \right\| \right) dt \\
&= \frac{1}{2}\omega\left(\left\| z_{\mathcal{S}}^k - z \right\| + \left\| z^* - z \right\| \right) \left\| z_{\mathcal{S}}^k - z^* \right\| .
\end{aligned}$$

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

$$\left\| z_{\mathcal{S}}^{k+1} - z^* \right\| \le \omega \hat{\rho} \left\| z_{\mathcal{S}}^k - z^* \right\| = \xi \left\| z_{\mathcal{S}}^k - z^* \right\|, \ k = 1, 2, \ldots \quad .$$

Using (10) for the first Newton step we obtain

$$\left\| z_{\mathcal{S}}^k - z^* \right\| \le \left\| z_{\mathcal{S}}^1 - z^* \right\| \xi^{k-1} \le \frac{2\beta\lambda^2}{1-\lambda^2}\, \xi^{k-1}, \ k = 1, 2, \ldots \quad .$$

□

It is known (see [36]) that the a-priori estimates (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 Theorem, sharp a-priori bounds were obtained in [34] and [12]. The equivalence between the bounds of [34] and [12] is proved in [36].

The a-priori estimates (12) are not sharp in the sense described above, but they are sufficiently good for our purpose. The sharp a-priori estimates for the simplified Newton method given in [36] are not in explicit form and therefore they are more difficult to use in complexity analysis.

In the "traditional form" of the Kantorovich Theorem, the Lipschitz condition (3) is replaced by the following two conditions:

$$\|F'(z)^{-1}\| \le \zeta,$$

$$\|(F'(y) - F'(w))\| \le \eta \|y - w\|, \ \forall y, w \in \overline{B}(z, \rho).$$

If these conditions hold then (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. On the other hand, it is interesting to note that the affine-invariant version of the Kantorovich Theorem can be obtained by applying the traditional form to the nonlinear operator $F'(z)^{-1}F$ instead of the operator F.

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

$$\begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\ge 0. \end{aligned} \tag{1}$$

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 HLCPs. Therefore HLCP provides a convenient and general framework for studying interior point methods.

Throughout this paper we assume that the HLCP (1) is monotone in the sense that:

$$Qu + Rv = 0 \text{ implies } u^T v \geq 0, \text{ for any } u, v \in \mathbb{R}^n . \tag{2}$$

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 [4]). If the HLCP is a reformulation of a LP then the following stronger condition holds:

$$Qu + Rv = 0 \text{ implies } u^T v = 0, \text{ for any } u, v \in \mathbb{R}^n . \tag{3}$$

In this case we say that the HLCP is skew-symmetric. Since in the skew-symmetric case we can often obtain sharper estimates, we analyze this case separately when it is beneficial to do so. The HLCP is uninteresting for $n = 1$, so we assume throughout this paper that $n \geq 2$.

It is known (see [18]) 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:

$$\begin{aligned} xs &= \tau e \\ Qx + Rs &= b \\ x, s &\geq 0 \end{aligned} \tag{4}$$

has a unique positive solution $z(\tau) = \left[\, x(\tau)^T,\ s(\tau)^T \,\right]^T$. 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 $\tau \to 0$. Therefore one can numerically solve the HLCP by following the central path as $\tau \to 0$. 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 $\tau > 0$. The algorithm determines a smaller value of the central path parameter

$$\tau_+ = (1 - \theta)\tau, \tag{5}$$

where the value $\theta \in (0, 1)$ is computed by some procedure. Then it computes an approximation (x^+, s^+) of $(x(\tau_+), s(\tau_+))$. The procedure is repeated with (x^+, s^+, τ_+) in place of (x, s, τ).

In what follows, we show how the basic step of a path following algorithm can be analyzed using the Kantorovich theorem. It is convenient to introduce the notations

$$z = \begin{bmatrix} x \\ s \end{bmatrix},\ z(\tau) = \begin{bmatrix} x(\tau) \\ s(\tau) \end{bmatrix},\ z^+ = \begin{bmatrix} x^+ \\ s^+ \end{bmatrix},\ z(\tau_+) = \begin{bmatrix} x(\tau_+) \\ s(\tau_+) \end{bmatrix},\ etc. \tag{6}$$

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

$$F_\sigma(z) = \begin{bmatrix} xs - \sigma e \\ Qx + Rs - b \end{bmatrix}. \tag{7}$$

The equation (4) defining $z(\tau)$ becomes

$$F_\tau(z) = 0, \tag{8}$$

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

$$F_{(1-\theta)\tau}(z) = 0\,. \tag{9}$$

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

$$\mathcal{F}^0 := \left\{ z = \left[x^T,\ s^T \right]^T \in \mathbb{R}^{2n}_{++}\ ;\ Qx + Rs = b \right\}. \tag{10}$$

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 (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 (9). In order to do so, we have to estimate the quantities α and ω that appear in Theorem 1.
Since the Jacobian of F_σ does not depend on σ, we denote

$$F'(z) := F'_\sigma(z) = \begin{bmatrix} S & X \\ Q & R \end{bmatrix}. \tag{11}$$

Using the monotonicity of the HLCP, it can be shown that $F'(z)$ is invertible for any $z \in \mathbb{R}^{2n}_{++}$. This fact is well known and can be obtained, for example, as an immediate consequence of Lemma 2.

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

$$L := F'(z)^{-1}(F'(y) - F'(\overline{y})) = F'(z)^{-1} \begin{bmatrix} V & U \\ 0 & 0 \end{bmatrix}, \tag{12}$$

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

$$h = \left[u^T,\ v^T \right]^T := y - \overline{y}.$$

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

$$t = L\overline{z}, \quad t = \left[t_x{}^T,\ t_s{}^T \right]^T, \quad \overline{z} = \left[\overline{z}_x{}^T,\ \overline{z}_s{}^T \right]^T. \tag{13}$$

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

$$\begin{aligned} st_x + xt_s &= u\overline{z}_s + v\overline{z}_x \\ Qt_x + Rt_v &= 0. \end{aligned} \tag{14}$$

The Lipschitz constant ω will be determined as an upper bound of the ratio $\|t\|/\|h\|$ for all $\overline{z}$ in $\mathbb{R}^{2n}$ and all y in some neighborhood of $\overline{y}$. In order to find a sharp estimate for ω, we need to bound $\|\, t \,\|$ in terms of $\|\, h \,\|$. Unfortunately, this is difficult to do for any of the standard norms, like the l_2-norm. However, it is easy to find sharp estimates in a local norm as is described below.

By multiplying the first equation in (14) by $(XS)^{-1/2}$ and by denoting

$$D = X^{-1/2}S^{1/2} \tag{15}$$

we obtain

$$\begin{aligned} Dt_x + D^{-1}t_s &= (XS)^{-1/2}(u\bar{z}_s + v\bar{z}_x) \\ Qt_x + Rt_v &= 0. \end{aligned} \tag{16}$$

From (2) and the second equation of (16) we have $t_x^T t_s \geq 0$ so that

$$\begin{aligned} \|Dt_x\|_2^2 + \|D^{-1}t_s\|_2^2 \leq \|Dt_x + D^{-1}t_s\|^2 &= \|(XS)^{-1/2}(u\bar{z}_s + v\bar{z}_x)\|^2 \\ &\leq (\min xs)^{-1}\|u\bar{z}_s + v\bar{z}_x\|_2^2 \end{aligned} \tag{17}$$

We can majorize (17) by using a two-dimensional Cauchy inequality as follows

$$\begin{aligned} \|u\bar{z}_s + v\bar{z}_x\|_2 &\leq \|u\bar{z}_s\|_2 + \|v\bar{z}_x\|_2 \\ &\leq \|Du\|_2\|D^{-1}\bar{z}_s\|_2 + \|D^{-1}v\|_2\|D\bar{z}_x\|_2 \\ &\leq \sqrt{\|Du\|_2^2 + \|D^{-1}v\|_2^2}\sqrt{\|D\bar{z}_x\|_2^2 + \|D^{-1}\bar{z}_s\|_2^2} \quad . \end{aligned} \tag{18}$$

If we introduce the following $\mathbb{R}^{2n}$-norm

$$\|q\|_z := \sqrt{\|Dq_x\|_2^2 + \|D^{-1}q_s\|_2^2}, \quad \forall q = \left[\, {q_x}^T,\, {q_s}^T \,\right]^T \in \mathbb{R}^{2n}, \tag{19}$$

then from (17) and (18) we deduce that

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

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

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

Thus we have obtained the following result:

Lemma 1. *If $z \in \mathcal{F}^0$ and $F'(z)$ is defined by (11) then*

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

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

$$\omega(z) = \frac{1}{\sqrt{\min xs}}. \tag{20}$$

The choice of the norm (19) was essential in obtaining this result. The 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 recent book of Renegar [44]. Local norms are also used in the analysis of Newton's method in the monograph of Nesterov and Nemirovskii [31].

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

$$\alpha(z, \tau) := \|F'(z)^{-1} F_\tau(z)\|_z \,. \tag{21}$$

It is convenient to introduce the Kantorovich measure of proximity

$$\kappa(z, \tau) := \alpha(z, \tau)\omega(z) \tag{22}$$

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

$$\mu := \mu(z) := \frac{x^T s}{n}. \tag{23}$$

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 sequence generated by Newton's method – and the sequence generated by the simplified Newton method – with starting point z will converge to the point $z(\tau)$ on the central path. We are now ready to describe 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 $\tau \leftarrow (1-\theta)\tau$;
  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 $(z^{k+1}, \tau_{k+1}) \leftarrow (z, \tau)$;
  Set $k \leftarrow (k+1)$;
until $(x^k)^T s^k \leq \epsilon$.

In Section 5 we prove that the parameter θ computed in the outer iteration is bounded below by a quantity of the form $\chi/\sqrt{n}$, where χ is a positive quantity depending only on κ_1 and κ_2. This implies that the number of outer iterations is $O(\sqrt{n}\log((x^{0\,T} s^0)/\epsilon)$. Since the hypothesis of the Kantorovich Theorem is satisfied, 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. There we will have $z \in \mathcal{F}_0$, since the interior point method is required to have a strictly feasible starting point. However, the following results hold for any $z \in \mathbb{R}^{2n}_{++}$.

First we bound the size of the solution of the following linear system.

$$\begin{aligned} su + xv &= a \\ Qu + Rv &= 0 \end{aligned} \tag{1}$$

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

$$\varsigma := \begin{cases} \sqrt{2}, & \text{if HLCP is monotone} \\ 1 & \text{if HLCP is skew-symmetric} \end{cases} . \tag{2}$$

Lemma 2. *If HLCP is monotone and $z \in \mathbb{R}^{2n}_{++}$ then (1) has a unique solution for any $a \in \mathbb{R}^n$ and the following inequalities are satisfied:*

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

Proof. By dividing the first equation in (1) with $\sqrt{xs}$ we obtain

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

where the diagonal matrix D is given by (15). It follows that

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

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

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

□

The above lemma allows us to establish a relationship in between $\alpha(z, \tau)$ defined in (21) and the proximity measure

$$\delta(z, \tau) := \left\| \sqrt{\frac{\tau}{xs}} - \sqrt{\frac{xs}{\tau}} \right\|_2 \tag{3}$$

considered in [14] (see also [46]).

Corollary 1. *For any $z \in \mathbb{R}^{2n}_{++}$ and any $\tau > 0$ we have*

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

where ς is given by (2).

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

Lemma 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 often be used in the sequel:

$$\omega(z) \ge (\mu(z))^{-1/2}\,. \tag{4}$$

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

$$(1 - \frac{\varsigma\,\kappa(z,\tau)}{\sqrt{n}})\mu(z) \le \tau \le (1 + \frac{\varsigma\,\kappa(z,\tau)}{\sqrt{n}})\mu(z)\,, \tag{5}$$

and

$$\delta^2(z,\mu(z)) \le (\varsigma\kappa(z,\tau))^2\left[1 + \frac{(\varsigma\kappa(z,\tau))^2}{n - (\varsigma\kappa(z,\tau))^2}\right]\,, \tag{6}$$

where ς is given by (2).

If HLCP (1) is skew-symmetric we also have

$$\delta^2(z,\mu(z)) \le \kappa^2(z,\mu(z))\,. \tag{7}$$

Proof. By using Corollary 1, Lemma 3 and (4) we deduce that

$$\begin{aligned}\varsigma^2\kappa^2(z,\tau) &= \varsigma^2\omega^2(z)\alpha^2(z,\tau) \ge \tau\,\omega^2(z)\delta^2(z,\tau) \ge \frac{\tau}{\mu}\,\delta^2(z,\tau)\\ &= \frac{\tau^2}{\mu^2}\,\delta^2(z,\mu(z)) + \frac{(\tau-\mu)^2}{\mu^2}\,n\,.\end{aligned} \tag{8}$$

The above inequality implies

$$\varsigma^2\kappa^2(z,\tau) \ge \frac{(\tau-\mu)^2}{\mu^2}\,n \tag{9}$$

and

$$\delta^2(z,\mu(z)) \le (\varsigma\kappa(z,\tau))^2 + \frac{\mu^2-\tau^2}{\tau^2}\,(\varsigma\kappa(z,\tau))^2 - \frac{(\tau-\mu)^2}{\tau^2}\,n\,. \tag{10}$$

Inequality (9) clearly implies (5). By denoting

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

we have

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

The derivative of f is given by

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

From the hypothesis of our corollary it follows that $t/n < 1$, and from the expression above it follows that f' is positive on the interval $(-\infty,\ t/n)$ and negative on the interval $(t/n,\ 1)$. 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 (6) follows from this observation and (10). Finally, (7) is obtained by setting $\varsigma = 1$ and $\tau = \mu$ in (8). □

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 [15], but we give a simple proof for the sake of completeness.

Lemma 4. *If*

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

then

$$\frac{1}{1+\eta+\sqrt{2\eta+\eta^2}}\,e \le \frac{xs}{\mu} \le (1+\eta+\sqrt{2\eta+\eta^2}\,)\,e\,.$$

Proof. We first notice 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

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

It follows that ν satisfies the following quadratic inequality:

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

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

$$1+\eta-\sqrt{2\eta+\eta^2-2\eta/n} \;\le \nu \le\; 1+\eta+\sqrt{2\eta+\eta^2-2\eta/n}\,.$$

This implies the following bounds are independent of n

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

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

As an immediate consequence of Corollary 2 and Lemma 4 we obtain the following result.

Corollary 3. *For any* $z=\left[x^T,\, s^T\right]^T \in \mathbb{R}^{2n}_{++}$ *and any* $\tau>0$ *with* $\varsigma\kappa(z,\tau)<\sqrt{n}$*, we have*

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

where

$$\psi=\psi(z,\tau)=1+\vartheta+\sqrt{2\vartheta+\vartheta^2}\,, \tag{11}$$

$$\vartheta=\vartheta(z,\tau)=\frac{1}{2}(\varsigma\kappa(z,\tau))^2\left[1+\frac{(\varsigma\kappa(z,\tau))^2}{n-(\varsigma\kappa(z,\tau))^2}\right], \tag{12}$$

and ς *is defined by (2).*

If HLCP (1) is skew-symmetric we also have

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

where

$$\varrho=\varrho(z)=1+.5\,\kappa^2(z,\mu(z))+\sqrt{\kappa^2(z,\mu(z))+.25\,\kappa^4(z,\mu(z))}\,. \tag{13}$$

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

$$\delta_K(z,\tau):=\frac{\|\, xs-\tau\, e\,\|_2}{\tau} \tag{14}$$

considered in [18].

Lemma 5. *For any* $(z,\tau)\in\mathbb{R}^{2n}_{++}\times\mathbb{R}_{++}$ *with* $\delta_K(z,\tau)<1$*, we have*

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

Proof. From the obvious relation

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

we deduce that

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

Also, by using (3) and Corollary 1 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 prove that Algorithm 1 is globally convergent for any choice of parameters $0 < \kappa_1 < \kappa_2 < .5$. More precisely, we 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 prove that each inner iteration terminates 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 1. *All points z generated by Algorithm 1 belong to $\mathcal{F}^0$. Moreover, if HLCP (1) is monotone, then*

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

and if HLCP (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 $\overline{z} > 0$ when $k = 0$. We also have $\kappa(\overline{z}) \leq \kappa_1$. According to the Kantorovich Theorem, all the points z generated in the inner iteration satisfy

$$\| z - \overline{z} \|_{\overline{z}} < \frac{1 - \sqrt{1 - 2\,\kappa_1}}{\omega(\overline{z}\,)} . \tag{2}$$

We prove that

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

If $z = \left[x^T ,\ s^T \right]^T$ and $\overline{z} = \left[\overline{x}^T ,\ \overline{s}^T \right]^T$ we can write

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

Since

$$\left\| \overline{x}^{-1}(\overline{x} - x) \right\|_\infty = \left\| (\overline{x}\,\overline{s})^{-1/2}\overline{D}(\overline{x} - x) \right\|_\infty$$
$$\leq \omega(\overline{z}) \left\| \overline{D}\,(\overline{x} - x) \right\|_\infty \leq \omega(\overline{z}) \, \| z - \overline{z} \|_{\overline{z}} < 1$$

we deduce that $x > 0$. The positivity of s is proved similarly.

We prove now that any $z = \left[\, x^T,\ s^T \,\right]^T$ produced during the inner iteration satisfies

$$Q\,x + R\,s = b\,.$$

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

$$su + xv = \tau e - xs$$
$$Qu + Rv = 0$$

$$z^+ = \left[\, x^{+T},\ s^{+T} \,\right]^T = z + \left[\, u^T,\ v^T \,\right]^T,\ z = \left[\, x^T,\ s^T \,\right]^T, \tag{3}$$

and we have

$$Q\,x^+ + R\,s^+ = Q\,x + R\,s + Q\,u + R\,v = Q\,x + R\,s = b\,.$$

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

$$\overline{s}u + \overline{x}v = \tau e - xs$$
$$Qu + Rv = 0$$

$$z^+ = \left[\, x^{+,T},\ s^{+T} \,\right]^T = z + \left[\, u^T,\ v^T \,\right]^T,\ z = \left[\, x^T,\ s^T \,\right]^T, \overline{z} = \left[\, \overline{x}^T,\ \overline{s}^T \,\right]^T. \tag{4}$$

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 (1) is an immediate consequence of Corollary 5 we only have to prove that if HLCP is skew-symmetric, then

$$\tau_k = \mu(z^k), \quad k = 0, 1, \ldots \tag{5}$$

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 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 (3) we have

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

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

After a simplified Newton step of the form (3) there holds

$$x^+s^+ = xs + (su + xv) + uv = \tau e + (s - \overline{s})u + (x - \overline{x})v + uv\,. \qquad (6)$$

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

$$0 = (u - \overline{u})^T(v - \overline{v}) = u^T v - (u^T\overline{v} + v^T\overline{u}) + \overline{u}^T\overline{v} = -(u^T\overline{v} + v^T\overline{u})\,.$$

The above equality and (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 $\kappa(z, \tau) \le \kappa_1$, and we want to determine the largest $\theta \in (0, 1)$ such that $\kappa(z, (1 - \theta)\tau) \le \kappa_2$ The following lemma enables us to find a lower bound for this θ.

Lemma 6. *For any* $(z, \tau, \theta) \in \mathbb{R}^{2n}_{++} \times \mathbb{R}_{++} \times (0, 1)$ *with* $\varsigma\,\kappa(z, \tau) < \sqrt{n}$, *we have*

$$\kappa(z, (1 - \theta)\tau) \le \kappa(z, \tau) + \theta\sqrt{\psi(z, \tau)}\sqrt{(\varsigma\,\kappa(z, \tau))^2 + (2\tau/\mu - 1)n}\,,$$

where ς *and* ψ *are defined by (2) and (11), respectively.*

In the skew-symmetric we have the slightly sharper estimate

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

where ϱ *is defined by (13).*

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

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

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

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

It follows that

$$\left[\, u(\theta)^T,\ v(\theta)^T \,\right]^T = \left[\, u^T,\ v^T \,\right]^T + \theta\left[\, \check{u}^T,\ \check{v}^T \,\right]^T\,,$$

where

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

and

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

Using (8) we obtain

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

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

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

Our claim is obtained by multiplying with $\omega(z)$ and using Corollary 3. We then substitute in $\varsigma = 1$ and $\tau = \mu$ to obtain the result in the skew-symmetric case. □

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 $i = 1, 2$:

$$\psi_i := \begin{cases} 1 + \vartheta_i + \sqrt{2\vartheta_i + \vartheta_i^2}\,, & \text{if HLCP is monotone} \\ 1 + \eta_i + \sqrt{2\eta_i + \eta_i^2}, & \text{if HLCP is skew-symmetric} \end{cases} \tag{7}$$

where

$$\vartheta_i = \kappa_i^2 \left(1 + \frac{\kappa_i^2}{1-\kappa_i^2} \right), \quad \eta_i = \frac{\kappa_i^2}{2}, \quad i = 1, 2\,. \tag{8}$$

Corollary 4. *The parameter θ determined at each outer iteration of Algorithm 1 satisfies*

$$\theta \geq \frac{\chi}{\sqrt{n}}\,,$$

where $\chi = \chi(\kappa_1, \kappa_2)$ is given by

$$\chi := \begin{cases} \dfrac{\sqrt{2}\,(\kappa_2 - \kappa_1)}{\sqrt{2 + \varsigma^2 \kappa_1{}^2}\,\sqrt{\psi_1}}\,, & \textit{if HLCP is skew-symmetric} \\ & \textit{or if no simplified Newton steps are performed} \\ \dfrac{\sqrt{2}(\kappa_2 - \kappa_1)}{(\sqrt{2} + \varsigma\kappa_1)\sqrt{\psi_1}}\,, & \textit{otherwise} \end{cases} \tag{9}$$

Proof. The result in the skew-symmetric case is an immediate consequence of Lemma 6. In the monotone HLCP case, if only Newton steps are performed, then

$$n\mu^+ = (x^+)^T (s^+) = (x+u)^T (s+v) = x^T s + u^T s + x^T v + u^T = n\tau + u^T v \geq n\tau\,.$$

Since in Algorithm 1 we begin with $\mu_0 = \tau_0$, we will always have $\mu \geq \tau$, and the result follows from Lemma 6. If simplified Newton directions are also performed, then according to Corollary 2 we have

$$\sqrt{(\varsigma\,\kappa(z,\tau))^2 + (2\tau/\mu - 1)n}$$
$$\leq \sqrt{(\varsigma\,\kappa(z,\tau))^2 + (2\varsigma\,\kappa(z,\tau)/\sqrt{n} + 1)n} = \varsigma\,\kappa(z,\tau) + \sqrt{n}$$

and by applying Lemma 6 we obtain the desired result. □

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

Theorem 2. *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*

$$\mathcal{N}(\kappa_1, \kappa_2) = \left\lceil \log_2 \left(\frac{\log_2(\zeta_{\mathcal{N}})}{\log_2\left((1 - \sqrt{1 - 2\kappa_2} - \kappa_2)/\kappa_2\right)} \right) \right\rceil, \tag{10}$$

and

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

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, where

$$\mathcal{S}(\kappa_1, \kappa_2) = \left\lceil \frac{\log_2(\zeta_{\mathcal{S}})}{\log_2(1 - \sqrt{1 - 2\kappa_2})} + 1 \right\rceil, \tag{11}$$

and

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

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

$$\kappa(\overline{z}, \tau) \leq \kappa_2 \tag{12}$$

and terminates when a point z is obtained such that

$$\kappa(z, \tau) \leq \kappa_1\,. \tag{13}$$

Note that $\overline{z}$ 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 $\overline{z}$, converge to the solution $z(\tau)$ of the equation $F_\tau(z) = 0$. In what follows we prove that (13) is satisfied provided $\|\, z - z(\tau)\,\|_{\overline{z}}$ is small enough.

Lemma 5 shows that (13) is satisfied if

$$\delta_K(z,\tau) \le \frac{\kappa_1}{1+\kappa_1}\,. \tag{14}$$

With

$$\overline{z} = \left[\,\overline{x}^T,\ \overline{s}^T\,\right]^T,\ z = \left[\,x^T,\ s^T\,\right]^T,\ \overline{D} = \overline{X}^{\,-1/2}\overline{S}^{\,1/2}$$

we can write

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

In order to bound $\|\,\overline{D}\,x\,\|_\infty$ we use Corollary 3 and the identity $\overline{D}\,\overline{x} = (\overline{x}\,\overline{s})^{1/2}$ to obtain

$$\|\,\overline{D}\,x\,\|_\infty \le \|\,\overline{D}\,\overline{x}\,\|_\infty + \|\,\overline{D}\,(x - \overline{x}\,)\,\|_\infty \le \sqrt{\psi_2\,\mu(\overline{z})}\ + \|\, z - \overline{z}\,\|_{\overline{z}}\,.$$

Similarly we get

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

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 $\overline{z}$, the Kantorovich Theorem ensures that

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

Taking into account (4) we deduce that

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

Therefore,

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

It follows that (14) is satisfied if

$$\|\, z - z(\tau)\,\|_{\overline{z}} \le \frac{\tau}{\sqrt{\mu}}\,\upsilon, \tag{15}$$

where

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

If Newton's method is used in the inner iteration then from the Kantorovich Theorem it follows that (15) is satisfied after m Newton steps provided

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

From Corollary 2 we obtain the following bound

$$\frac{\tau}{\sqrt{\bar{\mu}}} = \frac{\tau\sqrt{\bar{\mu}}}{\bar{\mu}} \ge (1-\varsigma\kappa_2/\sqrt{2})\,\sqrt{\bar{\mu}}\,. \tag{16}$$

Using (16) together with the fact

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

and with (4), we deduce that (15) is satisfied if

$$\lambda_2^{2^m} \le \frac{(1-\varsigma\kappa_2/\sqrt{2})(1-\lambda_2^2)\,\upsilon}{2\sqrt{1-2\kappa_2}}.$$

It is easily seen that the right hand side coincides with $\zeta_{\mathcal{N}}$. By taking the logarithm of both sides twice, 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 (12) we deduce in a similar fashion that the Kantorovich theorem implies (15) is satisfied after m steps if

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

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

Corollary 5. *Algorithm 1 terminates using at most* $O(\sqrt{n}\log(\frac{x^{0\,T}s^0}{\epsilon}))$ *Newton and simplified Newton steps.*

.

Proof. The argument is standard. From Corollary 4 it follows that

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

and the assertion follows by using $\tau_0 = \mu(z^0)$ together with Proposition 1 and Theorem 2. □

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 iterations $\mathcal{N}$ 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 duality 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 are only upper bounds. They are obtained by substituting numerical values in the formulae (9), (10) and (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, for example, (5)). 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 6. *If the HLCP is monotone and only Newton directions are performed, then:*

1. $\mathcal{N}(.12,\ .24) = 1, \quad \chi(.12,\ .24) > .1\,;$
2. $\mathcal{N}(.21,\ .42) = 2, \quad \chi(.21,\ .42) > .17\,;$
3. $\mathcal{N}(.24,\ .48) = 3, \quad \chi(.24,\ .48) > .196\,.$
4. $\mathcal{N}(.245,\ .49) = 4, \quad \chi(.245,\ .49) > .199\,.$

Corollary 7. *If the HLCP is monotone and simplified Newton directions are performed, then:*

1. $\mathcal{S}(.12,\ .24) = 1, \quad \chi(.12,\ .24) > .098\,;$
2. $\mathcal{S}(.21,\ .42) = 5, \quad \chi(.21,\ .42) > .149$
3. $\mathcal{S}(.24,\ .48) = 12, \quad \chi(.24,\ .48) > .162\,.$
4. $\mathcal{S}(.245,\ .49) = 18, \quad \chi(.245,\ .49) > .164\,.$

Corollary 8. *If the HLCP is skew-symmetric then:*

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

Acknowledgements. The original version of this paper [38] was written when the author was visiting the Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB). The author would like to thank Peter Deuflhard, 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. The author would also like to thank two anonymous referees whose comments and suggestions led to a better presentation of our results.

References

1. Alefeld, G.E., Potra, F.A., Shen, Z.: On the existence theorems of Kantorovich, Moore and Miranda. Computing Supplementum **15**, 21–28 (2001)
2. Amer, S.M.: On solution of nonlinear singular integral equations with shift in generalized Hölder space. Chaos Solitons Fractals **12** (7), 1323–1334 (2001)

3. Begehr, H., Efendiev, M.A.: On the asymptotics of meromorphic solutions for nonlinear Riemann-Hilbert problems. Math. Proc. Cambridge Philos. Soc. **127** (1), 159–172 (1999)
4. Cottle, R.W., Pang, J.-S., Stone, R.E.: m The Linear Complementarity Problem. Academic Press, Boston, MA, 1992
5. De Pascale, E., Zabrejko, P.P.: New convergence criteria for the Newton-Kantorovich method and some applications to nonlinear integral equations. Rend. Sem. Mat. Univ. Padova **100**, 211–230 (1998)
6. Dennis, J.E. Jr.: On the Kantorovich hypothesis for Newton's method. SIAM J. Numer. Anal. **6**, 493–507 (1969)
7. Dennis, J.E. Jr.: On the convergence of Newton-like methods. In: Numerical methods for nonlinear algebraic equations (Proc. Conf., Univ. Essex, Colchester, 1969), pp. 163–181. Gordon and Breach, London, 1970
8. Deuflhard, P., Heindl, G.: Affine invariant convergence theorems for Newton's method and extensions to related methods. SIAM J. Numer. Anal. **16**, 1–10 (1979)
9. Ezquerro, J.A., Hernández, M.A.: Multipoint super-Halley type approximation algorithms in Banach spaces. Numer. Funct. Anal. Optim., **21** (7-8), 845–858 (2000)
10. Ezquerro, J.A., Hernández, M.A.: A special type of Hammerstein integral equations. Int. Math. J. **1** (6), 557–566 (2002)
11. Ferreira, O.P., Svaiter, B.F.: Kantorovich's theorem on Newton's method in Riemannian manifolds. J. Complexity **18** (1), 304–329 (2002)
12. Gragg, W.B., Tapia, R.A.: Optimal error bounds for the Newton-Kantorovich theorem. SIAM J. Numer. Anal. **11**, 10–13 (1974)
13. Grigat, E., Sachs, G.: Predictor-corrector continuation method for optimal control problems. In: Variational calculus, optimal control and applications (Trassenheide, 1996), volume **124** of Internat. Ser. Numer. Math., pp. 223–232. Birkhäuser, Basel, 1998
14. Jansen, B., Roos, C., Terlaky, T., Vial, J.-P.: Primal-dual algorithms for linear programming based on the logarithmic barrier method. J. Optim. Theory Appl. **83** (1), 1–26 (1994)
15. Jansen, B., Roos, C., Terlaky, T., Vial, J.-Ph.: Primal-dual target-following algorithms for linear programming. Ann. Oper. Res. **62**, 197–231 (1996) Interior point methods in mathematical programming
16. Kantorovich, L.: On Newton's method for functional equations (Russian). Dokl. Akad. Nauk. SSSR **59**, 1237–1240 (1948)
17. Kantorovich, L.V., Akilov, G.P.: Functional analysis in normed spaces, volume **46** of International Series of Monographs in Pure and Applied Mathematics. Pergamon Press, Oxford, 1964. Translated from the Russian by D. E. Brown and Dr. A. P. Robertson
18. Kojima, M., Megiddo, N., Noma, T., Yoshise, A.: A unified approach to interior point algorithms for linear complementarity problems, volume **538** of Lecture Notes in Comput. Sci. Springer-Verlag, New York, 1991
19. Laumen, M.: A Kantorovich theorem for the structured PSB update in Hilbert space. J. Optim. Theory Appl. **105** (2), 391–415 (2000)
20. Matveev, A.F., Yunganns, P.: On the construction of an approximate solution of a nonlinear integral equation of permeable profile. Differ. Uravn. **33** (9), 1242–1252, 1295 (1997)
21. Mayer, J.: A generalized theorem of Miranda and the theorem of Newton-Kantorovich. Numer. Funct. Anal. Optim. 23(3-4), 333–357 (2002)
22. Miel, G.: An updated version of the Kantorovich theorem for Newton's method. Computing **27**, 237 (1981)
23. Miel, G.J.: The Kantorovich theorem with optimal error bounds. American Mathematical Monthly **86** (3), 212–215 (1979)
24. Moret, I.: A note on Newton type iterative methods. Computing **33**, 65–73 (1984)
25. Moret, I.: On a general iterative scheme for Newton-type methods. Numerical Functional Analysis and Optimization **9**, 1115–1137 (1987)
26. Moret, I.: A Kantorovich-type theorem for inexact Newton methods. Numer. Funct. Anal. Optim., **10** (3-4), 351–365 (1989)
27. Mukaidani, H., Shimomura, T., Mizukami, K.: Asymptotic expansions and a new numerical algorithm of the algebraic Riccati equation for multiparameter singularly perturbed systems. J. Math. Anal. Appl. **267** (1), 209–234 (2002)
28. Nagatou, K., Yamamoto, N., Nakao, M.T.: An approach to the numerical verification of solutions for nonlinear elliptic problems with local uniqueness. Numer. Funct. Anal. Optim. **20** (5-6), 543–565 (1999)
29. Nashed, M.Z., Chen, X.: Convergence of Newton-like methods for singular operator equations using outer inverses. Numer. Math. **66** (2), 235–257 (1993)
30. Nayakkankuppam, M.V., Overton, M.L.: Conditioning of semidefinite programs. Math. Program. **85** (3, Ser. A), 525–540 (1999)
31. Nesterov, Y., Nemirovsky, A.: Interior Point Polynomial Methods in Convex Programming. SIAM Publications. SIAM, Philadelphia, 1994

32. Ortega, J.M.: The Newton-Kantorovich theorem. Amer. Math. Monthly **75**, 658–660 (1968)
33. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970
34. Ostrowski, A.M.: Solution of Equations in Euclidean and Banach Spaces. Academic Press, New York, 1973
35. Potra, F.A.: An error analysis for the secant method. Numer. Math. **38**, 427–445 (1982)
36. Potra, F.A.: On the aposteriori error estimates for Newton's method. Beitraege Numer. Math. **12**, 125–138 (1984)
37. Potra, F.A.: Sharp error bounds for a class of Newton–like methods. Libertas Mathematica **5**, 71–84 (1985)
38. Potra, F.A.: A path-following method for linear complementarity problems based on the affine invariant Kantorovich theorem. ZIB-Report 00-30, Konrad-Zuse-Zentrum, Berlin, August 2000
39. Potra, F.A., Pták, V.: Sharp error bounds for Newton's process. Numer. Math. **34**, 63–72 (1980)
40. Potra, F.A., Pták, V.: Nondiscrete induction and iterative processes. Number 103 in Research Notes in Mathematics. John Wiley & Sons, Boston–London–Melbourne, 1984
41. Potra, F.A., Wright, S.J.: Interior-point methods. Journal of Computational and Applied Mathematics **124**, 281–302 (2000)
42. Rall, L.B.: Computational solution of nonlinear operator equations. Krieger Huntington, New York, 1979
43. Renegar, J.: A polynomial-time algorithm, based on Newton's method, for linear programming. Math. Programming **40** (1,Ser. A), 59–93 (1988)
44. Renegar, J.: A mathematical view of interior-point methods in convex optimization. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001
45. Renegar, J., Shub, M.: Unified complexity analysis for Newton LP methods. Math. Programming **53** (1, Ser. A), 1–16 (1992)
46. Roos, C., Vial, J.-Ph., Terlaky, T.: Theory and algorithms for linear optimization: an interior point approach. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997
47. Smale, S.: 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), pp. 185–196. Springer, New York, 1986
48. Stoffer, D., Palmer, K.J.: Rigorous verification of chaotic behaviour of maps using validated shadowing. Nonlinearity **12** (6), 1683–1698 (1999)
49. Tapia, R.A.: The Kantorovich theorem for Newton's method. American Mathematical Monthly **78**, 389–392 (1971)
50. Tsuchiya, T.: An application of the Kantorovich theorem to nonlinear finite element analysis. Numer. Math. **84** (1), 121–141 (1999)
51. Wang, X., Li, C., Lai, M.-J.: A unified convergence theory for Newton-type methods for zeros of nonlinear operators in Banach spaces. BIT **42** (1), 206–213 (2002)
52. Wright, S.J.: Primal–Dual Interior–Point Methods. SIAM Publications, Philadephia, 1997
53. Yamamoto, T.: A unified derivation of several error bounds for Newton's process. Journal of Computation and Applied Mathematics **12**, 179–191 (1985)
54. Yamamoto, T.: Error-bounds for Newton iterates derived from the Kantorovich-theorem. Numer. Math. **48**, 91–98 (1986)
55. Yamamoto, T.: A method for finding sharp error-bounds for Newton method under the Kantorivich assumptions. Numer. Math. **49**, 203–220 (1986)
56. Yamamoto, T.: Historical developments in convergence analysis for Newton's and Newton-like methods. J. Comput. Appl. Math. **124** (1-2), 1–23 (2000), Numerical analysis 2000, Vol. IV, Optimization and nonlinear equations
57. Ye, Y.: Interior Point Algorithms : Theory and Analysis. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997