

PRIMAL-DUAL AFFINE SCALING INTERIOR POINT METHODS FOR LINEAR COMPLEMENTARITY PROBLEMS *

FLORIAN A. POTRA †

Abstract. A first order affine scaling method and two m th order affine scaling methods for solving monotone linear complementarity problems (LCP) are presented. All three methods produce iterates in a wide neighborhood of the central path. The first order method has $O(nL^2(\log nL^2)(\log \log nL^2))$ iteration complexity. If the LCP admits a strict complementary solution then both the duality gap and the iteration sequence converge superlinearly with Q-order two. If $m = \Omega(\log(\sqrt{n}L))$, then both higher order methods have $O(\sqrt{n}L)$ iteration complexity. The Q-order of convergence of one of the methods is $(m + 1)$ for problems that admit a strict complementarity solution while the Q-order of convergence of the other method is $(m + 1)/2$ for general monotone LCPs.

Key words. linear complementarity, interior-point, affine scaling, large neighborhood, super-linear convergence

AMS subject classifications. 90C51, 65K05, 49M15, 90C05, 90C20

1. Introduction. The primal-dual affine scaling direction plays a special role in the theory and practice of interior point methods. It turns out that the search direction used by most primal-dual interior point methods is a convex combination $(1 - \gamma)w + \gamma\bar{w}$ of the primal-dual affine scaling method w and the centering direction \bar{w} (see the monographs [35, 48, 49]). Optimality is improved along the affine scaling direction, while centrality is improved on the centering direction. The first interior point method of this type was proposed, in the context of linear programming (LP), by Kojima, Mizuno and Yoshise [15]. They proved that the algorithm had $O(nL)$ iteration complexity, the same as Karmarkar's algorithm [12]. Shortly after that, they improved the algorithm, and generalized it for monotone linear complementarity problems (LCP) [14]. The improved algorithm has $O(\sqrt{n}L)$ iteration complexity. This iteration complexity was first obtained by Renegar [34] for an interior point method that follows the primal central path of LP and remains the best iteration complexity known to date. The algorithm of [14] follows the primal-dual central path. Starting with a point $z^0 \in \mathcal{N}_2(\alpha)$ in a (small) neighborhood of the primal-dual central path, the algorithm takes at each iteration a unit step along the direction $(1 - \gamma)w + \gamma\bar{w}$, where $\gamma = 1 - \alpha / ((1 - \alpha)\sqrt{n})$, producing a sequence of iterates (z^k) that remains in $\mathcal{N}_2(\alpha)$. A similar algorithm was independently proposed and investigated by Monteiro and Adler for LP [21], and for quadratic programming (QP) [22].

While the algorithms mentioned above attain the best known iteration complexity, their practical performance is not satisfactory because they generate points in a small neighborhood of the central path and use a fixed stepsize (of one) along a direction that is dominated by the centering direction. In order to alleviate some of these problems, Kojima, Mizuno and Yoshise [16], proposed a potential reduction algorithm for LCP with $O(\sqrt{n}L)$ iteration complexity. The search direction is of the same type with $\gamma = 1 / (1 + \sqrt{n})$. However, instead of taking a fixed stepsize along this direction, a line search is performed in order to ensure that the iterates remain strictly feasible,

*Work supported by the National Science Foundation under Grant No. 0139701, and by the National Institute of Health under Grant No. R01GM075298-01 .

†Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 22150 (potra@math.umbc.edu).

and that the Tanabe-Todd-Ye potential function [43, 44] is reduced by a fixed constant at each iteration.

Although the potential reduction method of [16] is more flexible than the short step path following methods of [14, 21, 22], it cannot attain superlinear convergence since its search direction has a fixed centering component. There are basically three ways to obtain superlinear convergence: to take $\gamma = 0$ every second step, to use a sequence of γ 's that converges to zero, or to take $\gamma = 0$ at every step. The methods from the first category are called predictor-corrector method, while the methods from the last category are called affine scaling methods and they form the subject of the present paper.

Zhang, Tapia and Dennis [52] gave sufficient conditions for a class of interior point methods to produce a sequence of iterates with duality gap converging superlinearly to zero. However, no example of an algorithm satisfying those conditions and having polynomial complexity was given. The results of [52] were generalized for LCP in [53]. The first interior point method having both polynomial complexity and superlinear convergence was the predictor-corrector method of Mizuno, Todd and Ye (MTY). This method was proposed for LP in [20], where it was shown to have $O(\sqrt{n}L)$ iteration complexity. Shortly after that, Ye et al. [51] proved that the duality gap of the iterates produced by MTY converges quadratically to zero. MTY was generalized to LCP in [11], and the resulting algorithm was proved to have $O(\sqrt{n}L)$ iteration complexity under general conditions, and superlinear convergence under the assumption that the LCP has a (perhaps not unique) strictly complementary condition (i.e., the LCP is nondegenerate) and the iteration sequence converges. From [3] it follows that the latter assumption always holds. Subsequently, Ye and Anstreicher [50] proved that MTY converges quadratically assuming only that the LCP is nondegenerate. The nondegeneracy assumption is not restrictive, since according to [24] a large class of interior point methods, which contains MTY, can have only linear convergence if this assumption is violated.

The largest step path-following method (LSPF) determines γ at each iteration such that the duality gap is minimized under the constraint that the point obtained by taking a unit step along the search direction $(1 - \gamma)w + \gamma\bar{w}$ remains in $\mathcal{N}_2(\alpha)$. McShane [17] proved that LSPF has $O(\sqrt{n}L)$ iteration complexity for general (monotone) LCP's and superlinear convergence under the assumption that the LCP is nondegenerate and the iteration sequence converges. The superlinear convergence is a consequence of the fact that the γ 's converge to zero. Gonzaga [6] proved superlinear convergence assuming only nondegeneracy, and showed that with the addition of a computationally trivial safeguard LSPF achieves Q-quadratic convergence. We mention that the convergence of the iteration sequence generated by LSPF follows from the general results of Bonnans and Gonzaga [3].

It took considerable effort to obtain superlinear convergence results for potential reduction algorithms. One of the first results of this type was obtained for LP by Tunçel [46]. His algorithm uses a direction of the form $(1 - \gamma)w + \gamma\bar{w}$ where γ converges to zero at the same rate as the duality gap. The stepsize along this direction is obtained by minimizing a special, nonsmooth, potential function. The algorithm has $O(nL)$ iteration complexity for general LP's and Q-quadratic convergence of the duality gap under the assumption that the LP is nondegenerate. This assumption is rather strong since it implies that the Jacobian of the corresponding LCP is nonsingular at the solution. This strong restriction was removed by Tütüncü [47], who proposed a potential reduction method with $O(nL)$ iteration complexity and Q-quadratic conver-

gence of both the duality gap and the iteration sequence for general LP's. The search direction is the reduced Newton direction of a multiplicative variant of the Tanabe-Todd-Ye potential function, which is not exactly of the form $(1 - \gamma)w + \gamma\bar{w}$, but it is shown that it asymptotically approaches the affine scaling direction thus ensuring superlinear convergence. The algorithm uses the largest stepsize along this direction such that the feasibility of the iterates is maintained and the original Tanabe-Todd-Ye potential function is reduced by a fixed quantity at each iteration.

All the superlinear convergence results mentioned above are obtained either for LP, which always has a strictly complementary solution, or for LCP's that admit such a solution. As mentioned before a large class of first order interior point methods cannot have superlinear convergence without this assumption [24]. In order to obtain superlinear convergence for degenerate LCP's one has to either use nonstandard interior point methods or to consider higher order methods. The first approach was first taken by Mizuno [18], who used the so called step variant of the Tapia indicator [5] to identify the variables that are not strictly complementary, and modified the MTY algorithm in order to accelerate the convergence to zero of those variables. Mizuno's result was refined in [33, 32]. The drawback of this approach is that it needs the estimation of the partition of the optimal face in order to achieve superlinear convergence. The second approach was taken by Sturm [42] who constructed a second order MTY type algorithm that has $O(\sqrt{n}L)$ iteration complexity and Q-superlinear convergence of order 1.5 for degenerate LCP's. By using m 'th order derivatives, Stoer, Wechs and Mizuno [40] constructed higher order MTY type algorithms with Q-order $m + 1$ for nondegenerate LCP's and $(m + 1)/2$ for degenerate LCP's. The complexity of the predictor-corrector algorithm for degenerate LCP's from [40] is analyzed in [37]. It follows that for monotone LCP's with feasible starting points the algorithm has $O(\sqrt{n}L)$ iteration complexity.

The path-following interior point algorithms described in the paragraph above use a small neighborhood of the central path. However, extensive numerical experiments show convincingly that predictor-corrector methods perform better when using large neighborhoods of the central path. Predictor-corrector methods of MTY type are more difficult to develop and analyze in such neighborhoods, since the correctors are rather inefficient there. For example, it is known that one needs $O(n)$ corrector steps in order to reduce the δ_∞ proximity measure by a factor of .5 (see [2]). Therefore a straightforward generalization of the MTY algorithm would have $O(n^{1.5}L)$ iteration complexity. Gonzaga [7] proposed an interior point method in the $\mathcal{N}_\infty(\alpha)$ neighborhood of the central path, where a predictor is followed by an a priori unknown number of correctors, with $O(nL)$ -iteration complexity [7]. No superlinear convergence results are known for this algorithm. An interior point method for LP, acting in a large neighborhood of the central path defined by a self-regular proximity measure, with $O(\sqrt{n}L \log n)$ iteration complexity and superlinear convergence, was proposed in [26]. The first order predictor-corrector method from [30] uses the wide $\mathcal{N}_\infty^-(\alpha)$ neighborhood, takes one predictor followed by exactly one corrector at each iteration, and has $O(nL)$ iteration complexity for general monotone LCP's and Q-quadratic convergence for nondegenerate LCP's. The m 'th order predictor-corrector methods of [30] also use the $\mathcal{N}_\infty^-(\alpha)$ neighborhood, and their Q-order of convergence is $m + 1$ for nondegenerate LCP's and $(m + 1)/2$ for degenerate LCP's. If $m = n^\omega$, for some $\omega > 0$, then they have $O(\sqrt{n}L)$ iteration complexity. We note that the m 'th order method uses two matrix factorizations and m backsolves per iteration, so that if $\omega < 1$ then the computational cost per iteration is dominated by the cost of the two matrix

factorizations. The results of [30] were generalized for sufficient LCP's in [31]. An m -th order interior point algorithm for sufficient linear complementarity problems in the $\mathcal{N}_\infty^-(\alpha)$ neighborhood, that requires only one matrix factorization and m back-solves per iteration was proposed by Stoer [41]. The Q-order of convergence is m for nondegenerate problems. The algorithm is modified in [36] so that its asymptotic Q-order of convergence is $m+1$ for nondegenerate LCP's and $(m+1)/2$ for degenerate LCP's. No complexity results have been proved for the algorithms in [36, 41].

Primal-dual affine scaling methods use the pure primal-dual affine scaling method at each iteration. The first method of this type was proposed by Monteiro, Adler and Resende [23]. Their algorithm takes a small fixed step (of length $1/(nL)$) along this direction, and has $O(nL^2)$ iteration complexity. Because of the fixed stepsize, the duality gap is reduced only by a factor of $(1 - 1/(nL))$ at each iteration and therefore the algorithm cannot be superlinearly convergent. Moreover, only the first nL^2 iterates are guaranteed to be feasible, so that the algorithm may not produce an infinite sequence. In the same paper, Monteiro, Adler and Resende [23] propose an m th order affine scaling method for QP with $O(n^{\frac{m+1}{2m}} L^{\frac{m+1}{2m}})$ iteration complexity. They use again a small fixed stepsize (depending on n, m and L) so that the algorithm cannot have superlinear convergence. In order to overcome the inherent inefficiency caused by the use of a fixed stepsize, Mizuno and Nagasawa [19] proposed an affine scaling method for LP, where at each iteration the stepsize is determined by a line search that ensures the feasibility of the iterates and the fact that the Tanabe-Todd-Ye potential function remains bounded. By choosing the parameter defining the potential function equal to $1/L$ they obtain $O(nL^2)$ iteration complexity. However, no superlinear convergence results are obtained. Tunçel [45] analyzes a variant of this affine scaling where Tunçel's potential function [46] is used instead of the Tanabe-Todd-Ye potential function. The iteration complexity of the resulting algorithm is again $O(nL^2)$ when the parameter δ defining this potential function is equal to $1/(2L)$. The first affine scaling method with polynomial complexity and superlinear convergence was obtained, to our knowledge, by Monteiro and Wright [25] for LCPs. The stepsize along the affine scaling direction is obtained in such a way that the iterates are feasible and remain in neighborhood of the central path depending on two parameters η and δ . This neighborhood, which is closely related to Tunçel's potential function, is relatively narrow for large values of the duality gap, but it widens considerably as the duality gap approaches zero. It is shown that by choosing $\delta = \Theta(1/L)$ one obtains $O(nL^2)$ iteration complexity for monotone LCPs. It is shown that if the LCP has a strictly complementary solution then the Q-order of convergence of the duality gap is $2 - \delta\xi$, where $\xi = 1$ for skew symmetric problems and $\xi = 2$ otherwise. As a byproduct Monteiro and Wright show that Tunçel's affine scaling method [45] for LP has Q-order $2 - \delta$.

We note that the primal-dual affine scaling direction considered in this paper is sometime called the classical primal-dual affine scaling direction, in order to distinguish it from the so-called Dikin primal-dual affine scaling direction considered by Jansen, Roos and Terlaky [8] for LP and generalized for LCP in [9]. The corresponding Dikin type primal-dual affine scaling algorithms have $O(nL)$ iteration complexity. By using an m th order Dikin primal-dual affine scaling direction Jansen et al. [10] have obtained an algorithm with $O(n^{\frac{m+1}{2m}} L)$ iteration complexity. Unlike the algorithms from [8, 9] which used a small fixed step size, the algorithm from [10] may take a larger step size which is determined at each iteration such that the new point remains in the $\mathcal{N}_\infty(\alpha)$ neighborhood of the central path. No superlinear results have been proved for Dikin type primal-dual affine scaling algorithms. In fact, since the

Dikin primal-dual affine scaling direction contains an important centering component, so superlinear convergence is expected for such algorithms.

In the present paper we present an affine scaling algorithm that uses the $\mathcal{N}_{\infty}^{-}(\alpha)$ neighborhood of the central path. It has $O(nL^2\phi(nL^2))$ iteration complexity for general monotone LCP's, where $\phi(t) = (\log t)(\log \log t)$. If the LCP admits a strictly complementary solution, then the duality gap converges Q-superlinearly to zero and the iteration sequence converges Q-superlinearly to a strictly complementary solution. The Q-orders of convergence of both sequences are equal to two. We also present two m th order affine scaling methods having $O((\sqrt{n}L)^{\frac{m+1}{m}}(\phi(\sqrt{n}L))^{\frac{1}{m}})$ iteration complexity for monotone LCPs. By taking $m = \Omega(\log(\sqrt{n}L))$ we obtain $O(\sqrt{n}L)$ iteration complexity, the best iteration complexity known so far for LP, QP, and monotone LCP. We note that both algorithms use one matrix factorization and m backsolves per iteration. This requires $O(n^3 + mn^2)$ arithmetic operations per iteration. Therefore if $\log(\sqrt{n}L) \ll n$, then the cost of an iteration is dominated by the cost of the matrix factorization.

If the LCP admits a strictly complementary solution, then the iterates produced by the first algorithm converges Q-superlinearly to a strictly complementary solution and the corresponding duality gaps converge Q-superlinearly to zero. The Q-orders of convergence of both sequences are equal to $m + 1$. If $m \geq 2$ then the second algorithm produces a sequence of iterates that converges Q-superlinearly to a maximal complementarity solution, even for degenerate LCPs, with Q-order $(m + 1)/2$. The sequence of duality gaps converges Q-superlinearly to zero with the same Q-order. To our knowledge this is the first affine scaling method acting in the $\mathcal{N}_{\infty}^{-}(\alpha)$ neighborhood of the central path that has $O(\sqrt{n}L)$ iteration complexity and Q-superlinear convergence in the absence of strict complementarity.

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 κ , $\lceil \kappa \rceil$ denotes the smallest integer greater or equal to κ . 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 $A(uv) \neq (Au)v$. 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 $v \in \mathbb{R}_+^n$ and $\lambda \in \mathbb{R}$, then \sqrt{v} denotes the vector with components $\sqrt{v_i}$, and $\lambda - v$ denotes the vector with components $\lambda - v_i$. Traditionally the vector $\lambda - v$ is written as $\lambda e - v$, 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 $\|U\|_2 = \|u\|_{\infty}$.

We frequently use the $O(\cdot)$ and $\Omega(\cdot)$ notation to express asymptotic relationships between functions. The most common usage will be associated with a sequence $\{x^k\}$ of vectors and a sequence $\{\tau_k\}$ of positive real numbers. In this case $x^k = O(\tau_k)$ means that there is a constant K (dependent on problem data) such that for every

$k \in \mathbb{N}$, $\|x^k\| \leq K\tau_k$. Similarly, if $x^k > 0$, $x^k = \Omega(\tau_k)$ means that $(x^k)^{-1} = O(1/\tau_k)$. If we have both $x^k = O(\tau_k)$ and $x^k = \Omega(\tau_k)$, we write $x^k = \Theta(\tau_k)$

If $x, s \in \mathbb{R}^n$, then the vector $z \in \mathbb{R}^{2n}$ obtained by concatenating x and s is denoted by $z = [x, s] = [x^T, s^T]^T$, and the mean value of xs is denoted by $\mu(z) = \frac{x^T s}{n}$.

2. The linear complementarity problem and its central path.

2.1. The horizontal 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 $z = [x, s]$ such that

$$(2.1) \quad \begin{aligned} xs &= 0 \\ Qx + Rs &= b \\ x, s &\geq 0. \end{aligned}$$

The standard (monotone) linear complementarity problem is obtained by taking $R = -I$, and Q positive semidefinite. There are other formulations of the linear complementarity problems as well but, as shown in [1], all popular formulations are equivalent, and the behavior of a large class of interior point methods is identical on those formulations, so that it is sufficient to prove results only for one of the formulations. We have chosen HLCP because of its symmetry. The linear programming problem, and the quadratic programming problem, can be formulated as an HLCP. Therefore, HLCP provides a convenient general framework for studying interior point methods.

Throughout this paper we assume that the HLCP is monotone, in the sense that:

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

This condition is satisfied if the HLCP is a reformulation of a QP [4]. If the HLCP is a reformulation of an LP then the following stronger condition holds

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

and HLCP is called skew-symmetric in this case. The following proposition contains two simple properties of a monotone HLCP.

PROPOSITION 2.1. *If HLCP is monotone then the $n \times 2n$ -matrix (Q, R) has full rank, and*

$$(u^T \bar{v} + \bar{u}^T v)^2 \leq 4(u^T v)(\bar{u}^T \bar{v}), \quad \forall [u, v], [\bar{u}, \bar{v}] \in \text{Ker}(Q, R).$$

Proof. The full rank property follows from [3]. If $Qu + Rv = Q\bar{u} + R\bar{v} = 0$, then $Q(u + \lambda\bar{u}) + R(v + \lambda\bar{v}) = 0$, for any $\lambda \in \mathbb{R}$, so that

$$0 \leq (u + \lambda\bar{u})^T (v + \lambda\bar{v}) = u^T v + \lambda(u^T \bar{v} + \bar{u}^T v) + \lambda^2 \bar{u}^T \bar{v}, \quad \forall \lambda \in \mathbb{R}.$$

Since the right hand side above is a nonnegative quadratic function in λ , its discriminant must be nonpositive, which proves the second part of the proposition. \blacksquare

In the skew-symmetric case we can often obtain sharper estimates, due to the following consequence of Proposition 2.1.

COROLLARY 2.2. *If HLCP is skew-symmetric then $u^T \bar{v} + \bar{u}^T v = 0$, for all u, v, \bar{u}, \bar{v} satisfying $Qu + Rv = Q\bar{u} + R\bar{v} = 0$.*

Let us denote the set of all feasible points of HLCP by

$$\mathcal{F} = \{z = [x, s] \in \mathbb{R}_+^{2n} : Qx + Rs = b\},$$

and the solution set (or the optimal face) of HLCP by

$$\mathcal{F}^* = \{z^* = [x^*, s^*] \in \mathcal{F} : x^* s^* = 0\}.$$

The structure of \mathcal{F}^* is very important in the analysis of interior point methods. Let us define three subsets \mathcal{B} , \mathcal{N} and \mathcal{J} of the index set $\{1, \dots, n\}$ by

$$\begin{aligned} \mathcal{B} &= \{i = 1, \dots, n \mid x_i^* > 0 \text{ for at least one } [x^*, s^*] \in \mathcal{F}^*\}, \\ \mathcal{N} &= \{i = 1, \dots, n \mid s_i^* > 0 \text{ for at least one } [x^*, s^*] \in \mathcal{F}^*\}, \\ \mathcal{J} &= \{i = 1, \dots, n \mid x_i^* = s_i^* = 0 \text{ for all } [x^*, s^*] \in \mathcal{F}^*\}. \end{aligned}$$

One can prove that \mathcal{B} , \mathcal{N} , and \mathcal{J} form a partition of $\{1, \dots, n\}$ and that there exists a solution $[x^*, s^*] \in \mathcal{F}^*$ such that $x_{\mathcal{B}}^* > 0$ and $s_{\mathcal{N}}^* > 0$. Such a solution is called a *maximal complementarity solution*, since one can prove that for any $[x^*, s^*] \in \mathcal{F}^*$, $x_i^* > 0 \Rightarrow i \in \mathcal{B}$ and $s_j^* > 0 \Rightarrow j \in \mathcal{N}$. If the solution of the HLCP is unique, then it is a maximal complementarity solution. Otherwise it can be shown that the relative interior of \mathcal{F}^* is composed of maximal complementarity solutions.

If the set \mathcal{J} is empty then a maximal complementarity solution is called a strictly complementary solution. Let us denote by \mathcal{F}^c the set of all such solutions, i.e.,

$$\mathcal{F}^c = \{z^* = [x^*, s^*] \in \mathcal{F}^* : x^* + s^* > 0\}.$$

We say that the HLCP is nondegenerate if it has a strictly complementary solution. If the set \mathcal{J} is nonempty, then we say that the HLCP is degenerate.

The set

$$\mathcal{F}^0 = \mathcal{F} \cap \mathbb{R}_{++}^{2n},$$

is called the set of strictly feasible points, or the set of interior points.

2.2. The central path and its analyticity. It is known (see [13]) that if \mathcal{F}^0 is nonempty, then, for any vector $p \in \mathbb{R}_{++}^n$ and any parameter $\tau > 0$ the nonlinear system,

$$(2.2) \quad \begin{aligned} xs &= \tau p \\ Qx + Rs &= b \end{aligned}$$

has a unique positive solution $z(\tau, p) = [x(\tau, p), s(\tau, p)]$. For a fixed p , the curve

$$\mathcal{C}(p) = \{z(\tau, p) : \tau > 0\}$$

is called the weighted central path of the HLCP with weight vector p . It turns out that if the HLCP is nondegenerate, then $z(\tau, p)$ is an analytic function in $\tau > 0$ and $p > 0$, that has an analytic continuation at $\tau = 0$. In case the HLCP is degenerate then $z(\tau, p)$ is analytic in $\rho = \sqrt{\tau}$ which also has an analytic continuation at $\rho = 0$. Moreover all the derivatives of z are bounded on any compact set contained in the

domain of analyticity of z . More precisely the following results follow from the more general results of [38, 39].

THEOREM 2.3. *If HLCP is monotone, and \mathcal{F}^0 is nonempty, then (2.2) has a unique positive solution $z(\tau, p) = [x(\tau, p), s(\tau, p)]$ for any $(\tau, p) \in \mathbb{R}_{++}^{n+1}$ and the following properties hold:*

- A. *If the HLCP is nondegenerate then $z(\tau, p)$ can be analytically extended to an open neighborhood of $\mathbb{R}_+ \times \mathbb{R}_{++}^n$ and for any compact set $\mathcal{K} \subset \mathbb{R}_+ \times \mathbb{R}_{++}^n$ and any integer $i \in \mathbb{N}$ there are constants $c(\mathcal{K}, i)$ such that*

$$\left\| \frac{\partial^i z(\tau, p)}{\partial \tau^i} \right\|_2 \leq c(\mathcal{K}, i), \quad \forall (\tau, p) \in \mathcal{K}, \quad i = 0, 1, 2, \dots \quad ;$$

- B. *If the HLCP is degenerate then $\bar{z}(\rho, p) := z(\rho^2, p)$ can be analytically extended to an open neighborhood of $\mathbb{R}_+ \times \mathbb{R}_{++}^n$ and for any compact set $\mathcal{K} \subset \mathbb{R}_+ \times \mathbb{R}_{++}^n$ and any integer $i \in \mathbb{N}$ there are constants $\bar{c}(\mathcal{K}, i)$ such that*

$$\left\| \frac{\partial^i \bar{z}(\rho, p)}{\partial \rho^i} \right\|_2 \leq \bar{c}(\mathcal{K}, i), \quad \forall (\rho, p) \in \mathcal{K}, \quad i = 0, 1, 2, \dots \quad .$$

If $p = e$, the vector of all ones, then $\mathcal{C} = \mathcal{C}(e)$ is simply called the central path of the HLCP. The distance of a point $z \in \mathcal{F}$ to the central path \mathcal{C} can be quantified by different proximity measures. The following proximity measures have been extensively used in the interior point literature:

$$\delta_2(z) := \left\| \frac{xs}{\mu(z)} - e \right\|_2, \quad \delta_\infty(z) := \left\| \frac{xs}{\mu(z)} - e \right\|_\infty, \quad \delta_\infty^-(z) := \left\| \left[\frac{xs}{\mu(z)} - e \right]^- \right\|_\infty,$$

where $[v]^-$ denotes the negative part of the vector v , i.e. $[v]^- = -\max\{-v, 0\}$.

By using the above proximity measures we can define the following neighborhoods of the central path

$$\mathcal{N}_2(\eta) = \{z \in \mathcal{F}^0 : \delta_2(z) \leq \eta\},$$

$$\mathcal{N}_\infty(\eta) = \{z \in \mathcal{F}^0 : \delta_\infty(z) \leq \eta\},$$

$$\mathcal{N}_\infty^-(\eta) = \{z \in \mathcal{F}^0 : \delta_\infty^-(z) \leq \eta\},$$

where $0 < \eta < 1$ is a given parameter. We have

$$(2.3) \quad \mathcal{C} \subset \mathcal{N}_2(\eta) \subset \mathcal{N}_\infty(\eta) \subset \mathcal{N}_\infty^-(\eta), \quad \lim_{\eta \downarrow 0} \mathcal{N}_\infty^-(\eta) = \mathcal{C}, \quad \lim_{\eta \uparrow 1} \mathcal{N}_\infty^-(\eta) = \mathcal{F}.$$

3. The first order affine scaling method.

3.1. The algorithm. In the remainder of this paper we will work with $\mathcal{N}_\infty^-(\eta)$. We note that this neighborhood can be written under the form:

$$\mathcal{N}_\infty^-(\eta) = \mathcal{D}(1 - \eta), \quad \text{where } \mathcal{D}(\beta) = \{z \in \mathcal{F}^0 : xs \geq \beta \mu(z)\}.$$

At each step of our algorithm we are given a point $z = [x, s] \in \mathcal{D}(\beta)$ and we compute the affine scaling direction $w = [u, v]$ by solving the following linear system

$$(3.1) \quad \begin{cases} su + xv &= -xs \\ Qu + Rv &= 0 \end{cases} .$$

This search direction coincides, up to a scalar factor, with the first derivative to central path $z(\tau, p)$ that passes through $p = xs/\mu$ at $\tau = \mu$. Indeed, by differentiating (2.2) we obtain

$$(3.2) \quad \begin{cases} s(\tau, p) \frac{\partial}{\partial \tau} x(\tau, p) + x(\tau, p) \frac{\partial}{\partial \tau} s(\tau, p) &= \frac{xs}{\mu} \\ Q \frac{\partial}{\partial \tau} x(\tau, p) + R \frac{\partial}{\partial \tau} s(\tau, p) &= 0 \end{cases} .$$

Given that $x(\mu, p) = x$ and $s(\mu, p) = s$, we obtain

$$(3.3) \quad u = -\mu \frac{\partial}{\partial \tau} x(\tau, p), \quad v = -\mu \frac{\partial}{\partial \tau} s(\tau, p).$$

It turns out that the duality gap has the fastest local decrease along this direction. Indeed if we denote by

$$(3.4) \quad z(\theta) = z + \theta w$$

a point along this direction we have

$$(3.5) \quad \begin{aligned} x(\theta)s(\theta) &= (x + \theta u)(s + \theta v) = (1 - \theta)xs + \theta^2 uv \\ \mu(\theta) &= \frac{(x + \theta u)^T (s + \theta v)}{n} = (1 - \theta)\mu + \theta^2 \frac{u^T v}{n} . \end{aligned}$$

In the skew-symmetric case we have $u^T v = 0$, while in the monotone case, according to Lemma 3.1, we have $0 \leq u^T v \leq .25n\mu$, so that we can write

$$(3.6) \quad \mu(\theta) \leq (1 - \theta + .25\theta^2)\mu \leq (1 - .5\theta)^2 \mu, \quad \mu'(\theta) \leq (.5\theta - 1)\mu .$$

Given that $\mu(\theta)$ is decreasing on the interval $[0, 2]$ we compute the steplength $\bar{\theta}$ as

$$(3.7) \quad \bar{\theta} = \max \left\{ \hat{\theta} \in [0, 1] : z(\theta) \in \mathcal{D}(\beta_+), \forall \theta \in [0, \hat{\theta}] \right\} ,$$

where

$$(3.8) \quad \beta_+ = \beta - \alpha .$$

We obtain thus a new point satisfying

$$(3.9) \quad z^+ = z(\bar{\theta}) \in \mathcal{D}(\beta_+)$$

and the process can be repeated. The initial value for β is $\bar{\beta}$ and the α 's are chosen as the terms of the series

$$(3.10) \quad \alpha_k = \frac{\nu(\bar{\beta} - \beta)}{(\exp(1) + k + 1) \log^{1+\nu}(\exp(1) + k + 1)}, \quad k = 0, 1, \dots ,$$

where $0 < \nu \leq 1$ is a given parameter. Since

$$\sum_{k=0}^{\infty} \alpha_k < \nu(\bar{\beta} - \beta) \int_e^{\infty} \frac{dt}{t \log^{1+\nu} t} = \bar{\beta} - \beta$$

it follows that all the points produced by the following algorithm will stay in $\mathcal{D}(\underline{\beta})$. Summing up we can formulate the following iterative procedure:

Algorithm 1

Given real parameters $0 < \underline{\beta} < \bar{\beta} < 1$, $0 < \nu \leq 1$, and a vector $z^0 \in \mathcal{D}(\bar{\beta})$:

Set $k \leftarrow 0$, $\beta_0 \leftarrow \bar{\beta}$ and consider the sequence (3.10);

repeat

Set $z \leftarrow z^k$, $\alpha \leftarrow \alpha_k$, $\beta \leftarrow \beta_k$, $\beta_+ \leftarrow \beta - \alpha$;

Compute direction $w = \lceil u, v \rceil$ by solving (3.1);

Compute steplength $\bar{\theta}$ from (3.7);

Compute z^+ from (3.9);

Set $\theta_k \leftarrow \bar{\theta}$, $z^{k+1} \leftarrow z^+$, $\mu_{k+1} \leftarrow \mu(z^+)$, $\beta_{k+1} \leftarrow \beta_+$;

Set $k \leftarrow k + 1$.

continue

3.2. Technical Results. In order to analyze Algorithm 1 we need to establish some properties of the solution of a linear system of the form

$$(3.11) \quad \begin{cases} su + xv = a \\ Qu + Rv = 0 \end{cases}.$$

By using the notation

$$(3.12) \quad D = X^{-1/2} S^{1/2},$$

this system can be written as

$$(3.13) \quad \begin{cases} Du + D^{-1}v = (xs)^{-1/2}a \\ QD^{-1}(Du) + RD(D^{-1}v) = 0 \end{cases}.$$

Since the pair (QD^{-1}, RD) is monotone one can easily prove the following results (see, for example, [29, Lemma 3.1]).

LEMMA 3.1. *If HLCP is monotone, then for any $z = \lceil x, s \rceil \in \mathbb{R}_{++}^{2n}$ and any $a \in \mathbb{R}^n$ the linear system (3.11) has a unique solution $w = \lceil u, v \rceil$ and the following properties are satisfied:*

- i) $\|Du\|_2^2 + \|D^{-1}v\|_2^2 \leq \|(xs)^{-1/2}a\|_2^2$;
- ii) $\left\| uv - \frac{u^T v}{n} e \right\|_2 \leq \|uv\|_2 \leq \frac{1}{2\sqrt{2}} \|(xs)^{-1/2}a\|_2^2$;
- iii) $\|uv\|_\infty \leq \frac{1}{4} \|(xs)^{-1/2}a\|_2^2$;
- iv) $u^T v \leq \frac{1}{4} \|(xs)^{-1/2}a\|_2^2$.

The proof of global convergence of Algorithm 1 requires explicit lower bounds for the series $\sum_{k=0}^{K-1} \alpha_k^{1/(m+1)}$ for K sufficiently large. Given that

$$(3.14) \quad \sum_{k=0}^{K-1} \alpha_k^{\frac{1}{m+1}} > (\nu(\bar{\beta} - \underline{\beta}))^{\frac{1}{m+1}} \int_{\exp(1)+1}^{K+\exp(1)+1} \frac{dt}{t^{\frac{1}{m+1}} \log^{\frac{(1+\nu)m}{m+1}} t},$$

it remains to find an explicit lower bound depending on K of the integral appearing on the right-hand-side of the above inequality. To this effect we will use the following results.

LEMMA 3.2. *Let f and g be two continuous, nonnegative, and nonincreasing real functions defined on the interval $[\gamma, \delta]$. If f is convex, then*

$$\int_{\gamma}^{\delta} f(t)g(t)dt \geq f\left(\frac{\gamma + \delta}{2}\right) \int_{\gamma}^{\delta} g(t)dt.$$

Proof. Let us denote $\zeta = (\gamma + \delta)/2$. From the convexity of f it follows that f has lateral derivatives $f'(\zeta - 0) \leq f'(\zeta + 0) \leq 0$ at ζ and that

$$f(t) \geq f(\zeta) + m(t - \zeta), \forall m \in [f'(\zeta - 0), f'(\zeta + 0)], \forall t \in [\gamma, \delta].$$

Since g is nonnegative and nonincreasing we deduce that

$$f(t)g(t) \geq f(\zeta)g(t) + mg(\zeta)(t - \zeta), \forall m \in [f'(\zeta - 0), f'(\zeta + 0)], \forall t \in [\gamma, \delta],$$

and the statement of our lemma follows by fixing m and taking the integral with respect to t of both sides of the above inequality. \blacksquare

COROLLARY 3.3. *If $0 < \omega < 1$, $\eta > 0$, and $\varsigma \geq 2\gamma > 2$, then*

$$\int_{\gamma}^{\gamma+\varsigma} \frac{dt}{t^{\omega} \log^{\eta} t} \geq \frac{\varsigma^{1-\omega}}{\log^{\eta} \varsigma}.$$

Proof. Given that the function $t \rightarrow \log^{-\eta}(t)$ is convex for $t > 1$, we can apply Lemma 3.2 to obtain

$$\int_{\gamma}^{\gamma+\varsigma} \frac{dt}{t^{\omega} \log^{\eta} t} \geq \frac{(\gamma + \varsigma)^{1-\omega} - \gamma^{1-\omega}}{(1-\omega) \log^{\eta}(\gamma + .5\varsigma)} \geq \frac{(\gamma + \varsigma)^{1-\omega} - \gamma^{1-\omega}}{(1-\omega) \log^{\eta} \varsigma}.$$

Therefore it remains to prove that

$$(\gamma + \varsigma)^{1-\omega} - \gamma^{1-\omega} \geq (1-\omega)\varsigma^{1-\omega}.$$

Dividing both sides of the above inequality by $\varsigma^{1-\omega}$ and denoting $\sigma = \gamma/\varsigma$, $\tau = 1 - \omega$, this reduces to proving that

$$\phi(\tau, \sigma) := (1 + \sigma)^{\tau} - \sigma^{\tau} - \tau \geq 0, \forall \tau \in [0, 1], \forall \sigma \in [0, .5].$$

Since

$$\frac{\partial \phi}{\partial \sigma} = \tau ((1 + \sigma)^{\tau-1} - \sigma^{\tau-1}) \leq 0,$$

we deduce that

$$\phi(\tau, \sigma) \geq \psi(\tau) := \phi(\tau, .5), \forall \tau \in [0, 1], \forall \sigma \in [0, .5].$$

We have

$$\begin{aligned} \psi(\tau) &= 1.5^{\tau} - .5^{\tau} - \tau, \psi(0) = 0, \psi(1) = 0 \\ \psi'(\tau) &= 1.5^{\tau} \log 1.5 - .5^{\tau} \log .5 - 1, \psi'(0) > 0, \psi'(1) < 0 \\ \psi''(\tau) &= 1.5^{\tau} \log^2 1.5 - .5^{\tau} \log^2 .5, \psi''(0) < 0, \psi''(1) > 0 \\ \psi'''(\tau) &= 1.5^{\tau} \log^3 1.5 - .5^{\tau} \log^3 .5 > 0, \forall \tau \in [0, 1]. \end{aligned}$$

The positivity of ψ''' implies that ψ'' has a unique zero $\tau_2 \in [0, 1]$, which in turn shows that ψ' is decreasing on $[0, \tau_2]$ and increasing on $[\tau_2, 1]$. Hence ψ' has a unique zero $\tau_1 \in [0, \tau_2]$. This shows that ψ increases on $[0, \tau_1]$ and decreases on $[\tau_1, 1]$. Therefore ψ is positive on $(0, 1)$. \blacksquare

Using this corollary together with the following lemma we will be able to obtain explicit upper bounds on the number of steps required by Algorithm 1 to reduce the duality gap below any desired tolerance.

LEMMA 3.4. *The function $h : [\exp(1), \infty) \rightarrow [\exp(1), \infty)$, given by $h(t) = t/\log t$ is increasing and bijective and its inverse $h^{-1} : [\exp(1), \infty) \rightarrow [\exp(1), \infty)$ satisfies the inequalities*

$$t \log t < h^{-1}(t) < t \left(1 + \frac{\log \log t}{\log t - 1}\right) \log t < 2t \log t, \quad \forall t \in (e, \infty).$$

Proof. The first part of the lemma follows immediately from the fact that

$$h'(t) = \frac{\log t - 1}{\log^2 t} > 0, \quad \forall t \in (e, \infty).$$

In order to prove the second part we denote by l and u the lower and upper bounds above and show that $h(l) < t < h(u)$. We have obviously

$$h(l) = \frac{t \log t}{\log(t \log t)} = \frac{t \log t}{\log t + \log \log t} < t,$$

which proves the first inequality. Finally, by using the inequality $\log(1 + \alpha) < \alpha$ we obtain

$$h(u) = \frac{t \left(1 + \frac{\log \log t}{\log t - 1}\right) \log t}{\log(t \log t) + \log \left(1 + \frac{\log \log t}{\log t - 1}\right)} > \frac{t \left(1 + \frac{\log \log t}{\log t - 1}\right) \log t}{\log t + \log \log t + \frac{\log \log t}{\log t - 1}} = t.$$

COROLLARY 3.5. *If $\sigma > \exp(1)$, $\nu > 0$, $m > 0$, then*

$$\tau \geq \left(2 \left(1 + \frac{1 + \nu}{m + 1} \log \frac{1 + \nu}{m}\right) \frac{m + 1}{m} \log \sigma\right)^{\frac{1 + \nu}{m}} \sigma^{\frac{m + 1}{m}} \text{ implies } \frac{\tau^{\frac{m}{m + 1}}}{\log^{\frac{1 + \nu}{m + 1}} \tau} \geq \sigma.$$

Proof. By noticing that

$$\frac{\tau^{\frac{m}{m + 1}}}{\log^{\frac{1 + \nu}{m + 1}} \tau} = \left(\frac{\tau^{\frac{m}{1 + \nu}}}{\frac{1 + \nu}{m} \log \tau^{\frac{m}{1 + \nu}}}\right)^{\frac{1 + \nu}{m + 1}},$$

we can write the last inequality in the statement of Corollary 3.5 under the equivalent form

$$\frac{\tau^{\frac{m}{1 + \nu}}}{\log \tau^{\frac{m}{1 + \nu}}} \geq \frac{(1 + \nu) \sigma^{\frac{m + 1}{1 + \nu}}}{m}.$$

According to Lemma 3.4 this inequality is satisfied if

$$\tau \geq \left(\frac{2(1 + \nu) \sigma^{\frac{m + 1}{1 + \nu}}}{m} \log \frac{(1 + \nu) \sigma^{\frac{m + 1}{1 + \nu}}}{m}\right)^{\frac{1 + \nu}{m}} =: rhs.$$

The right hand side of the above inequality can be majorized as follows:

$$\begin{aligned}
rhs &= \left(\frac{2(1+\nu)}{m} \right)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} \left(\log \frac{(1+\nu)\sigma^{\frac{m+1}{1+\nu}}}{m} \right)^{\frac{1+\nu}{m}} \\
&= \left(\frac{2(1+\nu)}{m} \right)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} \left(\left(1 + \frac{\log \frac{1+\nu}{m}}{\log \sigma^{\frac{m+1}{1+\nu}}} \right) \log \sigma^{\frac{m+1}{1+\nu}} \right)^{\frac{1+\nu}{m}} \\
&\leq \left(\frac{2(1+\nu)}{m} \right)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} \left(\left(1 + \frac{1+\nu}{m+1} \log \frac{1+\nu}{m} \right) \log \sigma^{\frac{m+1}{1+\nu}} \right)^{\frac{1+\nu}{m}} \\
&= \left(\frac{2(1+\nu)}{m} \right)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} \left(\left(1 + \frac{1+\nu}{m+1} \log \frac{1+\nu}{m} \right) \frac{m+1}{1+\nu} \log \sigma \right)^{\frac{1+\nu}{m}} \\
&= \left(2 \left(1 + \frac{1+\nu}{m+1} \log \frac{1+\nu}{m} \right) \frac{m+1}{m} \log \sigma \right)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}},
\end{aligned}$$

which proves the corollary. \blacksquare

LEMMA 3.6. *The function $g : [\exp(\exp(1)), \infty) \rightarrow [\exp(\exp(1)), \infty)$, given by $g(t) = t / \log \log t$ is increasing and bijective and its inverse satisfies*

$$t \log \log t < g^{-1}(t) < t \left(1 + \frac{\log \log \log t}{(\log \log t) \log t - 1} \right) \log \log t < 2.072 t \log \log t.$$

Proof. The first part of the lemma follows immediately from the fact that

$$g'(t) = \frac{-1 + (\log \log t) \log t}{(\log \log t)^2 \log t} > 0, \quad \forall t \in (5.8313, \infty).$$

In order to prove the second part we denote by l and u the lower and upper bounds above and show that $g(l) < t < g(u)$. We have obviously

$$g(l) = \frac{t \log \log t}{\log \log (t \log \log t)} = \frac{t \log \log t}{\log \log t + \log \log \log t} < t,$$

which proves the first inequality. By using the inequality $\log(1+\alpha) < \alpha$ we obtain

$$\begin{aligned}
g(u) &= \frac{t \left(1 + \frac{\log \log \log t}{(\log \log t) \log t - 1} \right) \log \log t}{\log \left(\log t + \log \log \log t + \log \left(1 + \frac{\log \log \log t}{(\log \log t) \log t - 1} \right) \right)} \\
&> \frac{t \left(1 + \frac{\log \log \log t}{(\log \log t) \log t - 1} \right) \log \log t}{\log \left(\log t + (\log \log \log t) \left(1 + \frac{1}{(\log \log t) \log t - 1} \right) \right)} \\
&> \frac{t \left(1 + \frac{\log \log \log t}{(\log \log t) \log t - 1} \right) \log \log t}{\log \log t + \frac{\log \log \log t}{\log t} \frac{(\log \log t) \log t}{(\log \log t) \log t - 1}} = t.
\end{aligned}$$

Finally, it is easily checked that

$$\max_{t \geq \exp(\exp(1))} \frac{\log \log \log t}{(\log \log t) \log t - 1} < 0.072. \quad \blacksquare$$

3.3. Polynomial complexity. In the next theorem we show that Algorithm 1 is well defined and we give bounds on the decrease of the duality gap at each iteration.

THEOREM 3.7. *If HLCP is monotone, then Algorithm 1 is well defined and produces a sequence of points $(z^k)_{k=0}^\infty$, with $z^k \in \mathcal{D}(\beta_k) \subset \mathcal{D}(\underline{\beta})$. If $n \geq 5$ then the following relations hold:*

$$\theta_k \geq \sqrt{2\alpha_k/n}, \quad \mu_{k+1} \leq \left(1 - \sqrt{\alpha_k/n}\right) \mu_k, \quad k = 0, 1, \dots,$$

$$\mu_k \leq \mu_0 \exp\left(-\sqrt{\frac{\nu(\bar{\beta} - \underline{\beta})k}{n \log^{1+\nu} k}}\right), \quad k = 8, 9, \dots,$$

Proof. The relation $z^k \in \mathcal{D}(\beta_k)$ is ensured by the line search (3.7). Let now $z \in \mathcal{D}(\beta)$ be given, and denote

$$p = \frac{xs}{\mu}, \quad q = \frac{uv}{\mu}.$$

From our hypothesis and from Lemma 3.1 it follows that

$$p \geq \beta e, \quad \|q\|_\infty \leq \frac{n}{4}, \quad \|q - (e^T q/n)e\|_2 \leq \|q\|_2 \leq \frac{n}{2\sqrt{2}}, \quad \frac{u^T v}{\mu} = e^T q \leq \frac{n}{4}.$$

Using (3.5) and (3.6) we deduce that

$$\begin{aligned} \frac{x(\theta)s(\theta)}{\mu(\theta)} &= \frac{(1-\theta)p + \theta^2 q}{1-\theta + (e^T q/n)\theta^2} \\ &= \frac{(1-\theta)p + \beta(e^T q/n)\theta^2 e + \beta\theta^2(q - (e^T q/n)e) + (1-\beta)\theta^2 q}{1-\theta + (e^T q/n)\theta^2} \\ &\geq \beta e - \frac{\|\beta(q - (e^T q/n)e) + (1-\beta)q\|_2 \theta^2}{1-\theta} e \\ (3.15) \quad &\geq \beta e - \frac{\|q\|_2 \theta^2}{1-\theta} e. \end{aligned}$$

If $n \geq 5$, then $\sqrt{2\alpha/n} \leq \sqrt{2\alpha_0/5} < 1 - 1/\sqrt{2}$ for any values of the parameters defining our algorithm, and in this case we can easily verify that

$$\frac{\|q\|_2 \theta^2}{1-\theta} \leq \frac{n\theta^2}{(1-\theta)2\sqrt{2}} \leq \alpha, \quad \forall \theta \in [0, \sqrt{2\alpha/n}].$$

It follows that $\bar{\theta} \geq \sqrt{2\alpha/n}$, and therefore

$$\begin{aligned} \mu_+ &= \mu(\bar{\theta}) \leq \mu\left(\sqrt{2\alpha/n}\right) \leq \left(1 - \sqrt{2\alpha/n} + .5\alpha/n\right) \mu \\ &\leq \left(1 - \left(1 - .25\sqrt{2\alpha_0/5}\right) \sqrt{2\alpha/n}\right) \mu < \left(1 - \sqrt{\alpha/n}\right) \mu. \end{aligned}$$

Hence we have proved that

$$\mu_{j+1} \leq \left(1 - \sqrt{\alpha_j/n}\right) \mu_j, \quad j = 0, 1, \dots$$

By repeatedly applying the above inequality we deduce that

$$\mu_k \leq \mu_0 \prod_{j=0}^{k-1} \left(1 - \sqrt{\alpha_j/n}\right).$$

Using the inequality $\log(1-t) < -t$, (3.14), $k \geq 8 > 2(\exp(1)+1)$, and Corollary 3.3, we obtain

$$\log\left(\frac{\mu_k}{\mu_0}\right) \leq \sum_{j=0}^{k-1} \log\left(1 - \sqrt{\alpha_j/n}\right) \leq -\sum_{j=0}^{k-1} \sqrt{\alpha_j/n} \leq -\sqrt{\frac{\nu(\bar{\beta} - \underline{\beta})k}{n \log^{1+\nu} k}},$$

which completes the proof of our theorem. \blacksquare

In the following corollary we give an explicit upper bound for the number of iterations required by Algorithm 1 to obtain a solution of the HLCP with prescribed accuracy. More precisely given any $\epsilon > 0$ we have to find an upper bound for the number

$$(3.16) \quad K_\epsilon := \min \{K : x^{kT} s^k \leq \epsilon, \forall k \geq K\}.$$

The upper bound will depend on n and

$$(3.17) \quad L_\epsilon := \log\left(\frac{x^{0T} s^0}{\epsilon}\right).$$

COROLLARY 3.8. *If $nL_\epsilon^2 > \exp(2)/(\nu(\bar{\beta} - \underline{\beta}))$, then*

$$K_\epsilon \leq \left\lceil \frac{7^{1+\nu}}{\nu(\bar{\beta} - \underline{\beta})} nL_\epsilon^2 \log^{1+\nu}(nL_\epsilon^2) \right\rceil.$$

Proof. From Theorem 3.2 we deduce that $x^{kT} s^k \leq \epsilon$ for any K with the property

$$\frac{K^{1/2}}{\log^{(1+\nu)/2} K} \geq \sigma := (\nu(\bar{\beta} - \underline{\beta}))^{-1/2} \sqrt{n} L_\epsilon.$$

Since $\sigma > \exp(1)$, we can use Corollary 3.5 to show that

$$K_\epsilon \leq \bar{K}_\epsilon := (4(1 + \log 2) \log \sigma)^{1+\nu} \sigma^2,$$

and the statement of our corollary follows by noticing that

$$\begin{aligned} 4(1 + \log 2) \log \sigma &= 2(1 + \log 2) (\log(nL_\epsilon^2) - \log(\nu(\bar{\beta} - \underline{\beta}))) \\ &\leq 4(1 + \log 2) (\log(nL_\epsilon^2) - 1) < 4(1 + \log 2) \log(nL_\epsilon^2) < 7 \log(nL_\epsilon^2). \end{aligned}$$

If ν is a positive constant independent of n and L_ϵ , then Algorithm 1 is independent of the dimension of the problem and the stopping criterium $x^{kT} s^k \leq \epsilon$. However, by letting ν depend on n and L_ϵ , we can slightly improve the computational complexity. \blacksquare

COROLLARY 3.9. *If*

$$nL_\epsilon^2 \geq \max \left\{ \exp(\exp(1)), \frac{2.072 \exp(2)}{\bar{\beta} - \underline{\beta}} \log \log \left(\frac{\exp(2)}{\bar{\beta} - \underline{\beta}} \right) \right\},$$

and take $\nu = (\log \log(nL_\epsilon^2))^{-1}$ in Algorithm 1. Then

$$K_\epsilon \leq \left\lceil \frac{134}{\bar{\beta} - \underline{\beta}} nL_\epsilon^2 \log(nL_\epsilon^2) \log \log(nL_\epsilon^2) \right\rceil.$$

Proof. From Lemma 3.6 we deduce that $nL_\epsilon^2 \geq \exp(2)/(\nu(\bar{\beta} - \underline{\beta}))$ and Corollary 3.8 can be applied. The desired result follows immediately by noticing that $\log^\nu(nL_\epsilon^2) = \exp(1)$ and $7^{1+\nu} \exp(1) \leq 49 \exp(1) < 134$. ■

We note that although for $\nu = (\log \log(nL_\epsilon^2))^{-1}$ Algorithm 1 depends on ϵ , it will produce an (infinite) sequence $z^k \in \mathcal{D}(\underline{\beta})$ with $\lim_{k \rightarrow \infty} \mu_k = 0$. As we mentioned in the introduction, this is not the case with the primal-dual affine-scaling algorithm of Monteiro, Adler and Resende [23]. Given a starting point $z^0 \in \mathcal{C}$ on the central path and a tolerance $\epsilon > 0$, their algorithm produces a finite sequence of points (z^k) by taking at each iteration a fixed stepsize $\theta_\epsilon = 1/(n \lceil \log L_\epsilon \rceil)$, i.e.,

$$z^{k+1} = z^k + \theta_\epsilon w^k, \quad k = 0, 1, \dots, \bar{K}_\epsilon := \frac{\lceil n \log L_\epsilon \rceil}{\theta_\epsilon}.$$

It is shown that $z^k \in \mathcal{N}_\infty(\bar{\alpha}_k)$, with $\bar{\alpha}_k < n\bar{K}_\epsilon \bar{\theta}_\epsilon^2 = 1, k = 0, 1, \dots, \bar{K}_\epsilon$, which implies the feasibility of z^k for $k \leq \bar{K}_\epsilon$. However the positivity of z^k is no longer guaranteed for $k > \bar{K}_\epsilon$, so that the algorithm is not defined for $k > \bar{K}_\epsilon$. Since the algorithm of [23] uses a fixed (and small) steplength it cannot have superlinearly convergence. Moreover since this algorithm produces only a finite sequence, we cannot even talk about its order of convergence. By contrast, in the next subsection we will show that the Q-order of convergence of Algorithm 1 is two.

3.4. Superlinear convergence. The main ingredient in the proof of superlinear convergence is provided by the following lemma which is easily obtained by applying Theorem 2.3 with

$$(3.18) \quad \mathcal{K} = \mathcal{K}(\beta, \bar{\mu}) = \{(\tau, p) : 0 \leq \tau \leq \bar{\mu}, e^T p = n, p \geq \underline{\beta} e\}$$

and equation (3.3).

LEMMA 3.10. *If HLCP 2.1 is monotone and has a strictly complementary solution then for any $\beta \in (0, 1)$ and any $\bar{\mu} > 0$ there is constant $c = c(\beta, \bar{\mu})$ such that the affine-scaling direction given by (3.1) satisfies*

$$\|u\|_2 \leq c\mu, \quad \|v\|_2 \leq c\mu, \quad \forall z \in \left\{ [x, s] \in \mathcal{D}(\beta) : \mu = \frac{x^T s}{n} \leq \bar{\mu} \right\}.$$

In the next theorem we prove that under the strict complementarity assumption the sequence (μ_k) converges to zero with Q-order 2. We recall that the Q-order of a sequence of positive numbers (η_k) that converges to zero is defined as

$$\mathcal{Q}(\eta_k) = \sup \{ \omega \in (1, \infty) : \exists \Gamma \in \mathbb{R}_{++}, \forall k \in \mathbb{N}, \eta_{k+1} \leq \Gamma \eta_k^\omega \}$$

It is known (see [27]) that for $\bar{\omega} > 1$, we have $\bar{\omega} = \mathcal{Q}(\eta_k)$ if and only if

$$\bar{\omega} = \liminf \frac{\log \eta_{k+1}}{\log \eta_k}.$$

THEOREM 3.11. *If HLCP is monotone and has a strictly complementary solution, then the sequence (z^k) produced by Algorithm 1 converges Q-superlinearly to a*

strictly complementary solution $z^* \in \mathcal{F}^c$, and the sequence of the corresponding complementarity gaps (μ_k) converges Q -superlinearly to zero. Moreover, the Q -orders of convergence of these sequences satisfy $\mathcal{Q}(z^k) = \mathcal{Q}(\mu_k) \geq 2$.

Proof. From Theorem 3.7 it follows that $p^k := (x^k s^k) / \mu_k \in \mathcal{K}(\underline{\beta}, \mu_0)$, $\forall k \geq 0$, so that by using Lemma 3.10 we deduce that there is a constant $c = c(\underline{\beta}, \mu_0)$ such that

$$(3.19) \quad \|u\|_2 \leq c\mu, \quad \|v\|_2 \leq c\mu, \quad \text{at each iteration } k = 0, 1, \dots$$

For k sufficiently large we have $\mu = \mu_k < 1$, so that $\log \mu_k < 0$, and by using Theorem 3.7 we obtain

$$0 \leq \frac{\log \alpha_k}{\log \mu_k} \leq \frac{\log(\nu(\bar{\beta} - \underline{\beta})) - \log((k + \exp(1) + 1) \log^{1+\nu}(k + \exp(1) + 1))}{\log \mu_0 - \sqrt{\frac{\nu(\bar{\beta} - \underline{\beta})k}{n \log^{1+\nu} k}}}.$$

It follows that $\lim_{k \rightarrow \infty} \frac{\log \alpha_k}{\log \mu_k} = 0$, which implies $\lim_{k \rightarrow \infty} \frac{\mu_k}{\alpha_k} = 0$, so that for k sufficiently large we have $\mu = \mu_k < \alpha_k c^{-2} = \alpha c^{-2}$. By using (3.15) and the notation from the proof of Theorem 3.7,

$$0 < \theta < 1 \quad \text{and} \quad \frac{\|q\|_2 \theta^2}{1 - \theta} \leq \alpha \quad \text{implies} \quad z(\theta) \in \mathcal{D}(\beta_+).$$

Given that

$$\frac{\|q\|_2 \theta^2}{1 - \theta} \leq \frac{\|u\|_2 \|v\|_2 \theta^2}{(1 - \theta)\mu} \leq \frac{c^2 \mu \theta^2}{1 - \theta} \leq \alpha, \quad \forall \theta \in \left(0, 1 - \frac{c^2 \mu}{\alpha}\right),$$

it follows that $\bar{\theta} \geq 1 - \frac{c^2 \mu}{\alpha}$, and therefore

$$\mu_+ = \mu(\bar{\theta}) \leq \mu \left(1 - \frac{c^2 \mu}{\alpha}\right) \leq \frac{c^2 \mu^2}{\alpha} + \frac{u^T v}{n} \leq \left(\frac{1}{\alpha} + \frac{1}{n}\right) c^2 \mu^2 \leq \frac{2c^2}{\alpha} \mu^2.$$

By taking logarithms we obtain

$$\frac{\log \mu_+}{\log \mu} \geq 2 + \frac{\log(2c^2)}{\log \mu} - \frac{\log \alpha}{\log \mu}.$$

Since the right-hand-side of the above inequality tends to zero as $k \rightarrow \infty$, we deduce that $\liminf \frac{\log \mu_{k+1}}{\log \mu_k} \geq 2$. Hence $\mathcal{Q}(\mu_k) \geq 2$. On the other hand, we have $\|z^{k+1} - z^k\|_2 \leq \|w^k\|_2 \leq \sqrt{2}c\mu_k$, and by applying Theorem 2 of [28] we deduce the convergence of the sequence (z^k) to a strictly complementary solution $z^* \in \mathcal{F}^c$ and the fact that $\mathcal{Q}(z^k) = \mathcal{Q}(\mu_k)$. \blacksquare

4. Higher order affine scaling methods.

4.1. The higher order affine scaling directions. The higher order affine scaling directions to be considered in this section are related to the higher order derivatives to the central path. As we have seen in Section 2, the central path $z(\tau, p) = [x(\tau, p), s(\tau, p)]$ passing through a positive vector $p \in \mathbb{R}_{++}^n$ is analytic in τ if HLCP has a strictly complementary solution, and in $\rho = \sqrt{\tau}$ in the general case. By repeatedly differentiating the equations of the central path

$$(4.1) \quad \begin{aligned} x(\tau, p)s(\tau, p) &= \tau p \\ Qx(\tau, p) + Rs(\tau, p) &= b \end{aligned}$$

with respect to τ we obtain

$$\begin{cases} s(\tau, p) \frac{\partial}{\partial \tau} x(\tau, p) + x(\tau, p) \frac{\partial}{\partial \tau} s(\tau, p) & = p \\ Q \frac{\partial}{\partial \tau} x(\tau, p) + R \frac{\partial}{\partial \tau} s(\tau, p) & = 0 \end{cases},$$

$$\begin{cases} s(\tau, p) \frac{\partial^i}{\partial \tau^i} x(\tau, p) + x(\tau, p) \frac{\partial^i}{\partial \tau^i} s(\tau, p) & = - \sum_{j=1}^{i-1} \binom{i}{j} \frac{\partial^j}{\partial \tau^j} x(\tau, p) \frac{\partial^{i-j}}{\partial \tau^{i-j}} s(\tau, p) \\ Q \frac{\partial^i}{\partial \tau^i} x(\tau, p) + R \frac{\partial^i}{\partial \tau^i} s(\tau, p) & = 0 \end{cases},$$

$$i = 2, 3, \dots .$$

If we re-parameterize the central path equations in terms of $\rho = \sqrt{\tau}$, then the derivatives with respect to ρ of $\bar{z}(\rho, p) = [\bar{x}(\rho, p), \bar{s}(\rho, p)] := z(\rho^2, p)$ are given by

$$\begin{cases} \bar{s}(\rho, p) \frac{\partial}{\partial \rho} \bar{x}(\rho, p) + \bar{x}(\rho, p) \frac{\partial}{\partial \rho} \bar{s}(\rho, p) & = 2\rho p \\ Q \frac{\partial}{\partial \rho} \bar{x}(\rho, p) + R \frac{\partial}{\partial \rho} \bar{s}(\rho, p) & = 0 \end{cases},$$

$$\begin{cases} \bar{s}(\rho, p) \frac{\partial^2}{\partial \rho^2} \bar{x}(\rho, p) + \bar{x}(\rho, p) \frac{\partial^2}{\partial \rho^2} \bar{s}(\rho, p) & = 2p - 2 \frac{\partial}{\partial \rho} \bar{x}(\rho, p) \frac{\partial}{\partial \rho} \bar{s}(\rho, p) \\ Q \frac{\partial^2}{\partial \rho^2} \bar{x}(\rho, p) + R \frac{\partial^2}{\partial \rho^2} \bar{s}(\rho, p) & = 0 \end{cases},$$

$$\begin{cases} \bar{s}(\rho, p) \frac{\partial^i}{\partial \rho^i} \bar{x}(\rho, p) + \bar{x}(\rho, p) \frac{\partial^i}{\partial \rho^i} \bar{s}(\rho, p) & = - \sum_{j=1}^{i-1} \binom{i}{j} \frac{\partial^j}{\partial \rho^j} \bar{x}(\rho, p) \frac{\partial^{i-j}}{\partial \rho^{i-j}} \bar{s}(\rho, p) \\ Q \frac{\partial^i}{\partial \rho^i} \bar{x}(\rho, p) + R \frac{\partial^i}{\partial \rho^i} \bar{s}(\rho, p) & = 0 \end{cases},$$

$$i = 3, 4, \dots .$$

If HLCP is a reformulation of a linear programming problem then it is known that it has a strictly complementary solution. In general it is very difficult to establish whether HLCP has a strictly complementary solution. However, if this is the case the algorithm should take advantage of this information. In order to treat the two cases in a unified manner we define

$$(4.2) \quad \vartheta = \begin{cases} 0, & \text{if HLCP is known to be nondegenerate} \\ 1, & \text{otherwise} \end{cases}.$$

At each step of our higher order affine-scaling interior point method we have a point $z = [x, s] \in \mathcal{D}(\beta)$ and we consider the vectors

$$(4.3) w^i = [u^i, v^i] := \begin{cases} \left(\frac{(-1)^i \mu^i}{i!} \left[\frac{\partial^i}{\partial \tau^i} x(\tau, p), \frac{\partial^i}{\partial \tau^i} s(\tau, p) \right] \right) \Bigg|_{\substack{p = \frac{xs}{\mu} \\ \tau = \mu}} & , \quad \text{if } \vartheta = 0 \\ \left(\frac{(-1)^i \mu^{i/2}}{i!} \left[\frac{\partial^i}{\partial \rho^i} \bar{x}(\rho, p), \frac{\partial^i}{\partial \rho^i} \bar{s}(\rho, p) \right] \right) \Bigg|_{\substack{p = \frac{xs}{\mu} \\ \rho = \sqrt{\mu}}} & , \quad \text{if } \vartheta = 1 \end{cases}$$

4.2. The higher order algorithm. It is easily seen that the vectors (4.3) can be obtained by solving the following m systems of linear equations:

$$(4.4) \quad \begin{cases} su^1 + xv^1 &= -(1 + \vartheta)xs \\ Qu^1 + Rv^1 &= 0 \end{cases}, \\ \begin{cases} su^2 + xv^2 &= \vartheta xs - u^1v^1 \\ Qu^2 + Rv^2 &= 0 \end{cases}, \\ \begin{cases} su^i + xv^i &= -\sum_{j=1}^{i-1} u^jv^{i-j} \\ Qu^i + Rv^i &= 0 \end{cases}, \quad i = 3, \dots, m.$$

The m linear systems above have the same matrix, so that their numerical solution requires only one matrix factorization and m backsolves. This involves $O(n^3) + mO(n^2)$ arithmetic operations.

We note that for $\vartheta = 0$, $w^1 = [u^1, v^1]$ is just the affine scaling direction considered in Section 3.

Given the vectors $w^i = [u^i, v^i]$ defined by (4.4), we consider the polynomial

$$(4.5) \quad z(\theta) = z + \sum_{i=1}^m w^i \theta^i,$$

which represents the m 'th order Taylor expansion around $\theta = 0$ of the function $\theta \rightarrow z((1 - \theta)\mu, xs/\mu)$ in case $\vartheta = 0$, and of the function $\theta \rightarrow \bar{z}(\sqrt{(1 - \theta)\mu}, xs/\mu)$ in case $\vartheta = 1$.

We have $z(0) = z \in \mathcal{D}(\beta)$ and we define

$$(4.6) \quad \check{\theta} = \sup \left\{ \tilde{\theta} > 0 : z(\theta) \in \mathcal{D}(\beta_+), \forall \theta \in [0, \tilde{\theta}] \right\},$$

where β_+ is given by (3.8). From (4.4)-(4.5) we deduce that

$$(4.7) \quad \begin{aligned} x(\theta)s(\theta) &= (1 - \theta)^{1+\vartheta}xs + \sum_{i=m+1}^{2m} \theta^i h^i, \\ \mu(\theta) &= (1 - \theta)^{1+\vartheta}\mu + \sum_{i=m+1}^{2m} \theta^i (e^T h^i / n), \\ \text{where } h^i &= \sum_{j=i-m}^m u^j v^{i-j}. \end{aligned}$$

Therefore the computation of (4.6) involves the solution of a system of polynomial inequalities of order $2m$ in θ . Good lower bounds of the exact solution can be obtained by a line search procedure. In what follows we will give simple lower bounds that will be used in the proof of global convergence.

Given $\check{\theta}$, or a suitable convenient lower bound, we compute

$$(4.8) \quad \bar{\theta} = \operatorname{argmin} \{ \mu(\theta) : \theta \in [0, \check{\theta}] \},$$

and obtain a new point

$$(4.9) \quad z^+ = z(\bar{\theta}).$$

We have $z^+ \in \mathcal{D}(\beta_+)$ by construction, and the process can be repeated. Our higher order affine-scaling method is thus defined by the following iterative procedure:

Algorithm 2

Given real parameters $0 < \underline{\beta} < \bar{\beta} < 1$, $0 < \nu \leq 1$, integer $m \geq 2$, boolean variable $\vartheta \in \{0, 1\}$, and a vector $z^0 \in \mathcal{D}(\bar{\beta})$:

Set $k \leftarrow 0$, $\beta_0 \leftarrow \bar{\beta}$ and consider the sequence (3.10) ;

repeat

Set $z \leftarrow z^k$, $\alpha \leftarrow \alpha_k$, $\beta \leftarrow \beta_k$;

Compute w^1, \dots, w^m by solving (4.4);

Set $\beta_+ = \beta - \alpha$;

Compute steplength $\bar{\theta}$ from (4.6), (4.8);

Compute z^+ from (4.9);

Set $\theta_k \leftarrow \bar{\theta}$, $z^{k+1} \leftarrow z^+$, $\mu_{k+1} \leftarrow \mu(z^+)$, $\beta_{k+1} \leftarrow \beta_+$;

Set $k \leftarrow k + 1$.

continue

4.3. Global convergence. The computational complexity of Algorithm 2 is the same for $\vartheta = 0$ and $\vartheta = 1$. One could eventually obtain slightly better constants if $\vartheta = 0$ and/or if HLCP is skew symmetric, but in what follows we will obtain bounds in the monotone case that are independent of ϑ .

LEMMA 4.1. *If HLCP is monotone and if $z = [x, s] \in \mathcal{D}(\beta)$ then the vectors h^i defined by (4.4), (4.7) satisfy*

$$\|h^i\|_2 \leq \frac{2\beta\mu}{i} \left(4\sqrt{n/\beta}\right)^i, \quad |e^T h^i| \leq \frac{\beta\mu}{i} \left(4\sqrt{n/\beta}\right)^i, \quad i = m+1, \dots, 2m.$$

Proof. First let us prove that the quantities $\eta_i := \|Du^i + D^{-1}v^i\|_2$ satisfy

$$\sqrt{\|Du^i\|_2^2 + \|D^{-1}v^i\|_2^2} \leq \eta_i \leq 2\bar{\alpha}_i \sqrt{\beta\mu} \left(\frac{1+\vartheta}{2} \sqrt{n/\beta}\right)^i, \quad i = 1, 2, \dots, m,$$

where the sequence

$$\bar{\alpha}_i := \frac{1}{i} \binom{2i-2}{i-1} \leq \frac{1}{i} 4^i$$

satisfies the following recurrence scheme

$$\bar{\alpha}_1 = 1, \quad \bar{\alpha}_i = \sum_{j=1}^{i-1} \bar{\alpha}_j \bar{\alpha}_{i-j}.$$

The first part of the inequality follows immediately, since by using (4.4) and the monotony of the HLCP we deduce that $u^i T v^i \geq 0$. Hence

$$\|Du^i + D^{-1}v^i\|_2^2 = \|Du^i\|_2^2 + 2u^i T v^i + \|D^{-1}v^i\|_2^2 \geq \|Du^i\|_2^2 + \|D^{-1}v^i\|_2^2.$$

By multiplying the first equations of (4.4) with $(xs)^{-1/2}$ we obtain

$$\begin{aligned} Du^1 + D^{-1}v^1 &= -(1+\vartheta)(xs)^{1/2} \\ Du^2 + D^{-1}v^2 &= \vartheta(xs)^{1/2} - (xs)^{-1/2}u^1v^1 \\ Du^i + D^{-1}v^i &= -(xs)^{-1/2} \sum_{j=1}^{i-1} Du^j D^{-1}v^{i-j}, \quad i = 3, \dots, m. \end{aligned}$$

Because $z \in \mathcal{D}(\beta)$ we have $(xs)^{-1/2} \leq 1/\sqrt{\beta\mu}$, and, using Lemma 3.1, we deduce that

$$\begin{aligned}\eta_1 &= (1 + \vartheta)\sqrt{n\mu}, \\ \eta_2^2 &= \vartheta^2 n\mu - 2\vartheta u^1 T v^1 + \left\| (xs)^{-1/2} u^1 v^1 \right\|_2^2 \leq \vartheta^2 n\mu + \frac{\|u^1 v^1\|_2^2}{\beta\mu} \\ &\leq \vartheta^2 n\mu + \frac{\eta_1^4}{8\beta\mu} \leq \vartheta^2 n\mu + \frac{(1 + \vartheta)^4 n^2 \mu}{8\beta} < \frac{(1 + \vartheta)^4 n^2 \mu}{4\beta}, \\ \eta_i &\leq \frac{1}{\sqrt{\beta\mu}} \sum_{j=1}^{i-1} \|Du^j\|_2 \|D^{-1}v^{i-j}\|_2, \quad i = 3, \dots, m.\end{aligned}$$

Since

$$\begin{aligned}&\|Du^j\|_2 \|D^{-1}v^{i-j}\|_2 + \|Du^{i-j}\|_2 \|D^{-1}v^j\|_2 \\ &\leq \left(\|Du^j\|_2^2 + \|D^{-1}v^j\|_2^2 \right)^{1/2} \left(\|Du^{i-j}\|_2^2 + \|D^{-1}v^{i-j}\|_2^2 \right)^{1/2} \leq \eta_j \eta_{i-j},\end{aligned}$$

we obtain

$$\eta_i \leq \frac{1}{2\sqrt{\beta\mu}} \sum_{j=1}^{i-1} \eta_j \eta_{i-j}, \quad i = 3, \dots, m.$$

The required inequalities are then proved by mathematical induction (see [55, 54]).

The upper bound for $\|h^i\|_2$, $i = m+1, m+2, \dots, 2m$ can be obtained by writing

$$\begin{aligned}\|h^i\|_2 &\leq \sum_{j=i-m}^m \|Du^j\|_2 \|D^{-1}v^{i-j}\|_2 \leq \sum_{j=1}^{i-1} \|Du^j\|_2 \|D^{-1}v^{i-j}\|_2 \\ &= \frac{1}{2} \sum_{j=1}^{i-1} \left(\|Du^j\|_2 \|D^{-1}v^{i-j}\|_2 + \|Du^{i-j}\|_2 \|D^{-1}v^j\|_2 \right) \\ &\leq \frac{1}{2} \sum_{j=1}^{i-1} \sqrt{\|Du^j\|_2^2 + \|D^{-1}v^j\|_2^2} \sqrt{\|Du^{i-j}\|_2^2 + \|D^{-1}v^{i-j}\|_2^2} \\ &\leq \frac{1}{2} \sum_{j=1}^{i-1} \eta_j \eta_{i-j} \leq 2\beta\mu \left(\frac{1 + \vartheta}{2} \sqrt{n/\beta} \right)^i \sum_{j=1}^{i-1} \bar{\alpha}_j \bar{\alpha}_{i-j} \\ &= 2\beta\mu \left(\frac{1 + \vartheta}{2} \sqrt{n/\beta} \right)^i \bar{\alpha}_i \leq \frac{2\beta\mu}{i} \left(2(1 + \vartheta) \sqrt{n/\beta} \right)^i \leq \frac{2\beta\mu}{i} \left(4\sqrt{n/\beta} \right)^i.\end{aligned}$$

Finally by using Proposition 2.1 and Lemma 3.1 *iv)* we have

$$\begin{aligned}|e^T h^i| &= \left| \sum_{j=i-m}^m u^j T v^{i-j} \right| \leq \frac{1}{2} \sum_{j=1}^{i-1} |u^j T v^{i-j} + u^{i-j} T v^j| \\ &\leq \sum_{j=1}^{i-1} \sqrt{u^j T v^j} \sqrt{u^{i-j} T v^{i-j}} \leq \frac{1}{4} \sum_{j=1}^{i-1} \eta_j \eta_{i-j} \leq \beta\mu \left(\frac{1 + \vartheta}{2} \sqrt{n/\beta} \right)^i \sum_{j=1}^{i-1} \bar{\alpha}_j \bar{\alpha}_{i-j} \\ &= \beta\mu \left(\frac{1 + \vartheta}{2} \sqrt{n/\beta} \right)^i \bar{\alpha}_i \leq \frac{\beta\mu}{i} \left(2(1 + \vartheta) \sqrt{n/\beta} \right)^i \leq \frac{\beta\mu}{i} \left(4\sqrt{n/\beta} \right)^i.\end{aligned}$$

In the following lemma we give a lower bound for the maximum stepsize along the higher order direction. ■

LEMMA 4.2. *If HLCP is monotone, $z \in \mathcal{D}(\beta)$, and $n \geq 5$, then the maximum stepsize $\check{\theta}$ defined in (4.6) satisfies the following inequality*

$$\check{\theta} \geq \hat{\theta} := \frac{\beta^{\frac{m-1}{2(m+1)}}}{4\sqrt{n}} \left(\frac{\alpha}{2}\right)^{\frac{1}{m+1}},$$

and

$$\mu(\hat{\theta}) \leq (1 - \hat{\theta}/2) \mu.$$

Proof. We have $z(\theta) \in \mathcal{D}(\beta_+)$ if and only if $x(\theta)s(\theta) - (\beta - \alpha)\mu(\theta)e \geq 0$. Using the fact that $z = z(0) \in \mathcal{D}(\beta)$ we can write

$$\begin{aligned} & x(\theta)s(\theta) - (\beta - \alpha)\mu(\theta)e \\ &= (1 - \theta)^{1+\vartheta}xs + \sum_{i=m+1}^{2m} \theta^i h^i - (\beta - \alpha) \left((1 - \theta)^{1+\vartheta}\mu + \frac{1}{n} \sum_{i=m+1}^{2m} \theta^i e^T h^i \right) e \\ &\geq (1 - \theta)^{1+\vartheta}\beta\mu e + \sum_{i=m+1}^{2m} \theta^i h^i - (\beta - \alpha) \left((1 - \theta)^{1+\vartheta}\mu + \sum_{i=m+1}^{2m} \theta^i \frac{e^T h^i}{n} \right) e \\ &= (1 - \theta)^{1+\vartheta}\alpha\mu e + \sum_{i=m+1}^{2m} \theta^i h^i - (\beta - \alpha) \sum_{i=m+1}^{2m} \theta^i \frac{e^T h^i}{n} e \\ &= (1 - \theta)^{1+\vartheta}\alpha\mu e + (\beta - \alpha) \sum_{i=m+1}^{2m} \theta^i \left(h^i - \frac{e^T h^i}{n} e \right) + (1 - \beta + \alpha) \sum_{i=m+1}^{2m} \theta^i h^i \\ &\geq (1 - \theta)^{1+\vartheta}\alpha\mu e - \sum_{i=m+1}^{2m} \theta^i \|h^i\|_2 e. \end{aligned}$$

Therefore,

$$(4.10) \quad z \in \mathcal{D}(\beta) \quad \text{and} \quad \sum_{i=m+1}^{2m} \theta^i \|h^i\|_2 \leq (1 - \theta)^{1+\vartheta}\alpha\mu \quad \text{implies} \quad z(\theta) \in \mathcal{D}(\beta - \alpha).$$

From Lemma 4.1 it follows that

$$\sum_{i=m+1}^{2m} \theta^i \|h^i\|_2 \leq 2\beta\mu \sum_{i=m+1}^{2m} \frac{1}{i} \left(4\theta\sqrt{n/\beta}\right)^i.$$

For any $t \in (0, 1]$ we have

$$\sum_{i=m+1}^{2m} \frac{t^i}{i} \leq t^{m+1} \sum_{i=m+1}^{2m} \frac{1}{i} < t^{m+1} \int_m^{2m} \frac{du}{u} = t^{m+1} \log 2 < .7 t^{m+1}.$$

Hence,

$$\theta \leq \frac{t\sqrt{\beta}}{4\sqrt{n}} \quad \text{and} \quad t \leq 1 \quad \text{implies} \quad \sum_{i=m+1}^{2m} \theta^i \|h^i\|_2 \leq 1.4\beta\mu t^{m+1}.$$

On the other hand if $n \geq 5$, then

$$\theta \leq \frac{t\sqrt{\beta}}{4\sqrt{n}} \text{ and } t \leq 1 \text{ implies } (1 - \theta)^{1+\vartheta} \geq \left(1 - \frac{t\sqrt{\beta}}{4\sqrt{n}}\right)^{1+\vartheta} \geq \left(1 - \frac{1}{4\sqrt{5}}\right)^2 > .7$$

By taking $t = (\alpha/(2\beta))^{1/(m+1)}$ we deduce that if $n \geq 5$, then

$$z(\theta) \in \mathcal{D}(\beta - \alpha), \forall \theta \in [0, \hat{\theta}],$$

which proves that $\check{\theta} \geq \hat{\theta}$. Taking again $t = (\alpha/(2\beta))^{1/(m+1)}$ we have $\hat{\theta} = \frac{t\sqrt{\beta}}{4\sqrt{n}}$, so that we can write

$$\begin{aligned} \frac{\mu(\hat{\theta})}{\mu} &= (1 - \hat{\theta})^{1+\vartheta} + \sum_{i=m+1}^{2m} \hat{\theta}^i \frac{e^T h^i}{n\mu} \leq 1 - \hat{\theta} + \frac{\beta}{n} \sum_{i=m+1}^{2m} \frac{t^i}{i} \\ &= 1 - \frac{t\sqrt{\beta}}{4\sqrt{n}} + \frac{.7\beta t^{m+1}}{n} = 1 - \frac{t\sqrt{\beta}}{4\sqrt{n}} \left(1 - \frac{2.8\sqrt{\beta}t^m}{\sqrt{n}}\right). \end{aligned}$$

Using the definition of t , $n \geq 5$, $0 < \nu \leq 1$, and the fact that

$$\alpha \leq \alpha_0 = \frac{\nu(\bar{\beta} - \beta)}{(\exp(1) + 1) \log^{1+\nu}(\exp(1) + 1)} < \frac{1}{(\exp(1) + 1) \log^2(\exp(1) + 1)},$$

we obtain

$$\frac{2.8\sqrt{\beta}t^m}{\sqrt{n}} \leq \frac{2.8\sqrt{\beta}}{\sqrt{5}} \left(\frac{\alpha}{2\beta}\right)^{\frac{m}{m+1}} \leq \frac{2.8\sqrt{\beta}}{\sqrt{5}} \left(\frac{\alpha}{2\beta}\right)^{\frac{1}{2}} \leq \frac{2\sqrt{\alpha}}{\sqrt{5}} \leq \frac{2\sqrt{\alpha_0}}{\sqrt{5}} < \frac{1}{2}.$$

It follows that $\mu(\hat{\theta}) < \left(1 - \frac{t\sqrt{\beta}}{8\sqrt{n}}\right)\mu$, which completes the proof. \blacksquare

THEOREM 4.3. *If HLCP is monotone, then Algorithm 2 is well defined and produces a sequence of points $(z^k)_{k=0}^\infty$, with $z^k \in \mathcal{D}(\beta_k) \subset \mathcal{D}(\beta)$. If $n \geq 5$ then the following relations hold:*

$$\begin{aligned} \mu_{k+1} &\leq \left(1 - \hat{\theta}_k/2\right) \mu_k, \hat{\theta}_k \geq \frac{\beta^{\frac{m-1}{2(m+1)}}}{4\sqrt{n}} \left(\frac{\alpha_k}{2}\right)^{\frac{1}{m+1}}, k = 0, 1, \dots, \\ \mu_k &\leq \mu_0 \exp\left(-\frac{\bar{\kappa}\nu^{\frac{1}{m+1}} k^{\frac{m}{m+1}}}{\sqrt{n} \log^{\frac{1+\nu}{m+1}} k}\right), \bar{\kappa} := \frac{\beta^{\frac{1}{6}} (\bar{\beta} - \beta)^{\frac{1}{3}}}{2^{\frac{10}{3}}}, k = 8, 9, \dots \end{aligned}$$

Proof. We have $z^k \in \mathcal{D}(\beta_k)$ by construction and $\mu_{k+1} \leq \left(1 - \hat{\theta}_k/2\right)\mu_k$ by virtue of Lemma 4.2. It follows that

$$\mu_k \leq \mu_0 \prod_{j=0}^{k-1} \left(1 - \hat{\theta}_j/2\right).$$

and by using inequality $\log(1 - t) < -t$, (3.14), $k \geq 8$, and Corollary 3.3, we obtain

$$\begin{aligned} \log\left(\frac{\mu_k}{\mu_0}\right) &\leq \sum_{j=0}^{k-1} \log\left(1 - \hat{\theta}_j/2\right) \leq -\frac{1}{2} \sum_{j=0}^{k-1} \hat{\theta}_j \leq -\frac{\beta^{\frac{m-1}{2(m+1)}}}{8\sqrt{n}} \sum_{j=0}^{k-1} \left(\frac{\alpha_j}{2}\right)^{\frac{1}{m+1}} \\ &< -\frac{\beta^{\frac{m-1}{2(m+1)}} (\nu(\bar{\beta} - \beta))^{\frac{1}{m+1}} k^{\frac{m}{m+1}}}{2^{3+\frac{1}{m+1}} \sqrt{n} \log^{\frac{1+\nu}{m+1}} k} \leq -\frac{\beta^{\frac{1}{6}} (\bar{\beta} - \beta)^{\frac{1}{3}} \nu^{\frac{1}{m+1}} k^{\frac{m}{m+1}}}{2^{\frac{10}{3}} \sqrt{n} \log^{\frac{1+\nu}{m+1}} k}, \end{aligned}$$

which completes the proof of our theorem. \blacksquare

In the next corollary we give an upper bound for the maximum number of iterations, K_ϵ required by Algorithm 2 to obtain an approximate solution of the HLCP with duality gap ϵ (see (3.16)). The upper bound is given in terms of the constants $\nu, \bar{\beta}, \underline{\beta}, m$ defining Algorithm 2, the dimension n and the quantity L_ϵ from (3.17) which depends on ϵ and the starting point.

COROLLARY 4.4. *If $\sqrt{n}L_\epsilon \geq \left(\bar{\kappa}\nu^{\frac{1}{m+1}}\right)^{-1} \exp(1)$, then*

$$K_\epsilon \leq \left\lceil \frac{6^{\frac{1+\nu}{m}}}{\bar{\kappa}^{\frac{m+1}{m}} \nu^{\frac{1}{m}}} (\sqrt{n}L_\epsilon)^{\frac{m+1}{m}} \log^{\frac{1+\nu}{m}} (\sqrt{n}L_\epsilon) \right\rceil .$$

Proof. From Theorem 4.3 we deduce that $x^k T s^k \leq \epsilon$ for any K with the property

$$\frac{K^{\frac{m}{m+1}}}{\log^{\frac{1+\nu}{m+1}} K} \geq \sigma := \left(\bar{\kappa}\nu^{\frac{1}{m+1}}\right)^{-1} \sqrt{n}L_\epsilon .$$

According to Corollary 3.5 it follows that

$$K_\epsilon \leq \bar{K}_\epsilon := \left(2 \left(1 + \frac{1+\nu}{m+1} \log \frac{1+\nu}{m} \right) \frac{m+1}{m} \log \sigma \right)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} .$$

By noticing that under our hypothesis $1 + \nu \leq 2 \leq m$, $0 < \bar{\kappa} < 1$, we obtain

$$\begin{aligned} \sigma &\geq \left(\bar{\kappa}\nu^{\frac{1}{m+1}}\right)^{-2} \exp(1) > \exp(1), \\ \log \sigma &\leq \log(\sqrt{n}L_\epsilon) + \log\left(\left(\bar{\kappa}\nu^{\frac{1}{m+1}}\right)^{-1}\right) \\ &\leq \log(\sqrt{n}L_\epsilon) + \log\left(\left(\bar{\kappa}\nu^{\frac{1}{m+1}}\right)^{-1} \exp(1)\right) \leq 2 \log(\sqrt{n}L_\epsilon), \\ \bar{K}_\epsilon &\leq (3 \log \sigma)^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} \leq (6 \log(\sqrt{n}L_\epsilon))^{\frac{1+\nu}{m}} \sigma^{\frac{m+1}{m}} \\ &\leq \frac{6^{\frac{1+\nu}{m}}}{\bar{\kappa}^{\frac{m+1}{m}} \nu^{\frac{1}{m}}} (\sqrt{n}L_\epsilon)^{\frac{m+1}{m}} \log^{\frac{1+\nu}{m}} (\sqrt{n}L_\epsilon) . \end{aligned}$$

If ν is a positive constant independent of n and L_ϵ , then Algorithm 1 is independent of the dimension of the problem and the stopping criterium $x^k T s^k \leq \epsilon$. However, by letting ν depend on n and L_ϵ , we can slightly improve the computational complexity. \blacksquare

COROLLARY 4.5. *If*

$$\sqrt{n}L_\epsilon \geq \max \left\{ \exp(\exp(1)), \frac{2.072 \exp(1)}{\bar{\kappa}} \log \log \left(\frac{\exp(1)}{\bar{\kappa}} \right) \right\} ,$$

and if we take $\nu = (\log \log(\sqrt{n}L_\epsilon))^{-1}$ in Algorithm 2, then

$$K_\epsilon \leq \left\lceil \frac{36^{\frac{1}{m}} \exp(1/m)}{\bar{\kappa}^{\frac{m+1}{m}}} (\sqrt{n}L_\epsilon)^{\frac{m+1}{m}} (\log(\sqrt{n}L_\epsilon) \log \log(\sqrt{n}L_\epsilon))^{\frac{1}{m}} \right\rceil .$$

Proof. Using Lemma 3.6 we deduce that

$$\frac{\sqrt{n}L_\epsilon}{\log^{\frac{1}{m}} \log(\sqrt{n}L_\epsilon)} \geq \frac{\sqrt{n}L_\epsilon}{\log \log(\sqrt{n}L_\epsilon)} \geq \frac{\exp(1)}{\bar{\kappa}},$$

which shows that Corollary 4.4 can be applied for our choice of ν . By noticing that $\log^{\frac{\nu}{m}}(\sqrt{n}L_\epsilon) = \exp(1/m)$, and $6^{1+\nu} \leq 6^2 = 36$, we obtain the desired result. \blacksquare

COROLLARY 4.6. *Assume that*

$$\sqrt{n}L_\epsilon \geq \max \left\{ \exp(\exp(1)), \frac{2.072 \exp(1)}{\bar{\kappa}} \log \log \left(\frac{\exp(1)}{\bar{\kappa}} \right) \right\},$$

and consider Algorithm 2 with $\nu = (\log \log(\sqrt{n}L_\epsilon))^{-1}$. Then the following implications hold:

$$\begin{aligned} m \geq \log \log \log(\sqrt{n}L_\epsilon) &\Rightarrow K_\epsilon \leq \left\lceil \frac{36^{\frac{1}{m}} \exp(1 + 1/m)}{\bar{\kappa}^{\frac{m+1}{m}}} (\sqrt{n}L_\epsilon)^{\frac{m+1}{m}} \log^{\frac{1}{m}}(\sqrt{n}L_\epsilon) \right\rceil, \\ m \geq \log \log(\sqrt{n}L_\epsilon) &\Rightarrow K_\epsilon \leq \left\lceil \frac{36^{\frac{1}{m}} \exp(2 + 1/m)}{\bar{\kappa}^{\frac{m+1}{m}}} (\sqrt{n}L_\epsilon)^{\frac{m+1}{m}} \right\rceil, \\ m \geq \log(\sqrt{n}L_\epsilon) &\Rightarrow K_\epsilon \leq \left\lceil \frac{36^{\frac{1}{m}} \exp(3 + 1/m)}{\bar{\kappa}^{\frac{m+1}{m}}} \sqrt{n}L_\epsilon \right\rceil. \end{aligned}$$

4.4. Higher order convergence. As seen in the previous subsection the computational complexity of Algorithm 2 is basically the same for $\vartheta = 1$ and $\vartheta = 0$. By contrast its asymptotic convergence properties depend on ϑ . In what follows we show that Algorithm 2 with $\vartheta = 1$ is superlinearly convergent for general problems. However, if the problem is known to have a strictly complementary solution it is advantageous to take $\vartheta = 0$ in order to obtain a higher order of convergence.

LEMMA 4.7. *Assume HLCP is monotone and consider the linear systems (4.4) where we take $\vartheta = 1$ for general problems, and $\vartheta = 0$ for problems that are known to have a strictly complementary solution. Then for any $\beta \in (0, 1)$, any $\bar{\mu} > 0$, and any integer $m \geq 2$, there is constant $c = c(\beta, \bar{\mu}, m)$ such that the solution $u^1, v^1, \dots, u^m, v^m$ of (4.4) satisfies*

$$\|u^i\|_2 \leq c \mu^{\frac{1}{1+\vartheta}}, \|v^i\|_2 \leq c \mu^{\frac{1}{1+\vartheta}}, \forall z \in \left\{ [x, s] \in \mathcal{D}(\beta) : \mu = \frac{x^T s}{n} \leq \bar{\mu} \right\}.$$

Proof. Use (4.3) and apply Theorem 2.3 with \mathcal{K} given by (3.18). \blacksquare

THEOREM 4.8. *Assume HLCP is monotone and consider Algorithm 2 where we take $\vartheta = 1$ for general problems, and $\vartheta = 0$ for problems that are known to have a strictly complementary solution. Then the sequence (z^k) produced by this algorithm converges Q -superlinearly to a maximal complementary solution $z^* \in \mathcal{F}^c$, and the sequence of the corresponding complementarity gaps (μ_k) converge Q -superlinearly to zero. Moreover, the Q -orders of convergence of these sequences satisfy*

$$\mathcal{Q}(z^k) = \mathcal{Q}(\mu_k) \geq \frac{m+1}{1+\vartheta}.$$

Proof. Since we are analyzing asymptotic properties we may assume $\mu_k < 1$. By using Theorem 4.3 we obtain

$$0 \leq \frac{\log \alpha_k}{\log \mu_k} \leq \frac{\log(\nu(\bar{\beta} - \underline{\beta})) - \log((k + \exp(1) + 1) \log^{1+\nu}(k + \exp(1) + 1))}{\log \mu_0 - \frac{\frac{1}{\bar{\kappa}\nu} \frac{m+1}{k} \frac{m}{m+1}}{\sqrt{n} \log^{\frac{1+\nu}{m+1}} k}}.$$

It follows that $\lim_{k \rightarrow \infty} \frac{\log \alpha_k}{\log \mu_k} = 0$, which implies

$$\lim_{k \rightarrow \infty} \frac{\mu_k^{\frac{m-\vartheta}{1+\vartheta}}}{\alpha_k} = 0.$$

From Theorem 4.3 it follows that $p^k := (x^k s^k) / \mu_k \in \mathcal{K}(\underline{\beta}, \mu_0)$, $\forall k \geq 0$, so that by using Lemma 4.7 we deduce that there is a constant $c = c(\underline{\beta}, \mu_0, m)$ such that

$$(4.11) \quad \|u^i\|_2 \leq c \mu^{\frac{i}{1+\vartheta}}, \quad \|v^i\|_2 \leq c \mu^{\frac{i}{1+\vartheta}}, \quad \text{at each iteration } k = 0, 1, \dots$$

From (4.7) it follows that

$$\|h^i\|_2 \leq \sum_{j=i-m}^m \|u^j\|_2 \|v^{i-j}\|_2 \leq mc^2 \mu^{\frac{i}{1+\vartheta}}.$$

For k sufficiently large we have $\mu^{\frac{1}{1+\vartheta}} \leq 1/2$, so that

$$\sum_{i=m+1}^{2m} \theta^i \|h^i\|_2 \leq \sum_{i=m+1}^{2m} \|h^i\|_2 \leq mc^2 \mu^{\frac{m+1}{1+\vartheta}} \sum_{i=0}^{m-1} \mu^{\frac{i}{1+\vartheta}} \leq 2mc^2 \mu^{\frac{m+1}{1+\vartheta}} = \bar{c} \mu^{\frac{m+1}{1+\vartheta}},$$

with $\bar{c} := 2mc^2$. For k sufficiently large we have $(\bar{c}/\alpha) \mu^{\frac{m-\vartheta}{1+\vartheta}} < 1$, and therefore

$$\sum_{i=m+1}^{2m} \theta^i \|h^i\|_2 \leq (1-\theta)^{1+\vartheta} \alpha \mu, \quad \forall \theta \in [0, \hat{\theta}], \quad \hat{\theta} = 1 - (\bar{c}/\alpha)^{\frac{1}{1+\vartheta}} \mu^{\frac{m-\vartheta}{(1+\vartheta)^2}}.$$

According to (4.6) and (4.10) we have $\check{\theta} \geq \hat{\theta}$, so that by using (4.8) we obtain

$$\begin{aligned} \mu_+ &= \mu(\bar{\theta}) \leq \mu(\hat{\theta}) = (1-\hat{\theta})^{1+\vartheta} \mu + \sum_{i=m+1}^{2m} \hat{\theta}^i (e^T h^i / n) \\ &\leq (1-\hat{\theta})^{1+\vartheta} \mu + \frac{1}{\sqrt{n}} \sum_{i=m+1}^{2m} \hat{\theta}^i \|h^i\|_2 \leq (1-\hat{\theta})^{1+\vartheta} \left(1 + \frac{\alpha}{\sqrt{n}}\right) \mu \\ &\leq \bar{c} \left(\frac{1}{\alpha} + \frac{1}{\sqrt{n}}\right) \mu^{\frac{m+1}{1+\vartheta}} \leq \frac{2\bar{c}}{\alpha} \mu^{\frac{m+1}{1+\vartheta}}. \end{aligned}$$

By taking logarithms we obtain

$$\frac{\log \mu_+}{\log \mu} \geq \frac{m+1}{1+\vartheta} + \frac{\log(2\bar{c})}{\log \mu} - \frac{\log \alpha}{\log \mu}.$$

Since the right-hand-side of the above inequality tends to zero as $k \rightarrow \infty$, we deduce that

$$\mathcal{Q}(\mu_k) = \liminf \frac{\log \mu_{k+1}}{\log \mu_k} \geq \frac{m+1}{1+\vartheta}.$$

For k sufficiently large we have $(c\mu)^{\frac{1}{1+\vartheta}} < 1 - \sqrt{2}$ so that by using Lemma 4.7 obtain

$$\|z^+ - z\|_2 \leq \sum_{i=1}^m \|w^i\|_2 \leq \sqrt{2} \sum_{i=1}^m (c\mu)^{\frac{i}{1+\vartheta}} = \sqrt{2} \frac{(c\mu)^{\frac{1}{1+\vartheta}} - (c\mu)^{\frac{m+1}{1+\vartheta}}}{1 - (c\mu)^{\frac{1}{1+\vartheta}}} \leq 2(c\mu)^{\frac{1}{1+\vartheta}}.$$

Finally, by applying Theorem 2 of [28] we deduce the convergence of the sequence (z^k) to a maximal complementary solution $z^* \in \mathcal{F}^*$ and the fact that $\mathcal{Q}(z^k) = \mathcal{Q}(\mu_k)$. ■

Acknowledgement. The author would like to thank Josef Stoer for carefully reading the manuscript and suggesting several improvements.

REFERENCES

- [1] M. Anitescu, G. Lesaja, and F. A. Potra. An infeasible–interior–point predictor–corrector algorithm for the P_* -Geometric LCP. *Applied Mathematics & Optimization*, 36(2):203–228, 1997.
- [2] K.M Anstreicher and R.A. Bosch. A new infinity-norm path following algorithm for linear programming. *SIAM J. Optim.*, 5(2):236–246, 1995.
- [3] J. F. Bonnans and C. C. Gonzaga. Convergence of interior point algorithms for the monotone linear complementarity problem. *Mathematics of Operations Research*, 21:1–25, 1996.
- [4] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Academic Press, Boston, MA, 1992.
- [5] A.S. El-Bakry, R.A. Tapia, and Y. Zhang. A study of indicators for identifying zero variables in interior-point methods. *SIAM Review*, 36(1):45–72, 1994.
- [6] C. C. Gonzaga. The largest step path following algorithm for monotone linear complementarity problems. *Math. Programming*, 76(2, Ser. A):309–332, 1997.
- [7] C. C. Gonzaga. Complexity of predictor-corrector algorithms for LCP based on a large neighborhood of the central path. *SIAM J. Optim.*, 10(1):183–194 (electronic), 1999.
- [8] B. Jansen, C. Roos, and T. Terlaky. A polynomial primal-dual Dikin-type algorithm for linear programming. *Math. Oper. Res.*, 21(2):341–353, 1996.
- [9] B. Jansen, C. Roos, and T. Terlaky. A family of polynomial affine scaling algorithms for positive semidefinite linear complementarity problems. *SIAM J. Optim.*, 7(1):126–140, 1997.
- [10] B. Jansen, C. Roos, T. Terlaky, and Y. Ye. Improved complexity using higher-order correctors for primal-dual Dikin affine scaling. *Math. Programming*, 76(1, Ser. B):117–130, 1997. Interior point methods in theory and practice (Iowa City, IA, 1994).
- [11] J. Ji, F. A. Potra, and S. Huang. A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence. *Journal of Optimization Theory and Applications*, 84(1):187–199, 1995.
- [12] N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [13] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, volume 538 of *Lecture Notes in Comput. Sci.* Springer-Verlag, New York, 1991.
- [14] M. Kojima, S. Mizuno, and A. Yoshise. A polynomial-time algorithm for a class of linear complementarity problems. *Math. Programming*, 44(1, (Ser. A)):1–26, 1989.
- [15] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point algorithm for linear programming. In *Progress in mathematical programming (Pacific Grove, CA, 1987)*, pages 29–47. Springer, New York, 1989.
- [16] M. Kojima, S. Mizuno, and A. Yoshise. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems. *Math. Programming*, 50(3, (Ser. A)):331–342, 1991.
- [17] K. A. McShane. Superlinearly convergent $O(\sqrt{n}L)$ -iteration interior-point algorithms for linear programming and the monotone linear complementarity problem. *SIAM Journal on Optimization*, 4(2):247–261, 1994.
- [18] S. Mizuno. A superlinearly convergent infeasible-interior-point algorithm for geometrical LCPs without a strictly complementary condition. *Math. Oper. Res.*, 21(2):382–400, 1996.
- [19] S. Mizuno and A. Nagasawa. A primal-dual affine-scaling potential-reduction algorithm for linear programming. *Math. Programming*, 62(1, Ser. B):119–131, 1993.
- [20] S. Mizuno, M. J. Todd, and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18(4):964–981, 1993.

- [21] R. D. C. Monteiro and I. Adler. Interior path following primal-dual algorithms. I. Linear programming. *Math. Programming*, 44(1, (Ser. A)):27–41, 1989.
- [22] R. D. C. Monteiro and I. Adler. Interior path following primal-dual algorithms. II. Convex quadratic programming. *Math. Programming*, 44(1, (Ser. A)):43–66, 1989.
- [23] R. D. C. Monteiro, I. Adler, and M. G. C. Resende. A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension. *Math. Oper. Res.*, 15(2):191–214, 1990.
- [24] R. D. C. Monteiro and S. J. Wright. Local convergence of interior-point algorithms for degenerate monotone LCP. *Computational Optimization and Applications*, 3:131–155, 1994.
- [25] R. D. C. Monteiro and S. J. Wright. Superlinear primal-dual affine scaling algorithms for LCP. *Math. Programming*, 69(2, Ser. A):311–333, 1995.
- [26] J. Peng, T. Terlaky, and Y. B. Zhao. A self-regularity based predictor-corrector algorithm for linear optimization. Preprint submitted for publication, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada.
- [27] F. A. Potra. On Q -order and R -order of convergence. *Journal of Optimization Theory and Applications*, 63:415–431, 1989.
- [28] F. A. Potra. Q -superlinear convergence of the iterates in primal-dual interior-point methods. *Math. Programming*, 91(Ser. A):99–115, 2001.
- [29] F. A. Potra. The Mizuno-Todd-Ye algorithm in a larger neighborhood of the central path. *European Journal of Operational Research*, 143:257–267, 2002.
- [30] F. A. Potra. A superlinearly convergent predictor-corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{nl})$ -iteration complexity. *Math. Programming*, 100:317–337, 2004.
- [31] F. A. Potra and X. Liu. Predictor-corrector methods for sufficient linear complementarity problems in a wide neighborhood of the central path. *Optimization Methods and Software*, 20(1):145–168, 2005.
- [32] F. A. Potra and R. Sheng. A path following method for LCP with superlinearly convergent iteration sequence. *Ann. Oper. Res.*, 81:97–114, 1998. Applied mathematical programming and modeling, III (APMOD95) (Uxbridge).
- [33] F. A. Potra and R. Sheng. Superlinearly convergent infeasible-interior-point algorithm for degenerate LCP. *Journal of Optimization Theory and Applications*, 97(2):249–269, 1998.
- [34] J. Renegar. A polynomial-time algorithm, based on Newton’s method, for linear programming. *Math. Programming*, 40(1, Ser. A):59–93, 1988.
- [35] 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.
- [36] J. Stoer. Improved high order long-step methods for solving linear complementarity problems. Technical Report 237, Institut für Angewandte Mathematik und Statistik Universität Würzburg, 97074 Würzburg, Germany, 1999.
- [37] J. Stoer and M. Wechs. The complexity of high-order predictor-corrector methods for solving sufficient linear complementarity problems. *Optim. Methods Softw.*, 10(2):393–417, 1998. Dedicated to Professor Masao Iri on the occasion of his 65th birthday.
- [38] J. Stoer and M. Wechs. Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity. *Math. Programming*, 83(3, Ser. A):407–423, 1998.
- [39] J. Stoer and M. Wechs. On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems. *Numer. Math.*, 81(4):631–645, 1999.
- [40] J. Stoer, M. Wechs, and S. Mizuno. High order infeasible-interior-point methods for solving sufficient linear complementarity problems. *Math. Oper. Res.*, 23(4):832–862, 1998.
- [41] Josef Stoer. High order long-step methods for solving linear complementarity problems. *Ann. Oper. Res.*, 103:149–159, 2001. Optimization and numerical algebra (Nanjing, 1999).
- [42] J. F. Sturm. Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists. *Mathematics of Operations Research*, 24:72–94, 1999.
- [43] K. Tanabe. Centered Newton method for mathematical programming. In *System modelling and optimization (Tokyo, 1987)*, volume 113 of *Lecture Notes in Control and Inform. Sci.*, pages 197–206. Springer, Berlin, 1988.
- [44] M. J. Todd and Y. Ye. A centered projective algorithm for linear programming. *Math. Oper. Res.*, 15(3):508–529, 1990.
- [45] L. Tunçel. Constant potential primal-dual algorithms: a framework. *Math. Programming*, 66(2, Ser. A):145–159, 1994.
- [46] L. Tunçel. On the convergence of primal-dual interior-point methods with wide neighborhoods. *Comput. Optim. Appl.*, 4(2):139–158, 1995.

- [47] Reha H. Tütüncü. Quadratic convergence of potential-reduction methods for degenerate problems. *Math. Program.*, 90(1, Ser. A):169–203, 2001.
- [48] S. J. Wright. *Primal–Dual Interior–Point Methods*. SIAM Publications, Philadelphia, 1997.
- [49] Y. Ye. *Interior Point Algorithms : Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.
- [50] Y. Ye and K. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP. *Mathematical Programming*, 62(3):537–551, 1993.
- [51] Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$ -iteration algorithm for linear programming. *Mathematical Programming*, 59(2):151–162, 1993.
- [52] Y. Zhang, R. A. Tapia, and J. E. Dennis. On the superlinear and quadratic convergence of primal–dual interior point linear programming algorithms. *SIAM Journal on Optimization*, 2(2):304–324, 1992.
- [53] Y. Zhang, R. A. Tapia, and F. A. Potra. On the superlinear convergence of interior point algorithms for a general class of problems. *SIAM J. on Optimization*, 3(2):413–422, 1993.
- [54] G. Zhao. On the relationship between the curvature integral and the complexity of path-following methods in linear programming. *SIAM J. Optim.*, 6(1):57–73, 1996.
- [55] G. Y. Zhao and J. S. Zhu. The curvature integral and the complexity of linear complementarity problems. *Math. Programming*, 70(1, Ser. A):107–122, 1995.