SIAM J. OPTIM.
Vol. 20, No. 3, pp. 1333–1363

ON A CLASS OF SUPERLINEARLY CONVERGENT POLYNOMIAL TIME INTERIOR POINT METHODS FOR SUFFICIENT LCP*

FLORIAN A. POTRA† AND JOSEF STOER‡

Abstract. A new class of infeasible interior point methods for solving sufficient linear complementarity problems (LCPs) requiring one matrix factorization and m backsolves at each iteration is proposed and analyzed. The algorithms from this class use a large ($\mathcal{N}_\infty^-$) neighborhood of an infeasible central path associated with the complementarity problem and an initial positive, but not necessarily feasible, starting point. The Q-order of convergence of the complementarity gap, the residual, and the iteration sequence is $m+1$ for problems that admit a strict complementarity solution and $(m+1)/2$ for general sufficient LCPs. The methods do not depend on the handicap κ of the sufficient LCP. If the starting point is feasible (or "almost" feasible), the proposed algorithms have $\mathcal{O}((1+\kappa)(1+\log \sqrt[m]{1+\kappa})\sqrt{n}\, L)$ iteration complexity, while if the starting point is "large enough," the iteration complexity is $\mathcal{O}((1+\kappa)^{2+1/m}(1+\log \sqrt[m]{1+\kappa})n\, L)$.

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

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

DOI. 10.1137/080716979

1. Introduction.

Interior point methods have provided the first polynomial time algorithms for solving linear programming (LP) and other classes of convex optimization problems. Numerical experiments performed over the past two decades show that the most efficient interior point methods for LP are primal-dual path-following interior point methods. The fact that these methods, as opposed to the ellipsoid method, perform much better in practice than indicated by the (worst case) computational complexity bounds is explained in part by their superlinear convergence. 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 [16], where it was shown to have $\mathcal{O}(\sqrt{n}L)$ iteration complexity. Shortly after that, Ye et al. [45], and independently Mehrotra [14], proved that the duality gap of the iterates produced by MTY converges quadratically to zero. MTY was generalized to linear complementarity problems (LCPs) in [9], and the resulting algorithm was proved to have $\mathcal{O}(\sqrt{n}L)$ iteration complexity under general conditions, and superlinear convergence under the assumption that the LCP has a (perhaps not unique) strictly complementary solution (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 [44] proved that MTY converges quadratically assuming only that the LCP is nondegenerate. The nondegeneracy assumption is not restrictive since according to [17] a large class of interior point methods, which contains MTY, can have only linear convergence if this

*Received by the editors February 27, 2008; accepted for publication (in revised form) May 28, 2009; published electronically October 16, 2009.
http://www.siam.org/journals/siopt/20-3/71697.html
†Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250 (potra@math.umbc.edu). The work of this author was supported by the National Science Foundation under grants 0139701, 0728878, and by the National Institute of Health under grant R01GM075298-01.
‡Institut für Angewandte Mathematik und Statistik Universität Würzburg Am Hubland D-97074 Würzburg Germany (jstoer@mathematik.uni-wuerzburg.de).

assumption is violated. In [9, 14, 16, 45, 44] one assumes that the starting point for the MTY algorithm is strictly feasible. A generalization of the MTY algorithm for infeasible starting points was proposed in [21, 20] for LP and in [22] for monotone LCPs. The proposed algorithm requires two matrix factorizations and at most three backsolves per iteration. Its computational complexity depends on the quality of the starting point. If the starting point is large enough, then the algorithm has $\mathcal{O}(nL)$ iteration complexity. If a certain measure of feasibility at the starting point is small enough, then the algorithm has $\mathcal{O}(\sqrt{n}L)$ iteration complexity. At each iteration, both "feasibility" and "optimality" are reduced exactly at the same rate. Moreover, the algorithm is quadratically convergent for problems having a strictly complementary solution. The generalizations of the MTY algorithm for LCPs with infeasible starting points from [10, 33] require two matrix factorizations and two backsolves per iteration and have the same properties.

According to Ostrowski [18], the asymptotic efficiency index is defined as $(ord)^{1/c}$, where *ord* is the Q-order of the iterative procedure and c is the cost of an iteration. Since the main cost associated with interior point methods is represented by matrix factorizations, it follows that the asymptotic efficiency index of the MTY type algorithms mentioned above is equal to $\sqrt{2}$. The problem arises then to find a superlinearly convergent interior point method with polynomial complexity that requires only one matrix factorization per step. The first result of this type was obtained by McShane [13] who proved that the so-called largest step path-following (LSPF) method has $\mathcal{O}(\sqrt{n}L)$ iteration complexity for general (monotone) LCPs and superlinear convergence under the assumption that the LCP is nondegenerate and the iteration sequence converges. No results about the Q-order were given. Gonzaga [8] 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]. The results of [3, 8] were extended for the infeasible case in [4]. An infeasible interior point method with polynomial and quadratic convergence for nondegenerate LCPs that uses only one matrix factorization per iteration was proposed by Wright [42]. It is based on the so-called fast step – safe step strategy. At each iteration the algorithm first tries to take a "fast step" along the primal-dual affine-scaling direction. If the fast step fails to reduce the complementarity gap by a given factor, then the algorithm reverts to a "safe step" along a direction consisting of a convex combination of the affine-scaling direction and the centering direction. Higher order variants of this strategy are given in [29, 43]. The "fast step – safe step" strategy is also used in a variant of the Mehrotra predictor-corrector method proposed by Zhang and Zhang [46], which has polynomial complexity and superlinear convergence and requires one matrix factorization per step.

The first superlinear convergence result for degenerate LCPs was obtained by Mizuno [15], who used the so-called step variant of the Tapia indicator [7] 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. The resulting algorithm has Q-order 1.25. Mizuno's result was refined in [30, 31, 32]. Subsequently, by using second order derivatives of weighted central paths, Sturm [40] obtained a MTY type algorithm of order 1.5 for degenerate LCPs. His approach was generalized in [39], where mth order derivatives were used to construct MTY type algorithms with Q-order $m+1$ for nondegenerate LCPs and $(m+1)/2$ for degenerate LCPs. The complexity of the predictor-corrector algorithm for degenerate LCPs from [39] is analyzed in [36]. The algorithms from [36, 39] require two matrix factorizations and $m+1$ backsolves per iteration. The algorithm proposed in [35] requires only one matrix factorization and m backsolves per iterations and has Q-order m in the nondegenerate case and $m/2$ in the degenerate case. The results of [35] were improved in the unpublished technical report [34], where an algorithm with the same computational cost is shown to have Q-order $m+1$ in the nondegenerate case and $(m+1)/2$ in the degenerate case. No computational complexity results are available for the algorithms of [34, 35]. We mention that the algorithms of [36, 39] use a small ($\mathcal{N}_2$) neighborhood of the central path, while the algorithms of [34, 35] use a large ($\mathcal{N}_\infty^-$) neighborhood of the central path.

Algorithms using large neighborhoods of the central path are more difficult to analyze, and in general their computational complexity estimates are worse than the corresponding estimates for algorithms using small neighborhoods. In fact this is one of the paradoxes of interior point methods, since it is known that the best practical results are achieved by interior point methods acting in a large neighborhood of the central path. For many years this fact has been accepted as an inherent difference between theory and practice. However, recent research has shown that it is possible to design superlinearly convergent interior point methods in the large neighborhood of the central path that have optimal, or close to optimal, computational complexity. In [1, 19] one improves the complexity by using new large neighborhoods of the central path and first order information, while in [24, 26] we use the classical $\mathcal{N}_\infty^-$ large neighborhood and the complexity is reduced by the use of higher order information. The latter approach ensures superlinear convergence even in the absence of strict complementarity. The algorithms of [24, 26] require two matrix factorizations and $m+1$ backsolves per iteration and have Q-order $m+1$ in the nondegenerate case and $(m+1)/2$ in the degenerate case. These algorithms have been generalized for sufficient LCPs in [12, 28]. The handicap $\kappa \geq 0$ of a sufficient LCP measures by how much the LCP differs from a monotone LCP (see the next section). Since there are no polynomial complexity results for the sufficient LCP, the iteration complexity of interior point methods for such problems will depend on the handicap κ. However, the interior point algorithms that do not depend explicitly on κ can be applied to any sufficient LCP. We mention that the algorithms from [36, 39] depend on κ, while the algorithm from [12, 34, 35] and Algorithms 1A and 2A from [28] do not depend on κ.

In a recent paper [27], the author has analyzed several feasible interior point methods for the monotone LCP acting in the $\mathcal{N}_\infty^-$ neighborhood of the central path that require one matrix factorization and m backsolves per iteration and have Q-order $m+1$ in the nondegenerate case and $(m+1)/2$ in the degenerate case, the same as the algorithm proposed by the second author of the present paper in [34]. The algorithms from [27] and [34] share the property that the $\mathcal{N}_\infty^-$ neighborhood of the central path is expanded at each iteration. The rate of expansion used in [34] ensures the high Q-order results mentioned above but does not guarantee polynomial complexity. However, by using the rate of expansion from [27] it is possible to modify the method of [34] in such a way that Q-order results are retained and polynomial complexity is guaranteed.

In the present paper we propose a class of polynomial time infeasible interior point methods for sufficient LCPs that have the same Q-order and computational cost per iteration as the methods from [34] and [27] (i.e., the same asymptotic efficiency index). The class depends on a parameter $\sigma \geq 0$. When $\sigma = 0$, the algorithms from this class reduce to the corresponding algorithms from [27], while for $\sigma > 0$ they reduce to a variant of the algorithm of [34]. The main contribution of the paper is that it generalizes the results of [27] to sufficient LCPs with infeasible starting points and it proposes a variant of the algorithm of [34] that has both superlinear convergence and polynomial complexity. Moreover, we hope that our analysis will also contribute to a better understanding of the behavior of interior point methods in the large neighborhood of the central path. The interior point methods to be presented in this paper do not depend on the handicap κ of the sufficient LCP. Their computational complexity depends on n, m, and κ. More precisely we will show that an approximate solution having both the complementarity gap and the norm of the residual less than ε can be obtained in at most $\mathcal{O}((\log((1+\kappa)^{1+\chi-\frac{1}{m+1}}n^{\frac{1+\chi}{2}}L))^{\frac{1+\nu}{m}}((1+\kappa)^{1+\chi-\frac{1}{m+1}}L)^{\frac{m+1}{m}}n^{\frac{1+\chi}{2}})$ iterations, where $L = \log(\varepsilon_0/\varepsilon)$, $0 < \nu \le 1$ is a parameter from the definition of the algorithm, $\chi = 0$ if the starting point is feasible or "almost feasible," and $\chi = 1$ if the starting point is "large enough." If we take $m \ge \max\{\log n\,,\ L\,,\ 2\}$, then the algorithm has $\mathcal{O}((1+\frac{\log(1+\kappa)}{m})(1+\kappa)^{\frac{(1+\chi)(m+1)-1}{m}}n^{\frac{1+\chi}{2}}L)$ iteration complexity. Since the amount of work per iteration of an mth order method ($\mathcal{O}(n^3+mn^2)$ arithmetic operations) is comparable, for small m/n, to that of a first order method ($\mathcal{O}(n^3)$ arithmetic operations), we consider that it is justified to compare all methods by means of their iteration complexity.

The paper is organized as follows. In section 2 we present some basic results on the sufficient horizontal LCP and the analyticity of its weighted infeasible central paths. In section 3 we analyze a class of interior point methods based on the first m derivatives of the weighted central paths. We prove that the methods are globally and superlinearly convergent under general assumptions. We also prove that they have polynomial complexity for starting points that are either "almost feasible" or "large enough." The paper ends with a very short section of conclusions.

Conventions. We denote by $I\!N$ the set of all nonnegative integers. $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_{++}$ denote the set of real, nonnegative real, and positive real numbers, respectively. 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 componentwise 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_iv_i, u_i/v_i, etc., respectively. This notation is consistent as long as componentwise 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 $A(uv) \ne (Au)v$. 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 $v \ge 3$ means that $v_i \ge 3,\ i = 1, \ldots, n$. Traditionally this is written as $v \ge 3e$. 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 $\|U\|_2 = \|u\|_\infty$.

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

2. The sufficient horizontal LCP. 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 $z = \lceil x,\, s \rfloor$ such that

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

Let us denote by Φ the null space of the matrix $[Q\ R] \in \mathbb{R}^{n\times 2n}$,

$$\Phi := \mathcal{N}([Q\ R]) = \{\lceil u,\, v \rfloor\ :\ Qu + Rv = 0\} \tag{2.2}$$

and by $\Phi^\perp$ its orthogonal space

$$\Phi^\perp = \left\{\lceil u,\, v \rfloor\ :\ u = Q^T x,\ v = R^T x \text{ for some } x \in \mathbb{R}^n\right\}\ . \tag{2.3}$$

We say that the pair (Q, R) is column sufficient if

$$\lceil u,\, v \rfloor \in \Phi, \quad uv \leq 0 \text{ implies } uv = 0\,,$$

and row sufficient if

$$\lceil u,\, v \rfloor \in \Phi^\perp, \quad uv \geq 0 \text{ implies } uv = 0\,.$$

(Q, R) is called a sufficient pair if it is both column and row sufficient, and in this case (2.1) is called a sufficient HLCP. The standard sufficient LCP is obtained by taking $R = -I$ and Q a sufficient matrix in the sense of Cottle, Pang, and Venkateswaran in [6] (see also [5]). The notion of a sufficient matrix is a far reaching generalization of the notion of a positive semidefinite matrix. The notion of a sufficient pair was introduced in [37, 39]. It turns out that (Q, R) is a sufficient pair if and only if, for any b, the HLCP (2.1) has a convex (perhaps empty) solution set and every KKT point of

$$\begin{array}{cc} \min_{x,s} & x^T s \\ \text{subject to} & Qx + Rs = b, \\ & x, s \geq 0, \end{array}$$

is a solution of (2.1).

By using the result of Väliaho [41] and the equivalence results from [2] (see also [35]) it follows that (Q, R) is a sufficient pair if and only if there is a $\kappa \geq 0$ such that

$$Qu + Rv = 0 \text{ implies } (1 + 4\kappa) \sum_{i\in\mathcal{I}^+} u_i v_i + \sum_{i\in\mathcal{I}^-} u_i v_i \geq 0\ \forall u, v \in \mathbb{R}^n, \tag{2.4}$$

where $\mathcal{I}^+ = \{i : u_i v_i > 0\}$ and $\mathcal{I}^- = \{i : u_i v_i < 0\}$. If the above condition is satisfied, we say that (Q, R) is a $P_*(\kappa)$ pair and we write $(Q, R) \in P_*(\kappa)$. In case $R = -I$, $(Q, -I)$ is a $P_*(\kappa)$ pair if and only if Q is a $P_*(\kappa)$ matrix in the sense that

$$(1 + 4\kappa) \sum_{i\in\hat{\mathcal{I}}^+} x_i [Qx]_i + \sum_{i\in\hat{\mathcal{I}}^-} x_i [Qx]_i \geq 0 \quad \forall x \in \mathbb{R}^n\,,$$

where $\hat{\mathcal{I}}^+ = \{i : x_i [Qx]_i > 0\}$ and $\hat{\mathcal{I}}^- = \{i : x_i [Qx]_i < 0\}$. Problem (2.1) is then called a $P_*(\kappa)$ LCP, and it is extensively discussed in [11]. Thus the class of sufficient pairs coincides with the class $P_* = \cup_{\kappa\geq 0} P_*(\kappa)$. In view of this equivalence the terms P_* pair, sufficient pair, P_* HLCP, and sufficient HLCP will be used interchangeably, although (as illustrated by the title of our paper) preference will be given to the latter. The following result was proved in [12].

LEMMA 2.1. *If $Q, R \in \mathbb{R}^{n\times n}$ are two matrices such that the pair (Q, R) is column sufficient, then the matrix $[Q\ R]$ has full rank.*

An example was also given showing that row sufficiency does not imply the full rank property. The full rank property is used in proving the existence of weighted infeasible central paths defined by nonlinear systems of the form

$$\begin{aligned} xs &= \tau p, \quad x,\, s \geq 0, \\ Qx + Rs &= b - \tau \bar{b}\,. \end{aligned} \tag{2.5}$$

A precise statement of this fact will be given later. For the time being, let us note that for $\tau = 0$, (2.5) reduces to (2.1). We denote the set of all feasible points of (2.1) by

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

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

$$\mathcal{F}^* = \{z^* = \lceil x^*,\, s^* \rfloor \in \mathcal{F} \;:\; x^* s^* = 0\}.$$

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 $\{1, \ldots, n\}$ by

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

One can prove that $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{J}$ form a partition of $\{1, \ldots, n\}$ and that there exists a solution $\lceil x^*,\, s^* \rfloor \in \mathcal{F}^*$ 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 $\lceil x^*,\, s^* \rfloor \in \mathcal{F}^*$, $x_i^* > 0 \Rightarrow i \in \mathcal{B}$ 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.,

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

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.

If the HLCP is a reformulation of a LP problem, then it is known that it has a strictly complementary solution. In general it is very difficult to establish whether the 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

$$\vartheta = \begin{cases} 1 & \text{in general (default option),} \\ 0 & \text{if HLCP is known to be nondegenerate.} \end{cases} \tag{2.6}$$

In the theory of interior point methods it is often assumed that the set

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

called the set of strictly feasible points, is nonempty. In the present paper we are not making this assumption. Instead we assume the nonemptyness of the set

$$\mathcal{F}_{\overline{b},\tau} := \{z = \lceil x,\, s \rfloor \in \mathbb{R}^{2n}_{++} \;:\; Qx + Rs = b - \tau\overline{b}\}\,, \tag{2.7}$$

which is related to the weighted infeasible central path defined by (2.5) at $\tau = \tau_0$.

The distinction between nondegenerate problems ($\vartheta = 0$) and (perhaps) degenerate problems ($\vartheta = 1$) plays a role in the following theorem.

THEOREM 2.2. *Assume that the sufficient HLCP* (2.1) *has a solution and that the set* (2.7) *is nonempty for some* $\tau = \tau_0 > 0$ *and* $\overline{b} \in \mathbb{R}^n$. *Then* (2.5) *has a unique positive solution* $z(\tau, p) = \lceil x(\tau, p),\, s(\tau, p) \rfloor$ *for any* $(\tau, p) \in (0, \tau_0] \times \mathbb{R}^n_{++}$ *and the following holds:*

The function $\overline{z}(q, p) := z(q^{1+\vartheta}, p)$ *is analytic on* $(0, q_0] \times \mathbb{R}^n_{++}$, $0 < q_0 := \tau_0^{1/(1+\vartheta)}$, *and can be analytically extended to an open neighborhood of* $[0, q_0] \times \mathbb{R}^n_{++}$. *For any compact set* $\mathcal{K} \subset \mathbb{R}^n_{++}$ *and any integer* $i \in \mathbb{N}$ *there are constants* $\overline{c}(\mathcal{K}, i)$ *such that*

$$\left\| \frac{\partial^i \overline{z}(q, p)}{\partial q^i} \right\|_2 \leq \overline{c}(\mathcal{K}, i)\, \forall q \in [0, q_0]\, \forall p \in \mathcal{K}\,.$$

The above theorem was first proved in [38] for $\tau_0 = 1$ under the additional assumption that the matrix $[Q\,R]$ has full rank. According to Lemma 2.1 this assumption is automatically satisfied. For any $p \in \mathbb{R}^n_{++}$, $\tau \in (0, \tau_0]$, and $\sigma \in \mathbb{R}$, we consider the function

$$\hat{z}(t, \tau, \sigma, p) := z(t, P(t, \tau, \sigma, p)), \text{ where } P(t, \tau, \sigma, p) := p + \sigma(\tau - t)(e - p). \tag{2.8}$$

Since $P(t, \tau, \sigma, p)$ is obviously analytic in t, τ, σ, and p, it follows from Theorem 2.2 that the function

$$\hat{z}(q^{1+\vartheta}, \tau, \sigma, p), \quad t = q^{1+\vartheta}, \tag{2.9}$$

is a well defined analytic function of q whenever $P(t, \tau, \sigma, p) > 0$, and $\lceil \hat{x},\, \hat{s} \rfloor := \hat{z}(t, \tau, \sigma, p)$ can be computed as the unique positive solution of the following nonlinear system:

$$\begin{aligned} \hat{x}\hat{s} &= q^{1+\vartheta} P(q^{1+\vartheta}, \tau, \sigma, p), \\ Q\hat{x} + R\hat{s} &= b - q^{1+\vartheta}\overline{b}. \end{aligned} \tag{2.10}$$

In this paper we will analyze an interior point method based on the derivatives of (2.9) with respect to q up to an arbitrary order $m \geq 1$. The parameter $\sigma \geq 0$ can be interpreted as a centering parameter: The size of σ determines how much the path $\hat{z}(t, \tau, \sigma, p)$ is tilted towards the central path as t decreases.

3. Infeasible interior point methods based on mth order derivatives.

3.1. The algorithm. By setting $p = e$ in (2.5) we obtain

$$\begin{aligned} xs &= \tau e, \\ Qx + Rs &= b - \tau\overline{b}. \end{aligned} \tag{3.1}$$

The curve $z = z(\tau, e)$ defined by this nonlinear system was called in [4] *the infeasible central path pinned on* $\overline{b}$. Such a path can be associated with any infeasible starting point $z^0 = \lceil x^0,\, s^0 \rfloor \in \mathbb{R}^{2n}_{++}$ by taking

$$\overline{b} := -\frac{r^0}{\tau_0}, \quad r^0 := Qx^0 + Rs^0 - b, \quad \tau_0 := \mu_0 := \mu(z^0) = \frac{x^{0\,T} s^0}{n}. \tag{3.2}$$

If the starting point is feasible, then $\overline{b} = 0$ and the curve $z = z(\tau, e)$ coincides with the classical primal-dual central path whose limit point $z(0, e)$ is the analytic center of the optimal face $\mathcal{F}^*$. If $\overline{b} \neq 0$, then the limit point $z(0, e)$, which exists by virtue of Theorem 2.2, is the so-called *shifted analytic center of the optimal face* (see [4, section 4]).

More generally, given any $\sigma \geq 0$ and any point $z := \lceil x, s \rfloor \in \mathbb{R}^n_{++}$, and setting $\tau := \mu(z) = x^T s/n$, we consider in this paper the weighted infeasible interior point path (see (2.8)) $\hat{z}(t, \tau, \sigma, xs/\tau)$, $t \leq \tau$, starting for $t = \tau$ at $\hat{z}(\tau, \tau, \sigma, xs/\tau) = z$, and given by the solution $\lceil \hat{x}, \hat{s} \rfloor := \hat{z}(q^{1+\vartheta}, \tau, \sigma, xs/\tau)$ of (2.10) for $t = q^{1+\vartheta}$.

The infeasible interior point methods to be presented in the present paper follow this weighted infeasible interior point path by generating a sequence of points belonging to the following large neighborhood of this central path:

$$\mathcal{D}(\beta) = \mathcal{N}^-_\infty(1-\beta) = \{z = \lceil x, s \rfloor \in \mathbb{R}^{2n}_{++} \;:\; xs \geq \beta\mu(z)e\} . \tag{3.3}$$

The initial value for β is β_0. At each iteration we set $\beta_{k+1} = \beta_k - \alpha_k$, where (α_k) is a given monotone sequence of positive numbers satisfying the property

$$\sum_{k=0}^{\infty} \alpha_k \leq \beta_0 - \beta_* , \quad \beta_* > 0 . \tag{3.4}$$

At the start of a typical iteration of our algorithm we have a point (z, τ) such that $z \in \mathcal{D}(\beta) \bigcap \mathcal{F}_{\overline{b},\tau}$ and we consider the mapping (2.8) with $p = xs/\tau$. Since our point $z = \lceil x, s \rfloor$ verifies (2.10) for $p = xs/\tau$ and $t = \tau = \mu(z)$, it follows that $z = \hat{z}(\tau, \tau, \sigma, xs/\tau)$. We wish to approximate $\hat{z}(t, \tau, \sigma, xs/\tau)$ around $t = \tau$ by the mth order Taylor polynomial around $\theta = 0$ of the function $\theta \mapsto \hat{z}\big(\big((1-\theta)\tau\big)^{1/(1+\vartheta)}, \tau, \sigma, xs/\tau\big)$. This Taylor polynomial has the form

$$z(\theta) = \lceil x(\theta), s(\theta) \rfloor = z + \sum_{i=1}^{m} \theta^i w^i, \quad w^i = \lceil u^i, v^i \rfloor, \tag{3.5}$$

where the vectors w^i are given by

$$w^i = \lceil u^i, v^i \rfloor := \begin{cases} \frac{(-1)^i \tau^i}{i!} \frac{\partial^i}{\partial t^i} \hat{z}\, \left(t, \tau, \sigma, \frac{xs}{\tau}\right)_{|t=\tau} & \text{if } \vartheta = 0, \\ \\ \frac{(-1)^i \tau^{i/2}}{i!} \frac{\partial^i}{\partial q^i} \hat{z}\, \left(q^2, \tau, \sigma, \frac{xs}{\tau}\right)_{|q=\sqrt{\tau}} & \text{if } \vartheta = 1. \end{cases} \tag{3.6}$$

It is easily seen that the vectors (3.6) can be obtained by solving the following m systems of linear equations:

$$\begin{aligned} &\begin{cases} su^1 + xv^1 &= (1+\vartheta)\left(\sigma\tau^2 e - (1+\sigma\tau)xs\right), \\ Qu^1 + Rv^1 &= (1+\vartheta)\tau\overline{b}, \end{cases} \\ &\begin{cases} su^2 + xv^2 &= \vartheta xs + (1+4\vartheta)\sigma\tau(xs - \tau e) - u^1 v^1, \\ Qu^2 + Rv^2 &= -\vartheta\tau\overline{b}, \end{cases} \\ &\begin{cases} su^i + xv^i &= \vartheta d^i - \sum_{j=1}^{i-1} u^j v^{i-j}, \\ Qu^i + Rv^i &= 0, \end{cases} \\ &i = 3, 4, \ldots, m\, , \text{where } d^4 = \sigma\tau(xs - \tau e),\ d^3 = -4d^4, \text{ and } d^5 = d^6 = \cdots = 0. \end{aligned} \tag{3.7}$$

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

For $\vartheta = 0$ we consider the most general case of arbitrary $m \geq 1$, but for $\vartheta = 1$ only the case $m \geq 2$, and we denote by $\mathcal{T} := \{(\vartheta, m) \neq (1,1) \mid \vartheta \in \{0,1\}, m \geq 1\}$ the set of these pairs. We deduce that

$$x(\theta)s(\theta) = (1-\theta)^{1+\vartheta} xs - \sigma\tau\theta\varphi_m(\vartheta,\theta)(xs - \tau e) + \sum_{i=m+1}^{2m} \theta^i h^i \,, \tag{3.8}$$

$$\mu(\theta) = (1-\theta)^{1+\vartheta}\mu - \sigma\tau\theta\varphi_m(\vartheta,\theta)(\mu - \tau) + \sum_{i=m+1}^{2m} \theta^i (e^T h^i/n)\,, \tag{3.9}$$

where

$$\begin{aligned} &\varphi_1(0,\theta) = 1, \quad \varphi_m(0,\theta) = 1-\theta, \quad m = 2,3,\ldots, \\ &\varphi_2(1,\theta) = 2 - 5\theta,\ \varphi_3(1,\theta) = 2 - 5\theta + 4\theta^2, \varphi_m(1,\theta) = (2-\theta)(1-\theta)^2,\ m \geq 4, \\ &h^i = \sum_{j=i-m}^{m} u^j v^{i-j},\ i = m+1, m+2, \ldots, 2m\,. \end{aligned} \tag{3.10}$$

For the excluded exceptional case $\vartheta = 1$, $m = 1$ we note only that the structure of the above formulas change:

$$\begin{aligned} x(\theta)s(\theta) &= (1-2\theta)xs - 2\theta\sigma\tau(xs - \tau e) + \theta^2 h^2, \\ \mu(\theta) &= (1-2\theta)\mu - 2\theta\sigma\tau(\mu - \tau) + \theta^2 e^T h^2/n\,. \end{aligned}$$

The quantity μ is a measure of optimality, while, as indicated by (2.7), the quantity τ is a measure of feasibility. The stepsize along the direction w is determined such that two different objectives are achieved.

First, we would like to keep the iterates in a neighborhood that is expanded in a controlled manner at each iteration. Therefore we define

$$\overline{\theta}_1 = \max\left\{\widehat{\theta} \in [0,1] \,:\, z(\theta) \in \mathcal{D}(\beta_+)\, \forall \theta \in \left[0, \widehat{\theta}\right]\right\}\,, \tag{3.11}$$

where $\beta_+ = \beta - \alpha$, and $\alpha > 0$ is a parameter that changes from iteration to iteration (the α's will be chosen from a sequence satisfying (3.4)).

Second, we would like to have $(\tau/\mu)^{\pm 1}$ bounded. We will achieve this by defining

$$\overline{\theta}_2 = \begin{cases} \max\limits_{\widehat{\theta} \leq 1} \left\{\widehat{\theta} \,:\, \gamma^{-\alpha} \geq \frac{\mu(\theta)}{(1-\theta)^{1+\vartheta}\mu} \geq \gamma^{\beta_0 - \beta_+}\ \forall\, 0 \leq \theta \leq \widehat{\theta}\right\} & \text{if } \tau \leq \mu, \\ \max\limits_{\widehat{\theta} \leq 1} \left\{\widehat{\theta} \,:\, \gamma^{\beta_+ - \beta_0} \geq \frac{\mu(\theta)}{(1-\theta)^{1+\vartheta}\mu} \geq \gamma^{\alpha}\ \forall\, 0 \leq \theta \leq \widehat{\theta}\right\} & \text{if } \tau > \mu, \end{cases} \tag{3.12}$$

where γ is some constant satisfying $0 < \gamma < 1$.

Finally, we denote

$$\overline{\theta} = \min\left\{\overline{\theta}_1, \overline{\theta}_2\right\} \tag{3.13}$$

and define the new iterate by

$$(z^+, \tau_+) = (z(\overline{\theta}), (1-\overline{\theta})^{1+\vartheta}\tau). \tag{3.14}$$

Note that in the formula for τ_+ the exceptional case $\vartheta = m = 1$ is excluded. Since $z^+ \in \mathcal{D}(\beta_+) \bigcap \mathcal{F}_{\overline{b},\tau_+}$, the process can be repeated. It follows that all the points produced by the following algorithm will stay in $\mathcal{D}(\beta_*)$.

The computation of the exact value of $\overline{\theta}$ for $m \geq 2$ is complicated since it 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 the proof of global convergence we will give simple lower bounds for $\overline{\theta}$. In order to simplify the presentation in the following algorithm we assume that the exact value of $\overline{\theta}$ is available.

ALGORITHM.

Given parameters $0 < \beta_* < \beta_0 < 1$, $0 < \gamma < 1$, $\vartheta \in \{0, 1\}$, $m \geq 1$, $(\vartheta, m) \neq (1, 1)$, a sequence (α_k) satisfying (3.4), and a starting point $z^0 \in \mathcal{D}(\beta_0)$:

Consider the notation from (3.2);
Set $k \leftarrow 0$;
repeat
Set $z \leftarrow z^k$, $\tau \leftarrow \tau_k$, $\alpha \leftarrow \alpha_k$, $\beta \leftarrow \beta_k$, $\beta_+ \leftarrow \beta - \alpha$;
Choose a scalar $\sigma \in \left[0, \min\left\{1, \gamma^{\beta_0 - \beta_*}/\tau_0\right\}\right]$;
Compute directions $w^i = \lceil u^i, v^i \rfloor$ from (3.7);
Compute steplength $\overline{\theta}$ from (3.11), (3.12), (3.13);
Compute (z^+, τ_+) from (3.14);
Set $\theta_k \leftarrow \overline{\theta}$, $z^{k+1} \leftarrow z^+$, $\tau_{k+1} \leftarrow \tau_+$, $\beta_{k+1} \leftarrow \beta_+$;
Set $k \leftarrow k + 1$.
continue

The algorithm is well defined because of the following.

THEOREM 3.1. *If HLCP is sufficient, then the algorithm is well defined. It generates an iteration sequence satisfying the following properties:*

$$
\begin{aligned}
&z^k = \lceil x^k, s^k \rfloor \in \mathcal{D}(\beta_k) \subset \mathcal{D}(\beta_*) \,;\\
&Qx^k + Rs^k = b + \frac{\tau_k}{\tau_0} r^0 \,;\\
&\tau_{k+1} = (1 - \theta_k)^{1+\vartheta} \tau_k \,;\\
&r^{k+1} = (1 - \theta_k)^{1+\vartheta} r^k \,, \ \textit{where } r^k = Qx^k + Rs^k - b \,;\\
&\tau_k/\gamma \geq \gamma^{\beta_k - \beta_0}\, \tau_k \geq \mu_k \geq \gamma^{\beta_0 - \beta_k}\, \tau_k \geq \gamma\, \tau_k.
\end{aligned}
$$

Proof. We use induction and prove first $r^k = (\tau_k/\tau^0) r^0 = -\tau_k \overline{b}$ for $k = 0, 1, \ldots$. This property is trivially satisfied for $k = 0$. Assume it is satisfied for some $k \geq 0$, and denote $z = z^k$, $r = r^k$, and $\tau = \tau_k$. Since we excluded the case $\vartheta = 1$, $m = 1$, we can write

$$
\begin{aligned}
r(\theta) &:= Qx(\theta) + Rs(\theta) - b\\
&= r + \theta\left(Qu^1 + Rv^1\right) + \theta^2\left(Qu^2 + Rv^2\right) + \sum_{i=3}^{m} \theta^i \left(Qu^i + Rv^i\right)\\
&= -\tau\overline{b} + \theta(1 + \vartheta)\tau\overline{b} - \theta^2\vartheta\tau\overline{b} = -(1 - \theta)^{1+\vartheta}\tau\overline{b} =: -\tau(\theta)\overline{b},
\end{aligned}
$$

which shows that $r^{k+1} = (\tau_{k+1}/\tau^0) r^0 = -\tau_{k+1}\overline{b}$. Also the last property is easily proved by induction. Indeed, since $\mu_0 = \tau_0$, the property is verified for $k = 0$. Let us assume that it is verified for a given k. If $\tau_k \leq \mu_k$, then by the induction hypothesis and (3.12) we have

$$\begin{aligned}&\gamma^{\beta_{k+1}-\beta_0}\tau_{k+1} = \gamma^{\beta_{k+1}-\beta_0}(1-\theta_k)^{1+\vartheta}\tau_k \geq \gamma^{-\alpha_k}(1-\theta_k)^{1+\vartheta}\mu_k \geq \mu_{k+1}\\ &\geq \gamma^{\beta_0-\beta_{k+1}}(1-\theta_k)^{1+\vartheta}\mu_k \geq \gamma^{\beta_0-\beta_{k+1}}(1-\theta_k)^{1+\vartheta}\tau_k = \gamma^{\beta_0-\beta_{k+1}}\tau_{k+1}.\end{aligned}$$

On the other hand, if $\tau_k > \mu_k$, then

$$\begin{aligned}&\gamma^{\beta_{k+1}-\beta_0}\tau_{k+1} = \gamma^{\beta_{k+1}-\beta_0}(1-\theta_k)^{1+\vartheta}\tau_k > \gamma^{\beta_{k+1}-\beta_0}(1-\theta_k)^{1+\vartheta}\mu_k \geq \mu_{k+1}\\ &\geq \gamma^{\alpha_k}(1-\theta_k)^{1+\vartheta}\mu_k \geq \gamma^{\beta_0-\beta_{k+1}}(1-\theta_k)^{1+\vartheta}\tau_k = \gamma^{\beta_0-\beta_{k+1}}\tau_{k+1}.\end{aligned}$$

Here, the first and second "$\geq$" follow by (3.12), and the third follows by the induction hypothesis. ∎

Remark. The algorithm is also well defined for an improper choice of ϑ violating (2.6), i.e., when $\vartheta = 0$ for a degenerate problem. Such an improper choice would not prevent the convergence of the algorithm but only its superlinear convergence (see Theorems 3.8 and 3.10).

The behavior of the iteration sequence (z^k) strongly depends on the choice of the sequence (α_k). Before specifying different choices for (α_k), we need to give some technical results that will be used in the analysis of the algorithm.

3.2. Technical results. We start by finding bounds for the solution of a linear system of the form

$$\left\{\begin{array}{rl} su+xv &= a,\\ Qu+Rv &= 0.\end{array}\right. \tag{3.15}$$

By using the notation

$$D = X^{-1/2}S^{1/2}, \quad \|\lceil u,\, v\rfloor\|_z^2 := \|Du\|_2^2 + \|D^{-1}v\|_2^2, \quad \tilde{a} = (xs)^{-1/2}a\,, \tag{3.16}$$

we obtain the following result.

LEMMA 3.2. *If HLCP is $P_*(\kappa)$, then for any $z = \lceil x,\, s\rfloor \in \mathbb{R}^{2n}_{++}$ and any $a \in \mathbb{R}^n$ the linear system* (3.15) *has a unique solution $w = \lceil u,\, v\rfloor$ for which the following estimate holds:*

$$\|\, w\,\|_z^2 \leq (1+2\kappa)\,\|\,\tilde{a}\,\|_2^2\,.$$

Using this lemma and the fact that the functional $\|\cdot\|_z$ is a norm on $\mathbb{R}^{2n}$ (see [25] for some properties of this norm) we can prove the following lemmas which establish bounds for the solutions of a linear system of the form

$$\left\{\begin{array}{rl} su+xv &= a,\\ Qu+Rv &= \tilde{b}.\end{array}\right. \tag{3.17}$$

LEMMA 3.3. *If HLCP is $P_*(\kappa)$, then for any $z = \lceil x,\, s\rfloor \in \mathbb{R}^{2n}_{++}$ and any $a, \tilde{b} \in \mathbb{R}^n$ the linear system* (3.17) *has a unique solution $w = \lceil u,\, v\rfloor$ and the following inequalities are satisfied:*

$$\|\, w\,\|_z \leq \sqrt{1+2\kappa}\;\|\,\tilde{a}\,\|_2 + (1+\sqrt{2+4\kappa})\,\zeta(z,\tilde{b}\,), \quad \|uv\|_2 \leq \frac{1}{2}\,\|\, w\,\|_z^2\,,$$

where $\tilde{a} = (xs)^{-1/2}a$ and

$$\zeta(z,\tilde{b}\,)^2 = \min\left\{\|\lceil\tilde{u},\, \tilde{v}\rfloor\|_z^2 \;:\; Q\tilde{u} + R\tilde{v} = \tilde{b}\right\} = \tilde{b}^T\left(QD^{-2}Q^T + RD^2R^T\right)^{-1}\tilde{b}\,.$$

Proof. Since (Q, R) is a sufficient pair, so is (QD^{-1}, RD), and therefore, according to Lemma 2.1, the matrix $A = [QD^{-1}\, RD]$ has full rank. It follows that the solution of the minimization problem defining $\zeta(z, \tilde{b}\,)$ is given by

$$\tilde{w} = \lceil\, \tilde{u},\, \tilde{v}\, \rfloor = A^T(AA^T)^{-1}\tilde{b},$$
$$\zeta(z, \tilde{b}\,)^2 = \|\, \tilde{w}\,\|_z^2 = \tilde{b}^T(AA^T)^{-1}\tilde{b} = \tilde{b}^T\left(QD^{-2}Q^T + RD^2R^T\right)^{-1}\tilde{b}\,.$$

If $w = \lceil\, u,\, v\, \rfloor$ is the solution of (3.17), then $\overline{w} = \lceil\, \overline{u},\, \overline{v}\, \rfloor = w - \tilde{w}$ satisfies

$$\begin{cases} s\overline{u} + x\overline{v} &= a - (s\tilde{u} + x\tilde{v}), \\ Q\overline{u} + R\overline{v} &= 0. \end{cases}$$

By applying Lemma 3.2 we deduce that

$$\begin{aligned}
\|\, w\,\|_z &\le \|\, \overline{w}\,\|_z + \|\, \tilde{w}\,\|_z \le \sqrt{1+2\kappa}\, \left\|\, \tilde{a} - (D\tilde{u} + D^{-1}\tilde{v})\, \right\|_2 + \|\, \tilde{w}\,\|_z \\
&\le \sqrt{1+2\kappa}\, \left(\|\, \tilde{a}\,\|_2 + \left\|\, D\tilde{u} + D^{-1}\tilde{v}\, \right\|_2\right) + \|\, \tilde{w}\,\|_z \\
&\le \sqrt{1+2\kappa}\left(\|\, \tilde{a}\,\|_2 + \sqrt{\|\, D\tilde{u}\,\|_2^2 + \|\, D^{-1}\tilde{v}\,\|_2^2 + 2\,\|\, D\tilde{u}\,\|_2\,\|\, D^{-1}\tilde{v}\,\|_2}\right) + \|\, \tilde{w}\,\|_z \\
&\le \sqrt{1+2\kappa}\left(\|\, \tilde{a}\,\|_2 + \sqrt{2}\,\|\, \tilde{w}\,\|_z\right) + \|\, \tilde{w}\,\|_z \\
&= \sqrt{1+2\kappa}\, \left(\|\, \tilde{a}\,\|_2 + \frac{1+\sqrt{2+4\kappa}}{\sqrt{1+2\kappa}}\, \|\, \tilde{w}\,\|_z\right).
\end{aligned}$$

Finally, we have $\|uv\|_2 \le \|Du\|_2\|D^{-1}v\|_2 \le (\|Du\|_2^2 + \|D^{-1}v\|_2^2)/2 = \|w\|^2/2$. ∎

In the remainder of this subsection we will find various bounds on the quantity $\zeta(z, \tilde{b}\,)$ appearing in Lemma 3.3.

LEMMA 3.4. *If HLCP is sufficient, then for any* $z = \lceil\, x,\, s\, \rfloor, z^0 = \lceil\, x^0,\, s^0\, \rfloor \in \mathbb{R}^{2n}_{++}$ *we have*

$$\zeta(z, \tilde{b}\,) \le \left\|\, (xsx^0s^0)^{-1/2}\, \right\|_\infty (x^Ts^0 + s^Tx^0)\zeta(z^0, \tilde{b}\,)\,,$$

where ζ *is defined in Lemma* 3.3.

Proof. Since (Q, R) is a sufficient pair, we deduce as in the proof of Lemma 3.3 that the matrices

$$B = QD^{-2}Q^T + RD^2R^T, \quad B_0 = QD_0^{-2}Q^T + RD_0^2R^T,$$

where

$$D = X^{-1/2}S^{1/2}, \quad D_0 = (X^0)^{-1/2}(S^0)^{1/2},$$

are symmetric positive definite and

$$\zeta(z, \tilde{b}\,)^2 = \tilde{b}^{\,T}B^{-1}\tilde{b} = \left\|\, B^{-1/2}\tilde{b}\, \right\|_2^2, \quad \zeta(z^0, \tilde{b}\,)^2 = \tilde{b}^{\,T}B_0^{-1}\tilde{b} = \left\|\, B_0^{-1/2}\tilde{b}\, \right\|_2^2.$$

Therefore

$$\begin{aligned}
\tilde{b}^{\,T}B^{-1}\tilde{b} &= \tilde{b}^{\,T}B_0^{-1/2}\left(B_0^{-1/2}BB_0^{-1/2}\right)^{-1}B_0^{-1/2}\tilde{b} \\
&\le \left\|\, \left(B_0^{-1/2}BB_0^{-1/2}\right)^{-1}\, \right\|_2 \left\|\, B_0^{-1/2}\tilde{b}\, \right\|_2^2.
\end{aligned}$$

For any $h \in \mathbb{R}^n$ we denote $f = Q^T B_0^{-1/2} h$, $g = R^T B_0^{-1/2} h$ and obtain successively

$$\begin{aligned} &h^T B_0^{-1/2} B B_0^{-1/2} h = f^T D^{-2} f + g^T D^2 g = \left\| D^{-1} f \right\|_2^2 + \| Dg \|_2^2 \\ &= \left\| D^{-1} D_0 D_0^{-1} f \right\|_2^2 + \left\| D D_0^{-1} D_0 g \right\|_2^2 \\ &\geq \left\| D_0^{-1} D \right\|_2^{-2} \left\| D_0^{-1} f \right\|_2^2 + \left\| D_0 D^{-1} \right\|_2^{-2} \| D_0 g \|_2^2 \\ &\geq \min \left\{ \left\| D_0^{-1} D \right\|_2^{-2}, \left\| D_0 D^{-1} \right\|_2^{-2} \right\} \left(\left\| D_0^{-1} f \right\|_2^2 + \| D_0 g \|_2^2 \right) \\ &= \min \left\{ \left\| (s^0 x)^{-1} (x^0 s) \right\|_\infty^{-1}, \left\| (x^0 s)^{-1} (s^0 x) \right\|_\infty^{-1} \right\} \| h \|_2^2 . \end{aligned}$$

It follows that

$$\begin{aligned} &\left\| \left(B_0^{-1/2} B B_0^{-1/2} \right)^{-1} \right\|_2 \leq \max \left\{ \left\| (s^0 x)^{-1} (x^0 s) \right\|_\infty, \left\| (x^0 s)^{-1} (s^0 x) \right\|_\infty \right\} \\ &= \max \left\{ \left\| (x^0 s^0 x s)^{-1} (x^0 s)^2 \right\|_\infty, \left\| (x^0 s^0 x s)^{-1} (s^0 x)^2 \right\|_\infty \right\} \\ &\leq \left\| (x^0 s^0 x s)^{-1} \right\|_\infty \max \left\{ \left\| (x^0 s)^2 \right\|_\infty, \left\| (s^0 x)^2 \right\|_\infty \right\} \\ &= \left(\left\| (x^0 s^0 x s)^{-1/2} \right\|_\infty \max \left\{ \left\| x^0 s \right\|_\infty, \left\| s^0 x \right\|_\infty \right\} \right)^2 \\ &\leq \left(\left\| (x^0 s^0 x s)^{-1/2} \right\|_\infty (x^T s^0 + s^T x^0) \right)^2 . \end{aligned}$$

The proof is complete. ∎

In what follows we show that if HLCP has a solution, then $x^T s^0 + s^T x^0$ is bounded.

LEMMA 3.5. *Assume that HLCP is $P_*(\kappa)$ and that $\mathcal{F}^*$ is nonempty. Let $(x^0, s^0) \in \mathbb{R}^{2n}_{++}$ be a starting point, and consider the notations from* (3.2) *and* (2.7). *Then for all $(x^*, s^*) \in \mathcal{F}^*$ and $z \in \mathcal{F}_{\overline{b},\tau}$, with $0 < \tau < \tau_0$, we have*

$$x^T s^0 + s^T x^0 + \frac{\tau_0 - \tau}{\tau_0} \left(x^T s^* + s^T x^* \right) \leq (1 + 4\kappa) \left(\max\{1, \zeta^*\} + \mu(z)/\tau \right) n \tau_0 ,$$

where $\zeta^* = ((x^0)^T s^* + (s^0)^T x^*)/((x^0)^T s^0)$.

Proof. By denoting $\psi = \tau/\tau_0,\ 0 < \psi < 1$, and using (2.7), (3.2), and $Qx^* + Rs^* = b$ we obtain

$$Q(\psi x^0 + (1-\psi) x^* - x) + R(\psi s^0 + (1-\psi) s^* - s) = \psi(r^0 + b) + (1-\psi) b - (b - \tau \overline{b}) = 0.$$

Using the inequalities $(x^*, s^*) \geq 0,\ (x, s) > 0$ and the $P_*(\kappa)$ property (2.4) we have

$$\begin{aligned} &[\psi x^0 + (1-\psi) x^* - x]^T [\psi s^0 + (1-\psi) s^* - s] \\ &\quad \geq -4\kappa \sum_{i \in \mathcal{I}_+} [\psi x^0 + (1-\psi) x^* - x]_i [\psi s^0 + (1-\psi) s^* - s]_i \\ &\quad \geq -4\kappa \sum_{i \in \mathcal{I}_+} \left(\psi^2 [x^0]_i [s^0]_i + (1-\psi)\psi([x^*]_i [s^0]_i + [x^0]_i [s^*]_i) + [x]_i [s]_i \right) \\ &\quad \geq -4\kappa (\psi^2 (x^0)^T s^0 + (1-\psi)\psi((x^*)^T s^0 + (x^0)^T s^*) + x^T s) , \end{aligned}$$

where

$$\mathcal{I}_+ = \{ i \ :\ [\psi x^0 + (1-\psi) x^* - x]_i [\psi s^0 + (1-\psi) s^* - s]_i > 0 \} .$$

Since $(x^*)^T s^* = 0,\ s^T x^* + x^T s^* \geq 0,\ s^T x^0 + x^T s^0 > 0,\ x^T s = n\mu$, and $0 < \psi < 1$, we deduce that

$$\begin{aligned}
&-4\kappa(\psi^2 n\mu_0 + (1-\psi)\psi((x^*)^T s^0 + (x^0)^T s^*) + n\mu) \\
&\leq [\psi x^0 + (1-\psi)x^* - x]^T[\psi s^0 + (1-\psi)s^* - s] \\
&= \psi^2 n\mu_0 + (1-\psi)\psi((x^0)^T s^* + (s^0)^T x^*) \\
&\quad -\psi(x^T s^0 + s^T x^0) + x^T s - (1-\psi)(s^T x^* + x^T s^*) + (1-\psi)^2(x^*)^T s^* \\
&\leq \psi^2 n\mu_0 + \psi(1-\psi)((x^0)^T s^* + (s^0)^T x^*) - \psi(x^T s^0 + s^T x^0) + n\mu.
\end{aligned}$$

The desired inequality follows by using the fact that $\psi\mu_0 = \psi\tau_0 = \tau$. ◻

We now apply these estimates to bound $\zeta(z, \tau\overline{b})$ for a typical iterate $\lceil x, s \rfloor$, τ found by the algorithm from a starting point $z^0 = \lceil x^0, s^0 \rfloor$, $\tau_0 := \mu(z^0)$. They satisfy for some $\tau < \tau_0$,

$$p := xs \geq \beta\,\mu\, e,\ \ \lceil x, s \rfloor \in \mathcal{F}_{\overline{b},\tau},\ \ \mu = x^T s/n,\ \ \gamma\,\mu \leq \tau \leq \mu/\gamma. \tag{3.18}$$

In order to avoid the introduction of too many constants in the following estimates and to exhibit the role of the parameter m, which determines the order of the algorithm, we use a special $\mathcal{O}_r\,(\ )$ notation (*order-restricted*) and denote by $\pi = \mathcal{O}_r\,(\rho)$ scalar quantities, which may vary with the iterations of the algorithm, $\pi = \pi_k,\ \rho = \rho_k$, for which there is a constant $\delta \geq 0$ *not* depending on m, but eventually depending on the problem instance, so that $|\pi_k| \leq \delta\rho_k$ for all iterations k. We write $\mathcal{O}(\ )$ instead of $\mathcal{O}_r\,(\ )$ if δ is *order-universal*, that is, a constant which is independent of the problem, of m, and of the starting point z_0 (it may still depend on the other constants β_*, β_0, and γ of the algorithm). Hence, e.g., $\beta_*^{\pm 1}$, $\gamma^{\pm 1} = \mathcal{O}(1)$ and $\kappa,\ n = \mathcal{O}_r\,(1)$ are problem dependent constants not depending on m as are the constants

$$\zeta_0^* := \max\{1, \zeta^*\}, \quad \zeta(z^0, r^0)$$

that are connected with the choice of z^0.

An important property is that $q = \mathcal{O}_r\,(p_1 + \cdots + p_j)$, $p_i = \mathcal{O}_r\,(p)$, $j = \mathcal{O}_r\,(1)$, implies $q = \mathcal{O}_r\,(p)$, and that $q = \mathcal{O}(f(m))$ implies $q = \mathcal{O}(1)$ if f is a bounded function of m.

We now derive bounds for $\zeta(z, \tau\overline{b})$ for the case of a general starting point z^0, for the case when z^0 is large enough in the sense that

$$x^0 \geq x^*,\ s^0 \geq s^* \text{ for some } \lceil x^*, s^* \rfloor \in \mathcal{F}^*, \tag{3.19}$$

and for the case when z^0 is almost feasible in the sense that

$$\zeta(z^0, r^0) \leq \frac{1}{\zeta_0^*\,(1+\kappa)\sqrt{n}}. \tag{3.20}$$

We note that while (3.20) is difficult to be verified in practice, it is always satisfied for strictly feasible starting points.

If the starting point is large enough in the sense of (3.19), then we obtain a simple bound by the following.

LEMMA 3.6. *Assume that HLCP is $P_*(\kappa)$ and that $\mathcal{F}^*$ is nonempty. Let $(x^0, s^0) \in \mathbb{R}^{2n}_{++}$ be a starting point satisfying* (3.19)*, and consider the notations from* (3.2) *and* (2.7)*. Then for all z, $0 < \tau < \tau_0$ satisfying* (3.18)*, we have*

$$\zeta(z, \tau\overline{b}) = \mathcal{O}((1+\kappa)\, n\sqrt{\mu})\,. \tag{3.21}$$

Proof. Since $\tilde{w} = \lceil \tilde{u}, \tilde{v} \rfloor = \frac{\tau}{\tau_0} \lceil x^0 - x^*, s^0 - s^* \rfloor$ satisfies $Q\tilde{u} + R\tilde{v} = -\tau\overline{b}$, we have

$$
\begin{aligned}
\zeta(z, \tau\overline{b}\,) &\le \| \tilde{w} \|_z = \frac{\tau}{\tau_0} \left(\left\| s^{1/2} x^{-1/2}(x^0 - x^*) \right\|_2^2 + \left\| x^{1/2} s^{-1/2}(s^0 - s^*) \right\|_2^2 \right)^{1/2} \\
&\le \frac{\tau}{\tau_0} \left(\left\| s^{1/2} x^{-1/2} x^0 \right\|_2^2 + \left\| x^{1/2} s^{-1/2} s^0 \right\|_2^2 \right)^{1/2} \\
&= \frac{\tau}{\tau_0} \left(\left\| (sx)^{-1/2} s x^0 \right\|_2^2 + \left\| (sx)^{-1/2} x s^0 \right\|_2^2 \right)^{1/2} \\
&\le \frac{\tau}{\tau_0} \left\| (sx)^{-1/2} \right\|_\infty \left(\| s x^0 \|_2^2 + \| x s^0 \|_2^2 \right)^{1/2} = \frac{\tau}{\tau_0} \left\| (sx)^{-1/2} \right\|_\infty \| \lceil s x^0, x s^0 \rfloor \|_2 \\
&\le \frac{\tau}{\tau_0} \left\| (sx)^{-1/2} \right\|_\infty \| \lceil s x^0, x s^0 \rfloor \|_1 = \frac{\tau}{\tau_0} \left\| (sx)^{-1/2} \right\|_\infty \left(s^T x^0 + x^T s^0 \right)
\end{aligned}
$$

and by (3.19),

$$
\zeta^* = ((x^0)^T s^* + (s^0)^T x^*)/((x^0)^T s^0) \le ((x^0)^T s^0 + (s^0)^T x^0)/((x^0)^T s^0) = 2.
$$

Therefore, by virtue of Lemma 3.5, we deduce that

$$
\zeta(z, \tau\overline{b}\,) \le \frac{\tau}{\tau_0} \left\| (sx)^{-1/2} \right\|_\infty \left(x^T s^0 + s^T x^0 \right) \le (1 + 4\kappa) \left\| (sx)^{-1/2} \right\|_\infty (2\tau + \mu(z))n
$$

so that by (3.18),

$$
\zeta(z, \tau\overline{b}\,) \le \frac{1 + 4\kappa}{\sqrt{\beta\mu}} \left(\frac{2}{\gamma}\mu + \mu \right) n = \mathcal{O}((1 + \kappa)\, n\sqrt{\mu})\,. \qquad \square
$$

The cases of a general starting point z^0 and of an almost feasible starting point (3.20) are treated in the following.

LEMMA 3.7. *Assume that HLCP is $P_*(\kappa)$ and that $\mathcal{F}^*$ is nonempty. Let $(x^0, s^0) \in \mathbb{R}^{2n}_{++}$ be the starting point of the algorithm and z a typical iterate satisfying* (3.18). *Then the following holds in general:*

$$
\zeta(z, \tau\overline{b}) = \mathcal{O}\big(\zeta_0^*\, \zeta(z^0, r^0)(1 + \kappa) n \sqrt{\mu}\big)\,. \tag{3.22}
$$

If the starting point z^0 satisfies (3.20)*, then*

$$
\zeta(z, \tau\overline{b}) = \mathcal{O}(\sqrt{n\mu})\,. \tag{3.23}
$$

Proof. Formula (3.22) follows from (3.18), Lemmas 3.4 and 3.5, and the estimate

$$
\begin{aligned}
\zeta(z, \tau\overline{b}\,) &= \frac{\tau}{\tau_0}\, \zeta(z, r^0\,) \le \frac{\tau}{\tau_0} \left\| (x s x^0 s^0)^{-1/2} \right\|_\infty (x^T s^0 + s^T x^0) \zeta(z^0, r^0\,) \\
&\le \frac{\tau (x^T s^0 + s^T x^0)}{\tau_0 \beta \sqrt{\mu\mu_0}}\, \zeta(z^0, r^0\,) \le \frac{1 + 4\kappa}{\beta\sqrt{\mu\mu_0}} \left(\max\{1, \zeta^*\}\tau + \mu \right) \zeta(z^0, r^0\,) n \\
&\le \frac{1 + 4\kappa}{\beta_* \sqrt{\mu_0}} \left(\max\{1, \zeta^*\}\gamma^{-1} + 1 \right) \zeta(z^0, r^0\,) n \sqrt{\mu} \\
&= \mathcal{O}\big(\zeta_0^*\, \zeta(z^0, r^0)(1 + \kappa) n \sqrt{\mu}\big)\,.
\end{aligned}
$$

The estimate (3.23) follows directly from the general estimates (3.22) and (3.20). $\square$

To prove the main results of the next subsection we need some estimates concerning the following vectors appearing in (3.7) and (3.17):

$$a = (1+\vartheta)[\sigma\tau^2 e - (1+\sigma\tau)xs\,], \quad \tilde{a} = (xs)^{-1/2}a, \quad \tilde{b} = (1+\vartheta)\tau\bar{b},$$

where $\lceil x, s \rfloor$ is a current iterate of the algorithm satisfying $p := xs \geq \beta\mu e \geq \beta_*\mu e$, $p > 0$, $e^T p = n\mu$, and $\sigma\tau \leq 1$, $\sigma\tau^2 \leq \gamma^{\beta_0-\beta_*}\tau \leq \mu$. Noting that

$$\left\| (xs)^{-1/2} \right\|_\infty \leq \frac{1}{\sqrt{\beta_*\mu}}, \quad \left\| (xs)^{-1/2} \right\|_2 = \frac{\sqrt{n}}{\sqrt{\beta^*\mu}}, \quad \left\| (xs)^{1/2} \right\|_2 = \sqrt{n\,\mu}, \tag{3.24}$$

we obtain

$$\begin{aligned} \|\tilde{a}\|_2 \leq &\quad 2\left\|\sigma\tau^2(xs)^{-1/2}\right\|_2 + 2\left\|(1+\sigma\tau)(xs)^{1/2}\right\|_2 \\ \leq &\quad \tfrac{2}{\sqrt{\beta_*}}\sqrt{n\,\mu} + 4\sqrt{n\,\mu} = \mathcal{O}\big(\sqrt{n\,\mu}\big) = \mathcal{O}_r\left(\sqrt{\mu}\right), \end{aligned} \tag{3.25}$$

$$\left\|\sigma\tau\left((xs)^{1/2} - \tau(xs)^{-1/2}\right)\right\|_2 \leq \quad \sqrt{n\mu}(1+\beta_*^{-1/2}) = \mathcal{O}\big(\sqrt{n\,\mu}\big) = \mathcal{O}_r\left(\sqrt{\mu}\right).$$

Since $a\,b + c\,d \leq \sqrt{a^2+d^2}\sqrt{b^2+c^2}$ holds for all real a, b, c, and d, we get for the solutions $w^i = \lceil u^i, v^i \rfloor$ of (3.7) for all i, $j \geq 1$,

$$\begin{aligned} \left\|u^i v^j + u^j v^i\right\|_2 &\leq \left\|Du^i\right\|_2 \left\|D^{-1}v^j\right\|_2 + \left\|Du^j\right\|_2 \left\|D^{-1}v^i\right\|_2 \\ &\leq \left\| w^i \right\|_z \left\| w^j \right\|_z. \end{aligned} \tag{3.26}$$

3.3. Global convergence. The computational complexity of the algorithm is basically 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 that are independent of ϑ. We will prove the global convergence of the algorithm under general assumptions, but we will give computational complexity results only when the starting point satisfies (3.19) or (3.20). In order to treat the two cases together it is convenient to define

$$\chi = \begin{cases} 1 & \text{if (3.19) is satisfied,} \\ 0 & \text{if (3.20) is satisfied.} \end{cases} \tag{3.27}$$

PROPOSITION 1. *If HLCP is sufficient and solvable, then for any starting point* $z^0 \in \mathcal{D}(\beta)$ *there is a constant* $0 < \Lambda = \mathcal{O}_r(1)$ *such that for all iterations and all* $m \geq 1$ *the vectors* h^i *(see* (3.10)*) produced at each iteration of the algorithm satisfy*

$$\frac{\left\| h^i \right\|_2}{\mu} \leq \frac{\Lambda^i}{i}, \quad i = m+1,\, m+2,\, \ldots,\, 2m.$$

Moreover, there are constants

$$0 < \Lambda_\chi = \mathcal{O}\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} n^{(1+\chi)/2}\Big), \quad \chi = 0, 1, \tag{3.28}$$

such that for any starting point $z^0 \in \mathcal{D}(\beta)$ *satisfying* (3.20) *(*$\chi = 0$*) or* (3.19) *(*$\chi = 1$*) the vectors* h^i *produced at each iteration of the algorithm satisfy (for all* $m \geq 1$*) the following inequalities:*

$$\frac{\left\| h^i \right\|_2}{\mu} \leq \frac{\Lambda_\chi^i}{i}, \quad i = m+1,\, m+2,\, \ldots,\, 2m.$$

Proof. In this proof we employ the $\mathcal{O}(\ )$ and $\mathcal{O}_r(\ )$ estimates of the previous subsection and use repeatedly the estimates (3.24), (3.25), and (3.26). We start by using (3.26) to write

$$\begin{aligned} \| h^i \|_2 &\leq \frac{1}{2} \sum_{j=i-m}^{m} \left(\| Du^j \|_2 \| D^{-1} v^{i-j} \|_2 + \| Du^{i-j} \|_2 \| D^{-1} v^j \|_2 \right) \\ &\leq \frac{1}{2} \sum_{j=i-m}^{m} \| w^j \|_z \| w^{i-j} \|_z \end{aligned} \tag{3.29}$$

and to find constants $\omega_1, \omega_2, \ldots, \omega_m$ such that

$$\| w^i \|_z \leq \omega_i \sqrt{\mu}\,, \quad i = 1, 2, \ldots, m\,. \tag{3.30}$$

From (3.7) and Lemma 3.3 we deduce that

$$\begin{aligned} \| w^1 \|_z &\leq \sqrt{1+2\kappa}\ \| \widetilde{a}^{\,1} \|_2 + (1+\vartheta)\left(1+\sqrt{2+4\kappa}\,\right)\zeta(z,\tau\overline{b})\,, \\ \| w^2 \|_z &\leq \sqrt{1+2\kappa}\ \| \widetilde{a}^{\,2} \|_2 + \vartheta\left(1+\sqrt{2+4\kappa}\,\right)\zeta(z,\tau\overline{b})\,, \\ \| w^i \|_z &\leq \sqrt{1+2\kappa}\ \| \widetilde{a}^{\,i} \|_2\,, \quad i = 3, 4, \ldots, m\,, \end{aligned} \tag{3.31}$$

where

$$\begin{aligned} \widetilde{a}^{\,1} &= -(1+\vartheta)\left((xs)^{1/2} + \sigma\tau\left((xs)^{1/2} - \tau(xs)^{-1/2}\right)\right), \\ \widetilde{a}^{\,2} &= \vartheta(xs)^{1/2} + (1+4\vartheta)\sigma\tau\left((xs)^{1/2} - \tau(xs)^{-1/2}\right) - (xs)^{-1/2}u^1v^1, \\ \widetilde{a}^{\,3} &= -4\vartheta\sigma\tau\left((xs)^{1/2} - \tau(xs)^{-1/2}\right) - (xs)^{-1/2}\left(u^1v^2 + u^2v^1\right), \\ \widetilde{a}^{\,4} &= \vartheta\sigma\tau\left((xs)^{1/2} - \tau(xs)^{-1/2}\right) - (xs)^{-1/2}\left(u^1v^3 + u^2v^2 + u^3v^1\right), \\ \widetilde{a}^{\,i} &= -(xs)^{-1/2}\sum_{j=1}^{i-1} u^j v^{i-j}, \quad i = 5, 6, \ldots, m\,. \end{aligned} \tag{3.32}$$

We first consider the general case of an arbitrary starting point z^0. Then by (3.22),

$$\zeta(t,\tau\overline{b}) = \mathcal{O}_r\left(\sqrt{\mu}\right). \tag{3.33}$$

Using (3.25) and applying repeatedly Lemma 3.3 (L3.3) and (3.24), (3.33), and (3.26) we deduce successively that

$$\| \widetilde{a}^{\,1} \|_2 \overset{(3.25)}{=} \mathcal{O}_r\left(\sqrt{\mu}\right) \overset{(3.33)}{\Rightarrow} \| w^1 \|_z = \mathcal{O}_r\left(\sqrt{\mu}\right) \Rightarrow \| u^1v^1 \|_2 \overset{\text{L3.3}}{\leq} 0.5\,\| w^1 \|_z^2 = \mathcal{O}_r\left(\mu\right)$$
$$\overset{(3.24)}{\Rightarrow} \left\|(xs)^{-1/2}u^1v^1\right\|_2 \leq \|u^1v^1\|_2 / \sqrt{\beta_*\mu} = \mathcal{O}_r\left(\sqrt{\mu}\right) \Rightarrow \|\widetilde{a}^2\|_2 = \mathcal{O}_r\left(\sqrt{\mu}\right) \overset{(3.33)}{\Rightarrow}$$
$$\| w^2 \|_z = \mathcal{O}_r\left(\sqrt{\mu}\right) \overset{(3.24)}{\Rightarrow} \left\|(xs)^{-1/2}(u^1v^2 + u^2v^1)\right\|_2 \leq \| w^1 \|_z \| w^2 \|_z / \sqrt{\beta_*\mu} = \mathcal{O}_r\left(\sqrt{\mu}\right)$$
$$\Rightarrow \|\widetilde{a}^3\|_2 = \mathcal{O}_r\left(\sqrt{\mu}\right) \overset{(3.33)}{\Rightarrow} \| w^3 \|_z = \mathcal{O}_r\left(\sqrt{\mu}\right) \Rightarrow \cdots \Rightarrow \| w^4 \|_z = \mathcal{O}_r\left(\sqrt{\mu}\right).$$

If $m \leq 4$, it follows that (3.30) is satisfied with some constants $\omega_i = \mathcal{O}_r(1)$, $i = 1, \ldots, m$. Assume now that $m \geq 5$. According to (3.24), (3.31), and (3.32), we have

$$\| w^i \|_z \leq \sqrt{1+2\kappa}\,\|\widetilde{a}^i\|_2 \leq \sqrt{\frac{1+2\kappa}{\beta_*\mu}} \sum_{j=i-m}^{m} \| Du^j \|_2 \| D^{-1}v^{i-j} \|_2\,,\ i = 5, \ldots, m\,,$$

and proceeding as in (3.29) we obtain

$$\| w^i \|_z \leq \frac{c}{\sqrt{\mu}} \sum_{j=1}^{i-1} \| w^j \|_z \| w^{i-j} \|_z , \quad c := \sqrt{\frac{1+2\kappa}{4\beta_*}}, \ i = 5, \ldots, m . \tag{3.34}$$

If we define recursively

$$\omega_i := c \sum_{j=1}^{i-1} \omega_j \omega_{i-j}, \quad i = \min\{5, m+1\}, 6, \ldots , \tag{3.35}$$

then by using (3.34) and the fact that (3.30) is satisfied for $i = 1, 2, 3, 4$ we can easily prove by induction that (3.30) holds for $i = 1, 2, \ldots, m$.

We show next that

$$\omega_i \leq \frac{\overline{\alpha}_i}{c} (0.25\Lambda)^i \leq \frac{\Lambda^i}{c\, i}, \quad i = 1, 2, \ldots, 2m, \tag{3.36}$$

where

$$\Lambda := 4 \max \left\{ \left(c\, \omega_i / \overline{\alpha}_i \right)^{1/i} \ : \ i = 1, \ldots, \min\{4, m\} \right\} = \mathcal{O}_r(1)$$

and $(\overline{\alpha}_i)$ is the sequence (considered in [47])

$$\overline{\alpha}_i := \frac{1}{i} \begin{pmatrix} 2i-2 \\ i-1 \end{pmatrix} \leq \frac{1}{i} 4^i ,$$

which satisfies the following recurrence:

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

From construction it follows that (3.36) holds for $i \leq \min\{4, m\}$, and using the recursions of the $\overline{\alpha}_i$ and the ω_i, it is easily proved by induction that it also holds for $\min\{5, m+1\} \leq i \leq 2m$.

By using (3.29), (3.30), and (3.36), we obtain for $i = m+1, \ldots, 2m$,

$$\begin{aligned} \left\| h^i \right\|_2 &\leq \frac{1}{2} \sum_{j=i-m}^{m} \| w^j \|_z \, \| w^{i-j} \|_z \leq \frac{\mu}{2} \sum_{j=i-m}^{m} \omega_j \omega_{i-j} \leq \frac{\mu}{2} \sum_{j=1}^{i-1} \omega_j \omega_{i-j} \\ &\leq \frac{\mu \, (0.25\Lambda)^i}{2c^2} \sum_{j=1}^{i-1} \overline{\alpha}_j \, \overline{\alpha}_{i-j} = \frac{\mu \overline{\alpha}_i \, (0.25\Lambda)^i}{2c^2} \leq \frac{\mu \Lambda^i}{2\, c^2\, i} = \mu \frac{2\beta_*}{1+2\kappa} \frac{\Lambda^i}{i}, \end{aligned} \tag{3.37}$$

which ends the proof of the first part of the proposition.

The more explicit order-universal bounds for starting points satisfying (3.19) or (3.20) are found by essentially the same arguments by just using instead of (3.33) the more precise $\mathcal{O}()$ bounds given in the previous subsection. So if (3.20) is satisfied ($\chi = 0$), we use (3.23), and if (3.19) is satisfied ($\chi = 1$), we use (3.21), that is,

$$\zeta(z, \tau \overline{b}) = \mathcal{O}\Big((1+\kappa)^{\chi} n^{\frac{1+\chi}{2}} \sqrt{\mu} \Big) . \tag{3.38}$$

By using Lemma 3.3 (L3.3), equations (3.25), (3.24), (3.38), and the inequality (3.29) we deduce successively that

$$\begin{aligned}
&\left\| \widetilde{a}^{\,1} \right\|_2 \overset{(3.25)}{=} \mathcal{O}(\sqrt{n\mu}) \overset{(3.38)}{\Rightarrow} \left\| w^1 \right\|_z = \mathcal{O}\Big((1+\kappa)^{.5+\chi} n^{\frac{1+\chi}{2}} \sqrt{\mu}\Big) \\
&\overset{(3.24)}{\Rightarrow} \left\| (xs)^{-1/2} u^1 v^1 \right\|_2 \overset{\text{L3.3}}{\leq} \frac{1}{2\sqrt{\beta_* \mu}} \left\| w^1 \right\|_z^2 = \mathcal{O}\big((1+\kappa)^{1+2\chi} n^{1+\chi} \sqrt{\mu}\big) \\
&\Rightarrow \left\| \widetilde{a}^2 \right\|_2 = \mathcal{O}\big((1+\kappa)^{1+2\chi} n^{1+\chi} \sqrt{\mu}\big) \overset{(3.38)}{\Rightarrow} \left\| w^2 \right\|_z = \mathcal{O}\big((1+\kappa)^{1.5+2\chi} n^{1+\chi} \sqrt{\mu}\big) \\
&\Rightarrow \left\| (xs)^{-1/2} (u^1 v^2 + u^2 v^1) \right\|_2 \overset{(3.29)}{\leq} \frac{1}{\sqrt{\beta_* \mu}} \left\| w^1 \right\|_z \left\| w^2 \right\|_z = \mathcal{O}\Big((1+\kappa)^{2+3\chi} n^{\frac{3+3\chi}{2}} \sqrt{\mu}\Big) \\
&\Rightarrow \left\| \widetilde{a}^3 \right\|_2 = \mathcal{O}\Big((1+\kappa)^{2+3\chi} n^{\frac{3+3\chi}{2}} \sqrt{\mu}\Big) \overset{(3.38)}{\Rightarrow} \left\| w^3 \right\|_z = \mathcal{O}\Big((1+\kappa)^{2.5+3\chi} n^{\frac{3+3\chi}{2}} \sqrt{\mu}\Big) \\
&\Rightarrow \left\| (xs)^{-1/2} (u^1 v^3 + u^2 v^2 + u^3 v^1) \right\|_2 \overset{(3.29)}{\leq} \frac{1}{\sqrt{\beta_* \mu}} (\left\| w^1 \right\|_z \left\| w^3 \right\|_z + .5 \left\| w^2 \right\|_z^2) \\
&= \mathcal{O}\big((1+\kappa)^{3+4\chi} n^{2+2\chi} \sqrt{\mu}\big) \Rightarrow \left\| \widetilde{a}^4 \right\|_2 = \mathcal{O}\big((1+\kappa)^{3+4\chi} n^{2+2\chi} \sqrt{\mu}\big) \\
&\overset{(3.38)}{\Rightarrow} \left\| w^4 \right\|_z = \mathcal{O}\big((1+\kappa)^{3.5+4\chi} n^{2+2\chi} \sqrt{\mu}\big) \,.
\end{aligned}$$

We have thus proved that

$$\left\| w^i \right\|_z \leq \omega_i^{(\chi)} \sqrt{\mu}, \quad \omega_i^{(\chi)} = \mathcal{O}\Big((1+\kappa)^{(1+\chi)i-0.5} n^{(1+\chi)i/2}\Big), \quad 1 \leq i \leq \min\{4, m\}.$$

Therefore we get as in the general case for all $i = m+1, \ldots, 2m$ a bound of the form

$$\frac{\left\| h^i \right\|_2}{\mu} \leq \frac{2\beta_* \Lambda_\chi^i}{(1+2\kappa)i}\,, \tag{3.39}$$

where

$$\Lambda_\chi = 4 \max \left\{ \left(\frac{c\,\omega_i^{(\chi)}}{\overline{\alpha}_i} \right)^{1/i} : i = 1, \ldots, \min\{4, m\} \right\} = \mathcal{O}\Big((1+\kappa)^{1+\chi} n^{(1+\chi)/2}\Big).$$

Now, $\delta := 2\beta^*/(1+2\kappa) = \mathcal{O}_r(1)$ satisfies $\sup_{i \geq m+1} \delta^{1/i} \leq 2(1+\kappa)^{-1/(m+1)} =: \delta^*$. So we may replace in (3.37) and (3.39), δ, $\bar{\Lambda}$, and Λ_χ by 1, 2Λ, and $\delta^* \Lambda_\chi$, respectively. ∎

In the next theorem we will prove the global convergence of the algorithm under the assumption that the sequence (α_k) satisfies the condition

$$\sum_{k=0}^{\infty} \sqrt[m+1]{\alpha_k} = \infty\,. \tag{3.40}$$

We mention that the conditions (3.4) and (3.40) are fairly restrictive, excluding any sequence with $\sup_k \alpha_{k+1}/\alpha_k < 1$. In subsection 3.4 we will give an explicit example of a sequence (α_k) that satisfies both (3.4) and (3.40) plus an additional condition that ensures superlinear convergence.

THEOREM 3.8. *If HLCP is sufficient and solvable and* (3.40) *is satisfied, then for all* $(\vartheta, \theta) \in \mathcal{T}$ *the duality gaps and the residuals of the iteration sequence generated by the algorithm converge to zero, i.e.,*

$$\lim_{k\to\infty} \mu_k = \lim_{k\to\infty} \tau_k = 0\,, \quad \lim_{k\to\infty} r^k = 0\,.$$

Proof. We will first find a lower bound for the steplength $\overline{\theta}$ computed by the algorithm and use repeatedly the following easily verified inequalities:

$$0 \le \varphi_m(\vartheta,\theta) \le 2, \quad 0 \le \theta\varphi_m(\vartheta,\theta) \le \frac{1}{4} \text{ for all } 0 \le \theta \le 1/4,\ (\vartheta, m) \in \mathcal{T}.$$

Similarly, one verifies for all $0 \le \sigma\tau \le 1$ and all $(\vartheta, m) \in \mathcal{T}$ the inequality

$$(1-\theta)^{1+\vartheta} - \sigma\tau\theta\varphi_m(\vartheta,\theta) \ge 0 \tag{3.41}$$

for all $0 \le \theta \le 1/2$.

We note that (3.41) even holds for all $0 \le \theta \le 1$ for all pairs $(\vartheta, m) \notin \{(0,1),\ (1,3)\}$. In addition, we use the following constants and estimates:

$$\gamma_0 := \min\{1, \gamma^{\alpha_0}|\log\gamma|\}, \quad \gamma^{-\alpha} - 1 \ge 1 - \gamma^{\alpha} \ge \alpha\gamma^{\alpha}|\log\gamma| \ge \gamma_0\alpha.$$

We have $z(\theta) \in \mathcal{D}(\beta_+)$ if and only if $x(\theta)s(\theta) - (\beta-\alpha)\mu(\theta)e \ge 0$. Using (3.8) and (3.9) and the fact that $z = z(0) \in \mathcal{D}(\beta)$ and assuming $\theta \in [0, 0.25]$ we get by (3.41),

$$\begin{aligned}
&x(\theta)s(\theta) - (\beta-\alpha)\mu(\theta)e \\
&= \left((1-\theta)^{1+\vartheta} - \sigma\tau\theta\varphi_m(\vartheta,\theta)\right) xs + \sum_{i=m+1}^{2m} \theta^i h^i + (1-\beta+\alpha)\sigma\tau^2\theta\varphi_m(\vartheta,\theta)e \\
&\quad -(\beta-\alpha)\left(\left((1-\theta)^{1+\vartheta} - \sigma\tau\theta\varphi_m(\vartheta,\theta)\right)\mu e + \sum_{i=m+1}^{2m} \theta^i \left(\frac{e^T h^i}{n}\right)\right) \\
&\ge \beta\left((1-\theta)^{1+\vartheta} - \sigma\tau\theta\varphi_m(\vartheta,\theta)\right)\mu e + \sum_{i=m+1}^{2m} \theta^i h^i + (1-\beta+\alpha)\sigma\tau^2\theta\varphi_m(\vartheta,\theta)e \\
&\quad -(\beta-\alpha)\left(\left((1-\theta)^{1+\vartheta} - \sigma\tau\theta\varphi_m(\vartheta,\theta)\right)\mu e + \sum_{i=m+1}^{2m} \theta^i \left(\frac{e^T h^i}{n}\right)\right) \\
&= \alpha\left((1-\theta)^{1+\vartheta} - \sigma\tau\theta\varphi_m(\vartheta,\theta)\right)\mu e + (1-\beta+\alpha)\sigma\tau^2\theta\varphi_m(\vartheta,\theta)e \\
&\quad + \sum_{i=m+1}^{2m} \theta^i \left(h^i - (\beta-\alpha)\frac{e^T h^i}{n}\, e\right) \ge \alpha\mu\psi_m(\alpha,\vartheta,\theta)\, e,
\end{aligned}$$

where

$$\begin{aligned}
\psi_m(\alpha,\vartheta,\theta) &:= (1-\theta)^{1+\vartheta} + \frac{(1-\beta)\tau + \alpha(\tau-\mu)}{\alpha\mu}\sigma\tau\theta\varphi_m(\vartheta,\theta) \\
&\quad - \frac{1}{\alpha\gamma_0\mu}\sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2 .
\end{aligned} \tag{3.42}$$

In the last inequality above and the definition of ψ_m, we used $0 < \gamma_0 \le 1$ and

$$-1 + \frac{(1-\beta+\alpha)\tau}{\alpha\mu} = \frac{(1-\beta)\tau + \alpha(\tau-\mu)}{\alpha\mu}, \quad h^i - \frac{\beta-\alpha}{n} e^T h^i\, e \ge -\left\| h^i \right\|_2 e,$$
$$\left\| h - \delta e^T h\, e \right\|_2^2 = \| h \|_2^2 - (e^T h)^2 \delta(2 - n\delta) \le \| h \|_2^2 \text{ if } 0 \le \delta n \le 2.$$

With the above notation it follows that

$$z \in \mathcal{D}(\beta),\ \psi_m(\alpha,\vartheta,\theta) \ge 0, \text{ and (3.41)} \quad \Longrightarrow \quad z(\theta) \in \mathcal{D}(\beta-\alpha)\,. \tag{3.43}$$

We also note that

$$-1 \le \frac{1-\beta+\alpha}{\alpha}\,\gamma^{\beta_0-\beta} - 1 \le \frac{(1-\beta)\tau + \alpha(\tau-\mu)}{\alpha\mu} \le \frac{1-\beta+\alpha}{\alpha}\,\gamma^{\beta-\beta_0} - 1,$$

and by using $\sigma\tau \le \sigma\tau_0 \le \gamma^{\beta_0-\beta_*}$, $\tau \ge \gamma^{\beta_0-\beta_*}\mu$, we deduce

$$\left(\frac{(1-\beta)\tau + \alpha(\tau-\mu)}{\alpha\mu}\right)\sigma\tau > \frac{\sigma\tau(\tau-\mu)}{\mu} \ge \left(\gamma^{\beta_0-\beta_*} - 1\right)\gamma^{\beta_0-\beta_*} \ge -\frac{1}{4}\,. \tag{3.44}$$

From Proposition 1 it follows that

$$\frac{1}{\mu}\sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2 \le \sum_{i=m+1}^{2m} \frac{1}{i}\,(\theta\Lambda)^i\,, \quad 0 \le \theta \le 1.$$

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

$$\sum_{i=m+1}^{2m} \frac{t^i}{i} \le t^{m+1} \sum_{i=m+1}^{2m} \frac{1}{i} < t^{m+1} \int_m^{2m} \frac{du}{u} = t^{m+1}\log 2 < 0.7t^{m+1}\,.$$

Hence for arbitrary $0 \le \theta \le \breve{\theta} \le \Lambda^{-1}$,

$$\frac{1}{\mu}\sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2 \le 0.7(\Lambda\breve{\theta})^{m+1}\,. \tag{3.45}$$

Since $0 \le \sigma\tau\theta\varphi_m(\vartheta,\theta) \le 0.25$ for all $0 \le \theta \le 0.25$, we get the estimate

$$\psi_m(\alpha,\vartheta,\theta) \ge (1-\breve{\theta})^{1+\vartheta} - \frac{1}{16} - \frac{0.7(\Lambda\breve{\theta})^{m+1}}{\alpha\gamma_0} \ge \frac{1}{2} - \frac{0.7(\Lambda\breve{\theta})^{m+1}}{\alpha\gamma_0}$$

for all $0 \le \theta \le \breve{\theta} \le \min\{0.25, \Lambda^{-1}\}$. Therefore $\psi_m(\alpha,\vartheta,\theta) \ge 0$ for all $\theta \le \breve{\theta}_1$, where

$$\breve{\theta}_1 := \min\left\{\frac{1}{4},\ \frac{1}{\Lambda},\ \frac{1}{\Lambda}\left(\frac{\gamma_0\alpha}{1.4}\right)^{\frac{1}{m+1}}\right\}. \tag{3.46}$$

Hence the quantity $\overline{\theta}_1$ defined by (3.11) satisfies $\overline{\theta}_1 \ge \breve{\theta}_1$.

Also we have

$$\begin{aligned}
&\mu(\theta) - \gamma^{\beta_0-\beta_+}(1-\theta)^{1+\vartheta}\mu \ge \mu(\theta) - \gamma^{\alpha_0}(1-\theta)^{1+\vartheta}\mu \ge \mu(\theta) - \gamma^{\alpha}(1-\theta)^{1+\vartheta}\mu \\
&= (1-\theta)^{1+\vartheta}\left(1-\gamma^{\alpha}\right)\mu + \sigma\tau(\tau-\mu)\theta\varphi_m(\vartheta,\theta) + \sum_{i=m+1}^{2m} \theta^i(e^T h^i/n) \\
&\ge \mu\gamma_0\alpha\,\breve{\psi}_m(\alpha,\vartheta,\theta),
\end{aligned}$$

where

$$\breve{\psi}_m(\alpha,\vartheta,\theta) = (1-\theta)^{1+\vartheta} + \frac{\sigma\tau(\tau-\mu)}{\mu\gamma_0\alpha}\theta\varphi_m(\vartheta,\theta) - \frac{1}{\gamma_0\alpha\mu}\sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2\,, \tag{3.47}$$

and

$$\mu(\theta) - \gamma^{\beta_+ - \beta_0}(1-\theta)^{1+\vartheta}\mu \le \mu(\theta) - \gamma^{-\alpha_0}(1-\theta)^{1+\vartheta}\mu \le \mu(\theta) - \gamma^{-\alpha}(1-\theta)^{1+\vartheta}\mu$$

$$= (1-\theta)^{1+\vartheta}\left(1-\gamma^{-\alpha}\right)\mu + \sigma\tau(\tau-\mu)\theta\varphi_m(\vartheta,\theta) + \sum_{i=m+1}^{2m} \theta^i (e^T h^i/n)$$

$$\le -\mu\gamma_0\alpha\,\widehat{\psi}_m(\alpha,\vartheta,\theta),$$

where

$$\widehat{\psi}_m(\alpha,\vartheta,\theta) = (1-\theta)^{1+\vartheta} - \frac{\sigma\tau(\tau-\mu)}{\mu\gamma_0\alpha}\theta\varphi_m(\vartheta,\theta) - \frac{1}{\gamma_0\alpha\mu}\sum_{i=m+1}^{2m}\theta^i \left\| h^i \right\|_2. \tag{3.48}$$

Therefore the following implications hold:

$$\breve{\psi}_m(\alpha_0,\vartheta,\theta) \ge 0 \text{ and } \widehat{\psi}_m(\alpha,\vartheta,\theta) \ge 0 \implies \gamma^{-\alpha} \ge \frac{\mu(\theta)}{(1-\theta)^{1+\vartheta}\mu} \ge \gamma^{\beta_0-\beta_+},$$

$$\breve{\psi}_m(\alpha,\vartheta,\theta) \ge 0 \text{ and } \widehat{\psi}_m(\alpha_0,\vartheta,\theta) \ge 0 \implies \gamma^{\beta_+-\beta_0} \ge \frac{\mu(\theta)}{(1-\theta)^{1+\vartheta}\mu} \ge \gamma^{\alpha}.$$

We will prove now that the quantity defined in (3.12) satisfies

$$\overline{\theta}_2 \ge \breve{\theta}_2 := \min\left\{\frac{\gamma_0\alpha_0}{4},\ \frac{1}{\Lambda},\ \frac{1}{\Lambda}\left(\frac{\gamma_0\alpha}{11.2}\right)^{\frac{1}{m+1}}\right\}. \tag{3.49}$$

For the remainder of this proof we assume that $0 \le \theta \le \breve{\theta}_2$. This implies $0 \le \theta \le 1/4$. If $\tau \le \mu$, then $0 \ge \sigma\tau(\tau-\mu)/\mu \ge -1/4$. Hence, by using $\varphi_m(\vartheta,\theta) \le 2$ and (3.45),

$$\breve{\psi}_m(\alpha_0,\vartheta,\theta) \ge (1-\theta)^{1+\vartheta} - \frac{\theta}{2\gamma_0\alpha_0} - \frac{1}{\mu\gamma_0\alpha_0}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2$$

$$\ge \frac{9}{16} - \frac{1}{8} - \frac{0.7(\Lambda\breve{\theta}_2)^{m+1}}{\gamma_0\alpha_0} > \frac{9}{16} - \frac{1}{8} - \frac{0.7}{11.2} = \frac{6}{16} > 0,$$

since $\alpha \le \alpha_0$ and $(\Lambda\breve{\theta}_2)^{m+1} = \gamma_0\alpha/11.2$. Moreover,

$$\widehat{\psi}_m(\alpha,\vartheta,\theta) \ge \left(1-\breve{\theta}_2\right)^{1+\vartheta} - \frac{1}{\mu\gamma_0\alpha}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 \ge \frac{9}{16} - \frac{0.7}{11.2} = \frac{1}{2} > 0.$$

If $\tau > \mu$, then

$$0 \le \sigma\tau\frac{(\tau-\mu)}{\mu} \le \sigma\tau_0\left(\frac{\tau}{\mu}-1\right) \le \gamma^{\beta_0-\beta_*}(\gamma^{\beta_*-\beta_0}-1) < 1$$

so that we find, as before, for all $\theta \le \breve{\theta}_2$,

$$\breve{\psi}_m(\alpha,\vartheta,\theta) \ge (1-\theta)^{1+\vartheta} - \frac{1}{\mu\gamma_0\alpha}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 > 0,$$

$$\widehat{\psi}_m(\alpha_0,\vartheta,\theta) \ge (1-\theta)^{1+\vartheta} - \frac{2\theta}{\alpha_0\gamma_0} - \frac{1}{\mu\gamma_0\alpha_0}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2$$

$$\ge \frac{9}{16} - \frac{8}{16} - \frac{1}{\mu\gamma_0\alpha_0}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 \ge 0.$$

This shows that (3.49) holds.

Since $\theta_1 \geq \breve{\theta}_1$, the step $\overline{\theta}$ computed by the algorithm satisfies $\overline{\theta} \geq \breve{\theta}_2'$ for all $\breve{\theta}_2' \leq \breve{\theta}_2$. To get a simple lower bound $\breve{\theta}_2'$ we use that Proposition 1 remains true if Λ is increased to some $\Lambda' = \mathcal{O}_r(1)$. As

$$\sup_m \left(\frac{\gamma_0}{11.2}\right)^{1/(m+1)} = 1,$$

it follows that $\Lambda' := \max(4/(\alpha_0\gamma_0), \Lambda)$ satisfies $\Lambda' = \mathcal{O}_r(1)$, $\Lambda' \geq \Lambda$, and

$$\frac{1}{\Lambda'}\left(\frac{\gamma_0\alpha}{11.2}\right)^{1/(m+1)} \leq \frac{\alpha^{1/(m+1)}}{\Lambda'} \leq \alpha^{1/(m+1)}\frac{\alpha_0\gamma_0}{4} \leq \frac{\alpha_0\gamma_0}{4}.$$

Hence, replacing Λ and $\hat{\theta}_2$ by Λ' and $\hat{\theta}_2'$, respectively, the new $\Lambda = \mathcal{O}_r(1)$ satisfies

$$\overline{\theta}_2 \geq \breve{\theta}_2 = \frac{1}{\Lambda}\left(\frac{\gamma_0\alpha}{11.2}\right)^{1/(m+1)}, \quad \Lambda \geq \frac{4}{\alpha_0\gamma_0}, \tag{3.50}$$

so that $\overline{\theta} \geq \breve{\theta}_2$. The algorithm therefore generates for all $k \geq 0$ steplengths θ_k satisfying

$$\theta_k \geq \breve{\omega} \sqrt[m+1]{\alpha_k}, \text{ where } \breve{\omega} := \frac{1}{\Lambda}\left(\frac{\gamma_0}{11.2}\right)^{\frac{1}{m+1}}. \tag{3.51}$$

Note that $\breve{\omega}^{-1}$ is a bounded function of m,

$$0 < \breve{\omega}^{-1} = \Lambda\left(\frac{11.2}{\gamma_0}\right)^{1/(m+1)} = \mathcal{O}_r(1)\,. \tag{3.52}$$

Using Theorem 3.1 we have

$$\frac{\tau_k}{\tau_0} = \prod_{j=0}^{k-1}(1-\theta_j)^{1+\vartheta} \leq \prod_{j=0}^{k-1}\left(1-\breve{\omega}\alpha_j^{1/(m+1)}\right)^{1+\vartheta}, \tag{3.53}$$

which by virtue of (3.40) shows that $\lim_{k\to\infty}\tau_k = 0$. The proof is complete by noticing that according to Theorem 3.1 we have $r^k = \frac{\tau_k}{\tau_0}r^0$ and $\mu_k \leq \tau_k/\gamma$. □

3.4. Higher order convergence. To investigate the superlinear convergence of the algorithm we need the further assumption that the sequence (3.40) diverges slow enough in the sense that

$$\lim_{k\to\infty}\frac{\log\alpha_k}{\sum_{i=0}^{k-1}\sqrt[m+1]{\alpha_i}} = 0\,. \tag{3.54}$$

Then, as we will see, the algorithm (with $\vartheta \in \{0,1\}$ satisfying (2.6), $m \geq 1$, and $(\vartheta, m) \neq (1,1)$) 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.

In subsection 3.5 a sequence (α_k) satisfying all requirements (3.4), (3.40), and (3.54) will be given explicitly.

The proof of superlinear convergence is based on Theorem 2.2. In order to be able to apply this theorem we need to show that the range of the function $t \mapsto \hat{z}(t,\tau,\sigma,xs/\tau) = z(t, P(t,\tau,\sigma,xs/\tau))$, $t \in (0,\tau]$, given by (2.8) lies in a compact set of $\mathbb{R}^n_{++}$. This follows from the fact that for any $z = \lceil x,\, s \rfloor \in \mathcal{D}(\beta_*)$ and $p = xs/\tau$, with $\gamma\mu \leq \tau \leq \mu/\gamma$, $0 \leq \sigma \leq \min\{1,\ \gamma^{\beta_0-\beta_*}/\tau_0\}$, and $t \in [0,\tau]$, we have, with $\lambda = \sigma\,(\tau - t)$,

$$P(t,\tau,\sigma,p) = (1-\lambda)\,p + \lambda e \geq (1-\lambda)\frac{\tau}{\mu}\,\beta_* e + \lambda\beta_*\gamma^{\beta_0-\beta_*} e \geq \beta_*\gamma^{\beta_0-\beta_*} e\,,$$
$$\|\,P(t,\tau,\sigma,p)\,\|_1 = (1-\lambda)\,\frac{n\mu}{\tau} + \lambda n \leq \gamma^{\beta_*-\beta_0} n\,.$$

By using (3.6) and Theorems 2.2 and 3.1 we obtain the following result.

LEMMA 3.9. *If HLCP is sufficient and solvable, then there is a constant c such that the vectors* $u^1, v^1, \ldots, u^m, v^m$ *computed at each iteration of the algorithm satisfy*

$$\left\|\,u^i\,\right\|_2 \leq c\,\tau^{\frac{i}{1+\vartheta}}\,, \quad \left\|\,v^i\,\right\|_2 \leq c\,\tau^{\frac{i}{1+\vartheta}}\,, \quad i = 1, 2, \ldots, m\,.$$

This lemma is crucial for proving the following theorem on the superlinear convergence of the algorithm.

THEOREM 3.10. *Let* $\vartheta \in \{0,1\}$ *be defined by* (2.6), *and let* $m \geq 1$ *such that* $(\vartheta, m) \neq (1,1)$. *If HLCP is sufficient and solvable and* (α_k) *is a monotone sequence of positive numbers satisfying* (3.4), (3.54), and (3.40), *then the sequence* $\left(z^k\right)$ *produced by the algorithm converges Q-superlinearly to a maximal complementarity solution* $z^* \in \mathcal{F}^c$. *Also the sequences of the corresponding complementarity gaps* (μ_k), *feasibility measures* (τ_k), *and residuals* $\left(r^k\right)$ *converge Q-superlinearly to zero. Moreover, the Q-orders of convergence of these sequences satisfy*

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

Proof. Since we are analyzing asymptotic properties, we may assume $\mu_k < 1$ and $\tau_k < 1$. By using $\alpha_k < 1$, Theorems 3.1 and 3.8, and (3.54) and (3.53) we obtain

$$0 \leq \frac{\log \alpha_k}{\log \tau_k} = \frac{\log \alpha_k}{\log(\tau_k/\tau_0) + \log \tau_0} \leq \frac{\log \alpha_k}{-(1+\vartheta)\,\breve{\omega}\sum_{j=0}^{k-1}\alpha_j^{1/(m+1)} + \log \tau_0} \to 0.$$

Hence there is a sequence (ε_k) so that

$$\alpha_k = \tau_k^{\varepsilon_k}, \quad \varepsilon_k \downarrow 0. \tag{3.55}$$

From (3.10) and Lemma 3.9 it follows for $i = m+1,\, m+2,\, \ldots,\, 2m$ that

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

For k sufficiently large we have $\tau^{\frac{1}{1+\vartheta}} \leq 1/2$ so that for all $\theta \in [0,1]$,

$$\sum_{i=m+1}^{2m} \theta^i \left\|\,h^i\,\right\|_2 \leq \sum_{i=m+1}^{2m} \left\|\,h^i\,\right\|_2 \leq mc^2\tau^{\frac{m+1}{1+\vartheta}} \sum_{i=0}^{m-1} \tau^{\frac{i}{1+\vartheta}} \leq 2mc^2\tau^{\frac{m+1}{1+\vartheta}}.$$

Hence, by $\mu \geq \gamma\tau$,

$$\frac{1}{\alpha\gamma_0\mu} \sum_{i=m+1}^{2m} \theta^i \left\|\,h^i\,\right\|_2 \leq \frac{2mc^2}{\alpha\gamma_0\gamma}\tau^{\frac{m-1}{1+\vartheta}} \leq \frac{\widetilde{c}}{\alpha}\tau^{\frac{m-1}{1+\vartheta}} \tag{3.56}$$

for any $\widetilde{c} \geq 2mc^2/(\gamma_0\gamma)$. A suitable $\widetilde{c}$ will be determined later. In what follows we will show that if τ is sufficiently small, then the stepsize $\overline{\theta} = \min\{\overline{\theta}_1, \overline{\theta}_2\}$ taken by the algorithm (determined by (3.11) and (3.12)) satisfies

$$\overline{\theta} \geq \hat{\theta} := 1 - \left(\frac{2\widetilde{c}}{\alpha}\tau^{\frac{m-1}{1+\vartheta}}\right)^{\frac{1}{1+\vartheta}}. \tag{3.57}$$

We now follow the reasoning as in the proof of Theorem 3.8. To prove (3.57) it is sufficient to show that for all $(\vartheta, m) \in \mathcal{T}$ and small τ (or μ),

(a) (3.41) holds for $0 \leq \theta \leq \hat{\theta}$;

(b) ψ_m (see (3.42)), $\breve{\psi}_m$ (see(3.47)), and $\widetilde{\psi}_m$ (see (3.48)) are nonnegative for $0 \leq \theta \leq \hat{\theta}$.

As to (a) (see the note following (3.41)), we have only to check cases $(\vartheta, m) = (0, 1)$ and $(\vartheta, m) = (1, 3)$, and we use (3.10) for explicit formulas for φ_m:

If $(\vartheta, m) = (0, 1)$, (3.41) follows from $\varphi_1(0, \theta) = 1$, $\sigma \leq 1$, $\tau \leq \mu/\gamma$, since for $\theta \leq \hat{\theta} < 1$,

$$(1-\theta) - \sigma\tau\theta \geq \frac{2\tilde{c}}{\alpha\gamma_0}\tau - \tau \geq \left(\frac{2\tilde{c}}{\alpha\gamma_0} - 1\right)\tau \geq 0 \quad \text{for small } \alpha > 0\,.$$

If $(\vartheta, m) = (1, 3)$, (3.41) follows similarly from $\varphi_3(1, \theta) = 2 - 5\theta + 4\theta^2 \leq 2$ since

$$(1-\theta)^2 - \sigma\tau\theta\varphi_3(1, \theta) \geq \frac{2\tilde{c}}{\alpha}\tau - 2\tau \geq 0 \quad \text{for small } \alpha\,.$$

As to (b), we first derive a common lower bound for ψ_m, $\breve{\psi}_m$, and $\widetilde{\psi}_m$, which is shown to be nonnegative for all $0 \leq \theta \leq \hat{\theta}$ if α or τ are small.

Using that $0 \leq \sigma \leq 1$, $0 \leq \theta \leq 1$, and $\gamma\tau \leq \mu \leq \tau/\gamma$ one can find a constant $\overline{c} > 0$ such that the middle terms in ψ_m, $\breve{\psi}_m$, and $\widetilde{\psi}_m$ are bounded below by

$$-\frac{\overline{c}}{\alpha}\tau|\varphi_m(\vartheta, \theta)|.$$

Therefore, by (3.56), a common lower bound *lbd* for ψ_m, $\breve{\psi}_m$, and $\widetilde{\psi}_m$ that is valid for all $0 \leq \theta \leq 1$ is given by

$$lbd := (1-\theta)^{1+\vartheta} - \frac{\overline{c}}{\alpha}\tau|\varphi_m(\vartheta, \theta)| - \frac{\tilde{c}}{\alpha}\tau^{\frac{m-\vartheta}{1+\vartheta}}, \tag{3.58}$$

where $\tilde{c} := \max(2mc^2/(\gamma_0\gamma), 2\overline{c})$.

Now, it can be shown that *lbd* is nonnegative for all $(\vartheta, m) \in \mathcal{T}$ and all $0 \leq \theta \leq \hat{\theta}$, provided that α and τ are small.

Since the proofs are similar for each $(\vartheta, m) \in \mathcal{T}$, we use (3.10): and give the typical arguments only for the cases $(\vartheta, m) = (0, 1)$ and $(\vartheta, m) = (1, m)$ with $m \geq 4$:

$\vartheta = 0$, $m = 1$, $\varphi_1(0, \theta) = 1$: Then, for $0 \leq \theta \leq \hat{\theta}$, *lbd* is bounded below by

$$\frac{2\tilde{c}}{\alpha}\tau - \frac{2\overline{c}}{\alpha}\tau - \frac{\tilde{c}}{\alpha}\tau \geq 0 \quad (\text{by } \tilde{c} \geq 2\overline{c}).$$

$\vartheta = 1$, $m \geq 4$: Then $|\varphi_m(1, \theta)| = |2 - \theta|(1-\theta)^2 \leq 2(1-\theta)^2$, and using $\alpha = \tau^\varepsilon$ (see (3.55)), *lbd* is bounded below for all $0 \leq \theta \leq \hat{\theta}$ by

$$\begin{aligned} lbd &\geq (1-\theta)^2\left(1 - \frac{2\overline{c}}{\alpha}\tau\right) - \frac{\tilde{c}}{\alpha}\tau^{\frac{m-1}{2}} \\ &\geq 2\tilde{c}\tau^{\frac{m-1}{2}-\varepsilon}(1 - 2\overline{c}\tau^{1-\varepsilon}) - \tilde{c}\tau^{\frac{m-1}{2}-\varepsilon} \geq 0 \end{aligned}$$

if τ is small.

Thus for sufficiently large k (i.e., $\alpha = \alpha_k$ and $\tau = \tau_k$ small) the steplength $\overline{\theta} = \theta_k$ used by the algorithm satisfies $\theta_k \geq \hat{\theta}$, which implies

$$\tau_{k+1} \leq (1 - \hat{\theta}\,)^{1+\vartheta} \tau_k \leq \frac{2\widetilde{c}}{\alpha_k} \tau_k^{\frac{m-1}{1+\vartheta}} \tau_k \leq \frac{2\widetilde{c}}{\alpha_k} \tau_k^{\frac{m+1}{1+\vartheta}} \leq 2\widetilde{c} \tau_k^{\frac{m+1}{1+\vartheta} - \varepsilon_k}.$$

Therefore

$$\mathcal{Q}\,(\tau_k) = \liminf \frac{\log \tau_{k+1}}{\log \tau_k} \geq \frac{m+1}{1+\vartheta}\,,$$

and by using $\gamma\tau_k \leq \mu_k \leq \tau_k/\gamma$ it follows that $\mathcal{Q}\,(\mu_k) = \mathcal{Q}\,(\tau_k)$. For k sufficiently large we have $(c\tau)^{\frac{1}{1+\vartheta}} < 1 - 1/\sqrt{2}$ so that by using (3.5) and Lemma 3.9 we obtain

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

and by applying Theorem 2 of [23] we deduce the convergence of the sequence $\left(z^k\right)$ to a maximal complementarity solution $z^* \in \mathcal{F}^*$ and the fact that $\mathcal{Q}\left(z^k\right) = \mathcal{Q}\,(\mu_k)$. ∎

3.5. Polynomial complexity. In Theorem 3.8 we have established global convergence for any sequence (α_k) satisfying (3.4) and (3.40). We will obtain polynomial complexity for starting points satisfying (3.20) or (3.19) by using the following sequence that was first considered in [27]:

$$\alpha_k = \frac{\nu(\beta_0 - \beta_*)}{(\exp(1) + k + 1) \log^{1+\nu}(\exp(1) + k + 1)}, \;\; k = 0, 1, \ldots\,, \tag{3.59}$$

where $0 < \nu \leq 1$ is a given parameter. From [27] it follows that this sequence satisfies (3.4) and

$$\sum_{k=0}^{K-1} \alpha_k^{\frac{1}{m+1}} > (\nu(\beta_0 - \beta_*))^{\frac{1}{m+1}} (\log K)^{-\frac{(1+\nu)\,m}{m+1}} K^{\frac{m}{m+1}}\,. \tag{3.60}$$

Using [27, Corollary 3.5] we obtain the following result.

LEMMA 3.11. *If* $M_* := (\nu(\beta_0 - \beta_*))^{\frac{-1}{m+1}} M > \exp(1)$ *and*

$$K \geq \left(2\left(1 + \frac{1+\nu}{m+1}\log\frac{1+\nu}{m}\right)\left(\frac{m+1}{m}\log M_*\right)\right)^{\frac{1+\nu}{m}} M_*^{\frac{m+1}{m}}\,,$$

then $\sum_{k=0}^{K-1} \alpha_k^{\frac{1}{m+1}} > M$.

In the remainder of this subsection we give explicit upper bounds for the number of iterations required by the algorithm to obtain a solution of the HLCP with prescribed accuracy in case the starting point satisfies (3.19) or (3.20). More precisely given any $\epsilon > 0$ we have to find an upper bound for the number

$$K_\epsilon := \min\left\{K \;:\; \max\left\{x^{k\,T} s^k, \left\|\, r^k \,\right\|_2\right\} \leq \epsilon \;\forall k \geq K\,\right\}\,. \tag{3.61}$$

The upper bound will depend on n, κ, m,

$$L_\epsilon := \log\max\left\{\frac{x^{0\,T} s^0}{\epsilon}, \frac{\left\|\, r^0 \,\right\|_2}{\epsilon}\right\}\,, \tag{3.62}$$

and

$$\eta_\chi := \frac{2\Lambda_\chi}{1+\vartheta}\left(\frac{11.2}{\nu\gamma_0(\beta_0-\beta_*)}\right)^{\frac{1}{m+1}},$$

where Λ_χ is the order-universal constant from Proposition 1 and χ is given by (3.27). We note that according to (3.28) we have $\eta_\chi = \mathcal{O}\big((1+\kappa)^{1+\chi-1/(m+1)}n^{(1+\chi)/2}\big)$.

THEOREM 3.12. *Assume that HLCP is sufficient and solvable, and consider the sequence produced by the algorithm, where the sequence* (α_k) *is given by* (3.59). *Assume also that the starting point satisfies* (3.20) *or* (3.19) *and that*

$$L_\epsilon \geq \max\left\{(\beta_0-\beta_*)|\log\gamma|\,,\ \ \exp(1)\right\}.$$

Then

$$\begin{aligned}K_\epsilon &\leq 46\left(\log\left(\eta_\chi L_\epsilon\right)\right)^{\frac{1+\nu}{m}}\left(\eta_\chi L_\epsilon\right)^{\frac{m+1}{m}}\\ &= \mathcal{O}\left(\left(\log\left((1+\kappa)^{\chi+\frac{m}{m+1}}n^{\frac{1+\chi}{2}}L_\epsilon\right)\right)^{\frac{1+\nu}{m}}\left((1+\kappa)^{1+\frac{\chi(m+1)}{m}}n^{\frac{1+\chi}{2}}L_\epsilon\right)^{\frac{m+1}{m}}\right),\end{aligned}$$

with χ *given by* (3.27).

Proof. In what follows we will use Proposition 1, Theorem 3.8, and their proofs. We note that the proofs of (3.50), (3.51), and (3.52) also hold with Λ replaced by Λ_χ so that $\theta_k \geq \breve{\omega}_\chi \sqrt[m+1]{\alpha_k}$, $k = 0, 1, \ldots,$ where

$$(3.63)\quad \breve{\omega}_\chi := \frac{1}{\Lambda_\chi}\left(\frac{\gamma_0}{11.2}\right)^{\frac{1}{m+1}},\ \ \frac{1}{\breve{\omega}_\chi} = \mathcal{O}\left((1+\kappa)^{1+\chi-\frac{1}{m+1}}\,n^{\frac{(1+\chi)}{2}}\right),\ \ \Lambda_\chi \geq \frac{4}{\alpha_0\gamma_0}.$$

By applying Theorem 3.1 and using (3.53) we deduce that

$$\log\frac{\|r^k\|_2}{\|r^0\|_2} = \log\frac{\tau_k}{\tau_0} \leq (1+\vartheta)\sum_{j=0}^{k-1}\log\left(1-\breve{\omega}_\chi\alpha_j^{\frac{1}{m+1}}\right) \leq -(1+\vartheta)\,\breve{\omega}_\chi\sum_{j=0}^{k-1}\alpha_j^{\frac{1}{m+1}},$$

$$\log\frac{\mu_k}{\mu_0} \leq (\beta_0-\beta_*)|\log\gamma| + \log\frac{\tau_k}{\tau_0} \leq L_\epsilon - (1+\vartheta)\,\breve{\omega}_\chi\sum_{j=0}^{k-1}\alpha_j^{\frac{1}{m+1}}.$$

Thus the following implication holds:

$$\sum_{k=0}^{K-1}\alpha_k^{\frac{1}{m+1}} \geq \breve{M} := \frac{2L_\epsilon}{(1+\vartheta)\breve{\omega}_\chi}\ \ \text{implies}\ \ K_\epsilon \leq K.$$

Since $L_\epsilon \geq \exp(1)$ and $\breve{\omega}_\chi \leq 0.25\gamma_0\alpha_0$, we have

$$\breve{M}_* := \frac{2L_\epsilon}{(1+\vartheta)\breve{\omega}_\chi\left(\nu(\beta_0-\beta_*)\right)^{1/(m+1)}} = \eta_\chi\,L_\epsilon > 4\exp(1).$$

Therefore we can use Lemma 3.11 and the inequality

$$\left(2\left(1+\frac{1+\nu}{m+1}\log\frac{1+\nu}{m}\right)\left(\frac{m+1}{m}\right)\right)^{\frac{1+\nu}{m}} \leq 46\ \ \forall m\geq 1\ \forall\nu\in[0,1]$$

to show that

$$K_\epsilon \leq \overline{K}_\epsilon := 46\left(\log\breve{M}_*\right)^{\frac{1+\nu}{m}}\breve{M}_*^{\frac{m+1}{m}},$$

which completes the proof of our theorem. ∎

For the simplest case $m = 1$, one obtains the following estimate:

$$K_\epsilon = \mathcal{O}\Big((\log \breve{M}_*)^{1+\nu} \breve{M}_*^2\Big) \ , \quad \breve{M}_* = \mathcal{O}\Big((1+\kappa)^{0.5+\chi} n^{(1+\chi)/2} L_\epsilon\Big) \ . \tag{3.64}$$

If ν and m are positive constants independent of n, κ, and L_ϵ, then the algorithm is also independent of n, κ, and L_ϵ. Since κ is generally unknown, we will not consider algorithms that depend on κ. However, by letting m depend on n or L_ϵ, we can slightly improve the computational complexity. We note that L_ϵ depends both on ϵ and the value of the complementarity gap and the norm of the residual at the starting point.

COROLLARY 3.13. *If the hypothesis of Theorem* 3.12 *is satisfied, then*

$$K_\epsilon = \mathcal{O}\bigg(\Big(\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)\Big)^{\frac{1+\nu}{m}} (1+\kappa)^{\frac{(1+\chi)(m+1)-1}{m}} L_\epsilon^{\frac{m+1}{m}} n^{\frac{1+\chi}{2}}\bigg)$$

if we take $m \geq \log n$ *in the algorithm, and*

$$K_\epsilon = \mathcal{O}\bigg(\Big(1 + \frac{\log(1+\kappa)}{m}\Big) (1+\kappa)^{\frac{(1+\chi)(m+1)-1}{m}} n^{\frac{1+\chi}{2}} L_\epsilon\bigg)$$

if we take $m \geq \max\{\log n\,,\ L_\epsilon\,,\ 2\}$ *in the algorithm.*

Proof. (i) If $m \geq \log n$, then $n = \mathcal{O}(\exp(m))$ and by (3.28),

$$\begin{aligned}
\eta_\chi L_\epsilon &= \mathcal{O}\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}}\, n^{\frac{(1+\chi)}{2}}\, L_\epsilon\Big) \\
&= \mathcal{O}\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}}\ \exp(m\,(1+\chi)/2)\, L_\epsilon\Big)\,, \\
(\eta_\chi L_\epsilon)^{\frac{1}{m}} &= \mathcal{O}\Big((1+\kappa)^{\frac{1+\chi}{m}-\frac{1}{m(m+1)}}\Big)\, L_\epsilon^{\frac{1}{m}}, \\
\log(\eta_\chi L_\epsilon) &= \mathcal{O}\Big(\max\{\log((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon), m\}\Big)\,, \\
\big(\log(\eta_\chi L_\epsilon)\big)^{\frac{1+\nu}{m}} &= \mathcal{O}\bigg(\Big(\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)\Big)^{\frac{1+\nu}{m}}\bigg)\,,
\end{aligned}$$

since $\mathcal{O}\big(m^{1/m}\big) = \mathcal{O}(1)$. The result then follows by (3.63):

$$\begin{aligned}
K_\epsilon &= \mathcal{O}\,\bigg(\Big(\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)\Big)^{\frac{1+\nu}{m}} \cdot (1+\kappa)^{\frac{1+\chi}{m}-\frac{1}{m(m+1)}} L_\epsilon^{\frac{1}{m}} \\
&\qquad\qquad \cdot\, (1+\kappa)^{1+\chi-\frac{1}{m+1}} n^{\frac{1+\chi}{2}} L_\epsilon\bigg) \\
&= \mathcal{O}\bigg(\Big(\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)\Big)^{\frac{1+\nu}{m}} \cdot \Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)^{\frac{m+1}{m}} n^{\frac{1+\chi}{2}}\bigg).
\end{aligned}$$

(ii) Since $m \geq \max\{\log n, L_\epsilon, 2\}$, we deduce that

$$\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)^{\frac{1}{m}} \leq (1+\kappa)^{\frac{1+\chi(m+1)-1}{m(m+1)}} \exp(1).$$

If $(1+\kappa)^{1+\chi-\frac{1}{m+1}} \geq \exp(m)$, then $\log(1+\kappa)^{1+\chi-\frac{1}{m+1}} \geq m > \log m \geq \log 2$ and

$$\begin{aligned}
&\Big(\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\Big)\Big)^{\frac{1+\nu}{m}} \leq \Big(\log\Big(m(1+\kappa)^{1+\chi-\frac{1}{m+1}}\Big)\Big)^{\frac{1+\nu}{m}} \\
&\leq \Big(2\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}}\Big)\Big)^{\frac{2}{m}} \leq 2\log\Big((1+\kappa)^{1+\chi-\frac{1}{m+1}}\Big)^{\frac{2}{m}} \\
&\leq 2\frac{\log(1+\kappa)^{1+\chi-\frac{1}{m+1}}}{m^{1-\frac{2}{m}}} \leq 12\frac{\log(1+\kappa)}{m} \qquad (\text{by } m \leq 3m^{1-\frac{2}{m}})\,,
\end{aligned}$$

while if $(1+\kappa)^{1+\chi-\frac{1}{m+1}} < \exp(m)$, then

$$\left(\log\left((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\right)\right)^{\frac{1+\nu}{m}} \le (2m)^{\frac{1+\nu}{m}} \le (2m)^{\frac{2}{m}} \le 4\,.$$

We have thus proved that in case (ii) we have

$$\left(\log\left((1+\kappa)^{1+\chi-\frac{1}{m+1}} L_\epsilon\right)\right)^{\frac{1+\nu}{m}} \le \left(4 + \frac{12\log(1+\kappa)}{m}\right)\,.$$

Therefore

$$K_\epsilon = \mathcal{O}\left(\left(1+\frac{\log(1+\kappa)}{m}\right)(1+\kappa)^{\frac{(1+\chi)(m+1)-1}{m(m+1)}} n^{\frac{1+\chi}{2}} L_\epsilon\right). \qquad \square$$

4. Conclusions. In this paper we have presented an infeasible interior point method of order m for solving sufficient LCPs. The algorithm does not depend on the handicap κ of the complementarity problem. It uses only one matrix factorization and m backsolves per step and produces iterates in the wide neighborhood of the central path. The iteration sequence produced by the higher order algorithm converges superlinearly to a maximal complementarity solution with Q-order at least $(m+1)/2$ under general assumptions. If the problem is nondegenerate, then the Q-order of convergence is $m+1$. The iteration complexity of the first order algorithm is $\mathcal{O}((1+\kappa)nL_\epsilon^2\log^2(\sqrt{(1+\kappa)n}L_\epsilon))$ for starting points that are feasible or "almost feasible," and $\mathcal{O}((1+\kappa)^3n^2L_\epsilon^2\log^2((1+\kappa)^{1.5}nL_\epsilon))$ for starting points that are "sufficiently large." The iteration complexity of the higher order algorithm with $m \ge \log n$ is $\mathcal{O}((1+\kappa)\sqrt{n}\;L_\epsilon^{1+\frac{1}{m}}\;\log^{\frac{2}{m}}((1+\kappa)^{\frac{m}{m+1}}L_\epsilon))$ for starting points that are feasible or "almost feasible," and $\mathcal{O}((1+\kappa)^{2+\frac{1}{m}}n\;L_\epsilon^{1+\frac{1}{m}}\;\log^{\frac{2}{m}}((1+\kappa)^{\frac{2m+1}{m+1}}L_\epsilon))$ for starting points that are "sufficiently large." If $m \ge \max\{\log n, L_\epsilon, 2\}$, then the corresponding iteration complexities are $\mathcal{O}((1+\kappa)(1+\log\sqrt[m]{1+\kappa})\sqrt{n}\;L_\epsilon)$ and $\mathcal{O}((1+\kappa)^{2+1/m}(1+\log\sqrt[m]{1+\kappa})n\;L_\epsilon)$, respectively. Although the above results are proved in a very general setting, they are new even for LP.

Acknowledgment. The authors thank an anonymous referee for carefully reading the manuscript and detailed comments that helped to improve the paper.

REFERENCES

[1] W. Ai and S. Zhang, *An $O(\sqrt{n}L)$ iteration primal-dual path-following method, based on wide neighborhoods and large updates, for monotone LCP*, SIAM J. Optim., 16 (2005), pp. 400–417.
[2] M. Anitescu, G. Lesaja, and F. A. Potra, *Equivalence between different formulations of the linear complementarity problem*, Optim. Methods Softw., 7 (1997), pp. 265–290.
[3] J. F. Bonnans and C. C. Gonzaga *Convergence of interior point algorithms for the monotone linear complementarity problem*, Math. Oper. Res., 21 (1996), pp. 1–25.
[4] J. F. Bonnans and F. A. Potra, *On the convergence of the iteration sequence of infeasible path following algorithms for linear complementarity problems*, Math. Oper. Res., 22 (1997), pp. 378–407.
[5] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, MA, 1992.
[6] R. W. Cottle, J.-S. Pang, and V. Venkateswaran, *Sufficient matrices and the linear complementarity problem*, Linear Algebra Appl., 114/115 (1989), pp. 231–249.
[7] A. S. El-Bakry, R. A. Tapia, and Y. Zhang, *A study of indicators for identifying zero variables in interior-point methods*, SIAM Rev., 36 (1994), pp. 45–72.
[8] C. C. Gonzaga, *The largest step path following algorithm for monotone linear complementarity problems*, Math. Program. Ser. A, 76 (1997), pp. 309–332.

[9] J. Ji, F. A. Potra, and S. Huang, *A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence*, J. Optim. Theory Appl., 84 (1995), pp. 187–199.
[10] J. Ji, F. A. Potra, and R. Sheng, *A predictor–corrector method for solving the P_*-matrix LCP from infeasible starting points*, Optim. Methods Softw., 6 (1995), pp. 109–126.
[11] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise, *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, Lecture Notes in Comput. Sci. 538, Springer-Verlag, New York, 1991.
[12] X. Liu and F. A. Potra, *Corrector-predictor methods for sufficient linear complementarity problems in a wide neighborhood of the central path*, SIAM J. Optim., 17 (2006), pp. 871–890.
[13] K. McShane, *Superlinearly convergent $O(\sqrt{n}L)$-iteration interior-point algorithms for linear programming and the monotone linear complementarity problem*, SIAM J. Optim., 4 (1994), pp. 247–261.
[14] S. Mehrotra, *Quadratic convergence in a primal-dual method*, Math. Oper. Res., 18 (1993), pp. 741–751.
[15] S. Mizuno, *A superlinearly convergent infeasible interior-point algorithm for geometrical LCPs without a strictly complementary condition*, Math. Oper. Res., 21 (1996), pp. 382–400.
[16] S. Mizuno, M. J. Todd, and Y. Ye, *On adaptive-step primal-dual interior-point algorithms for linear programming*, Math. Oper. Res., 18 (1993), pp. 964–981.
[17] R. D. C. Monteiro and S. J. Wright, *Local convergence of interior-point algorithms for degenerate monotone LCP*, Comput. Optim. Appl., 3 (1994), pp. 131–155.
[18] A. M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, Academic Press, New York, 1973.
[19] J. Peng, T. Terlaky, and Y. Zhao, *A predictor-corrector algorithm for linear optimization based on a specific self-regular proximity function*, SIAM J. Optim., 15 (2005), pp. 1105–1127.
[20] F. A. Potra, *Efficient hybrid algorithms for finding zeros of convex functions*, J. Complexity, 10 (1994), pp. 199–215.
[21] F. A. Potra, *An infeasible-interior-point predictor-corrector algorithm for linear programming*, SIAM J. Optim., 6 (1996), pp. 19–32.
[22] F. A. Potra, *An $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence*, Ann. Oper. Res., 62 (1996), pp. 81–102.
[23] F. A. Potra, *Q-superlinear convergence of the iterates in primal-dual interior-point methods*, Math. Program. Ser. A, 91 (2001), pp. 99–115.
[24] F. A. Potra, *A superlinearly convergent predictor-corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{n}L)$-iteration complexity*, Math. Program. Ser. A, 100 (2004), pp. 317–337.
[25] F. A. Potra, *The Kantorovich theorem and interior point methods*, Math. Program. Ser. A, 102 (2005), pp. 47–70.
[26] F. A. Potra, *Corrector-predictor methods for monontone linear complementarity problems in a wide neighborhood of the central path*, Math. Program. Ser. B, 111 (2008), pp. 243–272.
[27] F. A. Potra *Primal-dual affine scaling interior point methods for linear complementarity problems*, SIAM J. Optim., 19 (2008), pp. 114–143.
[28] F. A. Potra and X. Liu, *Predictor-corrector methods for sufficient linear complementarity problems in a wide neighborhood of the central path*, Optim. Methods Softw., 20 (2005), pp. 145–168.
[29] F. A. Potra and R. Sheng, *A large-step infeasible-interior-point method for the P_*-matrix LCP*, SIAM J. Optim., 7 (1997), pp. 318–335.
[30] F. A. Potra and R. Sheng, *Predictor-corrector algorithms for solving P_*-matrix LCP from arbitrary positive starting points*, Math. Program. Ser. B, 76 (1997), pp. 223–244.
[31] F. A. Potra and R. Sheng, *A path following method for LCP with superlinearly convergent iteration sequence*, Ann. Oper. Res., 81 (1998), pp. 97–114.
[32] F. A. Potra and R. Sheng, *Superlinearly convergent infeasible-interior-point algorithm for degenerate LCP*, J. Optim. Theory Appl., 97 (1998), pp. 249–269.
[33] R. Sheng and F. A. Potra, *A quadratically convergent infeasible-interior-point algorithm for LCP with polynomial complexity*, SIAM J. Optim., 7 (1997), pp. 304–317.
[34] J. Stoer, *Improved High Order Long-step Methods for Solving Linear Complementarity Problems*, Technical report 237, Institut für Angewandte Mathematik und Statistik Universität Würzburg, Würzburg, Germany, 1999.
[35] J. Stoer, *High order long-step methods for solving linear complementarity problems*, Ann. Oper. Res., 103 (2001), pp. 149–159.
[36] J. Stoer and M. Wechs, *The complexity of high-order predictor-corrector methods for solving sufficient linear complementarity problems*, Optim. Methods Softw., 10 (1998), pp. 393–417.

[37] J. Stoer and M. Wechs, *Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity*, Math. Program. Ser. A, 83 (1998), pp. 407–423.
[38] J. Stoer and M. Wechs, *On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems*, Numer. Math., 81 (1999), pp. 631–645.
[39] J. Stoer, M. Wechs, and S. Mizuno, *High order infeasible-interior-point methods for solving sufficient linear complementarity problems*, Math. Oper. Res., 23 (1998), pp. 832–862.
[40] J. F. Sturm, *Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists*, Math. Oper. Res., 24 (1999), pp. 72–94.
[41] H. Väliaho, *P_*-matrices are just sufficient*, Linear Algebra Appl., 239 (1996), pp. 103–108.
[42] S. J. Wright, *A path-following interior-point algorithm for linear and quadratic problems*, Ann. Oper. Res., 62 (1996), pp. 103–130.
[43] S. J. Wright and Y. Zhang, *A superquadratic infeasible-interior-point algorithm for linear complementarity problems*, Math. Program., 73 (1996), pp. 269–289.
[44] Y. Ye and K. Anstreicher, *On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP*, Math. Program., 62 (1993), pp. 537–551.
[45] Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang, *A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming*, Math. Program., 59 (1993), pp. 151–162.
[46] D. Zhang and Y. Zhang, *A Mehrotra-type predictor-corrector algorithm with polynomiality and Q-subquadratic convergence*, Ann. Oper. Res., 62 (1996), pp. 131–150.
[47] G. Y. Zhao and J. S. Zhu *The curvature integral and the complexity of linear complementarity problems*, Math. Program. Ser. A, 70 (1995), pp. 107–122.