Mathematical Programming manuscript No.
(will be inserted by the editor)

Florian A. Potra

Corrector-predictor methods for monotone linear complementarity problems in a wide neighborhood of the central path ⋆

Received: date / Revised version: date

Abstract. Two corrector-predictor interior point algorithms are proposed for solving monotone linear complementarity problems. The algorithms produce a sequence of iterates in the $\mathcal{N}_{\infty}^{-}$ neighborhood of the central path. The first algorithm uses line search schemes requiring the solution of higher order polynomial equations in one variable, while the line search procedures of the second algorithm can be implemented in $O(m\,n^{1+\alpha})$ arithmetic operations, where n is the dimension of the problems, $\alpha \in (0,1]$ is a constant, and m is the maximum order of the predictor and the corrector. If $m = \Omega(\log n)$ then both algorithms have $O(\sqrt{n}L)$ iteration complexity. They are superlinearly convergent even for degenerate problems.

Key words. linear complementarity problem, interior-point algorithm, large neighbourhood, superlinear convergence

1. Introduction

Predictor-corrector methods play a special role among interior point methods for linear programming (LP). They operate between two neighborhoods $\mathcal{N} \subset \overline{\mathcal{N}}$ of the primal-dual central path [21, 22]. The role of the predictor step is to increase optimality, while the role of the corrector is to increase proximity to the central path. At a typical iteration one is given a point $z \in \mathcal{N}$ with (normalized) duality gap $\mu = \mu(z)$. In the predictor step one computes a point $\overline{z} \in \overline{\mathcal{N}}$ with reduced duality gap, $\overline{\mu} = \mu(\overline{z}) < \mu$. The corrector step produces a point $z^+ \in \mathcal{N}$, with $\mu(z^+) = \mu(\overline{z})$, in the original neighborhood, so that the predictor-corrector scheme can be iterated. Predictor-corrector algorithms have "polynomial complexity", in the sense that for any $\varepsilon > 0$ and any starting point $z^0 \in \mathcal{N}$ they find a feasible point $z \in \mathcal{N}$ with duality gap $\mu(z) < \varepsilon$ in at most $O\left(n^{\iota} \log\left(\mu(z^0)/\varepsilon\right)\right)$ iterations, for some $\iota \geq .5$. In the case of a linear programming problem with integer data of bit-length L, this implies that a feasible point z with duality gap $\mu(z) = O(2^{-L})$ can be obtained in at most $O(n^{\iota}L)$ iterations. A rounding procedure involving $O(n^3)$ arithmetic operations can then be applied to obtain an exact solution of the LP. Therefore, in this case we say that the algorithm has "$O(n^{\iota}L)$ iteration complexity".

Extensive numerical experiments show convincingly that predictor-corrector methods perform better when using large neighborhoods of the central path.

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

Mathematics Subject Classification (1991): 49M15, 65K05, 90C33

⋆ Work supported by the National Science Foundation under Grant No. 0139701 .

Paradoxically, the best theoretical computational complexity results were obtained for small neighborhoods of the central path. For example, the Mizuno-Todd-Ye predictor corrector method (MTY) [10] uses small neighborhoods of the central path (defined by the δ_2 proximity measure) and has $O(\sqrt{n}L)$ iteration complexity – the best iteration complexity obtained so far *for any* interior point method. Predictor-corrector methods of MTY type are more difficult to develop and analyze in large neighborhoods of the central path (e.g., defined by the δ_∞ or the δ_∞^- proximity measures; see the definitions in the next section). This is due to the fact that correctors are rather inefficient in such neighborhoods. It is known that one needs $O(n)$ corrector steps in order to reduce the δ_∞ proximity measure by a factor of .5 (see [2]). Therefore a straightforward generalization of the MTY algorithm would have $O(n^{1.5}L)$ iteration complexity. Using a very elegant analysis, Gonzaga showed that a predictor-corrector method based on a δ_∞ neighborhood of the central path has $O(nL)$-iteration complexity [5]. However, his predictor-corrector method does no longer have the simple structure of MTY, where a predictor is followed by just one corrector. In Gonzaga's algorithm a predictor is followed by an apriori unknown number of correctors. In fact, the above mentioned complexity result is proved by showing that the total number of correctors is at most $O(nL)$.

Predictor-corrector methods are superlinearly convergent, a very important feature which is one of the reasons for their excellent practical performance. For example, Ye et al. [24] proved that the duality gap of the sequence produced by the MTY algorithm converges quadratically to zero. This result was extended to monotone linear complementarity problems (LCP) that are nondegenerate, in the sense that they have a strictly complementarity solution [7,23]. The nondegeneracy assumption is not restrictive, since according to [11] a large class of interior point methods, which contains MTY, can have only linear convergence if this assumption is violated. However, it is possible to obtain superlinear convergence for degenerate problems either by modifying the algorithm so that it detects "the variables that are not strictly complementary" and treats them differently [9,15–17], or by using higher order predictors [18,20]. All the superlinear convergence results mentioned above were proved for small neighborhoods of the central path.

The use of higher order methods also leads to interior point methods in the large neighborhood of the central path with improved iteration complexity, like those considered in [12,6,25]. The iteration complexity of the algorithms from [12,6] can be made arbitrarily close to $O(\sqrt{n}L)$, while the algorithm from [25], which is of order n^ω, $\omega > 0$, has $O(\sqrt{n}L)$ iteration complexity. These algorithms are not of a predictor-corrector type, and they are not superlinearly convergent. A superlinear interior point algorithm for sufficient linear complementarity problems in the δ_∞^- neighborhood was proposed by Stoer [19], but no complexity results have been proved for this algorithm. An interior point method for LP, acting in a large neighborhood of the central path defined by a self-regular proximity measure, with $O(\sqrt{n}L \log n)$ iteration complexity and superlinear convergence, was proposed in [13]. The relation between the neighborhood considered in that paper and the δ_∞^- neighborhood will be examined in Section 2. The method of [13] was developed only for LP, in which case a strictly complementary solution always exists. It appears that a straightforward generalization of this method for LCP is not likely to be superlinearly convergent on degenerate problems.

In a recent paper [14] we proposed a predictor-corrector method for monotone LCP using δ_∞^- neighborhoods of the central path, which has $O(\sqrt{n}L)$-iteration complexity and is superlinearly convergent even for degenerate problems. The method employed a predictor of order n^ω, $\omega > 0$, and a first order corrector. An anonymous referee pointed out that the algorithm could be presented as a corrector-predictor method. Coincidentally, the same suggestion was made by Clovis Gonzaga in a private conversation. In this variant one would start with a point in $\overline{\mathcal{N}}$, and one would use a corrector to produce a point closer to the central path, followed by a predictor step to obtain a point on the boundary of $\overline{\mathcal{N}}$. The advantage of this approach is that only one neighborhood of the central path needs to be considered, avoiding thus the explicit relation between the "radii" of the neighborhoods $\mathcal{N} \subset \overline{\mathcal{N}}$ assumed in [14]. Since the decrease of the duality gap along the predictor direction is faster if the point is closer to the central path, it makes sense to start the iteration with a corrector step. However, transforming the predictor-corrector algorithm of [14] into a corrector-predictor algorithm turned out to be a nontrivial task.

The classical centering direction is very efficient in the small neighborhood of the central path, so that no line search along that direction was employed in the original MTY algorithm [10]. Therefore, that algorithm can be trivially transformed into a corrector-predictor algorithm, by performing exactly the same operations, but starting with a corrector step. By contrast, the classical centering direction is very inefficient in the large neighborhood of the central path and a line search is always necessary. In the predictor-corrector setting from [14], a line search was used along the classical centering direction starting from a point $\overline{z} \in \overline{\mathcal{N}}$ in order to obtain a point $z^+ \in \mathcal{N}$. The line search is successful provided the radii of the neighborhoods $\mathcal{N}$ and $\overline{\mathcal{N}}$ differ only by a small factor. In a corrector-predictor setting, where only one neighborhood of the central path is used, the line search on the corrector direction has to be done in such a way that it optimizes the decrease of the δ_∞^- proximity measure to the central path. Designing such a line search is nontrivial since it involves minimizing a nonsmooth function. In the present paper we propose a very efficient implementation of the corrector line search. Also, since the pure centering direction is anyhow inefficient in the large neighborhood, we propose a corrector direction along which both the proximity to the central path and the duality gap can be reduced. In order to get additional efficiency we use higher order methods both in the corrector step and the predictor step. The resulting algorithm has very desirable properties.

More precisely, the corrector-predictor method presented in this paper uses a corrector based on a polynomial of order m_c followed by a predictor based on a polynomial of order m_p. The computation of the search directions involves two matrix factorizations and $m_c + m_p + 2$ backsolves. The line search procedures can be implemented in $O\left((m_c + m_p)n^{1+\alpha}\right)$ arithmetic operations, for some $0 < \alpha \leq$

1, or even in $O\left((m_c + m_p)n\, \log n\right)$ arithmetic operations. If $m_c + m_p = \Omega(\log n)$, then the algorithm has $O(\sqrt{n}L)$ iteration complexity. In case of full matrices, the cost of a factorization is $O(n^3)$ arithmetic operations, and the cost of a backsolve is $O(n^2)$ arithmetic operations, so that if $m_c + m_p = O(n^\omega)$, for some $0 < \omega < 1$, then one iteration of our algorithm can be implemented in at most $O(n^3)$ arithmetic operations, the cost being dominated by the cost of the two matrix factorizations.

The above mentioned complexity result is obtained by carefully analyzing the performance of the corrector at the current iteration in connection with the behavior of the predictor at the previous iteration. We show that we may fail to obtain a sufficient reduction of the duality gap in the corrector step, only if a significant reduction of the duality gap had already been obtained at the previous predictor step. Thus the complexity result is obtained from the joint contribution of the corrector and predictor steps. The superlinear convergence is due to the predictor steps. We show that the duality gap converges to zero with Q-order $m_p + 1$ in the nondegenerate case, and with Q-order $(m_p + 1)/2$ in the degenerate case. Hence if the order of the predictor is strictly greater than one, then our algorithm is superlinearly convergent even for degenerate problems.

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

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

We frequently use the $O(\cdot)$ and $\Omega(\cdot)$ notation to express asymptotic relationships between functions. The most common usage will be associated with a sequence $\{x^k\}$ of vectors and a sequence $\{\tau_k\}$ of positive real numbers. In this case $x^k = O(\tau_k)$ means that there is a constant K (dependent on problem data) such that for every $k \in \mathbb{N}$, $\left\|x^k\right\| \leq K\tau_k$. Similarly, if $x^k > 0$, $x^k = \Omega(\tau_k)$ means that $(x^k)^{-1} = O(1/\tau_k)$. If we have both $x^k = O(\tau_k)$ and $x^k = \Omega(\tau_k)$, we write $x^k = \Theta(\tau_k)$

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

2. The horizontal linear complementarity problem

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

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

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

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

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

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

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

In this case we say that the HLCP is skew-symmetric. In the skew-symmetric case we can often obtain sharper estimates, due to the following simple result.

Proposition 1. *If HLCP is skew-symmetric then* $u^T\overline{v}+\overline{u}^T v = 0$, *for all* $u, v, \overline{u}, \overline{v}$ *satisfying* $Qu + Rv = Q\overline{u} + R\overline{v} = 0$.

Proof. If $Qu + Rv = Q\overline{u} + R\overline{v} = 0$, then $Q(u + \overline{u}) + R(v + \overline{v}) = 0$, so that $u^T v = \overline{u}^T\overline{v} = (u + \overline{u})^T(v + \overline{v}) = 0$. Hence, $0 = (u + \overline{u})^T(v + \overline{v}) = u^T v + u^T\overline{v} + \overline{u}^T v + \overline{u}^T\overline{v} = u^T\overline{v} + \overline{u}^T v$. ∎

We remark that the corresponding statement fails in the monotone case, in the sense that there are examples where $Qu+Rv = Q\overline{u}+R\overline{v} = 0$ and $u^T\overline{v}+\overline{u}^T v < 0$. However, it is easily shown that if HLCP is monotone then the $n \times 2n$-matrix (Q, R) has full rank [3].

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

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

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

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

The relative interior of $\mathcal{F}$,

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

will be called the set of strictly feasible points, or the set of interior points. It is known (see [8]) that if $\mathcal{F}^0$ is nonempty, then, for any parameter $\tau > 0$ the nonlinear system,

$$\begin{aligned} xs &= \tau e \\ Qx + Rs &= b \end{aligned}$$

has a unique positive solution. The set of all such solutions defines the central path $\mathcal{C}$ of the HLCP. By considering the quadratic mapping $F_\tau : I\!R^{2n} \to I\!R^{2n}$

$$z = \begin{bmatrix} x \\ s \end{bmatrix} \to F_\tau(z) = \begin{bmatrix} xs - \tau e \\ Qx + Rs - b \end{bmatrix}$$

we can write

$$\mathcal{C} = \{z \in I\!R^{2n}_{++} \ : \ F_\tau(z) = 0, \, \tau > 0\} .$$

If $F_\tau(z) = 0$, then necessarily $\tau = \mu(z)$, where $\mu(z) = x^T s/n$. The distance of a point $z \in \mathcal{F}$ to the central path can be quantified by different proximity measures. The following proximity measures have been extensively used in the interior point literature:

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

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

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

$$\begin{aligned} \mathcal{N}_2(\alpha) &= \{z \in \mathcal{F}^0 \ : \ \delta_2(z) \le \alpha \} , \\ \mathcal{N}_\infty(\alpha) &= \{z \in \mathcal{F}^0 \ : \ \delta_\infty(z) \le \alpha \} , \\ \mathcal{N}^-_\infty(\alpha) &= \{z \in \mathcal{F}^0 \ : \ \delta^-_\infty(z) \le \alpha \} , \end{aligned}$$

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

$$\mathcal{C} \subset \mathcal{N}_2(\alpha) \subset \mathcal{N}_\infty(\alpha) \subset \mathcal{N}^-_\infty(\alpha), \ \lim_{\alpha \downarrow 0} \mathcal{N}^-_\infty(\alpha) = \mathcal{C}, \ \lim_{\alpha \uparrow 1} \mathcal{N}^-_\infty(\alpha) = \mathcal{F} . \tag{2}$$

A neighborhood that covers in the limit the set of all feasible points is called a wide neighborhood of the central path. In a recent paper, Peng et al. [13] considered a predictor-corrector method acting in a neighborhood of the form

$$\mathcal{N}(\alpha) = \left\{ z \in \mathcal{F}^0 \ : \ \sum_{i=1}^n \left(\frac{x_i s_i}{\mu(z)} \right)^{-\frac{1}{2}\log n} \le n^{\frac{1}{1-\alpha}} \right\} . \tag{3}$$

The algorithm has $O(\sqrt{n} L \log n)$ iteration complexity, where the constant hidden in the "$O(\cdot)$" notation is proportional to $\exp(1/(1-\alpha))$. We end this section by showing that this is also a wide neighborhood of the central path and by estimating its relation to $\mathcal{N}^-_\infty(\alpha)$.

Proposition 2. *The neighborhood $\mathcal{N}(\alpha)$ defined in (3) satisfies the following properties:*

$$\lim_{\alpha\downarrow 0}\mathcal{N}(\alpha)=\mathcal{C}\ ,\ \lim_{\alpha\uparrow 1}\mathcal{N}(\alpha)=\mathcal{F}, \tag{4}$$

$$\mathcal{N}(\alpha)\subset\mathcal{N}_\infty^-\left(1-\exp\frac{-2}{1-\alpha}\right)\ ,\ \mathcal{N}_\infty^-(\alpha)\subset\mathcal{N}\left(\frac{\log(1-\alpha)}{\log(1-\alpha)-2}\right)\ . \tag{5}$$

Proof. Since (4) is a consequence of (2) and (5) we only have to prove (5). Let us denote

$$p=\frac{xs}{\mu(z)},\quad \beta=\min_{1\le i\le n}p_i.$$

We have

$$\begin{aligned}
&z\in\mathcal{N}(\alpha)\ \Rightarrow\ \sum_{i=1}^n p_i^{-\frac12\log n}\le n^{\frac{1}{1-\alpha}}\ \Rightarrow\ \beta^{-\frac12\log n}\le n^{\frac{1}{1-\alpha}}\\
&\Rightarrow\ -\frac12\log\beta\le\frac{1}{1-\alpha}\ \Rightarrow\ \beta\ge\exp\left(\frac{-2}{1-\alpha}\right)\\
&\Rightarrow\ \|p-e\|_\infty^-\le 1-\exp\left(\frac{-2}{1-\alpha}\right)\ \Rightarrow\ z\in\mathcal{N}_\infty^-\left(1-\exp\frac{-2}{1-\alpha}\right),
\end{aligned}$$

and

$$\begin{aligned}
&z\in\mathcal{N}_\infty^-(\alpha)\ \Rightarrow\ \beta\ge 1-\alpha\ \Rightarrow\ \sum_{i=1}^n p_i^{-\frac12\log n}\le n(1-\alpha)^{-\frac12\log n}\\
&\Rightarrow\ z\in\mathcal{N}\left(\frac{\log(1-\alpha)}{\log(1-\alpha)-2}\right).
\end{aligned}$$

The proof is complete. ∎

3. A higher order corrector-predictor algorithm

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

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

At each step of our algorithm we are given a point $z=\lceil\, x,\, s\,\rfloor\in\mathcal{D}(\beta)$ and we consider an mth order vector valued polynomial of the form

$$z(\theta)=z+\sum_{i=1}^m w^i\theta^i\,, \tag{6}$$

where the vectors $w^i = \lceil u^i,\, v^i \rfloor$ are obtained as solutions of the following linear systems

$$\begin{aligned}
&\begin{cases} su^1 + xv^1 = \gamma\mu e - (1+\epsilon)xs \\ Qu^1 + Rv^1 = 0 \end{cases}, \\
&\begin{cases} su^2 + xv^2 = \epsilon xs - u^1 v^1 \\ Qu^2 + Rv^2 = 0 \end{cases}, \\
&\begin{cases} su^i + xv^i = -\sum_{j=1}^{i-1} u^j v^{i-j} \\ Qu^i + Rv^i = 0 \end{cases}, \quad i = 3, \ldots, m\,.
\end{aligned} \tag{7}$$

In a corrector step we choose $\epsilon = 0$ and $\gamma \in [\underline{\gamma}\,, \overline{\gamma}]$, where $0 < \underline{\gamma} < \overline{\gamma} < 1$ are given parameters, while in a predictor step we take

$$\gamma = 0\,, \qquad \text{and } \epsilon = \begin{cases} 0\,, & \text{if HLCP is nondegenerate} \\ 1\,, & \text{if HLCP is degenerate} \end{cases}. \tag{8}$$

We note that for $\epsilon = \gamma = 0$, w^1 is just the affine scaling direction, while for $\epsilon = 0$ and $\gamma = 1$, w^1 becomes the classical centering direction. For $\epsilon = 0$ and $0 < \gamma < 1$, w^1 is a convex combination of the affine scaling and the centering directions. The directions w^i are related to the higher derivatives of the central path [18]. The m linear systems above have the same matrix, so that their numerical solution requires only one matrix factorization and m backsolves. This involves $O(n^3)$ + $O(m\,n^2)$ arithmetic operations. We take $m = m_c$ in the corrector step, and $m = m_p$ in the predictor step.

It is easily seen that

$$\begin{aligned}
x(\theta)s(\theta) &= (1-\theta)^{1+\epsilon} xs + \gamma\theta\mu e + \sum_{i=m+1}^{2m} \theta^i h^i\,, \\
\mu(\theta) &= (1-\theta)^{1+\epsilon}\mu + \gamma\theta\mu + \sum_{i=m+1}^{2m} \theta^i (e^T h^i / n)\,, \\
&\text{where } h^i = \sum_{j=i-m}^{m} u^j v^{i-j}\,.
\end{aligned} \tag{9}$$

Since we want to preserve positivity of our iterates we will restrict the line search on the interval $[0\,, \theta_0]$, where

$$\theta_0 = \sup\{\hat{\theta}_0 \;:\; x(\theta) > 0,\; s(\theta) > 0,\; \forall \theta \in [0\,, \hat{\theta}_0]\}\,. \tag{10}$$

Determining θ_0 involves the computation of the smallest positive roots of the mth order polynomials $x_i(\theta),\; s_i(\theta),\; i = 1, \ldots, n$. The following notation will be used in describing both the corrector and predictor steps:

$$p(\theta) = \frac{x(\theta)s(\theta)}{\mu(\theta)}\,, \quad f(\theta) = \min_{i=1,\ldots,n} p_i(\theta)\,. \tag{11}$$

In the following lemma we give a lower bound for $f(\theta)$.

Lemma 1. *If $z \in \mathcal{D}(\beta)$ then*

$$f(\theta) \geq \beta + \frac{(1-\beta)\gamma\theta\mu - \sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2}{\mu(\theta)}\,, \quad \forall\, \theta \in [0,1]\,. \tag{12}$$

Proof. We have

$$p(\theta) = \frac{x(\theta)s(\theta)}{\mu(\theta)} = \frac{(1-\theta)^{1+\epsilon}xs + \gamma\theta\mu e + \sum_{i=m+1}^{2m} \theta^i h^i}{(1-\theta)^{1+\epsilon}\mu + \gamma\theta\mu + \sum_{i=m+1}^{2m} \theta^i e^T h^i / n}$$

$$\geq \frac{(1-\theta)^{1+\epsilon}\beta\mu e + \gamma\theta\mu e + \sum_{i=m+1}^{2m} \theta^i h^i}{\mu(\theta)}$$

$$= \frac{\beta\mu(\theta)e + (1-\beta)\gamma\theta\mu e + \beta\sum_{i=m+1}^{2m} \theta^i (h^i - \frac{e^T h^i}{n} e) + (1-\beta)\sum_{i=m+1}^{2m} \theta^i h^i}{\mu(\theta)}$$

$$\geq \beta e + \frac{(1-\beta)\gamma\theta\mu - \sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2}{\mu(\theta)}\, e\,,$$

where we have used the fact that $\left\| h^i - (e^T h/n)e \right\|_2 \leq \| h \|_2$ for any vector h. ∎

The corrector. The main purpose of the corrector step is to increase proximity to the central path. Our line search procedure ensures that the optimality measure $\mu(\theta)$ is also improved by the corrector step. If HLCP is monotone, but not skew-symmetric, then we choose $\sigma \in [\underline{\sigma}, \overline{\sigma}]$, where $0 < \underline{\sigma} < \overline{\sigma} < 1$ are given parameters, and define

$$\theta_1 = \sup\{\hat{\theta}_1 \ : \ 0 \leq \hat{\theta}_1 \leq \theta_0,\ \mu(\theta) \leq (1 - \sigma(1-\gamma)\theta)\mu,\ \forall\theta \in [0\,, \hat{\theta}_1]\}\,. \tag{13}$$

If HLCP is skew-symmetric, then according to Proposition 1 we have

$$\mu(\theta) = (1 - (1-\gamma)\theta)\mu\,,$$

so that in this case we take $\sigma = 1$ and $\theta_1 = \theta_0$.

The step-length of the corrector is obtained as

$$\theta_c = \operatorname{argmax}\{f(\theta) \ : \ \theta \in [0, \theta_1]\}\ . \tag{14}$$

As a result of the corrector step we obtain the point

$$\overline{z} = \lceil\, \overline{x}\,, \overline{s}\, \rfloor := z(\theta_c)\,. \tag{15}$$

We have clearly $\overline{z} \in \mathcal{D}(\beta_c)$ with $\beta_c > \beta$. While the parameter β is fixed during the algorithm, the positive quantity β_c varies from iteration to iteration. However, we will prove that there is a constant $\beta_c^* > \beta$, such that $\beta_c > \beta_c^*$ at all iterations.

The predictor. The predictor is obtained by taking $z = \overline{z}$, where $\overline{z}$ is the result of the corrector step, and $\gamma = 0$ in (6)-(7). The aim of the predictor step is to decrease the complementarity gap as much as possible while keeping the iterate in $\mathcal{D}(\beta)$. This is accomplished by defining the predictor step-length as

$$\theta_p = \operatorname{argmin}\{\mu(\theta)\ :\ \theta \in [\,0, \theta_2\,]\}\ , \tag{16}$$

where

$$\theta_2 = \max\{\hat{\theta}_2\ :\ z(\theta) \in \mathcal{D}(\beta),\ \forall \theta \in [0\,, \hat{\theta}_2]\}\,. \tag{17}$$

A standard continuity argument can be used to show that $z(\theta) > 0$, $\forall \theta \in [0\,, \theta_2]$. The computation of θ_2 and θ_p involves the solution of polynomial inequalities of order $2m$ in θ. A practical implementation of this line search algorithm will be sketched in Section 6. With the above line search the predictor step computes a point

$$z^+ = \lceil\, x^+,\, s^+\, \rfloor := z(\theta_p)\,. \tag{18}$$

By construction we have $z^+ \in \mathcal{D}(\beta)$, so that a new corrector step can be applied. Summing up we can formulate the following iterative procedure:

Algorithm 1
Given real parameters $0 < \beta < 1$, $0 < \underline{\gamma} < \overline{\gamma} < 1$, $0 < \underline{\sigma} < \overline{\sigma} < 1$,
integers $m_c, m_p \geq 1$, and a vector $z^0 \in \mathcal{D}(\beta)$:
Set $k \leftarrow 0$;
repeat
(corrector step)
Set $z \leftarrow z^k$;
Choose $\gamma \in [\,\underline{\gamma}\,, \overline{\gamma}\,]$ and set $\epsilon = 0$, $m = m_c$;
Compute directions $w^i = \lceil\, u^i,\, v^i\, \rfloor$, $i = 1, \ldots, m$, by solving (7);
Compute θ_0 from (10)
If HLCP is skew-symmetric, set $\sigma = 1$ and $\theta_1 = \theta_0$;
Else, choose $\sigma \in [\,\underline{\sigma}\,, \overline{\sigma}\,]$, and compute θ_1 from (13);
Compute corrector step-length θ_c from (14);
Compute $\overline{z}$ from (15);
Set $\overline{z}^{\,k} \leftarrow \overline{z}\ ,\ \ \overline{\mu}_{\,k} \leftarrow \overline{\mu}\ = \mu(\overline{z}\,)\,.$
(predictor step)
Set $z \leftarrow \overline{z}^{\,k}$, $m = m_p$ and choose γ, ϵ as in (8);
Compute directions $w^i = \lceil\, u^i,\, v^i\, \rfloor$, $i = 1, \ldots, m$, by solving (7);
Compute θ_p from (16);
Compute z^+ from (18);
Set $z^{k+1} \leftarrow z^+\ ,\ \mu_{k+1} \leftarrow \mu^+ = \mu(z^+)\ ,\ k \leftarrow k + 1.$
continue

4. Polynomial complexity

In this section we study the computational complexity of Algorithm 1. In the proof of our main theorem we will use the following technical result, which represents a slight improvement over the corresponding results of [6, 25].

Proposition 3. *If HLCP (1) is monotone and $z = \lceil x, s \rfloor \in \mathcal{D}(\beta)$ then the quantities h^i computed in (9) satisfy*

$$\zeta_i := \|h^i\|_2 \le \frac{2\beta\mu}{i}(\tau\sqrt{n}\,)^i\,,\ i = m+1,\ldots,2m\,, \tag{19}$$

where

$$\tau = \frac{2\sqrt{\beta(1+\epsilon-\gamma)^2 + (1-\beta)\gamma^2}}{\beta}\,. \tag{20}$$

Proof. First, we prove that the quantities

$$\eta_i = \left\| \, Du^i + D^{-1}v^i \, \right\|_2,\ D = X^{-1/2}S^{1/2}$$

satisfy the relations

$$\sqrt{\|\, Du^i \,\|_2^2 + \|\, D^{-1}v^i \,\|_2^2} \le \eta_i \le 2\alpha_i\sqrt{\beta\mu}\left(\frac{\tau\sqrt{n}}{4}\right)^i, \tag{21}$$

where

$$\alpha_1 = 1,\ \alpha_i = \frac{1}{i}\binom{2i-2}{i-1}, \quad i = 2,3,\ldots$$

The first inequality in (21) follows immediately, since by using (7) and the monotony of the HLCP we deduce that ${u^i}^T v^i \ge 0$. Hence

$$\left\| \, Du^i + D^{-1}v^i \, \right\|_2^2 = \left\| \, Du^i \, \right\|_2^2 + 2{u^i}^T v^i + \left\| \, D^{-1}v^i \, \right\|_2^2 \ge \left\| \, Du^i \, \right\|_2^2 + \left\| \, D^{-1}v^i \, \right\|_2^2 .$$

By multiplying the first equations of (7) with $(xs)^{-1/2}$ we obtain

$$\begin{aligned} Du^1 + D^{-1}v^1 &= -\left((1+\epsilon)(xs)^{1/2} - \gamma\mu(xs)^{-1/2}\right) \\ Du^2 + D^{-1}v^2 &= -\left(\epsilon(xs)^{1/2} - (xs)^{-1/2}u^1v^1\right) \\ Du^i + D^{-1}v^i &= -(xs)^{-1/2}\sum_{j=1}^{i-1} Du^j D^{-1}v^{i-j}\,, \quad 3 \le i \le m\,. \end{aligned}$$

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

$$\eta_1 = \left\| \, (1+\epsilon)(xs)^{1/2} - \gamma\mu(xs)^{-1/2} \, \right\|_2,\ \eta_2 = \left\| \, \epsilon(xs)^{1/2} - (xs)^{-1/2}u^1v^1 \, \right\|_2,$$

$$\eta_i \le \frac{1}{\sqrt{\beta\mu}}\sum_{j=1}^{i-1} \left\| \, Du^j \, \right\|_2 \left\| \, D^{-1}v^{i-j} \, \right\|_2,\ 3 \le i \le m\,.$$

We have

$$\begin{aligned} \eta_1^2 &= \left\| \, (1+\epsilon)(xs)^{1/2} - \gamma\mu(xs)^{-1/2} \, \right\|_2^2 = \sum_{j=1}^{n}\left((1+\epsilon)^2 x_js_j - 2(1+\epsilon)\gamma\mu + \frac{\gamma^2\mu^2}{x_js_j}\right) \\ &= ((1+\epsilon)^2 - 2(1+\epsilon)\gamma)\mu n + \gamma^2\mu^2\sum_{j=1}^{n}\frac{1}{x_js_j} \\ &\le \mu n\left((1+\epsilon)^2 - 2(1+\epsilon)\gamma + \frac{\gamma^2}{\beta}\right) = \frac{\beta\mu n\tau^2}{4}\,, \end{aligned}$$

which shows that the second inequality in (21) is satisfied for $i = 1$. Using the fact that ${u^1}^T v^1 \geq 0$ implies $\| u^1 v^1 \|_2^2 \leq \eta_1^4/8$ we deduce that

$$\eta_2^2 = \left\| \epsilon (xs)^{1/2} - (xs)^{-1/2} u^1 v^1 \right\|_2^2 = \sum_{j=1}^n \left(\epsilon^2 x_j s_j - 2\epsilon u_i^1 v_i^1 + \frac{(u_i^1 v_i^1)^2}{x_j s_j} \right)$$
$$\leq \epsilon^2 n\mu + \frac{\eta_1^4}{8\beta\mu} \leq \epsilon^2 n\mu + \frac{\beta\mu n^2 \tau^4}{128} .$$

We want to prove that the second inequality in (21) is satisfied for $i = 2$, i.e.,

$$\eta_2 \leq \frac{\sqrt{\beta\mu} n \tau^2}{8} . \tag{22}$$

This inequality holds provided $\epsilon^2 \leq \beta n \tau^4/128$. This is trivially satisfied for $\epsilon = 0$ and for $\epsilon = 1$ it reduces to

$$1 \leq \frac{n(\beta(2-\gamma)^2 + (1-\beta)\gamma^2)^2}{8\beta^3} .$$

Since the minimum over $0 \leq \beta, \gamma \leq 1$ of the right hand side above is attained for $\beta = \gamma = 1$ it follows that (22) is satisfied in case $\epsilon = 1$ whenever $n \geq 8$.

For $i \geq 3$ and $1 \leq j < i$, $j \neq i - j$ we have

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

In case $j = i - j$ there holds

$$\| Du^j \|_2 \| D^{-1} v^j \|_2 \leq \frac{1}{2} \left(\| Du^j \|_2^2 + \| D^{-1} v^j \|_2^2 \right) \leq \frac{1}{2} \eta_j^2 .$$

Therefore we obtain

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

Since $\alpha_i = \sum_{j=1}^{i-1} \alpha_j \, \alpha_{i-j}$, we deduce by induction that the second inequality in (21) holds for any $1 \leq i \leq m$. Finally, for any $m + 1 \leq i \leq 2m$, we have

$$\| h^i \|_2 \leq \sum_{j=i-m}^{m} \| Du^j \|_2 \| D^{-1} v^{i-j} \|_2 \leq \sum_{j=1}^{i-1} \| Du^j \|_2 \| D^{-1} v^{i-j} \|_2$$
$$= \frac{1}{2} \sum_{j=1}^{i-1} \left(\| Du^j \|_2 \| D^{-1} v^{i-j} \|_2 + \| Du^{i-j} \|_2 \| D^{-1} v^j \|_2 \right)$$
$$\leq \frac{1}{2} \sum_{j=1}^{i-1} \sqrt{\| Du^j \|_2^2 + \| D^{-1} v^j \|_2^2} \sqrt{\| Du^{i-j} \|_2^2 + \| D^{-1} v^{i-j} \|_2^2}$$

$$\le \frac{1}{2}\sum_{j=1}^{i-1}\eta_j\,\eta_{i-j} \le 2\beta\mu\left(\frac{\tau\sqrt{n}}{4}\right)^i\sum_{j=1}^{i-1}\alpha_j\,\alpha_{i-j}$$
$$= 2\beta\mu\left(\frac{\tau\sqrt{n}}{4}\right)^i\alpha_i \le \frac{2\beta\mu}{i}\left(\tau\sqrt{n}\right)^i\ ,$$

where we have used the fact that $\alpha_i \le \frac{1}{i}4^i$. The proof is complete. ∎
From the above proposition it follows that

Corollary 1. *If HLCP (1) is monotone and $z = \lceil x,\ s\rfloor \in \mathcal{D}(\beta)$ then the following relations hold for any $\alpha > 0$:*

$$\frac{\alpha}{\mu}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 < 1, \quad \forall\, 0 \le \theta \le \frac{1}{\tau\sqrt{n}}\min\left\{1\,,\ (1.4\,\alpha\beta)^{\frac{-1}{m+1}}\right\}\,, \tag{23}$$
$$\frac{\alpha}{\mu\sqrt{n}}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 < \theta\,, \quad \forall\, 0 \le \theta \le \frac{1}{\tau\sqrt{n}}\min\left\{1\,,\ (1.4\,\alpha\beta\tau\,)^{\frac{-1}{m}}\right\}\,. \tag{24}$$

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

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

Let us assume that $\theta \in (0,\ 1/(\tau\sqrt{n})]$. Using Proposition 3 we obtain

$$\frac{\alpha}{\mu}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 \le 2\alpha\beta\sum_{i=m+1}^{2m}\frac{1}{i}\left(\tau\sqrt{n}\theta\right)^i < 1.4\,\alpha\beta\left(\tau\sqrt{n}\theta\right)^{m+1}\,,$$

which, in turn, implies that

$$\frac{\alpha}{\mu\sqrt{n}}\sum_{i=m+1}^{2m}\theta^i\left\|h^i\right\|_2 < \frac{1.4\,\alpha\beta}{\sqrt{n}}\left(\tau\sqrt{n}\theta\right)^{m+1} = 1.4\,\alpha\beta\tau\theta\left(\tau\sqrt{n}\theta\right)^m\,.$$

The statement of our corollary follows now immediately from the above results. ∎
From the definition of τ (20) it follows that

$$\frac{2(1+\epsilon)\sqrt{1-\beta}}{\sqrt{\beta}} \le \tau \le 2\max\left\{\frac{1+\epsilon}{\sqrt{\beta}}\,,\ \frac{\sqrt{1-\beta+\beta\epsilon^2}}{\beta}\right\} < \frac{2(1+\epsilon)}{\beta}\,. \tag{25}$$

We note that the above bounds are sharp in the sense that the lower bound is attained for $\gamma = (1+\epsilon)\beta$, while the upper bound is attained either for $\gamma = 0$ or for $\gamma = 1$. In order to simplify proofs and to obtain estimates that are independent of γ in the analysis of the corrector step we will use the bound $\tau < 2/\beta$.

Theorem 1. *If HLCP (1) is monotone, then Algorithm 1 is well defined and the following relations hold for any integer $k \geq 0$:*

$$\overline{z}^{\,k}, z^k \in \mathcal{D}(\beta)\,,\ \mu_{k+1} \leq \left(1 - \frac{\chi}{n^{\frac{1}{2}+\frac{m_c+1}{2m_c(m_p+1)}}}\right)\overline{\mu}_{\,k}\ ,\ \overline{\mu}_{\,k+1} \leq \left(1 - \frac{\overline{\chi}}{n^{\frac{1}{2}+\upsilon}}\right)\overline{\mu}_{\,k}\,,$$

where χ, $\overline{\chi}$ are constants depending only on $\beta, \underline{\gamma}, \overline{\gamma}\,, \underline{\sigma}, \overline{\sigma}\,,$ and

$$\upsilon := \min\left\{\frac{1}{2m_c}\,,\ \frac{m_c+1}{2m_c(m_p+1)}\right\}. \tag{26}$$

Proof. **Analysis of the corrector.** On the corrector we have $\epsilon = 0$, $m = m_c$, $0 < \underline{\gamma} < \gamma < \overline{\gamma} < 1$, $0 < \underline{\sigma} < \sigma < \overline{\sigma} < 1$, and $\tau < 2/\beta$. First, we prove that if $z \in \overline{\mathcal{D}}(\beta)$, then the quantities θ_0 and θ_1, defined in (10) and (13), satisfy

$$\theta_0 \geq \theta_3 := \frac{\beta}{2\sqrt{n}}\left(\frac{1}{2.8}\right)^{\frac{1}{m_c+1}}\,, \tag{27}$$

$$\theta_1 \geq \theta_4 := \frac{\beta}{2\sqrt{n}}\left(\frac{(1-\overline{\sigma}\,)(1-\overline{\gamma}\,)}{2.8}\right)^{\frac{1}{m_c}}\,. \tag{28}$$

Using (23) with $\alpha = 2/\beta$, and the fact that $\theta_3 < 1/2$, it follows that for any $\theta \in [0,\,\theta_3]$ we have

$$\frac{x(\theta)s(\theta)}{\mu} > (1-\theta)\frac{xs}{\mu} + \frac{1}{\mu}\sum_{i=m+1}^{2m}\theta^i h^i \geq \frac{\beta}{2}e - \frac{1}{\mu}\sum_{i=m+1}^{2m}\theta^i\left\|\,h^i\,\right\|_2 e > 0\,.$$

Hence $x(\theta)s(\theta) > 0\,,\ \forall\,\theta \in [0\,,\theta_3]$. Since $x(0) > 0\,, s(0) > 0$, we can use a standard continuity argument to show that $x(\theta) > 0\,, s(\theta) > 0\,,\ \forall\,\theta \in [0\,,\theta_3]\,,$ which proves that $\theta_0 \geq \theta_3$. Similarly, by using (24) with $\alpha = 1/\left((1-\overline{\sigma}\,)(1-\overline{\gamma}\,)\right)$, we deduce that the following inequalities hold for any $\theta \in [0,\,\theta_4]$:

$$\begin{aligned}\frac{\mu(\theta) - (1-\sigma(1-\gamma)\theta)\mu}{\mu} &= \frac{1}{\mu\,n}\sum_{i=m+1}^{2m}\theta^i e^T h^i - (1-\sigma)(1-\gamma)\theta \\ \leq \frac{1}{\mu\sqrt{n}}\sum_{i=m+1}^{2m}\theta^i\left\|\,h^i\,\right\|_2 &- (1-\overline{\sigma}\,)(1-\overline{\gamma}\,)\,\theta \leq 0,\end{aligned}$$

which shows that $\theta_1 \geq \theta_4$.

Next, we show that if $z \in \mathcal{D}(\beta)$, then

$$f(\theta) \geq \beta + \frac{1}{2}(1-\beta)\gamma\,\theta \geq \beta + \frac{1}{2}(1-\beta)\underline{\gamma}\,\theta\,,\quad \forall\,\theta \in [0\,,\theta_5]\,, \tag{29}$$

where

$$\theta_5 := \min\left\{\theta_4\,,\ \frac{\beta}{2\,n^{\frac{1}{2}+\frac{1}{2m_c}}}\left(\frac{(1-\beta)\underline{\gamma}}{5.6}\right)^{\frac{1}{m_c}}\right\} \geq \frac{\chi_5}{n^{\frac{1}{2}+\frac{1}{2m_c}}}\,, \tag{30}$$

$$\chi_5 := \frac{\beta}{2}\left(\min\left\{\frac{(1-\overline{\sigma}\,)(1-\overline{\gamma}\,)}{2.8}\,,\ \frac{(1-\beta)\underline{\gamma}}{5.6}\right\}\right)^{\frac{1}{m_c}}\,. \tag{31}$$

According to (24), with $\tau = 2/\beta$ and $\alpha = 2\sqrt{n}/((1-\beta)\gamma)$, we have

$$\frac{1}{\mu}\sum_{i=m+1}^{2m}\theta^i\,\big\|\,h^i\,\big\|_2 < \frac{1}{2}(1-\beta)\gamma\,\theta\,,\quad \forall\theta \leq \frac{\beta}{2\sqrt{n}}\left(\frac{(1-\beta)\gamma}{5.6\sqrt{n}}\right)^{\frac{1}{m}}\,,$$

and (29) follows from the above inequality, (12), and the fact that

$$\mu(\theta) \leq (1-\sigma(1-\gamma)\theta)\mu \leq \mu\,,\ \forall\,\theta \in [0,\,\theta_4]\,.$$

Relation (29) shows that if $z \in \mathcal{D}(\beta)$, then the point $\overline{z}$ obtained in the corrector step of Algorithm 1 belongs to $\mathcal{D}(\beta+\delta)$, where

$$\delta = \frac{1}{2}(1-\beta)\gamma\,\theta_5\,. \tag{32}$$

As we mentioned before, the main purpose of the corrector is to increase proximity to the central path. However, it turns out that if the corrector step-length θ_c is large enough then we also obtain a significant reduction of the duality gap during the corrector step. In what follows we find a lower bound for θ_c in case the point $z \in \mathcal{D}(\beta)$ is not very well centered. More precisely we show that

$$\exists j \quad \text{such that } p_j := \frac{x_j s_j}{\mu} \leq \beta + .44\,\delta \ \Rightarrow\ \theta_c > .2\,\theta_5\,. \tag{33}$$

Let us denote

$$\lambda = .44\,\delta = .22(1-\beta)\gamma\,\theta_5\,,\ q^i = \frac{h^i}{\mu}\,,\ i = m+1,\ldots,2m\,.$$

For any $\theta \in [0,\,1]$ we have

$$\begin{aligned}
p_j(\theta) = \frac{x_j(\theta)s_j(\theta)}{\mu(\theta)} &= \frac{(1-\theta)p_j + \gamma\theta + \sum_{i=m+1}^{2m}\theta^i q^i_j}{(1-\theta) + \gamma\theta + \sum_{i=m+1}^{2m}\theta^i e^T q^i/n} \\
&< \frac{(1-\theta)(\beta+\lambda) + \gamma\theta + \sum_{i=m+1}^{2m}\theta^i q^i_j}{(1-\theta) + \gamma\theta + \sum_{i=m+1}^{2m}\theta^i e^T q^i/n} \\
&= \beta + \lambda + \frac{\gamma(1-\beta-\lambda)\theta - (\beta+\lambda)\sum_{i=m+1}^{2m}\theta^i e^T q^i/n + \sum_{i=m+1}^{2m}\theta^i q^i_j}{1-(1-\gamma)\theta + \sum_{i=m+1}^{2m}\theta^i e^T q^i/n} \\
&\leq \beta + \lambda + \frac{\gamma(1-\beta-\lambda)\theta + (1+\frac{\beta+\lambda}{\sqrt{n}})\sum_{i=m+1}^{2m}\theta^i\,\big\|\,q^i\,\big\|_2}{1-(1-\gamma)\theta - \frac{1}{\sqrt{n}}\sum_{i=m+1}^{2m}\theta^i\,\|\,q^i\,\|_2} \\
&\leq \beta + \lambda + \frac{\gamma(1-\beta)\theta + 2\sum_{i=m+1}^{2m}\theta^i\,\big\|\,q^i\,\big\|_2}{1-(1-\gamma)\theta - \sum_{i=m+1}^{2m}\theta^i\,\|\,q^i\,\|_2}\,.
\end{aligned}$$

Assume now that $\theta \in [0,\, .2\,\theta_5]$ and set $\theta = .2\phi$. Since $\phi \in [0,\, \theta_5]$, by virtue of (24), we can write

$$\sum_{i=m+1}^{2m} \theta^i \,\|\, q^i \,\|_2 = \sum_{i=m+1}^{2m} .2^i \phi^i \,\|\, q^i \,\|_2 \le\, .2^{m+1} \sum_{i=m+1}^{2m} \phi^i \,\|\, q^i \,\|_2$$
$$< \frac{.2^{m+1}}{2}\gamma(1-\beta)\phi = \frac{.2^m}{2}\gamma(1-\beta)\theta \le\, .1\gamma(1-\beta)\theta\,.$$

Using the fact that $\theta_5 < .5\,,\ \forall\, n \ge 1$, we obtain

$$p_j(\theta) < \beta + \lambda + \frac{1.2\gamma(1-\beta)\theta}{1-(1-\gamma+.1\,\gamma(1-\beta))\,\theta} \le \beta + \lambda + \frac{1.2\gamma(1-\beta)\theta}{1-\theta}$$
$$< \beta + \lambda + 1.4\gamma(1-\beta)\theta \le\, \beta + \lambda +\, .28\gamma(1-\beta)\theta_5 = \beta + \delta\,, \quad \forall\, \theta \in [0,\, .2\,\theta_5]\,.$$

It follows that $f(\theta_c) \ge \beta + \delta > \max_{0 \le \theta \le .2\,\theta_5} f(\theta)$, wherefrom we deduce that $\theta_c > .2\,\theta_5$.

Analysis of the predictor. In the predictor step we have $\gamma = 0$ and $m = m_p$, while ϵ can be either 0 or 1. Since the predictor step follows a corrector step, we have $z \in \mathcal{D}(\beta+\delta) \subset \mathcal{D}(\beta)$. In the predictor step we take $\gamma = 0$, so that from (20) we have $\tau = 2(1+\epsilon)/\sqrt{\beta} \le 4/\sqrt{\beta}$.

First, we study the behavior of the normalized duality gap in the predictor step. We start by proving that

$$(1-2.5\,\theta)\mu \le \mu(\theta) \le (1-.5\,\theta)\mu\,, \qquad (34)$$
$$\forall\, 0 \le \theta \le \theta_6 := \frac{\sqrt{\beta}}{4\sqrt{n}} \min\left\{1\,,\ \left(11.2\sqrt{\beta}\right)^{\frac{-1}{m_p}}\right\}\,.$$

We have

$$(1-2\theta)\mu + \sum_{i=m+1}^{2m} \theta^i (e^T h^i/n) \le \mu(\theta) \le (1-\theta)\mu + \sum_{i=m+1}^{2m} \theta^i (e^T h^i/n),$$

and the desired inequalities are obtained by noticing that (24), with $\alpha = 2$ and $\tau = 4/\sqrt{\beta}$, implies

$$\left|\sum_{i=m+1}^{2m} \theta^i (e^T h^i/n)\right| \le \frac{1}{\sqrt{n}} \sum_{i=m+1}^{2m} \theta^i \,\|\, h^i \,\|_2 < .5\,\theta\mu\,,$$

for all $\theta \in [0,\, \theta_6]$. Using Proposition 3 and the sum of a geometric series with ratio .1, we deduce that for any $\theta \in [0\,,\, \frac{\sqrt{\beta}}{40\sqrt{n}}]$ there holds

$$\mu'(\theta) = -(1+\epsilon-2\epsilon\theta)\mu + \sum_{i=m+1}^{2m} i\theta^{i-1}(e^T h^i/n) \le -\mu + \frac{1}{\sqrt{n}} \sum_{i=m+1}^{2m} i\theta^{i-1} \,\|\, h^i \,\|_2$$
$$\le -\mu + 8\mu\sqrt{\beta} \sum_{i=m}^{2m-1} \left(\frac{4\theta\sqrt{n}}{\sqrt{\beta}}\right)^i < -\mu + 8\mu\sqrt{\beta}\frac{.1^m}{1-.1} < -\mu + 8\mu\frac{.1^m}{.9} < 0\,.$$

Since $\theta_6 \geq \frac{\sqrt{\beta}}{44.8\sqrt{n}} > \frac{\sqrt{\beta}}{50\sqrt{n}}$, we conclude that

$$(1-2.5\,\theta)\mu \leq \mu(\theta) \leq (1-.5\,\theta)\mu \quad \text{and} \quad \mu'(\theta) < 0, \quad \forall\theta \in \left[0,\, \frac{\sqrt{\beta}}{50\sqrt{n}}\right]. \tag{35}$$

Next, we claim that the quantity θ_2 from (17) used in the computation of the predictor step-length satisfies

$$\theta_2 \geq \theta_7 := \frac{\sqrt{\beta}}{4\sqrt{n}} \min\left\{1\,,\left(11.2\sqrt{\beta}\right)^{\frac{-1}{m_p}}\,,\left(\frac{\delta}{2\beta}\right)^{\frac{1}{m_p+1}}\right\} \tag{36}$$

$$\geq \frac{\chi_7}{n^{\frac{1}{2}+\frac{m_c+1}{2m_c(m_p+1)}}}\,, \quad \chi_7 := \frac{1}{16}\left(\frac{(1-\beta)\underline{\gamma}\chi_5}{\beta}\right)^{\frac{1}{2}}. \tag{37}$$

Since $0 < \beta < 1$ and $n \geq 8$, (34) implies

$$\mu(\theta) \geq (1-2.5\theta_6)\mu \geq (1-\frac{2.5}{8\sqrt{2}})\mu \geq .7\mu,\ \forall\,\theta \in [0,\,\theta_6]\,.$$

By taking $\gamma = 0$, and $\beta+\delta$ instead of β, in (12), and using (23) with $\alpha = 1/(.7\delta)$, we deduce that

$$\begin{aligned} f(\theta) &\geq \beta + \delta - \frac{\sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2}{\mu(\theta)} \\ &\geq \beta + \delta - \frac{1}{.7\mu} \sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2 \geq \beta\,,\ \forall\,\theta \in [0,\,\theta_7]\,, \end{aligned}$$

which proves that $\theta_2 \geq \theta_7$. The second part of our claim follows by noticing that $\theta_5 \leq \theta_4 \leq \beta/(2\sqrt{n}) \leq \beta/(4\sqrt{2})$ and $11.2^{1/m_p} < 2(16\sqrt{2})^{1/(m_p+1)}$ implies

$$\begin{aligned} \theta_7 &\geq \frac{\sqrt{\beta}}{4\sqrt{n}} \min\left\{\left(\frac{1}{11.2}\right)^{\frac{1}{m_p}}\,,\left(\frac{\delta}{2\beta}\right)^{\frac{1}{m_p+1}}\right\} \\ &\geq \frac{\sqrt{\beta}}{4\sqrt{n}} \min\left\{\left(\frac{1}{11.2}\right)^{\frac{1}{m_p}}\,,\left(\frac{(1-\beta)\underline{\gamma}\theta_5}{4\beta}\right)^{\frac{1}{m_p+1}}\right\} \\ &\geq \frac{\sqrt{\beta}}{8\sqrt{n}}\left(\frac{(1-\beta)\underline{\gamma}\chi_5}{4\beta n^{\frac{m_c+1}{2m_c}}}\right)^{\frac{1}{m_p+1}} \geq \frac{\chi_7}{n^{\frac{1}{2}+\frac{m_c+1}{2m_c(m_p+1)}}}\,. \end{aligned}$$

Bounding the decrease of the duality gap. The duality gap clearly decreases during the corrector step, and the amount of decrease depends on the length of the corrector step-length, which in turn, as seen in (33), depends on the behavior of the preceding predictor step. Therefore a complete study of the decrease of the duality gap can be done by analyzing a succession of corrector-predictor-corrector steps. Assume that we are at iteration k and we have a point $z^k \in \mathcal{D}(\beta)$, with normalized duality gap μ_k. We follow the notations of Algorithm 1. The corrector step produces a point $\overline{z}^k \in \mathcal{D}(\beta+\delta)$, with δ given by (32). The corresponding normalized duality gap clearly satisfies $\overline{\mu}_k \le \mu_k$, but a bound on the decrease of the duality gap cannot be given at this stage. The corrector is followed by a predictor that produces a point $z^{k+1} = z(\theta_p) \in \mathcal{D}(\beta)$ with duality gap $\mu_{k+1} = \mu(\theta_p) = \min_{0\le\theta\le\theta_2} \mu(\theta)$. We have $\theta_7 \le \theta_2$ and $\theta_7 \le \theta_6$, so that according to (34)

$$\mu_{k+1} \le \mu(\theta_7) \le (1 - .5\,\theta_7)\,\overline{\mu}_k \le \left(1 - \frac{\chi_7}{2\,n^{\frac{1}{2}+\frac{m_c+1}{2m_c(m_p+1)}}}\right)\overline{\mu}_k\,, \quad \overline{\mu}_k \le \mu_k\,. \tag{38}$$

The above relation is sufficient for proving polynomial complexity, but it does not take into account the contribution of the corrector step. A finer analysis is needed in order to account for that. We distinguish two cases:

(a) $\theta_2 \ge \frac{\sqrt{\beta}}{50\sqrt{n}}$. According to (35), in this case we have

$$\mu_{k+1} = \min_{0\le\theta\le\theta_2} \mu(\theta)\mu\left(\frac{\sqrt{\beta}}{50\sqrt{n}}\right) \le \left(1 - \frac{\sqrt{\beta}}{100\sqrt{n}}\right)\overline{\mu}_k\,;$$

(b) $\theta_2 < \frac{\sqrt{\beta}}{50\sqrt{n}}$. In this case $\mu(\theta)$ is decreasing on the interval $[0\,,\,\theta_2]$, by virtue of (35), and by using (36) we deduce that $\theta_p = \theta_2, f(\theta_p) = \beta$. The latter equality must be true, since if $f(\theta_p) > \beta$, then, by a continuity argument, it follows that $\theta_2 > \theta_p$ which is a contradiction (see the definition of θ_2 (17)) . But if $f(\theta_p) = \beta$, then, according to (33), in the next corrector step we have $\theta_c > .2\,\theta_5$, so that

$$\overline{\mu}_{k+1} < (1 - .2\,\underline{\sigma}(1-\overline{\gamma})\theta_5)\,\mu_{k+1} \le \left(1 - \frac{\underline{\sigma}(1-\overline{\gamma})\chi_5}{5\,n^{\frac{1}{2}+\frac{1}{2m_c}}}\right)\mu_{k+1}\,.$$

In conclusion, for any $k \ge 0$ we have

$$\overline{\mu}_{k+1} \le \mu_{k+1} \le \left(1 - \frac{\sqrt{\beta}}{100\sqrt{n}}\right)\overline{\mu}_k\,,$$

or

$$\overline{\mu}_{k+1} < \left(1 - \frac{\underline{\sigma}(1-\overline{\gamma})\chi_5}{5\,n^{\frac{1}{2}+\frac{1}{2m_c}}}\right)\left(1 - \frac{\chi_7}{2\,n^{\frac{1}{2}+\frac{m_c+1}{2m_c(m_p+1)}}}\right)\overline{\mu}_k\,.$$

By taking

$$\overline{\chi} := \min\left\{\frac{\sqrt{\beta}}{100}\,,\,\frac{\underline{\sigma}(1-\overline{\gamma})\chi_5}{5}\,,\,\frac{\chi_7}{2}\right\}\,,$$

we deduce that

$$\overline{\mu}_{k+1} \le \left(1 - \frac{\overline{\chi}}{n^{\frac{1}{2}+\upsilon}}\right)\overline{\mu}_k\,, \quad k = 0, 1, \ldots$$

where υ is given by (26). The proof is complete ∎

As an immediate consequence of the above theorem we obtain the following complexity result:

Corollary 2. *Algorithm 1 produces a point* $z = \lceil x, s \rfloor \in \mathcal{D}(\beta)$ *with* $x^T s \le \varepsilon$*, in at most* $O\left(n^{1/2+\upsilon} \log\left(x^{0\,T} s^0/\varepsilon\right)\right)$ *iterations, where* υ *is given by (26).*

It follows that if the order of either the corrector or the predictor is larger than a multiple of $\log n$, then Algorithm 1 has $O\,(\sqrt{n}L)$-iteration complexity.

Corollary 3. *If* $\max\{m_c\,, m_p\} = \Omega(\log n)$*, then Algorithm 1 produces a point* $z = \lceil x, s \rfloor \in \mathcal{D}(\beta)$ *with* $x^T s \le \varepsilon$*, in at most* $O\left(\sqrt{n} \log\left(x^{0\,T} s^0/\varepsilon\right)\right)$*.*

Proof. Under the hypothesis of the corollary there is a constant ϑ, such that $\upsilon \le \vartheta/\log n$. Hence $n^{1/2+\upsilon} \le n^{\frac{\vartheta}{\log n}}\sqrt{n} = e^{\vartheta}\sqrt{n}$. ∎

In applications we can take $\max\{m_c\,, m_p\} = \lceil n^{\omega} \rceil$, for some $\omega \in (0,1)$. Since $\lim_{n\to\infty} n^{\frac{1}{n^{\omega}}} = 1$, this choice does not change the asymptotic constant related to the $O(.)$ notation in corollaries 2 and 3. For $\omega = 0.1$, the values of $\lceil n^{\omega} \rceil$ corresponding to $n = 10^6$, $n = 10^7$, $n = 10^8$, and $n = 10^9$ are 4, 6, 7, and 8 respectively.

5. Superlinear convergence

In this section we show that the duality gap of the sequence produced by Algorithm 1 is superlinearly convergent. More precisely we will prove that $\mu_{k+1} = O(\mu^{m_p+1})$ if the HLCP (1) is nondegenerate, and $\mu_{k+1} = O(\mu^{(m_p+1)/2})$ otherwise. The main ingredient of our proof is provided by the following lemma, which is an immediate consequence of the results of [18] about the analyticity of the central path:

Lemma 2. *If* $\gamma = 0$*, then the solution of (7) satisfies*

$$u^i = O(\mu^i), \quad v^i = O(\mu^i), \quad i = 1, \dots, m\,, \text{ if HLCP (1) is nondegenerate}\,,$$

and

$$u^i = O(\mu^{i/2}), \quad v^i = O(\mu^{i/2}), \quad i = 1, \dots, m\,, \text{ if HLCP (1) is degenerate}\,.$$

Theorem 2. *The sequence* μ_k *produced by Algorithm 1 satisfies*

$$\mu_{k+1} = O(\mu_k^{m_p+1})\,, \text{ if HLCP (1) is nondegenerate}\,,$$

and

$$\mu_{k+1} = O(\mu_k^{(m_p+1)/2})\,, \text{ if HLCP (1) is degenerate}\,.$$

Proof. For simplicity we denote $m = m_p$ and

$$\nu = \frac{m+1}{1+\epsilon} = \begin{cases} m+1\,, & \text{if HLCP (1) is nondegenerate}\,, \\ (m+1)/2\,, & \text{if HLCP (1) is degenerate}\,. \end{cases}$$

The superlinear convergence is due to the predictor step. In what follows we use the notation and the results from the proof Theorem 1 to analyze the asymptotic behavior of the predictor step. Since the predictor step follows a corrector step, we have $z \in \mathcal{D}(\beta+\delta)$. From Lemma 2 and equation (9) it follows that there is a constant ϖ such that

$$\sum_{i=m+1}^{2m} \theta^i \left\| h^i \right\|_2 \le \theta^{m+1}\varpi\mu^\nu \le \varpi\mu^\nu \,, \quad \forall\, \theta \in [\,0\,,\,1\,]\,.$$

It follows that

$$(1-\theta)^{1+\epsilon}\mu - \frac{\varpi}{\sqrt{n}}\,\mu^\nu \;\le\; \mu(\theta) \;\le\; (1-\theta)^{1+\epsilon}\mu + \frac{\varpi}{\sqrt{n}}\,\mu^\nu \,.$$

Using (12) with $\gamma = 0$ and $\beta+\delta$ instead of β we deduce that if $\mu^{\nu-1} < 1/\left(1/\delta + \varpi/\sqrt{n}\right)$, then for any $\theta \in [\,0\,,\,\theta_8\,]$, where

$$\theta_8 := 1 - \left(\left(\frac{1}{\delta} + \frac{\varpi}{\sqrt{n}}\right)\mu^{\nu-1}\right)^{\frac{1}{1+\epsilon}} \,, \tag{39}$$

we have

$$f(\theta) \ge \beta + \delta - \frac{\varpi\mu^\nu}{(1-\theta)^{1+\epsilon}\mu - \frac{\varpi}{\sqrt{n}}\mu^\nu} = \beta$$

It follows that $\theta_2 \ge \theta_8$, and therefore

$$\mu^+ = \min_{0\le\theta\le\theta_2} \mu(\theta) \le \mu(\theta_8) \le (1-\theta_8)^{1+\varepsilon}\mu + \frac{\varpi}{\sqrt{n}}\,\mu^\nu \le \left(\frac{1}{\delta} + \frac{2\varpi}{\sqrt{n}}\right)\mu^\nu \,. \tag{40}$$

The proof is complete. ∎

6. Line search procedures

The line search in both the corrector and the predictor step of the interior point method presented in Section 3 involves computing roots of higher order polynomials in one variable, for which no finite algorithms are known. In this section we revisit those line searches, and we propose a variant of the interior point method that uses line searches with polynomial computational complexity.

The corrector step we want to maximize $f(\theta)$ over the set

$$\Phi = \{\theta \in I\!R_+ \;:\; x(\theta) \ge 0,\, s(\theta) \ge 0,\, f(\theta) \ge \beta,\, \mu(\theta) \le (1-\sigma(1-\gamma)\theta)\mu\}\,.$$

Since $x(\theta)$ and $s(\theta)$ are polynomials of order m_c in θ, it follows that there are points

$$0 = \underline{\phi}_0 < \overline{\phi}_0 < \underline{\phi}_1 \le \overline{\phi}_1 < \cdots < \underline{\phi}_K \le \overline{\phi}_K, \text{ such that } \Phi = \bigcup_{l=0}^{K} \left[\,\underline{\phi}_l\,,\,\overline{\phi}_l\,\right]\,.$$

It turns out that the corrector step-size considered in Section 3 satisfies

$$\theta_c = \operatorname{argmax}\{f(\theta) \;:\; \theta \in \left[\,0\,,\,\overline{\phi}_0\,\right]\}\,.$$

For large values of m_c it is likely that $K > 0$, so that we may improve the performance of the corrector by eventually extending the line search beyond the interval $\left[\,0\,,\,\overline{\phi}_0\,\right]$. We do our line search over a discrete partition of the interval $\left[\,0\,,\,\overline{\phi}\,\right]$, where $\overline{\phi}$ is a parameter chosen by the user. Since in the skew-symmetric case we have $\mu(\theta) < 0$ for $\theta > 1/(1-\gamma)$, we should have $\overline{\phi} \le 1/(1-\gamma)$. A reasonable choice is $\overline{\phi} = 1$, although other values of $\overline{\phi}$ could also be considered. We construct a partition of the interval $\left[\,0\,,\,\overline{\phi}\,\right]$ of the form

$$\theta_5 = \tilde{\phi}_1 < \tilde{\phi}_2 < \cdots < \tilde{\phi}_{J_\phi} = \overline{\phi}\,, \quad \tilde{\phi}_j = \theta_5 \phi(j)\,,\; j = 2, \ldots, J_\phi - 1\,, \tag{41}$$

where θ_5 is defined in (30), and $\phi : [\,1\,,\,J_\phi\,] \bigcap I\!\!N \to I\!\!R$ is an increasing function satisfying the following properties

$$\phi(1) = 1,\; \phi(J_\phi - 1) < \frac{\overline{\phi}}{\theta_5} \le \phi(J_\phi),\; J_\phi = O(n^\alpha),\; \text{for some } \alpha > 0\,. \tag{42}$$

The last property ensures the polynomial complexity of the corrector line search which computes the step-length

$$\tilde{\theta}_c = \text{ argmax }_{\theta \in \tilde{\Phi}} f(\theta)\,, \tag{43}$$

where

$$\begin{aligned}\tilde{\Phi} = \{\ \tilde{\phi}_j\ &: 1 \le j \le J_\phi, \\ &x(\tilde{\phi}_j) > 0,\; s(\tilde{\phi}_j) > 0,\; f(\tilde{\phi}_j) \ge \beta,\; \mu(\tilde{\phi}_j) \le \Big((1 - \sigma(1-\gamma))\tilde{\phi}_j\Big)\,\mu\}\,.\end{aligned} \tag{44}$$

From the proof of Theorem 1 it follows that $\theta_5 = \tilde{\phi}_1 \in \tilde{\Phi}$, which guarantees the polynomial complexity of the resulting interior point method.

Since evaluating f at one point requires $O(nm_c)$ arithmetic operations it follows that the corrector line search can be done in $O(n^{1+\alpha}m_c)$ arithmetic operations. Here are some examples of functions satisfying (42).

– the uniform partition

$$\phi(j) = j,\; J_\phi = \left\lceil \overline{\phi}\,/\theta_5 \right\rceil = O(n^{\frac{1}{2} + \frac{1}{2m_c}})\,; \tag{45}$$

– the power graded partition

$$\phi(j) = j^\kappa,\; \text{for some } \kappa > 0,\; J_\phi = \left\lceil \left(\overline{\phi}\,/\theta_5\right)^{\frac{1}{\kappa}} \right\rceil = O(n^{\frac{1}{2\kappa} + \frac{1}{2\kappa m_c}})\,; \tag{46}$$

– the exponentially graded partition

$$\phi(j) = \varrho^{j-1},\; \text{for some } \varrho > 1,\; J_\phi = \left\lceil \frac{\log(\overline{\phi}\,/\theta_5)}{\log \varrho} + 1 \right\rceil = O(\log n)\,. \tag{47}$$

It follows that by taking $m_c = \Theta(\log n)$ the cost of the line search is $O(n^{\frac{3}{2}} \log n)$ arithmetic operations for the uniform partition, $O(n^{1+\frac{1}{2\kappa}} \log n)$ arithmetic operations for the power graded partition, and $O(n \log^2 n)$ arithmetic operations for the exponentially graded partition. Similarly, by taking $m_c = \Theta(n^{\omega})$, with $\omega \in (0, 1)$, the cost of the line search is $O(n^{\frac{3}{2}+\omega})$ arithmetic operations for the uniform partition, $O(n^{1+\frac{1}{2\kappa}+\omega} \log n)$ arithmetic operations for the power graded partition, and $O(n^{1+\omega} \log n)$ arithmetic operations for the exponentially graded partition. Of course, other functions ϕ, satisfying (42) can be considered to generate a partition of the interval $\left[\, 0\,,\, \overline{\phi}\, \right]$. For any such a partition the line search can be efficiently implemented by the following procedure.

Procedure 1
INPUT: $z \in \mathcal{D}(\beta),\ \mu.$ (* $\mu = \mu(z)$ *)
  set $\overline{\beta} = \beta$, $\overline{\theta} = 0$; (* $(\overline{\beta}, \overline{\theta})$ keeps track of $(\max f(\theta), \text{argmax } f(\theta))$ *)
  For $j = J_\phi, \ldots, 1$
    set $\xi = \infty$; (* ξ keeps track of $\min x(\tilde{\phi}_j)s(\tilde{\phi}_j)$ *)
    For $l = 1, \ldots, n$:
      If $\tilde{x}_l := x_l(\tilde{\phi}_j) \le 0$, **Return**;
      If $\tilde{s}_l := s_l(\tilde{\phi}_j) \le 0$, **Return**;
      If $\tilde{x}_l\tilde{s}_l < \xi$, set $\xi = \tilde{x}_l\tilde{s}_l$;
    End For;
    set $\tilde{\mu} = \mu(\tilde{z})$;
    If $\tilde{\mu} > (1 - \sigma(1-\gamma)\tilde{\phi}_j)\mu$, **Return**;
    If $\xi \le \overline{\beta}\,\tilde{\mu}$, **Return**;
    set $\overline{z} = \tilde{z}$, $\overline{\mu} = \tilde{\mu}$, $\overline{\beta} = \xi/\tilde{\mu}$, $\overline{\theta} = \tilde{\phi}_j$;
  End For;
OUTPUT $\overline{z} \in \mathcal{D}(\overline{\beta}),\ \overline{\mu},\ \overline{\theta}.$ (* $\overline{\beta} > \beta$, $\overline{\mu} = \mu(\overline{z})$, $\overline{\theta} = \tilde{\theta}_c$ from (43) *)

In the skew-symmetric case we have $\sigma = 1$ and $\mu(\theta) = (1 - (1-\gamma)\theta)\mu$, so that we can introduce the instruction "**If** $\xi \le \overline{\beta}\,(1-(1-\gamma)\tilde{\phi}_j)\mu$, **Return**" before finishing the computation of $z(\tilde{\phi}_j)$, since if $\xi \le \overline{\beta}\,(1-(1-\gamma)\tilde{\phi}_j)\mu$, then $f(\tilde{\phi}_j)$, will be less than the current maximum of f. Therefore in the skew-symmetric case we will use the following procedure:

Procedure 2
INPUT: $z \in \mathcal{D}(\beta),\ \mu.$ (* $\mu = \mu(z)$ *)
  set $\overline{\beta} = \beta$, $\overline{\theta} = 0$; (* $(\overline{\beta}, \overline{\theta})$ keeps track of $(\max f(\theta), \text{argmax } f(\theta))$ *)
  For $j = J_\phi, \ldots, 1$
    set $\xi = \infty$, $\tilde{\mu} = (1 - (1-\gamma)\tilde{\phi}_j)\mu$; (* $\xi \leftarrow \min x(\tilde{\phi}_j)s(\tilde{\phi}_j)$ *)
    For $l = 1, \ldots, n$:
      If $\tilde{x}_l := x_l(\tilde{\phi}_j) \le 0$, **Return**;
      If $\tilde{s}_l := s_l(\tilde{\phi}_j) \le 0$, **Return**;
      If $\tilde{x}_l\tilde{s}_l < \xi$, set $\xi = \tilde{x}_l\tilde{s}_l$;
      If $\xi < \overline{\beta}\,\tilde{\mu}$, **Return**;
    End For;

set $\overline{z} = \tilde{z}$, $\overline{\mu} = \tilde{\mu}$, $\overline{\beta} = \xi/\tilde{\mu}$, $\overline{\theta} = \tilde{\phi}_j$;
End For;
OUTPUT $\overline{z} \in \mathcal{D}(\overline{\beta})$, $\overline{\mu}$, $\overline{\theta}$. (* $\overline{\beta} > \beta$, $\overline{\mu} = \mu(\overline{z})$, $\overline{\theta} = \tilde{\theta}_c$ from (43) *)

The line search on the predictor has to be done in such a way as to preserve the superlinear convergence of our interior point method. We construct a partition of the interval $[\,0\,,\,1\,]$ of the form

$$\theta_7 = \tilde{\psi}_1 < \tilde{\psi}_2 < \cdots < \tilde{\psi}_{J_\psi - 1} < \tilde{\psi}_{J_\psi}\,, \quad \tilde{\psi}_j = \theta_7\psi(j)\,,\ j = 2, \ldots, J_\psi - 1\,,$$
$$\tilde{\psi}_{J_\psi} = \begin{cases} (1 + \tilde{\psi}_{J_\psi - 1})/2 & \text{if } \mu^{(\nu-1-\varsigma)/(1+\epsilon)} \geq 1 - \tilde{\psi}_{J_\psi - 1} \\ 1 - \mu^{(\nu-1-\varsigma)/(1+\epsilon)} & \text{if } \mu^{(\nu-1-\varsigma)/(1+\epsilon)} < 1 - \tilde{\psi}_{J_\psi - 1} \end{cases}, \tag{48}$$

where θ_7 is given by (36), $\varsigma \in (0, .5)$ is a given parameter and $\psi\,:\,[\,1\,,\,J_\psi\,] \bigcap I\!N \to I\!R$ is an increasing function satisfying the following properties

$$\psi(1) = 1,\ \psi(J_\psi - 1) < \frac{1}{\theta_7} \leq \psi(J_\psi),\ J_\psi = O(n^\alpha),\ \text{for some } \alpha > 0\,. \tag{49}$$

The above choice of $\tilde{\psi}_{J_\psi}$ is motivated by the fact that, according to the proof of Theorem 2, if $\mu^{\nu-1} < 1/\left(1/\delta + \varpi/\sqrt{n}\right)$, then

$$\theta_2 \geq \theta_8\,, \quad \text{and} \quad \mu(\theta) \leq (1-\theta)^{1+\epsilon}\mu + \frac{\varpi}{\sqrt{n}}\mu^\nu\,,\ \forall\,\theta \in [\,0\,,\,\theta_8\,]\,.$$

The constant ϖ is very difficult to estimate, but if $\mu^\varsigma < 1/\left(1/\delta + \varpi/\sqrt{n}\right)$, then

$$\tilde{\psi}_{J_\psi} \leq \theta_8\,,\ \mu(\tilde{\psi}_{J_\psi}) \leq (1 - \tilde{\psi}_{J_\psi})^{1+\epsilon}\mu + \frac{\varpi}{\sqrt{n}}\mu^\nu \leq \mu^{\nu-\varsigma} + \frac{\varpi}{\sqrt{n}}\mu^\nu \leq 2\mu^{\nu-\varsigma}\,. \tag{50}$$

This ensures superlinear convergence of order $\nu - \varsigma$ for the resulting interior-point method that computes the step-length of the predictor as

$$\tilde{\theta}_p = \ \text{argmin}_{\ \theta \in \tilde{\Psi}}\mu(\theta)\,, \tag{51}$$

where

$$\tilde{\Psi} = \left\{\tilde{\psi}_j\ : 1 \leq j \leq J_\psi,\ x(\tilde{\psi}_j) > 0,\ s(\tilde{\psi}_j) > 0,\ f(\tilde{\psi}_j) \geq \beta\right\}\,. \tag{52}$$

The line search can be efficiently implemented by the following procedure.

Procedure 3
INPUT: $z \in \mathcal{D}(\overline{\beta})$, μ, ς. (* $\overline{\beta} > \beta$, $\mu = \mu(z)$, $0 < \varsigma < .5$ *)
set $\mu^+ = \mu, \theta^+ = 0$; (* $(\mu^+, \theta^+) \leftarrow (\min \mu(\theta), \text{argmin}\ \mu(\theta))$ *)
For $j = J_\psi, \ldots, 1$
set $\xi = \infty$; (* $\xi \leftarrow \min x(\tilde{\psi}_j)s(\tilde{\psi}_j)$ *)
For $l = 1, \ldots, n$:
If $\tilde{x}_l := x_l(\tilde{\psi}_j) \leq 0$, **Return**;
If $\tilde{s}_l := s_l(\tilde{\psi}_j) \leq 0$, **Return**;
If $\tilde{x}_l\tilde{s}_l < \xi$, set $\xi = \tilde{x}_l\tilde{s}_l$;

End For;
set $\tilde{\mu} = \mu(\tilde{z})$;
If $\xi < \beta\tilde{\mu}$, **Return**;
If $\tilde{\mu} \geq \mu^+$, **Return**;
set $\mu^+ = \tilde{\mu}$, $\theta^+ = \tilde{\psi}_j$, $z^+ = z(\tilde{\psi}_j)$;
End For;
OUTPUT $z^+ \in \mathcal{D}(\beta),\ \mu^+,\ \theta^+$. (* $\mu^+ = \mu(z^+)$, $\overline{\theta} = \tilde{\theta}_p$ from (51) *)

In the skew-symmetric case we have $\mu(\theta) = (1-\theta)\mu$, so that we can slightly increase the efficiency of the above procedure, by eliminating early the points $\tilde{\psi}_j$ for which $(1-\tilde{\psi}_j)\mu$ is greater or equal to the current minimum of $\mu(\theta)$.

Procedure 4
INPUT: $z \in \mathcal{D}(\overline{\beta})$, μ, ς. (* $\overline{\beta} > \beta$, $\mu = \mu(z)$, $0 < \varsigma < .5$ *)
set $\mu^+ = \mu, \theta^+ = 0$; (* $(\mu^+, \theta^+) \leftarrow (\min \mu(\theta), \text{argmin}\ \mu(\theta))$ *)
For $j = J_\psi, \ldots, 1$
set $\xi = \infty$; (* $\xi \leftarrow \min x(\tilde{\psi}_j)s(\tilde{\psi}_j)$ *)
If $(1-\tilde{\psi}_j)\mu \geq \mu^+$, **Return**;
For $l = 1, \ldots, n$:
If $\tilde{x}_l := x_l(\tilde{\psi}_j) \leq 0$, **Return**;
If $\tilde{s}_l := s_l(\tilde{\psi}_j) \leq 0$, **Return**;
If $\tilde{x}_l\tilde{s}_l < \xi$, set $\xi = \tilde{x}_l\tilde{s}_l$;
End For;
set $\tilde{\mu} = \mu(\tilde{z})$;
If $\xi < \beta\tilde{\mu}$, **Return**;
set $\mu^+ = \tilde{\mu}$, $\theta^+ = \tilde{\psi}_j$, $z^+ = z(\tilde{\psi}_j)$;
End For;
OUTPUT $z^+ \in \mathcal{D}(\beta),\ \mu^+,\ \theta^+$. (* $\mu^+ = \mu(z^+)$, $\overline{\theta} = \tilde{\theta}_p$ from (51) *)

By using the line search procedures described above in Algorithm 1, we obtain the following interior point algorithm for solving monotone linear complementarity problems.

Algorithm 2
Given real parameters $0 < \beta < 1$, $0 < \underline{\gamma} < \overline{\gamma} < 1$, $0 < \underline{\sigma} < \overline{\sigma} < 1$, $0 < \varsigma < .5$, integers $m_c, m_p \geq 1$, and a vector $z^0 \in \mathcal{D}(\beta)$:
Set $k \leftarrow 0$;
repeat
(corrector step)
Set $z \leftarrow z^k$;
Choose $\gamma \in [\,\underline{\gamma}\,,\,\overline{\gamma}\,]$ and set $m = m_c$;
Compute directions $w^i = \lceil\, u^i,\ v^i \,\rfloor$, $i = 1, \ldots, m$, by solving (7);
If HLCP is skew-symmetric, compute $\overline{z}$, $\overline{\mu}$ by Procedure 2;
Else, choose $\sigma \in [\,\underline{\sigma}\,,\,\overline{\sigma}\,]$, compute $\overline{z}$, $\overline{\mu}$ by Procedure 1;
Set $\overline{z}^{\,k} \leftarrow \overline{z}$, $\overline{\mu}_k \leftarrow \overline{\mu}$.

(predictor step)
Set $z \leftarrow \overline{z}^{\,k}$, $\gamma = 0$, and $m = m_c$;
Compute directions $w^i = \lceil u^i, v^i \rfloor$, $i = 1, \ldots, m$, by solving (7);
If HLCP is skew-symmetric, compute z^+, μ^+ by Procedure 4;
Else, compute z^+, μ^+ by Procedure 3;
Set $z^{k+1} \leftarrow z^+$, $\mu_{k+1} \leftarrow \mu^+$, $k \leftarrow k+1$.
continue

Algorithm 2 has the same worst case iteration bounds as Algorithm 1.

Theorem 3. *If HCLP (1) is monotone then Algorithm 2 is well defined and the following relations hold for any integer $k \geq 0$:*

$$\overline{z}^{\,k}, z^k \in \mathcal{D}(\beta)\,,\ \mu_{k+1} \leq \left(1 - \frac{\chi}{n^{\frac{1}{2}+\frac{m_c+1}{2m_c(m_p+1)}}}\right)\overline{\mu}_k\,,\ \overline{\mu}_k \leq \left(1 - \frac{\overline{\chi}}{n^{\frac{1}{2}+\frac{1}{2m_c}}}\right)\mu_k\,,$$

where χ, $\overline{\chi}$ are constants depending only on $\beta, \underline{\gamma}, \overline{\gamma}, \underline{\sigma}, \overline{\sigma}$.

Proof. From the proof of Theorem 1 it follows that

$$f(\theta_5) \geq \beta + \delta,\ x(\theta_5) > 0,\ s(\theta_5) > 0,\ \mu(\theta_5) \leq (1 - \sigma(1-\gamma)\theta_5)\mu,$$

which shows that $\theta_5 \in \tilde{\Phi}$. Therefore the corrector step of Algorithm 2 produces a point $\overline{z} \in \mathcal{D}(\beta + \delta)$ such that

$$\overline{\mu} \leq (1 - \sigma(1-\gamma)\theta_5)\,\mu \leq \left(1 - \frac{\underline{\sigma}(1-\overline{\gamma})\chi_5}{n^{\frac{1}{2}+\frac{1}{2m_c}}}\right)\mu\,.$$

Similarly, from the proof of Theorem 1 we have $\theta_7 \in \tilde{\Psi}$, and therefore (38) holds. The proof is complete. ∎

Corollary 4. *Algorithm 2 produces a point $z = \lceil x, s \rfloor \in \mathcal{D}(\beta)$ with $x^T s \leq \varepsilon$, in at most $O\left(n^{1/2+\upsilon}\log\left(x^{0\,T} s^0/\varepsilon\right)\right)$ iterations, where υ is given by (26).*

It follows that if the order of either the corrector or the predictor is larger than a multiple of $\log n$, then Algorithm 1 has $O\,(\sqrt{n}L)$-iteration complexity.

Corollary 5. *If $\max\{m_c, m_p\} = \Omega(\log n)$, then Algorithm 2 produces a point $z = \lceil x, s \rfloor \in \mathcal{D}(\beta)$ with $x^T s \leq \varepsilon$, in at most $O\left(\sqrt{n}\log\left(x^{0\,T} s^0/\varepsilon\right)\right)$.*

The superlinear convergence of Algorithm 2 follows from (50). More precisely we have the following result.

Theorem 4. *The sequence μ_k produced by Algorithm 2 satisfies*

$$\mu_{k+1} = O(\mu_k^{m_p+1-\varsigma})\ ,\ \textit{if HLCP (1) is nondegenerate}\,,$$

and

$$\mu_{k+1} = O(\mu_k^{.5(m_p+1)-\varsigma})\ ,\ \textit{if HLCP (1) is degenerate}\,.$$

7. Conclusions

We have presented a corrector-predictor interior point algorithm for monotone homogenous linear complementarity problems acting in a wide neighborhood of the central path.

The corrector, based on a polynomial of order m_c, increases both centrality and optimality. Its numerical implementation requires one matrix factorization, $m_c + 1$ backsolves, and one line search. The line search procedure can be implemented in $O(m_c n^{1+\alpha})$ arithmetic operation, for some $0 < \alpha \leq 1$, or even in $O(m_c n \log n)$ arithmetic operations. In case of full matrices the cost of the factorization is $O(n^3)$ arithmetic operations, and the cost of a backsolve is $O(n^2)$ arithmetic operations. It follows that if $m_c = O(n^\omega)$, for some $0 < \omega < 1$, then the cost of the corrector step is dominated by the cost of the matrix factorization, so the corrector step can be implemented in at most $O(n^3)$ arithmetic operations.

The predictor step further increases optimality and eventually ensures superlinear convergence. It is based on a polynomial of order m_p. If $m_p = O(n^\omega)$, for some $0 < \omega < 1$, its numerical implementation requires at most $O(n^3)$ arithmetic operations, the cost being dominated by the cost of a matrix factorization.

If $\max\{m_c, m_p\} = \Omega(\log n)$, then the iteration complexity of the algorithms is $O(\sqrt{n}L)$. The algorithms are superlinearly convergent even for degenerate problems.

Acknowledgement. This paper is dedicated to Clovis Gonzaga, a master of the elegant proof, whose friendship I cherish.

References

1. M. Anitescu, G. Lesaja, and F. A. Potra. An infeasible–interior–point predictor–corrector algorithm for the P_*-Geometric LCP. *Applied Mathematics & Optimization*, 36(2):203–228, 1997.
2. K.M Anstreicher and R.A. Bosch. A new infinity-norm path following algorithm for linear programming. *SIAM J. Optim.*, 5(2):236–246, 1995.
3. J. F. Bonnans and C. C. Gonzaga. Convergence of interior point algorithms for the monotone linear complementarity problem. *Mathematics of Operations Research*, 21:1–25, 1996.
4. R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem.* Academic Press, Boston, MA, 1992.
5. C. C. Gonzaga. Complexity of predictor-corrector algorithms for LCP based on a large neighborhood of the central path. *SIAM J. Optim.*, 10(1):183–194 (electronic), 1999.
6. P-F. Hung and Y. Ye. An asymptotical $(O\sqrt{n}L)$-iteration path-following linear programming algorithm that uses wide neighborhoods. *SIAM Journal on Optimization*, 6(3):159–195, August 1996.
7. J. Ji, F. A. Potra, and S. Huang. A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence. *Journal of Optimization Theory and Applications*, 84(1):187–199, 1995.
8. M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, volume 538 of *Lecture Notes in Comput. Sci.* Springer-Verlag, New York, 1991.
9. S. Mizuno. A superlinearly convergent infeasible-interior-point algorithm for geometrical LCPs without a strictly complementary condition. *Math. Oper. Res.*, 21(2):382–400, 1996.

10. S. Mizuno, M. J. Todd, and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18(4):964–981, 1993.
11. R. D. C. Monteiro and S. J. Wright. Local convergence of interior-point algorithms for degenerate monotone LCP. *Computational Optimization and Applications*, 3:131–155, 1994.
12. R.C. Monteiro, I. Adler, and M.G. Resende. A polynomial-time primal-dual affine scalling algorithm for linear and convex quadratic programming and its power series extension. *Mathematics of Operations Research*, 15:191–214, 1990.
13. J. Peng, T. Terlaky, and Y. B. Zhao. A self-regularity based predictor-corrector algorithm for linear optimization. AdvOL-Report, 2002/4, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada, September 2002, to appear in SIAM J. on Optimization, 2005 .
14. 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. Programming*, 100:317–337, 2004.
15. F. A. Potra and R. Sheng. A large-step infeasible–interior–point method for the P_*-matrix LCP. *SIAM Journal on Optimization*, 7(2):318–335, 1997.
16. F. A. Potra and R. Sheng. Predictor-corrector algorithms for solving P_*-matrix LCP from arbitrary positive starting points. *Mathematical Programming, Series B*, 76(1):223–244, 1997.
17. F. A. Potra and R. Sheng. Superlinearly convergent infeasible–interior–point algorithm for degenerate LCP. *Journal of Optimization Theory and Applications*, 97(2):249–269, 1998.
18. J. Stoer, M. Wechs, and S. Mizuno. High order infeasible-interior-point methods for solving sufficient linear complementarity problems. *Math. Oper. Res.*, 23(4):832–862, 1998.
19. Josef Stoer. High order long-step methods for solving linear complementarity problems. *Ann. Oper. Res.*, 103:149–159, 2001. Optimization and numerical algebra (Nanjing, 1999).
20. J. F. Sturm. Superlinear convergence of an algorithm for monotone linear complementarity problems, when no strictly complementary solution exists. *Mathematics of Operations Research*, 24:72–94, 1999.
21. S. J. Wright. *Primal–Dual Interior–Point Methods*. SIAM Publications, Philadephia, 1997.
22. Y. Ye. *Interior Point Algorithms : Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.
23. Y. Ye and K. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP. *Mathematical Programming*, 62(3):537–551, 1993.
24. Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. *Mathematical Programming*, 59(2):151–162, 1993.
25. G. Zhao. Interior point algorithms for linear complementarity problems based on large neighborhoods of the central path. *SIAM J. Optim.*, 8(2):397–413 (electronic), 1998.