A predictor-corrector algorithm for linear optimization based on a specific self-regular proximity function

J.M. Peng*, T. Terlaky* and Y.B. Zhao†

May 22, 2003

Abstract. It is well known that the so-called first-order predictor-corrector methods working in a large neighborhood of the central path are among the most efficient interior-point methods (IPMs) for linear optimization (LO) problems. However, the best known iteration complexity of this type of methods is $O\left(n \log \frac{(x^0)^T s^0}{\varepsilon}\right)$. It is of interests to investigate whether the complexity of the first-order predictor-corrector type methods can be further improved. In this paper, based on a specific self-regular proximity function, we define a new neighborhood of the central path. In particular, we show that the new neighborhood matches the standard large neighborhood that is defined by the infinity norm and widely used in the IPM literature. A new first-order predictor-corrector method for LO that uses a self-regularity induced search direction in the corrector steps is proposed. We prove that our predictor-corrector algorithm, working in a large neighborhood, has an $O\left(\sqrt{n} \log n \log \frac{(x^0)^T s^0}{\varepsilon}\right)$ iteration bound. Local quadratic convergence of the algorithm is also established.

Key words. Linear Optimization, Interior-Point Methods (IPMs), Self-Regular Proximity Function, Large Neighborhoods, Polynomial Complexity

*Department of Computing and Software, McMaster University, Hamilton, Ontario L8S 4L7, Canada (Email: pengj@mcmaster.ca, terlaky@mcmaster.ca). The research of the first author was partially supported by the grant # RPG 249635-02 of the National Sciences and Engineering Research Council of Canada (NSERC) and a PREA award. The research of the second author was supported by NSERC grant #RPG 0048923. This research was also supported by the MITACS project "New Interior Point Methods and Software for Convex Conic-Linear Optimization and Their Application to Solve VLSI Circuit Layout Problems" and an FPP grant from IBM Watson Research Laboratory.

†The Fields Institute for Research in Mathematical Sciences, University of Toronto, Toronto, Ontario M5T 3J1, Canada / Institute of Applied Mathematics, AMSS, Chinese Academy of Science, Beijing, China (Email: ybzhao@amss.ac.cn). The research of the third author was partially supported by grant NSFC 10201032 of the Natural Science Foundation of China.

1 Introduction

To begin, we first clarify the notation used in this paper. R^n denotes the n-dimensional Euclidean space. $x \geq 0$ ($x > 0$) means all components of x are nonnegative (positive). e denotes the vector with all components equal to 1. For any positive vectors $x = (x_1, ..., x_n)^T$ and $s = (s_1, ..., s_n)^T \in R^n$, the symbols $xs, \sqrt{x}$ and x^r where r is a real number denote, respectively, the vectors whose components are $x_i s_i, \sqrt{x_i}$ and x_i^r $(i = 1, ..., n)$. In particular, we denote by x^{-1} the vector $(\frac{1}{x_1}, ..., \frac{1}{x_n})^T$.

Started from Karmarkar's seminal paper [7], a large body of research has been done in the field of IPMs for LO problems. Several recent books on this subject [18, 23, 24] contain substantial materials about various aspects of IPMs for LO. In the present paper, we consider the standard LO problem that takes the following form:

$$\min\{c^T x : \ Ax = b, \ x \geq 0\}.$$

Here $x \in R^n, c \in R^n, b \in R^m$, and A is an $m \times n$ matrix satisfying $\text{Rank}(A) = m$. Throughout the whole paper, we also assume that there exists a strictly feasible primal-dual pair (x^0, s^0, y^0) such that

$$Ax^0 = b, \ A^T y^0 + s^0 = c, \ (x^0, s^0) > 0.$$

In the present paper, among various IPMs for LO, we focus on the so-called first-order predictor-corrector IPMs that work in a wide neighborhood of the central path.

We start with a brief review of predictor-corrector type methods. The idea of predictor-corrector type algorithms is very natural and elegant. These algorithms follow the central path by alternatively taking predictor steps and corrector steps. A predictor step aims at reducing the duality gap as much as possible, thus the resulting iterate might move to the boundary of the feasible set, deviating from the central path, while a corrector step tries to bring the iterate back to a certain neighborhood of the central path. The best known representative of this class of methods is the Mizuno-Todd-Ye (MTY) algorithm for LO [10], which works in a small neighborhood of the central path. The quadratic convergence of the duality gap was proved in [26], and convergence of the iterates generated by the algorithm was proved in [4]. The MTY method was later extended to complementarity problems by Ye and Andersen [25] and to semidefinite optimization by Luo, Sturm and Zhang [9]. According to numerical experiments, IPMs working in a large neighborhood of the central path perform much better than their counterparts in small neighborhoods. Several authors have investigated IPMs that work in certain large neighborhoods of the central path [1, 3, 5, 8, 14, 15, 16, 21, 22, 28, 29]. It should be noted that the so-far best known worst-case complexity for the predictor-corrector algorithms with large neighborhoods is $O(n \log \frac{(x^0)^T s^0}{\epsilon})$, which is higher than that of those small neighborhood based algorithms. As remarked by Renegar[17], this becomes one of the ironies of the interior-point literature that algorithms which are more efficient in practice often have somewhat-worse complexity bounds.

Several strategies have already been proposed for improving the theoretical complexity of IPMs in large neighborhoods. For example, by using higher-order approximation of the central path, the iteration complexity of IPMs can become arbitrarily close to $O(\sqrt{n}\log\frac{(x^0)^T s^0}{n})$ [5, 6, 11, 16, 19, 20, 27, 28]. In IPMs based on high-order approximation, we usually need to solve multiple linear systems. Recently, Peng, Roos and Terlaky [12] proposed a class of IPMs in large neighborhoods, and proved that such an algorithm has an $O\left(\sqrt{n}\log n\log\frac{(x^0)^T s^0}{\varepsilon}\right)$ iteration bound. The approach in [12] is based on the notion of self-regular functions. One merit of the self-regular approach is that IPMs based on self-regular functions solve only one linear system at each step. Inspired by the results in [12], we consider in the present paper the issue whether the iterations bound of first-order predictor-corrector type IPMs within a large neighborhood can be improved. This is a far from easy task, as pointed in [16], the predictor-corrector methods of MTY type are much more difficult to develop and analyze in a wide neighborhoods of the central path. The difficulty for developing MTY-type methods within large neighborhoods comes from two points. First, theoretically, corrector steps based on the classical first order centering direction (first order correctors) are rather inefficient in a large neighborhood. A remedy for this point is to employ the search direction based on self-regular proximity functions introduced in [12]. However, this gives rise to another issue, namely whether the neighborhood defined by a self-regular proximity function matches the large neighborhoods used in the IPM literature. Secondly, once we use a new search direction in the corrector step, can we still obtain the quadratic convergence for the new algorithm? The purpose of this paper is to address the above-mentioned questions and show that the iteration bound of first-order predictor-corrector methods within a large neighborhood can be improved to $O\left(\sqrt{n}\log n\log\frac{(x^0)^T s^0}{\varepsilon}\right)$. To this end, a proximity measure function, induced by a specific self-regular function introduced in [12], will be explored. We then use this proximity function to define a new neighborhood of the central path and this new neighborhood matches the neighborhood $\mathcal{N}_\infty^-$ that has been widely used in most practical implementations of IPMs. Correspondingly, we also use the self-regularity induced search direction in the corrector steps.

This paper is organized as follows. In Section 2, we introduce the new neighborhood of the central path induced by a self-regular proximity function and discuss the properties of the proximity function. In Section 3, we describe the algorithm and establish $O\left(\sqrt{n}\log n\log\frac{(x^0)^T s^0}{\varepsilon}\right)$ complexity of the algorithm. Local quadratic convergence of the algorithm is proved in Section 4. Finally, we close the paper by a few concluding remarks.

2 Proximity functions and neighborhoods

This section consists of two parts. In the first subsection, we compare various large neighborhoods and proximity functions defining the neighborhoods. In particular,

we discuss a specific self-regular proximity function and show that the neighborhood induced by this proximity function matches the large neighborhood defined by infinity norm as used in many practical implementations of IPMs. In the second subsection, we explore some properties of the proximity function.

2.1 Large neighborhoods and proximity functions

We note that under the interior-point assumption, for every given parameter $\mu > 0$ the following system

$$\begin{aligned} Ax &= b,\ x > 0, \\ A^T y + s &= c,\ s > 0, \\ xs &= \mu e, \end{aligned} \tag{1}$$

has a unique solution denoted by $(x(\mu), s(\mu), y(\mu))$. The union of solutions $\{x(\mu) : \mu > 0\}$ forms the so-called primal central path, while $\{(y(\mu), s(\mu)) : \mu > 0\}$ is called the central path of the dual problem in the dual space. The iterates, generated by IPMs, are confined to be in certain neighborhoods of the central path, which are defined by certain proximity functions.

To describe these proximity functions, we need to introduce some notation. For any strictly feasible primal-dual pair (x, s) and any positive number μ, we define

$$v := \sqrt{\frac{xs}{\mu}}, \quad v^{-1} := \sqrt{\frac{\mu e}{xs}} \tag{2}$$

to be the vectors whose i^{th} components are $\sqrt{\frac{x_i s_i}{\mu}}$ and $\sqrt{\frac{\mu}{x_i s_i}}$, respectively. A typical proximity function can be defined by:

$$\Phi(x, s, \mu) := \Psi(v) = \sum_{i=1}^{n} \psi(v_i), \tag{3}$$

where $\psi(t)$ is a so-called kernel function. In the definition of neighborhoods, the parameter μ is usually associated with the present duality gap

$$\mu_{gap} = \frac{x^T s}{n}.$$

A typical neighborhood of the central path is usually defined by

$$\mathcal{N} := \{(x, s) > 0 | Ax = b, A^T y + s = c, \Phi(x, s, \mu_{gap}) \leq \eta(\tau, n)\}, \tag{4}$$

where $\eta(\tau, n)$ is a function of a parameter τ and the dimension of the underlying problem n.

There are many choices for the proximity function $\Phi(x, s, \mu)$ and the function $\eta(\tau, n)$ in (4). For instance, if we choose $\Phi(x, s, \mu) = \|v^2 - e\|$ and $\eta(\tau, n) = \rho$ where

$\rho \in (0,1)$ (i.e., $\eta(\tau,n) = 0.25$), then we obtain a so-called small neighborhood. If we choose $\Phi(x,s,\mu) = \|v^2 - e\|_\infty$ (or $\Phi(x,s,\mu) = \|(v^2-e)^-\|_\infty$ with $w^- = \min(w,0)$) and $\eta(\tau,n) = \rho$ with $\rho \in (0,1)$, then we get a so-called large neighborhood. Let us denote the neighborhoods generated via using the proximity functions $\|v^2 - e\|_\infty$ and $\|(v^2-e)^-\|_\infty$ by $\mathcal{N}_\infty(\rho)$ and $\mathcal{N}_\infty^-(\rho)$, respectively. It worths pointing out that the neighborhood $\mathcal{N}_\infty^-(\rho)$ is frequently used in the practical implementation of IPMs, although the corresponding proximity function does not have a barrier property. On the other hand, as observed by many researchers, various IPMs are tightly associated with some proximity functions that have some barrier and coercive properties to prevent the iterate from moving to infinity and to the boundary of the feasible set. For instance, Peng, Roos and Terlaky [12] introduced a large class of proximity functions based on the notion of self-regular functions. For self-regular proximity functions, if we still choose $\eta(\tau,n)$ as a small constant, then the region defined by (4) becomes a small neighborhood. If $\eta(\tau,n) = \tau n = \mathcal{O}(n)$, then we obtain a large neighborhood.

A natural question is whether a large neighborhood defined by a self-regular proximity function can match the neighborhood $\mathcal{N}_\infty^-(\rho)$. It should be noted that the key in the definition of a neighborhood (4) is the relation $\Phi(x,s,\mu_{gap}) \le \eta(\tau,n)$. Let $\mathcal{I}$ be the index set $\mathcal{I} = \{1,\cdots,n\}$. We consider a specific proximity function defined by

$$\Phi_{lb}(x,s,\mu) = \Psi_{lb}(v) = \frac{1}{2}\|v\|^2 - \frac{1}{2}n - \sum_{i\in\mathcal{I}} \log v_i,$$

whose kernel function is the classical logarithmic barrier function

$$\psi_{lb}(t) = \frac{1}{2}t^2 - \frac{1}{2} - \log t.$$

This proximity function has been widely used in the analysis of IPMs (see [12, 18] and the reference therein). It is well-known that the function $\Phi_{lb}(x,s,\mu)$, as a function of μ, has an optimal value at the point $\mu = \mu_{gap}$. Therefore, we have

$$\Phi_{lb}(x,s,\mu_{gap}) = -\frac{1}{2}\sum_{i\in\mathcal{I}} \log \frac{nx_is_i}{x^Ts}.$$

If we choose $\eta(\tau,n) = \frac{\rho n}{2(1-\rho)}$, then we can write the inequality used in the definition (4) of the neighborhood as

$$-\sum_{i\in\mathcal{I}} \frac{1}{2}\log \frac{nx_is_i}{x^Ts} \le \frac{\rho n}{2(1-\rho)}. \tag{5}$$

Note that if the primal-dual pair (x,s) is in $\mathcal{N}_\infty^-(\rho)$, then for a fixed pair (x,s), from the optimality condition and the convexity of $\Phi(x,s,\mu)$ we can conclude

$$\begin{aligned}\Phi_{lb}(x,s,\mu_{gap}) &= -\sum_{i\in\mathcal{I}} \frac{1}{2}\log \frac{nx_is_i}{x^Ts} \le -\frac{n}{2}\log(1-\rho) \\ &= \frac{n}{2}\log\left(1 + \frac{\rho}{1-\rho}\right) \le \frac{\rho n}{2(1-\rho)},\end{aligned} \tag{6}$$

where the last inequality follows from

$$\log(1+t) \leq t, \quad \forall t \geq -1.$$

The relation (6) indicates that, if the kernel function $\psi_{lb}(t)$ is used and $\eta(\tau, n) = \frac{\rho n}{2(1-\rho)}$, then the neighborhood $\mathcal{N}_\infty^-(\rho)$ is usually smaller than the neighborhood $\mathcal{N}$ defined by (4). It is worthwhile to consider how much is the difference between these two neighborhoods. Let us consider the following example with $x = e$ and $s = (1, \ldots, 1, \frac{1}{2n})^T$. In this case one has

$$\mu_{gap} = 1 - \frac{1}{n} + \frac{1}{2n^2}.$$

Suppose that $\rho = \frac{3}{4}$. For sufficiently large n that satisfies

$$1 - \frac{n}{2n^2 - 2n + 1} > \frac{3}{4},$$

one can show that $(x, s) \notin \mathcal{N}_\infty^-(3/4)$. On the other hand, one can also verify easily that $\Phi_{lb}(x, s, \mu_{gap}) \leq \frac{3n}{2}$. This implies that, in some extreme cases, the neighborhood $\mathcal{N}$ defined by (4) with $\psi_{lb}(t)$ is too large and significantly different from the neighborhood $\mathcal{N}_\infty^-(\rho)$ as used in the practical implementation of IPMs. This explains partially why the existing theoretical complexity results of IPMs with large neighborhoods are not as good as what is observed in practice.

Now let us recall the definition of self-regular functions [12].

Definition 2.1 *A function $\psi(t) \in C^2 : (0, \infty) \to R$ is said be self-regular if it satisfies the following conditions:*
(i) $\psi(t)$ is strictly convex with respect to $t > 0$ and vanished at $t = 1$, i.e., $\psi(1) = \psi'(1) = 0$. Further, there exists positive constants $\nu_2 \geq \nu_1 > 0$ and $p \geq 1, q \geq 1$ such that

$$\nu_1(t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq (t^{t-1} + t^{-1-q}), \ \forall t \in (0, \infty);$$

(ii) For any $t_1, t_2 > 0$,

$$\psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1-t)\psi(t_2), \ r \in [0, 1].$$

As showed in [12], among various large-update IPMs based on self-regular proximity functions, the algorithm with a barrier degree $q = O(\log n)$ has the best worst-case complexity result. This motivates us to consider the kernel following function:

$$\psi_1(t) = \frac{t^2 - 1}{2} + \frac{t^{-\log n} - 1}{\log n} = \frac{t^2 - 1}{2} + \frac{n^{-\log t} - 1}{\log n}. \tag{7}$$

The corresponding proximity function becomes

$$\Phi_1(x, s, \mu) = \Psi_1(v) = \frac{1}{2}\|v\|^2 - \frac{n}{2} + \sum_{i \in \mathcal{I}} \frac{n^{-\log v_i} - 1}{\log n}. \tag{8}$$

Because this proximity function has an extremely strong barrier property, we need to change correspondingly the function $\eta(\tau, n)$ in (4) so that the neighborhood given by (4) can match the neighborhood $\mathcal{N}_\infty^-(\rho)$ with $\rho \in (0,1)$. Note that if $(x,s) \in \mathcal{N}_\infty^-(\rho)$ and $\mu = \mu_{gap}$, then from (8) we obtain

$$\Phi_1(x, s, \mu_{gap}) = \Psi_1(v) = \sum_{i \in \mathcal{I}} \frac{n^{-\log v_i} - 1}{\log n} < \frac{n^{1-\log(1-\rho)} - n}{\log n}.$$

This inequality implies that if we choose $\tau = 1 - \log(1-\rho)$ and define

$$\eta(\tau, n) := \frac{n^\tau - n}{\log n}, \tag{9}$$

then the neighborhood $\mathcal{N}$ defined by (4) contains the neighborhood $\mathcal{N}_\infty^-(\rho)$. It worths mentioning that $\eta(\tau, n)$ is an increasing function of τ. In the rest of this paper, $\eta(\tau, n)$ is defined by (9) and thus our algorithm will operate in a large neighborhood.

2.2 Properties of the proximity function

In this section we investigate various properties of the proximity function $\Phi_1(x, s, \mu)$. We first consider the case that $\mu = \mu_{gap}$. Recall that the value of $\Phi_1(x, s, \mu_{gap})$ is essentially determined by the sum $\sum_{i \in \mathcal{I}} n^{-\log v_i}$. To estimate this sum, we need the following technical result.

Lemma 2.2 *Suppose that* $\log n > 2$. *Then one has*

$$\frac{n^{\log t} - 1}{\log n} \geq \frac{1}{2}\big(t^2 - 1\big), \quad \forall t > 0.$$

Proof: To prove the lemma, we define

$$f(t) = \frac{n^{\log t} - 1}{\log n} - \frac{1}{2}\big(t^2 - 1\big) = \frac{t^{\log n} - 1}{\log n} - \frac{1}{2}\big(t^2 - 1\big).$$

By simple calculus, one gets

$$f'(t) = t^{\log n - 1} - t.$$

Because $\log n > 2$, one can readily verify that $f'(t) < 0$ when $t < 1$ and $f'(t) > 0$ if $t \geq 1$. Thus $f(t)$ has a global minimum $f(1) = 0$ at the point $t = 1$. We thus obtain the inequality in the lemma as desired. □

It follows from Lemma 2.2 that

$$n^{-\log t} \geq \frac{\log n}{2}\big(t^{-2} - 1\big) + 1, \quad \forall t > 0.$$

This relation, and the fact that $\|v\| = n$ whenever $\mu = \mu_{gap}$ imply

$$\sum_{i\in\mathcal{I}} n^{-\log v_i} \geq \frac{\log n}{2}\Big(\big\|v^{-1}\big\|^2 - n\Big) + n = \frac{\log n}{2}\big\|v - v^{-1}\big\|^2 + n \geq n,$$

and equality holds if and only if $v = e$.

We proceed to explore more properties of the proximity function $\Phi_1(x, s, \mu)$. For fixed $(x, s) > 0$, let θ^* denote the point at which the proximity function $f(\theta) = \Phi_1(x, s, \theta\mu_{gap})$ assume its global minimum. We also assume that there exists a positive number $\tau > 1$ such that

$$\Phi_1(x, s, \mu_{gap}) = \eta(\tau, n) = \frac{n^\tau - n}{\log n}.$$

Now it follows from the optimality condition of $f(\theta)$ that

$$\frac{n}{2(\theta^*)^2} - \frac{n^{\frac{1}{2}\log\theta^*}}{2\theta^*}\sum_{i\in\mathcal{I}} n^{-\log v_i} = 0. \tag{10}$$

From this equality and the fact $\sum_{i\in\mathcal{I}} n^{-\log v_i} = n^\tau$ we obtain

$$n^{1-\tau} = \theta^* n^{\frac{1}{2}\log\theta^*},$$

which yields

$$(1-\tau)\log n = \log\theta^*\Big(1 + \frac{1}{2}\log n\Big).$$

Thus we have

$$\theta^* = \exp^{-\frac{(2\tau-2)\log n}{2+\log n}}. \tag{11}$$

This gives one of the main results of this section, which follows directly from the choice (11) of θ^*.

Lemma 2.3 *Suppose that $\Phi_1(x, s, \mu_{gap}) = \eta(\tau, n)$ for some $\tau > 1$ where $\eta(\tau, n)$ is given by (9), and θ^* is given by (11). Then the proximity function $\Phi_1(x, s, \mu)$ attains its global minimum at the point*

$$\mu^* = \theta^*\mu_{gap}.$$

It worths exploring the relation between μ_{gap} and μ^*, which is dominated by θ^*. If $n \geq \exp^2$, then from (11) we obtain

$$\exp^{2-2\tau} \leq \exp^{(1-\tau)\frac{2\log n}{2+\log n}} \leq \exp^{1-\tau}.$$

This shows that for sufficiently large n, θ^* is uniformly bounded above and below and these bounds depend only on τ. Since we shall use this $\mu = \theta^*\mu_{gap}$ update of the parameter μ in our algorithm, therefore, we cast θ^* as a function of τ, i.e.,

$$\theta^* := \theta(\tau) = \exp^{-\frac{(2\tau-2)\log n}{2+\log n}}. \tag{12}$$

When n is sufficiently large, we can easily see that we have $\theta^* \cong \exp^{2-2\tau}$. Take for example, if $\tau = 2.5$ and n is sufficiently large (saying $n \geq 100$), then one has $\theta^* \leq 0.1$. This indicates that for sufficiently large n, if $\Phi_1(x, s, \mu_{gap}) = \eta(\tau, n)$ and we use $\mu = \theta(\tau)\mu_{gap}$ (which is indeed a large update) as our targeted μ in the algorithm, then the proximity function $\Phi_1(x, s, \mu)$ will not increase after the update. It is straightforward to verify that $\theta(\tau)$ is a decreasing function of τ.

For fixed $(x, s) > 0$ and $\theta(\tau)$ given by (12), let us denote by $\eta_1(\tau, n)$ the value of the proximity function $\Phi_1(x, s, \theta(\tau)\mu_{gap})$ when $\Phi_1(x, s, \mu_{gap}) = \eta(\tau, n)$. It follows directly from (10)

$$\begin{aligned}\eta_1(\tau, n) &= \Phi_1(x, s, \theta(\tau)\mu_{gap}) = \frac{n}{2\theta(\tau)} - \frac{n}{2} + \frac{1}{\log n}\left(\frac{n}{\theta(\tau)} - n\right) \\ &= \left(\frac{n}{2} + \frac{n}{\log n}\right)\left(\frac{1}{\theta(\tau)} - 1\right). \end{aligned} \tag{13}$$

One can easily verify that $\eta_1(\tau, n)$ is strictly increasing with respect to τ.

We point out here that in the above discussion, we implicitly assume that the parameter τ is dependant on the present iterate (x, s). However, in most IPMs, we usually fix the parameter in the definition of the neighborhood (such as ρ in $\mathcal{N}_\infty^-(\rho)$) and only require the iterate to be in a certain neighborhood. Therefore, in the remaining part of this paper, we assume that $\tau > 1$ is fixed. If the present iterate is in a certain neighborhood(say $\Phi_1(x, s, \mu_{gap}) \leq \eta(\tau, n)$), then we update the parameter μ and compute a search direction regarding to the targeted μ. In this case, we need to investigate the change of the proximity function. Our following result provides an upper bound for $\Phi_1(x, s, \theta(\tau)\mu_{gap})$ whenever $\Phi_1(x, s, \mu_{gap}) < \eta(\tau, n)$.

Lemma 2.4 *If $\Phi_1(x, s, \mu_{gap}) = \eta(\tau_0, n) < \eta(\tau, n)$, then one has*

$$\Phi_1(x, s, \theta(\tau)\mu_{gap}) < \eta_1(\tau, n).$$

Proof: First we observe that $\tau_0 \leq \tau$ and thus $\theta(\tau_0) > \theta(\tau)$. Now recall the definition of $\theta(\tau_0)$, we have

$$\frac{x^T s}{\theta(\tau_0)\mu_{gap}} = \sum_{i \in \mathcal{I}} n^{-\log \sqrt{\frac{x_i s_i}{\theta(\tau_0)\mu_{gap}}}}.$$

It follows that

$$\begin{aligned}\sum_{i \in \mathcal{I}} n^{-\log \sqrt{\frac{x_i s_i}{\theta(\tau)\mu_{gap}}}} &= n^{\frac{1}{2}\log \frac{\theta(\tau)}{\theta(\tau_0)}} \sum_{i \in \mathcal{I}} n^{-\log \sqrt{\frac{x_i s_i}{\theta(\tau_0)\mu_{gap}}}} < \sum_{i \in \mathcal{I}} n^{-\log \sqrt{\frac{x_i s_i}{\theta(\tau_0)\mu_{gap}}}} \\ &= \frac{x^T s}{\theta(\tau_0)\mu_{gap}} < \frac{x^T s}{\theta(\tau)\mu_{gap}} = \frac{n}{\theta(\tau)}. \end{aligned} \tag{14}$$

Applying this inequality to the proximity function $\Phi_1(x, s, \mu)$ yields the desired inequality in the lemma. □

We close this section by the following lemma.

Lemma 2.5 *Let $(x, s) > 0$ and $\tau \geq 1$. Then the relation $\Phi_1(x, s, \mu_{gap}) \leq \eta(\tau, n)$ holds if and only if there exists a parameter $\mu > 0$ such that $\Phi_1(x, s, \mu) \leq \eta_1(\tau, n)$.*

Proof: The necessary part follows directly from Lemma 2.4, thus it remains to consider the sufficient part. Let us assume that $\Phi_1(x, s, \mu_1) \leq \eta_1(\tau, n)$ holds for some $\mu_1 > 0$. Suppose that the statement of the lemma does not hold, i.e.,

$$\Phi_1(x, s, \mu_{gap}) = \eta(\tau_1, n) > \eta(\tau, n), \quad \text{for some } \tau_1 > \tau.$$

Then from Lemma 2.3 we can claim that the proximity function $\Phi_1(x, s, \mu)$ has a global minimum $\Phi_1(x, s, \theta(\tau_1)\mu_{gap}) = \eta_1(\tau_1, n)$ at the point $\mu^* = \theta(\tau_1)\mu_{gap}$. Now by recalling the fact that $\eta_1(\tau, n)$ is strictly increasing in τ, one has $\eta_1(\tau_1, n) > \eta_1(\tau, n)$. It follows immediately that

$$\Phi_1(x, s, \mu_1) \geq \Phi_1(x, s, \theta(\tau_1)\mu_{gap}) > \eta_1(\tau, n),$$

contradicting to our assumption $\Phi_1(x, s, \mu_1) \leq \eta_1(\tau, n)$. This completes the proof of the lemma. □

3 A pseudo predictor-corrector method

In this section, we describe our algorithm and establish its complexity result. The section is divided into three parts. In the first subsection, we describe our new algorithm. The second subsection is devoted to investigating the behavior of the proximity function in the corrector step. In the last part, we will estimate the step size used in the predictor step and summarize the complexity result.

3.1 The algorithm

We start with a brief description of standard IPMs. Suppose the present iterate (x, s) is strictly feasible. For any given duality gap parameter μ, we can apply Newton method to system (1) and obtain a search direction via solving the following linear equation system

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ x\Delta s + s\Delta x &= \mu e - xs. \end{aligned} \tag{15}$$

Let us consider the case that v is defined by (2) with $\mu = \theta(\tau)\mu_{gap}$. Therefore the relation $\Phi_1(x, s, \theta(\tau)\mu_{gap}) = \Psi_1(v)$ holds. For ease of reference, we also denote the search direction in the scaled v-space as

$$d_x := \frac{v\Delta x}{x}, \quad d_s := \frac{v\Delta s}{s}. \tag{16}$$

Using this notation and (2), we can write system (15) in the scaled v-space as

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned} \tag{17}$$

where $\bar{A} = \frac{1}{\mu}AV^{-1}X,\ V = \mathrm{diag}\,(v),\ X = \mathrm{diag}\,(x)$. Recall that the so-called primal-dual affine scaling search direction is the solution of the system

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= -v, \end{aligned} \tag{18}$$

while the search direction induced by the self-regular proximity function $\Psi_1(v)$ can be computed by solving [12]

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1}n^{-\log v} - v = -\nabla\Phi_1(v). \end{aligned} \tag{19}$$

In what follows we give a technical result about the search direction based on the self-regular proximity function $\Psi_1(v)$. We have

Proposition 3.1 *Let (x, s) be a strictly feasible pair and $\Psi_1(v) \leq \eta_1(\tau, n)$, where $\eta_1(\tau, n)$ is defined by (13). Then the solution (d_x, d_s) of system (19) satisfies*

$$v^T(d_x + d_s) \leq 0.$$

Proof: The proposition follows immediately from (14). □

Proposition 3.1 indicates that if the present iterate is in the neighborhood $\mathcal{N}$ defined by (4), then the duality gap will not increase along the search direction obtained by solving (19).

Now we begin to outline the key steps of our algorithm. We assume that a strictly feasible starting point is available and the starting point is in the neighborhood $\mathcal{N}$ where $\eta(\tau, n)$ is defined by (9). Then we solve system (19) to get a search direction. By taking a step $x(\alpha) = x + \alpha\Delta x, s(\alpha) = s + \alpha\Delta s, y(\alpha) = y + \alpha\Delta y$ along the search direction, we can decrease the proximity function. We call such a step a pseudo-corrector step to distinguish it from the standard corrector step in the IPM literature where the search direction is always obtained by solving (15) with $\mu = \mu_{gap}$. Followed the pseudo-corrector is a predictor step, where the search direction is obtained by solving system (18). We take a step along the predictor search direction while keeping the iterate in the neighborhood. This process is repeated until the duality gap is small enough. The algorithmic scheme is stated as follows.

Algorithm 3.1

Input:
Proximity parameters $\tau > 2, \theta(\tau)$ given by (12) and $\eta(\tau, n)$ given by (9) ;
an accuracy parameter $\varepsilon > 0$;
$(x, s) = (x^0, s^0)$ such that $\Phi_1(x, s, \mu_{gap}) \leq \eta(\tau, n)$.

begin
while $x^T s \geq \varepsilon$ **do**
begin
begin
Pseudo-Corrector Step
Compute μ_{gap};
Update μ to $\mu := \theta(\tau)\mu_{gap}$;
Compute $\sigma = \left\| v - v^{-1} n^{-\log v} \right\|$;
Solve the system (19) for $\Delta x, \Delta y, \Delta s$,
Find a sufficiently large feasible step size
such that $\Phi_1(x(\alpha), s(\alpha), \mu) \leq \Phi_1(x(0), s(0), \mu) - \frac{\theta_3(\tau)}{2 \log n} \max\{\sigma, v_{\min}^{-1}(\alpha) n^{-\log v_{\min}(\alpha)}\}$;
$x = x(\alpha); s = s(\alpha); y = y(\alpha)$;
end
begin
Predictor step
Solve the system (18) for $\Delta x, \Delta y, \Delta s$,
Find the maximal step size $\alpha \in (0, 1]$
such that $\Phi_1(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \eta(\tau, n)$;
end
end
end

Remark 3.2 *$\theta_3(\tau)$ is an independent constant defined by (22).*

Remark 3.3 *One significant difference between our pseudo-corrector step and the standard corrector step lies in the changing behavior of the duality gap along the corresponding search direction. As stated in Proposition 3.1, the duality gap never increases but might decrease after a pseudo-corrector step. For instance, if the present iterate (x, s) is on the central path, then one can easily verify that that the search direction (up to a scalar factor) in our pseudo-corrector step is essentially the affine scaling search direction used in the predictor step.*

3.2 The pseudo-corrector step

In this subsection we estimate the change of the proximity function in the pseudo-corrector step. For this we first estimate how much is the decrease of the proximity function Φ_1 in the corrector step. Suppose that the current point (x, s) is in the neighborhood $\mathcal{N}$ given by (4) where $\eta(\tau, n)$ is given by (9) and let $\sigma = \|\nabla\Phi_1(v)\|$. In what follows we estimate the magnitudes of σ and $v_{\min}$. In particular, we will derive a lower bound of $v_{\min}$, from which we can further get an estimation of the maximal feasible step size.

Lemma 3.4 [1] *Suppose that the present iterate (x, s) is in the neighborhood $\mathcal{N}$ defined by (4) and v is defined by (2) with $\mu = \theta(\tau)\mu_{gap}$. Then one has*

$$v_{\min} \geq \exp^{-\frac{2\tau+\log n}{2+\log n}} \geq \exp^{-\tau}, \tag{20}$$

$$v_{\min} n^{\log v_{\min}} \sigma \geq \theta_1(\tau). \tag{21}$$

where

$$\theta_1(\tau) = \min\left\{1 - \exp^{-\frac{1}{2}}, n^{\frac{1}{4}} \exp^{-\frac{1}{4}} \sqrt{\frac{1}{\theta(\tau)} - 2}\right\}.$$

Proof: Using the notation v, we can write $\Phi_1(x, s, \theta(\tau)\mu_{gap}) = \Psi_1(v)$. When $(x, s) \in \mathcal{N}$, from Lemma 2.4 we obtain $\Psi_1(v) \leq \eta_1(\tau, n)$ where $\eta_1(\tau, n)$ is given by (13). It follows from (14) that

$$n^{-\log v_{\min}} < \sum_{i\in\mathcal{I}} n^{-\log v_i} \leq \sum_{i\in\mathcal{I}} (v_i)^2 = \frac{x^T s}{\theta(\tau)\mu_{gap}} = n \exp^{\frac{(2\tau-2)\log n}{2+\log n}}.$$

Therefore, one has

$$\log v_{\min}^{-1} < 1 + \frac{2\tau - 2}{2 + \log n} \leq \tau,$$

which further yields (20).

To prove (21), we observe first that

$$\sigma^2 > \left\|v - v^{-1}\right\|^2 > \left(\frac{1}{\theta(\tau)} - 2\right) n.$$

If $v_{\min} \geq \exp^{-\frac{1}{4}}$, then one gets

$$v_{\min} n^{\log v_{\min}} \sigma \geq n^{\frac{1}{4}} \exp^{-\frac{1}{4}} \left(\frac{1}{\theta(\tau)} - 2\right)^{\frac{1}{2}}.$$

[1] For sufficiently large n, we can get much sharper estimations than (20) and (21). For instance, if $\log n \geq 10\tau - 10$, then relation (20) can be improved to $v_{\min} \geq \exp^{-1.2}$, and inequality (21) can be sharpen to $v_{\min} n^{\log v_{\min}} \sigma \geq 0.9$.

If $v_{\min} < \exp^{-\frac{1}{4}}$, then we have

$$\begin{aligned} v_{\min} n^{\log v_{\min}} \sigma &\geq v_{\min} n^{\log v_{\min}} \left(v_{\min}^{-1} n^{-\log v_{\min}} - v_{\min} \right) \\ &= 1 - v_{\min}^2 n^{\log v_{\min}} \geq 1 - v_{\min}^2 n^{-\frac{1}{4}} \\ &\geq 1 - \exp^{-\frac{1}{2}}. \end{aligned}$$

Inequality (21) follows from the above two inequalities. □

We next proceed to show that there exists a step size that satisfies the condition imposed in the pseudo-corrector step of Algorithm 3.1. We have

Theorem 3.5 *Suppose that the present iterate (x, s) is in the neighborhood $\mathcal{N}$ defined by (4) and v is defined by (2) with $\mu = \theta(\tau)\mu_{gap}$. Let (d_x, d_s) be the solution of system (19) and*

$$\theta_3(\tau) = \frac{\exp^{-\tau} \log n}{2 + 2\log n} \frac{\theta_1(\tau)}{1 + \theta_1(\tau)} \log \frac{2 + \theta_1(\tau)}{2}. \tag{22}$$

Then the step size $\alpha = \frac{\theta_3(\tau)}{\sigma \log n}$ is strictly feasible. Moreover, for this step size, we have

$$\Phi_1(x(\alpha), s(\alpha), \mu) \leq \Phi_1(x(0), s(0), \mu) - \frac{\theta_3(\tau)}{2\log n} \max\{\sigma, v_{\min}^{-1}(\alpha) n^{-\log v_{\min}(\alpha)}\}.$$

Proof: To prove the theorem, we need to estimate the maximal feasible step size $\alpha_{\max}$. By using the orthogonality of d_x and d_s, we have

$$\max\{\|d_x\|^2, \|d_s\|^2\} \leq \|d_x\|^2 + \|d_s\|^2 = \|d_x + d_s\|^2 = \sigma^2. \tag{23}$$

Now recall that the strict feasibility of $(x + \alpha\Delta x, s + \alpha\Delta s)$ can be retained if and only if $(v + \alpha d_x, v + \alpha d_s)$ is strictly feasible. Using (23) and Lemma 3.4, we obtain the following lower bound for the maximal feasible step size

$$\alpha_{\max} \geq v_{\min} \sigma^{-1} \geq \sigma^{-1} \exp^{-\tau}.$$

We progress to investigate the behavior of the function $\Phi_1(x(\alpha), s(\alpha), \theta(\tau)\mu_{gap})$ along the search direction $(\Delta x, \Delta s)$ obtained by solving (19).

Let us define

$$v(\alpha) = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}.$$

Using this, we can write the proximity function $\Phi_1(x(\alpha), s(\alpha), \theta(\tau)\mu_{gap})$ in the scaled v-space as

$$\Phi_1(x(\alpha), s(\alpha), \theta(\tau)\mu_{gap}) = \Psi_1(v(\alpha)).$$

Note that from the choice of $\Psi_1(v(\alpha))$, we have

$$\begin{aligned} \Psi_1(v(\alpha)) &= \frac{(v + \alpha d_x)^T (v + \alpha d_s)}{2} - \frac{n}{2} + \frac{1}{\log n} \sum_{i \in \mathcal{I}} \left(n^{-\log v_i(\alpha)} - 1 \right) \\ &\leq \frac{1}{2} \|v\|^2 + \alpha v^T (d_x + d_s) - \frac{n}{2} + \frac{1}{2\log n} \sum_{i \in \mathcal{I}} \left(n^{-\log(v_i + \alpha d_x^i)} + n^{-\log(v_i + \alpha d_s^i)} - 2 \right), \end{aligned}$$

where the inequality follows from the orthogonality of d_x and d_s as well as the self-regularity of the function $\psi_1(t)$. Therefore, one has

$$\Psi_1(v(\alpha)) - \Psi_1(v(0)) \leq -\alpha \sum_{i\in\mathcal{I}} \left(v_i^2 - n^{-\log v_i}\right) + \frac{1}{2\log n} \sum_{i\in\mathcal{I}} \left(n^{-\log \frac{v_i+\alpha d_x^i}{v_i}} + n^{-\log \frac{v_i+\alpha d_s^i}{v_i}}\right).$$

We can cast the expression at the right hand side of the above inequality as a function of α, saying

$$g(\alpha) := -\alpha \sum_{i\in\mathcal{I}} \left(v_i^2 - n^{-\log v_i}\right) + \frac{1}{2\log n} \sum_{i\in\mathcal{I}} \left(n^{-\log \frac{v_i+\alpha d_x^i}{v_i}} + n^{-\log \frac{v_i+\alpha d_s^i}{v_i}}\right).$$

One can easily verify that $g(\alpha)$ is convex regarding to α. Furthermore, by simple calculus, one gets

$$g(0) = 0, g'(0) = -\sigma^2.$$

For any $\alpha \leq \alpha_{\max}$, we have

$$\begin{aligned} g''(\alpha) &= (1+\log n) \sum_{i\in\mathcal{I}} \left(\frac{(d_x^i)^2}{(v_i+\alpha d_x^i)^2} n^{-\log(v_i+\alpha d_x^i)} + \frac{(d_s^i)^2}{(v_i+\alpha d_s^i)^2} n^{-\log(v_i+\alpha d_s^i)}\right) \\ &\leq (1+\log n) \sum_{i\in\mathcal{I}} \left(\frac{(d_x^i)^2}{(v_i-\alpha\sigma)^2} n^{-\log(v_i-\alpha\sigma)} + \frac{(d_s^i)^2}{(v_i-\alpha\sigma)^2} n^{-\log(v_i-\alpha\sigma)}\right) \\ &\leq \frac{(1+\log n)\sigma^2}{(v_{\min}-\alpha\sigma)^2} n^{-\log(v_{\min}-\alpha\sigma)}, \end{aligned}$$

where the inequalities follow from the fact $\sigma = \|d_x + d_s\|$ and

$$\min_{i\in\mathcal{I}}\{v_i + \alpha d_x^i, v_i + \alpha d_s^i\} \geq \min_{i\in\mathcal{I}} v_i - \alpha\sigma \geq v_{\min} - \alpha\sigma.$$

Let

$$g_1(\alpha) = -\alpha\sigma^2 + \int_0^\alpha \int_0^\zeta \frac{(1+\log n)\sigma^2}{(v_{\min}-t\sigma)^2} n^{-\log(v_{\min}-t\sigma)} dt d\zeta.$$

It is easy to see that $g_1(\alpha)$ is also convex in the feasible region with

$$g_1(0) = 0, g_1'(0) = -\sigma^2.$$

Moreover, for any strictly feasible step size α, the relation

$$g(\alpha) \leq g_1(\alpha)$$

holds. We now proceed to estimate the step size α^* at which $g_1(\alpha)$ has a global minimum. From the convexity of $g_1(\alpha)$ it follows that α^* must satisfy the following optimality condition

$$-\sigma^2 + \frac{1+\log n}{\log n}\left(\frac{\sigma}{v_{\min}-\alpha^*\sigma} n^{-\log(v_{\min}-\alpha^*\sigma)} - \frac{\sigma}{v_{\min}} n^{-\log v_{\min}}\right) = 0, \tag{24}$$

which is equivalent to

$$-\sigma v_{\min} n^{\log v_{\min}} + \frac{1+\log n}{\log n}\left(\frac{1}{1-\alpha^* v_{\min}^{-1}\sigma} n^{-\log(1-\alpha^* v_{\min}^{-1}\sigma)} - 1\right) = 0.$$

Using (21), we derive

$$-\theta_1(\tau) + 2\left(\frac{1}{1-\alpha^* v_{\min}^{-1}\sigma} n^{-\log(1-\alpha^* v_{\min}^{-1}\sigma)} - 1\right) \geq 0. \tag{25}$$

Let

$$\theta_2(\tau) = \frac{\exp^{-\tau}\log n}{2+2\log n}\log\frac{2+\theta_1(\tau)}{2}.$$

It is easy to see that for reasonable large n, we can cast $\theta_2(\tau)$ as a constant depending only on τ. We now claim that

$$\alpha^* \geq \frac{\theta_2(\tau)}{\sigma\log n} \tag{26}$$

To prove (26), we observe that if $\alpha^* > \frac{v_{\min}}{\sigma(1+\log n)}$, then by using (20), we have

$$\alpha^* \geq \frac{v_{\min}}{\sigma(1+\log n)} \geq \frac{\exp^{-\tau}}{\sigma(1+\log n)} \geq \frac{\theta_2(\tau)}{\sigma\log n}.$$

It remains to consider the case where $\alpha^* \leq \frac{v_{\min}}{\sigma(1+\log n)}$. Let $t = (1-\alpha^* v_{\min}^{-1}\sigma)^{-1}$, we can rewrite (25) as

$$t n^{\log t} \geq 1 + \frac{\theta_1(\tau)}{2}.$$

Hence,

$$(1+\log n)\log t \geq \log\frac{2+\theta_1(\tau)}{2}.$$

Note that for sufficiently large n, one has

$$-\log(1-\alpha v_{\min}^{-1}\sigma) \leq \log(1+2\alpha v_{\min}^{-1}\sigma) \leq 2\alpha v_{\min}^{-1}\sigma, \quad \text{if} \quad \alpha v_{\min}^{-1}\sigma \leq \frac{1}{1+\log n}.$$

When $\alpha^* \leq \frac{v_{\min}}{\sigma(1+\log n)}$, it must hold

$$2\alpha^* v_{\min}^{-1}\sigma \geq \frac{1}{1+\log n}\log\frac{2+\theta_1(\tau)}{2}.$$

The above relation yields

$$\alpha^* \geq \frac{v_{\min}}{2\sigma(1+\log n)}\log\frac{2+\theta_1(\tau)}{2} \geq \frac{\exp^{-\tau}}{2\sigma(1+\log n)}\log\frac{2+\theta_1(\tau)}{2} = \frac{\theta_2(\tau)}{\sigma\log n},$$

which yields inequality (26). Let us choose the step size

$$\alpha = \frac{\theta_2(\tau)}{\sigma \log n}. \tag{27}$$

By using Lemma 12 in [12], we have

$$\Psi_1(v(\alpha)) \leq \Psi_1(v) + g_1(\alpha) \leq \Psi_1(v) - \frac{\alpha\sigma^2}{2} \leq \Psi_1(v) - \frac{\theta_2(\tau)\sigma}{2\log n}. \tag{28}$$

Since $\alpha \leq \alpha^*$, it must hold

$$-\sigma^2 + \frac{1+\log n}{\log n}\left(\frac{\sigma}{v_{\min} - \alpha\sigma} n^{-\log(v_{\min}-\alpha\sigma)} - \frac{\sigma}{v_{\min}} n^{-\log v_{\min}}\right) \leq 0. \tag{29}$$

It follows

$$\frac{n^{-\log(v_{\min}-\alpha\sigma)}}{v_{\min}-\alpha\sigma} = \frac{\sigma\log n}{1+\log n} + \frac{n^{-\log v_{\min}}}{v_{\min}} \leq \sigma + \frac{n^{-\log v_{\min}}}{v_{\min}} \leq \frac{\sigma(1+\theta_1(\tau))}{\theta_1(\tau)},$$

where the last inequality is given by (21). Recall (22), one has

$$\theta_3(\tau) = \frac{\theta_1(\tau)\theta_2(\tau)}{1+\theta_1(\tau)} < \theta_2(\tau).$$

From (28) we immediately obtain

$$\Psi_1(v(\alpha)) \leq \Psi_1(v) - \frac{\theta_2(\tau)\sigma}{2\log n} \leq \Psi_1(v) - \frac{\theta_3(\tau) n^{-\log(v_{\min}-\alpha\sigma)}}{2\log n(v_{\min}-\alpha\sigma)}.$$

Now let us recall the fact that

$$v_{\min}(\alpha) \geq v_{\min} - \alpha\sigma.$$

Using these results, we can conclude

$$\Psi_1(v(\alpha)) \leq \Psi_1(v) - \frac{\theta_3(\tau)}{2\log n} v_{\min}^{-1}(\alpha) n^{-\log v_{\min}(\alpha)}.$$

Because $\theta_3(\tau) < \theta_2(\tau)$, combining the above inequality with (28) yields the desired inequality in the theorem. □

3.3 The predictor step and complexity of the algorithm

In this subsection we are going to estimate the step size used in the predictor step that keeps the new iterate in the neighborhood $\mathcal{N}$ defined by (4). To guarantee the resulting iterate in a certain neighborhood, it suffices to show that the value of the proximity function for a strictly feasible step size is bounded above by $\eta(\tau, n)$ or $\eta_1(\tau, n)$, depending on the duality parameter μ used in the definition of the proximity function. In order to distinguish the iterates before and after a pseudo-corrector step, we denote (x^+, s^+) as the primal-dual pair after the pseudo-corrector step. It should be mentioned that in the predictor step, we have the freedom of choosing the duality gap parameter μ because in the predictor step, the resulting search directions Δx and Δs in the original x- and s-spaces are completely independent of the parameter μ. Let μ^+ be the duality gap parameter used in the predictor step. Correspondingly, the scaled vector v^+ can be defined as

$$v^+ = \sqrt{\frac{x^+ s^+}{\mu^+}}; (v^+)^{-1} = \sqrt{\frac{\mu^+ e}{x^+ s^+}}. \tag{30}$$

In what follows we will discuss how to choose an appropriate duality gap parameter μ^+ such that the resulting scaled vector v^+ satisfies the following properties:

A1:

$$\left\|v^+\right\|^2 \leq \frac{n}{\theta(\tau)};$$

A2:

$$\sum_{i\in\mathcal{I}} n^{-\log v_i^+} - \left\|v^+\right\|^2 \leq 0;$$

A3:

$$\Psi_1(v^+) \leq \Phi_1(x, s, \theta(\tau)\mu_{gap}) - \frac{\theta_3(\tau)}{2\log n}\left(v_{\min}^+\right)^{-1} n^{-\log v_{\min}^+}.$$

We have

Lemma 3.6 *Let (x^+, s^+) be the iterate after the pseudo-corrector step and $\mu_{gap}^+ = (x^+)^T s^+/n$. Then there exists a duality gap parameter $\mu^+ \geq \theta(\tau)\mu_{gap}^+$ such that the resulting vector v^+ given by (30) satisfies properties A1 to A3.*

Proof: To prove the lemma, we first observe that if

$$\sum_{i\in\mathcal{I}} n^{-\log\sqrt{\frac{x_i^+ s_i^+}{\theta(\tau)\mu_{gap}}}} - \frac{(x^+)^T s^+}{\theta(\tau)\mu_{gap}} \leq 0, \tag{31}$$

then we can set $\mu^+ = \theta(\tau)\mu_{gap}$, the same as in the corrector step. In this case, one can easily see that both properties A2 and A3 are satisfied. Note that, as indicated by Proposition 3.1, the duality gap will not increase in the pseudo-corrector step, this implies $\mu^+ \geq \theta(\tau)\mu^+_{gap}$ and property A1 is also satisfied.

Next we consider the situation when (31) fails. In this case, our discussion becomes a bit more sophisticated. However, because the duality gap might decrease while the summation $\sum_{i\in\mathcal{I}} n^{-\log\sqrt{\frac{x_i^+ s_i^+}{\theta(\tau)\mu_{gap}}}}$ increases in the pseudo-corrector step, relation (31) might be violated. This implies that instead of (31), we have

$$\sum_{i\in\mathcal{I}} n^{-\log\sqrt{\frac{x_i^+ s_i^+}{\theta(\tau)\mu_{gap}}}} - \frac{(x^+)^T s^+}{\theta(\tau)\mu_{gap}} > 0 \tag{32}$$

after the pseudo-corrector step. Now let us cast $\Phi_1(x^+, s^+, \mu)$ as a function of μ. By following a similar process as in the proof of Lemma 2.3, we can show that the function $\Phi_1(x^+, s^+, \mu)$ attains its global minimum at some point μ^* satisfying

$$\sum_{i\in\mathcal{I}} n^{-\log\sqrt{\frac{x_i^+ s_i^+}{\mu^*}}} - \frac{(x^+)^T s^+}{\mu^*} = 0.$$

When (32) holds, we can conclude that

$$\mu^* < \theta(\tau)\mu_{gap}.$$

On the other hand, if we choose $\mu^+ = \mu^*$ and define v^+ by (30), we have

$$\begin{aligned}
\Psi_1(v^+) &\leq \Phi_1(x^+, s^+, \theta(\tau)\mu_{gap}) \\
&\leq \Phi_1(x, s, \theta(\tau)\mu_{gap}) - \frac{\theta_3(\tau)}{2\log n}\max_{i\in\mathcal{I}}\left\{\sqrt{\frac{\theta(\tau)\mu_{gap}}{x_i^+ s_i^+}}\, n^{\log\sqrt{\frac{\theta(\tau)\mu_{gap}}{x_i^+ s_i^+}}}\right\} \\
&\leq \Phi_1(x, s, \theta(\tau)\mu_{gap}) - \frac{\theta_3(\tau)}{2\log n}\max_{i\in\mathcal{I}}\left\{\sqrt{\frac{\mu^*}{x_i^+ s_i^+}}\, n^{\log\sqrt{\frac{\mu^*}{x_i^+ s_i^+}}}\right\} \\
&= \Phi_1(x, s, \theta(\tau)\mu_{gap}) - \frac{\theta_3(\tau)}{2\log n}\left(v^+_{\min}\right)^{-1} n^{-\log v^+_{\min}}.
\end{aligned}$$

It remains to verify that A1 is also true for this special choice. Note that because $\Phi_1(x^+, s^+, \mu^+) < \eta_1(\tau, n)$, by Lemma 5, $\Phi_1(x^+, s^+, (x^+)^T s^+/n) = \eta(\tau_0, n) < \eta(\tau, n)$ for some $\tau_0 < \tau$. By the definition of $\theta(\cdot)$, we have $\mu^+ = \theta(\tau_0)(x^+)^T s^+/n$. Since $\theta(\cdot)$ is a decreasing function, we have $\theta(\tau_0) > \theta(\tau)$ and $\mu^+ \geq \theta(\tau)(x^+)^T s^+/n$. Moreover, one has

$$\frac{(x^+)^T s^+}{\mu^+} = \frac{n}{\theta(\tau_0)} < \frac{n}{\theta(\tau)},$$

which implies A1. This finishes the proof of the lemma. □

Now we can progress to estimate the increase of the proximity function in the predictor step. To release the notational load, we drop the subscript + used in our previous discussion and simply denote the present iterate by (x, s) and correspondingly the scaled vector by v. We emphasize here again that the scaled vector v satisfies all the properties A1, A2 and A3. We star by estimating $v_{\min}$ for the scaled vector v satisfying A1 to A3. We have

Lemma 3.7 *Let $n \geq 2$ and v satisfy A1, A2 and A3. Then*

$$v_{\min} \geq \exp^{\frac{\log \theta(\tau)}{\log 2} - 1}.$$

Proof: Combination of (A1) and (A2) yields

$$n^{-\log v_{\min}} \leq \sum_{i\mathcal{I}} n^{-\log v_i} \leq \frac{n}{\theta(\tau)}.$$

It follows

$$\log v_{\min} \geq \frac{\log \theta(\tau)}{\log n} - 1 \geq \frac{\log \theta(\tau)}{\log 2} - 1,$$

where the last inequality follows from the fact that $\log \theta(\tau) < 0$ because $\theta(\tau) < 1$ and the assumption that $\log n \geq \log 2$. □

Now we can state the main theorem in this section.

Theorem 3.8 *Let (d_x, d_s) be the solution of system (18). For reasonably large n, there exists an independent constant $\theta_4(\tau) > 0$ such that the step size*

$$\alpha = \frac{\theta_4(\tau) \exp^{\frac{\log \theta(\tau)}{\log 2} - 1}}{\sqrt{n} \log n}$$

is strictly feasible in the predictor step. Moreover, the new iterate updated with this step size is still in the neighborhood defined by (4).

Proof: We start our discussion on the value of the function $\Psi_1(v(\alpha))$ for a strictly feasible step size α. Note that in the predictor step, the conditions A1, A2 and A3 hold. Therefore, we have

$$\max\{\|d_x\|, \|d_s\|\} \leq \|d_x + d_s\| = \|v\| \leq \sqrt{\frac{n}{\theta(\tau)}},$$

which indicates that the maximal step size satisfies

$$\alpha_{\max} \geq v_{\min} \sqrt{\frac{\theta(\tau)}{n}}.$$

By following a similar reasoning chain as in our discussion in the pseudo-corrector step, we can show that

$$\Psi_1(v(\alpha)) - \Psi_1(v(0)) \leq -\alpha \left\|v\right\|^2 + \frac{1}{2\log n} \sum_{i\in\mathcal{I}} \left(n^{-\log \frac{v_i+\alpha d_x^i}{v_i}} + n^{-\log \frac{v_i+\alpha d_s^i}{v_i}} \right) =: h(\alpha).$$

It is easy to verify that $h(\alpha)$ is a convex function of α. Furthermore, by simple calculus, one gets

$$h(0) = 0, h'(0) = \sum_{i\in\mathcal{I}} n^{-\log v_i} - \left\|v\right\|^2.$$

Using property A2, we can claim that $h'(0) \leq 0$. Moreover, for any $\alpha \leq \alpha_{\max}$, one has

$$\begin{aligned} h''(\alpha) &= (1+\log n) \sum_{i\in\mathcal{I}} \left(\frac{(d_x^i)^2}{(v_i+\alpha d_x^i)^2} n^{-\log(v_i+\alpha d_x^i)} + \frac{(d_s^i)^2}{(v_i+\alpha d_s^i)^2} n^{-\log(v_i+\alpha d_s^i)} \right) \\ &\leq (1+\log n) \sum_{i\in\mathcal{I}} \left(\frac{(d_x^i)^2}{(v_i-\alpha\left\|v\right\|)^2} n^{-\log(v_i-\alpha\|v\|)} + \frac{(d_s^i)^2}{(v_i-\alpha\left\|v\right\|)^2} n^{-\log(v_i-\alpha\|v\|)} \right) \\ &\leq \frac{(1+\log n)\left\|v\right\|^2}{(v_{\min}-\alpha\left\|v\right\|)^2} n^{-\log(v_{\min}-\alpha\|v\|)}. \end{aligned}$$

It follows that

$$\begin{aligned} \Psi_1(v(\alpha)) &\leq \Psi_1(v) + \int_0^\alpha \int_0^\zeta \frac{(1+\log n)\left\|v\right\|^2}{(v_{\min}-t\left\|v\right\|)^2} n^{-\log(v_{\min}-t\|v\|)} dt d\zeta \\ &\leq \Psi_1(v) + \frac{(1+\log n)\alpha^2\left\|v\right\|^2}{2(v_{\min}-\alpha\left\|v\right\|)^2} n^{-\log(v_{\min}-\alpha\|v\|)}. \end{aligned}$$

By using property A3, we obtain

$$\Psi_1(v) \leq \eta_1(\tau, n) - \frac{\theta_3(\tau)}{2\log n} v_{\min}^{-1} n^{-\log v_{\min}}.$$

Therefore, for any step size α satisfying

$$\frac{(1+\log n)\alpha^2\left\|v\right\|^2}{2(v_{\min}-\alpha\left\|v\right\|)^2} n^{-\log(v_{\min}-\alpha\|v\|)} \leq \frac{\theta_3(\tau)}{2\log n} v_{\min}^{-1} n^{-\log v_{\min}}, \tag{33}$$

one has

$$\Psi_1(v(\alpha)) \leq \eta_1(\tau, n),$$

which, by Lemma 2.5, further implies that

$$\Phi_1(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \eta(\tau, n).$$

Thus for any step size α satisfying (33), the new iterate $(x(\alpha), s(\alpha))$ is still in the neighborhood defined by (4). Note we can solve the equation

$$\frac{(1+\log n)\alpha^2 \|v\|^2}{2(1-\alpha v_{\min}^{-1}\|v\|)^2} n^{-\log(1-\alpha v_{\min}^{-1}\|v\|)} = \frac{\theta_3(\tau) v_{\min}}{2\log n} \tag{34}$$

to find a step size with the required property. Let $t = \alpha \|v\| v_{\min}^{-1} \log n$. We can write equation (34) as

$$\frac{t^2 n^{-\log(1-t/\log n)}}{1-t/\log n} = \frac{\theta_3(\tau) v_{\min}^{-1} \log n}{1+\log n}.$$

Let us assume that n is sufficiently large, saying $n \geq \exp^2$. Then for any $t \leq 1$, we have

$$-\log\left(1-\frac{t}{\log n}\right) \leq \log\left(1+\frac{2t}{\log n}\right) \leq \frac{2t}{\log n}.$$

The above relation implies that

$$n^{-\log(1-t/\log n)} \leq \exp^{2t} \quad \text{if} \quad n \geq \exp^2, t \leq 1.$$

Using this relation, one can easily show that there exists a constant $\theta_4(\tau)$ depending only on τ such that

$$t \geq \theta_4(\tau).$$

By using Lemma 3.7, we can further claim

$$\alpha \geq \frac{v_{\min}\theta_4(\tau)}{\|v\|\log n} \geq \frac{\theta_4(\tau)\exp^{\frac{\log\theta(\tau)}{\log 2}-1}}{\sqrt{n}\log n}$$

for some constant $\theta_5(\tau)$ relevant to τ only. This finishes the proof of the theorem. □

We close this section by giving the total iteration bound of Algorithm 3.1. As indicated in Proposition 3.1, the duality gap will not increase in the pseudo-corrector step. Theorem 3.8 shows that the duality gap will decrease at least at a rate $1 - \frac{\theta_4(\tau)\exp^{\frac{\log\theta(\tau)}{\log 2}-1}}{\sqrt{n}\log n}$ in each predictor step. As a direct consequence of these two results, we have

Theorem 3.9 *After at most*

$$\left\lceil \frac{\sqrt{n}\log n}{\theta_4(\tau)\exp^{\frac{\log\theta(\tau)}{\log 2}-1}} \log\frac{(x^0)^T s^0}{\epsilon} \right\rceil$$

iterations, Algorithm 3.1 will find an approximate solution satisfying $x^T s \leq \epsilon$.

4 Local quadratic convergence

In this section we prove the local quadratic convergence of the algorithm. Let (x_-, s_-) denote the primal-dual pair before the pseudo-corrector step, and (x, s) the present iterate after a pseudo-corrector step. Recall that for the predictor direction, Ye and Anstreicher [25] has proved the following relation for monotone complementarity problems

$$|\Delta x_i \Delta s_i| = O((\mu_{gap})^2), \ i = 1, ..., n, \tag{35}$$

when the product $x^T s$ (or μ_{gap}) is sufficiently small. Since an LO problem can be cast as a special class of monotone linear complementarity problem, the above relation remains valid for LO if μ_{gap} is small enough. To simplify the analysis, we allow us to abuse a bit the big-O notation and ignore the hidden constant in the big-O relation. Now let us recall the choice of the parameter μ in the predictor step in our algorithm. We then have

$$|d_x^i d_s^i| = \frac{|\Delta x_i \Delta s_i|}{\mu} \leq \frac{|\Delta x_i \Delta s_i|}{\theta(\tau)\mu_{gap}} = O(\mu_{gap}), \tag{36}$$

where the inequality follows from Lemma 3.6. We next give a technical result about the step size used in the predictor step.

Lemma 4.1 *Let (x, s) be an iterate in Algorithm 3.1. If the present duality gap $x^T s$ is sufficiently small so that the relation (35) holds, then the step size α used in the predictor step satisfies $\alpha \geq 1 - O(\mu_{gap})$.*

Proof: We start with an estimation of the maximal feasible step size. As we mentioned in Section 3, the new iterate is strictly feasible if and only if both $v + \alpha d_x$ and $v + \alpha d_s$ are strictly feasible. Further, the maximal feasible step size $\alpha_{\max}$ satisfies the following condition

$$(v + \alpha d_x)(v + \alpha d_s) = v^2 - \alpha v^2 + \alpha^2 d_x d_s \geq 0, \quad \alpha \in [0, \alpha_{\max}],$$

or

$$e - \alpha + \alpha^2 v^{-2} d_x d_s \geq 0, \quad \forall \alpha \in [0, \alpha_{\max}].$$

Since $v_{\min} \geq \exp^{-2}$ and $|d_x d_s| = O(\mu_{gap})$, we can conclude that $\alpha_{\max} \geq 1 - O(\mu_{gap})$.

Now we are going to show that the step size α used in the predictor step also satisfies the relation

$$\alpha \geq 1 - O(\mu_{gap}). \tag{37}$$

To prove (37), we notice that for any strictly feasible step size $\alpha \in (0, 1]$, from the choice of μ and the corresponding scaled vector v we obtain

$$\Phi_1(x(\alpha), s(\alpha), (1-\alpha)\mu) - \Phi_1(x(0), s(0), \mu) = \frac{1}{\log n}\left(\sum_{i \in \mathcal{I}} n^{-\log \frac{v_i(\alpha)}{\sqrt{1-\alpha}}} - \sum_{i \in \mathcal{I}} n^{-\log v_i}\right)$$

$$= \frac{1}{\log n} \sum_{i \in \mathcal{I}} n^{-\log v_i} \left(n^{-\log \frac{v_i(\alpha)}{\sqrt{1-a} v_i}} - 1 \right) = \frac{1}{\log n} \sum_{i \in \mathcal{I}} n^{-\log v_i} \left(n^{-\frac{1}{2} \log \frac{v_i^2(\alpha)}{(1-a) v_i^2}} - 1 \right)$$
$$= \frac{1}{\log n} \sum_{i \in \mathcal{I}} n^{-\log v_i} \left(n^{-\frac{1}{2} \log \frac{(1-\alpha) v_i^2 + \alpha^2 d_x^i d_s^i}{(1-a) v_i^2}} - 1 \right)$$
$$= \frac{1}{\log n} \sum_{i \in \mathcal{I}} n^{-\log v_i} \left(n^{-\frac{1}{2} \log \left(1 + \frac{\alpha^2}{1-\alpha} v_i^{-2} d_x^i d_s^i \right)} - 1 \right).$$

Let $\mathcal{I}_-$ be the index set

$$\mathcal{I}_- = \{ i \in \mathcal{I} : d_x^i d_s^i < 0 \}.$$

It follows immediately that

$$\begin{aligned}
&\Phi_1(x(\alpha), s(\alpha), (1-\alpha)\mu) - \Phi_1(x(0), s(0), \mu) \\
&\quad \leq \frac{1}{\log n} \sum_{i \in \mathcal{I}_-} n^{-\log v_i} \left(n^{-\frac{1}{2} \log \left(1 + \frac{\alpha^2}{1-\alpha} v_i^{-2} d_x^i d_s^i \right)} - 1 \right) \\
&\quad \leq \frac{n^{-\log v_{\min}}}{\log n} \sum_{i \in \mathcal{I}_-} \left(n^{-\frac{1}{2} \log \left(1 + \frac{\alpha^2}{1-\alpha} v_i^{-2} d_x^i d_s^i \right)} - 1 \right) \\
&\quad \leq \frac{n^{-\log v_{\min}}}{\log n} \left(n^{-\frac{1}{2} \log \left(1 + \sum_{i \in \mathcal{I}_-} \frac{\alpha^2}{1-\alpha} v_i^{-2} d_x^i d_s^i \right)} - 1 \right) \\
&\quad \leq \frac{n^{-\log v_{\min}}}{\log n} \left(n^{-\frac{1}{2} \log \left(1 - \frac{\alpha^2}{1-\alpha} O(\mu_{gap}) \right)} - 1 \right).
\end{aligned}$$

This inequality indicates that, for sufficiently small μ_{gap}, there exists a step size $\alpha \geq 1 - O(\mu_{gap})$ such that

$$\Phi_1(x(\alpha), s(\alpha), (1-\alpha)\mu) - \Phi_1(x(0), s(0), \mu) \leq \frac{\theta_4(\tau)}{2 \log n} (v_{\min})^{-1} n^{-\log v_{\min}}.$$

This finishes the proof of the lemma. □

It is straightforward to verify that if the step size used in the predictor step satisfies $\alpha \geq 1 - O(\mu_{gap})$, then after the predictor step we have

$$\mu_{gap}^+ = (1-\alpha)\mu_{gap} = O(\mu_{gap}^2).$$

Now we can state the main result in this section.

Theorem 4.2 *Let (x^k, s^k) be generated by Algorithm 3.1. The algorithm is quadratically convergent in the sense that $\mu_{gap}^{k+1} = O((\mu_{gap}^k)^2)$ and every accumulation point of the sequence (x^k, s^k) is a strictly complementary solution of the problem.*

Proof: The quadratic convergence of the algorithm follows from Lemma 4.1. The convergence properties of every accumulation point of the iterates can be proved by using the properties of the central path, which converges to a strictly complementarity solution of the underlying problem. The details of the proof are omitted here and we refer the readers to [24] for analogous discussions. □

5 Conclusions

A new predictor-corrector algorithm working in a large neighborhood of the central path is proposed for linear optimization. This algorithm maintains the simple structure of the MTY method, that is, each iteration consists of one predictor step and only one corrector step, and uses only first-order information of the problem. We prove that the algorithm retains local quadratic convergence and has an $O\left(n^{\frac{1}{2}} \log n \log \frac{(x^0)^T s^0}{\varepsilon}\right)$ iteration bound. This complexity result improves the so-far best known complexity $O\left(n \log \frac{(x^0)^T s^0}{\varepsilon}\right)$ for predictor-corrector algorithms working in large neighborhoods.

It is also possible to extend the results in this paper to other cases such as predictor-corrector methods for complementarity problems. The details for such an extension are left for the interested reader.

References

[1] K.M. Anstreicher and R.A. Bosch, A new infinity-norm path following algorithm for linear programming, *SIAM J. Optimization,* 5 (1995), pp. 236-246.

[2] H. Frenk, C. Roos, T. Terlaky and S. Zhang, Eds,*High performance optimiztion*, Kluwer Academic publishers, 2000.

[3] C.C. Gonzaga, Complexity of predictor-corrector algorithms for LCP based on a large neighborhood of the central path, *SIAM J. Optim.*, 10 (1999), pp. 183-194.

[4] C.C. Gonzaga and R.A. Tapia, On the convergence of the Mizuno-Todd-Ye algorithm to the analytical center of the soluton set, *SIAM J. Optim.*, 7 (1997), pp. 47-65.

[5] P.F. Huang and Y. Ye, An asymptotical $O(\sqrt{n}L)$-iteration path-following linear programming algorihtm that use wide neighborhoods, *SIAM J. Optim.*, 6 (1996), pp. 570-586.

[6] B. Jansen, C. Roos, T. Terlaky and Y. Ye, Improved complexity using higer order corrections for primal-dual Dikin affine scaling, *Math. Programming*, Series B, 76 (1997), 117-130.

[7] N.K. Karmarkar, A new polymonimal-time algorithm for linear programming, *Combinatorica,* 4 (1984), pp. 373-395.

[8] M. Kojima, N. Megiddo and S. Mizuno, Theoretical convergence of large-step primal-dual interior point algorithms for linear programming, *Math. Programming*, 59 (1993), pp. 1–21.

[9] Z.-Q. Luo, J.F. Sturm and S. Zhang, Superlinear convergence of a symmetric primal-dual path-following alghorithm for semidefinite programming, *SIAM J. Optimization*, 8(1), 1998, pp. 59–81.

[10] 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.

[11] R.C. Monterio, I. Adler and M.O.C. Resende, A polynomial-time primal-dual affine scaling algorithm for linera and convex quadratic programming and its power series extension, *Math. Oper. Res.,* 15 (1990), pp. 191–214.

[12] J.M. Peng, C. Roos and T. Terlaky, Self-regular proximities and new search directions for linear and semidefinite optimization, *Math. Programming*, 93(2002), pp. 129-171.

[13] J. Peng, C. Roos and T. Terlaky, *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*, Princeton University Press, Princeton, New Jersy, 2002.

[14] F.A. Potra, The Mizuno-Todd-Ye algorithm in a large neiborhood of the central path, *European J. Oper. Res.*, 143(2), 2002, pp. 257–267.

[15] 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.

[16] F.A. Potra, A superliner covergent predictor-corrector method for degenerate LCP in a wide neighborhood of the central path with $O(\sqrt{n}L)$-iteration complexity, Technical reports, Department of Mathematics and Statistics, University of Maryland, USA, April, 2001.

[17] J. Renegar, *A Mathematical View of Interior-Point Methods in Convex Optimization*, MPS/SIAM Series on Optimization, SIAM, Philadelphia, PA, 2001.

[18] C. Roos, T. Terlaky, and J. Ph., Vial, *Theory and Algorithms for Linear Optimization: An Interior Approach*, John Wiley & Sons, Chichester, UK, 1997.

[19] 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.

[20] J. Stoer, W. Wechs and S. Mizuno, High order infeasible-interior-point methods for solving sufficent linear complementarity problems, *Math. Oper. Res.*, 23 (1998), pp. 832–862.

[21] J. Sun and G. Zhao, Global linear and local quadratic convergence of a long-step adaptive-mode interior point method for some monotone variational inequality problems, *SIAM J. Optim,* 8 (1998), pp. 123–139.

[22] L. Tüncel, On the convergence of primal-dual interior-point methods with wide neighborhoods. *Comput. Optim. Appl.,* 4 (1995), pp. 139–158.

[23] S.J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, USA, 1997.

[24] Y. Ye, *Interior-Point Algorithms, Theory and Analysis*, John Wiley and Sons, Chichester, UK, 1997.

[25] Y. Ye and K. Anstreicher, On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP, *Math. Programming*, 62 (1993), pp. 537–551.

[26] Y. Ye, O. Güler, R.A. Tapia, and Y. Zhang, A quadratically convergent $O(\sqrt{n}L$-iteration algorithm for linear programming, *Math. Programming*, 59 (1993), pp. 151–162.

[27] Y. Zhang and D. Zhang, On polynomiallity of the Mehrotra-type predictor-corrector interior point algorithms, *Math. Programming*, 68 (1995), pp. 303–317.

[28] G. Zhao, Interior-point algorithms for linear complementarity problems based on large neighborhoods of the central path, *SIAM J. Optimization,* 8 (1998), pp. 397–413.

[29] Y.B. Zhao and J. Han, Two interior-point methods for nonlinear P_*-complementarity problems, *J. Optim. Theory Appl.,* 102 (1999), pp. 659–679.