

# McMaster University

## Advanced Optimization Laboratory



**Title:**

**A Dynamic Large-Update  
Primal-Dual Interior-Point Method  
for Linear Optimization**

**Authors:**

J. Peng and T. Terlaky,

**AdvOL-Report No. 2001/7**  
December 2001, Hamilton, Ontario, Canada

Jiming Peng and Tamás Terlaky  
McMaster University  
Department of Computing and Software  
1280 Main Street West, Hamilton  
Ontario, Canada, L8S, 4L7  
E-mail: pengj@mcmaster.ca,  
terlaky@mcmaster.ca

## Abstract

Primal-dual interior-point methods (IPMs) have shown their power in solving large classes of optimization problems. However, at present there is still a gap between the practical behavior of these algorithms and their theoretical worst-case complexity results, with respect to the strategies of updating the duality gap parameter in the algorithm. The so-called small-update IPMs enjoy the best known theoretical worst-case iteration bound but work very poorly in practice, while the so-called large-update IPMs perform much better in practice but with relatively weaker theoretical results. In this paper, by restricting us to linear optimization (LO), we first exploit some interesting properties of a proximity measure function that has a key role in defining the neighborhood of the central path. These simple but important features of the proximity measure function indicate that, when the current iterate is in a large neighborhood of the central path, then the large-update IPM emerges to be the only natural choice. Then, we apply these results to a specific self-regularity based IPM proposed recently by the authors of this work and C. Roos. Among others, we show that this self-regularity based IPM can also predict precisely the change of the duality gap as the standard IPM does. Therefore, we can directly apply the modified IPM to the simplified self-dual homogeneous model for LO. This provides a remedy for an implementation issue of the proposed IPMs in [11]. A dynamic large-update IPM in large neighborhood is proposed. Different from the algorithm in [11], the new dynamic IPM always takes large-update and does not utilize any inner iteration to get centered. An  $\mathcal{O}\left(n^{\frac{2}{3}} \log \frac{n}{\varepsilon}\right)$  iteration bound of the algorithm is established.

**Keywords:** Linear Optimization, Primal-Dual Interior-Point Method, Proximity Function, Polynomial Complexity, Self-Regular Function.

*AMS Subject Classification:* 90C05

# 1 Introduction

Since Karmarkar's seminal paper [5], many researchers have proposed and analyzed various interior-point methods (IPMs) for solving large classes of optimization problems and numerous exciting results have been reported. For a survey of these results we refer to the recent books on the subject ([13, 16, 17]). In this paper, we deal with primal-dual IPMs for solving the following standard LO problem:

$$(P) \quad \min\{c^T x : Ax = b, x \geq 0\},$$

where  $A \in \Re^{m \times n}$  satisfies  $\text{rank}(A) = m$ ,  $b \in \Re^m$ ,  $c \in \Re^n$ , and its dual problem

$$(D) \quad \max\{b^T y : A^T y + s = c, s \geq 0\}.$$

Without loss of generality, we assume that both (P) and (D) satisfy the interior point condition (IPC), i.e., there exists an  $(x^0, s^0, y^0)$  such that

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

For this and some other properties mentioned below, see, e.g., [13]. If the IPC holds, then finding an optimal solution of (P) and (D) is equivalent to solving the following system.

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

Here  $xs$  denotes the coordinatewise product of the vectors  $x$  and  $s$ . The basic idea of primal-dual IPMs is to replace the third equation in (1) by the parameterized equation  $xs = \mu e$  with  $e = (1, \dots, 1)^T$ . This leads to the following system

$$\begin{aligned} Ax &= b, & x &\geq 0, \\ A^T y + s &= c, & s &\geq 0, \\ xs &= \mu e. \end{aligned} \tag{2}$$

If the IPC holds, then for each  $\mu > 0$ , system (2) has a unique solution. This solution (denoted by  $(x(\mu), s(\mu))$ ) is called the  $\mu$ -center of the primal-dual pair (P) and (D). The set of  $\mu$ -centers with all  $\mu > 0$  gives *the central path* of (P) and (D) [6]. It has been shown that the limit of the central path (as  $\mu$  goes to zero) exists. Because the limit point satisfies the complementarity condition, it naturally yields optimal solutions for both (P) and (D) [13].

Primal-dual IPMs follow the central path  $(x(\mu), s(\mu))$  approximately and approach the optimal solution set of the underlying LO problem as  $\mu$  goes to zero. Let us briefly indicate how this goes. Without loss of generality we assume that the present point  $(x, y, s)$  is in a certain neighborhood of the central path for some positive  $\mu$ . We first update  $\mu$  to  $\mu_+ := (1 - \theta)\mu$ , for some  $\theta \in (0, 1)$ . Then we solve the following Newton system

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu_+ e - xs \end{aligned} \tag{3}$$

and get the unique search direction  $(\Delta x, \Delta y, \Delta s)$ . By taking a step along the search direction where the step size is defined by some line search rules, one constructs a new triple  $(x, y, s)$  that

is ‘closer’ to  $(x(\mu_+), y(\mu_+), s(\mu_+))$ . We repeat this procedure until the present iterate is ‘close enough’ to  $(x(\mu_+), y(\mu_+), s(\mu_+))$  and thus we can set  $\mu := \mu_+$ . Then  $\mu$  is reduced again by the factor  $1 - \theta$  and we apply Newton’s method again targeting at the new  $\mu$ -center, and so on. This process is repeated until  $\mu$  is small enough.

Note that our primary target is to reduce the duality gap as fast as possible, which is done by subsequently decreasing the parameter  $\mu$  with a fixed ratio  $1 - \theta$  at each outer iteration of the algorithm. As a consequence, the choice of the parameter  $\theta$  has an important role in the design and analysis of IPMs. Usually, if  $\theta$  is a constant independent of  $n$ , the dimension of the problem, for instance  $\theta = \frac{1}{2}$ , then we call the algorithm a large-update (or long-step) method. If  $\theta$  depends on the problem dimension such as  $\theta = \frac{1}{\sqrt{n}}$ , then the algorithm is named a small-update (or short-step) method. At present there is still a gap between the practical performance of IPMs and their theoretical worst-case complexity results with respect to different choices of  $\theta$ . The small-update method has the best known  $\mathcal{O}(\sqrt{n} \log \frac{n}{\varepsilon})$  iteration bound, while the large-update method has a worse  $\mathcal{O}(n \log \frac{n}{\varepsilon})$  iteration bound [13, 16, 17]. However, large-update IPMs perform much better in practice than small-update methods [1].

Several strategies have been proposed by various investigators aiming at improving the theoretical complexity of large-update IPMs [3, 4, 8, 11]. In this paper, we will focus on the approach suggested in [11] where new IPMs are induced based on the so-called self-regular proximity functions. As we described earlier, in both the analysis and implementation of IPMs we need to use some proximity functions to keep control on the ‘distance’ from the current iterates to the current  $\mu$ -centers and to the boundary of the feasible set. The basic idea in [11] is to employ a specific class of proximity functions and consequently its induced search directions in the algorithm. To describe briefly the algorithm in [11], we need to introduce some notation first. For any strictly feasible primal-dual pair  $(x, s)$  and any positive number  $\mu$ , we define

$$v := \sqrt{\frac{xs}{\mu}}, \quad v^{-1} := \sqrt{\frac{\mu e}{xs}} \quad (4)$$

to be the vectors whose  $i^{\text{th}}$  components are  $\sqrt{\frac{x_i s_i}{\mu}}$  and  $\sqrt{\frac{\mu}{x_i s_i}}$ , respectively. The proximity functions used in [11] are defined as follows:

$$\Phi(x, s, \mu) := \Psi(v) = \sum_{i=1}^n \psi(v), \quad (5)$$

where  $\psi(t)$ , the kernel function of the proximity function is a so-called univariate self-regular function. For ease of reference, we also denote the search direction in the scaled  $v$ -space as

$$d_x := \frac{v \Delta x}{x}, \quad d_s := \frac{v \Delta s}{s}. \quad (6)$$

Using this notation and (4), we can write the new search direction introduced in [11] as a solution of the following new system

$$\begin{aligned} \bar{A} d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= -\nabla \Psi(v), \end{aligned} \quad (7)$$

where  $\bar{A} = \frac{1}{\mu} A V^{-1} X$ ,  $V = \text{diag}(v)$ ,  $X = \text{diag}(x)$ . The algorithm in [11] can be outlined as follows.

---

**Algorithm 1 for LO**

---

**Input:**

A proximity parameter  $\tau$ ;  
 an accuracy parameter  $\varepsilon > 0$ ;  
 a fixed barrier update parameter  $\theta$ ,  $0 < \theta < 1$ ;  
 $(x^0, s^0)$  and  $\mu^0 = 1$  such that  $\Phi(x^0, s^0, \mu^0) \leq \tau n$ .

**begin**

$x := x^0$ ;  $s := s^0$ ;  $\mu := \mu^0$ ;

**while**  $n\mu \geq \varepsilon$  **do**

**begin**

$\mu := (1 - \theta)\mu$ ;

**while**  $\Phi(x, s, \mu) \geq n\tau$  **do**

Solve the system (7) for  $\Delta x, \Delta y, \Delta s$ ,

**begin**

Determine a step size  $\alpha$ ;

$x := x + \alpha\Delta x$ ;

$s := s + \alpha\Delta s$ ;

$y := y + \alpha\Delta y$ .

**end**

**end**

**end**

---

It should be noted that in the above algorithmic scheme,  $\mu$  is treated as an independent parameter. This is different from what implemented in most IPM solvers, where  $\mu$  is always chosen as the duality gap  $\frac{x^T s}{n}$ . As we observed in the thesis [10, Section 7.1], this might increase the cost for computing a search direction when the algorithm is applied to the self-dual embedding model for LO. A popular strategy used in most IPM solvers is that they always employ the large-update algorithm in the outer iteration but take only one or very few inner iterations to get recentered. Although such a strategy works very efficiently in practice, there is no theoretical explanation for this phenomenon. On the other hand, although the proximity function  $\Psi(v)$  has a dominant role in Algorithm 1, it does not give any indication why we should use large-update IPMs and why they work so well numerically.

The main motivation of this paper is to elaborate on the approach in [11] and then to provide some partial answers to these issues. For this we need to explore the roles of the parameter  $\mu$  and the duality gap in a self-regularity based IPM and their relation. For simplicity of discussion, we will focus only on a specific self-regular proximity measure function with the kernel function  $\psi(t) = \frac{1}{2} \left( t - \frac{1}{t} \right)^2$ . Among others, we will show that if the present iterate is not in a small neighborhood of the central path, then the proximity function itself will naturally lead to a large-update IPM. Moreover, a particular dynamic large-update version of the self-regularity based algorithm can predict the change of the duality gap in the same way as the IPM based on the standard Newton search direction does. This provides a remedy for an implementation issue of the IPMs in [11].

The paper is organized as follows. First, in Section 2, we explore the role of the parameter  $\mu$  w.r.t. this specific proximity measure function. In particular, we will present an interesting observation that, when the current iterate is in a large neighborhood of the central path and the parameter  $\mu$  is defined by the duality gap  $\frac{x^T s}{n}$ , the proximity measure function will not increase after a large update of  $\mu$ . For a special large-update IPM, based on our self-regular proximity measure function, we will discuss how the duality gap changes along the search direction. In Section 3, we propose a new IPM and establish its complexity. Finally, we close this paper by some concluding remarks in Section 4.

A few words about our notation. Throughout the paper  $\|\cdot\|$  denotes the 2-norm of vectors. We denote by  $\mathcal{I}$  the index set  $\mathcal{I} = \{1, 2, \dots, n\}$  and  $x^{-T} s^{-1} = \sum_{i \in \mathcal{I}} x_i^{-1} s_i^{-1}$ . For any  $x = (x_1, x_2, \dots, x_n)^T \in \mathfrak{R}^n$ ,  $x_{\min} = \min\{x_1, x_2, \dots, x_n\}$  (or  $x_{\max}$ ) is the smallest (or largest) component of  $x$ .

## 2 Properties of the Proximity Measure Function

In this section we investigate the properties of the proximity measure function  $\Phi(x, s, \mu)$  or  $\Psi(v)$ , with respect to the argument  $\mu$ . We are particularly interested in the case that the present iterate  $(x, s)$  is far away from the central path. For notational convenience,  $\mu_{gap} := \frac{x^T s}{n}$  denotes the current duality gap and  $v^{gap}$  denotes the vector  $v^{gap} := \sqrt{\frac{xs}{\mu_{gap}}}$ . Note that when the point  $(x, s)$  is fixed, then we can cast the proximity measure function  $\Phi(x, s, \mu)$  as a function of  $\mu$ , i.e.,

$$\Phi(x, s, \mu) := \frac{x^T s}{2\mu} - n + \frac{\mu}{2} x^{-T} s^{-1}.$$

If  $\mu = \mu_{gap}$ , from the choice of the kernel function we know that the value of the proximity measure function is determined by the product  $x^T s x^{-T} s^{-1}$ . If  $(x, s)$  is not in a small neighborhood of the central path, we can assume without loss of generality that there exists a constant  $\tau > 1$  satisfying  $x^T s x^{-T} s^{-1} = \tau n^2$ . It is easy to verify the following interesting relation that plays a crucial role later in the design of our algorithmic scheme.

**Proposition 2.1** *Suppose that the iterate  $(x, s) > 0$  is fixed. Let  $\mu_h = \frac{n}{x^{-T} s^{-1}}$ , then we have*

$$\Phi(x, s, \mu_{gap}) = \Phi(x, s, \mu_h).$$

Because both  $x$  and  $s$  are positive vectors, the function  $\Phi(x, s, \mu)$  is convex with respect to  $\mu$ . Using the optimality condition for convex optimization problems, we can easily prove the following results.

**Proposition 2.2** *For any fixed iterate  $(x, s) > 0$ , the proximity measure function  $\Phi(x, s, \mu)$  has a global minimizer at the point*

$$\mu^* = \sqrt{\frac{x^T s}{x^{-T} s^{-1}}} = \sqrt{\mu_{gap} \mu_h}.$$

Obviously, when the iterate  $(x, s)$  is not on the central path, then  $\mu_h$  is always smaller than  $\mu_{gap}$ , and  $\mu^*$  is the geometric mean of  $\mu_{gap}$  and  $\mu_h$ . Now, let us recall that when the primal-dual

pair  $(x, s)$  is in a large neighborhood of the central path, then usually  $x^T s x^{-T} s^{-1} = \tau n^2$  holds for some positive constant  $\tau > 1$ . This further implies that

$$\frac{\mu_{gap}}{\mu_h} = \tau.$$

Therefore, when  $\tau$  is a relatively large constant, then we can set our target in the outer iteration as

$$\mu_h := \left(1 - \frac{\tau - 1}{\tau}\right) \mu_{gap},$$

which is equivalent to choosing the constant  $\theta = \frac{\tau - 1}{\tau}$  in the update of the parameter  $\mu$  in Algorithm 1. When  $\tau > 2$ , we have  $\theta > \frac{1}{2}$ . This leads naturally to a large-update method. However, we know from Proposition 2.1 that the proximity measure function will still have the same value.

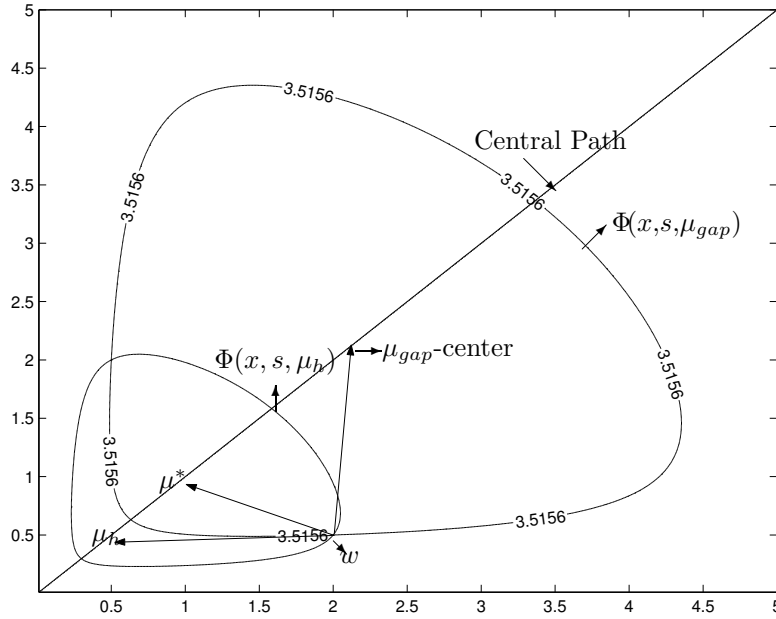


Figure 1: The neighborhoods of  $\mu_{gap}$  and  $\mu_h$  for a given point  $w = \sqrt{xs}$

It should be noted that in Proposition 2.1 we only discuss a specific case of the proximity function when the iterate is far from the central path. In practical implementations of IPMs, we also use a large-update strategy if the iterate is close to the central path, or equivalently the proximity function  $\Phi(x, s, \mu_{gap})$  or the ratio  $\frac{\mu_{gap}}{\mu_h}$  is bounded above by a certain constant. Thus, after one update of  $\mu$ , we need also to investigate the growth behavior of the proximity function, which is demonstrated by the following lemma.

**Lemma 2.3** *If*

$$\frac{\mu_{gap}}{\mu_h} = \frac{x^T s x^{-T} s^{-1}}{n^2} \leq \tau,$$

*then*

$$\Phi\left(x, s, \frac{\mu_{gap}}{\tau}\right) \leq \frac{(\tau - 1)n}{2}.$$

**Proof:** By the assumption of the lemma we can write  $\mu_{gap} = \bar{\tau}\mu_h$  for some  $\bar{\tau} \leq \tau$ . It follows that

$$\begin{aligned}\Phi\left(x, s, \frac{\mu_{gap}}{\tau}\right) &= \frac{\tau n}{2} - n + \frac{n\mu_{gap}}{2\tau\mu_h} \\ &= \frac{(\tau-1)n}{2} - \frac{n}{2} + \frac{n\bar{\tau}}{2\tau} \leq \frac{(\tau-1)n}{2},\end{aligned}$$

which further concludes the lemma.  $\square$

In fact, for any positive  $\tau > 1$ , if we define

$$\mu_t := \frac{2x^T s}{(\tau+1)n + \sqrt{(\tau+1)^2 n^2 - 4x^T s \sum_{i \in \mathcal{I}} x_i^{-1} s_i^{-1}}} = \frac{2\mu_{gap}}{\tau+1 + \sqrt{(\tau+1)^2 - 4\mu_{gap}/\mu_h}}, \quad (8)$$

then it holds obviously  $\Phi(x, s, \mu_t) = \frac{(\tau-1)n}{2}$ . It is easy to verify that  $\mu_t = \mu_h$  if and only if  $\mu_{gap}/\mu_h = \tau$ . In particular, we have  $\mu_h > \mu_t$  whenever  $\mu_{gap}/\mu_h < \tau$ .

Now we proceed to discuss the properties of the search direction based on our specific self-regular proximity function for different updates of  $\mu$ . Note that, due to the specific choice of the kernel function  $\psi(t)$ , we can rewrite system (7) in the original space as

$$\begin{aligned}A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu^2 x^{-1} s^{-1} - xs.\end{aligned} \quad (9)$$

Let us denote by  $(\Delta x(\mu), \Delta y(\mu), \Delta s(\mu))$  the solution of system (9). The following two lemmas discuss the change of the duality gap along the search direction  $(\Delta x(\mu), \Delta s(\mu))$  for  $\mu = \mu^*$  and  $\mu = \mu_h$ .

**Lemma 2.4** *Let  $(\Delta x(\mu^*), \Delta s(\mu^*))$  be the solution of system (9) with  $\mu = \mu^*$ . Then the relation*

$$x^T \Delta s(\mu^*) + s^T \Delta x(\mu^*) = 0$$

*holds.*

**Proof:** The Lemma follows immediately from the choice of  $\mu^*$ .  $\square$

The lemma indicates that if  $\mu = \mu^*$ , then the step won't change the duality gap, i.e.,

$$(x + \alpha \Delta x(\mu^*))^T (s + \alpha \Delta s(\mu^*)) = x^T s$$

for any feasible step size  $\alpha$ .

Similarly, by the choice of  $\mu_h$  we readily get

**Lemma 2.5** *Let  $(\Delta x(\mu_h), \Delta s(\mu_h))$  be the solution of system (9) with  $\mu = \mu_h$ . Then the relation*

$$x^T \Delta s(\mu_h) + s^T \Delta x(\mu_h) = \frac{n^2}{x^{-T} s^{-1}} - x^T s$$

*holds.*

Recall that in traditional IPMs based on the standard Newton direction, we need to solve equation system (3) at each iteration. In this case, if we set the target to  $\mu_+ = \mu_h$ , then the solution of system (3) will satisfy

$$x^T \Delta s + s^T \Delta x = n\mu_h - x^T s.$$

This implies that if the targeted parameter is  $\mu_h$ , then the search direction based on our specific self-regular proximity function and the standard Newton direction will predict the change of the duality gap in the same way.

### 3 A Dynamic Large-Update IPM

In this section we consider a specific variant of Algorithm 1. This variant is more flexible and closer to what is implemented in IPM solvers, since we use large-update at each iterate and do not employ any inner iterations to recenter. It is worthwhile to mention that in practical implementations of IPMs, we always stipulate that the iterate satisfies the condition  $\min(x_i s_i) \geq \ell \mu_{gap}$ , where  $\ell$  is a small constant. Otherwise we utilize corrector steps to obtain a point satisfying such a requirement. Note that if we choose  $\ell = \frac{1}{\tau}$ , then the above condition implies  $\Phi(x, s, \min(x_i s_i)) \leq \frac{(\tau-1)n}{2}$ . On the other hand, one can easily see that

$$\Phi(x, s, \mu_{gap}) = \Phi(x, s, \mu_h) \leq \Phi(x, s, \min(x_i s_i)).$$

This relation demonstrates that the neighborhood defined by  $\Phi(x, s, \mu_{gap}) \leq \tau n$  is larger than the neighborhood generated by  $\Phi(x, s, \min(x_i s_i)) \leq \tau n$ . For example, let us consider the case that  $x = e$  and  $s = (1, \dots, 1, \frac{1}{2n})^T$ . For sufficiently large  $n$ , one can show that  $\Phi(x, s, \mu_{gap}) \leq 2n$ , but the proximity function  $\Phi(x, s, \min(x_i s_i))$  becomes as large as  $\mathcal{O}(n^2)$ . This implies that, in some extreme cases, imposing the condition  $\Phi(x, s, \mu_{gap}) = \tau n$  on the proximity function is probably too aggressive. Let us further consider an implementation issue in the algorithm. Suppose that the present point  $(x, s)$  is in a certain neighborhood of the central path. We then solve the linear system (9) for the search direction  $(\Delta x, \Delta s)$ , from which we can estimate the maximal feasible step size  $\alpha_{\max}$ . A popular heuristics for choosing the step size in IPM solvers is to use a damped factor of  $\alpha_{\max}$ , saying  $0.995\alpha_{\max}$  as a step size. Of course, if the value of the corresponding proximity function is too large for this step size, then we can reduce the step size appropriately so that the value of the proximity function at the new iterate is below the threshold. Note that it is also possible that the proximity function has a relatively small value for the step size  $0.995\alpha_{\max}$ . In this case, theoretically we can still increase the step size such that the value of the proximity function remains below a prescribed bound. However, a practical issue is that, since the step size is already quite close to the maximal feasible step size, even a small increase of the step size might cause some numerical problems or drive the iterate too close to the boundary of the feasible region. In this situation, it is better to use the default value  $0.995\alpha_{\max}$  as the step size. Note that after such a step, we obtain a point pair  $(x, s)$  with a small proximity function value.

Motivated by the above observation, we change slightly the procedure of Algorithm 1 as follows. We utilize a constant  $\tau \geq 10$  to keep control on the update of the duality gap parameter  $\mu$ , namely force the value of the proximity function to satisfy the following relation

$$\Phi(x, s, \mu_{gap}) = \Phi(x, s, \mu_h) \leq \frac{(\tau-1)n}{2}.$$

We also stipulate that when the proximity measure function  $\Phi(x, s, \mu_{gap})$  have a relatively small value, for instance, when  $\Phi(x, s, \mu_{gap}) = \Phi(x, s, \mu_h) \leq \frac{(\tau-2)n}{4}$  or equivalently  $\mu_{gap} \leq \frac{\tau}{2}\mu_h$ , we choose  $\mu_t$  defined by (8) as our targeted center parameter to solve the system (9) for the search direction. Otherwise, we choose the targeted parameter  $\mu_h$ . Further, the step size is carefully chosen so that the relation

$$\Phi(x, s, \mu_{gap}) \leq \frac{(\tau-1)n}{2}$$

holds for all the iterates. Note that in both cases, we have  $\mu_{gap} \geq \frac{\tau}{2}\mu$  where  $\mu$  is the targeted parameter. Hence, our algorithm is indeed a large-update one when  $\tau \geq 10$ . At each step, we also stipulate that the step size should be chosen such that the proximity function  $\Phi(x(\alpha), s(\alpha), \mu_t)$  has a sufficient decrease while the proximity function  $\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha))$  is less than or equal to  $\frac{(\tau-1)n}{2}$  (or equivalently  $\Phi(x(\alpha), s(\alpha), \mu^*(\alpha)) \leq (\sqrt{\tau}-1)n$ ). For simplicity we use the notation  $x(\alpha) := x + \alpha\Delta x$ ,  $y(\alpha) := y + \alpha\Delta y$  and  $s(\alpha) := s + \alpha\Delta s$ . The algorithm can be outlined as follows.

---

### Algorithm 2 for LO

---

**Input:**

A proximity parameters  $10 < \tau$ ;  
 an accuracy parameter  $\varepsilon > 0$ ;  
 $(x, s) = (x^0, s^0)$  such that  $x^T s x^{-T} s^{-1} \leq \tau n^2$ .

**begin**

**while**  $x^T s \geq \varepsilon$  **do**

**begin**

$$\tau_0 = \frac{x^T s x^{-T} s^{-1}}{n^2};$$

Compute  $\mu_t$  by (8);

If  $\tau \leq 2\tau_0$  then  $\mu := \mu_h$ ;

otherwise  $\mu := \mu_t$ ;

Solve the system (9) for  $\Delta x, \Delta y, \Delta s$ ,

**begin**

Determine a step size that

decreases  $\Phi(x(\alpha), s(\alpha), \mu_t)$  sufficiently and

satisfies  $\mu_{gap}(\alpha) \leq \tau\mu_h(\alpha)$ ;

$x = x(\alpha); s = s(\alpha); y = y(\alpha)$ .

**end**

**end**

**end**

---

**Remark 3.1** *In the algorithm it would be desirable that the step size  $\alpha$  minimizes the proximity function  $\Psi(x(\alpha), s(\alpha), \mu_t)$  while the constraint  $x(\alpha)^T s(\alpha) x(\alpha)^{-T} s(\alpha)^{-1} \leq \tau n^2$  is satisfied. However, this would require the exact solution of the subproblem. Instead of this, in the present algorithm we solve the subproblem approximately so that the value of the function  $\Phi(x(\alpha), s(\alpha), \mu_t)$  is decreased sufficiently and the constraint is satisfied. Theorem 3.7 will give a default value for such a step size.*

We proceed to analyze the complexity of the algorithm. The key of the analysis is to estimate the value of the step size  $\alpha$  in Algorithm 2. For this we need to estimate the function  $\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha))$  with  $\mu_{gap}(\alpha) = \frac{x(\alpha)^T s(\alpha)}{n}$ , and hence  $x(\alpha)^T s(\alpha) x(\alpha)^{-T} s(\alpha)^{-1}$ , as well as the function  $\Phi(x(\alpha), s(\alpha), \mu_t)$ .

To simplify the analysis, we use the notation  $v, d_x, d_s$  for the system (7) when  $\mu$  is chosen as described in Algorithm 2. Let us further denote  $\sigma = \|v - v^{-3}\|$ . We discuss first the decreasing behavior of the proximity measure function  $\Phi(x(\alpha), s(\alpha), \mu_t)$  when  $\mu_t$  is used as the targeted duality gap parameter. By specifying the kernel function to the special case in this paper and modifying slightly the proof of Theorem 3.3.4 in [10] or Theorem 3.6 in [11], we have

**Theorem 3.2** *Let  $\tau \geq 10$  and  $(\Delta x, \Delta s)$  be the solution of system (9) with  $\mu = \mu_t$  in Algorithm 2. Then the step size  $\alpha_1 = \frac{9}{10}\sigma^{-\frac{4}{3}} < 1$  is strictly feasible. Moreover, for any step size  $\alpha \leq \frac{1}{9}\sigma^{-\frac{4}{3}}$ , the relation*

$$\Phi(x(\alpha), s(\alpha), \mu_h) \leq \Phi(x, s, \mu_h) - \frac{\alpha}{4}\sigma^2$$

holds.

**Proof:** We outline only the key steps in the proof. The reader is referred to [10] or [11] for the detailed proof. In order to estimate the decrease of the proximity measure function after a step, we first give some bounds for the minimal component  $v_{\min}$  of  $v$  and the maximal feasible step size  $\alpha_{\max}$ . Recall that by using Lemma 2.13 and Lemma 3.2 of [11], we have

$$v_{\min} \geq (1 + \sigma)^{-\frac{1}{3}} \geq \frac{3\sigma}{1 + 3\sigma}\sigma^{-\frac{1}{3}} \geq \frac{9}{10}\sigma^{-\frac{1}{3}}, \quad (10)$$

where the second inequality follows from Lemma A.1 of [11] or Lemma 1.3.1 in [10], and the last inequality is given by the fact that when  $\tau \geq 10$  is used in Algorithm 2, the relation

$$\sigma^2 \geq \|v - v^{-1}\|^2 \geq (\tau - 1)n \geq 9 \quad (11)$$

holds. It follows immediately that

$$\alpha_{\max} \geq v_{\min}\sigma^{-1} > \frac{9}{10}\sigma^{-\frac{4}{3}}. \quad (12)$$

This proves the first statement of the theorem. Now let us define

$$f(\alpha) = \Psi(v_+(\alpha)) - \Psi(v) = -\frac{1}{2}\|v^{-1}\|^2 + \frac{1}{2}\alpha v^T(d_x + d_s) + \frac{1}{2}\sum_{i \in \mathcal{I}} \frac{1}{[v + \alpha d_x]_i [v + \alpha d_s]_i},$$

with  $v_+(\alpha) = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}$ , where the second equality follows from the orthogonality of  $d_x$  and  $d_s$ . For any feasible step size  $\alpha$ , by following a similar chain of reasoning as in the proof of Lemma 3.3 in [11], we can show that

$$f''(\alpha) \leq \frac{3\sigma^2}{2}(v_{\min} - \alpha\sigma)^{-4}.$$

Moreover, via simple calculus, we have

$$f(0) = 0; \quad f'(0) = -\frac{\sigma^2}{2},$$

which further gives

$$f(\alpha) \leq f_1(\alpha) := -\frac{\sigma^2}{2}\alpha + \frac{3\sigma^2}{2} \int_0^\alpha \int_0^\xi (v_{\min} - \zeta\sigma)^{-4} d\zeta d\xi.$$

Obviously, the function  $f_1(\alpha)$  has the global minimum at the point

$$\bar{\alpha} := \frac{v_{\min}}{\sigma} \left( 1 - \left( 1 + \sigma v_{\min}^3 \right)^{-\frac{1}{3}} \right).$$

Moreover, for any  $\alpha \leq \bar{\alpha}$ , by using Lemma 3.5 of [11] or Lemma 1.3.3 of [10], we have

$$f(\alpha) \leq f_1(\alpha) \leq -\frac{\alpha}{4}\sigma^2, \quad \forall \alpha \leq \bar{\alpha}.$$

Therefore, the proof of the theorem will be finished if we can show that

$$\bar{\alpha} \geq \frac{1}{9}\sigma^{-\frac{4}{3}}. \quad (13)$$

By using Lemma A.1 of [11] and following a similar process as in the proof of Lemma 3.4 in [11], we can conclude that

$$\bar{\alpha} \geq \frac{v_{\min}^3 \sigma}{3(1 + v_{\min}^3 \sigma)} v_{\min} \sigma^{-1} > \frac{1}{3(2\sigma + 1)} v_{\min} > \frac{1}{6} \left( 1 + \sigma^{-1} \right)^{-\frac{4}{3}} \sigma^{-\frac{4}{3}}, \quad (14)$$

where the last two inequalities follow from (10). Recall that  $\sigma \geq 3$  whenever  $\tau \geq 10$ , it follows

$$\frac{1}{6} \left( 1 + \sigma^{-1} \right)^{-\frac{4}{3}} \geq \frac{1}{9},$$

which, together with (14) yields inequality (13).  $\square$

We mention that the bound (12) for the maximal step size  $\alpha_{\max}$  is sharper than the one presented in [11], because we could use some special features of the kernel function  $\frac{1}{2}(t - t^{-1})^2$  and the orthogonality of  $d_x$  and  $d_s$ . However, such a slight improvement does not allow to improve the final complexity of the algorithm.

Next we proceed to estimate the proximity function  $\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha))$  or equivalently the function  $\Phi(x(\alpha), s(\alpha), \mu^*(\alpha))$  for a feasible step size when  $\mu_t$  is used in Algorithm 2 as the targeted parameter. For this it suffices to consider the function  $\Phi(x(\alpha), s(\alpha), \mu^*)$ , because the inequality  $\Phi(x(\alpha), s(\alpha), \mu^*(\alpha)) \leq \Phi(x(\alpha), s(\alpha), \mu^*)$  holds.

**Theorem 3.3** *Let  $\tau \geq 10$  and  $(\Delta x, \Delta s)$  be the solution of system (9) with  $\mu = \mu_t$  in Algorithm 2. Then the step size  $\alpha_2 = \min\{\frac{2}{3(\tau+1)}, \frac{1}{9}\}\sigma^{-\frac{4}{3}} < 1$  is strictly feasible. Moreover, for any step size  $\alpha \leq \alpha_2$ , the relation*

$$\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \Phi(x, s, \mu_t)$$

*holds.*

Before we prove the theorem, it should be noted that there exist two cases when  $\mu_t$  is used as the targeted duality gap parameter: (i)  $\mu_h = \mu_t$ ; (ii)  $\mu_{gap} < \frac{\tau\mu_h}{2}$ .

We start by discussing case (i).

**Lemma 3.4** *Let  $(\Delta x, \Delta s)$  be the solution of the system (9) with  $\mu = \mu_h = \mu_t$  in Algorithm 2. Then the step size  $\bar{\alpha}_2 = \frac{2}{3(\tau+1)}\sigma^{-\frac{4}{3}} < 1$  is strictly feasible. Moreover, for any step size  $\alpha \leq \bar{\alpha}_2$ , the relation*

$$\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \Phi(x, s, \mu)$$

holds.

**Proof:** First we observe that from the assumption that  $\mu_{gap} = \tau\mu_h = \tau\mu_t$  we obtain  $\mu^* = \sqrt{\tau}\mu_h$ . Let us define

$$\begin{aligned} g(\alpha) &:= \Phi(x(\alpha), s(\alpha), \mu^*) - \Phi(x, s, \mu^*) \\ &= -\frac{\sqrt{\tau}}{2} \|v^{-1}\|^2 + \frac{1}{2\sqrt{\tau}} \alpha v^T (d_x + d_s) + \frac{\sqrt{\tau}}{2} \sum_{i \in \mathcal{I}} \frac{1}{[v + \alpha d_x]_i [v + \alpha d_s]_i}. \end{aligned}$$

It is straightforward to verify that

$$g(0) = 0, \quad g''(\alpha) = \sqrt{\tau} f''(\alpha) \leq \frac{3\sqrt{\tau}\sigma^2}{2} (v_{\min} - \alpha\sigma)^{-4}.$$

Moreover, from the definition of  $v$ , we have  $\|v^{-1}\|^2 = n$ , which further implies

$$\|v^{-3}\|^2 - \|v^{-1}\|^2 = \|v^{-3} - v^{-1}\|^2 + 2\|v^{-2} - e\|^2 \geq 0.$$

This inequality, together with the fact that  $\|v\|^2 = \tau \|v^{-1}\|^2$ , gives

$$g'(0) = -\frac{\sigma^2}{2\sqrt{\tau}} - \frac{\tau-1}{2\sqrt{\tau}} \sum_{i \in \mathcal{I}} v_i^{-3} (v_i^{-3} - v_i) \leq -\frac{\sigma^2}{2\sqrt{\tau}}.$$

It follows that

$$g(\alpha) \leq g_1(\alpha) := -\frac{\sigma^2}{2\sqrt{\tau}}\alpha + \frac{3\sqrt{\tau}\sigma^2}{2} \int_0^\alpha \int_0^\xi (v_{\min} - \zeta\sigma)^{-4} d\zeta d\xi.$$

By some cumbersome calculus, one can show that the equation  $g_1(\alpha) = 0$  has two roots:  $\alpha = 0$  and  $\alpha = \frac{v_{\min}^3 \sigma}{\tau + v_{\min}^3 \sigma} v_{\min} \sigma^{-1}$ . This further implies that

$$g(\alpha) \leq g_1(\alpha) \leq 0; \quad \forall \alpha \leq \frac{v_{\min}^3 \sigma}{\tau + v_{\min}^3 \sigma} v_{\min} \sigma^{-1}.$$

On the other hand, from (10) and (11) we obtain

$$\frac{v_{\min}^3 \sigma}{\tau + v_{\min}^3 \sigma} v_{\min} \geq \frac{\sigma}{\tau(\sigma+1) + \sigma} (1 + \sigma)^{-\frac{1}{3}} \geq \frac{1}{\tau+1} (1 + \sigma^{-1})^{-\frac{4}{3}} \sigma^{-\frac{1}{3}} \geq \frac{2}{3(\tau+1)} \sigma^{-\frac{1}{3}}.$$

Hence, whenever  $\alpha \leq \frac{2}{3(\tau+1)}\sigma^{-\frac{4}{3}}$ , the inequality  $\Phi(x(\alpha), s(\alpha), \mu^*(\alpha)) \leq \Phi(x(\alpha), s(\alpha), \mu^*(0))$  holds. Now, by using the inequality  $\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \Phi(x(\alpha), s(\alpha), \mu^*(\alpha))$ , we can conclude the lemma.  $\square$

Next we progress to discuss case (ii) where  $\mu_{gap} = \tau_0 \mu_h$  with  $\tau_0 < \frac{\tau}{2}$ .

**Lemma 3.5** *Let  $\tau \geq 10$  and  $(\Delta x, \Delta s)$  be the solution of system (9) with  $\mu = \mu_t$ . If  $\mu_{gap} = \tau_0 \mu_h$  and  $\tau_0 < \frac{1}{2}\tau$ . Then the step size  $\hat{\alpha}_2 = \frac{1}{9}\sigma^{-\frac{4}{3}} < 1$  is strictly feasible. Moreover, for any step size  $\alpha \leq \hat{\alpha}_2$ , we have*

$$\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \Phi(x, s, \mu_t).$$

**Proof:** It suffices to estimate the interval in which the proximity function satisfies

$$\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \frac{(\tau - 1)n}{2}.$$

We start by considering the function

$$g_1(\alpha) = \Phi(x(\alpha), s(\alpha), \mu^*) = \frac{1}{2} \left( \frac{\mu_t}{\mu^*} \|v_+(\alpha)\|^2 - 2n + \frac{\mu^*}{\mu_t} \|v_+(\alpha)^{-1}\|^2 \right).$$

Note that, from the definitions of  $\mu^*$  and  $\mu_t$ , we know that

$$\frac{\mu_t}{\mu^*} \|v_+(0)\|^2 = \frac{\mu^*}{\mu_t} \|v_+(0)^{-1}\|^2 = n\sqrt{\tau_0}.$$

On the other hand, by the choice of the search direction we have

$$\|v_+(\alpha)\|^2 < \|v_+(0)\|^2$$

for any strictly feasible step size. It is easy to verify that

$$\|v_+(\alpha)^{-1}\|^2 \leq (1 - \alpha v_{\min}^{-1} \sigma)^{-2} \|v^{-1}\|^2.$$

Combining the above two inequalities together, we can claim that

$$g_1(\alpha) \leq n(\sqrt{\tau} - 1)$$

for all  $\alpha$  satisfying  $(1 - \alpha v_{\min}^{-1} \sigma)^{-2} \leq \sqrt{2} \leq \sqrt{\frac{\tau}{\tau_0}}$ . Because

$$1 - 2^{-\frac{1}{4}} \geq \frac{1}{8},$$

one can further prove that if

$$\alpha \leq \frac{1}{8} v_{\min} \sigma^{-1},$$

then  $g_1(\alpha) \leq n(\sqrt{\tau} - 1)$ . Note that from (10)

$$v_{\min} \geq \frac{9}{10} \sigma^{-\frac{1}{3}},$$

follows. This inequality implies that  $g_1(\alpha) \leq n(\sqrt{\tau} - 1)$  whenever  $\alpha \leq \frac{1}{9} \sigma^{-\frac{4}{3}}$ . Recall that from Theorem 3.2 we already know that this step size is strictly feasible. This completes the proof of the lemma.  $\square$

Theorem 3.3 follows immediately from Lemma 3.4 and Lemma 3.5.

It remains to consider the behaviors of  $\Phi(x(\alpha), s(\alpha), \mu_t)$  and  $\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha))$  when  $\mu_h$  is used as the targeted duality gap parameter in Algorithm 2.

**Theorem 3.6** Let  $\tau \geq 10$  and  $(\Delta x, \Delta s)$  be the solution of system (9) with  $\mu = \mu_h$ ,  $\frac{\tau}{2}\mu_h \leq \mu_{gap} < \tau\mu_h$  and  $\mu_t < \mu_h$ . Then the step size  $\alpha_3 = \min\left\{\frac{1}{9}, \frac{2}{3(\tau+1)}\right\} \sigma^{-\frac{4}{3}} < 1$  is strictly feasible. Moreover, for any step size  $\alpha \leq \alpha_3$ , we have

$$\Phi(x(\alpha), s(\alpha), \mu_t) \leq \Phi(x, s, \mu_t) - \frac{\alpha}{4} \max\left\{\frac{\mu_t}{\mu_h} \sigma^2, 2\Phi(x, s, \mu_t)\right\}; \quad (15)$$

$$\Phi(x(\alpha), s(\alpha), \mu_{gap}(\alpha)) \leq \Phi(x, s, \mu_t). \quad (16)$$

In particular,

$$\Phi(x(\alpha_3), s(\alpha_3), \mu_t) \leq \Phi(x, s, \mu_t) - \min\left\{\frac{1}{54}, \frac{1}{9(\tau+1)}\right\} \Phi(x, s, \mu_t)^{\frac{1}{3}} \quad (17)$$

holds.

**Proof:** We first point out that the proof of Lemma 3.4 can be easily adapted to prove relation (16) in the theorem by substituting the parameter  $\tau$  by  $\tau_1 := \frac{\mu_{gap}}{\mu_h}$ . Therefore, in what follows we focus only on inequalities (15) and (17).

In order to investigate the behavior of the proximity function  $\Phi(x(\alpha), s(\alpha), \mu_t)$  for a feasible step size  $\alpha$ , we define

$$\begin{aligned} h(\alpha) &:= \Phi(x(\alpha), s(\alpha), \mu_t) - \Phi(x(0), s(0), \mu_t) \\ &= \frac{\alpha\mu_h}{2\mu_t} v^T(d_x + d_s) + \frac{\mu_t}{2\mu_h} \sum_{i \in \mathcal{I}} \left( \frac{1}{[v + \alpha d_x]_i [v + \alpha d_s]_i} - \frac{1}{v_i^2} \right) \\ &= \frac{1}{2} \left( \frac{\mu_h}{\mu_t} - \frac{\mu_t}{\mu_h} \right) \alpha v^T(d_x + d_s) + \frac{\mu_t}{2\mu_h} \left( \alpha v^T(d_x + d_s) - \|v^{-1}\|^2 + \sum_{i \in \mathcal{I}} \frac{1}{[v + \alpha d_x]_i [v + \alpha d_s]_i} \right). \end{aligned}$$

Now, by applying a procedure similar to the proof of Theorem 3.2 to the second term in the above formulae, we can prove that for  $\alpha \leq \frac{1}{9}\sigma^{-\frac{4}{3}}$ , the following relation

$$h(\alpha) \leq \frac{1}{2} \left( \frac{\mu_h}{\mu_t} - \frac{\mu_t}{\mu_h} \right) \alpha v^T(d_x + d_s) - \frac{\alpha\mu_t}{4\mu_h} \sigma^2 \quad (18)$$

holds. Because

$$v^T(d_x + d_s) = -\|v\|^2 + \|v^{-1}\|^2 = -\left(\frac{\mu_{gap}}{\mu_h} - 1\right) \|v^{-1}\|^2 \leq -\|v^{-1}\|^2,$$

for any  $\alpha \leq \frac{1}{9}\sigma^{-\frac{4}{3}}$ , we have

$$\begin{aligned} h(\alpha) &\leq -\frac{\alpha}{4} \left( \frac{\mu_h}{\mu_t} \left( \|v\|^2 - \|v^{-1}\|^2 \right) + \frac{\mu_t}{\mu_h} \left( \|v^{-3}\|^2 - \|v^{-1}\|^2 \right) \right) - \left( \frac{\alpha\mu_h}{4\mu_t} - \frac{\alpha\mu_t}{4\mu_h} \right) \|v^{-1}\|^2 \\ &= -\frac{\alpha}{4} \left( \frac{\mu_h}{\mu_t} \|v\|^2 + \frac{\mu_t}{\mu_h} \left( \|v^{-3}\|^2 - 2\|v^{-1}\|^2 \right) \right) \\ &\leq -\frac{\alpha}{4} \left( \frac{\mu_h}{\mu_t} \|v\|^2 + \frac{\mu_t}{\mu_h} \|v^{-1}\|^2 - 2n \right) \\ &= -\frac{\alpha}{2} \Phi(x, s, \mu_t). \end{aligned}$$

On the other hand, from (18) we easily derive

$$h(\alpha) \leq -\frac{\alpha\mu_t}{4\mu_h} \sigma^2, \quad \forall \alpha \leq \frac{1}{9}\sigma^{-\frac{4}{3}}.$$

Combining the above two bounds, inequality (15) follows immediately, and we thus get the required bound for the proximity function  $\Phi(x(\alpha), s(\alpha), \mu_t)$ . To prove (17), we observe that (15) implies

$$\Phi(x(\alpha_3), s(\alpha_3), \mu_t) \leq \Phi(x, s, \mu_t) - \min \left\{ \frac{1}{36}, \frac{1}{6(\tau+1)} \right\} \max \left\{ \frac{\mu_t}{\mu_h} \sigma^{\frac{2}{3}}, 2\Phi(x, s, \mu_t) \sigma^{-\frac{4}{3}} \right\}. \quad (19)$$

With little effort, one can show that

$$\max \left\{ \frac{\mu_t}{\mu_h} \sigma^{\frac{2}{3}}, 2\Phi(x, s, \mu_t) \sigma^{-\frac{4}{3}} \right\} \geq \left( \frac{\sqrt{2}\mu_t}{\mu_h} \right)^{\frac{2}{3}} \Phi(x, s, \mu_t)^{\frac{1}{3}}. \quad (20)$$

Now, let us recall the assumptions in the theorem. If we define  $\mu_{gap} = \tau_2 \mu_t$ , then we have  $\frac{\tau}{2} \leq \tau_1 < \tau < \tau_2$ . Moreover, by the definition of  $\mu_t$  one can easily see that  $\mu_{gap} \leq (\tau+1)\mu_t$ , which further implies

$$\frac{\mu_h}{\mu_t} = \frac{\tau_2}{\tau_1} < \frac{2\tau+2}{\tau} < 2 + \frac{2}{\tau} \leq 2.2, \quad (21)$$

where the last inequality follows from the assumption that  $\tau \geq 10$ . It follows from (20) that

$$\max \left\{ \frac{\mu_t}{\mu_h} \sigma^{\frac{2}{3}}, 2\Phi(x, s, \mu_t) \sigma^{-\frac{4}{3}} \right\} \geq \left( \frac{\sqrt{2}\mu_t}{\mu_h} \right)^{\frac{2}{3}} \Phi(x, s, \mu_t)^{\frac{1}{3}} \geq \frac{2}{3} \Phi(x, s, \mu_t)^{\frac{1}{3}}.$$

Combining this with (19) yields

$$\Phi(x(\alpha), s(\alpha), \mu_t) \leq \Phi(x, s, \mu_t) - \min \left\{ \frac{1}{54}, \frac{1}{9(\tau+1)} \right\} \Phi(x, s, \mu_t)^{\frac{1}{3}}, \quad (22)$$

which further completes the proof of the theorem.  $\square$

Finally, summarizing Theorem 3.2, Theorem 3.3 and Theorem 3.6 together, we have

**Theorem 3.7** *Let  $\tau \geq 10$  and  $(\Delta x, \Delta s)$  be the solution of system (9) used in Algorithm 2. Then the step size  $\alpha^* = \min\{\frac{1}{16}, \frac{2}{3\tau}\} \sigma^{-\frac{4}{3}} < 1$  is strictly feasible. Moreover, we have*

$$\Phi(x(\alpha^*), s(\alpha^*), \mu_{gap}(\alpha^*)) \leq \Phi(x, s, \mu_t); \quad (23)$$

$$\Phi(x(\alpha^*), s(\alpha^*), \mu_t) \leq \Phi(x, s, \mu_t) - \min \left\{ \frac{1}{54}, \frac{1}{9(\tau+1)} \right\} \Phi(x, s, \mu_t)^{\frac{1}{3}}. \quad (24)$$

To obtain an upper bound for the total number of iterations of the algorithm, we need to estimate the value of the step size  $\alpha^*$  or the change of the parameter  $\mu_t$  before and after one iterate. The following technical lemma will be used in our estimation about  $\mu_t$ . The lemma is a direct consequence of [13, Lemma IV.36].

**Lemma 3.8** *Let  $v_+ = \frac{v}{\sqrt{1-\theta}}$  for some  $\theta \in (0, 1)$ . Then we have*

$$\Psi(v_+) \leq \frac{1}{1-\theta} \Psi(v) + \frac{\theta \sqrt{2n\Psi(v)}}{1-\theta} + \frac{n\theta^2}{1-\theta}.$$

By applying Lemma 3.8 to Theorem 3.7, we can prove the following theorem.

**Theorem 3.9** *Let  $\tau \geq 10$  and  $(\Delta x, \Delta s)$  be the solution of system (9) and  $\alpha^*$  is the default step size defined in Theorem 3.7. Then there exists a constant  $\rho > 0$  such that*

$$\Phi\left(x(\alpha^*), s(\alpha^*), \frac{\mu}{1-\theta}\right) \leq \Phi(x, s, \mu_t),$$

where  $\theta = \rho n^{-\frac{2}{3}}$ .

By the choice of  $\mu_t$  we know that the proximity function  $\Phi(x, s, \mu_t)$  always keeps as a constant at each iterate. Let us denote by  $\mu_t^+$  the target parameter value after one step. Then we have

$$\Phi(x(0), s(0), \mu_t) = \Phi(x(\alpha^*), s(\alpha^*), \mu_t^+).$$

On the other hand, we know that

$$\begin{aligned} \Phi(x(\alpha^*), s(\alpha^*), \mu_{gap}(\alpha^*)) &\leq \Phi(x(\alpha^*), s(\alpha^*), \mu_t^+); \\ \Phi(x(\alpha^*), s(\alpha^*), \mu_t) &\leq \Phi(x(\alpha^*), s(\alpha^*), \mu_t^+). \end{aligned}$$

Because  $\mu_t^+ \leq \mu_{gap}(\alpha^*)$ , from the above two inequalities we get

$$\mu_t > \mu_t^+.$$

Therefore, by using Theorem 3.7, we can claim that

$$\mu_t^+ \leq (1 - \rho n^{-\frac{2}{3}})\mu_t.$$

In light of this relation, we know that after at most  $\mathcal{O}(n^{\frac{2}{3}} \log \frac{n}{\varepsilon})$  iterations, we have  $\mu_t \leq \varepsilon$ . Therefore, the  $\mathcal{O}(n^{\frac{2}{3}} \log \frac{n}{\varepsilon})$  complexity of the algorithm follows directly.

## 4 Concluding Remarks

Some interesting properties of the proximity function induced by the kernel function  $\frac{1}{2}(t - t^{-1})^2$  have been explored. In particular, these attractive features of the proximity function indicate that if the present iterate is far from the central path, then the large-update algorithm appears to be a natural choice for finding a good search direction and keeping control on the value of the proximity function. Furthermore, in some important cases, this self-regularity based search direction can predict the change of the duality gap along the search direction in the same way as the standard Newton direction does. Based on these observations, a dynamic primal-dual large-update IPM for solving LO is proposed and the complexity of the algorithm match that of its analogue presented in [11]. It worths mentioning that the dynamic IPM in this paper does not use any inner process to get recentered. This is very close to the algorithm implemented in many IPM solvers and different from the algorithmic scheme used in [11]. We mention that we have already had some preliminary numerical tests for our new algorithm based on a variant of LIPSOL [18]. Compared with the standard large-update IPM, the number of iterations of the dynamic algorithm is usually less than or equal to that of the large-update IPM based on the standard Newton direction. There is another difference between these two algorithms. We utilize a line search routine to find a suitable step size for our search direction. It is interesting

to note that in most cases, we can really use the default step size  $0.95\alpha_{\max}$ . For this step size, the proximity function at the new iterate still satisfies  $\Phi(x, s, \mu_{\text{gap}}) \leq \frac{(\tau-1)n}{2}$ . Nevertheless, much more extensive numerical tests are needed to further check the efficiency of the dynamic algorithm.

There are several ways to extend our results. The first is to consider dynamic large-update IPMs for semidefinite optimization and second-order conic optimization. For this we need to investigate the properties of proximity function in the cone of semidefinite matrices and second-order cone. Then we can combine the idea and techniques in this paper with those in papers [11] and [12] together to carry out transparently the analysis in the conic cases.

Another way to extend the results in this paper is to consider dynamic large-update IPMs based on general self-regular functions. However, we would like to point out that it seems very hard to design dynamic IPMs for the whole family of self-regular functions. Recall our results in Section 2. Some properties of the proximity function rely on the strict convexity of  $\Phi(x, s, \mu)$  with respect to  $\mu$ . By the definition of  $\Phi(x, s, \mu)$ , we know that if the function  $\psi(1/\sqrt{t})$  is strictly convex, then  $\Phi(x, s, \mu)$  is strictly convex in  $t$ . For the specific choice in this paper, one can easily verify that the function  $\frac{t}{2} - 2 + \frac{2}{t}$  is strictly convex. However, we can not obtain the strict convexity of  $\psi(1/\sqrt{t})$  from the self-regularity of  $\psi(t)$ . It is worthwhile to explore for which subclass of the self-regular functions, the strict convexity of  $\psi(1/\sqrt{t})$  holds. Because the analysis in this paper is already considerably involved, we leave such extensions to the interested reader.

Finally, we remark that the neighborhood used in this paper and many other papers in the literature are usually determined by a proximity function, like  $\Phi(x, s, \mu_{\text{gap}})$ . However, as we discussed earlier at the beginning of Section 3, the neighborhood defined by  $\Phi(x, s, \mu_{\text{gap}})$  is possibly too large and very different from what implemented in most IPM solvers. This also indicates that there exists still a big gap between the pure theoretical analysis of IPMs and their practical implementation. We hope this paper can lead to some new investigations on IPMs that can help us to bridge this gap.

**Acknowledgement** The fact that a self-regular proximity function might have an optimal value at some point much less than the duality gap was first observed by Prof. Y. Nesterov during his visit to McMaster University in December 2000, and the present work is inspired by his observation.

## References

- [1] E.D. Andersen, J. Gondzio, Cs. Mészáros, and X. Xu. Implementation of interior point methods for large scale linear programming. In T. Terlaky, editor, *Interior Point Methods of Mathematical Programming*, pages 189–252. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
- [2] R.A. Horn, C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [3] P. Hung and Y. Ye. An asymptotically  $O(\sqrt{n}L)$ -iteration path-following linear programming algorithm that uses long steps. *SIAM J. on Optimization*, 6:570-586,1996.
- [4] B. Jansen, C. Roos, T. Terlaky, and Y. Ye. Improved complexity using higher order correctors for primal-dual Dikin affine scaling, *Mathematical Programming, Series B*, 76:117–130, 1997.

- [5] N.K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [6] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior Point and Related Methods*, pages 131–158. Springer Verlag, New York, 1989. Identical version in: *Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan*, pages 1–35, 1986.
- [7] S. Mehrotra. On the implementation of a (primal-dual) interior point method. *SIAM J. on Optimization*, 2(4):575–601, 1992.
- [8] R.D.C. Monteiro, I. Adler and M.G.C. Resende. A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series. *Mathematics of Operations Research*, 15:191–214, 1990.
- [9] Y.E. Nesterov and A.S. Nemirovskii. Interior point polynomial algorithms in convex programming. *SIAM Studies in Applied Mathematics*, Vol(13), Philadelphia, USA, 1994.
- [10] J. Peng. New Design and Analysis of Interior-Point Methods. Ph. D. Thesis, Faculty of Information Technology and Systems, Delft University of Technology, The Netherlands, June, 2001.
- [11] J. Peng, C. Roos and T. Terlaky. Self-regular functions and new search directions for linear and semidefinite optimization. Technical Report, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands, 2000.
- [12] J. Peng, C. Roos and T. Terlaky. New polynomial primal-dual methods for second-order conic optimization based on self-regular proximities. Technical Report, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands, 2000.
- [13] C. Roos, T. Terlaky, and J.-Ph.Vial. *Theory and Algorithms for Linear Optimization. An Interior Approach*. John Wiley & Sons, Chichester, UK, 1997.
- [14] W. Rudin, *Principles of Mathematical Analysis*. Mac-Graw Hill Company, New York, 1978.
- [15] G. Sonnevend. An “analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In A. Prékopa, J. Szelezsán, and B. Strazicky, editors, *System Modelling and Optimization: Proceedings of the 12th IFIP-Conference held in Budapest, Hungary, September 1985*, volume 84 of *Lecture Notes in Control and Information Sciences*, pages 866–876. Springer Verlag, Berlin, West-Germany, 1986.
- [16] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, USA, 1997.
- [17] Y. Ye. *Interior Point Algorithms, Theory and Analysis*. John Wiley & Sons, Chichester, UK, 1997.
- [18] Y. Zhang. User’s guide to LIPSOL: linear-programming interior point solvers V0.4. Interior point methods. *Optimization Methods and Software*, 11/12:385–396, 1999.