McMaster University

Advanced Optimization Laboratory

Title:

Adaptive Large Neighborhood Self-Regular Predictor-Corrector IPMs for LO

Authors:

Maziar Salahi and Tamás Terlaky

AdvOl-Report No. 2004/7

July 2004, Hamilton, Ontario, Canada

Adaptive Large Neighborhood Self-Regular Predictor-Corrector IPMs for LO

Maziar Salahi*† and Tamás Terlaky†

July 20, 2004

Abstract

It is known that predictor-corrector methods in a large neighborhood of the central path are among the most efficient interior point methods (IPMs) for linear optimization (LO) problems. The best iteration bound based on the classical logarithmic barrier function is $\mathcal{O}\left(n \log \frac{n}{\epsilon}\right)$. In this paper we propose a family of self-regular proximity based predictor-corrector (SR-PC) IPM for LO in a large neighborhood of the central path. Like all predictor-corrector IPMs, our new SR-PC algorithms have a predictor step and a corrector step. In the predictor step we use either the affine scaling or a self-regular direction, while in the corrector step we always use a self-regular direction. Our new algorithms use a special proximity function with different search directions and thus allows us to improve the so far best theoretical iteration complexity for a family of SR-PC IPMs. An $\mathcal{O}\left(\sqrt{n}\exp(\frac{1-q+\log n}{2})\log n \log \frac{n}{\epsilon}\right)$ worst-case iteration bound with quadratic convergence is established, where q is the barrier degree of the SR proximity function. If $q = 1 + \log n$, then we have the so far best iteration complexity for the first order predictor-corrector method in a large neighborhood, and if $q = 1 + 2\log(\log n)$ our algorithm has an $\mathcal{O}(n \log \frac{n}{\epsilon})$ iteration complexity. For the case $q = 1$, the result is a factor $\log n$ worse than the exist recent results.

Keywords: Linear Optimization, Predictor-Corrector Method, Superlinear Convergence, Self-regular Proximity Function, Large Neighborhood, Polynomial Complexity.

1 Introduction

Since Karmarkar's seminal paper [6], many researchers have proposed and analyzed various Interior-Point Methods (IPMs) for solving large classes of optimization problems with numerous new exciting results. For a survey of these results we refer to the recent books [18, 21, 22] on the subject. In this paper, we deal with primal-dual IPMs for solving the following standard Linear Optimization (LO) problem:

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

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

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

*Department of Mathematics and Statistics, McMaster University, L8S 4K1 , salahim@math.mcmaster.ca

†Advanced Optimization Lab, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada, L8S 4K1. terlaky@mcmaster.ca, msalahi@optlab.mcmaster.ca

We may assume without loss of generality [18, 22] that both (P) and (D) satisfy the interior point condition (IPC), i.e., there exists an (x^0, s^0, y^0) such that

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

In this paper we focus on predictor-corrector IPMs for LO in a large neighborhood of the central path. First we start with a brief review of the predictor-corrector method. This method follows the central path by taking a predictor step and then a corrector step. In the predictor step our aim is to reduce the duality gap and in the corrector step we are trying to bring the iterate back to a certain neighborhood of the central path. The best known predictor-corrector algorithm is the Mizuno-Todd-Ye (MTY) algorithm for LO, that works in a small neighborhood of the central path [9]. In 1991 Ji et al. proposed a predictor-corrector algorithm for LCP with superlinear convergence [5]. Quadratic convergence of the duality gap was proved in [23]. Due to the fact that corrector steps based on the first order centering direction are rather inefficient in a large neighborhood, analyzing the MTY type predictor-corrector algorithm in a large neighborhood is more difficult. Different variants of the predictor-corrector algorithms in large neighborhood have been proposed. Some of the important developments are summarized in the next paragraph.

Anstreicher and Bosch [1] proved that the iteration complexity of a straightforward implementation of the MTY predictor-corrector method in a large neighborhood is $O(n^{\frac{3}{2}} \log \frac{n}{\epsilon})$. Using a very elegant analysis, Gonzaga [3] proposed a predictor-corrector algorithm in a wide neighborhood using multiple centering steps in the corrector step. He showed that the maximum number of centering steps is $\mathcal{O}\left(n \log \frac{n}{\epsilon}\right)$. In [16] Potra proposed a predictor-corrector algorithm for LCP that has the same structure as the MTY algorithm. He proved that under general conditions his algorithm has an $\mathcal{O}(n \log \frac{n}{\epsilon})$ iteration complexity, and quadratic convergence of the duality gap under the assumption that the LCP is nondegenerate. Recently, Potra and Liu have [17] proposed a predictor-corrector algorithm for sufficient linear complementarity problems with the same structure as the MTY method in the wide neighborhood $\mathcal{N}_\infty^-(\rho)$, and proved an $\mathcal{O}\left(n \log \frac{n}{\epsilon}\right)$ iteration complexity with quadratic convergence. Several strategies have been proposed for improving the theoretical complexity of IPMs in a large neighborhood. For example, Hung and Ye [4] proposed an asymptotically $\mathcal{O}(\sqrt{n} \log \frac{n}{n})$ complexity algorithm that is based on higher-order approximation of the central path. A more practical predictor-corrector algorithm, proposed by Mehrotra [8], is based on the power series algorithm of Bayer and Lagarias [2] and the primal-dual version of Monteiro, Adler and Resende [10]. Mehrotra's technique is used widely in all IPM implementations. A polynomial time version of a Mehrotra type predictor-corrector algorithm has been presented in [24].

The authors of [11, 14, 19] proposed new classes of IPMs in a large neighborhood based on the novel concept of self-regular (SR) proximity functions, and proved that such an algorithm admits an $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$ iteration bound. Various large-update algorithms with improved complexity result have been proposed [14, 19] by using SR-functions. Recently, Peng et al., [15] proposed a new variant of the predictor-corrector algorithm based on a specific SR-proximity function with $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$ iteration complexity. In this paper we propose a novel variant of the predictor-corrector algorithm that is using different search directions in the predictor and corrector steps. In the predictor step it makes either a SR step with adaptive choice of the target value or an affine scaling step. In the corrector step it uses a SR search direction. When $q = 1 + \log n$ our algorithm has the same polynomial complexity as Peng et al.'s [15] algorithm, and when $q = 1 + 2\log(\log n)$ it has an $\mathcal{O}(n \log \frac{n}{\epsilon})$ iteration complexity. In the case of $q = 1$, our result is a factor $\log n$ worse than the result in [17]. We also prove the quadratic convergence of the new algorithm.

The paper is organized as follows. First, in Section 2, we give a brief description of IPMs and then we introduce the family of SR functions and SR search directions for LO. In Section 3, first we introduce our SR neighborhood, then we discuss the role of the parameter μ w.r.t. a specific proximity function [19]. At the end of this section we discuss how the duality gap changes along the search direction [19]. In Section 4, a new SR predictor-corrector IPM in a large neighborhood of the central path is proposed and its polynomial complexity is established. In Section 5, we prove quadratic convergence of the duality gap. Finally, the paper is closed by some conclusions in Section 6.

A few words about our notation. Throughout the paper $\|\cdot\|$ denotes the 2-norm of vectors. We denote by $\mathcal{I}$ the index set $\mathcal{I} = \{1, 2, \ldots, n\}$ and $x^{-T}s^{-1} = \sum_{i\in\mathcal{I}} x_i^{-1}s_i^{-1}$. For any $x = (x_1, x_2, \ldots, x_n)^T \in R^n$, $x_{\min} = \min\{x_1, x_2, \ldots, x_n\}$ is the smallest component of x and $x_{\max}$ is defined analogously to denote the largest component. For any two vectors x and s, xs denotes the componentwise product of two vectors and e denotes a vector with all components equal to one. Throughout R^n_{++} denotes the positive orthant and R^n_+ denotes the nonnegative orthant in R^n.

2 IPMs and Self-Regular Functions

Before getting to the main theme of the paper first we give a brief introduction to IPMs. Finding optimal solutions of (P) and (D) is equivalent to solving the following system:

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

The basic idea of primal-dual IPMs is to replace the third equation in (1) by the parameterized equation $xs = \mu e$. This leads to the following system:

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

If the IPC holds, then for each $\mu > 0$, system (2) has a unique solution. This solution, denoted by $(x(\mu), y(\mu), s(\mu))$, is called the μ-center of the primal-dual pair (P) and (D). The set of μ-centers with all $\mu > 0$ gives *the central path* of (P) and (D) [7, 20]. It has been shown that the limit of the central path (as μ goes to zero) exists. Because the limit point satisfies the complementarity condition, it naturally yields optimal solutions for both (P) and (D), respectively [18].

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

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

to obtain the search direction $(\Delta x, \Delta y, \Delta s)$. Under the assumption that $\operatorname{rank}(A) = m$, the Newton system (3) has a unique solution. By taking a step along the search direction, where

the step size is chosen so that the new iterate has significantly smaller proximity value w.r.t. $(x(\mu_+), y(\mu_+), s(\mu_+))$. 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 μ is reduced again by the factor $1-\theta$ and we apply Newton's method again with targeting the new μ-center, and so on. This process is repeated until μ is small enough.

Now let us recall the definition of SR functions form [11]. The family of univariate SR functions is defined as follows.

Definition 2.1 *A twice continuously differentiable function $\psi(t) : (0, \infty) \to R$ is SR if it satisfies the following two conditions:*

SR.1 The function $\psi(t)$ is strictly convex with respect to $t > 0$ and vanishes at its global minimal point $t = 1$, i.e., $\psi(1) = \psi'(1) = 0$. Further, there exist positive constants $\nu_2 \geq \nu_1 > 0$ and $p \geq 1,\ q \geq 1$ such that

$$\nu_1(t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq \nu_2(t^{p-1} + t^{-1-q}), \quad \forall t \in (0, \infty); \tag{4}$$

SR.2 For any $t_1, t_2 > 0$,

$$\psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1-r)\psi(t_2), \quad \forall r \in [0, 1]. \tag{5}$$

If $\psi(t)$ is SR, then parameter q is called the *barrier degree* and parameter p is called the *growth degree* of the SR function $\psi(t)$.

There are two popular families of SR functions. The first family is given by

$$\Upsilon_{p,q}(t) = \frac{t^{p+1} - 1}{p(p+1)} + \frac{t^{1-q} - 1}{q(q-1)} + \frac{p-q}{pq}(t-1), \quad p \geq 1,\ q > 1, \tag{6}$$

with $\nu_1 = \nu_2 = 1$. The second family is defined as

$$\Gamma_{p,q}(t) = \frac{t^{p+1} - 1}{p+1} + \frac{t^{1-q} - 1}{q-1}, \quad p \geq 1,\ q > 1, \tag{7}$$

with $\nu_1 = \min(p, q)$ and $\nu_2 = \max(p, q)$. For $p, q = 1$ in both cases the classical logarithmic barrier function

$$\Gamma_{11}(t) = \Upsilon_{11}(t) = \frac{t^2 - 1}{2} - \log t$$

is obtained.

Let $v \in R^n_{++}$. Then an SR-proximity function $\Psi : R^n_{++} \to R_+$ measures the discrepancy between the vectors v and e, and is defined as $\Psi(v) = \sum_{i=1}^{n} \psi(v_i)$, where $\psi(t)$ is a univariate SR function, called the kernel function of the SR-proximity. Further, we recall that a new paradigm of IPMs was recently introduced by Peng et al. in [11]. They used SR-proximity measures to define search directions and to control the iterative process.

In classical primal-dual IPMs, we need to solve (3), while in SR-IPMs the Newton system (3) is modified. To define the modified system we need to introduce the notation

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

where the i^{th} components of the vector v and v^{-1} are $\sqrt{\frac{x_i s_i}{\mu}}$ and $\sqrt{\frac{\mu}{x_i s_i}}$, respectively. Then, the Newton system for SR-IPMs for LO is given as:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\mu v \nabla \Psi(v) \end{bmatrix}, \tag{8}$$

where $v\nabla\Psi(v) = (v_1\nabla\psi(v_1), \cdots, v_n\nabla\psi(v_n))^T$, $X = \text{diag}\,(x)$, and $S = \text{diag}\,(s)$ (see [11]). For ease of reference, we also scale the search directions Δx and Δs in the scaled v-space as

$$d_x := \frac{v\Delta x}{x} \quad \text{and} \quad d_s := \frac{v\Delta s}{s}. \tag{9}$$

Using this notation the Newton system (8) can be written as

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= -\nabla\Psi(v), \end{aligned} \tag{10}$$

where $\bar{A} = \frac{1}{\mu}AV^{-1}X$, and $V^{-1} = \text{diag}\, v^{-1}$.

3 Properties of a SR-Proximity and a SR-Neighborhood

In this section we investigate some properties of the SR-proximity function

$$\Phi_\ell(x, s, \mu) := \Psi_\ell(v) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{-\log n} - n}{\log n} \tag{11}$$

with respect to the argument μ. This proximity function is induced by the kernel function $\Gamma_{1q}(t)$, given by (7), where $q = 1 + \log n$. For notational convenience $\mu_g := \frac{x^T s}{n}$ denotes the parameter value associated with the current duality gap.

Predictor-corrector algorithms originally use an infinity neighborhood, which is defined as

$$\mathcal{N}_\infty^-(\rho) := \{(x, y, s) : (x, s) > 0,\ Ax = b,\ A^T y + s = c,\ \left\|(v^2 - e)^-\right\|_\infty \le \rho\}, \tag{12}$$

where $a^- = \min(a, 0)$ and $\rho \in (0, 1)$ is a constant independent of n. We define the SR neighborhood in a way that it contains the infinity neighborhood and these two neighborhoods almost match each other [15]. Our SR neighborhood is defined as

$$\mathcal{N}(\eta(n, \tau)) := \{(x, y, s) | (x, s) > 0,\ Ax = b,\ A^T y + s = c,\ \Phi_\ell(x, s, \mu_g) \le \eta(n, \tau)\}, \tag{13}$$

where $\eta(n, \tau)$ is a positive function that depends on a constant τ and the dimension of the underlying problem. Assume that $(x, s) \in \mathcal{N}_\infty^-(\rho)$, then for this (x, s) pair with $\mu = \mu_g$ we have

$$\Phi_\ell(x, s, \mu_g) = \frac{e^T v^{-\log n} - n}{\log n} \le \frac{n(1-\rho)^{\frac{-\log n}{2}} - n}{\log n} = \frac{n^\tau - n}{\log n} = \eta(n, \tau),$$

where $\tau = 1 - \frac{1}{2}\log(1 - \rho)$. One can see easily that with this choice of $\eta(n, \tau)$ the neighborhood $\mathcal{N}(\eta(n, \tau))$ contains the neighborhood $\mathcal{N}_\infty^-(\rho)$.

In the original predictor-corrector method we use the primal-dual affine scaling direction in the predictor step, which is the solution of the following system of equations

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= -xs. \end{aligned} \tag{14}$$

It is well known that close to optimality, the step size of the affine scaling direction is converging to one [18]. This fact implies that close to optimality the affine scaling direction is the best choice. In this paper in the first iterations we change this scheme and we use SR directions for the predictor step. To be more precise, we compare the duality gap reduction of the affine scaling direction with the warranted duality gap reduction of the SR direction and then we move in the direction that gives larger duality gap reduction.

Now we introduce the target value that we use in the predictor step. Since we operate in a large neighborhood, we define $\frac{(\tau_1-1)n}{2}$ as the maximum allowed value of the proximity function w.r.t. the target μ, where $\tau_1 = (1-\rho)^{-1}$. One can show that $\Phi_\ell(x,s,\mu_\ell^t) = \frac{(\tau_1-1)n}{2}$ if and only if μ_ℓ^t solves the equation

$$2(x^{\frac{-\log n}{2}})^T s^{\frac{-\log n}{2}} \mu^{\frac{2+\log n}{2}} - (2n + \tau_1 n \log n)\,\mu + x^T s \log n = 0. \tag{15}$$

This equation has two positive roots, one is less than or equal to μ_g, and the other one is larger than or equal to μ_g. In the algorithm we will use the smaller positive root μ_ℓ^t as the target in the SR-predictor step. One can easily prove that $\mu_\ell^t \le \mu_\ell^h$, holds when $\mu_g \le \tau_1\mu_\ell^h$, with $\mu_\ell^t = \mu_\ell^h$ if and only if $\mu_g = \tau_1\mu_\ell^h$, where $\mu_\ell^h = \left(\frac{n}{x^{\frac{-\log n}{2}} s^{\frac{-\log n}{2}}}\right)^{\frac{2}{\log n}}$ is the generalized harmonic mean [19] of the components of the vector xs. The following lemma gives a useful relation between μ_g and μ_ℓ^t.

Lemma 3.1 *Let μ_ℓ^t be defined by equation (15). Then the inequality*

$$\tau_1\mu_\ell^t \le \mu_g \le \left(\tau_1 + \frac{2}{\log n}\right)\mu_\ell^t$$

holds.

Proof: We give a proof for the right hand side inequality. The left hand side can be proved analogously. The function in (15) is a convex function w.r.t. μ and for $\mu = \mu_g$ it is negative. If we replace μ by $\frac{\mu_g}{\tau_1+\frac{2}{\log n}}$ in that function and simplify, it suffices to prove that

$$2\left(\frac{\mu_g}{(\tau_1+\frac{2}{\log n})\mu_\ell^h}\right)^{\frac{\log n}{2}} \frac{1}{\tau_1+\frac{2}{\log n}} - \frac{2+\tau_1\log n}{\tau_1+\frac{2}{\log n}} + \log n \ge 0, \tag{16}$$

which obviously is true. This completes the proof of the right hand side inequality. That completes the proof of the lemma. □

The following result [19] shows that the global minimum of the proximity function as a function of μ, unlike the case of the primal-dual logarithmic barrier function, does not happen at μ_g.

Proposition 3.2 *For any fixed point $(x, s) > 0$, the proximity function $\Phi_\ell(x, s, \mu)$, as a function of μ, has its global minimizer at*

$$\mu_\ell^* = \left(\frac{x^T s}{(x^{\frac{-\log n}{2}})^T s^{\frac{-\log n}{2}}} \right)^{\frac{2}{2+\log n}}.$$

Moreover, $\Phi_\ell(x, s, \mu)$ is a decreasing function w.r.t. μ when $\mu \leq \mu_\ell^*$, and it is an increasing function of μ if $\mu > \mu_\ell^*$.

Proof: It can be easily proved that the proximity function as a function of μ is strictly convex. Using the strict convexity of the proximity function and the optimality condition, one can easily verify that μ_ℓ^* is the global minimizer of the proximity function. The proof of the second part is straightforward. □

The following remark is crucial for the rest of the paper.

Remark 3.3 *Without loss of generality we may always assume that after a predictor step the new iteration is on the boundary of the SR neighborhood, or equivalently $\Phi(x, s, \mu_g) = \eta(n, \tau)$.*

The relation between μ_g and μ_ℓ^* plays an important role in the analysis of our new algorithm.

Lemma 3.4 *Let $\mu_\ell^* = \mu_{1+\log n}^*$ as it is defined in Proposition 3.2. If $\Phi_\ell(x, s, \mu_g) = \eta(n, \tau)$, then $\mu_\ell^* = \theta_1(\tau)\mu_g$, where $\theta_1(\tau) = \exp(\frac{2(1-\tau)\log n}{2+\log n})$, and we have $\exp(2(1-\tau)) \leq \theta_1(\tau) \leq \exp(1-\tau)$.*

Proof: Let $\mu_\ell^* = \theta_1(\tau)\mu_g$ and then we need to derive the specified formula for $\theta_1(\tau)$. The global minimum of the proximity function is the root of its derivative with respect to μ i.e.,

$$-\frac{x^T s}{2(\mu_\ell^*)^2} + \frac{(\mu_\ell^*)^{\frac{-2+\log n}{2}} \left(x^{-\frac{\log n}{2}}\right)^T s^{-\frac{\log n}{2}}}{2} = 0,$$

that is equivalent to the following equation if we let $\mu_\ell^* = \theta_1(\tau)\mu_g$

$$\frac{n}{\theta_1(\tau)^2 \mu_g} = \theta_1(\tau)^{\frac{-2+\log n}{2}} \mu_g^{\frac{-2+\log n}{2}} \left(x^{-\frac{\log n}{2}}\right)^T s^{-\frac{\log n}{2}}.$$

This implies that

$$n\theta_1(\tau)^{\frac{-2-\log n}{2}} = \left\| v^{\frac{-\log n}{2}} \right\|^2 = n^\tau = \exp(\tau \log n),$$

thus we have

$$\theta_1(\tau) = \exp\left(\frac{2(1-\tau)\log n}{2+\log n}\right).$$

The reader can verify easily that

$$\exp(2(1-\tau)) \leq \theta_1(\tau) \leq \exp(1-\tau).$$

□

The search directions throughout this section are based on the SR-proximity function

$$\Psi_q(v) = \Phi_q(x, s, \mu) := \frac{e^T v^2 - n}{2} + \frac{e^T v^{1-q} - n}{q-1}, \tag{17}$$

where $1 \leq q \leq 1 + \log n$.

One can prove easily the following technical lemma that is useful in predicting the reduction of duality gap.

Lemma 3.5 *Let $1 \leq q \leq 1 + \log n$, then the global minimum of $\Psi_q(v)$ in (17), μ_q^*, is decreasing w.r.t. q.*

Note that our SR-PC algorithm uses the specific SR-proximity function $\Phi_\ell(x, s, \mu)$ to define the neighborhood, while the search directions are given by the SR-proximity function $\Phi_q(x, s, \mu)$, where $1 \leq q \leq 1 + \log n$. Due to the choice of the proximity function, we can write system (8) as:

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu^{\frac{q+1}{2}}(x^{\frac{1-q}{2}})s^{\frac{1-q}{2}} - xs. \end{aligned} \tag{18}$$

Let us denote the solution of system (18) by $(\Delta x(\mu), \Delta y(\mu), \Delta s(\mu))$. Let $\mu = \mu_\ell^*$. The following two lemmas discuss the change of the duality gap along the search direction

$$(\Delta x(\mu_\ell^*), \Delta y(\mu_\ell^*), \Delta s(\mu_\ell^*)).$$

Lemma 3.6 *Let $(\Delta x(\mu_\ell^*), \Delta y(\mu_\ell^*), \Delta s(\mu_\ell^*))$ be the solution of system (18) with $\mu = \mu_\ell^*$ and let $1 \leq q \leq 1 + \log n$. Then the relation*

$$x^T\Delta s(\mu_\ell^*) + s^T\Delta x(\mu_\ell^*) \leq 0$$

holds. If $q = 1 + \log n$ then equality holds in this relation.

Proof: Using the third equation in (18) one has

$$x^T\Delta s(\mu_\ell^*) + s^T\Delta x(\mu_\ell^*) = (\mu_\ell^*)^{\frac{q+1}{2}}(x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}} - x^T s \leq 0,$$

where the last inequality follows from Lemma 3.5. This completes the proof of the lemma. □

Corollary 3.7 *If $\mu = \mu_\ell^*$, then the duality gap decreases, i.e.,*

$$(x + \alpha\Delta x(\mu_\ell^*))^T(s + \alpha\Delta s(\mu_\ell^*)) = x^T s + \alpha\left(x^T\Delta s(\mu_\ell^*) + s^T\Delta x(\mu_\ell^*)\right) \leq x^T s.$$

When μ_ℓ^t is the target value, analogue to Lemma 3.6, we can get the following result.

Lemma 3.8 *Let $(\Delta x(\mu_\ell^t), \Delta y(\mu_\ell^t), \Delta s(\mu_\ell^t))$ be the solution of system (18), where $\mu = \mu_\ell^t$ is defined by equation (15) and let $1 \leq q \leq 1 + \log n$. Then*

$$x^T\Delta s(\mu_\ell^t) + s^T\Delta x(\mu_\ell^t) = (\mu_\ell^t)^{\frac{q+1}{2}}(x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}} - x^T s,$$

and

$$(x + \alpha\Delta x(\mu_\ell^t))^T(s + \alpha\Delta s(\mu_\ell^t)) = (x + \alpha\Delta x)^T(s + \alpha\Delta s) = x^T s\left(1 - \alpha + \frac{(\mu_\ell^t)^{\frac{q+1}{2}}\alpha}{\mu_g(\mu_q^h)^{\frac{q-1}{2}}}\right), \tag{19}$$

where

$$\mu_q^h = \left(\frac{n}{(x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}}}\right)^{\frac{2}{q-1}}.$$

Remark 3.9 *If $\mu_\ell^t \sim \mu_q^h$, then (19) implies that the search direction based on our specific SR-proximity function and the standard Newton direction will predict the change of the duality gap almost in the same way. However, if $\mu_q^t \ll \mu_q^h$, then the ratio $\frac{(\mu_q^t)^{\frac{q+1}{2}}}{\mu_g(\mu_q^h)^{\frac{q-1}{2}}}$ is very small and for the SR search direction the duality gap reduction is much larger than it would be for the standard Newton direction.*

For notational convenience, we define

$$\sigma_\ell = \left(\sum_{i=1}^{n}(v_i - {v_i}^{-1-\log n})(v_i - {v_i}^{-q})\right)^{\frac{1}{2}} \tag{20}$$

and

$$\sigma_q = \left\|v - v^{-q}\right\| = \|d_x + d_s\|,$$

where $1 \le q \le 1 + \log n$. It is easy to check that σ_ℓ and σ_q are well defined.

The following two lemmas specify the relation between σ_ℓ, σ_q and $\|(d_x, d_s)\|$.

Lemma 3.10 *For $1 \le q \le 1 + \log n$ one has $\sigma_q \le \sigma_\ell$.*

Proof: If $v_i \ge 1$ then $v_i - {v_i}^{-q} \le v_i - {v_i}^{-\log n - 1}$, otherwise ${v_i}^{-q} - v_i \le {v_i}^{-1-\log n} - v_i$, thus $\sigma_q \le \sigma_\ell$. □

Lemma 3.11 *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of (18), where $\mu = \mu_\ell^*$ as it is defined in Proposition 3.2 and $1 \le q \le 1 + \log n$. Then*

$$\sigma_\ell \ge \|(d_x, d_s)\|.$$

Proof: From the definition of the search directions we have

$$\sigma_\ell^2 = \sum_{i=1}^{n}(v_i - {v_i}^{-1-\log n})(v_i - {v_i}^{-q}) \ge \sum_{i=1}^{n}(v_i - {v_i}^{-q})(v_i - {v_i}^{-q}) = \|(d_x, d_s)\|^2,$$

where the last inequality follows from the orthogonality of d_x and d_s. □

Now in what follows we derive bounds for the smallest coordinate of v.

Lemma 3.12 *Suppose that the present iterate (x, s) is on the boundary of the neighborhood $\mathcal{N}(\eta(n, \tau))$ (defined by (13)), where $\mu = \mu_\ell^*$ as it is defined in Proposition 3.2. Then for $\tau \ge 2$ one has*

$$v_{\min} \ge \exp(-\tau), \tag{21}$$

$$v_{\min}^{1+\log n}\sigma_\ell \ge \theta_2(\tau)\exp\left(\frac{q - 1 - \log n}{2}\right), \tag{22}$$

where $\theta_2(\tau) = \min\left\{(\exp(\tau - 1) - 2)^{\frac{1}{2}}, \exp(1 - \tau)(1 - \exp(-\frac{1}{2}))\right\}$.

Proof: Using the definition of μ_ℓ^* we have

$$v_{\min}^{-\log n} \leq \left\| v^{-\frac{\log n}{2}} \right\| = \|v\|^2 = \frac{x^T s}{\mu_\ell^*} = n\theta_1(\tau),$$

that by Lemma 3.4 implies

$$v_{\min} \geq \exp(-1)\theta_1(\tau)^{\frac{1}{\log n}} \geq \exp\left(\frac{-2\tau - \log n}{2 + \log n}\right) \geq \exp(-\tau).$$

To prove (22), first assume that $v_{\min} \leq \exp(-\frac{1}{4})$, then by using (20) we have

$$\begin{aligned} v_{\min}^{1+\log n}\sigma_\ell &\geq v_{\min}^{\frac{-q+1+\log n}{2}}(1 - v_{\min}^{2+\log n})^{\frac{1}{2}}(1 - v_{\min}^{q+1})^{\frac{1}{2}} \\ &\geq \left(1 - \exp(-\frac{1}{2})\right) v_{\min}^{\frac{-q+1+\log n}{2}} \geq \exp(1-\tau)\left(1 - \exp(-\frac{1}{2})\right)\exp\left(\frac{q-1-\log n}{2}\right). \end{aligned}$$

If $v_{\min} \geq \exp(-\frac{1}{4})$, then

$$v_{\min}^{1+\log n}\sigma_\ell \geq \left(\frac{1}{\theta_1(\tau)} - 2\right)^{\frac{1}{2}} n^{\frac{1}{4}} \geq (\exp(\tau - 1) - 2)^{\frac{1}{2}} n^{\frac{1}{4}},$$

where the first inequality follows from the inequality

$$\sigma_\ell \geq \left\| v - v^{-1} \right\| \geq \left(\frac{1}{\theta_1(\tau)} - 2\right)^{\frac{1}{2}} n^{\frac{1}{2}},$$

and the second inequality follows from the bound given for $\theta_1(\tau)$ in Lemma 3.4. This completes the proof. □

Remark 3.13 *If $q = 1 + \log n$ and $\tau \geq 2$, then*

$$v_{\min}^{1+\log n}\sigma_\ell \geq \theta_2(\tau).$$

Remark 3.14 *If $q = 1$ and $\tau \geq 2$, then*

$$v_{\min}^{1+\log n}\sigma_\ell \geq \frac{\theta_2(\tau)}{\sqrt{n}}.$$

The following lemma gives a lower bound for the maximal feasible step size.

Lemma 3.15 *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (18) where $1 \leq q \leq 1 + \log n$ and $\mu = \mu_\ell^*$ is defined by (15). Then the maximal feasible step size, $\alpha_{\max}$, satisfies*

$$\alpha_{\max} \geq \bar{\alpha} = \frac{1}{\sigma_\ell \exp(\tau)}.$$

Proof: We know that

$$v(\alpha_{\max}) = (v + \alpha_{\max} d_x)^{\frac{1}{2}}(v + \alpha_{\max} d_s)^{\frac{1}{2}} = v(e + \alpha_{\max} v^{-1} d_x)^{\frac{1}{2}}(e + \alpha_{\max} v^{-1} d_s)^{\frac{1}{2}},$$

that is nonnegative if

$$e + \alpha_{\max} v^{-1} d_x \geq 0, \text{ and } \quad e + \alpha_{\max} v^{-1} d_s \geq 0.$$

These inequalities imply

$$\alpha_{\max} \geq \frac{1}{\|(v^{-1}d_x, v^{-1}d_s)\|}.$$

We also know that

$$\left\|(v^{-1}d_x, v^{-1}d_s)\right\| \leq \frac{\|(d_x, d_s)\|}{v_{\min}} \leq \sigma_\ell \exp(\tau).$$

This completes the proof of the lemma. □

4 An Adaptive Large Neighborhood SR-PC-IPM

In this section we present our new algorithm. At each iteration we have a predictor step and a corrector step. In the corrector step we recenter to a smaller neighborhood. In the predictor step we make either an adaptive SR step or an affine scaling step in order to reduce the duality gap as much as possible in the given large neighborhood. In the predictor step we compare the decrease of the duality gap for the SR and the affine scaling steps. If the reduction of the duality gap for the affine scaling step is larger than the one theory guarantees for the SR step, then we make an affine scaling step. Otherwise we do an adaptive SR step. With this adaptive choice of the predictor step we preserve the best known polynomial complexity of large-update SR-IPMs for the case $q = 1 + \log n$ and quadratic convergence. Our adaptive large neighborhood SR-PC algorithm is as follows.

Algorithm SR-PC:
Adaptive Large Neighborhood SR-PC-IPM

Input:
A proximity parameter $\tau \geq 2$ and $\eta(n,\tau) = \frac{n^\tau - n}{\log n}$;
an accuracy parameter $\varepsilon > 0$; $1 \leq q \leq 1 + \log n$;
$(x,s) = (x^0, s^0)$ such that $\Phi_\ell(x,s,\mu_g) \leq \eta(n,\tau)$.

begin
while $x^T s \geq \epsilon$ **do**
begin
Corrector step
Solve (18) with $\mu = \mu_\ell^*$ and choose a step size α_1 such that
$\alpha_1 = \text{argmin}_\alpha \Phi_\ell(x(\alpha), s(\alpha), \mu_\ell^*)$.
Let $(x,y,s) = (x(\alpha_1), y(\alpha_1), s(\alpha_1))$.
end
begin
Predictor step
(**Affine scaling step**) Solve (14) and choose the maximal step size α_3 such that
$(x(\alpha_3), y(\alpha_3), s(\alpha_3)) \in \mathcal{N}(\eta(n,\tau))$.
If $1 - \alpha_3 \leq 1 - \alpha_3^* + \alpha_3^* \frac{(\mu_\ell^t)^{\frac{q+1}{2}}}{\mu_g (\mu_q^h)^{\frac{q-1}{2}}}$, then let[a] $\alpha = \alpha_3$ and accept the affine scaling step.
else
(**SR-step**) Solve (18) with $\mu = \mu_\ell^t$ derived from (15).
Determine the maximal step size α_2 such that
$(x(\alpha_2), y(\alpha_2), s(\alpha_2)) \in \mathcal{N}(\eta(n,\tau))$ and let $\alpha = \alpha_2$.
end
Let $(x,y,s) = (x(\alpha), y(\alpha), s(\alpha))$.
end

[a] The value of α_3^* is given in Theorem 4.2 as the warranted step size in the SR step. The duality gap reduction formula is given in Lemma 3.8 for the SR step.

4.1 The Corrector Step

In this subsection we estimate the decreasing of the proximity function in the corrector step when μ_ℓ^* is the target value, and we compute a lower bound for the step size that guarantees sufficient reduction of the proximity function. The proof of the following theorem, except the last two inequalities at the end, is analogous to the proof of Lemmas 3.3.2 and 3.3.3 in [11].

Theorem 4.1 *Let us assume that the present iterate is in the neighborhood $\mathcal{N}(\eta(n,\tau))$, and let $(\Delta x, \Delta y, \Delta s)$ be the solution of (18) with $\mu = \mu_\ell^*$ and $1 \leq q \leq 1 + \log n$. Then, for*

$$\alpha_1^* = \frac{\theta_2(\tau) v_{\min}}{4\sigma_\ell \log n \exp(\frac{-q+1+\log n}{2})} \quad \text{and} \quad \alpha_2^* = \frac{\theta_2^2(\tau)}{16 \log n \exp(-q+1+\log n)}$$

we have

$$\Phi_\ell(x(\alpha_1^*), s(\alpha_1^*), \mu_\ell^*) \leq \Phi_\ell(x, s, \mu_\ell^*) - \max\{\alpha_1^* \sigma_\ell^2,\ \alpha_2^* v(\alpha)_{\min}^{-\log n}\}. \tag{23}$$

Proof: Let us examine the reduction of the proximity value as a function of the step length. For this purpose let us define

$$\begin{aligned} g(\alpha) :\ &= \Phi_\ell\left(x(\alpha), s(\alpha), \mu_\ell^*\right) - \Phi_\ell(x, s, \mu_\ell^*) \\ &= \frac{\|v(\alpha)\|^2 - n}{2} + \frac{\left\|v(\alpha)^{\frac{-\log n}{2}}\right\|^2 - n}{\log n} - \frac{\|v\|^2 - n}{2} - \frac{\left\|v^{\frac{-\log n}{2}}\right\|^2 - n}{\log n}, \end{aligned} \tag{24}$$

where $v(\alpha) = \sqrt{\frac{x(\alpha)s(\alpha)}{\mu_\ell^*}} = (v + \alpha d_x)^{\frac{1}{2}}(v + \alpha d_s)^{\frac{1}{2}}$. After simplifying the expression, and the fact that $(v_i + \alpha(d_x)_i)^{\frac{-\log n}{2}}(v_i + \alpha(d_s)_i)^{\frac{-\log n}{2}}$ satisfies condition SR.2 of Definition 2.1, we have

$$\begin{aligned} g(\alpha)\ &\leq \frac{1}{2}v^T(d_x + d_s)\alpha + \frac{1}{2\log n}\sum_{i=1}^{n}[(v_i + \alpha(d_x)_i)^{-\log n} + (v_i + \alpha(d_s)_i)^{-\log n}] - \frac{\left\|v^{\frac{-\log n}{2}}\right\|^2}{\log n} \\ &:= g_1(\alpha). \end{aligned} \tag{25}$$

Using (20), the definition of σ_ℓ, the derivative of $g_1(\alpha)$ at zero can be given as

$$g_1{}'(0) = \frac{1}{2}v^T(d_x + d_s) - \frac{1}{2}(v^{-1-\log n})^T(d_x + d_s) = -\frac{\sigma_\ell^2}{2}.$$

One can prove, the same way as Lemma 3.3.3 in [11] is proved, that

$$g_1(\alpha) \leq -\frac{\sigma_\ell^2}{2}\alpha + \frac{(1 + \log n)\sigma_\ell^2}{2}\int_0^\alpha \int_0^\xi (v_{\min} - \eta\sigma_\ell)^{-2-\log n} d\eta d\xi := g_2(\alpha).$$

It is easy to see, via making use of simple calculus, that $g_2(\alpha)$ is convex and twice continuously differentiable for all α. Let $\bar{\alpha}_1$ denote the global minimizer of $g_2(\alpha)$ in the interval $(0, \bar{\alpha}]$ (see Lemma 3.15). Now as in Lemma 3.3.3 of [11] we have $\bar{\alpha}_1$ as the solution of the equation

$$-\sigma_\ell^2 + \sigma_\ell\left((v_{\min} - \alpha\sigma_\ell)^{-1-\log n} - v_{\min}^{-1-\log n}\right) = 0. \tag{26}$$

Using Lemma 1.3.1 of [11] one can derive that

$$\begin{aligned} \bar{\alpha}_1\ &\geq \frac{v_{\min}\sigma_\ell v_{\min}^{1+\log n}}{\sigma_\ell(1 + \log n)(1 + \sigma_\ell v_{\min}^{1+\log n})} \geq \frac{v_{\min}\sigma_\ell v_{\min}^{1+\log n}}{\sigma_\ell(1 + \log n)(1 + \sigma_\ell v_{\min}^{1+\log n})} \\ &\geq \frac{\theta_2(\tau)v_{\min}}{4\sigma_\ell \log n \exp(\frac{-q+1+\log n}{2})} = \alpha_1^*, \end{aligned}$$

where the last inequality follows from Lemma 3.12.

By using Lemma 1.3.3 of [11], for $\alpha = \alpha_1^*$ we have

$$\Phi_\ell(x(\alpha), s(\alpha), \mu_\ell^*) \leq \Phi_\ell(x, s, \mu_\ell^*) - \frac{1}{2}\alpha_1^*\sigma_\ell^2.$$

From (26) and Lemma 3.1 of [12] we have

$$(v_{\min} - \alpha\sigma_\ell)^{-1-\log n} \leq v_{\min}^{-1-\log n} + \sigma_\ell \leq \left(1 + \frac{1}{\theta_2(\tau)}\exp\left(\frac{-q + 1 + \log n}{2}\right)\right)\sigma_\ell.$$

Then we can write

$$\begin{aligned}\Phi_\ell(x(\alpha), s(\alpha), \mu_\ell^*) &\leq \Phi_\ell(x, s, \mu_\ell^*) - \frac{\theta_2^2(\tau)(v_{\min} - \alpha\sigma_\ell)^{-\log n}}{16 \log n \exp(-q + 1 + \log n)} \\ &\leq \Phi_\ell(x, s, \mu_\ell^*) - \alpha_2^* v(\alpha)_{\min}^{-\log n}, \end{aligned} \tag{27}$$

where $\alpha_2^* = \dfrac{\theta_2^2(\tau)}{16 \log n \exp(-q + 1 + \log n)}$ and the last inequality follows from

$$v_{\min} - \alpha\sigma_\ell \leq v(\alpha)_{\min}.$$

This completes the proof of the theorem. □

4.2 The Predictor Step

In this subsection we discuss the behavior of the search direction with the different μ values that we choose in the predictor step. We use $\mu = 0$ (affine scaling) whenever the reduction of the duality gap is at least as good as theory guarantees for the $q-$degree SR step with $\mu = \mu_\ell^t$. If the reduction is not satisfactory, we make an SR step with $\mu = \mu_\ell^t$. Then the worst case that can happen is the $q-$degree SR step. This implies that for the worst case iteration complexity analysis it suffices to discus the behavior of the $q-$degree SR step. the In what follows we compute the step size for the affine scaling step and the $q-$degree SR step with $\mu = \mu_\ell^t$.

Theorem 4.2 *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (18), where $\mu = \mu_\ell^t$ is defined by equation (15) and $1 \leq q \leq 1 + \log n$. Then, for the step size $\alpha_3^* = \dfrac{\theta_2(\tau)\exp(-\tau)}{4\sigma_\ell \log n \exp(\frac{-q+1+\log n}{2})}$, the relation*

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^t) \leq \Phi_\ell(x, s, \mu_\ell^t) - \frac{1}{2}\alpha_3^* \sigma_\ell^2$$

holds.

Proof: The proof is analogous to the proof of Theorem 4.1 except that at the end of the proof we use the bound for $v_{\min}$ from Lemma 3.12. □

Theorem 4.3 *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (18), where $\mu = \mu_\ell^t$ is defined by equation (15) and $1 \leq q \leq 1 + \log n$. Then for step size α_3^*, the relation*

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_g(\alpha_3^*)) \leq \eta(n, \tau) \tag{28}$$

holds.

Proof: By Theorem 4.2, we know that the step size α_3^* is strictly feasible and

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^*(\alpha_3^*)) \leq \Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^t) \leq \Phi_\ell(x, , s, \mu_\ell^t) = \frac{(\tau_1 - 1)n}{2}.$$

Thus, $\mu_\ell^t(\alpha_3^*)$ is well defined and so by the definition of τ_1 (see page 6) we have

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^t(\alpha_3^*)) = \frac{(\tau_1 - 1)n}{2}.$$

Using Lemma 3.2, and that $\mu_\ell^t(\alpha_3^*) \leq \mu_\ell^h(\alpha_3^*)$, we have

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^h(\alpha_3^*)) \leq \frac{(\tau_1 - 1)n}{2},$$

that by Lemma 2.3 of [19] is equivalent to

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_g(\alpha_3^*)) \leq \eta(n, \tau).$$

This completes the proof of the theorem. □

Remark 4.4 *If we make an affine scaling step in the predictor step then the duality gap reduction is at least as big as the warranted duality gap reduction of the SR step, so for the worst case iteration complexity it suffices to consider that in all iterations we are making an SR step.*

The following technical lemma is an important tool to prove the polynomial complexity of our new algorithm. One can find a slightly different version of this lemma in [19].

Lemma 4.5 *Let $v_+ = \frac{v}{\sqrt{1-\theta}}$ for some $\theta \in (0,1)$. Then we have:*

$$\Psi_\ell(v_+) \leq \frac{\Psi_\ell(v)}{1-\theta} + \frac{n\theta}{(1-\theta)}.$$

Proof: From the definition of the proximity function we have

$$\begin{aligned}
\Psi_\ell(v_+) &= \frac{\|v_+\|^2 - n}{2} + \frac{\left\|v_+^{\frac{-\log n}{2}}\right\|^2 - n}{\log n} \\
&= \frac{\frac{1}{1-\theta}\|v\|^2 - n}{2} + \frac{(1-\theta)^{\frac{\log n}{2}}\left\|v^{\frac{-\log n}{2}}\right\|^2 - n}{\log n} \\
&= \frac{1}{1-\theta}\left(\frac{\|v\|^2 - n}{2} + \frac{\left\|v^{\frac{-\log n}{2}}\right\|^2 - n}{\log n}\right) + \frac{n\theta}{2(1-\theta)} + \frac{n\theta}{(\log n)(1-\theta)} \\
&\quad + \left((1-\theta)^{\frac{\log n}{2}} - \frac{1}{1-\theta}\right)\frac{\left\|v^{\frac{-\log n}{2}}\right\|^2}{\log n} \qquad (29) \\
&\leq \frac{\Psi_\ell(v)}{1-\theta} + \frac{n\theta}{(1-\theta)}. \qquad (30)
\end{aligned}$$

This completes the proof of the lemma. □

By applying Lemma 4.5 to Theorem 4.2, we can prove the following theorem.

Theorem 4.6 *Let $\tau \geq 2$ and $(\Delta x, \Delta y, \Delta s)$ be the solution of system (18) as defined in Algorithm 2, and let α_3^* be the default step size as defined in Theorem 4.2. Then*

$$\Phi_\ell\left(x(\alpha_3^*), s(\alpha_3^*), (1-\bar{\theta})\mu_\ell^t\right) \leq \Phi_\ell(x, s, \mu_\ell^t),$$

where

$$\bar{\theta} = \frac{2\alpha_3^*}{\tau_1}\sqrt{\frac{\tau - 1}{n}},$$

and τ_1 is defined on page 6.

Proof: From Lemma 4.5, it can be seen that to prove the theorem it suffices to choose θ satisfying the inequality

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^t) + \frac{n\theta}{2} \leq (1-\theta)\Phi_\ell(x, s, \mu_\ell^t).$$

Using Theorem 4.2 we conclude that the above inequality will certainly be satisfied if

$$\theta\Phi_\ell(x, s, \mu_\ell^t) + \frac{n\theta}{2} \leq \frac{1}{2}\alpha_3^*\sigma_\ell^2. \tag{31}$$

Since $\sigma_\ell \geq \sigma_q$ for $1 \leq q \leq 1 + \log n$, we have

$$\theta\Phi_\ell(x, s, \mu_\ell^t) + \frac{n\theta}{2} \leq \frac{1}{2}\alpha_4^*\sqrt{2\Psi_\ell(v)}, \tag{32}$$

where $\alpha_4^* = \sigma_\ell\alpha_3^*$. Recalling the fact that $\Phi_\ell(x, s, \mu_\ell^t) = \frac{(\tau_1-1)n}{2}$ (the bound for the proximity measure on page 6), we can rewrite inequality (31) as:

$$\theta\left(\frac{(\tau_1-1)n}{2} + \frac{n}{2}\right) \leq \frac{1}{2}\alpha_4^*\sqrt{(\tau_1-1)n}.$$

This relation implies that if we choose

$$\bar{\theta} = \frac{\alpha_4^*}{\tau_1}\sqrt{\frac{\tau-1}{n}},$$

then

$$\Phi_\ell\left(x(\alpha^*), s(\alpha^*), (1-\bar{\theta})\mu_\ell^t\right) \leq \Phi_\ell(x, s, \mu_\ell^t),$$

that completes the proof. □

Now we can proceed to discuss the complexity of Algorithm SR-PC that follows from the reduction of the duality gap in the predictor step by the $q-$degree SR directions. By the choice of μ_ℓ^t we know that the proximity function $\Phi_\ell(x, s, \mu_\ell^t)$ keeps invariant for all the iterates. Let us denote by $(\mu_\ell^t)^+$ the target parameter value after one step. Then we have

$$\Phi_\ell(x, s, \mu_\ell^t) = \Phi_\ell(x(\alpha^*), s(\alpha^*), (\mu_\ell^t)^+).$$

On the other hand, form Theorem 4.6 we have

$$\Phi_\ell(x(\alpha^*), s(\alpha^*), (1-\theta)\mu_\ell^t) \leq \Phi_\ell(x(\alpha^*), s(\alpha^*), (\mu_\ell^t)^+).$$

Since the proximity function is a convex function w.r.t. μ, we have

$$(\mu_\ell^t)^+ \leq (1-\bar{\theta})\,\mu_\ell^t. \tag{33}$$

Now we are ready to give the complexity of Algorithm SR-PC.

Theorem 4.7 *After at most*

$$\left\lceil \frac{1}{\bar{\theta}} \log \frac{(\exp(\tau-1)+1)n}{\epsilon} \right\rceil.$$

iterations Algorithm SR-PC terminates with a solution for which $x^T s \leq \epsilon$.

Proof: Since we ensured that the worst case duality gap reduction is the one that the SR direction gives, then without loss of generality we may assume that in the predictor step we always makes a SR step. Using inequality (33) and Corollary 3.8 after each iterations (predictor-corrector step), after at most

$$\left\lceil \frac{1}{\bar{\theta}} \log \frac{(\exp(\tau-1)+1)n}{\epsilon} \right\rceil$$

iterations we have $\mu_\ell^t \leq \dfrac{\epsilon}{(\exp(\tau-1)+1)\,n}$, and then Lemma 3.1 implies $x^T s \leq \epsilon$, that completes the proof of the theorem. □

Corollary 4.8 *If $q = 1 + \log n$, then the number of iteration of Algorithm SR-PC is at most*

$$O\left(\sqrt{n}\log n \log \frac{(\exp(\tau-1)+1)n}{\epsilon}\right).$$

Corollary 4.9 *If $q = 1$, then the number of iterations of Algorithm SR-PC is at most*

$$O\left(n\log n \log \frac{(\exp(\tau-1)+1)n}{\epsilon}\right).$$

5 Quadratic Convergence

In this section we prove the quadratic convergence of Algorithm SR-PC. In case of monotone LCPs, Ye and Anstreicher [23] have proved for predictor-corrector algorithms for sufficiently small μ_g the relation

$$|(\Delta x)_i(\Delta s)_i| = \mathcal{O}({\mu_g}^2),\ i = 1, \cdots, n \tag{34}$$

holds. Since LO is a special case of monotone LCP, relation (34) is valid for LO as well. Since in the predictor step we have two choices for the target value, it suffices to discuss the case when the affine scaling direction is used as the predictor direction. The reason is that close to optimality affine scaling is always dominant to the SR direction. Then we have

$$|(d_x)_i(d_s)_i| = \frac{(\Delta x)_i(\Delta s)_i}{\mu^*} = \mathcal{O}(\mu_g).$$

In the following lemma we give a lower bound for the step size of the affine scaling predictor step.

Lemma 5.1 *Let (x, s) be an iterate in Algorithm SR-PC. If the present duality gap is sufficiently small, such that (34) holds, then the step size α in the predictor step satisfies $\alpha \geq 1 - O(\mu_g)$.*

Proof: First we give an estimate for the maximal feasible step size in a predictor step. The new iterate is strictly feasible if both $v+\alpha d_x$ and $v+\alpha d_s$ are strictly feasible. Then, the maximal feasible step size $\alpha_{\max}$ satisfies the following inequality

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

or, equivalently

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

Since $|d_x d_s| = \mathcal{O}(\mu_g)$, we can conclude that $\alpha_{\max} \geq 1 - \mathcal{O}(\mu_g)$.

Now we prove that the step size in the affine scaling step also satisfies

$$\alpha \geq 1 - \mathcal{O}(\mu_g).$$

Let us define $v_+(\alpha) = \sqrt{\frac{x(\alpha)s(\alpha)}{(1-\alpha)\mu}}$, then we have

$$\begin{aligned}
\Phi(x(\alpha), s(\alpha),\ (1-\alpha)\mu) - \Phi(x,s,\mu) &= \frac{\left\| v_+(\alpha)^{\frac{-\log n}{2}} \right\|^2}{\log n} - \frac{\left\| v^{\frac{-\log n}{2}} \right\|^2}{\log n} \\
&= \frac{1}{\log n} \sum_{i=1}^{n} v_i^{-\log n} \left(\left(1 + \frac{\alpha^2}{1-\alpha} v_i^{-2} (d_x)_i (d_s)_i \right)^{\frac{-\log n}{2}} - 1 \right) \\
&\leq \frac{v_{\min}^{-\log n}}{\log n} \left(\left(1 + \frac{\alpha^2}{1-\alpha} \sum_{(d_x)_i (d_s)_i \leq 0} v_i^{-2} (d_x)_i (d_s)_i \right)^{\frac{-\log n}{2}} - 1 \right) \\
&\leq \frac{v_{\min}^{-\log n}}{\log n} \left(\left(1 - \frac{\alpha^2}{1-\alpha} \mathcal{O}(\mu_g) \right)^{\frac{-\log n}{2}} - 1 \right).
\end{aligned}$$

This inequality shows that, for sufficiently small μ_g we have

$$\Phi(x(\alpha), s(\alpha), (1-\alpha)\mu) - \Phi(x,s,\mu) \leq \alpha_2^* v_{\min}^{-\log n}, \tag{35}$$

where α has to satisfy the following inequality

$$\left(1 - \frac{\alpha^2}{1-\alpha} \mathcal{O}(\mu_g) \right)^{\frac{-\log n}{2}} - 1 \leq \frac{\theta_2^2(\tau)}{16 \exp(-q + 1 + \log n)}.$$

That is equivalent to find an α such that

$$\begin{aligned}
O(\mu_g)\alpha^2 \ &+ \left(1 - \left(1 + \frac{\theta_2^2(\tau)}{16 \exp(-q+1+\log n)} \right)^{\frac{-2}{\log n}} \right) \alpha \\
&-1 + \left(1 + \frac{\theta_2^2(\tau)}{16 \exp(-q+1+\log n)} \right)^{\frac{-2}{\log n}} \leq 0.
\end{aligned}$$

This equation has two roots. Based on the positive root, analogous to the proof of Lemmas II.64 and II.65 of [18] one can show that $\alpha \geq 1 - O(\mu_g)$. Finally, by Lemma 4.1 the proof is completed. □

Using Lemma 5.1 we have

$$\mu_g^+ = (1-\alpha)\mu_g = \mathcal{O}\left((\mu_g^k)^2 \right),$$

where $\mu_g^+ = \frac{x(\alpha)^T s(\alpha)}{n}$. Our quadratic convergence result is the following.

Theorem 5.2 *Let the iterate (x^k, y^k, s^k) generated by Algorithm SR-PC. When μ_g is sufficiently small, the algorithm is quadratically convergent in the sense that $\mu_g^{k+1} = \mathcal{O}\left((\mu_g^k)^2 \right)$ and any accumulation point of the sequence (x^k, s^k) is a strictly complementary optimal solution of the problem.*

Proof: The quadratic convergence of the algorithm follows from Lemma 5.1. The proof of the convergence to an accumulation point and the proof of their properties is analogous to the proof of Theorem 5.14 in [21] and we omitted it here. □

6 Conclusion

A new family of predictor-corrector SR-IPM's in a large neighborhood of the central path for LO is proposed. Our new algorithms use different search directions with a single proximity function that allows to use wide neighborhoods. An $\mathcal{O}\left(\sqrt{n}\log n\exp(\frac{-q+1+\log n}{2})\log\frac{n}{\epsilon}\right)$ worst case iteration complexity result with quadratic convergence of the algorithm is proved. For the case $q = 1+\log n$ this complexity matches the result of [15] and for the case an $q = 1+2\log(\log n)$ it has an $\mathcal{O}(n\log\frac{n}{\epsilon})$ iteration complexity.

Acknowledgments: This research was supported by MITACS project, an NSERC discovery grant, and the CRC program. The first author would also like to thank the Iranian Ministry of Science, Research and Technology for supporting his research. The authors are grateful to F.A. Potra for construction comments on an earlier draft of this paper.

References

[1] K.M. Anstreicher and R.A. Bosch. A new infinity-norm path following algorithm for linear programming, *SIAM J. on Optimization*, **5(2)** (1995), 236-246.

[2] D.A. Bayer and J.C. Lagarias. Karmarkar's linear programming algorithm and Newton's method, *Math. Programming*, **50** (1991), 291-330.

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

[4] 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** (1996), 570-586.

[5] J. Ji, F.A. Potra, and S. Huang. A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence, *Journal of Optimization Theory and Applications,* **84(1)** (1995), 187-199.

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

[7] N. Megiddo. Pathways to the optimal set in linear programming. In: N. Megiddo, editor, *Progress in Mathematical Programming: Interior Point and Related Methods*, 1989, 131-158. Springer Verlag, New York. Identical version in: *Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan*, 1986, 1-35.

[8] S. Mehrotra. On the implementation of a (primal-dual) interior point method. *SIAM J. on Optimization*, **2** (1992), 575-601.

[9] S. Mizuno. M.J. Todd, and Y. Ye. On adaptive step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, **18** (1993), 964-981.

[10] 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 extensions. *Mathematics of Operations Research*, **15** (1990), 191-214.

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

[12] J. Peng, C. Roos, and T. Terlaky. A new and efficient large-update interior-point method for linear optimization. *J. of Computational Technology*, **6** (2001), 61-80.

[13] J. Peng, C. Roos, and T. Terlaky. New polynomial primal-dual interior point methods for second-order conic optimization based on self-regular proximities. *SIAM J. on Optimization*, **13** (2002), 179-203.

[14] J. Peng and T. Terlaky. A dynamic large-update primal-dual interior-point method for linear optimization. *Optimization Methods and Software*, **17** (2002), 1077-1104.

[15] J. Peng, T. Terlaky, and Y.B. Zhao. A predictor-corrector algorithm for linear optimization based on a specific self-regular proximity function. Submitted to: *SIAM J. on Optimization,* 2003.

[16] F.A. Potra. A superlinearly convergent predictor-corrector method for degenrate LCP in a wide neighborhood of the central path, *Math. Programming,* series A, **100(2)** (2004), 317 - 337.

[17] F.A. Potra and X. Liu. Predictor-corrector methods for sufficient linear complementarity problems in a wide neighborhood of the central path, *Technical Report* 2003*, Department of Mathematics and Statistics, University of Maryland, USA.*

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

[19] M. Salahi and T. Terlaky. An adaptive self-regular proximity based large-update interior point method for linear optimization. *Technical Report* 2003/4*, Advanced Optimization Lab. Department of Computing and Software, McMaster University..*

[20] 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 Modeling and Optimization: Proceeding of the 12th IFIP Conference*, Budapest, Hungary, September 1985, Volume 84, Lecture Notes in Control and Information Sciences, pp. 866-876. Springer Verlag, Berlin, 1986.

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

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

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

[24] Y. Zhang. On the polynomiality of the Mehrotra-type predictor-corrector interior point algorithms. *Math. Programming*, **68** (1998), 303-317.