A SELF-REGULAR NEWTON BASED ALGORITHM FOR LINEAR OPTIMIZATION

M. SALAHI[1]

Abstract

In this paper, using the framework of self-regularity, we propose a hybrid adaptive algorithm for the linear optimization problem. If the current iterates are far from a central path, the algorithm employs a self-regular search direction, otherwise the classical Newton search direction is employed. This feature of the algorithm allows us to prove a worst case iteration bound. Our result matches the best iteration bound obtained by the pure self-regular approach and improves on the worst case iteration bound of the classical algorithm.

1. Introduction

Since Karmarkar's landmark paper [3], interior point methods (IPMs) have become a highly active research area, producing a large quantity of research results, and several powerful optimization packages based on IPMs have been developed; see [1, 2, 5, 7, 8, 10–12] and references therein. By introducing the class of self-regular (SR) IPMs [6], Peng *et al.* significantly reduced the worst case iteration bound of primal–dual IPMs. In the best case, their new approach achieves an $\mathcal{O}(\sqrt{n} \log n \log(x^0)^T s^0/\epsilon)$ worst case iteration bound compared to the $\mathcal{O}(n \log(x^0)^T s^0/\epsilon)$ iteration bound of the classical approaches. In this paper, using the framework of self-regularity, we propose a hybrid adaptive algorithm for linear optimization (LO) which achieves an $\mathcal{O}(\sqrt{n} \log n \log(x^0)^T s^0/\epsilon)$ worst case iteration bound by utilizing both the classical and SR search directions. Our result improves on classical algorithms while matching the order of the pure SR approach. The search directions in our new algorithm are defined based on the current iterate, which is not the case in the pure SR or classical approaches.

---

[1]Department of Mathematics, University of Guilan, Rasht, Iran; e-mail: salahim@guilan.ac.ir, salahi.maziar@gmail.com.

Throughout the paper we deal with the standard form of the LO problem,

$$\min_x \left\{ c^T x : Ax = b, \, x \ge 0 \right\}, \tag{P}$$

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

$$\max_{y,s} \left\{ b^T y : A^T y + s = c, \, s \ge 0 \right\}. \tag{D}$$

It is customary in IPM theory to assume that both (P) and (D) satisfy the interior point condition (IPC) [7], that is, there exists an (x^0, s^0, y^0) such that

$$Ax^0 = b, \; x^0 > 0 \quad \text{and} \quad A^T y^0 + s^0 = c, \, s^0 > 0.$$

Finding an optimal solution of (P) and (D) is equivalent to solving the system

$$\begin{aligned} Ax &= b, & x &\ge 0, \\ A^T y + s &= c, & s &\ge 0, \\ xs &= \mathbf{0}, \end{aligned} \tag{1.1}$$

where xs is the componentwise product of vectors x and s. The basic idea of primal–dual IPMs is to replace the third equation in (1.1) by the parameterized equation $xs = \mu \mathbf{1}$, where $\mathbf{1}$ is a vector of 1s. This leads to the system

$$\begin{aligned} Ax &= b, & x &\ge 0, \\ A^T y + s &= c, & s &\ge 0, \\ xs &= \mu \mathbf{1}. \end{aligned} \tag{1.2}$$

If the IPC holds, then System (1.2) has a unique solution for each $\mu > 0$. This solution, denoted by $(x(\mu), y(\mu), s(\mu))$, is called the μ-centre of the primal–dual pair (P) and (D). The set of μ-centres gives *the central path* of (P) and (D) [4, 9] and it has been shown that the limit of the central path (as μ goes to zero) exists and it converges to an optimal solution [7]. Primal–dual IPMs follow the central path $(x(\mu), s(\mu))$ approximately by staying in a certain neighbourhood of it and approach the optimal set as μ goes to zero. To do so, starting from a strictly feasible point (x, y, s), they solve the following system for $\mu = \beta \mu_g$:

$$\begin{aligned} A\Delta x &= 0, \quad A^T \Delta y + \Delta s = 0, \\ s\Delta x + x\Delta s &= \mu \mathbf{1} - xs, \end{aligned} \tag{1.3}$$

where $\mu_g = x^T s / n$ denotes the current central path parameter value corresponding to the duality gap and $0 < \beta < 1$. Under the assumption that $\text{rank}(A) = m$, the Newton system (1.3) has a unique solution. The step size is chosen so that the new iterates $(x + \alpha \Delta x, y + \alpha \Delta y, s + \alpha \Delta s)$ give a sufficient complementarity gap reduction, while staying in a predefined neighbourhood of the central path. We repeat this procedure until the complementarity gap equals or decreases below a prescribed tolerance ϵ, after which we have reached an ϵ-solution of (P) and (D) [7, 10].

The rest of the paper is organized as follows. In Section 2, we introduce the family of SR functions and SR-IPMs for LO. In Section 3, we first introduce a specific SR proximity measure, and then we discuss the role of the parameter μ in the proximity function and how we adaptively choose this parameter. Finally, we outline our new algorithm. In Section 4, we prove the polynomial iteration bound of the new algorithm. To improve the readability of the paper we have relegated some of the proofs to an appendix.

1.1. Conventions

- $(x^{-q})^T s^{-q} = \sum_{i \in \mathcal{I}} x_i^{-q} s_i^{-q}$, where $q \geq 1$ and $\mathcal{I} = \{1, \ldots, n\}$.
- For any $x = (x_1, x_2, \ldots, x_n)^T \in R^n$, $x_{\min} = \min\{x_1, x_2, \ldots, x_n\}$.
- R^n_+ denotes the nonnegative orthant and R^n_{++} denotes the interior of R^n_+.

2. SR functions and SR-IPMs

The family of univariate SR functions is defined as follows.

DEFINITION 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 for $t > 0$ and $\psi(1) = \psi'(1) = 0$. Furthermore, there exist positive constants $\nu_2 \geq \nu_1 > 0$ and $p \geq 1, q \geq 1$ such that

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

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

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

If $\psi(t)$ is SR, then the parameter q is called the *barrier degree* and the parameter p is called the *growth degree* of $\psi(t)$. Two popular families of SR functions are known: the first is given by

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

with $\nu_1 = \nu_2 = 1$; the second is given by

$$\Gamma_{p,q}(t) = \frac{t^{p+1} - 1}{p+1} + \frac{t^{1-q} - 1}{q-1}, \quad p \geq 1, q > 1, \tag{2.1}$$

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

$$\Gamma_{11}(t) = \Upsilon_{11}(t) = \frac{t^2 - 1}{2} - \log t \tag{2.2}$$

is obtained. See [6] for properties of these functions.

Let $v \in R^n_{++}$. Then an SR proximity function $\Psi \colon R^n_{++} \to R_+$ 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. In classical primal–dual IPMs one solves (1.3), while in SR-IPMs the Newton system (1.3) is modified. To define the modified system we introduce the notation

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

Then the Newton system for SR-IPMs for LO is given by

$$A\Delta x = 0, \quad A^T \Delta y + \Delta s = 0,$$
$$s\Delta x + x\Delta s = -\mu v \nabla \Psi(v), \tag{2.3}$$

where $v\nabla\Psi(v) = (v_1\psi'(v_1), \ldots, v_n\psi'(v_n))^T$ (see [6]). For ease of reference, we also scale the search directions Δx and Δs in the scaled v-space as

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

Using this, the Newton system (2.3) can be written as

$$\bar{A}d_x = 0, \quad \bar{A}^T \Delta y + d_s = 0,$$
$$d_x + d_s = -\nabla\Psi(v), \tag{2.4}$$

where $\bar{A} = (1/\mu)AV^{-1}X$, $X = \text{diag}(x)$ and $V^{-1} = \text{diag}(v^{-1})$. It is worth noting that the right-hand side of the third equation in (2.4) is the negative gradient of the SR proximity function, similar to the classical case which is the negative gradient of the classical logarithmic barrier function.

3. Proximity measure and barrier parameter

Throughout this paper we use the following SR proximity measure:

$$\Phi(x, s, \mu) := \frac{\mathbf{1}^T v^2 - n}{2} + \frac{\mathbf{1}^T v^{-\log n} - n}{\log n}. \tag{3.1}$$

This proximity measure is induced by the kernel function $\Gamma_{1q}(t)$, given by (2.1), with $q = 1 + \log n$. The generalized harmonic mean of the vectors x and s is denoted by μ_h and given by

$$\mu_h = \left(\frac{n}{(x^{-\log n/2})^T s^{-\log n/2}} \right)^{2/\log n}.$$

The following result shows that the global minimum of the proximity measure (3.1) as a function of μ, unlike the case of the primal–dual logarithmic barrier function, does not occur at μ_g.

PROPOSITION 3.1. *For any fixed point* $(x, s) \in R_{++}^n \times R_{++}^n$, *the proximity measure* $\Phi(x, s, \mu)$, *as a function of* μ, *has its global minimizer at*

$$\mu^* = \left(\frac{x^T s}{(x^{-\log n/2})^T s^{-\log n/2}} \right)^{2/(2+\log n)} = \left(\mu_g \mu_h^{\log n/2} \right)^{2/(2+\log n)}.$$

Moreover, $\Phi(x, s, \mu)$ *is a decreasing function of* μ *when* $\mu \leq \mu^*$, *and increasing when* $\mu > \mu^*$.

PROOF. By differentiating the proximity function with respect to μ we have

$$\Phi'_\mu(x, s, \mu) = -\frac{x^T s}{2\mu^2} + \frac{\mu^{(\log n/2)-1}(x^{-\log n/2})^T s^{-\log n/2}}{2}.$$

Now the statements of the proposition are obvious.   □

COROLLARY 3.2. *For any* $(x, s) \in R_{++}^n \times R_{++}^n$, *one has* $\mu_h \leq \mu^* \leq \mu_g$.

PROOF. Using the generalized harmonic-geometric mean inequality, $\mu_h \leq \mu_g$. By the definition of μ^* given in Proposition 3.1 one has the desired results.   □

In order to prove the polynomiality of the IPM based algorithm, we require keeping the iterate in a certain neighbourhood of the central path. Most practical algorithms use the negative infinity norm neighbourhood, which is defined as

$$\mathcal{N}_\infty^-(\rho) := \left\{ (x, s) > 0 \,\middle|\, Ax = b, A^T y + s = c, \left\| (v^2 - \mathbf{1})^- \right\|_\infty \leq \rho \right\},$$

where $a^- = \min(a, 0)$ and $\rho \in (0, 1)$ is a constant independent of n. The closer the ρ value to one, the larger the neighbourhood is. To achieve our goal, we define the SR neighbourhood in such a way that it contains the negative infinity norm neighbourhood and these two neighbourhoods almost match each other. Our SR neighbourhood is defined as

$$\mathcal{N}(n, \tau) := \left\{ (x, s) > 0 \,\middle|\, Ax = b, A^T y + s = c, \Phi(x, s, \mu_g) \leq \eta(n, \tau) \right\},$$

where $\eta(n, \tau)$ is a positive function that depends on a constant τ and the dimension of the underlying problem, and it will be determined in the sequel. Assume that $(x, s) \in \mathcal{N}_\infty^-(\rho)$; then, for $\mu = \mu_g$,

$$\Phi(x, s, \mu_g) = \frac{\mathbf{1}^T v^{-\log n} - n}{\log n} \leq \frac{n(1-\rho)^{-\log n/2} - n}{\log n} = \frac{n(\tau^{\log n/2} - 1)}{\log n} =: \eta(n, \tau),$$

where $\tau = 1/(1 - \rho)$. One can easily see that by this choice of $\eta(n, \tau)$ the neighbourhood $\mathcal{N}(n, \tau)$ contains the neighbourhood $\mathcal{N}_\infty^-(\rho)$. The reverse statement also holds, but for a bigger ρ value. This is shown in the next lemma.

LEMMA 3.3. *Let (x, y, s) be a point in $\mathcal{N}(n, \tau)$. Then it belongs to the set $\mathcal{N}_\infty^-(1 - e^{-2}\tau^{-1})$.*

PROOF. Let $(x, y, s) \in \mathcal{N}(n, \tau)$. Then from the definition of the proximity function, we have

$$\Phi(x, s, \mu_g) = \frac{\mathbf{1}^T v^{-\log n} - n}{\log n} \le \eta(n, \tau).$$

This further implies that for all $i \in \mathcal{I}$, $v_i^{-\log n} \le n\tau^{\log n/2}$ or $v_i \ge e^{-1}\tau^{-1/2}$. Therefore $(x, y, s) \in \mathcal{N}_\infty^-(1 - e^{-2}\tau^{-1})$. □

REMARK 2. From our discussion it follows that operating in a SR neighbourhood is like operating in a negative infinity norm neighbourhood. Moreover, we note that just the starting point of the algorithm is taken from this neighbourhood; for the later iteration we use the solvability of (3.2).

COROLLARY 3.4. *The inequalities $\Phi(x, s, \mu_g) \le \eta(n, \tau)$ and $\mu_g \le \tau\mu_h$ are equivalent.*

PROOF. The proof follows from the definition of the proximity function. □

Henceforth, without loss of generality, let us assume that there exists a feasible solution for problems (P) and (D) for which $\mu_g \le \tau\mu_h$. We now describe the adaptive choice of the barrier parameter, which is an important ingredient of our algorithm. We define the barrier parameter as the smallest positive root of the equation

$$\Phi(x, s, \mu) = \frac{(\tau - 1)n}{2}.$$

After expanding this equation we have

$$2(x^{-\log n/2})^T s^{-\log n/2}\mu^{1+\log n/2} - (2 + \tau\log n)n\mu + n\log n\mu_g = 0,$$

or

$$2\left(\frac{\mu}{\mu_h}\right)^{\log n/2}\mu - (2 + \tau\log n)\mu + \log n\mu_g = 0. \tag{3.2}$$

In the next lemma some conditions are given under which (3.2) is solvable. Its smallest positive root is denoted by μ_t.

LEMMA 3.5. *For all $(x, s) \in R_{++}^n \times R_{++}^n$ for which $\mu_g \le \tau\mu_h$, Equation (3.2) has two positive solutions, one less than or equal to μ^* and the other greater than or equal to μ^*.*

PROOF. See Appendix A. □

The following technical lemma plays a crucial role in our later analysis.

LEMMA 3.6. *For any $(x, s) \in R_{++}^n \times R_{++}^n$ for which Equation (3.2) is solvable,*

$$\tau^{\log n/(2+\log n)} \le \frac{\mu_g}{\mu_t} \le \left(\tau + \frac{2}{\log n}\right).$$

PROOF. See Appendix A. □

Although we are utilizing a specific proximity function to find the barrier parameter, the search directions throughout this paper are either based on the SR proximity function (3.1) as

$$
A\Delta x = 0, \quad A^T \Delta y + \Delta s = 0,
$$
$$
s\Delta x + x\Delta s = \mu^{1+\log n/2}(x^{-\log n/2})s^{-\log n/2} - xs,
$$

(3.3)

or the classical Newton step (1.3).

LEMMA 3.7. *Let* $(\Delta x(\mu_t), \Delta y(\mu_t), \Delta s(\mu_t))$ *be the solution of System* (3.3) *with* $\mu = \mu_t$. *Then*

$$
(x + \alpha\Delta x(\mu_t))^T(s + \alpha\Delta s(\mu_t)) = \left(1 - \alpha + \alpha\frac{\mu_t^{(2+\log n)/2}}{\mu_g(\mu_h)^{\log n/2}}\right)x^T s, \quad (3.4)
$$

where

$$
\mu_h = \left(\frac{n}{(x^{-\log n/2})^T s^{-\log n/2}}\right)^{2/\log n}.
$$

PROOF. The proof follows from the third equation of (3.3). □

LEMMA 3.8. *Let* $(\Delta x(\mu_t), \Delta y(\mu_t), \Delta s(\mu_t))$ *be the solution of System* (1.3) *with* $\mu = \mu_t$. *Then*

$$
(x + \alpha\Delta x(\mu_t))^T(s + \alpha\Delta s(\mu_t)) = \left(1 - \alpha + \alpha\frac{\mu_t}{\mu_g}\right)x^T s. \quad (3.5)
$$

PROOF. This follows from the third equation of (1.3). □

Now the question is: does the complementarity gap (3.4) or (3.5) decrease for $\alpha > 0$?

REMARK 3. Since by our construction $\mu_t < \mu_g$, then for the Newton directions, the complementarity gap always decreases. For the SR directions also by our construction we will have $\mu_t < \mu^*$, which implies that the complementarity gap decreases.

In the following lemma we give a lower bound for the smallest coordinate of v that motivates the way we determine which value of q should be utilized in the definition of the search direction, and subsequently we present our new algorithm.

LEMMA 3.9. *Let* (x, y, s) *be a feasible point, and* μ_g *and* μ_h *the corresponding parameter values. Suppose that for the given point Equation* (3.2) *has two positive roots and denote its smaller one by* μ_t. *Then*

$$
v_{\min} \geq e^{-1}\left(\tau + \frac{2}{\log n}\right)^{-1/\log n}. \quad (3.6)
$$

PROOF. Using the fact that $\mu_t \leq \mu^*$ and Lemma 3.6,

$$v_{\min}^{-\log n} \leq \left\| v^{-\log n/2} \right\|^2 \leq \|v\|^2 = \frac{\mu_g}{\mu_t} n \leq \left(\tau + \frac{2}{\log n} \right) n.$$

Then (3.6) holds.   □

In our algorithm the classical Newton direction is used whenever the actual iterate is not too far from the central path, while if the actual point is farther away, the SR search direction with barrier degree $1 + \log n$ is employed.

• **Newton step.** We employ the Newton search direction when

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

• **SR step.** We employ the SR search direction otherwise.

Now we can outline our hybrid adaptive algorithm. The term K_q^* is the reduction proved for the proximity measure in one of Theorems 4.4 and 4.5 for the corresponding q value.

Hybrid adaptive algorithm

Input:
  A neighbourhood parameter $\tau \geq 6$ and $n \geq 4$;
  an accuracy parameter $\epsilon > 0$;
  $(x^0, y^0, s^0) \in \mathcal{N}(n, \tau)$.
begin
  while $x^T s \geq \epsilon$ **do**
  Compute μ_t, the smallest positive root of Equation (3.2).
  If $v_{\min} \geq e^{-1/2} \left(\tau + \frac{2}{\log n} \right)^{-1/\log n}$, then solve (1.3) with $\mu = \mu_t$,
  otherwise solve (3.3).
  Find the maximum step size α for which
  $\Phi(x(\alpha), s(\alpha), \mu_t) \leq \Phi(x, s, \mu_t) - K_q^*$.
  Set $(x, y, s) = (x(\alpha), y(\alpha), s(\alpha))$.
  end

end

---

4. Iteration bound

In this section we derive the worst case iteration bound of our hybrid adaptive algorithm. To do so, first we prove some necessary technical lemmas. For notational

convenience, we define

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

where ℓ stands for $1 + \log n$ and $\sigma_q = \|v - v^{-q}\| = \|d_x + d_s\|$, where $q = 1$ or $1 + \log n$. It is easy to check that $\sigma_{\ell q}$ and σ_q are well defined. The following two lemmas specify the relation between $\sigma_{\ell q}, \sigma_q$ and $\|(d_x, d_s)\|$.

LEMMA 4.1. *For $q = 1$ or $q = 1 + \log n$, one has $\sigma_q \leq \sigma_{\ell q}$.*

PROOF. If $v_i \geq 1$ then $v_i - v_i^{-q} \leq v_i - v_i^{-1-\log n}$, otherwise $v_i^{-q} - v_i \leq v_i^{-1-\log n} - v_i$, thus $\sigma_q \leq \sigma_{\ell q}$. □

LEMMA 4.2. $\sigma_{\ell q} \geq \sqrt{(\tau^{\log n/(2+\log n)} - 2)n}$.

PROOF. From the definition of $\sigma_{\ell q}$,

$$\sigma_{\ell q}^2 \geq \left\| v - v^{-1} \right\|^2 = \frac{\mu_g}{\mu_t} n - 2n + \left\| v^{-1} \right\|^2 \geq \frac{\mu_g}{\mu_t} n - 2n \geq \left(\tau^{\log n/(2+\log n)} - 2 \right)n,$$

where the last inequality follows from Lemma 3.6. □

REMARK 4. It is worth noting that $\tau^{\log n/(2+\log n)} - 2 > 0$ for all $n \geq 4$ and $\tau \geq 6$.

The following lemma also is useful in proving the reduction of the proximity measure for different q values.

LEMMA 4.3. *Suppose that for the current iterates Equation* (3.2) *is solvable and $\mu = \mu_t$ is its smallest positive root. Then*

$$v_{\min} \geq e^{-1/2}\left(\tau + \frac{2}{\log n}\right)^{-1/\log n} \Rightarrow v_{\min}^{1+\log n}\sigma_{\ell 1} \geq \frac{(\tau^{\log n/(2+\log n)} - 2)^{1/2}}{e^{1/2}\left(\tau + \frac{2}{\log n}\right)^{1+1/\log n}} := C_1 \tag{4.1}$$

and

$$v_{\min} \geq e^{-1}\left(\tau + \frac{2}{\log n}\right)^{-1/\log n} \quad \Rightarrow \quad v_{\min}^{1+\log n}\sigma_{\ell\ell} \geq \frac{1}{2}. \tag{4.2}$$

PROOF. See Appendix A. □

In the next two theorems we discuss the reduction of the proximity measure after each iteration for different q values.

THEOREM 4.4. *Suppose that for the current iterates Equation* (3.2) *has two positive roots and $v_{\min} \geq e^{-1/2}(\tau + 2/\log n)^{-1/\log n}$. Furthermore, let $(\Delta x, \Delta y, \Delta s)$ be the solution of* (3.3) *with $\mu = \mu_t$ as the smaller positive root of* (3.2) *and $q = 1$. Then*

$$\Phi(x(\alpha_1^*), s(\alpha_1^*), \mu_t) \leq \Phi(x, s, \mu_t) - \frac{\alpha_1^* \sigma_{\ell 1}^2}{4},$$

where

$$\alpha_1^* = C_1 \left[e^{1/2} \left(\tau + \frac{2}{\log n} \right)^{1/\log n} (1 + C_1) \sigma_{\ell 1} \log n \right]^{-1}.$$

PROOF. See Appendix A.   □

THEOREM 4.5. *Suppose that for the current iterates Equation* (3.2) *has two positive roots and* $v_{\min} \geq e^{-1} (\tau + 2/\log n)^{-1/\log n}$. *Moreover, let* $(\Delta x, \Delta y, \Delta s)$ *be the solution of* (3.3) *with* $\mu = \mu_t$ *and* $q = 1 + \log n$. *Then*

$$\Phi(x(\alpha_2^*), s(\alpha_2^*), \mu_t) \leq \Phi(x, s, \mu_t) - \frac{\alpha_2^* \sigma_{\ell\ell}^2}{4},$$

where

$$\alpha_2^* = \left[6e \left(\tau + \frac{2}{\log n} \right)^{1/\log n} \sigma_{\ell\ell} \log n \right]^{-1}.$$

PROOF. By a proof similar to that of Theorem 4.4 and using (A.2) and inequality (4.2),

$$\bar{\alpha}_1 \geq \frac{v_{\min} \sigma_{\ell\ell} v_{\min}^{1+\log n}}{\sigma_{\ell\ell} (1 + \log n)(1 + \sigma_{\ell\ell} v_{\min}^{1+\log n})} \geq \frac{1}{6e \left(\tau + \frac{2}{\log n} \right)^{1/\log n} \sigma_{\ell\ell} \log n} =: \alpha_2^*.$$

Moreover, by [6, Lemma 1.3.3], for $\alpha = \alpha_2^*$,

$$\Phi(x(\alpha), s(\alpha), \mu_t) \leq \Phi(x, s, \mu_t) - \frac{1}{4} \alpha_2^* \sigma_{\ell\ell}^2. \qquad \Box$$

REMARK 5. Following the results of Theorems 4.4 and 4.5, and the fact that the value of the proximity function approaches infinity when the barrier parameter approaches zero or infinity, we can conclude that Equation (3.2) for the next iterate with step size $\alpha = \alpha_1^*, \alpha_2^*$, respectively, has two positive roots. We denote the smaller one by $\mu_t(\alpha)$.

REMARK 6. Following the statements of Theorems 4.4 and 4.5, and using the fact that $\mu^*(\alpha)$ is the global minimizer of the proximity function as a function of $\mu(\alpha)$ for $(x(\alpha), s(\alpha)) > 0$, one has $\mu_t(\alpha) < \mu^*(\alpha)$. Moreover, Lemma 3.7 implies that the complementarity gap after each iteration is strictly decreasing.

The following technical lemma plays a crucial role in deriving an upper bound for the number of iterations for our hybrid adaptive algorithm.

LEMMA 4.6. *Let* $0 < \theta < 1$ *and* $v^+ = v/\sqrt{1 - \theta}$. *Then*

$$\Phi(x, s, (1 - \theta)\mu_t) \leq \frac{\Phi(x, s, \mu_t)}{1 - \theta} + \frac{2n\theta}{1 - \theta}.$$

PROOF. From the definition of the proximity function,

$$
\begin{aligned}
\Phi(x, s, (1-\theta)\mu_t) &= \frac{\|v_+\|^2 - n}{2} + \frac{\left\|v_+^{-\log n/2}\right\|^2 - n}{\log n} \\
&= \frac{\frac{1}{1-\theta}\|v\|^2 - n}{2} + \frac{(1-\theta)^{\log n/2}\left\|v^{-\log n/2}\right\|^2 - n}{\log n} \\
&= \frac{1}{1-\theta}\left(\frac{\|v\|^2 - n}{2} + \frac{\left\|v^{-\log n/2}\right\|^2 - n}{\log n}\right) \\
&\quad + \frac{n\theta}{2(1-\theta)} + \frac{n\theta}{\log n(1-\theta)} \\
&\quad + \left((1-\theta)^{\log n/2} - \frac{1}{1-\theta}\right)\frac{\left\|v^{-\log n/2}\right\|^2}{\log n} \\
&\leq \frac{\Phi(x, s, \mu_t)}{1-\theta} + \frac{2n\theta}{1-\theta}.
\end{aligned}
$$

This completes the proof of the lemma. $\qquad\square$

LEMMA 4.7. *Suppose that $\tau \geq 6$, $n \geq 4$, and $(\Delta x, \Delta y, \Delta s)$ is the solution of System* (3.3) *with $\mu = \mu_t$ as the target value. Then for*

$$
\theta = \frac{\alpha^* \sigma_{\ell q}\sqrt{\tau^{\log n/(2+\log n)} - 2}}{(\tau + 3)n^{1/2}},
$$

we have

$$
\Phi(x(\alpha), s(\alpha), (1-\theta)\mu_t) \leq \Phi(x, s, \mu_t), \quad \forall \alpha \leq \alpha^* \text{ where } \alpha^* = \alpha_1^* \text{ or } \alpha_2^*.
$$

PROOF. By Lemma 4.6 we know that

$$
\Phi(x(\alpha), s(\alpha), (1-\theta)\mu_t) \leq \frac{1}{1-\theta}\Phi(x(\alpha), s(\alpha), \mu_t) + \frac{2n\theta}{1-\theta}.
$$

Therefore the statement of the lemma holds whenever

$$
\frac{1}{1-\theta}\Phi(x(\alpha), s(\alpha), \mu_t) + \frac{2n\theta}{1-\theta} \leq \Phi(x, s, \mu_t),
$$

or

$$
\Phi(x(\alpha), s(\alpha), \mu_t) + 2n\theta \leq (1-\theta)\Phi(x, s, \mu_t).
$$

Now using Theorem 4.4 or 4.5,

$$
\frac{(\tau + 3)\theta n}{2} \leq \frac{\alpha^* \sigma_{\ell q}^2}{4},
$$

where $\alpha^* = \alpha_1^*$, or α_2^*, and $q = 1$ or $1 + \log n$. Using Lemma 4.2, this inequality definitely holds for

$$\theta = \frac{\alpha^* \sigma_{\ell q} \sqrt{\tau^{\log n / (2 + \log n)} - 2}}{2(\tau + 3) n^{1/2}}. \qquad \Box$$

COROLLARY 4.8. *After each iteration with step size $\alpha^* = \alpha_1^*$ or α_2^*, Equation (3.2) has two positive roots. Moreover,*

$$\mu_t(\alpha) \leq (1 - \theta)\mu_t, \quad \forall \alpha \leq \alpha^*. \qquad (4.3)$$

PROOF. Since the value of the proximity function $\Phi(x(\alpha^*), s(\alpha^*), \mu)$ goes to infinity when μ approaches zero, then by Lemma 4.7, Equation (3.2) has a root smaller than $(1 - \theta)\mu_t$. Furthermore, let $\mu^*(\alpha^*)$ be the global minimizer of $\Phi(x(\alpha^*), s(\alpha^*), \mu)$. Then by Theorem 4.4 or 4.5,

$$\Phi(x(\alpha^*), s(\alpha^*), \mu^*(\alpha^*)) < \frac{(\tau - 1)n}{2}.$$

Therefore, since the left-hand side of (3.2) as a function of μ is strictly convex, it has another root. We denote the smaller one by $\mu_t(\alpha^*)$, which is less than $(1 - \theta)\mu_t$. $\qquad \Box$

In the following lemma we give lower bounds for θ for different choices of q.

LEMMA 4.9. *We have*

$$q = 1 \qquad \Rightarrow \theta \geq \frac{C_1 \sqrt{\tau^{\log n / (2 + \log n)} - 2}}{2 e^{1/2} (\tau + 3) \left(\tau + \frac{2}{\log n} \right)^{1 / \log n} (1 + C_1) n^{1/2} \log n},$$

$$q = 1 + \log n \quad \Rightarrow \theta \geq \frac{\sqrt{\tau^{\log n / (2 + \log n)} - 2}}{12 e (\tau + 3) \left(\tau + \frac{2}{\log n} \right)^{1 / \log n} n^{1/2} \log n}.$$

PROOF. This follows from the definition of θ and choice of α^*. $\qquad \Box$

In the following theorem we give an upper bound for the number of iterations required by the hybrid adaptive algorithm in the worst case to find an ϵ-approximate solution.

THEOREM 4.10. *The hybrid adaptive algorithm stops after at most $\mathcal{O}(\sqrt{n} \log n \log(x^0)^T s^0 / \epsilon)$ iterations with a solution (x, y, s) for which $x^T s \leq \epsilon$.*

PROOF. In light of inequality (4.3) and Lemma 4.9, after at most $\mathcal{O}(\sqrt{n} \log n \log(x^0)^T s^0 / \epsilon)$ iterations one has $\mu_t \leq \epsilon / ((\tau + 2 / \log n)n)$. Then from Lemma 3.6, we have $x^T s \leq \epsilon$. $\qquad \Box$

5. Concluding remarks

In this paper, using the framework of self-regularity, we have proposed an adaptive hybrid algorithm which employs both the classical and SR search directions depending on the position of the current iterate with respect to the central path. The hybrid structure has enabled us to achieve the best iteration bound that can be achieved by pure SR search directions and improve on the classical algorithms.

Appendix A. Proofs

In this section we present the proofs of some lemmas and a theorem presented in the paper. The following two technical lemmas are used in the proof of Theorem 4.4.

LEMMA A.1. *Suppose that $\alpha \in [0, 1]$. Then $(1 + t)^{\alpha} \leq 1 + \alpha t$, for all $t \geq -1$.*

PROOF. See [6, Lemma 1.3.1]. □

LEMMA A.2. *Suppose $h(\cdot)$ is a twice differentiable convex function with $h(0) = 0$ and $h'(0) < 0$. Suppose that $h(\cdot)$ attains its global minimum at its stationary point $t^* > 0$ and $h''(t)$ is increasing with respect to t. Then for any $t \in [0, t^*]$, $h(\cdot) \leq h'(0)t/2$.*

PROOF. See [6, Lemma 1.3.3]. □

PROOF OF LEMMA 3.5. If $\mu_g = \tau \mu_h$, then from Equation (3.2) one has $\mu_t = \mu_h$ which is strictly less than μ^* as $\mu_g = \tau \mu_h > \mu_h$. Moreover, since Φ is a strictly decreasing function of μ for $\mu < \mu^*$, we have that $\Phi(x, s, \mu^*) < (\tau - 1)n/2$ and (3.2) has exactly one root less than μ^*. Furthermore, since Φ as a function of μ going to infinity for $\mu > \mu^*$, then (3.2) has exactly one root that is greater than μ^*. Now let us assume that $\mu_g = \tau_1 \mu_h$, where $1 \leq \tau_1 < \tau$. Then

$$\Phi(x, s, \mu_h) = \frac{(\tau_1 - 1)n}{2} < \frac{(\tau - 1)n}{2}.$$

We also know that the value of the proximity measure strictly increases to infinity when μ approaches zero. All these together imply that (3.2) has exactly one solution which is strictly less than μ_h. Similar to the previous case, another root which is greater than μ^* must also exist. □

PROOF OF LEMMA 3.6. The right-hand-side inequality easily follows from (3.2). For the left-hand-side inequality, since μ^* is the global minimizer of the proximity measure $\Phi(x, s, \mu)$ as a function of μ and the proximity measure approaches infinity when μ approaches zero, $\mu_t \leq \mu^*$. Now let us consider the case where $\mu_g \leq \tau \mu_h$. Then $\Phi(x, s, \mu_g/\tau) \leq (\tau - 1)n/2$. This itself implies that $\mu_t \leq \mu_g/\tau$. For the case where $\mu_g > \tau \mu_h$, using the definition of μ^* leads to $\mu^* \leq \tau^{-\log n/(2+\log n)} \mu_g$. Now, since $\mu_t \leq \mu^*$, then

$$\mu_t \leq \max\left(\frac{\mu_g}{\tau}, \tau^{-\log n/(2+\log n)} \mu_g\right) = \tau^{-\log n/(2+\log n)} \mu_g. \qquad \square$$

PROOF OF LEMMA 4.3. If $v_{\min} \geq e^{-1/2}(\tau + 2/\log n)^{-1/\log n}$, then by Lemma 4.2,

$$v_{\min}^{1+\log n}\sigma_{\ell 1} \geq e^{-1/2}n^{-1/2}\big(\tau^{\log n/(2+\log n)} - 2\big)^{1/2}\Big(\tau + \frac{2}{\log n}\Big)^{-1-1/\log n} n^{1/2}$$

$$= \frac{(\tau^{\log n/(2+\log n)} - 2)^{1/2}}{e^{1/2}\big(\tau + \frac{2}{\log n}\big)^{1+1/\log n}} := C_1,$$

completing the proof of (4.1). Finally, (4.2) follows from [6, Proposition 3.1.5]. □

PROOF OF THEOREM 4.4. Let us define

$$g(\alpha) := \Phi\,(x(\alpha), s(\alpha), \mu_t) - \Phi(x, s, \mu_t)$$

$$= \frac{\|v(\alpha)\|^2 - n}{2} + \frac{\big\|v(\alpha)^{-\log n}\big\|^2 - n}{\log n} - \frac{\|v\|^2 - n}{2} - \frac{\big\|v^{-\log n}\big\|^2 - n}{\log n},$$

where

$$v(\alpha) = \sqrt{\frac{x(\alpha)s(\alpha)}{\mu_t}} = (v + \alpha d_x)^{1/2}(v + \alpha d_s)^{1/2}.$$

After simplifying the expression, and using the arithmetic-geometric mean inequality, we obtain

$$g(\alpha) \leq \frac{1}{2}v^T(d_x + d_s)\alpha + \frac{1}{2\log n}\sum_{i=1}^{n}\big[(v_i + \alpha(d_x)_i)^{-\log n} + (v_i + \alpha(d_s)_i)^{-\log n}\big]$$

$$- \frac{\big\|v^{-\log n}\big\|^2}{\log n} := g_1(\alpha).$$

By the definition of $\sigma_{\ell q}$, the derivative of $g_1(\alpha)$ at zero is

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

Moreover, for the second derivative of $g_1(\alpha)$, using Lemma 4.1 gives

$$g_1''(\alpha) \leq \frac{(1 + \log n)\sigma_{\ell 1}^2}{2}(v_{\min} - \alpha\sigma_{\ell 1})^{-2-\log n}.$$

Now using the fact that

$$g_1(0) = 0 \quad \text{and} \quad g_1(\alpha) = g_1(0) + g_1'(0)\alpha + \int_0^\alpha \int_0^\zeta g_1''(\eta)\,d\eta\,d\zeta,$$

one further has

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

It is easy to see, via use of simple calculus, that $g_2(\alpha)$ is convex and twice continuously differentiable for all $\alpha \in [0, \alpha_{\max})$. Let $\bar{\alpha}_1$ denote the global minimizer of $g_2(\alpha)$. Then it is the solution of the equation

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

From (A.1),

$$\bar{\alpha}_1 = \frac{v_{\min}}{\sigma_{\ell 1}}\Big(1 - (\sigma_{\ell 1}v_{\min}^{1+\log n} + 1)^{-1/(1+\log n)}\Big), \quad \text{or}$$

$$\bar{\alpha}_1 = \frac{v_{\min}}{\sigma_{\ell 1}}\left(1 - \left(1 - \frac{\sigma_{\ell 1}v_{\min}^{1+\log n}}{1 + \sigma_{\ell 1}v_{\min}^{1+\log n}}\right)^{1/(1+\log n)}\right).$$

Now using Lemma A.1 for the inner parentheses leads to

$$\bar{\alpha}_1 \geq \frac{v_{\min}\sigma_{\ell 1}v_{\min}^{1+\log n}}{\sigma_{\ell 1}(1 + \log n)(1 + \sigma_{\ell 1}v_{\min}^{1+\log n})}$$

$$\geq \frac{C_1}{e^{1/2}\Big(\tau + \frac{2}{\log n}\Big)^{1/\log n}(1 + C_1)\sigma_{\ell 1}\log n} =: \alpha_1^*, \tag{A.2}$$

where the last inequality follows from part one of Lemma 4.3. Now, since $g_2(\alpha)$ satisfies all conditions of Lemma A.2,

$$\Phi(x(\alpha_1^*), s(\alpha_1^*), \mu_t) \leq \Phi(x, s, \mu_t) - \frac{1}{4}\alpha_1^*\sigma_{\ell 1}^2.$$

Acknowledgement

References

[1] E. D. Andersen and K. D. Andersen, "The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm", in: *High performance optimization*, (eds H. Frenk, K. Roos, T. Terlaky and S. Zhang), (Kluwer Academic Publishers, Dordrecht, 2000) 197–232.

[2] CPLEX: ILOG optimization. http://www.ilog.com.

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

[4] N. Megiddo, "Pathways to the optimal set in linear programming", in: *Progress in mathematical programming: interior point and related methods*, (ed. N. Megiddo), (Springer, New York, 1989) 131–158. (Identical version in: *Proc. 6th Mathematical Programming Symp. of Japan*, Nagoya, Japan, 1986, 1–35.)

[5] S. Mehrotra, "On the implementation of a (primal-dual) interior point method", *SIAM J. Optim.* **2** (1992) 575–601.

[6]   J. Peng, C. Roos and T. Terlaky, *Self-regularity: a new paradigm for primal-dual interior-point methods* (Princeton University Press, Princeton, NJ, 2002).

[7]   C. Roos, T. Terlaky and J.-Ph. Vial, *Interior point algorithms for linear optimization*, 2nd edn (Springer, New York, 2006).

[8]   M. Salahi, "New adaptive interior point algorithms for linear optimization", Ph. D. Thesis, Department of Mathematics and Statistics, McMaster University, 2006.

[9]   G. Sonnevend, "An 'analytic center' for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming", in: *System modeling and optimization: proceedings of the 12th IFIP conference, Budapest, Hungary, September 1985*, Volume 84 of *Lecture Notes in Control and Information Sciences* (eds A. Prékopa, J. Szelezsán and B. Strazicky), (Springer, Berlin, 1986) 866–876.

[10] S. J. Wright, *Primal-dual interior-point methods* (SIAM, Philadelphia, PA, 1997).

[11] Y. Ye, *Interior point algorithms, theory and analysis* (Wiley, Chichester, 1997).

[12] Y. Zhang, "Solving large-scale linear programs by interior point methods under the MATLAB environment", *Optim. Methods Softw.* **10** (1999) 1–31.