

AN INTERIOR POINT METHOD FOR LINEAR PROGRAMMING BASED ON A CLASS OF KERNEL FUNCTIONS

K. AMINI AND M.R. PEYGHAMI

Interior point methods are not only the most effective methods for solving optimisation problems in practice but they also have polynomial time complexity. However, there is still a gap between the practical behavior of the interior point method algorithms and their theoretical complexity results. In this paper, by focusing on linear programming problems, we introduce a new family of kernel functions that have some simple and easy to check properties. We present a simplified analysis to obtain the complexity of generic interior point methods based on the proximity functions induced by these kernel functions. Finally, we prove that this family of kernel functions leads to improved iteration bounds of the large-update interior point methods.

1. INTRODUCTION

After the seminal paper of Karmarkar [3], linear optimisation, that is linear programming, revitalised as an active area of research. Currently, the interior point methods not only are the most effective methods in practice but also have polynomial time complexity [9, 11, 12]. In this paper, we deal with primal-dual interior point methods for solving the standard linear optimisation problem

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

where $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are fixed data and $x \in \mathbb{R}^n$ is the vector of unknowns. The dual problem of (P) is given by

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

where $s \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. We assume that both (P) and (D) satisfy the interior point condition, that is, there exists (x^0, y^0, s^0) such that

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

Received 20th September, 2004

The authors would like to thank the Research Council of Sharif University of Technology. They also would like to thank A. Ghaffari Hadigheh for his helpful comments.

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9727/05 \$A2.00+0.00.

It is well known that the interior point condition can be assumed without loss of generality. In fact, we assume that $x^0 = s^0 = e$, where e denotes the all-one vector of length n ([9]). To find a primal-dual optimal solution of (P) and (D) , it is sufficient to solve the following system of equations:

$$(1) \quad \begin{aligned} Ax &= b, & x &\geq 0, \\ A^T y + s &= c, & s &\geq 0, \\ x_i s_i &= 0, & i &= 1, \dots, n. \end{aligned}$$

The first and second equations denote primal and dual feasibility while the third equation is the so-called *complementarity condition* of problems (P) and (D) .

We use the following notational conventions. Throughout the paper, $\|\cdot\|$ denotes the 2-norm of a vector. For $x, s \in \mathbb{R}^n$, xs denotes the coordinate-wise (Hadamard) product of the vectors x and s , that is, $(xs)_i = x_i s_i$ for all $i \in \{1, \dots, n\}$. The nonnegative and positive orthants are denoted by \mathbb{R}_+^n and \mathbb{R}_{++}^n , respectively. We say $f(x) = \Theta(g(x))$ if there exist some positive constants c_1 and c_2 such that $c_1 g(x) \leq f(x) \leq c_2 g(x)$ holds for all $x > 0$. Further, $f(x) = \mathcal{O}(g(x))$ if there exists a positive constant c such that $f(x) \leq cg(x)$ holds for all $x > 0$.

The basic idea behind the primal-dual interior point methods is replacing the third equation in (1) by the parameterised equation $xs = \mu e$, with some $\mu > 0$. This substitution leads us to the following system:

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

Without loss of generality, we assume that A has full row rank and the interior point condition holds. Under these assumptions, system (2) has a unique solution for each value $\mu > 0$. The set of unique solutions $\{(x(\mu), y(\mu), s(\mu)) \mid \mu > 0\}$, is referred to as the *central path* of (P) and (D) . The central path for linear optimisation was first recognised independently by Sonnevend [10] and Megiddo [4]. As $\mu \rightarrow 0$, the limit of the central path exists and converges to an optimal solution of (P) and (D) .

Let us briefly explain how a primal-dual interior point method works. Let $(x(\mu), y(\mu), s(\mu))$ be known for some $\mu > 0$. One may assume that $x(\mu) = s(\mu) = e$, for $\mu = 1$ ([7]). First, we decrease μ to $\mu_+ := (1 - \theta)\mu$, for some $\theta \in (0, 1)$ and then, solve the system (2) to obtain an approximate solution by applying Newton’s method. The Newton direction for (2) is determined by the following system:

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

Since A has full row rank for any $x > 0$ and $s > 0$, system (3) defines a Newton search direction $(\Delta x, \Delta y, \Delta s)$, uniquely. The third equation in (3) is referred to as the *centring equation*. Taking a step along the Newton direction determined by (3), one constructs a triple (x_+, y_+, s_+) as

$$x_+ = x + \alpha \Delta x, \quad y_+ = y + \alpha \Delta y, \quad s_+ = s + \alpha \Delta s.$$

By repeating this procedure, we find an iterate “close” enough to $(x(\mu), y(\mu), s(\mu))$. In this way, an ε -solution of the problems (P) and (D) is found if μ is small enough, say $n\mu \leq \varepsilon$.

The choice of θ , the so-called *barrier update parameter*, plays an important role in the theory and practice of interior point methods. Usually, if θ is a constant independent of the problem dimension n , say $\theta = 1/2$, then we refer to the algorithm as *large-update* (or *long-step*) method. If θ depends on the problem dimension n , say $\theta = 1/\sqrt{n}$, then the algorithm is referred to as *small-update* (or *short-step*) method. Recall that small-update methods have the best iteration bound in theory ([9]). They require $\mathcal{O}(\sqrt{n} \log(n/\varepsilon))$ iterations to produce a ε -solution. On the other hand, large-update methods based on the Newton direction determined by (3) are much more efficient than small-update methods in practice ([1]), but have a worst-case iteration bound, that is, $\mathcal{O}(n \log(n/\varepsilon))$ ([9, 11, 12]). This phenomenon is called “*The gap between theory and practice*”. Recently, to resolve this discrepancy, Peng, Roos and Terlaky [7] introduced a family of Self-Regular proximity functions and for a special member of the self-regular family, they established an $\mathcal{O}(\sqrt{n} \log n \log(n/\varepsilon))$ iteration bound for the large-update interior point methods.

In this paper, we present another family of proximity functions. They enjoy some mild and easy to check properties, that is, exponential-convexity, superconvexity and monotonicity of the second derivatives. We provide some powerful tools for the complexity analysis of the generic primal-dual interior point methods. We also prove that the approach constructed by these kernel functions can achieve the worst case iteration bound $\mathcal{O}\left(q\sqrt{n}(\log n)^{1+(1/q)} \log(n/\varepsilon)\right)$ for the large-update methods, where $q \geq 1$ is the so-called *barrier degree* of these kernel functions.

The paper is organised as follows. In Section 2, we review some necessary concepts and introduce some simplifying notation. Section 3 is devoted to define the new family of kernel functions and their properties. In Section 4, the growth behaviour of this family is investigated and a default value for the step size is derived. We estimate the decrease of the proximity function during a damped step in Section 5. From these results, the iteration bound is obtained in this section as well. The last section contains some concluding remarks.

2. PRELIMINARY

In this section we describe the idea underlying the approach of this paper. First, we associate to any triple (x, y, s) the vectors

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

whose i^{th} components are $\sqrt{(x_i s_i)/\mu}$ and $\sqrt{\mu/(x_i s_i)}$, respectively. Introduce the *scaled search directions* d_x and d_s ,

$$(5) \quad d_x := \frac{v \Delta x}{x} \quad \text{and} \quad d_s := \frac{v \Delta s}{s}.$$

Then the system (3), with μ_+ replaced by μ , can be rewritten as

$$(6) \quad \begin{aligned} \bar{A} d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned}$$

where

$$(7) \quad \bar{A} := \frac{1}{\mu} A V^{-1} X, \quad V := \text{diag}(v), \quad X := \text{diag}(x).$$

Note that d_x and d_s are orthogonal vectors, since $d_x \in \text{null}(\bar{A})$ and $d_s \in \text{range}(\bar{A}^T)$. Consequently, $d_x = d_s = 0$ if and only if $v^{-1} - v = 0$, or equivalently $v = e$. Thus, $d_x = d_s = 0$ holds if and only if $(x, s) = (x(\mu), s(\mu))$.

An important observation is that the right hand side of the third equation in (6) equals to minus of the gradient of the *scaled barrier (proximity) function*

$$\Psi_c(v) := \sum_{i=1}^n \left(\frac{v_i^2 - 1}{2} - \log v_i \right).$$

By replacing $\Psi_c(v)$ with any other strictly convex function $\Psi(v)$ with $v \in \mathbb{R}_{++}^n$, where $\Psi(v)$ is minimal at $v = e$ with $\Psi(e) = 0$, the general scaled centring equation is

$$(8) \quad d_x + d_s = -\nabla \Psi(v).$$

We reassert that in (8), $d_x = d_s = 0$ holds if and only if $\nabla \Psi(v) = 0$. Thus, $d_x = d_s = 0$ holds if and only if $(x, s) = (x(\mu), s(\mu))$, as it should. To simplify matters, we restrict our interest to the case where $\Psi(v)$ is separable with identical coordinate functions, that is,

$$(9) \quad \Psi(v) := \sum_{i=1}^n \psi(v_i),$$

where $\psi : D \rightarrow \mathbb{R}_+$ with $\mathbb{R}_{++} \subseteq D$, is a strictly convex function with $\psi(t)$ minimised at $t = 1$, and $\psi(1) = 0$. We refer to the univariate function $\psi(t)$ as the *kernel function* of the unscaled barrier (proximity) function Ψ introduced in (9). In general, it is supposed that the kernel function is twice differentiable. With this assumption and considering (5), the new search direction $(\Delta x, \Delta y, \Delta s)$ is defined by solving the following system for $d_x, \Delta y, d_s$.

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

The aforementioned discussion is summarised in the following *generic* primal-dual algorithm ([9]) for linear optimisation.

Algorithm 1. Generic Primal-Dual Algorithm for Linear Optimisation

Input:

- A proximity function $\Psi(v)$;
- A threshold parameter $\tau > 0$;
- An accuracy parameter $\varepsilon > 0$;
- A fixed barrier update parameter θ , $0 < \theta < 1$;

begin

$x := e$; $s := e$; $\mu := 1$;

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

begin

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

$v := \sqrt{(xs)/\mu}$;

while $\Psi(v) \geq \tau$ **do**

begin

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

$y_+ = y + \alpha \Delta y$;

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

end

end

end

In Algorithm 1, the inner “while loop” is called the *inner iteration* and the outer “while loop” is called the *outer iteration*. Each outer iteration consists of an update of parameter μ and a sequence of (one or more) inner iterations. The total number of inner iterations is referred to as *iteration complexity* of the algorithm. Usually, this number is described as a function of n and ε . The choice of parameters $\Psi(v)$, τ , θ and step size α plays an important role in controlling the complexity of the algorithm.

3. NEW FAMILY OF KERNEL FUNCTIONS

In this section, we introduce a new class of kernel functions that helps us to narrow “the gap between theory and practice”. Consider

$$(11) \quad \psi(t) = \frac{t^2 - 1}{2} - \frac{\int_1^t e^{x^{-q}} dx}{e},$$

where $q \geq 1$. The parameter q is called the *barrier degree*. Obviously, $\psi(t)$ is a kernel function, since $\psi(1) = \psi'(1) = 0$, and it can be determined by its second derivative as follows:

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi.$$

The proximity functions induced by kernel function (11) are

$$(12) \quad \Psi(v) = \frac{e^T v^2 - n}{2} - \frac{1}{e} \sum_{i=1}^n \int_1^{v_i} e^{x^{-q}} dx, \quad q \geq 1.$$

In the complexity analysis of Algorithm 1, we use the norm-based proximity measure $\delta(v)$ defined by

$$(13) \quad \delta(v) := \frac{1}{2} \|\nabla \Psi(v)\| = \frac{1}{2} \sqrt{\sum_{i=1}^n (\psi'(v_i))^2} = \frac{1}{2} \|d_x + d_s\|.$$

Note that $\Psi(v)$ is a strictly convex function with a minimum of 0 at $v = e$. Thus, we have

$$\Psi(e) = 0 \Leftrightarrow \delta(v) = 0 \Leftrightarrow v = e.$$

Let us verify some properties of the kernel function $\psi(t)$ introduced by (11). These properties lead to an upper bound for the growth behaviour of the proximity function and let us define a default value for step size. First, we need the first three derivatives of $\psi(t)$ with respect to t . They are

$$(14) \quad \begin{aligned} \psi'(t) &= t - e^{t^{-q}-1}, \\ \psi''(t) &= 1 + \frac{qe^{t^{-q}-1}}{t^{q+1}}, \\ \psi'''(t) &= -\left(\frac{q^2 e^{t^{-q}-1} + qe^{t^{-q}-1}}{t^{q+2}} + \frac{q^2 e^{t^{-q}-1}}{t^{2q+2}}\right). \end{aligned}$$

It can be directly concluded from (11) and (14) that $\psi''(t) \geq 1$, for all $t > 0$ and

$$\lim_{t \rightarrow 0} \psi(t) = \lim_{t \rightarrow +\infty} \psi(t) = +\infty.$$

According to these facts, one can easily prove the following lemma.

LEMMA 3.1. $\psi''(t)$ is monotonically decreasing for all $t > 0$ and $q \geq 1$.

Remember that if the kernel function satisfies the so-called *exponential convexity*, that is,

$$(15) \quad \psi(\sqrt{t_1 t_2}) \leq \frac{1}{2} [\psi(t_1) + \psi(t_2)], \quad \forall t_1, t_2 > 0,$$

then the complexity analysis of the algorithm is greatly simplified ([2, 6]). The following lemma establishes the exponential convexity property of the kernel functions (11) that plays an important role to identify the default value for the step size.

LEMMA 3.2. *Let $\psi(t)$ be defined as (11). Then, $\psi(t)$ has the exponential convexity property.*

PROOF: According to [8, Lemma 2.1.2], we know that (15) holds if and only if $\psi'(t) + t\psi''(t) \geq 0$, for $t > 0$. Since $\psi'(t) \geq 0$, for $t \geq 1$, then $\psi'(t) + t\psi''(t) \geq 0$ holds. On the other hand, from (14), the following relation holds for $0 < t < 1$,

$$\psi'(t) + t\psi''(t) = 2t + e^{t^{-q}-1}(-1 + qt^{-q}) \geq 0,$$

which completes the proof. □

We also need some properties of the kernel function (11) stated in the following lemma. One can find its proof in [2].

LEMMA 3.3. *Let $\psi(t)$ be defined by (11). Then*

- (i) $(t - 1)^2/2 \leq \psi(t) \leq \psi'(t)^2/2, \quad \forall t > 0.$
- (ii) $\Psi(v) \leq 2\delta(v)^2$, and
- (iii) $\|v\| \leq \sqrt{n} + \sqrt{2\Psi(v)} \leq \sqrt{n} + 2\delta(v).$

4. GROWTH BEHAVIOR OF THE PROXIMITY FUNCTION AND A DEFAULT VALUE FOR STEP SIZE

In this section, we verify the growth behaviour of the proximity function (12) and investigate its decreasing property during a feasible step size.

4.1. GROWTH BEHAVIOR. Before updating μ in the generic interior point method, we have $\Psi(v) \leq \tau$. In updating μ in an outer iteration, the vector v is divided by the factor $\sqrt{1 - \theta}$, which generally leads to an increase of the value of $\Psi(v)$. Thus, during the inner iterations, the value of $\Psi(v)$ decreases until it passes the threshold τ . We proceed by giving an upper bound to the increase of $\psi(t)$, when t increases to βt , with $\beta \geq 1$.

LEMMA 4.1 *Let $\beta \geq 1$, Then*

$$(16) \quad \psi(\beta t) \leq \psi(t) + \frac{1}{2}(\beta^2 - 1)t^2.$$

PROOF: From (11), we have

$$\psi(\beta t) - \psi(t) = \frac{(\beta t)^2 - t^2}{2} + \frac{\int_{\beta t}^t e^{x^{-q}} dx}{e} \leq \frac{(\beta t)^2 - t^2}{2},$$

where the last inequality follows from the fact $\int_{\beta t}^t e^{x^{-q}} dx \leq 0$, for $\beta \geq 1$. □

Lemma 4.1 enables us to estimate the effect of an update of the barrier parameter μ on the value of the proximity function (12). The following lemma presents an upper bound for the value of $\Psi(v)$ after the μ -update.

LEMMA 4.2. *Let $0 < \theta < 1$ and $v_+ := v/\sqrt{1-\theta}$. Then,*

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

PROOF: From Lemma 4.1 with $\beta = 1/\sqrt{1-\theta}$, and Lemma 3.3, we have

$$\begin{aligned} \Psi(v_+) &= \sum_{i=1}^n \psi(\beta v_i) \\ &\leq \sum_{i=1}^n \left(\psi(v_i) + \frac{1}{2}(\beta^2 - 1)v_i^2 \right) \\ &= \Psi(v) + \frac{\theta\|v\|^2}{2(1-\theta)} \\ &\leq \Psi(v) + \frac{\theta(\sqrt{n} + \sqrt{2\Psi(v)})^2}{2(1-\theta)}. \end{aligned}$$

□

4.2. ESTIMATING OF A DEFAULT VALUE FOR THE STEP SIZE. In order to estimate the decrease of the proximity function (12) during one step, we need a feasible step size. Taking a step along the search direction given by (5) and (10), with a step size defined by a line search rule, one constructs a triple (x_+, y_+, s_+) as

$$x_+ = x + \alpha\Delta x, \quad y_+ = y + \alpha\Delta y \quad \text{and} \quad s_+ = s + \alpha\Delta s.$$

By (5), we have

$$x_+ = \frac{x}{v}(v + \alpha d_x), \quad \text{and} \quad s_+ = \frac{s}{v}(v + \alpha d_s).$$

Thus,

$$v_+^2 = \frac{x_+ s_+}{\mu} = (v + \alpha d_x)(v + \alpha d_s),$$

and by Lemma 3.2

$$\Psi(v_+) = \Psi(\sqrt{(v + \alpha d_x)(v + \alpha d_s)}) \leq \frac{1}{2}[\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)].$$

Let $f(\alpha) = \Psi(v_+) - \Psi(v)$, and

$$(17) \quad f_1(\alpha) = \frac{1}{2}[\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)] - \Psi(v).$$

It is easy to verify that $f(\alpha) \leq f_1(\alpha)$ and $f_1(\alpha)$ is a convex function with $f(0) = f_1(0) = 0$. The first and second derivatives of $f_1(\alpha)$ with respect to α are

$$\begin{aligned}
 f'_1(\alpha) &= \frac{1}{2} \sum_{i=1}^n (\psi'(v_i + \alpha d_{x_i})d_{x_i} + \psi'(v_i + \alpha d_{s_i})d_{s_i}), \\
 f''_1(\alpha) &= \frac{1}{2} \sum_{i=1}^n (\psi''(v_i + \alpha d_{x_i})d_{x_i}^2 + \psi''(v_i + \alpha d_{s_i})d_{s_i}^2).
 \end{aligned}
 \tag{18}$$

Using (8), equation (18) can be rewritten as

$$f'_1(0) = \frac{1}{2} \nabla \Psi(v)^T (d_x + d_s) = -\frac{1}{2} \nabla \Psi(v)^T \nabla \Psi(v) = -2\delta(v)^2.
 \tag{19}$$

Assume that v_* is the minimum component of the vector v , that is, $v_* = \min_{1 \leq i \leq n} v_i$. Recall the definition of $\delta(v)$ in (13). For the sake of simplicity, we denote it by $\delta := \delta(v)$. According to the orthogonality of d_x and d_s , and using (8), we conclude that $\|(d_x, d_s)\| = 2\delta$. Therefore, for $1 \leq i \leq n$, we have

$$v_i + \alpha d_{x_i} \geq v_* - 2\alpha\delta, \quad \text{and} \quad v_i + \alpha d_{s_i} \geq v_* - 2\alpha\delta.$$

Using these facts, one can easily prove the following lemma.

LEMMA 4.3. *Let $f_1(\alpha)$ be defined as (17). Then, $f'_1(\alpha) \leq 2\delta^2\psi''(v_* - 2\alpha\delta)$.*

In order to find an appropriate value for the step size α , we use the convexity of $f_1(\alpha)$. One knows that for a convex function $h(t)$, the first derivative $h'(t)$ is nonpositive for all values of t less than or equal to the minimiser of $h(t)$. Thus, using this property of $f_1(\alpha)$ and Lemma 4.3, we establish the following lemma that allows us to define the default value for the step size.

LEMMA 4.4 *For any α that satisfies the inequality*

$$-\psi'(v_* - 2\alpha\delta) + \psi'(v_*) \leq 2\delta,
 \tag{20}$$

we have $f'_1(\alpha) \leq 0$.

PROOF: Using Lemma 4.3 and (19), we have

$$\begin{aligned}
 f'_1(\alpha) &= f'_1(0) + \int_0^\alpha f''_1(\xi)d\xi \leq -2\delta^2 + 2\delta^2 \int_0^\alpha \psi''(v_* - 2\xi\delta)d\xi \\
 &= -2\delta^2 - \delta \int_0^\alpha \psi''(v_* - 2\xi\delta)d(v_* - 2\xi\delta) = -2\delta^2 - \delta(\psi'(v_* - 2\alpha\delta) - \psi'(v_*)).
 \end{aligned}$$

Therefore, $f'_1(\alpha) \leq 0$ holds if α satisfies (20). □

We note that $\psi''(t) \geq 1$, and therefore the function $-(\psi'(t))/2$ is invertible on \mathbb{R}_{++} . Suppose that $\rho : [0, \infty) \rightarrow (0, 1]$ is the inverse function of the restriction of $-(\psi'(t))/2$ on the interval $(0, 1]$. We solve inequality (20) for the largest possible α , irrespective

to the value of v_* . Since $\psi''(t)$ is a decreasing function, the derivative of the expression at the left hand side in (20) with regard to v_* is negative. In order to have equality in (20) with fixed δ , we must find the minimum value of $\psi'(v_*)$ and the maximum value of $\psi'(v_* - 2\alpha\delta)$. Thus,

$$\delta = \frac{1}{2} \|\nabla \Psi(v)\| \geq \frac{1}{2} |\psi'(v_*)| \geq -\frac{1}{2} \psi'(v_*),$$

and equality holds if and only if v_* is the only coordinate in v that differs from one, and $v_* \leq 1$. Hence the worst situation for the step size occurs when v_* satisfies

$$(21) \quad -\frac{1}{2} \psi'(v_*) = \delta.$$

Observe that in this case, $\psi'(v_*)$ has the minimum value, and the inequality (20) reduces to

$$-\frac{1}{2} \psi'(v_* - 2\alpha\delta) \leq 2\delta.$$

The first derivative of the left hand side of this inequality with respect to α is positive. Thus, the largest possible value of α satisfying the inequality (21) must satisfy

$$(22) \quad -\frac{1}{2} \psi'(v_* - 2\alpha\delta) = 2\delta.$$

Due to the definition of ρ , equations (21) and (22) can be rewritten as

$$(23) \quad v_* = \rho(\delta), \quad \text{and} \quad v_* - 2\alpha\delta = \rho(2\delta).$$

Let $\bar{\alpha}$ be the largest possible solution of the inequality (20). Thus, (23) implies

$$(24) \quad \bar{\alpha} = \frac{1}{2\delta} [v_* - \rho(2\delta)] = \frac{1}{2\delta} [\rho(\delta) - \rho(2\delta)] = \frac{1}{2\delta} \int_{2\delta}^{\delta} \rho'(\sigma) d\sigma.$$

In the worst case, this step size is the largest possible solution of (20). From the definition of ρ , we have

$$-\psi'(\rho(\delta)) = 2\delta.$$

Taking the derivative of this equality with regard to δ , $\rho'(\delta) = -2 / (\psi''(\rho(\delta))) / 2 < 0$ is obtained. Then, using (24), we have

$$(25) \quad \frac{1}{\psi''(\rho(2\delta))} \leq \bar{\alpha} \leq \frac{1}{\psi''(\rho(\delta))}.$$

The following lemma summarises this discussion.

LEMMA 4.5 *Let $\rho : [0, \infty) \rightarrow (0, 1]$ denote the inverse function of the restriction of $-(\psi'(t))/2$ to the interval $(0, 1]$. Then, $\bar{\alpha}$, the largest possible solution of the inequality (20), satisfies (25).*

5. DECREASE OF THE PROXIMITY FUNCTION AND THE COMPLEXITY OF THE ALGORITHM

In this section, we first aim to estimate the decrease of the proximity function $\Psi(v)$ defined by (12) during the default step size given by

$$(26) \quad \tilde{\alpha} = \frac{1}{\psi''(\rho(2\delta))}.$$

Note that $\tilde{\alpha} \leq \bar{\alpha}$. Then, we present the complexity of Algorithm 1 which employs the proximity function defined by (12).

5.1. ESTIMATED VALUE OF $f(\tilde{\alpha})$. To obtain an estimated value for $f(\tilde{\alpha})$, we need the following technical lemma ([7]).

LEMMA 5.1. *Let $h(t)$ be a twice differentiable convex function with $h(0) = 0$, $h'(0) < 0$ that attains its (global) minimum at $t^* > 0$. If $h''(t)$ is an increasing function on $[0, t^*]$, then*

$$h(t) \leq \frac{1}{2}h'(t), \quad 0 \leq t \leq t^*.$$

The following lemma presents an upper bound for $f(\alpha)$.

LEMMA 5.2. *If the step size α satisfies $\alpha \leq \tilde{\alpha}$, then $f(\alpha) \leq -\alpha\delta^2$.*

PROOF: Let the function $h(t)$ satisfy

$$h(0) = f_1(0) = 0, \quad h'(0) = f_1'(0) = -2\delta^2 \quad \text{and} \quad h''(\alpha) = 2\delta^2\psi''(v_* - 2\alpha\delta).$$

From Lemma 4.3, we have $f_1''(\alpha) \leq h''(\alpha)$ and consequently, $f_1'(\alpha) \leq h'(\alpha)$ and $f_1(\alpha) \leq h(\alpha)$. Thus, $h'(\alpha)$ can be rewritten as:

$$h'(\alpha) = -2\delta^2 + 2\delta^2 \int_0^\alpha \psi''(v_* - 2\xi\delta) d\xi = -2\delta^2 - \delta(\psi'(v_* - 2\alpha\delta) - \psi'(v_*)).$$

Therefore, from Lemma 4.4, $h'(\alpha) \leq 0$ holds, for all $\alpha \leq \tilde{\alpha}$. On the other hand, since $\psi''(t)$ is a decreasing function, $h''(\alpha)$ is increasing in α . From Lemma 5.1,

$$f_1(\alpha) \leq h(\alpha) \leq \frac{1}{2}\alpha h'(0) = -\alpha\delta^2,$$

which completes the proof, since $f(\alpha) \leq f_1(\alpha)$. □

As a consequence of Lemma 5.2, we have

$$(27) \quad f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}.$$

Now, we apply the results obtained so far on proximity functions (12). First, we need to compute $\rho(2\delta)$. Assuming $s = \rho(2\delta)$, we have $-\psi'(s) = 4\delta$. Therefore, from (14) and (26), we may write

$$(28) \quad \tilde{\alpha} = \frac{1}{\psi''(s)} = \frac{1}{1 + qs^{-1-q}e^{s^{-q}-1}},$$

$$(29) \quad e^{s^{-q}-1} = 4\delta + s \leq 4\delta + 1,$$

and from (29), $s^{-q} \leq \log(4\delta + 1) + 1$ implies

$$s^{-1-q} \leq (1 + \log(4\delta + 1))^{(1+q)/q}.$$

Substituting this inequality and (29) in (28), we conclude that

$$\tilde{\alpha} \geq \frac{1}{1 + q(4\delta + 1)(1 + \log(4\delta + 1))^{(q+1)/q}} = \Theta\left(\frac{1}{4q\delta(\log \delta)^{(q+1)/q}}\right).$$

Thus, using Lemma 5.2, we obtain the following estimate for $f(\tilde{\alpha})$:

$$(30) \quad f(\tilde{\alpha}) = \Theta\left(-\frac{\delta}{4q(\log \delta)^{(q+1)/q}}\right).$$

5.2. ITERATION COMPLEXITY. According to Lemma 4.2, after the update of μ to $(1 - \theta)\mu$, we have

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

At the start of an outer iteration, inequality $\Psi(v) \leq \tau$ holds. We assume that $\tau = \mathcal{O}(n)$ and $\theta = \Theta(1)$. To count the inner iterations, we must return to the situation $\Psi(v) \leq \tau$ after the update of μ to $(1 - \theta)\mu$. Let Ψ_0 be the value of $\Psi(v)$ after the μ -update, with the subsequent values denoted by $\Psi_k, k = 1, 2, \dots, K$; where K denotes the total number of inner iterations in an outer iteration. Lemma 4.2 implies

$$(31) \quad \Psi_0 \leq \tau + \frac{\theta}{2(1 - \theta)}(2\tau + 2\sqrt{2n\tau} + n) = \mathcal{O}\left(n + \frac{n\theta}{1 - \theta}\right) = \mathcal{O}(n).$$

The decrease on each inner iteration is given by (30), that is,

$$\Psi_{k+1} \leq \Psi_k - \bar{\kappa} \frac{\delta}{4q(\log \delta)^{(q+1)/q}}, \quad k = 1, 2, \dots, K - 1,$$

where $\bar{\kappa}$ is some positive constant. Thus, the decrease depends monotonically on δ . Consequently, from Lemma 3.3 and assuming

$$1 \leq \tau \leq \Psi_{k+1} \leq \Psi_k, \quad k = 0, 1, \dots, K - 1,$$

we may express the decrease in terms of Ψ . In other words,

$$(32) \quad \Psi_{k+1} \leq \Psi_k - \kappa \Delta \Psi_k, \quad k = 0, 1, \dots, K - 1,$$

where

$$\Delta \Psi_k = \Theta\left(\frac{\Psi_k^{1/2}}{q(\log \Psi_k)^{1+1/q}}\right),$$

and κ is some positive constant. To complete the complexity analysis, We need the following technical lemma ([5]).

LEMMA 5.3. *If $\alpha \in [0, 1]$, then*

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

The following lemma is important for deriving the number of inner iterations of generic primal-dual interior point methods using the proximity function defined in (12).

LEMMA 5.4. *Considering (32), we have*

$$K = \mathcal{O}\left(\frac{\Psi_0}{(\kappa\Psi_0)/2}\right),$$

where K denotes the total number of inner iterations in an outer iteration.

PROOF: According to (32), we have

$$(33) \quad \Delta\Psi_k = f(\Psi_k)\Psi_k^{1/2},$$

where $f(\Psi) = \Theta(1/(q(\log \Psi)^{1+(1/q)}))$ is a monotonically decreasing function. Substituting (33) in (32) leads to

$$\Psi_{k+1} \leq \Psi_k - \kappa f(\Psi_k)\Psi_k^{1/2}, \quad k = 0, 1, \dots, K - 1,$$

which implies

$$\begin{aligned} 0 \leq \Psi_{k+1}^{1/2} &\leq (\Psi_k - \kappa f(\Psi_k)\Psi_k^{1/2})^{1/2} = \Psi_k^{1/2}(1 - \kappa f(\Psi_k)\Psi_k^{-(1/2)})^{1/2} \\ &\leq \Psi_k^{1/2}(1 - \frac{1}{2}\kappa f(\Psi_k)\Psi_k^{-(1/2)}) = \Psi_k^{1/2} - 1/2\kappa f(\Psi_k); \end{aligned}$$

where the last inequality holds by Lemma 5.3. Since $f(\Psi)$ is a monotonically decreasing function, we have

$$0 \leq \Psi_{k+1}^{1/2} \leq \Psi_k^{1/2} - \frac{1}{2}\kappa f(\Psi_k) \leq \Psi_k^{1/2} - \frac{1}{2}\kappa f(\Psi_0), \quad k = 0, 1, \dots, K - 1,$$

and consequently

$$\Psi_k^{1/2} \leq \Psi_0^{1/2} - \frac{1}{2}\kappa k f(\Psi_0), \quad k = 0, 1, \dots, K.$$

With $k = K$, this inequality leads to

$$0 \leq \Psi_0^{1/2} - \frac{1}{2}\kappa K f(\Psi_0).$$

Thus,

$$K \leq \frac{\Psi_0^{1/2}}{\kappa f(\Psi_0)/2} = \frac{\Psi_0}{(\kappa\Psi_0)/2},$$

where the equality follows from (33). That completes the proof. □

Relation (31) proves $\Psi_0 = \mathcal{O}(n)$. Thus, Lemma 5.4 implies that the number of inner iterations is bounded above by

$$K \leq \left\lfloor \frac{\Psi_0^{1/2}}{\kappa f(\Psi_0)/2} \right\rfloor = \left\lfloor \mathcal{O}(q\sqrt{n}(\log n)^{1+(1/q)}) \right\rfloor.$$

The iteration complexity of the algorithm is obtained by multiplying this number by the number of outer iterations, which is bounded above by $\mathcal{O}(\log n/\varepsilon)$ ([9]). Neglecting the integer brackets, which does not change the order of complexity, the iteration complexity is

$$(34) \quad \mathcal{O}(q\sqrt{n}(\log n)^{1+(1/q)} \log \frac{n}{\varepsilon}).$$

6. CONCLUSION

In this paper, we provided an analysis for the complexity of the generic primal-dual interior point methods based on the proximity functions induced by the new family of kernel functions. The complexity has been improved in comparison with the classical results given in [9] with the assumption $q \geq 1$. Under the conditions

1. $\psi''(t)$ is monotonically decreasing;
2. $\psi(t)$ is exponentially convex;
3. $\psi(t)$ is superconvex, that is, $\psi''(t) \geq 1$ for all $t > 0$,

the resulting large-update primal-dual interior point method can be easily analyzed. The results in Sections 2 and 3 were based only on these three properties. In particular, we presented an estimate for the decrease of the proximity function in terms of the induced kernel functions based on their first and second derivatives only. We also obtained the iteration complexity for these kernel functions.

REFERENCES

- [1] E.D. Andersen, J. Gondzio, Cs. Mészáros and X. Xu, 'Implementation of interior point methods for large scale linear programming', in *Interior Point Methods of Mathematical Programming*, (T. Terlaky, Editor) (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996), pp. 189–252.
- [2] Y.Q. Bai, M. El Ghami and C. Roos, 'A new efficient large-update primal-dual interior-point method based on a finite barrier', *SIAM J. Optim. (electronic)* **13** (2003), 766–782.
- [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*, (N. Megiddo, Editor) (Springer-Verlag, New York, 1989), pp. 131–158.
- [5] J. Peng, C. Roos and T. Terlaky, 'A new class of polynomial primal-dual methods for linear and semidefinite optimization', *European J. Oper. Res.* **143** (2002), 234–256.

- [6] J. Peng, C. Roos and T. Terlaky, 'A new and efficient large-update interior-point method for linear optimization', *Vychisl. Technol.* **6** (2001), 61–80.
- [7] J. Peng, C. Roos and T. Terlaky, 'Self-regular functions and new search directions for linear and semidefinite optimization', *Math. Program.* **93** (2002), 129–171.
- [8] J. Peng, C. Roos and T. Terlaky, *Self-regularity: A new paradigm for primal-dual interior-point algorithms*, Princeton Series in Applied Mathematics (Princeton University Press, Princeton, N.J., 2002).
- [9] C. Roos, T. Terlaky and J.-Ph.Vial, *Theory and algorithms for linear optimization: An interior-point approach* (John Wiley & Sons, Chichester, UK, 1997).
- [10] G. Sonnevend, 'An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming', (A. Prékopa, J. Szelezsán, and B. Strazicky, Editors), *Lecture Notes in Control and Information Sciences* **84** (Springer-Verlag, Berlin, 1986), pp. 866–876.
- [11] S.J. Wright, *Primal-dual interior-point methods* (SIAM, Philadelphia, USA, 1997).
- [12] Y. Ye, *Interior point algorithms, theory and analysis* (John Wiley & Sons, Chichester, UK, 1997).

Department of Mathematical Sciences
Sharif University of Technology
Tehran
Iran
e-mail: amini@mehr.sharif.edu
peyghami@mehr.sharif.edu