

GENERIC PRIMAL-DUAL INTERIOR POINT METHODS BASED ON A NEW KERNEL FUNCTION

M. EL GHAMI¹ AND C. ROOS²

Abstract. In this paper we present a generic primal-dual interior point methods (IPMs) for linear optimization in which the search direction depends on a univariate kernel function which is also used as proximity measure in the analysis of the algorithm. The proposed kernel function does not satisfy all the conditions proposed in [2]. We show that the corresponding large-update algorithm improves the iteration complexity with a factor $n^{\frac{1}{6}}$ when compared with the method based on the use of the classical logarithmic barrier function. For small-update interior point methods the iteration bound is $O(\sqrt{n} \log \frac{n}{\epsilon})$, which is currently the best-known bound for primal-dual IPMs.

Keywords. Linear optimization, primal-dual interior-point algorithm, large and small-update method.

Mathematics Subject Classification. 90C05, 90C31.

1. INTRODUCTION

The purpose of this work is to present a primal-dual interior-point method (IPM) based on a new barrier function for solving the standard linear optimization problem

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

Received February 9, 2007. Accepted November 28, 2007.

¹ Department of Informatics, University of Bergen, Thormøblensgate 55, 5008 Bergen, Norway; melghami@ii.uib.no

² Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, PO Box 5031, 2600 GA Delft, The Netherlands; C.Roos@ewi.tudelft.nl

where $A \in R^{m \times n}$ is a real $m \times n$ matrix with rank m , and $c, x \in R^n$, $b \in R^m$. The dual problem of (P) is given by

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

with $y \in R^m$ and $s \in R^n$.

Without loss of generality [13] we assume that (P) and (D) satisfy the interior-point condition (IPC), *i.e.*, there exist x^0, y^0 , and s^0 such that

$$Ax^0 = b, \quad x^0 > 0, \quad A^T y^0 + s^0 = c, \quad s^0 > 0. \quad (1)$$

It is well known that finding an optimal solution of (P) and (D) is equivalent to solving the non-linear system of equations

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

The first equation requires that x is feasible for (P), and the second equation that the pair (y, s) is feasible for (D), whereas the third equation is the so-called *complementarity condition* for (P) and (D); here and a long this paper xs denotes the coordinatewise product of the vectors x and s . Similarly, this notation is extended to quotients, for example $\frac{x}{s}$.

The basic idea underlying primal-dual IPMs is to replace the third equation in (2) by the nonlinear equation $xs = \mu \mathbf{1}$, with parameter $\mu > 0$ and with $\mathbf{1}$ denoting the all-one vector $(1, 1, \dots, 1)^T$. The system (2) now becomes:

$$\begin{aligned} Ax &= b, & x &\geq 0, \\ A^T y + s &= c, & s &\geq 0, \\ xs &= \mu \mathbf{1}. \end{aligned} \quad (3)$$

If the IPC holds the parameterized system (3) has a unique solution $(x(\mu), y(\mu), s(\mu))$ for each $\mu > 0$; $x(\mu)$ is called the μ -center of (P) and $(y(\mu), s(\mu))$ is the μ -center of (D). The set of μ -centers (with $\mu > 0$) defines a homotopy path, which is called *the central path* of (P) and (D) [7,8,14]. If $\mu \rightarrow 0$, then the limit of the central path exists ($\lim_{\mu \rightarrow 0} x(\mu)$, $\lim_{\mu \rightarrow 0} y(\mu)$, $\lim_{\mu \rightarrow 0} s(\mu)$, exist) and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D) [13].

Let $(x, s) \in \mathbf{R}^{2n}$ be a fixed primal-dual solution. We define the vector

$$v := \sqrt{\frac{xs}{\mu}}. \quad (4)$$

Note that the pair (x, s) coincides with the μ -center $(x(\mu), s(\mu))$ if and only if $v = \mathbf{1}$.

Let $\Psi : \mathbf{R}_{++}^n \rightarrow \mathbf{R}_+$ be a smooth, strictly convex function, which is minimal at $v = \mathbf{1}$, with $\Psi(\mathbf{1}) = 0$. Following [2–4,10,11] we define search directions from a given primal-dual feasible solution $\Delta x, \Delta y, \Delta s$ by

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= -\mu v \nabla \Psi(v). \end{aligned} \tag{5}$$

Because A has full row rank, the system (5) uniquely defines $(\Delta x, \Delta s, \Delta y)$ for any feasible $x > 0$ and $s > 0$. Note that $\Delta x = 0, \Delta s = 0, \Delta y = 0$ if and only if $v = \mathbf{1}$, because the right-hand sides in (5) vanish if and only if $\nabla \Psi(v) = 0$, and this occurs if and only if $v = \mathbf{1}$. Introducing the notations

$$\bar{A} := \frac{1}{\mu} AV^{-1}X = AS^{-1}V, \text{ where } V := \text{diag}(v), X := \text{diag}(x), S := \text{diag}(s), \tag{6}$$

and

$$d_x := \frac{v\Delta x}{x}, \quad d_s := \frac{v\Delta s}{s}, \tag{7}$$

system (5) can be reformulated as

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= -\nabla \Psi(v). \end{aligned} \tag{8}$$

Knowing d_x and d_s , the vectors Δx and Δs can be computed from (7). The algorithm considered in this paper is described in Figure 1.

The algorithm consists of *inner iterations* and *outer iterations*. Each outer iteration performs an update of the barrier parameter and a sequence of inner iterations. It is generally agreed that the total number of inner iterations required by the algorithm is an appropriate measure for its efficiency. This number will be referred to as the *iteration complexity* of the algorithm; it is usually described as a function of the dimension n and the accuracy parameter ϵ . A crucial question is, of course, how to choose the proximity function Ψ , the threshold parameter τ , the barrier update parameter θ , and the step size α , so as to minimize the iteration complexity of the algorithm. So far researchers have considered only separable proximity functions:

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

where ψ is called the kernel function of $\Psi(v)$. Table 1 gives some examples of kernel functions that have been analyzed in earlier papers, with the complexity results for the corresponding algorithms. For ψ_6 the bound is minimal if we choose $q = \frac{1}{2} \log n$. This gives the best bound known so far for large-update ($\theta = O(1), \tau = O(n)$) interior-point methods: $O(\sqrt{n} \log n \log \frac{n}{\epsilon})$.

Generic Primal-Dual Algorithm for LO

Input:
 a proximity function $\Psi(v)$;
 a threshold parameter $\tau > 0$;
 an accuracy parameter $\epsilon > 0$;
 a barrier update parameter θ , $0 < \theta < 1$;

begin
 $x := \mathbf{1}$; $s := \mathbf{1}$; $\mu := 1$; $v := \mathbf{1}$;
while $n\mu \geq \epsilon$ **do**
begin
 $\mu := (1 - \theta)\mu$;
while $\Psi(v) > \tau$ **do**
begin
 Solve $(\Delta x, \Delta s, \Delta y)$ from (5)
 $x := x + \alpha\Delta x$;
 $s := s + \alpha\Delta s$;
 $y := y + \alpha\Delta y$;
 $v := \sqrt{\frac{xs}{\mu}}$;
end
end
end

FIGURE 1. The generic primal-dual interior-point algorithm or LO.

Recently Bai *et al.* [2] introduced a new class of kernel functions which is defined by some simple conditions. However, an important question remained open, namely whether there exists a kernel function that does not satisfy the conditions that are needed in the analysis used in [2], and which gives rise to an efficient primal-dual IPM. This paper offers a positive answer to that question.

In this paper we investigate the kernel function

$$\psi(t) = \psi_\tau(t) = 8t^2 - 11t + 1 + \frac{2}{\sqrt{t}} - 4 \log t. \quad (9)$$

This kernel function is mentioned in [2], but it has not been analyzed there because it does not satisfy the conditions that are needed in the analysis used in [2]. To make this clear we proceed with the next lemma.

Lemma 1.1. *Let ψ be as defined in (9). Then,*

$$t\psi''(t) + \psi'(t) > 0, \quad \text{if } t > 0, \quad (10\text{-a})$$

$$\psi'''(t) < 0, \quad \text{if } t > 0, \quad (10\text{-b})$$

$$t\psi''(t) - \psi'(t) > 0, \quad \text{if } t > 0. \quad (10\text{-c})$$

TABLE 1. Examples of kernel functions and its iteration bound for large-update methods.

i	Kernel functions ψ_i	Iteration bound	References
1	$\frac{t^2-1}{2} - \log t$	$O(n \log \frac{n}{\epsilon})$	[1,5,6,15]
2	$\frac{t^2-1}{2} + \frac{(e-1)^2}{e} \frac{1}{e^t-1} - \frac{e-1}{e}$	$O(n^{\frac{3}{4}} \log \frac{n}{\epsilon})$	[4]
3	$\frac{1}{2} (t - \frac{1}{t})^2$	$O(n^{\frac{2}{3}} \log \frac{n}{\epsilon})$	[9]
4	$\frac{t^2-1}{2} + e^{\frac{1}{t}-1} - 1$	$O(\sqrt{n} \log^2 n \log \frac{n}{\epsilon})$	[2]
5	$\frac{t^2-1}{2} - \int_1^t e^{\frac{1}{\xi}-1} d\xi$	$O(\sqrt{n} \log^2 n \log \frac{n}{\epsilon})$	[2]
6	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1}, q > 1$	$O(qn^{\frac{q+1}{2q}} \log \frac{n}{\epsilon})$	[10,12]
7	$8t^2 - 11t + 1 + \frac{2}{\sqrt{t}} - 4 \log t$	$O(n^{\frac{5}{6}} \log \frac{n}{\epsilon})$	New

Proof. The first three derivatives of $\psi(t)$ with respect to t are given by

$$\psi'(t) = 16t - 11 - \frac{1}{t^{\frac{3}{2}}} - \frac{4}{t}, \tag{11}$$

$$\psi''(t) = 16 + \frac{3}{2t^{\frac{5}{2}}} + \frac{4}{t^2}, \tag{12}$$

$$\psi'''(t) = -\frac{15}{4t^{\frac{7}{2}}} - \frac{8}{t^3}. \tag{13}$$

Using (11) and (12) we write

$$\psi'(t) + t\psi''(t) = 32t - 11 + \frac{1}{2t^{\frac{3}{2}}},$$

and one may easily check that this is positive for all $t > 0$. Thus (10-a) follows. Inequality (10-b) immediately follows from (13). By (11) and (12),

$$t\psi''(t) - \psi'(t) = \frac{2}{t^{\frac{3}{2}}} + \frac{8}{t} + 11 > 0, \quad t > 0,$$

which gives (10-c). □

In [2] the authors imposed four conditions on the kernel functions, namely (10-a)–(10-c) and one additional condition, namely:

$$2\psi''(t)^2 - \psi'(t)\psi'''(t) > 0, \quad t > 0. \quad (14)$$

Taking $t = \frac{1}{6}$, the left hand side expression is less than -3734 , showing that (9) does not satisfy (14). The aim of this paper is to show that despite the fact that $\psi(t)$ as given by (9) does not satisfy the conditions used in [2], a slightly different analysis makes clear that this kernel function is good enough to be the kernel function of a polynomial-time interior-point algorithm.

It is worth mentioning that the properties in Lemma 1.1 admit a nice geometric interpretation. Due to Lemma 1.6 in [11], (10-a) implies that ψ is e -convex, *i.e.*, $\psi(e^\xi)$, $\xi \geq 0$, is convex, which is equivalent to $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2))$ for all $t_1, t_2 > 0$. Also, (10-b) means that $\psi''(t)$ is decreasing and by Lemma 2.2 in [2], (10-c) implies that $\psi(\sqrt{\xi})$, $\xi \geq 0$, is convex.

The paper is organized as follows. In Section 2 we start by deriving some properties of the kernel function $\psi(t)$, as well as the corresponding properties of the barrier function $\Psi(v)$. The estimate of the step size and the decrease behavior of the barrier function are discussed in Section 3. The inner iteration bound and the total iteration bound of the algorithm are derived in Section 4. Finally, some concluding remarks follow in Section 5.

We use the following notational conventions. Throughout the paper, $\|\cdot\|$ denotes the 2-norm of a vector. The nonnegative and the positive orthants are denoted as \mathbf{R}_+^n and \mathbf{R}_{++}^n , respectively. Finally, if $z \in \mathbf{R}_+^n$ and $f: \mathbf{R}_+ \rightarrow \mathbf{R}_+$, then $f(z)$ denotes the vector in \mathbf{R}_+^n whose i th component is $f(z_i)$, with $1 \leq i \leq n$.

2. PROPERTIES OF NEW KERNEL FUNCTION

In this section, we focus on some properties of $\psi(t)$ that will be used in the analysis of the algorithm.

Since $\psi(1) = \psi'(1) = 0$ and $\psi''(t) \geq 0$, $\psi(t)$ is determined by $\psi''(t)$:

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi. \quad (15)$$

Lemma 2.1 (Lemma 2.4 in [2]). *If $\psi(t)$ satisfies (10-b) and (10-c), then $\psi(t)$ satisfies*

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, \quad \text{if } t > 1, \quad \beta > 1. \quad (16)$$

Lemma 2.2. *One has*

$$\psi(t) < \frac{1}{2}\psi''(1)(t-1)^2, \quad t > 1.$$

Proof. By using Taylor’s theorem and $\psi(1) = \psi'(1) = 0$, we obtain

$$\psi(t) = \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{6}\psi'''(\xi)(\xi-1)^3,$$

where $1 < \xi < t$ if $t > 1$. Since $\psi'''(\xi) < 0$, the lemma follows. □

Following [2], we introduce a norm-based proximity measure $\delta(v)$, according to

$$\delta(v) := \frac{1}{2}\|\nabla\Psi(v)\| = \frac{1}{2}\sqrt{\sum_{i=1}^n(\psi'(v_i))^2}, \quad v \in \mathbf{R}_{++}^n. \tag{17}$$

2.1. RELATION BETWEEN $\Psi(v)$ AND $\delta(v)$

For the analysis of the algorithm in Section 3 we need to establish the relation between $\Psi(v)$ and $\delta(v)$. The inverse function of $\psi(t)$, for $t \geq 1$, plays an important role in this relation.

The next theorem, which is one of main results in [2], gives a lower bound on $\delta(v)$ in term of $\Psi(v)$. This is due to the fact that $\psi(t)$ satisfies (10-b).

Theorem 2.3 (Theorem 4.9 in [2]). *Let $\varrho : [0, \infty) \rightarrow [1, \infty)$ be the inverse function of ψ on $[0, \infty)$. One has*

$$\delta(v) \geq \frac{1}{2}\psi'(\varrho(\Psi(v))).$$

Note that at the start of each inner iteration we have $\Psi(v) \geq \tau$. By Theorem 2.3 this implies that $\delta(v) \geq \frac{1}{2}\psi'(\varrho(\tau))$. We always assume that $\tau \geq 1$, to ensure that $\delta(v) \geq \frac{2}{5}$.

Theorem 2.4. *If $\Psi(v) \geq 1$, then*

$$\delta(v) \geq \frac{2}{5}\Psi(v)^{\frac{1}{2}}. \tag{18}$$

Proof. The inverse function of $\psi(t)$ for $t \in [1, \infty)$ is obtained by solving t from the equation

$$\psi(t) = 8t^2 - 11t + 1 + \frac{2}{\sqrt{t}} - 4 \log t = s, \quad t \geq 1.$$

Since it is hard to solve this equation explicitly, we derive a lower bound for t , as this suffices for our goal. One has

$$8t^2 = s - 1 - \frac{2}{\sqrt{t}} + 11t + 4 \log t \geq s - 1 - 2 + 11 \geq s + 8,$$

whence

$$\varrho(s) = t \geq \left(1 + \frac{s}{8}\right)^{\frac{1}{2}}.$$

We have

$$\varrho(\Psi(v)) \geq \left(1 + \frac{\Psi(v)}{8}\right)^{\frac{1}{2}}.$$

Now using that $\delta(v) \geq \frac{1}{2}\psi'(\varrho(\Psi(v)))$, and assuming $\Psi(v) \geq \tau \geq 1$, we obtain

$$\begin{aligned} \delta &\geq \frac{1}{2} \left(16 \left(1 + \frac{\Psi(v)}{8} \right)^{\frac{1}{2}} - 11 - \frac{1}{\left(1 + \frac{\Psi(v)}{8} \right)^{\frac{3}{2}}} - \frac{4}{\left(1 + \frac{\Psi(v)}{8} \right)} \right) \\ &\geq 8 \left(\left(1 + \frac{\Psi(v)}{8} \right)^{\frac{1}{2}} - 1 \right) = \frac{\Psi(v)}{\left(1 + \frac{\Psi(v)}{8} \right)^{\frac{1}{2}} + 1} \geq \frac{2\sqrt{2}\Psi(v)^{\frac{1}{2}}}{3 + 2\sqrt{2}} \geq \frac{2}{5}\Psi(v)^{\frac{1}{2}}. \end{aligned}$$

This proves the theorem. □

Note that if $\Psi(v) \geq 1$, substitution in (18) gives

$$\delta(v) \geq \frac{2}{5}. \tag{19}$$

2.2. GROWTH BEHAVIOR OF THE BARRIER FUNCTION

Note that at the start of each outer iteration of the algorithm, just before the update of μ with the factor $1 - \theta$, we have $\Psi(v) \leq \tau$. Due to the update of μ the vector v is divided by the factor $\sqrt{1 - \theta}$, with $0 < \theta < 1$, which in general leads to an increase in the value of $\Psi(v)$. Then, during the subsequent inner iterations, $\Psi(v)$ decreases until it passes the threshold τ again. Hence, during the course of the algorithm the largest values of $\Psi(v)$ occur just after the updates of μ . That is why in this section we derive an estimate for the effect of a μ -update on the value of $\Psi(v)$. We start with an important theorem. This is due to the fact that $\psi(t)$ satisfies (16).

Theorem 2.5 (cf. [2]). *With ϱ as defined in Theorem 2.3, we have for any positive vector v and any $\beta > 1$,*

$$\Psi(\beta v) \leq n\psi \left(\beta\varrho \left(\frac{\Psi(v)}{n} \right) \right).$$

Corollary 2.6. *One has*

$$\Psi(\beta v) \leq \frac{43n}{4} \left(\beta\varrho \left(\frac{\Psi(v)}{n} \right) - 1 \right)^2. \tag{20}$$

Proof. Since $\beta > 1$ and $\varrho \left(\frac{\Psi(v)}{n} \right) \geq 1$, the corollary follows from Lemma 2.2, Theorem 2.5 and $\psi''(1) = \frac{43}{2}$. □

Corollary 2.7. *Let $0 \leq \theta \leq 1$ and $v_+ = \frac{v}{\sqrt{1-\theta}}$. If $\Psi(v) \leq \tau$, then*

$$\Psi(v_+) \leq \frac{43n}{4} \left(\frac{\varrho \left(\frac{\tau}{n} \right)}{\sqrt{1-\theta}} - 1 \right)^2. \tag{21}$$

Proof. Set $\beta = \frac{1}{\sqrt{1-\theta}}$ in relation (20) and use the fact that $\Psi(v) \leq \tau$. \square

Suppose that the barrier update parameter θ and threshold value τ are given. According to the algorithm, at the start of each outer iteration we have $\Psi(v) \leq \tau$. By Corollary 2.7, after each μ -update the growth of $\Psi(v)$ is limited by (21). Therefore, we define

$$L(n, \theta, \tau) := \frac{43n}{4} \left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}} - 1 \right)^2, \quad (22)$$

$L(n, \theta, \tau)$ is an upper bound of $\Psi(v_+)$, the value of $\Psi(v)$ after the μ -update.

3. ANALYSIS OF THE ALGORITHM

In this section, we determine a default step size which not only keeps the iterations feasible but also gives rise to a sufficiently large decrease of the barrier function $\Psi(v)$ in each inner iteration. Apart from the necessary adaptations to the present context and some simplifications, the analysis below follows the same line of arguments that was used first in [12], and later in [2,3].

3.1. DECREASE OF THE PROXIMITY DURING A (DAMPED) NEWTON STEP

After a damped step, with step size α , using (4) we have

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

Thus we obtain

$$v_+^2 = \frac{x_+ s_+}{\mu} = (v + \alpha d_x) (v + \alpha d_s). \quad (23)$$

Since ψ satisfies (10-a), it is e -convex as introduced in [11]. This implies

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

Thus we have $f(\alpha) := \Psi(v_+) - \Psi(v) \leq f_1(\alpha)$, where

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

is a convex function of α , since $\Psi(v)$ is convex.

Obviously, $f(0) = f_1(0) = 0$. The derivative of f is

$$f_1'(\alpha) = \frac{1}{2} \sum_{i=1}^n (\psi'(v_i + \alpha d_{x_i}) d_{x_i} + \psi'(v_i + \alpha d_{s_i}) d_{s_i}).$$

This gives, using last equation in (8) and (17),

$$f_1'(0) = \frac{1}{2} \nabla \Psi(v)^T (d_x + d_s) = -\frac{1}{2} \nabla \Psi(v)^T \nabla \Psi(v) = -2\delta(v)^2. \quad (24)$$

Differentiating once more, we obtain

$$f_1''(\alpha) = \frac{1}{2} \sum_{i=1}^n (\psi''(v_i + \alpha d_{x_i}) d_{x_i}^2 + \psi''(v_i + \alpha d_{s_i}) d_{s_i}^2). \quad (25)$$

Below we use the following notation:

$$v_1 := \min(v).$$

Lemma 3.1. *One has*

$$f_1''(\alpha) \leq 2\delta(v)^2 \psi''(v_1 - 2\alpha\delta(v)).$$

Proof. The last equation in (8) and (17) imply that $\|(d_x, d_s)\| = 2\delta(v)$. Since d_x and d_s are orthogonal, we have $\|d_x\| \leq 2\delta(v)$ and $\|d_s\| \leq 2\delta(v)$. Therefore,

$$v_i + \alpha d_{x_i} \geq v_1 - 2\alpha\delta(v), \quad v_i + \alpha d_{s_i} \geq v_1 - 2\alpha\delta(v), \quad 1 \leq i \leq n.$$

Recall from (10-b) that ψ'' is monotonically decreasing, so using the above inequalities and (25) we obtain

$$f_1''(\alpha) \leq \frac{1}{2} \psi''(v_1 - 2\alpha\delta(v)) \sum_{i=1}^n (d_{x_i}^2 + d_{s_i}^2) = 2\delta(v)^2 \psi''(v_1 - 2\alpha\delta(v)).$$

This proves the lemma. \square

Since $f_1(\alpha)$ is convex, we will have $f_1'(\alpha) \leq 0$ for all α less than or equal to the value where $f_1(\alpha)$ is minimal, and *vice versa*. In this respect the next result is important.

Lemma 3.2. *One has $f_1'(\alpha) \leq 0$ if α satisfies the inequality*

$$-\psi'(v_1 - 2\alpha\delta(v)) + \psi'(v_1) \leq 2\delta(v). \quad (26)$$

Proof. We may write, using Lemma 3.1, and also (24),

$$\begin{aligned} f_1'(\alpha) &= f_1'(0) + \int_0^\alpha f_1''(\xi) d\xi \leq -2\delta(v)^2 + 2\delta(v)^2 \int_0^\alpha \psi''(v_1 - 2\xi\delta(v)) d\xi \\ &= -2\delta(v)^2 - \delta(v) (\psi'(v_1 - 2\alpha\delta(v)) - \psi'(v_1)) \leq 0, \end{aligned}$$

which proves the lemma. \square

Lemma 3.3. *Let $\rho : [0, \infty) \rightarrow (0, 1]$ denote the inverse function of $-\frac{1}{2}\psi'(t)$ restricted to the interval $(0, 1]$. Then, in the worst case, the step size*

$$\bar{\alpha} := \frac{1}{2\delta} [\rho(\delta(v)) - \rho(2\delta(v))] \quad (27)$$

is the largest possible solution of inequality (26).

Proof. Given $\delta(v)$, we want to find the largest possible α such that (26) holds, irrespective the value of v_1 . Since ψ'' is decreasing, the derivative to v_1 of the expression at the left in (26) (i.e. $-\psi''(v_1 - 2\alpha\delta(v)) + \psi''(v_1)$) is negative. Hence, the smaller v_1 is, the smaller α will be. One has

$$\delta = \frac{1}{2} \|\nabla\Psi(v)\| \geq \frac{1}{2} |\psi'(v_1)| \geq -\frac{1}{2}\psi'(v_1).$$

Equality holds if and only if v_1 is the only coordinate in v that differs from 1, and $v_1 \leq 1$ (in which case $\psi'(v_1) \leq 0$). Hence, the worst situation for the step size occurs when v_1 satisfies

$$-\frac{1}{2}\psi'(v_1) = \delta(v). \tag{28}$$

The derivative to α of the expression at the left in (26) equals $2\delta\psi''(v_1 - 2\alpha\delta(v)) \geq 0$, and hence the left hand side is increasing in α . So the largest possible value of α satisfying (26), satisfies

$$-\frac{1}{2}\psi'(v_1 - 2\alpha\delta(v)) = 2\delta(v). \tag{29}$$

Due to the definition of ρ , Equations (28) and (29) can be written as

$$v_1 = \rho(\delta(v)), \quad v_1 - 2\alpha\delta(v) = \rho(2\delta(v)),$$

proving the lemma. □

Lemma 3.4. *Let ρ and $\bar{\alpha}$ be as defined in Lemma 3.3. Then, in the worst case situation (where $v_1 = \rho(\delta(v))$),*

$$\frac{1}{\psi''(\rho(2\delta(v)))} \leq \bar{\alpha} \leq \frac{1}{\psi''(\rho(\delta(v)))}. \tag{30}$$

Proof. As $-\frac{1}{2}\psi'$ is monotonically decreasing, on $[0, 1]$, its inverse ρ has the same property.

An immediate consequence of (27) is

$$\bar{\alpha} = \frac{1}{2\delta(v)} \int_{2\delta(v)}^{\delta(v)} \rho'(\sigma) d(v)\sigma = \frac{1}{\delta(v)} \int_{\delta(v)}^{2\delta(v)} \frac{d\sigma}{\psi''(\rho(\sigma))}. \tag{31}$$

Since ψ'' and ρ are monotonically decreasing, the argument of the last integral is maximal if $\sigma = \delta$ and minimal if $\sigma = 2\delta(v)$. Hence the inequalities in the lemma immediately follow. □

In the sequel we use

$$\tilde{\alpha} = \frac{1}{\psi''(\rho(2\delta(v)))} \tag{32}$$

as the default step size. By Lemma 3.4 we have $\tilde{\alpha} \leq \bar{\alpha}$.

Lemma 3.5 (Lem. 3.12 in [11]). *Let h be a twice differentiable convex function with $h(0) = 0$, $h'(0) < 0$, which attains its minimum at $t^* > 0$. If h'' is increasing for $t \in [0, t^*]$ then*

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

Lemma 3.6. *If the step size α is such that $\alpha \leq \bar{\alpha}$ then*

$$f(\alpha) \leq -\alpha \delta(v)^2. \tag{33}$$

Proof. Let the univariate function h be such that

$$h(0) = f_1(0) = 0, \quad h'(0) = f_1'(0) = -2\delta^2, \quad h''(\alpha) = 2\delta^2 \psi''(v_1 - 2\alpha\delta).$$

Due to Lemma 3.1, $f_1''(\alpha) \leq h''(\alpha)$. As a consequence, $f_1'(\alpha) \leq h'(\alpha)$ and $f_1(\alpha) \leq h(\alpha)$. We may write

$$\begin{aligned} h'(\alpha) &= -2\delta(v)^2 + 2\delta(v)^2 \int_0^\alpha \psi''(v_1 - 2\xi\delta(v)) \, d\xi \\ &= -2\delta^2 - \delta(v) (\psi'(v_1 - 2\alpha\delta(v)) - \psi'(v_1)). \end{aligned}$$

Since $\alpha \leq \bar{\alpha}$, lemma:estimate alpha 1, we have inequality (26) is certainly satisfied. Thus it follows that $h'(\alpha) \leq 0$, for all $\alpha \leq \bar{\alpha}$. Since ψ'' is decreasing, as a function of t , h'' is increasing in α . Hence Lemma 3.5 applies and we obtain

$$f(\alpha) \leq f_1(\alpha) \leq h(\alpha) \leq \frac{1}{2}\alpha h'(0) = -\alpha\delta(v)^2. \quad \square$$

Theorem 3.7. *Let ρ be as defined in Lemma 3.3 and $\tilde{\alpha}$ as in (32) and $\Psi(v) \geq 1$. Then*

$$f(\tilde{\alpha}) \leq -\frac{\delta(v)^2}{\psi''(\rho(2\delta(v)))} \leq -\frac{\delta(v)^{\frac{1}{3}}}{662}. \tag{34}$$

Proof. Since $\tilde{\alpha} \leq \bar{\alpha}$, Lemma 3.6 gives $f(\tilde{\alpha}) \leq -\tilde{\alpha} \delta(v)^2$, where $\tilde{\alpha} = \frac{1}{\psi''(\rho(2\delta(v)))}$. Thus the first inequality follows. To obtain the inverse function $t = \rho(s)$ of $-\frac{1}{2}\psi'(t)$ for $t \in (0, 1]$ we need to solve t from the equation

$$-\left(16t - 11 - \frac{1}{t^{\frac{3}{2}}} - \frac{4}{t}\right) = 2s.$$

This gives $\frac{1}{t^{\frac{3}{2}}} = 2s + 16t - 11 - \frac{4}{t} \leq 2s + 5$, whence

$$\rho(s) \geq \frac{1}{(2s + 5)^{\frac{2}{3}}}.$$

It follows that

$$\begin{aligned} \tilde{\alpha} &= \frac{1}{\psi''(\rho(2\delta(v)))} = \frac{1}{16 + \frac{4}{(\rho(2\delta(v)))^2} + \frac{3}{2(\rho(2\delta(v)))^{\frac{5}{2}}}} \\ &\geq \frac{1}{16 + 4(4\delta(v) + 5)^{\frac{4}{3}} + \frac{3}{2}(4\delta(v) + 5)^{\frac{5}{3}}}. \end{aligned}$$

Since $(4\delta(v) + 5)^{\frac{4}{3}} \leq (4\delta(v) + 5)^{\frac{5}{3}}$, using (19) (i.e., $\frac{5}{2}\delta(v) \geq 1$), it follows that

$$\tilde{\alpha} \geq \frac{1}{16 + \frac{11}{2}(4\delta(v) + 5)^{\frac{5}{3}}} \geq \frac{1}{662\delta(v)^{\frac{5}{3}}}.$$

Hence

$$f(\tilde{\alpha}) \leq -\frac{\delta(v)^2}{662\delta(v)^{\frac{5}{3}}} = -\frac{\delta(v)^{\frac{1}{3}}}{662}.$$

Thus the theorem follows. \square

Using (18), substitution gives

$$f(\tilde{\alpha}) \leq -\frac{\Psi(v)^{\frac{1}{6}}}{1655}.$$

4. ITERATION COMPLEXITY

In this section we derive the complexity bounds for large-update methods and small-update methods. An upper bound for the total number of iterations is obtained by multiplying (the upper bound for) the number of inner iterations K by the number of barrier parameter updates, which is bounded above by (cf. [13] Lem. II.17, p. 116)

$$\frac{1}{\theta} \log \frac{n}{\epsilon}.$$

Lemma 4.1 (Prop. 2.2 in [9]). *Let t_0, t_1, \dots, t_K be a sequence of positive numbers such that*

$$t_{k+1} \leq t_k - \kappa t_k^{1-\gamma}, \quad k = 0, 1, \dots, K-1, \quad (35)$$

where $\kappa > 0$ and $0 < \gamma \leq 1$. Then $K \leq \left\lceil \frac{t_0^\gamma}{\kappa\gamma} \right\rceil$.

Lemma 4.2. *If K denotes the number of inner iterations, we have*

$$K \leq 1986\Psi_0^{\frac{5}{6}}.$$

Proof. The definition of K implies $\Psi_{K-1} > \tau$ and $\Psi_K \leq \tau$ and

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

with $\kappa = \frac{1}{1655}$ and $\gamma = \frac{5}{6}$. Application of Lemma 4.1, with $t_k = \Psi_k$ yields the desired inequality. \square

Let $L = L(n, \theta, \tau)$, as defined in (22). Using $\psi_0 \leq L$, and Lemma 4.2 we obtain the following upper bound on the total number of iterations:

$$\frac{1986L^{\frac{5}{6}}}{\theta} \log \frac{n}{\epsilon}. \quad (36)$$

4.1. COMPLEXITY FOR LARGE-UPDATE AND SMALL-UPDATE METHODS

We finally have to estimate L , *i.e.*, to derive an upper bound for $\Psi(v)$ just after a μ -update. To do this we need to estimate an upper bound for ϱ .

From (15) and $\psi''(t) \geq 16$, we have

$$s = \psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \geq \int_1^t \int_1^\xi 16 d\zeta d\xi = 8(t-1)^2,$$

which implies

$$t = \varrho(s) \leq 1 + \sqrt{\frac{s}{8}}.$$

Using Lemma 2.2, with $\psi''(1) = \frac{43}{2}$, and (22), to estimate Ψ_0 we obtain

$$\Psi_0 \leq L(n, \theta, \tau) = \frac{43n}{4} \left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}} - 1 \right)^2 \leq \frac{43n}{4} \left(\frac{1 + \sqrt{\frac{\tau}{8n}}}{\sqrt{1-\theta}} - 1 \right)^2.$$

Using $1 - \sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \leq \theta$, this leads to

$$\psi_0 \leq \frac{43n}{4(1-\theta)} \left(\theta + \sqrt{\frac{\tau}{8n}} \right)^2 = \frac{43(\theta\sqrt{n} + \sqrt{\frac{\tau}{8}})^2}{4(1-\theta)}.$$

We conclude that the total number of iterations is bounded above by

$$\frac{K}{\theta} \log \frac{n}{\epsilon} \leq \frac{1986}{\theta} \left(\frac{43(\theta\sqrt{n} + \sqrt{\frac{\tau}{8}})^2}{4(1-\theta)} \right)^{\frac{5}{6}} \log \frac{n}{\epsilon}.$$

A large-update methods uses $\tau = O(n)$ and $\theta = \Theta(1)$. Then the right hand side expression is $O\left(n^{\frac{5}{6}} \log \frac{n}{\epsilon}\right)$.

For small-update methods use $\tau = \Theta(1)$ and $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$. Then the right hand side expression is $O\left(\sqrt{n} \log \frac{n}{\epsilon}\right)$.

5. CONCLUDING REMARKS

In this paper we prove that the iteration bound of a large-update interior-point method based on the kernel function considered in this paper is $O\left(n^{\frac{5}{6}} \log \frac{n}{\epsilon}\right)$, which improves the classical iteration complexity with a factor $n^{\frac{1}{6}}$. For small-update methods we obtain the best know iteration bound, namely $O\left(\sqrt{n} \log \frac{n}{\epsilon}\right)$.

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*, edited by T. Terlaky, Kluwer Academic Publishers, The Netherlands (1996) 189–252.
- [2] Y.Q. Bai, M. El Ghami, and C. Roos, A comparative study of new barrier functions for primal-dual interior-point algorithms in linear optimization. *SIAM J. Optim.* **15** (2004) 101–128.
- [3] 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.* **13** (2003) 766–782.
- [4] Y.Q. Bai, C. Roos, and M. El Ghami, A primal-dual interior-point method for linear optimization based on a new proximity function. *Optim. Methods Softw.* **17** (2002) 985–1008.
- [5] C.C. Gonzaga, Path following methods for linear programming. *SIAM Rev.* **34** (1992) 167–227.
- [6] D. den Hertog, Interior Point Approach to Linear, Quadratic and Convex Programming, *Mathematics and its Applications* **277**. Kluwer Academic Publishers, Dordrecht, 1994.
- [7] N. Megiddo, Pathways to the optimal set in linear programming. In *Progress in Mathematical Programming: Interior Point and Related Methods*, edited by N. Megiddo, Springer Verlag, New York, 1989, 131–158. Identical version in: *Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan*, (1986) 1–35.
- [8] R.D.C. Monteiro and I. Adler, Interior-path following primal-dual algorithms: Part I: Linear programming. *Math. Program.* **44** (1989) 27–41.
- [9] J. Peng, C. Roos, and T. Terlaky, A new class of polynomial primal-dual methods for linear and semidefinite optimization. To appear in *Eur. J. Oper. Res.*
- [10] J. Peng, C. Roos, and T. Terlaky, A new and efficient large-update interior-point method for linear optimization. *J. Comput. Tech.* **6** (2001) 61–80.
- [11] 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.
- [12] J. Peng, C. Roos, and T. Terlaky, *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*. Princeton University Press, 2002.
- [13] C. Roos, T. Terlaky, and J.-P. Vial, *Theory and Algorithms for Linear Optimization. An Interior-Point Approach*. John Wiley & Sons, Chichester, UK, 1997.
- [14] G. Sonnevend, An “analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In *System Modelling and Optimization : Proceedings of the 12th IFIP- Conference held in Budapest, Hungary, September 1985*, edited by A. Prékopa, J. Szelezsán, and B. Strazicky, *Lecture Notes in Control and Information Sciences*, Springer Verlag, Berlin, West-Germany, **84** (1986) 866–876.
- [15] M.J. Todd, Recent developments and new directions in linear programming. In *Mathematical Programming: Recent Developments and Applications*, edited by M. Iri and K. Tanabe, Kluwer Academic Press, Dordrecht (1989) 109–157.
- [16] Y. Ye, *Interior Point Algorithms, Theory and Analysis*, John Wiley and Sons, Chichester, UK, 1997.