J Optim Theory Appl (2008) 138: 341–359
DOI 10.1007/s10957-008-9389-z

A Class of Large-Update and Small-Update Primal-Dual Interior-Point Algorithms for Linear Optimization

Y.Q. Bai · G. Lesaja · C. Roos · G.Q. Wang · M. El Ghami

Published online: 7 May 2008

Abstract In this paper we present a class of polynomial primal-dual interior-point algorithms for linear optimization based on a new class of kernel functions. This class is fairly general and includes the classical logarithmic function, the prototype self-regular function, and non-self-regular kernel functions as special cases. The analysis of the algorithms in the paper follows the same line of arguments as in Bai et al. (SIAM J. Optim. 15:101–128, 2004), where a variety of non-self-regular kernel functions were considered including the ones with linear and quadratic growth terms. However, the important case when the growth term is between linear and quadratic was not considered. The goal of this paper is to introduce such class of kernel functions and to show that the interior-point methods based on these functions have favorable complexity results. They match the currently best known iteration bounds for the prototype self-regular function with quadratic growth term, the simple non-self-regular function with linear growth term, and the classical logarithmic kernel

Communicated by F.A. Potra.

This research is partially supported by the grant of National Science Foundation of China 10771133 and the Program of Shanghai Pujiang 06PJ14039.

Y.Q. Bai · G.Q. Wang
Department of Mathematics, Shanghai University, Shanghai, 200444, China

G. Lesaja (✉)
Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460-8093, USA
e-mail: goran@GeorgiaSouthern.edu

C. Roos
Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, Netherlands

M. El Ghami
Department of Computer Science, University of Bergen, Bergen, Norway

function. In order to achieve these complexity results, several new arguments had to be used.

Keywords Linear optimization · Interior-point methods · Primal-dual methods · Complexity · Kernel functions

1 Introduction

We consider the linear optimization (LO) problem in the standard form

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

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

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

We assume that both (P) and (D) satisfy the interior-point condition (IPC), that is, there exists a point (x^0, s^0, y^0) such that

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

which means that the interiors of the feasible regions of (P) and (D) are not empty. It is well known that the IPC can be assumed without loss of generality. In fact, we may assume that $x^0 = s^0 = e$, where e denotes the all-one vector. We will adopt this assumption in the sequel. For details on the IPC and some other properties mentioned below, for example, see [2].

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

$$Ax = b, \quad x \geq 0, \tag{1a}$$

$$A^T y + s = c, \quad s \geq 0, \tag{1b}$$

$$xs = 0, \tag{1c}$$

where the vector xs denotes the componentwise product of the vectors x and s. The basic idea of primal-dual interior-point algorithms is to replace the third equation in (1), which is commonly known as *complementarity condition* for (P) and (D), by the parameterized equation $xs = \mu e$, with $\mu > 0$. Thus, we consider the system

$$Ax = b, \quad x \geq 0, \tag{2a}$$

$$A^T y + s = c, \quad s \geq 0, \tag{2b}$$

$$xs = \mu e. \tag{2c}$$

Since $\mathrm{rank}(A) = m$ and the IPC holds, the parameterized system (2) has a unique solution, for each $\mu > 0$. This solution is denoted as $(x(\mu), y(\mu), s(\mu))$ and we call $x(\mu)$ the μ-center of (P) and $(y(\mu), s(\mu))$ the μ-center of (D). The set of μ-centers

(with μ running through all positive real numbers) gives a homotopy path, which is called *the central path* of (P) and (D). The relevance of the central path for LO was first recognized by Sonnevend [3] and Megiddo [4]. The main property of the central path can be summarized as follows. If $\mu \to 0$, then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D).

The limiting property of the central path mentioned above leads naturally to the main idea of the iterative methods for solving (P) and (D): Trace the central path while reducing μ at each iteration. However, tracing the central path exactly would be too costly and inefficient. It has been shown that it is sufficient to trace the central path approximately while still maintaining good properties of the algorithms.

The general outline of the generic interior-point primal-dual method is as follows. First, without loss of generality, it is assumed that a point $(x(\mu), y(\mu), s(\mu))$ is known for some positive parameter μ. For example, due to the IPC assumption we may set $\mu = 1$, with $x(1) = s(1) = e$. Then μ is decreased to $\mu_+ := (1 - \theta)\mu$, for some $\theta \in (0, 1)$ and, redefining $\mu := \mu_+$, we solve the following Newton system:

$$A\Delta x = 0, \tag{3a}$$

$$A^T \Delta y + \Delta s = 0, \tag{3b}$$

$$s\Delta x + x\Delta s = \mu e - xs. \tag{3c}$$

Because A has a full row rank, system (3) has a unique solution for any $(x, s) > 0$. The solution $(\Delta x, \Delta y, \Delta s)$ is known as the Newton direction and this direction is used in all existing implementations of the primal-dual methods. By taking a step along the search direction, one constructs a new triple (x_+, y_+, s_+) with

$$x_+ = x + \alpha\Delta x, \qquad y_+ = y + \alpha\Delta y, \qquad s_+ = s + \alpha\Delta s, \tag{4}$$

where α denotes the stepsize, $\alpha \in (0, 1)$, which has to be chosen appropriately. If necessary, we repeat the procedure until we find iterates that are in a certain neighborhood of the μ-center $(x(\mu), y(\mu), s(\mu))$. Then μ is again reduced by the factor $1 - \theta$ and Newton's method is applied targeting at the new μ-centers, and so on. This process is repeated until μ is small enough, for example, until $n\mu \le \epsilon$, where ϵ is a small positive number. At this stage we have found ϵ-approximate solutions of (P) and (D).

For the analysis of primal-dual interior-point algorithms it is convenient to associate to any triple (x, s, μ) with $x > 0$ primal feasible, $s > 0$ dual feasible and $\mu > 0$, the vector

$$v := \sqrt{\frac{xs}{\mu}}. \tag{5}$$

Note that the pair (x, s) coincides with the μ-center $(x(\mu), s(\mu))$ if and only if $v = e$. The *scaled search directions* d_x and d_s are introduced according to

$$d_x := \frac{v\Delta x}{x}, \qquad d_s := \frac{v\Delta s}{s}, \tag{6}$$

where the operations are componentwise product and division. Then, the system (3) can be rewritten as

$$\bar{A}d_x = 0, \tag{7a}$$

$$\bar{A}^T \Delta y + d_s = 0, \tag{7b}$$

$$d_x + d_s = v^{-1} - v, \tag{7c}$$

where

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

Given a vector x, $X = \operatorname{diag}(x)$ represents a diagonal matrix where the components of the vector x are placed on the main diagonal in their natural order. Note that d_x and d_s are orthogonal vectors, since the vector d_x belongs to the null space and d_s to the row space of the matrix $\bar{A}$. Hence, $d_x = d_s = 0$ if and only if $v^{-1} - v = 0$, which is equivalent to $v = e$. We conclude that $d_x = d_s = 0$ holds if and only if the pair (x, s) coincides with the μ-center $(x(\mu), s(\mu))$. A crucial observation is that the right-hand side $v^{-1} - v$ in the third equation of (7) equals minus the gradient of the function

$$\Psi_c(v) := \sum_{i=1}^{n} \left(\frac{v_i^2 - 1}{2} - \log v_i \right), \tag{8}$$

where v_i represents the ith component of the vector v. In other words,

$$d_x + d_s = -\nabla \Psi_c(v). \tag{9}$$

One may easily verify that $\nabla^2 \Psi_c(v) = \operatorname{diag}(e + v^{-2})$. Since this matrix is positive definite, $\Psi_c(v)$ is strictly convex. Moreover, since $\nabla \Psi_c(e) = 0$, it follows that $\Psi_c(v)$ attains its minimal value at $v = e$, with $\Psi_c(e) = 0$. Thus, it follows that $\Psi_c(v)$ is nonnegative everywhere and vanishes if and only if $v = e$, that is, if and only if $x = x(\mu)$ and $s = s(\mu)$. Hence, we see that the μ-centers $x(\mu)$ and $s(\mu)$ can be characterized as the minimizer of the function $\Psi_c(v)$.

The equation (9) is called the *scaled centering equation*. Its importance arises from the fact that it essentially defines the search directions. The above observation regarding the function $\Psi_c(v)$ leads to an obvious generalization: we can replace $\Psi_c(v)$ by any strictly convex function $\Psi(v)$, $v \in \mathrm{R}^n_{++}$, such that $\Psi(v)$ is minimal at $v = e$ and $\Psi(e) = 0$. Thus, the new scaled centering equation becomes

$$d_x + d_s = -\nabla \Psi(v). \tag{10}$$

Note that since d_x and d_s are orthogonal, we will still have $d_x = 0$ and $d_s = 0$ if and only if $v = e$, that is, if and only if $x = x(\mu)$ and $s = s(\mu)$, as it should. Function $\Psi(v)$ is called a scaled *barrier function*. Of course, different barrier functions lead to different Newton directions.

To simplify matters we will restrict ourselves to the case where the barrier function $\Psi(v)$ is separable with identical coordinate functions $\psi(v_i)$. Thus,

$$\Psi(v) = \sum_{i=1}^{n} \psi(v_i), \tag{11}$$

where $\psi(t) : (0, +\infty) \to [0, +\infty)$ is twice differentiable and attains its minimum at $t = 1$, with $\psi(1) = 0$. Following the terminology introduced in [5–7], we call the univariate function $\psi(t)$ the *kernel function* of the barrier function $\Psi(v)$. Obviously, in the case

$$\psi_c(t) = \frac{t^2 - 1}{2} - \log t, \tag{12}$$

and $\psi(t) = \psi_c(t)$, we have $\Psi(v) = \Psi_c(v)$, and then the new search direction coincides with the classical Newton direction for primal-dual methods. One may easily verify that $\Phi_c(x, s, \mu) := \Psi_c(v)$ is the 'classical' primal-dual logarithmic barrier function. We call $\psi_c(t)$ the *classical kernel function*. Note that the term $-\log t$ dominates the behavior of this kernel function if t approaches zero, whereas the term $\frac{t^2-1}{2}$ dominates the behavior of the kernel function if t tends to infinity. We call the first term the *barrier term* and the second term the *growth term* of the classical kernel function.

The new search directions $(\Delta x, \Delta y, \Delta s)$ are obtained by first solving the system that consists of the first two equations in (7) and the scaled centering equation (10). Once d_x and d_s are found we apply (6) to find Δx and Δs. They can also be obtained directly by solving the following system:

$$A\Delta x = 0, \tag{13a}$$

$$A^T \Delta y + \Delta s = 0, \tag{13b}$$

$$s\Delta x + x\Delta s = -\mu v \nabla \Psi(v). \tag{13c}$$

In principle, each kernel function gives rise to a primal-dual algorithm. The generic form of this algorithm is shown in Fig. 1. The parameters τ, θ and the stepsize α in the algorithm should be tuned in such a way that the number of iterations required by the algorithm is as small as possible. The specific default stepsize α will be discussed in Sect. 3. Obviously, the resulting iteration bound will depend on the kernel function, and our main task becomes to find a kernel function that minimizes the iteration bound.

In the literature two types of methods are distinguished: small-update methods and large-update methods, according to the value of the *barrier-update parameter* θ. Large-update methods are characterized by the fact that θ is a fixed constant ($\theta \in (0, 1)$), independent of the dimension n of the problem, whereas small-update methods use a value of θ that depends of the dimension of the problem, with $\theta = O(\frac{1}{\sqrt{n}})$. The iteration bounds for large-update and small-update methods based on the classical kernel function are $O(n \log \frac{n}{\epsilon})$ and $O(\sqrt{n} \log \frac{n}{\epsilon})$, respectively.

The first contribution in the design and analysis of the primal-dual IPMs based on the use of kernel functions different than the classical logarithmic kernel function (12)

Generic Primal-Dual Algorithm for LO

Step 1. Determine input parameters: a threshold parameter $\tau \geq 1$, an accuracy parameter $\varepsilon > 0$, and a fixed barrier update parameter θ, $0 < \theta < 1$.
Step 2. Set $x := e, s := e, \mu := 1$.
Step 3. Until $n\mu \leq \varepsilon$ perform the following steps:
Step 4. Calculate $\mu := (1-\theta)\mu$ and $v := \sqrt{\frac{xs}{\mu}}$.
Step 5. Until $\Psi(v) \leq \tau$ perform the following steps:
Step 6. Calculate direction $(\Delta x, \Delta y, \Delta s)$ by solving (13).
Step 7. Calculate stepsize α.
Step 8. Update $x := x + \alpha\Delta x$, $s := s + \alpha\Delta s$, $y := y + \alpha\Delta y$.

Fig. 1 Generic algorithm

was presented in [7]. The recently published monograph [8] can be considered as an extensive version of [7]. The kernel functions introduced and analyzed in [7, 8] are so-called self-regular functions. The iteration bound of $O(\sqrt{n}\log n \log\frac{n}{\epsilon})$ for large-update methods was obtained, which is currently the best iteration bound known for these type of methods and is a considerable improvement over the $O(n\log\frac{n}{\epsilon})$ bound for large-update methods based on the classical logarithmic barrier function.

Subsequently Bai et al. proposed IPMs for LO using classes of kernel functions that are not self-regular. For some of the classes they managed to match the best known iteration bounds obtained by using self-regular functions. The comprehensive overview of these results can be found in [1].

Until recently almost all existing kernel functions considered in the literature have had at a least quadratic growth term. The only exception has been the simple kernel function with the linear growth term considered first in [9]. However, an important question remained open, namely whether there exist a kernel function with a growth term whose degree is between linear and quadratic, and which gives rise to an efficient primal-dual IPM. This paper offers a positive answer to that question.

The proposed new class of kernel functions is

$$\psi_{p,q}(t) = \frac{t^{p+1}-1}{p+1} + \frac{t^{1-q}-1}{q-1}, \quad t > 0,\ p \in [0,1],\ q > 1, \tag{14}$$

where p is a *growth* parameter and q is a *barrier* parameter. Note that $\frac{t^{1-q}-1}{q-1} \to -\log t$ when $q \to 1$. Thus, for $q = 1$, we define $\psi_{p,q}(t) = \frac{t^{p+1}-1}{p+1} - \log t$, $t > 0$, $p \in [0,1]$.

The class of kernel functions introduced above contains a few known kernel functions:

- For $p = 1$ and $q = 1$, $\psi(t) = \psi_c(t) = \frac{t^2-1}{2} - \log t$, the classical logarithmic kernel function, see [2].
- For $p = 1$ and $q > 1$, $\psi(t) = \frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1}$, the prototype self-regular kernel function, see [8].
- For $p = 0$ and $q = 2$, $\psi(t) = t - \frac{1}{t} - 2$, the kernel function with linear growth rate, which from an algebraic point of view is the simplest kernel function, see [9].

It is worth mentioning that for $0 \leq p < 1$ and $q > 1$, $\psi_{p,q}(t)$ is a non-self-regular function. Our goal is to analyze the large-update and small-update primal-dual IPMs based on $\psi_{p,q}(t)$, as given in (14), with the scaled barrier function defined by (11), and of course with the intention to derive as good as possible iterations bounds.

The paper is organized as follows. In Sect. 2, we present some properties of the new kernel function, as well as several properties of the barrier function based on the kernel function. The estimate of the stepsize is discussed in Sect. 3. The iteration bounds of the algorithm with large-update and small-updates, are derived in Sect. 4. Finally, some concluding remarks follow in Sect. 5.

In addition to the notation already introduced, some additional notation will be used throughout the paper. If $x = (x_1; x_2; \ldots; x_n) \in \mathrm{R}^n$, then $\|x\| = \sqrt{x^T x}$ denotes the standard 2-norm of the vector x and $x_{\min}$ denotes the smallest and $x_{\max}$ the largest value of the components of x. Finally, if $g(x) \geq 0$ is a real valued function of a real nonnegative variable, the notation $g(x) = O(x)$ means that $g(x) \leq \bar{c}x$ for some positive constant $\bar{c}$ and $g(x) = \Theta(x)$ means that $\bar{c}_1 x \leq g(x) \leq \bar{c}_2 x$ for two positive constants $\bar{c}_1$ and $\bar{c}_2$.

2 Properties of the Kernel and Barrier Functions

In this section we discuss some properties of the new kernel function $\psi(t)$ defined in (14) and the corresponding barrier function that will be used in the complexity analysis of the algorithm. According to (14), if $q > 1$, then the scaled barrier function $\Psi(v)$ is given by

$$\Psi(v) = \sum_{i=1}^{n} \psi(v_i) = \sum_{i=1}^{n} \left(\frac{v_i^{p+1} - 1}{p+1} + \frac{v_i^{1-q} - 1}{q-1} \right), \tag{15}$$

where $0 \leq p \leq 1$, $q > 1$, $v \in \mathrm{R}^n_{++}$, and v_i are components of the vector v. In the analysis of the algorithm, we also use the norm-based *proximity measure* $\delta(v)$ defined by

$$\delta(v) := \frac{1}{2}\|\nabla\Psi(v)\| = \frac{1}{2}\sqrt{\sum_{i=1}^{n} (\psi'(v_i))^2}. \tag{16}$$

Since $\Psi(v)$ is strictly convex and attains its minimum value of zero at $v = e$, we have

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

The derivatives of $\psi(t)$ play a crucial role in our analysis. Thus, we write down the first three derivatives:

$$\psi'(t) = t^p - t^{-q}, \tag{17a}$$

$$\psi''(t) = pt^{p-1} + qt^{-q-1}, \tag{17b}$$

$$\psi'''(t) = p(p-1)t^{p-2} - q(q+1)t^{-q-2}. \tag{17c}$$

Note that these expressions are also valid in the limiting case, that is, if $q = 1$. It is quite straightforward to verify the following:

$$\psi(1) = \psi'(1) = 0, \qquad \lim_{t\to 0} \psi(t) = \lim_{t\to\infty} \psi(t) = +\infty. \tag{18}$$

Moreover, from (17) we conclude that $\psi(t)$ is strictly convex and $\psi''(t)$ is monotonically decreasing on the interval $t \in (0, +\infty)$.

Several results stated below have been discussed in [1] and references therein. The proofs that are simple modifications of the existing ones are omitted.

Lemma 2.1 *If $t_1 > 0$ and $t_2 > 0$, then*

$$\psi(\sqrt{t_1 t_2}) \le \frac{1}{2}(\psi(t_1) + \psi(t_2)).$$

Proof This result follows easily by using Lemma 2.1.2 in [8], which states that the above inequality holds if and only if $t\psi''(t) + \psi'(t) \ge 0$ for all $t > 0$. Hence, since

$$\begin{aligned} t\psi''(t) + \psi'(t) &= t\left(pt^{p-1} + \frac{q}{t^{q+1}}\right) + t^p - \frac{1}{t^q} \\ &= pt^p + \frac{q}{t^q} + t^p - \frac{1}{t^q} = (p+1)t^p + \frac{(q-1)}{t^q} > 0, \end{aligned}$$

the proof is complete. □

Lemma 2.2 *If $t \ge 1$, then*

$$\frac{\psi'(t)}{2}(t-1) \le \psi(t) \le \frac{p+q}{2}(t-1)^2.$$

Proof If $f(t) = 2\psi(t) - (t-1)\psi'(t)$, then $f'(t) = \psi'(t) - (t-1)\psi''(t)$ and $f''(t) = -(t-1)\psi'''(t)$. Also $f(1) = 0$ and $f'(1) = 0$. Since $\psi'''(t) < 0$, it follows that if $t \ge 1$ then $f''(t) \ge 0$ whence $f'(t) \ge 0$ and $f(t) \ge 0$. This implies the first inequality. The second inequality follows from Taylor's theorem and the fact that $\psi''(1) = p + q$. □

The next lemma combines the results of Lemma 4.8 and Lemma 3.1 in [1] and is stated without proof.

Lemma 2.3 *Suppose that* $\psi(t_1) = \psi(t_2)$, *with* $t_1 \leq 1 \leq t_2$. *The following statements hold*:

(i) *One has* $\psi'(t_1) \leq 0$, $\psi'(t_2) \geq 0$, *and*

$$-\psi'(t_1) \geq \psi'(t_2). \tag{19}$$

(ii) *If* $\beta \geq 1$, *then*

$$\psi(\beta t_1) \leq \psi(\beta t_2); \tag{20}$$

equality holds if and only if $\beta = 1$ *or* $t_1 = t_2 = 1$.

Lemma 2.4 *If* $t \geq 1$, *then*

$$\psi'(t)^2 \geq 2\psi(t)\psi''(t).$$

Proof Defining $f(t) = \psi'(t)^2 - 2\psi(t)\psi''(t)$, one has $f(1) = 0$ and

$$f'(t) = 2\psi'(t)\psi''(t) - 2\psi'(t)\psi''(t) - 2\psi(t)\psi'''(t) = -2\psi(t)\psi'''(t) > 0.$$

This proves the lemma. □

Lemma 2.5 *Let* $\rho(s) : [0, \infty) \to (0, 1]$ *be the inverse function of* $-\frac{1}{2}\psi'(t)$ *for* $t \leq 1$. *The following inequality holds*:

$$\rho(s) \geq \frac{1}{(1+2s)^{\frac{1}{q}}}. \tag{21}$$

Proof Since $s = -\frac{1}{2}\psi'(t)$, we have

$$-2s = t^p - t^{-q} \quad \Longrightarrow \quad t^{-q} = t^p + 2s \leq 1 + 2s.$$

Since $t = \rho(s)$, this implies the lemma. □

Lemma 2.6 *If* $t \geq 1$ *and* $q \geq 2 - p$, *then*

$$t \leq 1 + \sqrt{t\psi(t)}.$$

Proof Defining $f(t) = t\psi(t) - (t-1)^2$, we have $f(1) = 0$ and $f'(t) = \psi(t) + t\psi'(t) - 2(t-1)$. Moreover, it is clear that $f'(1) = 0$ and

$$\begin{aligned} f''(t) &= 2\psi'(t) + t\psi''(t) - 2 = (2+p)t^p + (q-2)t^{-q} - 2 \\ &\geq pt^p + (q-2)t^{-q} \geq p(t^p - t^{-q}) \geq 0. \end{aligned}$$

The second inequality above is due to the fact that $q \geq 2 - p$. Thus, we obtain $t\psi(t) \geq (t-1)^2$, which implies the lemma. □

The lemma below is important because it provides a base for obtaining very good iterations bounds for both small-update and large-update algorithms.

Lemma 2.7 *Let $\varrho : [0, \infty) \to [1, \infty)$ be the inverse function of $\psi(t)$ for $t \geq 1$. The following inequalities hold*:

$$(1 + (p+1)s)^{\frac{1}{p+1}} \leq \varrho(s) \leq 1 + s + \sqrt{s^2 + 2s}. \tag{22}$$

Moreover, if $q \geq 2 - p$, then

$$\varrho(s) \leq 1 + \sqrt{s + s^2 + s\sqrt{s^2 + 2s}}. \tag{23}$$

Proof Since $q > 1$ and $t \geq 1$, we have

$$s = \psi(t) = \frac{t^{p+1} - 1}{p+1} + \frac{t^{1-q} - 1}{q-1} \leq \frac{t^{p+1} - 1}{p+1}.$$

Hence, the first inequality in (22) follows. The second inequality in (22) follows by using the first inequality of Lemma 2.2,

$$\begin{aligned} s = \psi(t) &\geq \frac{1}{2}(t-1)\psi'(t) = \frac{1}{2}(t-1)(t^p - t^{-q}) \\ &\geq \frac{1}{2}(t-1)\left(1 - \frac{1}{t}\right) = \frac{1}{2}\left(t + \frac{1}{t} - 2\right). \end{aligned}$$

Hence, solving the following inequality:

$$t^2 - 2(1+s)t + 1 \leq 0,$$

leads to

$$t = \varrho(s) \leq 1 + s + \sqrt{s^2 + 2s}. \tag{24}$$

Finally, let $q \geq 2 - p$. By Lemma 2.6, one has

$$t \leq 1 + \sqrt{t\psi(t)} \leq 1 + \sqrt{ts}.$$

Substitution of the upper bound for t given by (24) leads to

$$\varrho(s) \leq 1 + \sqrt{s + s^2 + s\sqrt{s^2 + 2s}}.$$

This completes the proof of the lemma. □

The theorem below provides a bound for $\delta(v)$ in terms of $\Psi(v)$ which will play an important role in the analysis of the algorithm. The theorem is a special case of Theorem 3.2 in [1] and is therefore stated without proof.

Theorem 2.1 *The following inequality holds*:

$$\delta(v) \geq \frac{1}{2}\psi'(\varrho(\Psi(v))). \tag{25}$$

Corollary 2.1 *If* $\Psi(v) \geq \tau \geq 1$, *then*

$$\delta(v) \geq \frac{1}{6}(\Psi(v))^{\frac{p}{1+p}}. \tag{26}$$

Proof Using Theorem 2.1 and the fact that $\Psi(v) \geq \tau \geq 1$, we have

$$\begin{aligned}\delta(v) &\geq \frac{1}{2}\psi'(\varrho(\Psi(v))) = \frac{1}{2}((\varrho(\Psi(v)))^p - (\varrho(\Psi(v)))^{-q}) \\ &\geq \frac{1}{2}((\varrho(\Psi(v)))^p - (\varrho(\Psi(v)))^{-1}).\end{aligned}$$

Note that $t^p - \frac{1}{t}$ is monotonically increasing in t. Thus, by using the first inequality in (22), we obtain

$$\begin{aligned}\delta(v) &\geq \frac{1}{2}\frac{(\varrho(\Psi(v)))^{p+1} - 1}{\varrho(\Psi(v))} \geq \frac{1}{2}\frac{(1+(p+1)\Psi(v))^{\frac{p+1}{p+1}} - 1}{(1+(p+1)\Psi(v))^{\frac{1}{p+1}}} \\ &= \frac{1}{2}\frac{(p+1)\Psi(v)}{(1+(p+1)\Psi(v))^{\frac{1}{p+1}}} \geq \frac{1}{2}\frac{\Psi(v)}{(3\Psi(v))^{\frac{1}{p+1}}} \geq \frac{1}{6}(\Psi(v))^{\frac{p}{p+1}},\end{aligned}$$

which proves the theorem. □

According to the algorithm, at the start of each inner iteration we have $\Psi(v) > \tau \geq 1$. Therefore, the above corollary is very helpful for deriving the upper bound on the number of inner iteration, which will be discussed in Sect. 4.

3 Analysis of the Algorithm

3.1 Growth Behavior of the Barrier Function at the Start of Outer Iteration

At the start of each outer iteration of the algorithm, just before the update of the parameter 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 subsection we derive an estimate for the effect of a μ-update on the value of $\Psi(v)$.

The following theorem yields an upper bound for $\Psi(v)$ after the μ-update in terms of the inverse function of $\psi(t)$ for $t \geq 1$. It is equivalent to Theorem 3.2 in [1]. Since the proof is not significantly affected by the introduction of the new kernel function, it is omitted.

Theorem 3.1 *Let $\varrho : [0, \infty) \to [1, \infty)$ be defined as in Lemma 2.7. Then, for any positive vector v and any $\beta \geq 1$, the following inequality holds*:

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

Corollary 3.1 *Let $0 \leq \theta \leq 1$ and $v_+ = \frac{v}{\sqrt{1-\theta}}$. If $\Psi(v) \leq \tau$, then*

$$\Psi(v_+) \leq n\psi\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}\right) \leq \frac{(p+q)n}{2}\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}} - 1\right)^2. \tag{27}$$

Proof With $\beta \geq 1$ and $\Psi(v) \leq \tau$ the first inequality follows from Theorem 3.1. The second inequality follows by using Lemma 2.2 and $\psi''(1) = \frac{p+q}{q}$. □

Lemma 3.1 *We have the following upper bounds on the value of $\Psi(v_+)$ after a μ-update*:

$$\Psi(v_+) \leq L_1 := n\psi\left(\frac{1 + \frac{\tau}{n} + \sqrt{(\frac{\tau}{n})^2 + \frac{2\tau}{n}}}{\sqrt{1-\theta}}\right), \quad q > 1, \tag{28}$$

and

$$\Psi(v_+) \leq L_2 := n\psi\left(\frac{1 + \sqrt{\frac{\tau}{n} + \frac{\tau^2}{n^2} + \frac{\tau}{n}\sqrt{\frac{\tau^2}{n^2} + \frac{2\tau}{n}}}}{\sqrt{1-\theta}}\right), \quad q \geq 2 - p. \tag{29}$$

Proof At the start of each outer iteration we have $\Psi(v) \leq \tau$. By Corollary 3.1, the value of $\Psi(v)$ is bounded above by (27) after the update of parameter μ to $(1-\theta)\mu$. Using two different upper bounds for $\varrho(s)$ stated in (22) and (23), respectively, we obtain the bounds stated in the lemma. □

3.2 Determining the Stepsize

In this section, we determine a default stepsize which not only keeps the iterations feasible but also gives rise to a sufficiently large decrease of $\Psi(v)$, as defined in (15), 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 were first used in [7], and later in [5, 6].

In each inner iteration we first compute the search direction $(\Delta x, \Delta y, \Delta s)$ from the system (13). After a stepsize α is determined the new iterate (x_+, y_+, s_+), is calculated by (4).

Recall that during an inner iteration the parameter μ is fixed. Hence, after the step in the direction $(\Delta x, \Delta y, \Delta s)$ with the stepsize α, the new v-vector is given by

$$v_+ = \sqrt{\frac{x_+ s_+}{\mu}}. \tag{30}$$

Since

$$x_+ = x\left(e + \alpha\frac{\Delta x}{x}\right) = x\left(e + \alpha\frac{d_x}{v}\right) = \frac{x}{v}(u + \alpha d_x),$$
$$s_+ = s\left(e + \alpha\frac{\Delta s}{s}\right) = s\left(e + \alpha\frac{d_s}{v}\right) = \frac{s}{v}\left(v + \alpha d_s\right),$$
$$xs = \mu v^2,$$

we obtain

$$v_+ = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}.$$

Next, we consider the decrease in Ψ as a function of α. We define two functions

$$f(\alpha) = \Psi(v_+) - \Psi(v), \tag{31}$$

and

$$f_1(\alpha) := \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)) - \Psi(v). \tag{32}$$

Lemma 2.1 implies that

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

The above inequality shows that $f_1(\alpha)$ is an upper bound of $f(\alpha)$. Obviously,

$$f(0) = f_1(0) = 0.$$

Taking the derivative with respect to α, we get

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

From the above equation and using (10), we obtain

$$f_1'(0) = \frac{1}{2}\nabla\Psi(v)^T(d_x + d_s) = -\frac{1}{2}\nabla\Psi(v)^T\nabla\Psi(v) = -2\delta(v)^2. \tag{33}$$

Differentiating once again, we get

$$f_1''(\alpha) = \frac{1}{2}\sum_{i=1}^{n}(\psi''(v_i + \alpha d_{xi})d_{xi}^2 + \psi''(v_i + \alpha d_{si})d_{si}^2) > 0, \tag{34}$$

unless $d_x = d_s = 0$. It is worthwhile to point out that during an inner iteration x and s are not both at the μ-center since $\Psi(v) \ge \tau > 0$, so we may conclude that $f_1(\alpha)$ is strictly convex in α.

In what follows we state several lemmas that will be used in obtaining a suitable lower bound on the stepsize α. They are variants of similar lemmas in [1, 6, 9] and are stated without the proofs. Also, in the sequel we simplify the notation $\delta(v)$ to just δ.

Lemma 3.2 (Lemma 3.1 in [6]) *The following inequality holds*:

$$f_1''(\alpha) \leq 2\delta^2 \psi''(v_{\min} - 2\alpha\delta). \tag{35}$$

Lemma 3.3 (Lemma 3.2 in [6]) *If the stepsize* α *satisfies*

$$-\psi'(v_{\min} - 2\alpha\delta) + \psi'(v_{\min}) \leq 2\delta, \tag{36}$$

then

$$f_1'(\alpha) \leq 0. \tag{37}$$

Lemma 3.4 (Lemma 3.3 in [9]) *The largest possible value of the stepsize* α *satisfying* (37) *is given by*

$$\bar{\alpha} := \frac{1}{2\delta}(\rho(\delta) - \rho(2\delta)). \tag{38}$$

Lemma 3.5 (Lemma 4.4 in [1]) *Let* ρ *and* $\bar{\alpha}$ *be defined by* (21) *and* (38) *respectively. The following inequality holds*:

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}. \tag{39}$$

Theorem 3.2 *We have*

$$\bar{\alpha} \geq \tilde{\alpha} := \frac{1}{(p+q)(1+4\delta)^{\frac{q+1}{q}}}. \tag{40}$$

Proof Using Lemma 3.5, (21) and the fact that $\psi''(t)$ is monotonically decreasing for $t \in (0, +\infty)$, we have

$$\begin{aligned}
\bar{\alpha} &\geq \frac{1}{\psi''(\rho(2\delta))} \\
&\geq \frac{1}{p(1+4\delta)^{\frac{1-p}{q}} + q(1+4\delta)^{\frac{1+q}{q}}} \\
&\geq \frac{1}{(p+q)(1+4\delta)^{\frac{q+1}{q}}} \\
&= \tilde{\alpha},
\end{aligned}$$

proving the theorem. □

We will use $\tilde{\alpha}$ as default stepsize.

3.3 Decrease of the Barrier Function During an Inner Iteration

Using the lower bound on the stepsize obtained in (40) we can obtain results on the decrease of the barrier function.

Lemma 3.6 (Lemma 4.5 in [1]) *If the stepsize α is such that $\alpha \leq \bar{\alpha}$, then*

$$f(\alpha) \leq -\alpha\delta^2. \tag{41}$$

Theorem 3.3 *The following inequality holds*:

$$f(\tilde{\alpha}) \leq -\frac{1}{60(p+q)}\Psi(v)^{\frac{p(q-1)}{q(p+1)}}.$$

Proof According to Lemma 3.6, if the stepsize α is such that $\alpha \leq \bar{\alpha}$, then $f(\alpha) \leq -\alpha\delta^2$. By (40), the default stepsize $\tilde{\alpha}$ satisfies $\tilde{\alpha} \leq \alpha$, hence, the following upper bound for $f(\alpha)$ is obtained $f(\tilde{\alpha}) \leq -\tilde{\alpha}\delta^2$. Using Corollary 2.1, after some elementary reductions, we obtain

$$\begin{aligned} f(\tilde{\alpha}) &\leq -\frac{\delta^2}{(p+q)(1+4\delta)^{\frac{q+1}{q}}} \\ &\leq -\frac{\Psi(v)^{\frac{2p}{p+1}}}{36(p+q)(1+\frac{2}{3}\Psi(v)^{\frac{p}{p+1}})^{\frac{q+1}{q}}} \\ &\leq -\frac{1}{60(p+q)}\Psi(v)^{\frac{p(q-1)}{q(p+1)}}. \end{aligned}$$

Here we used the fact that the first upper bound for $f(\tilde{\alpha})$ is monotonically decreasing in δ. Thus, the proof is complete. □

4 Complexity of the Algorithm

We have all the ingredients to derive an upper bound on the number of iterations needed by the algorithm to obtain an ϵ-approximate solution for (P) and (D). In the first section we derive an iteration bound for large-update methods and in the second one we derive an iteration bound for small-update methods.

4.1 Iteration Bound for the Large-Update Method

We need to count how many inner iterations are required to return to the situation where $\Psi(v) \leq \tau$ after a μ-update. We denote the value of $\Psi(v)$ after the μ-update as Ψ_0, and the subsequent values in the same outer iteration are denoted as Ψ_k, $k = 1, 2, \ldots, K$, where K denotes the total number of inner iterations in the outer

iteration. By using (28) in Lemma 3.1, we have

$$\Psi_0 \leq n\psi\left(\frac{1+\frac{\tau}{n}+\sqrt{(\frac{\tau}{n})^2+\frac{2\tau}{n}}}{\sqrt{1-\theta}}\right).$$

Since $\psi(t) \leq \frac{t^{p+1}-1}{p+1}$ when $t \geq 1$, and $1-(1-\theta)^{\frac{p+1}{2}} \leq \theta$, after some elementary reductions we obtain

$$\Psi_0 \leq \frac{n\theta+(p+1)\tau+n(p+1)\sqrt{(\frac{\tau}{n})^2+\frac{2\tau}{n}}}{(p+1)(1-\theta)^{\frac{p+1}{2}}}. \tag{42}$$

Now, Theorem 3.3 leads to

$$\Psi_{k+1} \leq \Psi_k - \beta(\Psi_k)^{1-\gamma}, \quad k=0,1,\ldots,K-1, \tag{43}$$

where $\beta = \frac{1}{60(p+q)}$ and $\gamma = \frac{p+q}{q(p+1)}$. Using Lemma A.3, (42) and (43) we obtain the following upper bound on the number K of inner iterations.

Lemma 4.1 *The following inequality holds*:

$$K \leq 60q(p+1)\left(\frac{n\theta+(p+1)\tau+n(p+1)\sqrt{(\frac{\tau}{n})^2+\frac{2\tau}{n}}}{(p+1)(1-\theta)^{\frac{p+1}{2}}}\right)^{\frac{p+q}{q(p+1)}}. \tag{44}$$

Now, we can derive an upper bound on the total number of iterations needed by the large-update version of the algorithm in Fig. 1.

Theorem 4.1 *Given that $\theta = \Theta(1)$ and $\tau = O(n)$, which are characteristics of the large-update methods, the generic algorithm described in Fig. 1 will obtain an ϵ-approximate solution of problems* (P) *and* (D) *in at most $O(q\, n^{\frac{p+q}{q(p+1)}} \log\frac{n}{\epsilon})$ iterations.*

Proof It is well known that the number of outer iterations is bounded above by

$$\frac{1}{\theta}\log\frac{n}{\epsilon} \tag{45}$$

[2, Lemma Π.17, p. 116]. By multiplying this number and the upper bound for the number of inner iterations per outer iteration (as given by Lemma 4.1), we get an upper bound for the total number of iterations, namely

$$\frac{60q(p+1)}{\theta}\left(\frac{n\theta+(p+1)\tau+n(p+1)\sqrt{(\frac{\tau}{n})^2+\frac{2\tau}{n}}}{(p+1)(1-\theta)^{\frac{p+1}{2}}}\right)^{\frac{p+q}{q(p+1)}}\log\frac{n}{\epsilon}. \tag{46}$$

Using that $\theta = \Theta(1)$, and $\tau = O(n)$, some elementary transformations reduces this iteration bound to the bound in the theorem. □

The obtained complexity result contains several previously known complexity results as special cases. When $p = 1$ and $q > 1$, the kernel function $\psi(t)$ becomes the prototype self-regular function. If in addition, $q = \log n$ the iteration bound reduces to the best known bound for self-regular functions, $O(\sqrt{n}\log n \log \frac{n}{\epsilon})$. Letting $q = 1$ and $p = 1$, the iteration bound becomes $O(n \log \frac{n}{\epsilon})$ and $\psi(t)$ represents the classical logarithmic kernel function. For $q = 2$ and $p = 0$, the iteration bound is the same as the one obtained in [6] for the simple kernel function $\psi(t) = t - \frac{1}{t} - 2$.

4.2 Iteration Bound for the Small-Update Method

It is not hard to show that if the above analysis were used for small-update methods the iteration bound would not be as good as it can be for these types of methods. We need to use (29) in Lemma 3.1 to get the improved iteration bound, albeit that this only holds for $q \geq 2 - p$. We then have

$$\begin{aligned}\Psi_0 &\leq n\psi\left(\frac{1+\sqrt{\frac{\tau}{n}+\frac{\tau^2}{n^2}+\frac{\tau}{n}\sqrt{\frac{\tau^2}{n^2}+\frac{2\tau}{n}}}}{\sqrt{1-\theta}}\right)\\ &\leq \frac{(p+q)n}{2}\left(\frac{1+\sqrt{\frac{\tau}{n}+\frac{\tau^2}{n^2}+\frac{\tau}{n}\sqrt{\frac{\tau^2}{n^2}+\frac{2\tau}{n}}}}{\sqrt{1-\theta}}-1\right)^2,\end{aligned}$$

where the second inequality is due to Lemma 2.2.

Using $1-\sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \leq \theta$, the above inequality can be simplified to

$$\Psi_0 \leq \frac{p+q}{2(1-\theta)}\left(\theta\sqrt{n}+\sqrt{\tau+\frac{\tau^2}{n}+\tau\sqrt{\frac{\tau^2}{n^2}+\frac{2\tau}{n}}}\right)^2. \tag{47}$$

Following the same line of arguments as in the proof of Lemma 4.1, we obtain the following lemma.

Lemma 4.2 *If $q \geq 2 - p$, the following inequalities hold*:

$$K \leq \frac{60q(p+q)}{(1-\theta)}\left(\theta\sqrt{n}+\sqrt{\tau+\frac{\tau^2}{n}+\tau\sqrt{\frac{\tau^2}{n^2}+\frac{2\tau}{n}}}\right)^{\frac{2(p+q)}{q(p+1)}}. \tag{48}$$

Given the upper bound on the number of outer iterations (45), as mentioned in the previous section, the upper bound on the total number of iterations is

$$\frac{60q(p+q)}{\theta(1-\theta)}\left(\theta\sqrt{n}+\sqrt{\tau+\frac{\tau^2}{n}+\tau\sqrt{\frac{\tau^2}{n^2}+\frac{2\tau}{n}}}\right)^{\frac{2(p+q)}{q(p+1)}}\log\frac{n}{\epsilon}. \tag{49}$$

For small-update methods we have $\theta=\Theta(\frac{1}{\sqrt{n}})$ and $\tau=O(1)$. After some elementary reductions one easily obtains that the iteration bound is $O(q^2\sqrt{n}\log\frac{n}{\varepsilon})$. We summarize this result in the theorem below.

Theorem 4.2 *Given that* $\theta=\Theta(\frac{1}{\sqrt{n}})$ *and* $\tau=O(1)$, *which are characteristics of the small-update methods, the generic algorithm described in the Fig.* 1 *will obtain* ϵ*-approximate solutions of* (P) *and* (D) *in at most* $O(q^2\sqrt{n}\log\frac{n}{\varepsilon})$ *iterations.*

5 Conclusions

In this paper we have analyzed large-update and small-update versions of the primal-dual interior-point algorithm described in Fig. 1 that are based on the new class of kernel functions (14). This class is fairly general and includes the classical logarithmic, the prototype self-regular and non-self-regular kernel functions as special cases.

The analysis of the algorithm in the paper follows the same line of arguments as in [1], where a variety of non-self-regular kernel functions were considered, including the ones with linear and quadratic growth terms. However, the important case, when the growth term is between linear and quadratic, was not considered. In this paper, we have introduced such a class of kernel functions and showed that IPMs based on it have very good iteration bounds.

In order to achieve these complexity results several new arguments had to be developed regarding the new class of kernel functions, most notable Lemma 2.7, Lemma 2.5, and Corollary 2.1. Lemma 2.7 is essential in obtaining tight upper bounds on the barrier function at the start of the outer iteration stated in Lemma 3.1. Lemma 2.5 is important in obtaining a suitable lower bound on the stepsize stated in the Theorem 3.2. Together with Corollary 2.1 it leads to the good upper bound on the decrease of the barrier function during the inner iteration stated in the Theorem 3.3.

Combination of these results enabled us to obtain very good complexity results for both small-update and large-update IPMs. Moreover, our iteration bounds match the currently best known bounds. Examples are prototype self-regular function with a quadratic growth term (see [2]), simple non-self-regular function with a linear growth term (see [5]), and the classical logarithmic kernel function, (see [1] for an overview).

The above results do not necessarily imply that the numerical behavior of the method will be competitive with existing methods. Computational studies have to clarify this matter, and will be the subject of future research.

Appendix: Technical Lemmas

We list three simple technical lemmas because they are used in the complexity analysis of the algorithms. The original proofs of these lemmas can be found in [5–8], respectively.

Lemma A.1 (Lemma 20 in [5]) *If $\alpha \in [0,1]$ and $t \geq -1$, then $(1+t)^\alpha \leq 1+\alpha t$.*

Lemma A.2 (Lemma 12 in [7]) *Let $h(t)$ be a twice differentiable convex function with $h(0)=0$, $h'(0)<0$ and let $h(t)$ attain its (global) minimum at $t^*>0$. If $h''(t)$ is monotonically increasing for $t \in [0,t^*]$, then*

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

Lemma A.3 (Proposition 2.2 in [8]) *Let $t_0, t_1, \ldots, t_K$ be a sequence of positive numbers such that*

$$t_{k+1} \leq t_k - \beta t_k^{1-\gamma}, \quad k = 0, 1, \ldots, K-1,$$

where $\beta > 0$ and $0 < \gamma \leq 1$. Then $K \leq \lfloor \frac{t_0^\gamma}{\beta\gamma} \rfloor$.

References

1. Bai, Y.Q., Ghami, M.E., Roos, C.: A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. SIAM J. Optim. **15**, 101–128 (2004)
2. Roos, C., Terlaky, T., Vial, J.-P.: Interior-Point Methods for Linear Optimization. Springer, New York (2006), 2nd edn. Theory and Algorithms for Linear Optimization Wiley, New York (1997)
3. Sonnevend, G.: An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In: Prékopa, A., Szelezsán, J., Strazicky, B. (eds.) System Modeling and Optimization. Proceedings of the 12th IFIP-Conference, Budapest, Hungary, September 1985. Lecture Notes in Control and Information Sciences, vol. 84, pp. 866–876. Springer, Berlin (1986)
4. Megiddo, N.: Pathways to the optimal set in linear programming. In: Megiddo, N. (ed.) Progress in Mathematical Programming: Interior Point and Related Methods, pp. 131–158. Springer, New York (1989). Identical version in: Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan, pp. 1–35 (1986)
5. Bai, Y.Q., Ghami, M.E., Roos, C.: A new efficient large-update primal-dual interior-point method based on a finite barrier. SIAM J. Optim. **13**, 766–782 (2003)
6. Bai, Y.Q., Roos, C., Ghami, M.E.: A primal-dual interior-point method for linear optimization based on a new proximity function. Optim. Methods Softw. **17**, 985–1008 (2002)
7. Peng, J., Roos, C., Terlaky, T.: Self-regular functions and new search directions for linear and semidefinite optimization. Math. Program. **93**, 129–171 (2002)
8. Peng, J., Roos, C., Terlaky, T.: Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms. Princeton University Press, Princeton (2002)
9. Bai, Y.Q., Roos, C.: A polynomial-time algorithm for linear optimization based on a new simple kernel function. Optim. Methods Softw. **18**, 631–646 (2003)