Iranian Journal of Operations Research
Vol. 3, No. 2, 2012, pp. 1-23

Local Self-concordance of Barrier Functions Based on Kernel-functions

Y.Q. Bai[1], G. Lesaja[2], H. Mansouri[3], C. Roos[*,4], M. Zangiabadi[5]

Many efficient interior-point methods (IPMs) are based on the use of a self-concordant barrier function for the domain of the problem that has to be solved. Recently, a wide class of new barrier functions has been introduced in which the functions are not self-concordant, but despite this fact give rise to efficient IPMs. Here, we introduce the notion of locally self-concordant barrier functions and we prove that the new barrier functions are locally self-concordant. In many cases, the (local) complexity numbers of the new barrier functions along the central path are better than the complexity number of the logarithmic barrier function by a factor between 0.5 and 1.

Keywords: *Linear optimization, Self-dual embedding, Primal-dual interior-point method, Self-concordance, Kernel function, Polynomial complexity.*

AMS Subject Classification: **90C05, 90C51.**

Manuscript received on 30/6/2012 and accepted for publication on 31/7/2012.†

1. Introduction

Since the seminal work of Karmarkar [11] in 1984, the field of Interior-Point Methods (IPMs) has been one of the most active areas of research in optimization. IPMs are among the most effective methods for solving linear optimization (*LO*) problems and wide classes of more general convex optimization problems. They enjoy a polynomial-time theoretical complexity and behave very well in practice.

Another milestone in the development of IPMs is the development of the theory of *self-concordant* (SC) functions and SC barrier functions (SCBFs), introduced by Nesterov and Nemirovski [17] in the early 90 s of the past century. This theory was a breakthrough that provided a unified framework for the design and analysis of IPMs for a large class of important convex optimizations problems, yielding the best known polynomial complexity results.

†Invited paper.

*Corresponding Author.
[1] Department of Mathematical Sciences, Shanghai University, Shanghai, China.
[2] Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460-8093, U.S.A.
[3] Department of Applied Mathematics, Faculty of Mathematical Sciences, Shahrekord University, P. O. Box 115, Shahrekord, Iran.
[4] Department of Software Technology and Algorithms, Delft University of Technology, P. O. Box 5031, 2600, Delft, The Netherlands.
E-mail: c.roos@tudelft.nl
[5] Department of Applied Mathematics, Faculty of Mathematical Sciences, Shahrekord University, P. O. Box 115, Shahrekord, Iran

For a survey of the theory and practice of IPMs we refer the reader to recent books related to the subject Boyd and Vandenberghe [4], Nesterov[16], Nesterov and Nemirovski [17], Renegar [24], Roos et al. [25], Terlaky [28], Vanderbei [29], Wright [30], and Ye [31].
Many IPMs are so-called path-following algorithms. They are based on approximately following the so-called *central path* of a problem as a guide to its optimal set. The 'engine' in such methods is usually some variant of Newton's method. Another key ingredient in the design and analysis of IPMs
is a barrier function for the domain of the problem. Until recently, most IPMs were based on the use of the classical logarithmic barrier function. The best known iteration bound for linear optimization
for LO is attained by a full-Newton step IPM (FNS-IPM) based on the logarithmic barrier function, namely

$$O(\sqrt{n}\log\frac{n}{\varepsilon}). \tag{1}$$

Here, n denotes the number of inequalities in the problem and ε is the desired accuracy. The bound in (1) has been achieved by several authors; e.g., Gonzaga [10], Kojima et al. [13], Mizuno [15], Renegar [23], Roos and Vial [27]. The seminal work of Nesterov and Nemirovski [17] has made clear that the reason for the prominent role of the logarithmic barrier function in LO and convex optimization is that this function is self-concordant.

The special role of the logarithmic barrier function in linear and convex optimization has been questioned many times (see, e.g., Lasserre [14]). Recently, it has become clear that the theoretical complexity of large-update methods, which are the most efficient methods in practice, can be significantly improved by using other barrier functions than the logarithmic barrier function. These other barrier functions are the so-called kernel-function based barrier functions Bai et al. [1], which include the self-regular barrier functions of Peng et al. [20, 21, 22] as a subclass. Although the new barrier functions were designed for large-update methods, they can also be used for small-update methods and, surprisingly enough, it turned out that in many cases they give rise to methods with the iteration bound (1). Since the new functions are (in general) not self-concordant, this was not at all expected.

Here, we aim to get a better understanding of this nice behavior of the new barrier functions. We do this by defining a localized version of self-concordance. Let us recall that a barrier function $f : D \to \mathbb{R}$ is a self-concordant barrier function (SCBF) if some expression in the second and third order directional derivatives in any point x of the domain D is bounded above by a uniform constant (i.e., κ), and similarly, another expression, in the first and second order directional derivatives, in any point x of the domain D is bounded above by a uniform constant (i.e., ν). We localize this definition by saying that f is a locally SCBF (LSCBF) if these expressions are bounded above in each point $x \in D$ by (local) parameters $\kappa(x)$ and $\nu(x)$, respectively. The main result of our work here is that the new barrier functions are LSC. Moreover, along the central path the parameters $\kappa(x)$ and $\nu(x)$ are indeed constant, and finally, the so-called complexity number, which is an important measure for the efficiency of IPMs, is in many cases smaller than for the logarithmic barrier function.

The remainder of our work is organized as follows. In Section 2, we start with a preparatory section in which we recall known material from the theory of SCBFs. We deal with the problem of minimizing a linear function over a closed convex set with nonempty (relative) interior. The formal definition of an SCBF is given and we outline the related FNS-IPM. Then, we briefly describe how the iteration bound for the algorithm is obtained, showing that this bound depends linearly on the complexity number of the SCBF. We restrict ourselves here to LO problems, although the extension to the more general class of symmetric optimization problems is nowadays more or less straightforward. Every such problem can be embedded in a self-dual problem; a strictly complementary solution of this self-dual problem can be used to construct an optimal solution of the original problem. In Section 3, we describe the self-dual LO problem and some of its properties, including the definition of its central path and the availability of a point on the central path that can be used to start any IPM. In Section 4, we introduce new barrier functions for the self-dual problem. These functions are based on a special class of kernel functions, called b-kernel functions; these are defined by three simple conditions on the signs of the first three derivatives of the barrier term. We show that these functions are strictly convex and give some examples of such functions from the literature. Section 5 is the main part of the paper. We give a formal definition of the local self-concordant parameters $\kappa(x)$ and $\nu(x)$ and the parameters are computed for a generic kernel function. This turns out to be a tedious task, specially for $\kappa(x)$, but with a remarkable outcome. The resulting values are used in Section 6 to compute the local complexity number along the central path for the kernel functions given in Section 4. In Section, 7 we conclude with some final remarks.

Notations:

We briefly mention some notational conventions that are used throughout the paper. The 2-norm of a vector is denoted by $\|.\|$, whereas $\|.\|_\infty$ denotes the infinity norm. If $x,s \in \mathbb{R}^n$, then xs denotes the coordinate-wise (or Hadamard) product of the vectors x and s. Furthermore, e denotes the all-one vector and 0 a zero vector (or matrix) of appropriate size. The nonnegative orthant and positive orthant are denoted by $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$, respectively. If $z \in \mathbb{R}^n_+$ $(\mathbb{R}^n_{++})$, we write $z \geq 0$ $(z > 0)$. As usual, we write $f(x) = O(g(x))$, if there exists a positive constant c such that $f(x) \leq cg(x)$, for all $x \in dom(f)$.

2. Short Introduction to Self-concordant Barrier Functions

We first recall the notion of a self-concordant barrier function (SCBF) $\varphi : D \to \mathbb{R}$, where D is an open convex subset of $\mathbb{R}^n$. We start by considering the case where $n = 1$. Then, φ is a univariate function and its domain D an open interval in $\mathbb{R}^n$. One calls φ a (κ,ν)-SCBF, if there exist nonnegative numbers κ and ν such that

$$\left|\varphi'''(x)\right| \leq 2\kappa\varphi''(x)^{\frac{3}{2}}, \qquad (\varphi'(x))^2 \leq \nu\varphi''(x), \qquad \forall x \in D. \tag{2}$$

Note that this definition assumes that φ is three times differentiable. Moreover, it implies that $\varphi''(x)$ is nonnegative, and hence φ is convex. The most well-known example of an SCBF is $-log\ x$, Which is a (1,1)-SCBF on the positive real axis.

Now, suppose that φ is a multivariate function, i.e., n >1. Then, φ is called a (κ,ν)-SCBF, if its restriction to an arbitrary line that intersects D is a (κ,ν)-SCBF. In other words, φ is a (κ,ν)-SCBF, if (2) holds when we replace all derivatives in (2) by directional derivatives, for every direction $h \in \mathbb{R}^n$. More precisely, denoting these directional derivatives by $\nabla\varphi(x)[h]$, $\nabla^2\varphi(x)[h,h]$ and $\nabla^3\varphi(x)[h,h,h]$, respectively, φ is a (κ,ν)-SCBF if and only if[1]

$$\nabla^3\varphi(x)[h,h,h] \leq 2\kappa\,(\nabla^2\varphi(x)[h,h])^{\frac{3}{2}}, \qquad \forall x \in D, \forall h \in \mathbb{R}^n, \tag{3}$$

$$(\nabla\varphi(x)[h])^2 \leq \nu\nabla^2\varphi(x)[h,h], \qquad \forall x \in D, \forall h \in \mathbb{R}^n. \tag{4}$$

[1] There is no need to take the absolute value of the first expression in (3), because $\nabla^3\varphi(x)[h,h,h]$ changes sign, if we replace h by $-h$.

The importance of this notion becomes clear when considering the problem of minimizing a linear function $c^T x$ over the closure $\overline{\mathcal{D}}$ of D. So, we consider the following problem:

$$(P) \quad min\,\{c^T x : x \in \overline{\mathcal{D}}\}.$$

For $\mu > 0$, we define

$$\varphi_\mu(x) := \frac{c^T x}{\mu} + \varphi(x), \qquad x \in D.$$

If μ is fixed, then $\varphi_\mu(x)$ and $\varphi\,(x)$ differ only by a linear function; therefore, all their second and third order directional derivatives are the same. It follows that $\varphi_\mu(x)$ satisfies (3). As is known from the theory of self-concordant functions, this implies that Newton's method will be very efficient when using it to obtain a minimizer of $\varphi_\mu(x)$, provided that we have a starting point that is close enough to the minimizer, which is denoted by $x(\mu)$ [2].Before stating the related result, we need to discuss how this 'closeness' is measured. For this, we use the norm of the Newton step with respect to the Hessian matrix of $\varphi_\mu(x)$ at x. For each $x \in D$, the Hessian matrix of φ at x is denoted as $H(x)$. So, $H(x) = \nabla^2\varphi_\mu(x) = \nabla^2\varphi\,(x)$. Under a mild assumption, namely that D does not contain a straight line, $H(x)$ is positive definite, and hence defines a norm; see Nesterov [16, Theorem 4.1.3]. Denoting the gradient of $\varphi_\mu(x)$ at x by $g_\mu(x)$, we have

$$g_\mu(x) := \nabla\varphi_\mu(x) = \frac{c}{\mu} + \nabla\varphi\,(x) = \frac{c}{\mu} + g(x),$$

where $g(x)$ denotes the gradient of $\varphi\,(x)$ at x. Then, the Newton step at x is given by

$$\Delta x = -H(x)^{-1} g_\mu(x),$$

and our measure for the distance of x to $x(\mu)$ is given by

$$\lambda_\mu(x) := \|\Delta x\|_{H(x)} = \sqrt{\Delta x^T H(x) \Delta x} = \sqrt{g_\mu(x)^T H(x)^{-1} g_\mu(x)} = \|g_\mu(x)\|_{H(x)^{-1}}.$$

The following lemma very nicely quantifies the behavior of a full Newton step. For its proof, we refer to the existing literature Glineur [9], Nesterov [16], Nesterov and Nemirovskii [18].

Lemma 2.1. *If φ satisfies (3) and $\kappa\lambda_\mu(x) < 1$, then*

$$\lambda_\mu(x + \Delta x) \le \kappa\left(\frac{\lambda_\mu(x)}{1 - k\lambda_\mu(x)}\right)^2.$$

A major question is what the effect is on $\lambda_\mu(x)$ when μ is reduced to $\mu^+ = (1-\theta)\mu$, where $\theta \in [0,1)$. If $\lambda := \lambda_\mu(x)$, then we may write

[2] When μ runs though all positive numbers, then $x(\mu)$ runs through the so-called *central path* of (P). When μ approaches 0 then $x(\mu)$ converges to an optimal solution of (P). Therefore, in IPMs the central path is used as a guideline to the set of optimal solutions of (P).

$$g_{\mu^+}(x) = \frac{c}{\mu^+} + \nabla\varphi(x) = \frac{1}{1-\theta}\left(\frac{c}{\mu} + (1-\theta)\nabla\varphi(x)\right) = \frac{1}{1-\theta}\left(g_\mu(x) - \theta\nabla\varphi(x)\right).$$

Hence, denoting $H(x)$ shortly by H, and using the triangle inequality, we get

$$\begin{aligned}\lambda_{\mu^+}(x) &= \frac{1}{1-\theta}\left\|g_\mu(x) - \theta\nabla\varphi(x)\right\|_{H^{-1}}\\ &\le \frac{1}{1-\theta}\left(\left\|g_\mu(x)\right\|_{H^{-1}} + \theta\|g(x)\|_{H^{-1}}\right)\\ &= \frac{1}{1-\theta}\left(\lambda\mu(x) + \theta\|g(x)\|_{H^{-1}}\right)\\ &= \frac{1}{1-\theta}\left(\lambda_\mu(x) + \theta\lambda(x)\right).\end{aligned}$$

At this stage, we need that (4) implies that $\lambda(x) \le \sqrt{v}$ (cf. Glineur [9, Thm. 2.2]). The proof goes as follows. By using (4) with $h = \Delta x$, we have

$$\lambda(x)^2 = \Delta x^T H(x)\Delta x = -\Delta x^T g_\mu(x) = -\nabla\varphi_\mu(x)[\Delta x] \le |\nabla\varphi_\mu(x)[\Delta x]|$$

$$\le \sqrt{v\nabla^2\varphi_\mu(x)[\Delta x, \Delta x]} = \sqrt{v\Delta x^T H(x)\Delta x} = \sqrt{v\lambda(x)^2} = \sqrt{v}\lambda(x).$$

Dividing by $\lambda(x)$, we obtain $\lambda(x) \le \sqrt{v}$, as desired. Hence the following result has been shown.

Lemma 2.2. If φ satisfies (4) and $\mu^+ := (1-\theta)\mu$, then

$$\lambda_{\mu^+}(x) \le \frac{\lambda_\mu(x) + \theta\sqrt{v}}{1-\theta}.$$

The above two lemmas are all we need for the analysis of the simple algorithm described in Algorithm 2.1. Recall that Lemma 2.1 only uses property (3), and Lemma 2.2 only (4). The algorithm starts at a point $x^0 \in D$ that is close enough to a point $x(\mu_0)$ on the central path, and simply repeats doing a μ-update and then a full Newton step, until it gets close enough to an optimal solution of (P). For the purpose of our work here, the following convergence result is of utmost importance. We include the proof, as given in Glineur [8], because it makes clear that the iterates move in a narrow neighborhood of the central path, which is relevant for the rest of the paper.

Theorem 2.3. *If* $\tau = \frac{1}{9\kappa}$ and $\theta = \frac{5}{9+36\kappa\sqrt{v}}$, then the number of iterations of the algorithm is at most

$$\left\lceil 2 + (1 + 4\kappa\sqrt{v})\ln\frac{2v\mu_0}{\varepsilon}\right\rceil.$$

Proof. At the start of the first iteration we have $x \in D$ and $\mu = \mu_0$ such that $\lambda_\mu(x) \le \tau$. When the barrier parameter is updated to $\mu^+ = (1-\theta)\mu$, Lemma 2.2 gives

$$\lambda_{\mu^+}(x) \le \frac{\lambda_\mu(x) + \theta\sqrt{v}}{1-\theta} \le \frac{\tau + \theta\sqrt{v}}{1-\theta}. \tag{5}$$

Algorithm 2.1: Full Newton Step Interior-Point Method

Input:
accuracy parameter $\varepsilon \in (0,1)$;
proximity parameter $0 < \tau < \frac{1}{\kappa}$;
update parameter $\theta \in (0,1)$;
$x^0 \in D$ and $\mu_0 > 0$ such that $\lambda_{\mu^0}(x^0) \leq \tau$.
begin
$x := x^0$, $\mu := \mu_0$;
while $\nu\mu \geq \varepsilon$ *do*
μ-update: $\mu := (1-\theta)\mu$;
Newton's step: $x := x + \Delta x$;
endwhile
end

Then, after the Newton step, the new iterate is $x^+ = x + \Delta x$ and, by Lemma 2.1,

$$\lambda_{\mu^+}(x^+) \leq \kappa \left(\frac{\lambda_{\mu^+}(x)}{1-\kappa\lambda_{\mu^+}(x)} \right)^2$$

The algorithm is well defined if we choose τ and θ such that $\lambda_{\mu^+}(x^+) \leq \tau$. To get the lowest iteration bound, we need at the same time to maximize θ. From (6), we deduce that $\lambda_{\mu^+}(x^+) \leq \tau$ certainly holds, if

$$\frac{\lambda_{\mu^+}(x)}{1-\kappa\lambda_{\mu^+}(x)} \leq \frac{\sqrt{\tau}}{\sqrt{\kappa}}, \tag{6}$$

which is equivalent to

$$\lambda_{\mu^+}(x^+) \leq \frac{\sqrt{\tau}}{\kappa\sqrt{\tau}+\sqrt{\kappa}}.$$

According to (5), this will hold, if $\frac{\tau+\theta\sqrt{\nu}}{1-\theta} \leq \frac{\sqrt{\tau}}{\kappa\sqrt{\tau}+\sqrt{\kappa}}$. This leads to the following condition on θ:

$$\theta \leq \sqrt{\tau}\,\frac{1-\kappa\tau-\sqrt{\kappa\tau}}{\sqrt{\tau}+\sqrt{\nu\kappa}\left(1+\sqrt{\kappa\tau}\right)}.$$

If $\tau = \frac{1}{9\kappa}$, then this upper bound for θ gets the value $\frac{5}{9+36\kappa\sqrt{\nu}} \geq \frac{1}{2+8\kappa\sqrt{\nu}}$. This clarifies the choice of the values of τ and θ in the theorem. It follows that after each iteration the property $\lambda_\mu(x) \leq \tau$ now easily be obtained as follows. Since after the kth iteration we have $\mu = \mu_0(1-\theta)^k$, the algorithm stops if k is such that $\nu\mu_0(1-\theta)^k \leq \varepsilon$. One easily verifies that this certainly holds, if

$$k \geq \frac{1}{\theta} \ln \frac{\nu\mu_0}{\varepsilon}.$$

So, the number of iterations is at most

$$\left\lceil \frac{1}{\theta} ln \frac{\nu\mu_0}{\varepsilon} \right\rceil. \tag{7}$$

This completes the proof. □

Note that the order of magnitude of the iteration bound in Theorem 2.3 depends linearly on the quantity $\kappa\sqrt{\nu}$. Following Glineur [8], we call this the complexity number of φ, and denote it as γ. It may be worth pointing out that one easily deduces from the above proof that after the μ-update, we always have

$$\lambda_{\mu^+}(x) \leq \frac{1}{4k}. \tag{8}$$

Since during the Newton step the proximity value $\lambda_\mu(x)$ decreases, we may conclude that during the course of the algorithm $\lambda_\mu(x)$ never exceeds the value $\frac{1}{4k}$. It means that the iterates move in a narrow neighborhood of the central path to the set of optimal solutions. The results of this section are valid for every convex problem whose domain has an SCBF. In the rest of the paper, we restrict ourselves to linear optimization problems. As we argue in the next section, it suffices to deal with a specific self-dual linear optimization problem.

3. Self-dual LO Problem

In this section, we recall the fact that a solution of any LO problem can be obtained by embedding the given problem and its dual problem in a problem of the form

$$min\{q^T x: Mx \geq -q, x \geq 0\} \tag{SP}$$

where M is a skew-symmetric matrix (i.e., $M^T = -M$) of size $n \times n$ and the vector q is

$$q = (0_{n-1}; n). \tag{9}$$

One easily verifies that the dual problem of (SP) has the same feasible region, whereas it maximizes $-q^T x$. This means that it is essentially the same problem (SP). We therefore say that (SP) is self-dual. Of course, n and M depend on the LO problem that we want to solve. We assume $n \geq 2$. For details, see Roos et al. [26, Part I], where it is also made clear that we may assume that the all-one-vector e is feasible for (SP) and, moreover, $Me + q = e$. In other words, defining

$$s(x) := Mx + q, \qquad x \in \mathbb{R}^n,$$

we have

$$s(e) = e. \tag{10}$$

Note that x is a feasible solution of (SP) if and only if $x \geq 0$ and $s(x) \geq 0$, and also that

$$q^T x = (s(x) - Mx))^T x = s(x)^T x - x^T Mx = x^T s(x), \tag{11}$$

where we used that $x^T Mx = 0$, which holds since M is skew-symmetric. In order to obtain an optimal solution for the given LO problem, we need a strictly complementary solution of (SP), i.e., a feasible solution x such that

$$xs(x) = 0, \ x + s(x) > 0.$$

Yet, we describe how such a solution can be obtained. For details we refer again to Roos et al. [26, Part I]. Due to (10), the system

$$s = Mx + q, x > 0, s > 0, \quad (12)$$
$$xs = \mu e \quad (13)$$

admits the solution $x = e, \ s = e, \ \mu = 1$. This means that (SP) has a strictly feasible point, i.e., (SP) satisfies the interior-point condition (IPC). As is well-known, this implies that the above system has a unique solution for every μ>0. The x-part of the solution is precisely the point $x(\mu)$ on the central path (SP) with respect to the logarithmic barrier function. Moreover, if μ approaches zero, then $x(\mu)$ converges to a strictly complementary solution of (SP).

As we saw in the previous section, an IPM approximately follows the central path until the objective value is small enough. As a starting point, we can use $x = e$, which is the μ-center for μ= 1. Let us mention that it has been shown in Roos et al. [26, Section 3.3] that if the objective value of the final iterate is small enough and this iterate is close enough to the central path, then there exists a simple rounding procedure that yields a strictly complementary solution of (SP) in polynomial time.

4. Kernel-function Based Barrier Functions for (SP)

In this section, we introduce a wide class of new barrier functions for the (interior of the) domain of (SP), which is given by

$$\mathcal{P}^0 := \{x : x > 0, \ s(x) > 0\}.$$

Let $x \in \mathcal{P}^0$. So, x is strictly feasible for (SP). For any $\mu > 0$ we define the vector $v(x, \mu)$ as follows:

$$v(x, \mu) := \sqrt{\frac{xs(x)}{\mu}}. \quad (14)$$

We call $v(x, \mu)$ the variance vector of x with respect to μ. In the sequel, the values for x and μ often follow from the context. In such cases, when this gives no rise to confusion, we feel free to omit the arguments x and/or μ in $s(x)$ and $v(x, \mu)$.
From (11), we derive that

$$q^T x = x^T s(x) = \sum_{i=1}^n x_i s_i(x) = \mu \sum_{i=1}^n v_i^2 = \mu ||v||^2. \quad (15)$$

Let $\psi : (0, \infty) \to [0, \infty)$ be a three times differentiable strictly convex univariate function such that $\Psi(t)$ is minimal at $t = 1$, and $\Psi(1) = 0$, whereas $\Psi(t)$ goes to infinity if t approaches zero.
For any $\mu > 0$, we then define the function $\mu(x)$ as follows:

---

[1] Note that $s(0) = q \geq 0$, meaning that the zero vector is feasible for (SP). Since the objective value of any feasible solution x is simply nx_n, which is nonnegative, it follows that the zero vector is an optimal solution of (SP). But this solution is not strictly complementary, since $0 + s(0) = q$, and q has $n - 1$ zero entries.

$$\Psi_{\mu}(x) := 2\,\mu \sum_{i=1}^{n} \psi(v_i)\,, \qquad x\ \in \mathcal{P}^0 \tag{16}$$

with v as defined in (14). Note that if $x \in \mathcal{P},^0$ then $xs(x) > 0$, and hence $v > 0$, which makes clear that $\Psi_{\mu}(x)$is well-defined. Moreover, if x approaches the boundary of $\mathcal{P}$, then at least one of the coordinates of x or $s(x)$ goes to zero and hence one of the entries of v approaches zero, which means that $\Psi_{\mu}(x)$ goes to infinity. This makes clear that Ψ_{μ} is a barrier function for the domain of (SP). We call Ψ_{μ} a kernel-based barrier function, and ψ its kernel function.

Also, note that if x is the μ-center, for some μ>0, i.e., if $x = x(\mu)$, then $xs(x) = \mu e$, whence $v = e$. Hence, if x is a μ-center, then $\psi(v_i) = \psi(1) = 0$, for each i, which implies $\Psi_{\mu}(x) = 0$. On the other hand, since $\psi(t)$ is minimal at $t = 1$ and $\psi(1) = 0$, we have $\psi(t) \geq 0$, for every $t > 0$. It thus follows from (16) that $\Psi_{\mu}(x) \geq 0,$ for every $x \in \mathcal{P}^0$. This makes clear that $x(\mu)$ is a minimizer of $\Psi_{\mu}(x)$. As we show below (Theorem 4.1), $\Psi_{\mu}(x) \geq 0$ is strictly convex. Hence, this minimizer is unique. As a consequence, we have $x = x(\mu)$ if and only if $v = e$.

Indeed, an important question is whether $\Psi_{\mu}(x)$ is strictly convex in x. To address this, we need to make some assumptions on the barrier term $\psi_b(t)$ of $\psi(t)$, which is defined by the following relation:

$$\psi(t) \equiv (t^2 - 1) + \psi_b\,(t),\ t > 0. \tag{17}$$

Obviously, $\psi_{\mathrm{b}}(t)$ dominates the behavior of $\psi(t)$ when t approaches zero. We assume that the following conditions are satisfied:

$$\psi'_b(t) < 0, \qquad t > 0 \tag{18}$$

$$\psi''_b(t) > 0, \qquad t > 0, \tag{19}$$

$$\psi'''_b(t) < 0, \qquad t > 0. \tag{20}$$

A kernel function $\psi(t)$ having these properties is called a *b*-kernel function. We have the following result.

Theorem 4.1. If $\psi(t)$ satisfies (18) and (19), then $\Psi_{\mu}(x)$ is strictly convex in x, for each $\mu > 0$.

Proof. Using (15), which gives $\sum_{i=1}^{n} {v_i}^2 = q^T x \ \ /_{\mu}$, we may write

$$\Psi_{\mu}(x) = 2\sum_{i=1}^{n} \psi(v_i) = \sum_{i=1}^{n} (\,{v_i}^2 - 1 + 2\,\psi_b(v_i) = \frac{q^T x}{\mu} - n + 2\sum_{i=1}^{n} \psi_b(v_i). \tag{21}$$

Since $q^T x$ is linear in x, it follows from (21) that $\Psi_{\mu}(x)$ is strictly convex in x if and only if

$$\Psi(x) := 2\,\sum_{i=1}^{n} \psi_b(v_i)\ - n \tag{22}$$

is strictly convex in x.[4] We show this by proving that the Hessian matrix of $\Psi(x)$, denoted by $H(x)$, is positive definite. We therefore start by computing $H(x)$. One has

$$\frac{\partial \Psi}{\partial x_j} = 2\sum_{i=1}^{n} \psi'_b(v_i)\frac{\partial v_j}{\partial x_j}, \qquad 1 \le j \le n. \tag{23}$$

Using this, we find that the (k, j)-entry of $H(x)$ is given by

$$\frac{\partial^2 \Psi}{\partial x_k \partial x_j} = 2\sum_{i=1}^{n}\left[\psi''_b(v_i)\frac{\partial v_i}{\partial x_k}\frac{\partial v_i}{\partial x_j} + \psi'_b(v_i)\frac{\partial^2 v_i}{\partial x_k \partial x_j}\right], \qquad 1 \le j, k \le n. \tag{24}$$

We proceed by computing the first and second order partial derivatives of v_j. Since $s_i(x) = \sum_{j=1}^{n} M_{ij}x_j + q_j$, one has

$$\frac{\partial s_i(x)}{\partial x_j} = M_{ij}.$$

The definition (14) of the variance vector v at x implies that

$$s_i x_i(x) = \mu v_i^2, \qquad 1 \le i \le n. \tag{25}$$

Taking the derivative with respect to x_j at both sides, we obtain

$$\delta_{ij} s_i(x) + x_i M_{ij} = 2\mu v_i \frac{\partial v_j}{\partial x_j}.$$

Thus we get

$$\frac{\partial v_i}{\partial x_j} = \frac{1}{2\mu v_i}(\delta_{ij} s_i(x) + x_i M_{ij}) = v_i\left(\frac{\delta_{ij}}{2x_i} + \frac{M_{ij}}{2s_i}\right). \tag{26}$$

Using this, it follows that

$$\frac{\partial^2 v_i}{\partial x_k \partial x_j} = \frac{\partial}{\partial x_k}\left(v_i\left(\frac{\delta_{ij}}{2x_i} + \frac{M_{ij}}{2s_i}\right)\right) = v_i\left(\frac{\delta_{ik}}{2x_i} + \frac{M_{ik}}{2s_i}\right)\left(\frac{\delta_{ij}}{2x_i} + \frac{M_{ij}}{2s_i}\right) + v_i\left(\frac{\delta_{ij}\delta_{ik}}{2x_i^2} - \frac{M_{ij}M_{ik}}{2s_i^2}\right)$$

$$= -v_i\left(\frac{\delta_{ik}}{2x_i} - \frac{M_{ik}}{2s_i}\right)\left(\frac{\delta_{ij}}{2x_i} - \frac{M_{ij}}{2s_i}\right).$$

The last equality holds, since $(\alpha_1 + \beta_1)(\alpha_2 + \beta_2) - 2\beta_1\beta_2 = -(\alpha_1 - \beta_1)(\alpha_2 - \beta_2)$, where αi, βi are arbitrary real scalars. Defining the vectors $a_i, b_i \in \mathbb{R}^n$, for $i = 1, \dots, n$, as

$$a_{ij} = \frac{\delta_{ij}}{2x_i} + \frac{M_{ij}}{2s_i}, \quad b_{ij} = \frac{\delta_{ij}}{2x_i} - \frac{M_{ij}}{2s_i}, \qquad 1 \le j \le n, \tag{27}$$

we thus have shown that

[4] Note that $\Psi(x)$ can be considered as a barrier function for (SP).

$$\frac{\partial v_i}{\partial x_j} = v_i a_{ij}, \quad \frac{\partial^2 v_i}{\partial x_k \partial x_j} = -v_i b_{ik} b_{ij}. \tag{28}$$

Now, one easily verifies that the gradient of $\psi_\mu(x)$ at x equals $\frac{q}{\mu}$ plus the gradient of $\psi(x)$. Using (26), this gives

$$\nabla\Psi_\mu(x) = \frac{q}{\mu} + 2\sum_{i=1}^{n} \psi'_b(v_i) v_i\, a_i. \tag{29}$$

Substituting (28) into (24), we obtain

$$H(x) = 2\sum_{i=1}^{n} \left[\psi''_b(v_i) v^2{}_i\, a_i a_i{}^T + \psi'_b(v_i) v_i b_i b_i{}^T\right]. \tag{30}$$

Of course, the matrices $a_i\, a_i^{\mathrm{T}}$ and $b_i\, b_i^{\mathrm{T}}$ are positive semidefinite. Moreover, due to the assumptions (18) and (19), the coefficients $\psi''_b(v_i)v_i^2$ and $\psi'_b(v_i)v_i$ are positive. Therefore, we may already conclude that $H(x)$ is positive semidefinite.
To complete the proof, let $z \in \mathbb{R}^{\mathrm{n}}$ and $z^{\mathrm{T}}H(x)z = 0$. Then, it follows from (30) that

$$\sum_{i=1}^{n} \left[\psi''_b(v_i) v_i{}^2 (a_i{}^T z)^2 + (-\psi'_b(v_i) v_i)\left(b_i{}^T z\right)^2\right] = 0.,$$

Since all terms in the above sum are nonnegative, it follows that each term vanishes. Thus, it follows that $a_i^T z = 0$ and $b_i^T z = 0$, for $i = 1, \dots, n$. This implies $(a_i + b_i)^T z = 0$, for each i. Due to the definitions of the vectors a_i and b_i, the vector $a_i + b_i$ equals e_i / x_i, where e_i is the ith standard unit vector. So, we have also $(a_i + b_i)^T z = z_i / x_i$ Since $x_i > 0$, it follows that $z_i = 0$, for each i, proving that $z = 0$. This completes the proof. □

We conclude this section by presenting some examples of *b*-kernel functions in Figure 1.

These include, apart from the logarithmic kernel function $\psi_1(t)$ (in the first line of the table), some so-called self-regular kernel functions proposed by Peng et al. [19, 21, 22] and some kernel functions proposed by Bai et al. [1, 2, 3], ElGhami et al. [5]. These kernel functions were used to significantly improve the iteration bound of large-update IPMs. However, the conditions that were imposed on the kernel functions in these references are much more technical than the conditions (18)–(20), which are simple and natural conditions on the signs of the derivatives of the barrier term.
Up till now the analysis of algorithms based on kernel-function based barrier functions is quite tedious. On the other hand, Bai et al. [1] presented a scheme which simplifies this analysis, but, despite this, the average length of papers that present a new kernel function is still 15 to 20 pages.

Therefore, it would be a nice achievement if we could find a way to further simplify and unify the analysis. In fact, if kernel-function based barrier functions were SCBFs, then this would be the solution. But as we will see, this the case only for the logarithmic kernel function (line 1 in Fig. 1). In the rest of this paper, we will show that the notion of local self-concordance might be a promising first step in this respect.

i	kernel function $\psi_i\ (t)$	reference
1	$\frac{t^2-1}{2} - \log t$	Bai et al. [1], Frisch [6, 7]
2	$\frac{t^2-1}{2} + \frac{1}{t} - 1$	Bai et al. [3]
3	$\frac{1}{2}\left(t - \frac{1}{t}\right)^2$	Bai et al. [1], Peng et al. [19]
4	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1},\ q > 1$	Bai et al. [1], Peng et al. [21, 22]
5	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q(q-1)} - \frac{q-1}{q}(t-1),\ q > 1$	Bai et al. [1], Peng et al. [21, 22]
6	$\frac{t^2-1}{2} + e^{\frac{1}{t}-1} - 1$	Bai et al. [1]
7	$\frac{t^2-1}{2} + \frac{e^{q\left(\frac{1}{t}-1\right)}}{q} - 1,\ q > 0$	
8	$\frac{t^2-1}{2} - \int_1^t e^{q\left(\frac{1}{\xi}-1\right)}\, d\xi,\ q > 1$	Bai et al. [2]
9	$\frac{t^2-1}{2} + \frac{6}{\pi}\cot\frac{3\pi t}{4t+2}$	Bai et al. [1], ElGhami et al. [5]
10	$\frac{t^2-1}{2} + e^{e^{\frac{4\cot\frac{\pi t}{t+1}}{\pi}}-1} - 1$	Kheirfam [12]

Figure 1: Examples of b-kernel functions

5. Local Self-concordance

We say that a barrier function $\phi : D \to \mathbb{R}$ is locally self-concordant (LSC) at $x \in D$, if there exist nonnegative (and finite!) numbers κ and ν such that

$$\nabla^3\phi(x)[h, h, h] \leq 2\kappa\, (\nabla^2 \phi\, x)\, [h, h])^{3/2}, \quad \forall\, h \in \mathbb{R}^n \tag{31}$$

$$(\nabla\, \phi(x)[h])^2 \leq \nu \nabla^2\phi(x)[h, h], \qquad \forall\, h \in \mathbb{R}^n \tag{32}$$

Note that if ϕ is locally (κ,ν)-self-concordant at x, then it is also locally (κ',ν')-self-concordant at $x \in D$, for all $\kappa' \geq \kappa$ and $\nu' \geq \nu$. Given $x \in$ D , the smallest possible values for κ and ν are

denoted as $\kappa(x)$ and $\nu(x)$, respectively. These values are given by

$$\nu(x) := \max_{h\in\mathbb{R}^n} \frac{(\nabla\phi(x)[\,h]\,)^2}{\nabla^2\phi(x)[\,h,h]}, \qquad x \in D, \tag{33}$$

$$\kappa(x) := \max_{h\in\mathbb{R}^n} \frac{\nabla^3\phi(x)[\,h,h,h]}{2(\nabla^2\phi(x)[h,h])^{\frac{3}{2}}}, \qquad x \in D. \tag{34}$$

Now, it is obvious that if ν and κ are such that $\kappa(x) \le \kappa$ and $\nu(x) \le \nu$, for all $x \in$ D, then ϕ is an SCBF. In the next subsections, we will compute the values of $\nu(x)$ and $\kappa(x)$ for barrier functions based on b-kernel functions. This will make clear that these barrier functions are LSC, and that these functions are not self-concordant, with only the logarithmic barrier function as an exception. We will also compute the complexity number on the central path. As a preparation for these computations, the next subsection is devoted to the computation of the first three directional derivatives of $\psi(x)$.

5.1. Computation of the First Three Directional Derivatives

From (28) we deduce that the first and second directional derivatives of v_i in the direction h are given by

$$\nabla v_i(x)[h] = v_i a_i^T h, \qquad \nabla^2 v_i(x)[h,h] = -v_i(b_i^T h)^2, \tag{35}$$

where the vectors a_i and b_i are as defined in (27). We also need the directional derivatives of $a_i^T\, h$ and $b_i^T\, h$ in the direction h. Since

$$a_i^T\, h = \sum_{k=1}^n \left(\frac{\delta_{i\kappa}}{2x_i} + \frac{M_{i\kappa}}{2s_i}\right) h_k, \qquad b_i^T\, h = \sum_{k=1}^n \left(\frac{\delta_{i\kappa}}{2x_i} - \frac{M_{i\kappa}}{2s_i}\right) h_k, \tag{36}$$

one has

$$\begin{aligned}\nabla a_i^T\, h[h] &= \sum_{j=1}^n \frac{\partial\, a_i^T h}{\partial x_j}\, h_j = \sum_{j=1}^n \sum_{k=1}^n \left(-\frac{\delta_{i\kappa}\delta_{ij}}{2x_i^2} - \frac{M_{i\kappa}M_{ij}}{2s_i^2}\right) h_\kappa\, h_j \\ &= -\sum_{j=1}^n \sum_{k=1}^n [a_{ij}\, a_{ik} + b_{ij}\, b_{ik}\,]\; h_\kappa\, h_j = -(a_i^T\, h)^2 - (b_i^T\, h)^2\end{aligned} \tag{37}$$

and

$$\begin{aligned}\nabla b_i^T\, h[h] &= \sum_{j=1}^n \frac{\partial\, b_i^T h}{\partial x_j}\, h_j = \sum_{j=1}^n \sum_{k=1}^n \left(-\frac{\delta_{i\kappa}\delta_{ij}}{2x_i^2} + \frac{M_{i\kappa}M_{ij}}{2s_i^2}\right) h_\kappa\, h_j \\ &= -\sum_{j=1}^n \sum_{k=1}^n [a_{ij}\, a_{ik} + b_{ij}\, b_{ik}\,]\; h_\kappa\, h_j = -2(a_i^T\, h)(b_i^T\, h),\end{aligned} \tag{38}$$

where the last equality is due to the identity $(\alpha_1+\beta_1)(\alpha_2-\beta_2) + (\alpha_1 - \beta_1)(\alpha_2 + \beta_2) = 2\alpha_1\alpha_2 - 2\beta_1\beta_2$, with α_i, β_i being arbitrary real scalars. Yet, we are ready to compute the first three directional derivatives of $\psi(x)$ as given by (22). Due to (35), we have

$$\nabla\psi(x)[h] = 2\sum_{i=0}^n \psi_b'\,(v_i) v_i a_i^T\, h.$$

Also, using (37), we obtain

$$\nabla^2 \Psi(x)[h,h]=2\sum_{i=1}^{n}[\psi_b''(v_i){v_i}^2(a_i^T h)^2+\psi_b'(v_i)v_i(a_i^T h)^2-\psi_b'(v_i)v_i((a_i^T h)^2+(b_i^T h)^2)]$$

$$=2\sum_{i=1}^{n}[\psi_b''(v_i){v_i}^2(a_i^T h)^2-\psi_b'(v_i)v_i(b_i^T h)^2.$$

Finally, also using (38), we get

$$\nabla^3 \Psi(x)[h,h,h]$$

$$=2\sum_{i=1}^{n}[\psi_b'''(v_i){v_i}^3(a_i^T h)^3+2\psi_b''(v_i){v_i}^2(a_i^T h)^3\text{-}2{\psi''}_b(v_i){v_i}^2(a_i^T h)((a_i^T h)^2+(b_i^T h)^2)]+$$
$$2\sum_{i=1}^{n}[-\psi_b''(v_i){v_i}^2(a_i^T h)(b_i^T h)^2-\psi_b'(v_i)v_i(a_i^T h)(b_i^T h)^2+4\psi_b'(v_i)v_i(a_i^T h)(b_i^T h)^2]$$

$$=2\sum_{i=1}^{n}[\psi_b'''(v_i){v_i}^3(a_i^T h)^3-3\psi_b''(v_i){v_i}^2(a_i^T h)(b_i^T h)^2+3\psi_b'(v_i)v_i(a_i^T h)(b_i^T h)^2].$$

It will be convenient to introduce the following notation:

$$\xi(t):=\psi_b''(t)-\frac{\psi_b'(t)}{t}\ , \quad t>0 \tag{39}$$

Note that $(t)>0$, because of (18) and (19). We then may write

$$\nabla^3 \Psi(x)[h,h,h]=2\sum_{i=1}^{n}[\psi_b'''(v_i){v_i}^3(a_i^T h)^3-\ 3\xi(v_i){v_i}^2(a_i^T h)(b_i^T h)^2]. \tag{40}$$

To simplify the notation we define vectors $\eta := (\eta_1,\ldots,\eta_n)$ and $\zeta := (\zeta_1,\ldots,\zeta_n)$ according to

$$\eta_i := a_i^T h=\frac{h_i}{2x_i}+\frac{(Mh)_i}{2s_i}, \qquad \zeta_i := b_i^T h=\frac{h_i}{2x_i}\text{-}\frac{(Mh)_i}{2s_i}, \quad 1\le i\le n,$$

where we used (36). It follows that

$$\eta+\zeta=\frac{h}{x}, \qquad \eta-\zeta=\frac{Mh}{s}. \tag{41}$$

With $X=\operatorname{diag}(x)$ and $S=\operatorname{diag}(s)$, [41] can be written as

$$h=X(\eta+\zeta), \qquad Mh=S(\eta-\zeta). \tag{42}$$

Since M is skew-symmetric, we have $h^TMh=0$. Hence, it follows that $(\eta+\zeta)^T XS(\eta-\zeta)=0$. This means that

$$\sum_{i=1}^{n}(\eta_i+\zeta_i)^T\ {v_i}^2(\eta_i-\zeta_i)=0,$$

or equivalently,

$$\sum_{i=1}^{n}{v_i}^2{\eta_i}^2=\sum_{i=1}^{n}{v_i}^2{\zeta_i}^2.$$

Using the vectors η and ζ, the results of this section can be summarized as follows:

$$\nabla\Psi(x)[h]=2\sum_{i=1}^{n}\psi_b'(v_i)v_i\eta_i, \tag{43}$$

$$\nabla^2 \Psi(x)[h, h]=2\sum_{i=0}^{n}[\psi_b''(v_i)v_i^2\eta_i^2 - \psi_b'(v_i)v_i^2\zeta_i^2], \tag{44}$$

$$\nabla^3 \Psi(x)[h, h, h]= 2\sum_{i=0}^{n}[\psi_b'''(v_i)v_i^3\eta_i^3 - 3\xi(v_i)v_i^2\eta_i\zeta_i^2]. \tag{45}$$

5.2. Upper Bound for $\nu(x)$

We are now ready for the computation of $\nu(x)$. Substituting (43) and (44) into the definition (33) of $\nu(x)$ we get

$$\nu(x), = max_{h\in R^n} \frac{(2\sum_{i=1}^{n}\psi_b'(v_i)v_i\eta_i)^2}{2\sum_{i=1}^{n}(\psi_b''(v_i)v_i^2\eta_i^2 - \psi_b'(v_i)v_i\zeta_i^2)}. \tag{46}$$

When solving the above maximization problem exactly, we should use the fact that the vectors η and ζ satisfy (41). Then, it becomes hard to get a solution, however, because the solution will then depend in a cumbersome way on the matrix M. We therefore consider a relaxation by allowing η and ζ to be arbitrary vectors. This way, we obtain an upper bound for $\nu(x)$. We therefore have

$$\nu(x) \le max_{h\in R^n} \frac{2(\sum_{i=1}^{n}\psi_b'(v_i)v_i\eta_i)^2}{2\sum_{i=0}^{n}(\psi_b''(v_i)v_i^2\eta_i^2 - \psi_b'(v_i)v_i\zeta_i^2)}. \tag{47}$$

Note that both terms in the sum in the denominator are nonnegative. For any value of η, the choice $\zeta = 0$ will give the largest value of the right-hand side in (47). Thus, we are left with

$$\nu(x) \le max_{h\in R^n} \frac{2(\sum_{i=1}^{n}\psi_b'(v_i)v_i\eta_i)^2}{\sum_{i=1}^{n}\psi_b''(v_i)v_i^2\eta_i^2)}.$$

By the Cauchy-Schwarz inequality, we have

$$\left(\sum_{i=1}^{n}\psi_b'(v_i)v_i\eta_i\right)^2 = \left(\sum_{i=1}^{n}\frac{\psi_b'(v_i)}{\sqrt{\psi_b''(v_i)}} \cdot \sqrt{\psi_b''(v_i)}\, v_i\eta_i\right)^2$$
$$\le \sum_{i=1}^{n}\frac{\psi_b'(v_i)^2}{\psi_b''(v_i)} \sum_{i=1}^{n}\psi_b''(v_i)v_i^2\eta_i^2,$$

and hence we obtain

$$\nu(x) \le 2\sum_{i=1}^{n}\frac{\psi_b'(v_i)^2}{\psi_b''(v_i)}. \tag{48}$$

5.3. Upper Bound for $\kappa(x)$

Next, we turn to the computation of $\kappa(x)$. As we will see, this is much more complicated. Substituting (44) and (45) in the definition (34) of $\kappa(x)$, we get

$$\kappa(x) = max_{h\in \mathbb{R}^n} \frac{2\sum_{i=1}^{n}[\psi_b'''(v_i)v_i^3\eta_i^3 - 3\xi(v_i)v_i^2\eta_i\zeta_i^2]}{2[2\sum_{i=1}^{n}(\psi_b''(v_i)v_i^2\eta_i^2 - \psi_b'(v_i)v_i\zeta_i^2)]^{\frac{3}{2}}}. \tag{49}$$

Just as in the previous case we consider a relaxation of this problem by allowing η and ζ to be arbitrary vectors. This way, we obtain an upper bound for $\kappa(x)$. Note that if η changes sign then the objective function changes sign, as well. We therefore have

$$2\sqrt{2}\ \kappa(x) \le max_{\eta,\zeta\in\mathbb{R}^n} \frac{\sum_{i=1}^{n}[3\xi(v_i){v_i}^2\eta_i{\zeta_i}^2 - \psi_b'''(v_i){v_i}^3{\eta_i}^3]}{[2\sum_{i=1}^{n}(\psi_b''(v_i){v_i}^2{\eta_i}^2 - \psi_b'(v_i)v_i{\zeta_i}^2)]^{\frac{3}{2}}}. \tag{50}$$

The last expression is homogeneous in (η, ζ) and the denominator is positive. Hence, we have

$$2\sqrt{2}\kappa(x) \le$$

$$max_{\eta,\zeta}\left\{\sum_{i=1}^{n}\left[3\xi(v_i){v_i}^2\eta_i{\zeta_i}^2 - \psi_b'''(v_i)\,{v_i}^3\ {\eta_i}^3\right] : \sum_{i=1}^{n}(\psi_b''(v_i){v_i}^2{\eta_i}^2 - \psi_b'(v_i)v_i{\zeta_i}^2) = 1\right\}. \tag{51}$$

We may assume without loss of generality that $\eta \ge 0$ (because if $\eta_i < 0$, then replacing η_i by $-\eta_i$ leaves the solution feasible and increases the objective value) and $\zeta \ge 0$ (since only squares of ζ_i occur in the problem formulation). After dividing the objective function by 3, the optimality conditions are, for some multiplier λ,

$$\xi(v_i){v_i}^2{\zeta_i}^2 - \psi_b'''(v_i)\,{v_i}^3\ {\eta_i}^2 = 2\,\lambda\,\psi_b''(v_i){v_i}^2\eta_i, \qquad 1 \le i \le n, \tag{52}$$

$$2\,\xi(v_i){v_i}^2\eta_i\zeta_i = -2\,\lambda\,\psi_b'(v_i)v_i\zeta_i, \qquad 1 \le i \le n. \tag{53}$$

By adding these equations, after multiplying (52) by η_i and (53) by ζ_i we get a surprise, namely that the contribution of each index i to the objective function is equal to 2λ times its contribution to the constraint function:

$$2\lambda\left(\psi_b''(v_i){v_i}^2{\eta_i}^2 - \psi_b'(v_i)v_i{\zeta_i}^2\right) = \xi(v_i){v_i}^2{\zeta_i}^2\eta_i - \psi_b'''(v_i){v_i}^3{\eta_i}^3 + 2\xi(v_i){v_i}^2\eta_i{\zeta_i}^2$$

$$= 3\xi(v_i){v_i}^2{\zeta_i}^2\eta_i - \psi_b'''(v_i){v_i}^3{\eta_i}^3.$$

It follows from this that for each feasible solution the objective value is given by 2λ. Due to assumption (20), we have $-\psi_b'''(v_i) > 0$. Since also $\xi(v_i) > 0$, (52) implies that if λ = 0 then $\eta_i = \zeta_i = 0$, for each i, which is infeasible for (51). From this, we conclude that λ must be positive. The conditions (52) and (53) are satisfied by all optimal solutions of (51), but as is well-known, not every solution may be optimal. In fact, the solutions of (52) and (53) are the so-called stationary points of (51). So, we need to find all the stationary points and find out which one of these points has the maximal objective value. As we already made clear, hereby we may restrict ourselves to stationary points with $\eta \ge 0$ and $\zeta \ge 0$.

Table 1: Contribution to objective value and contribution to function in four cases

Case	η_i	ζ_i	contribution to objective value	contribution to constraint function
I: $i \in I_1$	$= 0$	$= 0$	0	0
II: $i \in I_2$	> 0	$= 0$	$8\lambda^3 \frac{\psi_b''(v_i)^3}{\psi_b'''(v_i)^2}$	$4\lambda^2 \frac{\psi_b''(v_i)^3}{\psi_b'''(v_i)^2}$
III: $i \in I_2$	> 0	> 0	$2\lambda^3 \frac{\psi_b''(v_i)^3}{\psi_b'''(v_i)^2} \bar{\varphi}(v_i)^2\big(3 - \bar{\varphi}(v_i)\big)$. .	$\lambda^2 \frac{\psi_b''(v_i)^3}{\psi_b'''(v_i)^2} \bar{\varphi}(v_i)^2\big(3 - \bar{\varphi}(v_i)\big)$
IV: $i \in I_2$	$= 0$	> 0	n.a.	n.a.

It will be convenient to distinguish four cases for each index i , according to the signs of η_i and ζ_i , as shown in the second and third columns of Table 1. Recall that the values in the fourth column must be just 2λ times the values in the last column. So, we need to the values in one of these columns.

The set of indices i of type I is denoted by I_1 , and in a similar way, we define the sets I_2, I_3 *and* I_4 for indices of types II, III, and IV, respectively. The function $\varphi(t)$ in Table 1 is defined as follows:

$$\varphi(t) = min\big(2, \bar{\varphi}(t)\big), \qquad \bar{\varphi}(t) := \frac{\psi_b'(t)\psi_b'''(t)}{\psi_b''(t)\xi(t)}, \qquad t > 0. \tag{54}$$

Note that due to our assumptions (18)–(20), and (39), $\varphi(t)$> 0, for each $t > 0$. We proceed by justifying the entries in the last two columns of Table 1. The line for case I needs no further explanation. Also case IV is easy, since (53) makes clear that this case cannot occur. In other words, $I_4 = \emptyset$. Fixing i, we now consider first case II and then case III. In case II we have $\zeta_i = 0$; hence, optimality condition (53) is certainly satisfied. From (52), we derive that

$$\eta_i = \frac{2\lambda\, \psi_b''(v_i) {v_i}^2}{-\, \psi_b'''(v_i) {v_i}^3} = \frac{2\lambda\, \psi_b''(v_i)}{-\psi_b'''(v_i) v_i}.$$

Hence, the contribution to the constraint function is

$$\psi_b''(v_i){v_i}^2{\eta_i}^2 - \psi_b'(v_i) v_i {\zeta_i}^2 = \psi_b''(v_i){v_i}^2 \left(\frac{2\lambda\, \psi_b''(v_i)}{-\, \psi_b'''(v_i) v_i}\right)^2 = 4\lambda^2 \frac{\psi_b''(v_i)^3}{\psi_b'''(v_i)^2}.$$

This fixes case II. Next, we consider case III. In this case, (53) implies that

$$\eta_i = \frac{-2\lambda\psi_b'(v_i)v_i}{2\xi(v_i){v_i}^2} = \frac{-\lambda\psi_b'(v_i)}{\xi(v_i)v_i}. \tag{55}$$

Using this and (52), we get

$$\begin{aligned}\xi(v_i){v_i}^2{\zeta_i}^2 &= \eta_i {v_i}^2\big(2\lambda\psi_b''(v_i) + \psi_b'''(v_i)v_i\eta_i\big)\\ &= \eta_i {v_i}^2\left(2\lambda\psi_b''(v_i) - \psi_b'''(v_i)v_i\frac{\lambda\psi_b'(v_i)}{\xi(v_i)v_i}\right)\\ &= \lambda\eta_i {v_i}^2\left(2\lambda\psi_b''(v_i) - \psi_b'''(v_i)v_i\frac{\psi_b'(v_i)}{\xi(v_i)}\right)\\ &= \lambda\eta_i {v_i}^2\psi_b''(v_i)\big(2-\bar{\varphi}(t)\big).\end{aligned}$$

Since the left-hand side is positive, the last expression must be positive as well. Since λ > 0 and $\eta_i > 0$, we conclude that if $i \in I_3$, then

$$\bar{\varphi}(v_i) < 2.$$

Using (52) once more, and also (55), the contribution to the objective function can be deduced as follows:

$$3\xi(v_i)\, {v_i}^2\eta_i\, {\zeta^2}_i - {\psi'''}_b(v_i){v_i}^3{\eta_i}^3 = 3\big[2\lambda{\psi''}_b(v_i){v_i}^2\eta_i + {\psi'''}_b(v_i){v_i}^3\big]\eta_i - {\psi'''}_b(v_i){v_i}^3{\eta_i}^3$$

$$= 6\lambda{\psi''}_b(v_i){v_i}^2{\eta_i}^2 + 2{\psi'''}_b(v_i){v_i}^3{\eta_i}^3 = 6\lambda^3{\psi''}_b(v_i)\frac{{\psi'}_b(v_i)^2}{\xi(v_i)^2} - 2\lambda^3{\psi'''}_b(v_i)\frac{{\psi'}_b(v_i)^3}{\xi(v_i)^3}. \tag{56}$$

After dividing the expression in (56) by 2λ, we get the contribution of index i to the constraint function, which is given by

$${\psi''}_b(v_i){v_i}^2{\eta_i}^2 - {\psi'}_b(v_i)v_i{\zeta_i}^2 = 3\lambda^2\frac{{\psi'}_b(v_i)^2}{\xi(v_i)^2}{\psi''}_b(v_i) - \lambda^2\frac{{\psi'}_b(v_i)^3}{\xi(v_i)^3}{\psi'''}_b(v_i)$$

$$= \lambda^2\frac{{\psi''}_b(v_i)^3}{{\psi'''}_b(v_i)^2}\frac{{\psi'}_b(v_i)^2\,{\psi'''}_b(v_i)^2}{\xi(v_i)^2{\psi''}_b(v_i)^2}\left(3 - \frac{{\psi'}_b(v_i){\psi'''}_b(v_i)}{{\psi''}_b(v_i)\xi(v_i)}\right) = \lambda^2\frac{{\psi''}_b(v_i)^3}{{\psi'''}_b(v_i)^2}\bar{\varphi}(v_i)^2(3-\bar{\varphi}(v_i)),$$

where the last equality is due to (54). Thus, we have justified all the entries in Table 1. By adding all entries in the fourth column we get the objective value in a stationary point. As we established before, this value equals 2λ. Thus, we obtain

$$2\lambda = 2\lambda^3\left(\sum_{i\in I_2} 4\frac{{\psi''}_b(v_i)^3}{{\psi'''}_b(v_i)^2} + \sum_{i\in I_3}\bar{\varphi}(v_i)^2\big(3-\bar{\varphi}(v_i)\big)\frac{{\psi''}_b(v_i)^3}{{\psi'''}_b(v_i)^2}\right).$$

Since $\lambda > 0$, this implies

$$2\lambda = \frac{2}{\sqrt{\sum_{i\in I_2} 4\frac{{\psi''}_b(v_i)^3}{{\psi'''}_b(v_i)^2} + \sum_{i\in I_3}\bar{\varphi}(v_i)^2(3-\bar{\varphi}(v_i))\frac{{\psi''}_b(v_i)^3}{{\psi'''}_b(v_i)^2}}}. \tag{57}$$

Note that v is fixed in the present context, and hence also the expressions $\psi''_b(v_i)^3/\psi'''_b(v_i)^2$ and $\bar{\varphi}(v_i)^2(3-\bar{\varphi}(v_i))$. Recall that the above expression represents the objective value at one specific stationary point. In fact, we have found many such points, each of which being characterized by a partition $\{I_1, I_2, I_3\}$ of the index set. Our final task is to find the partition that maximizes 2λ. Note that we cannot put all indices in I_1, because this would mean that $\eta_i = \zeta_i = 0$ for all i; as we observed before, this is infeasible. We conclude that the largest objective value is obtained if all indices but one are in I_1, and this index is either in I_2 or I_3. In other words, we should find the smallest term in the sum under the square root.

One easily verifies that $z^2(3-z) \le 4$ for all $z > 0$, with equality being satisfied only if $z = 2$. Recall that an index i can be in I_3 only if $\bar{\varphi}(\boldsymbol{v_i}) \in (0, 2)$. Since

$$0 \le \bar{\varphi}(v_i)^2(3-\bar{\varphi}(v_i)) < 4, \qquad 1 \le i \le n,$$

it follows that if $\bar{\varphi}(\boldsymbol{v_i}) \in (0, 2)$ holds then we get the highest value for 2λ, if $i \epsilon I_3$, and otherwise the highest value for 2λ occurs, if $i \in I_2$. Since 2^2 (2 − 1) = 4, and z^2 (3 − z) is increasing for $z \in (0, 2)$, in both cases (57) reduces to

$$2\lambda = \frac{2}{\sqrt{\varphi(v_i)^2(3-\varphi(v_i))\frac{\psi''_b(v_i)^3}{\psi'''_b(v_i)^2}}} = \frac{2}{\varphi(v_i)\sqrt{(3-\varphi(v_i))}}\frac{-\psi'''_b(v_i)}{\psi''_b(v_i)^{\frac{3}{2}}},$$

where we used $\varphi(x)$ as defined in (54). By maximizing over i, we get the largest possible value for 2λ, which is an upper bound for $2\sqrt{2}\,\kappa(x)$, according to (50). Thus, we obtain the following upper bound for $\kappa(x)$:

$$\kappa(x) \le max_i \frac{1}{\sqrt{2}\varphi(v_i)\sqrt{3-\varphi(v_i)}}\frac{-\psi'''_b(v_i)}{\psi''_b(v_i)^{\frac{3}{2}}}. \tag{58}$$

6. Complexity Number along the Central Path

Yet, we are ready to compute the complexity number on the central path. Because the variance vector of a point on the central path is the all-one vector, it follows that if $x = x(\mu)$, then $v_i = 1$, for all i. Hence, we obtain from (48) and (58) that

$$v(x(\mu)) \le 2\sum_{i=1}^{n}\frac{\psi'_b(1)^2}{\psi''_b(1)} = 2n\frac{\psi'_b(1)^2}{\psi''_b(1)},$$

$$\kappa(x(\mu)) \le \frac{1}{\sqrt{2}\varphi(1)\sqrt{3-\varphi(1)}}\frac{-\psi'''_b(1)}{\psi''_b(1)^{\frac{3}{2}}}.$$

Hence, the complexity number on the central path satisfies

$$\gamma(x(\mu)) \le \frac{1}{\sqrt{2}\varphi(1)\sqrt{3-\varphi(1)}}\frac{-\psi'''_b(1)}{\psi''_b(1)^{\frac{3}{2}}}\sqrt{2n\frac{\psi'_b(1)^2}{\psi''_b(1)}} = \frac{1}{\sqrt{2}\varphi(1)\sqrt{3-\varphi(1)}}\frac{\psi'_b(1)\psi'''_b(1)}{\psi'_b(1)^2}\sqrt{2n}.$$

Using the definition (54) of $\bar{\varphi}$, this can be written as

$$\gamma\big(x(\mu)\big) \leq \begin{cases} \dfrac{1}{\sqrt{2}\sqrt{\left(3-\overline{\varphi}\,(1)\right)}}\dfrac{\xi(1)}{\psi_b^{''}(1)}\sqrt{2n}, & \text{if } \overline{\varphi}(1) \, < \, 2, \\ \dfrac{\xi(1)\,\overline{\varphi}\,(1)}{2\sqrt{2}\psi_b^{''}(1)}\sqrt{2n}, & \text{if } \overline{\varphi}(1) \, \geq \, 2. \end{cases} \tag{59}$$

Table 2 shows the resulting values for the (classes of) kernel functions in Fig. 1.

7. Conclusions

We conclude with a few remarks on the results presented in Table 2. First, we point out the remarkable fact that (x) , as well as $\kappa(x)$ and $\nu(x)$, is constant along the central path. Of course, this is due to the fact that these numbers depend only on the variance vector v of x, which is already clear from (48) and (58). Second, the complexity number of the logarithmic barrier function turns out to be $\sqrt{2n}$, which is in agreement with the theory of SCBFs. Also, using (48) and (58), one easily verifies that in that case we have for every $x \in \mathcal{P}^0$ that $\kappa(x) = 1$ and $\nu(x) = 2n$, proving that this function is an SCBF, whereas for the other kernel-based barrier functions $\kappa(x)$ and $\nu(x)$ are unbounded on $\mathcal{P}^0$. Third, in order to ease comparison with the logarithmic barrier function, we wrote the other complexity numbers as a multiple of $\sqrt{2n}$. It turns out that in many cases these complexity numbers are lower than $\sqrt{2n}$. The only exceptions are ψ_5 for $q \geq 1$, and ψ_7 for $q \leq 1.2999$. The lowest complexity numbers arise for ψ_4, ψ_7 and ψ_8, when q increases; asymptotically, these complexity numbers become $\frac{1}{2}\sqrt{2n}$. Fourth and finally, except for ψ_5 (with $q > 3$), we have in all cases that $\overline{\varphi}(1) < 2$. An interesting question is whether the notion of local self-concordance can be used to design efficient algorithms. Computational experiments in Matlab provided clear evidence that Algorithm 2.1 works fine for each of the new barrier functions that we dealt with here. But, the theoretical analysis will require further research.

Acknowledgments

The authors thank the referees for their useful comments. The first author was supported by a grant from the National Natural Science Foundation of China (No. 11071158). The third and fifth authors were partially supported by the Center of Excellence for Mathematics, University of Shahrekord, Shahrekord, Iran.

Table 2: Complexity numbers on the central path

i	kernel function $\psi_i(t)$	$\psi_b''(1)$	$\xi(1)$	$\bar{\varphi}(1)$	$\gamma(x(\mu)) \leq$
1	$\frac{t^2-1}{2} - \log t$	1	2	1	$\sqrt{2n}$
2	$\frac{t^2-1}{2} + \frac{1}{t} - 1$	2	3	1	$\frac{3}{4}\sqrt{2n}$
3	$\frac{1}{2}\left(t - \frac{1}{t}\right)^2$	3	4	1	$\frac{2}{3}\sqrt{2n}$
4	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1}, \quad q > 1$	q	$q+1$	1	$\frac{q+1}{2q}\sqrt{2n}$
$5a$	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q(q-1)} - \frac{q-1}{q}(t-1), \quad q \leq 3$	1	2	$\frac{q+1}{2}$	$\frac{2}{\sqrt{5-q}}\sqrt{2n}$
$5b$	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q(q-1)} - \frac{q-1}{q}(t-1), \quad q \geq 3$	1	2	$\frac{q+1}{2}$	$\frac{q+1}{2\sqrt{2}}\sqrt{2n}$
6	$\frac{t^2-1}{2} + e^{\frac{1}{t}-1} - 1$	3	4	$\frac{13}{12}$	$\sqrt{\frac{32}{69}}\sqrt{2n}$
7	$\frac{t^2-1}{2} + \frac{e^{q\left(\frac{1}{t}-1\right)}-1}{q}, \quad q > 0$	$q+2$	$q+3$	$\frac{q^2+6q+6}{q^2+5q+6}$	$\sqrt{\frac{(q+3)^3}{2(q+2)(2q^2+9q+12)}}\sqrt{2n}$
8	$\frac{t^2-1}{2} - \int_1^t e^{q(\frac{1}{\xi}-1)} d\xi, \quad q \geq 1$	q	$q+1$	$\frac{q+2}{q+1}$	$\sqrt{\frac{(q+1)^3}{2q^2(2q+1)}}\sqrt{2n}$
9	$\frac{t^2-1}{2} + \frac{6}{\pi}\cot\frac{3\pi t}{4t+2}$	$\frac{4}{3}$	$\frac{7}{3}$	$\frac{48+\pi^2}{56}$	$\frac{7}{2}\sqrt{\frac{7}{120-\pi^2}}\sqrt{2n}$
10	$\frac{t^2-1}{2} + e^{e^{\frac{4\cot\frac{\pi t}{t+1}}{\pi}}-1} - 1$	3	4	$\frac{100+\pi^2}{96}$	$\frac{16}{\sqrt{3(188-\pi^2)}}\sqrt{2n}$

References

[1] Bai, Y., Ghami, M.E., and Roos, C. (2004), A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization, *SIAM Journal on Optimization*, 15(1), 101–128.

[2] Bai, Y., Guo, J. and Roos, C. (2009), A new kernel function yielding the best known iteration bounds for primal-dual interior-point algorithms, *Acta Mathematica Sinica*, 25(12), 2169–2178.

[3] Bai, Y., Lesaja, G., Roos, C., Wang, G.Q. and El Ghami, M. (2008), A class of large-update and small-update primal-dual interior-point algorithms for linear optimization, *Journal of Optimization Theory and Applications*, 138(3), 341–359.

[4] Boyd, S. and Vandenberghe, L. (2004), Convex Optimization, Cambridge University Press, Cambridge, UK.

[5] ElGhami, M., Guennoun, Z., Bouali, S. and Steihaug, T. (2011), Interior-point methods for linear optimization based on a kernel function with trigonometric barrier term, Submitted to JCAM.

[6] Frisch, K. (1955a), The logarithmic potential method for convex programming, Unpublished manuscript, Institute of Economics, University of Oslo, Oslo, Norway.

[7] Frisch, K. (1955b), The logarithmic potential method for solving linear programming problems, Memorandum, University Institute of Economics, Oslo.

[8] Glineur, F. (2000), Topics in Convex Optimization: Interior Point Methods, Conic Duality and Approximation, MPS/SIAM Series on Optimization.

[9] Glineur, F. (2001), Topics in convex optimization: interior-point methods, conic duality and approximations, PhD. thesis, Faculté Polytechnique de Mons.

[10] Gonzaga, C. (1989), An algorithm for solving linear programming problems in $O(n^3L)$ operations, In Megiddo, N., editor, Progress in Mathematical Programming : Interior Point and Related Methods, pages 1–28. Springer Verlag, New York.

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

[12] Kheirfam, B. (2012), A generic interior-point algorithm for monotone linear complementarity problems based on a new kernel function, Manuscript.

[13] Kojima, M., Mizuno, S. and Yoshise, A. (1989). A primal-dual interior point algorithm for linear Programming, In Megiddo, N., editor, Progress in Mathematical Programming: Interior Point and Related Methods, pages 29–47, Springer Verlag, New York.

[14] Lasserre, J. (2000), Why the logarithmic barrier function in convex and linear programming? *Operations Research Letters*, 27, 149–152.

[15] Mizuno, S. (1990), An $O(n^3L)$ algorithm using a sequence for linear complementarity problems, *Journal of the Operations Research Society of Japan*, 33, 66–75.

[16] Nesterov, Y. (2004), Introductory Lectures on Convex Optimization. A Basic Course, volume 87 of Applied Optimization, Kluwer Academic Publishers, Boston, MA.

[17] Nesterov, Y. and Nemirovski, A. (1993), Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms, SIAM Publications. SIAM, Philadelphia, USA.

[18] Nesterov, Y. and Nemirovskii, A. (1994), Interior-point polynomial algorithms in convex programming, volume 13 of *SIAM Studies in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

[19] Peng, J., Roos, C. and Terlaky, T. (2000), New complexity analysis of the primal-dual Newton method for linear optimization, *Ann. Oper Res.*, 99, 23–39, Applied mathematical programming and modeling, IV (Limassol, 1998).

[20] Peng, J., Roos, C. and Terlaky, T. (2002a), A new class of polynomial primal-dual methods for linear and semidefinite optimization, *European Journal of Operations Research*, 143(2), 234–256.

[21] Peng, J., Roos, C. and Terlaky, T. (2002b), Self-regular functions and new search directions for linear and semidefinite optimization, *Mathematical Programming*, 93, 129–171.

[22] Peng, J., Roos, C. and Terlaky, T. (2002c), Self-Regularity. A New Paradigm for Primal-Dual Interior-Point Algorithms, Princeton University Press.

[23] Renegar, J. (1988), A polynomial-time algorithm, based on Newton's method, for linear programming, *Mathematical Programming*, 40, 59–93.

[24] Renegar, J. (2001), A Mathematical View of Interior-Point Methods in Convex Optimization, MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia,PA.

[25] Roos, C., Terlaky, T. and Vial, J.-P. (2005a), Theory and Algorithms for Linear Optimization, Springer, Chichester, UK (1st Edition, Theory and Algorithms for Linear Optimization, An Interior-Point Approach. John Wiley & Sons, 1997).

[26] Roos, C., Terlaky, T. and Vial, J.-P. (2005b*),*Theory and Algorithms for Linear Optimization. Springer, Chichester, UK (1st Edition, Theory and Algorithms for Linear Optimization*,* An Interior-Point Approach. John Wiley & Sons, 1997).

[27] Roos, C. and Vial, J.-P. (1992), A polynomial method of approximate centers for linear programming, *Mathematical Programming*, 54, 295–305.

[28] Terlaky, T., (1996), Interior Point Methods of Mathematical Programming, volume 5 of Applied Optimization, Kluwer Academic Publishers, Dordrecht.

[29] Vanderbei, R. (1996), Linear Programming: Foundations and Extensions, Kluwer Academic Publishers, Boston, USA.

[30] Wright, S. J. (1997), Primal-dual Interior-Point Methods, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

[31] Ye, Y. (1997), Interior Point Algorithms, Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York. Theory and analysis, A Wiley-Interscience Publication.