Filomat 34:12 (2020), 3957–3969
https://doi.org/10.2298/FIL2012957T

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Primal-Dual Interior Point Methods for Semidefinite Programming Based on a New Type of Kernel Functions

Imene Touil[a], Wided Chikouche[a]

[a] *LMPA, Mohamed Seddik Ben Yahia University, Jijel, Algeria*

Abstract. In this paper, we propose the first hyperbolic-logarithmic kernel function for Semidefinite programming problems. By simple analysis tools, several properties of this kernel function are used to compute the total number of iterations. We show that the worst-case iteration complexity of our algorithm for large-update methods improves the obtained iteration bounds based on hyperbolic [24] as well as classic kernel functions. For small-update methods, we derive the best known iteration bound.

1. Introduction

In this paper, we deal with the semidefinite programming (SDP) problem given in the following standard form

$$(P)\min\{\langle C, X\rangle \ : \ \langle A_i, X\rangle = b_i,\ i = 1, ..., m,\ X \in \mathcal{S}^n_+\},$$

and its dual problem

$$(D)\max\{b^t y \ : \ \sum_{i=1}^{m} y_i A_i + S = C,\ S \in \mathcal{S}^n_+\},$$

where $b \in \mathcal{R}^m$, $C, A_i \in \mathcal{S}^n$ are given, and $y \in \mathcal{R}^m$, $X, S \in \mathcal{S}^n$. Here $\mathcal{S}^n_+$ denotes the cone of positive semidefinite matrices in the real space of $(n \times n)$ symmetrical matrices $\mathcal{S}^n$. We assume that the A_i's are linearly independent. Interior point methods (IPMs) provide a powerful approach for solving SDP problems. The pioneering works in this direction are due to Alizadeh [1, 2] and Nesterov and Nemirovskii [14], we can find a comprehensive list of publications on this topic in the SDP homepage maintained by Alizadeh [1]. Most IPMs for SDP can be viewed as natural extensions of IPMs for linear programming (LP) and have similar polynomial complexity results. However, obtaining a valid search direction in the SDP case is much more difficult than in the LP case.

IPMs for LP and SDP are generally based on the logarithmic kernel function [10, 21] with complexity $O(n \log \frac{n}{\epsilon})$, for large-update methods and $O(\sqrt{n} \log \frac{n}{\epsilon})$ for small-update methods, where n is the size of the problem and ϵ is the accuracy parameter. Kernel functions play an important role in design and complexity analysis of interior point algorithms for solving convex optimization problems. In each iteration of these

2010 *Mathematics Subject Classification*. Primary 90C22; Secondary 90C51, 90C31
Keywords. Linear semidefinite programming, Primal-dual interior point methods, Hyperbolic-logarithmic kernel function, Complexity analysis, Large- and small-update methods
Received: 14 December 2019; Revised: 29 July 2020; Accepted: 02 December 2020
Communicated by Predrag Stanimirović
Email addresses: i_touil@yahoo.fr (Imene Touil), w_chikouche@yahoo.com (Wided Chikouche)

methods, the search direction is determined and the distance between the current point and the central path is measured by a proximity function which is typically induced from a kernel function.
In the last two decades, many new kernel functions have been introduced with the ultimate goal to find a kernel function that improves the complexity analysis of linear and also nonlinear optimization problems. Kernel functions were introduced first in 2001 by Peng et al. [15–17]. The authors proposed new variants of IPMs for LP and SDP problems based on a class of self-regular kernel functions with a non logarithmic barrier term. They improved the iteration bound for large-update methods from the classical iteration bound $O(n\log\frac{n}{\epsilon})$ to $O(\sqrt{n}\log n\log\frac{n}{\epsilon})$, which almost closes the gap between the iteration bounds for large- and small-update methods. After [15–17], researchers looked for other kernel functions with even better bounds, may be even better than the bound for small-update methods, since in practice large-update behave much more efficient than small-update methods.

In 2005, Wang et al. [25] presented a primal-dual interior point algorithm for SDP problems based on a simple non self-regular kernel function which was first introduced in [6] for LP. Later on, Qian et al. [20] proposed a new kernel function with simple algebraic expression and established the iteration complexity as $O(n^{\frac{3}{4}}\log\frac{n}{\epsilon})$ for large-update methods for SDP. In 2004, Bai et al [4] introduced first new kernel function with a trigonometric barrier term. The IPM with this function has been further analyzed by El Ghami et al [11] in 2012. They established the iterations complexity as $O(n^{\frac{3}{4}}\log\frac{n}{\epsilon})$ for large-update methods.

Since then, a number of various kernel functions with a trigonometric barrier term has been proposed and analyzed. For these, we refer the reader to [7, 12, 13, 19]. The authors in [7, 19] are the first to reach the best known complexity bound for large-update methods for trigonometric kernel functions. In 2014, Peyghami et al [18] and Cai et al. [8] proposed a new kernel function with trigonometric-logarithmic barrier term for LP, which has $O(n^{\frac{2}{3}}\log\frac{n}{\epsilon})$ complexity bounds for large-update methods. This improves the complexity bounds obtained in [11].

Motivated by their works, the purpose of this paper is to deal with the primal-dual IPMs for SDP based on a new kind of kernel functions. Compared to the existing ones, the proposed function has a hyperbolic-logarithmic barrier term. Furthermore, we present a primal-dual IPM for SDP based on this new kernel function. The obtained iteration bound for large-update methods, namely, $O\left(n^{\frac{2}{3}}\log\frac{n}{\epsilon}\right)$, improves the classical iteration complexity with a factor $n^{\frac{1}{3}}$. For small-update methods, we derive the iteration bound $O\left(\sqrt{n}\log\frac{n}{\epsilon}\right)$, which matches the currently best known iteration bound for small-update methods.

Note that, kernel function with a hyperbolic barrier term was recently introduced, for the first time, by Touil and Chikouche [24]. They established that the complexity bound is $O\left(n^{\frac{3}{4}}\log\frac{n}{\epsilon}\right)$ for large-update methods. The obtained iteration bound is not as good as the best existing ones, but it is of interest that it belongs to a new class that so far has not been investigated.
The paper is organized as follows. In section 2, we recall some fundamental concepts about central path and search direction for SDP. The new kernel function and some of its properties as well as the corresponding barrier function are described in section 3. The estimate of the step size and the decrease behavior of the new barrier function are discussed in section 4, where we present the iteration bound of our algorithm for large- and small-update methods. Finally, some conclusions and remarks follow in section 5.
Let us finish this introduction with some notations used in the whole paper: The set of all $(n\times n)$ matrices with real entries is denoted by $\mathcal{R}^{n\times n}$. Given $M\in\mathcal{R}^{n\times n}$, M^t denotes the transpose of M. $\mathcal{S}^n_{++}$ denotes the cone of positive definite matrices in $\mathcal{S}^n$. The scalar product of two matrices A and B in $\mathcal{S}^n$ is the trace of their product i.e., $\langle A,B\rangle=\mathrm{tr}(AB)=\sum\limits_{i,j=1}^{n}a_{ij}b_{ij}$. For any $M\in\mathcal{S}^n$, we denote by $\lambda_i(M), i=1,...,n$, the eigenvalues of the matrix M. The diagonal matrix with diagonal entries $\lambda_i(M),\ i=1,...n$, is denoted by $diag(\lambda_1(M),...,\lambda_n(M))$. The Frobenius norm of $M\in\mathcal{S}^n$ is $\|M\|=\langle M,M\rangle^{\frac{1}{2}}=\sqrt{\sum\limits_{i,j=1}^{n}m_{ij}^2}=\sqrt{\sum\limits_{i=1}^{n}\lambda_i^2(M)}$. For any $M\in\mathcal{S}^n_{++}$, the expression $M^{\frac{1}{2}}$ denotes its symmetric square root. Finally, if $f(x)\geq 0$ is a real valued function of a real nonnegative variable, the notation $f(x)=O(g(x))$ means that $f(x)\leq Cg(x)$ for some positive constant C and $f(x)=\Theta(g(x))$ means that $C_1g(x)\leq f(x)\leq C_2g(x)$ for two positive constants C_1 and C_2.

2. Preliminaries

2.1. Framework of kernel-based IPMs for SDP

We assume that the problems (P) and (D) satisfy the interior point condition (IPC), i.e., there exists (X^0, y^0, S^0) such that

$$\langle A_i, X^0\rangle = b_i,\ i = 1, ..., m,\ X^0 \in \mathcal{S}^n_{++},\ \sum_{i=1}^{m} y_i^0 A_i + S^0 = C,\ S^0 \in \mathcal{S}^n_{++}.$$

It is well known that finding an optimal solution of (P) and (D) is equivalent to solve the following system

$$\langle A_i, X\rangle = b_i,\ i = 1, ..., m,\ X \in \mathcal{S}^n_{+},\ \sum_{i=1}^{m} y_i A_i + S = C, S \in \mathcal{S}^n_{+},\ XS = 0. \tag{1}$$

The core idea of primal-dual IPMs is to replace the complementarity condition $XS = 0$ by the parameterized equation $XS = \mu I$, where $\mu > 0$.

Under the IPC, the modified system has a unique solution denoted by (X_μ, y_μ, S_μ), for each $\mu > 0$. The set of all solutions (X_μ, y_μ, S_μ), with $\mu > 0$ is known as the *central path* or *central trajectory* [23, 26]. The principal idea of IPMs is to follow this central path and approach the optimal set as μ goes to zero. As is well known, the definition of search directions for SDP requires some symmetrization scheme. In this paper, we use the scheme due to Nesterov and Todd (NT) [22], which uses the positive definite matrix

$$P = \left[X^{\frac{1}{2}}\left(X^{\frac{1}{2}}SX^{\frac{1}{2}}\right)^{-\frac{1}{2}}X^{\frac{1}{2}}\right]^{-\frac{1}{2}} = \left[S^{-\frac{1}{2}}\left(S^{\frac{1}{2}}XS^{\frac{1}{2}}\right)^{\frac{1}{2}}S^{-\frac{1}{2}}\right]^{-\frac{1}{2}}.$$

The matrix P can be used to scale X and S to the same matrix V because

$$V = \frac{1}{\sqrt{\mu}}PXP = \frac{1}{\sqrt{\mu}}P^{-1}SP^{-1}. \tag{2}$$

Obviously the matrices P and V are symmetric and positive definite.

For fixed $\mu > 0$, a direct application of the Newton method to the system (1) with the parameterized equation $XS = \mu I$ instead of $XS = 0$, gives

$$\begin{cases} \langle \overline{A_i}, D_X\rangle = 0,\ i = 1, ..., m, \\ \sum\limits_{i=1}^{m} \Delta y_i \overline{A_i} + D_S = 0, \\ D_X + D_S = V^{-1} - V, \end{cases} \tag{3}$$

where

$$\overline{A_i} = \frac{1}{\sqrt{\mu}}P^{-1}A_iP^{-1}, i = 1, ..., m, D_X = \frac{1}{\sqrt{\mu}}P\Delta XP, D_S = \frac{1}{\sqrt{\mu}}P^{-1}\Delta SP^{-1}. \tag{4}$$

The solution of (3) defines the (scaled) NT search direction $(D_X, \Delta y, D_S)$.
Now, we define the matrix function $\psi(V)$ obtained from $\psi(t)$.

Definition 2.1. *Let $V \in \mathcal{S}^n_{++}$ and*

$$V = Q^t diag(\lambda_1(V), \lambda_2(V), ..., \lambda_n(V))Q,$$

where Q is any orthonormal matrix ($Q^t = Q^{-1}$) that diagonalizes V.
The matrix function $\psi(V) : \mathcal{S}^n_{++} \to \mathcal{S}^n$ is defined by

$$\psi(V) = Q^t diag(\psi(\lambda_1(V)), \psi(\lambda_2(V)), ..., \psi(\lambda_n(V)))Q.$$

Remark 2.2. *In this paper, when we use the function $\psi(.)$ and its first three derivatives $\psi'(.)$, $\psi''(.)$, and $\psi'''(.)$ without any specification, it denotes a matrix function if the argument is a matrix and a univariate function (from $\mathcal{R}$ to $\mathcal{R}$) if the argument is in $\mathcal{R}$.*

Corresponding to our kernel function, we define the proximity (barrier) function $\Psi(V) : \mathcal{S}^n_{++} \to \mathcal{R}_+$ as follows

$$\Phi(X,S;\mu) := \Psi(V) := \operatorname{tr}(\psi(V)) = \sum_{i=1}^{n} \psi(\lambda_i(V)). \tag{5}$$

Just as in [9], we replace the right-hand side in the last equation of (3) by $-\psi'(V)$, we get

$$\begin{cases} \langle \overline{A_i}, D_X \rangle = 0, \; i = 1, ..., m, \\ \sum\limits_{i=1}^{m} \Delta y_i \overline{A_i} + D_S = 0, \\ D_X + D_S = -\psi'(V). \end{cases} \tag{6}$$

It is easy to verify that this system has a unique solution $(D_X, \Delta y, D_S)$. From (4), having D_X and D_S, we can compute ΔX and ΔS.

Due to the orthogonality of ΔX and ΔS, it is trivial to see that D_X and D_S are orthogonal, and so, $\langle D_X, D_S \rangle = 0$. Then, we can easily verify that the matrix function $\psi(V)$ determines in a natural way an interior point algorithm.

$$D_X = D_S = 0_{n\times n} \Leftrightarrow \psi'(V) = 0_{n\times n} \Leftrightarrow V = I \Leftrightarrow \Psi(V) = 0 \Leftrightarrow XS = \mu I,$$

i.e., if and only if $X = X_\mu$ and $S = S_\mu$, as it should. Otherwise $\Psi(V) > 0$. Thus we conclude that $\Delta X, \Delta y$ and ΔS all vanish if and only if $V = I$, that is, if and only if $(X, y, S) = (X_\mu, y_\mu, S_\mu)$. Otherwise, we will use $(\Delta X, \Delta y, \Delta S)$ as the new search direction. The new iterate (X_+, y_+, S_+) is given by

$$X_+ = X + \alpha\Delta X, y_+ = y + \alpha\Delta y, S_+ = S + \alpha\Delta S, \tag{7}$$

where α denotes the default step size, $\alpha \in (0, 1]$, which has to be chosen appropriately.

For the analysis of the interior point algorithm, we define the norm-based proximity measure $\sigma(V)$, as follows

$$\sigma(V) = \frac{1}{2}\|D_X + D_S\| = \frac{1}{2}\left\|\psi'(V)\right\| = \frac{1}{2}\sqrt{\operatorname{tr}\left(\psi'(V)^2\right)}. \tag{8}$$

2.2. The primal-dual interior point algorithm for SDP

In general, each kernel function gives rise to a primal-dual interior point algorithm. For the description of our algorithm, it is clear that closeness of (X, y, S) to (X_μ, y_μ, S_μ) is measured by the value of $\Psi(V)$ with τ as a threshold value: if $\Psi(V) \le \tau$, then we decrease μ to $\mu_+ := (1-\theta)\mu$, for some fixed $\theta \in]0, 1[$, and solve the Newton system (6) to obtain the unique search direction $(\Delta X, \Delta y, \Delta S)$. Then, we apply (7) to get the new iterate. This procedure is repeated until we find a new iterate (X_+, y_+, S_+) that is $\Psi(V) \le \tau$ and then we let $\mu := \mu_+$ and $(X, y, S) := (X_+, y_+, S_+)$. Then, μ is again reduced by the factor $(1-\theta)$ and we solve the Newton system targeting at the new μ_+-center, and so on. This process is repeated until μ is small enough, say until $n\mu < \epsilon$ for a certain accuracy parameter ϵ, at this stage we have found an $\epsilon-$optimal solution of (P) and (D). The algorithm of primal-dual IPM based on our kernel function is given in Algorithm 1.

3. Properties of the new proximity function

In this section, we introduce the new kernel function and study its properties, which are essential to our complexity analysis.

Algorithm 1 : Primal-dual algorithm for SDP

Input
a threshold parameter $\tau \geq 1$;
an accuracy parameter $\epsilon > 0$;
a fixed barrier update parameter $\theta \in]0, 1[$;
a kernel function $\psi(t)$;
(X^0, y^0, S^0) satisfy the IPC and $\mu^0 = 1$ such that $\Phi\left(X^0, S^0; \mu^0\right) \leq \tau$;
begin
$X: = X^0; y: = y^0; S: = S^0; \mu: = \mu^0$;
while $n\mu \geq \epsilon$ **do**
  begin (outer iteration)
  $\mu: = (1-\theta)\mu$;
  while $\Phi(X, S; \mu) = \Psi(V) > \tau$ **do**
    begin (inner iteration)
    Solve system (6) and use (4) to obtain $(\Delta X, \Delta y, \Delta S)$;
    Choose a suitable step size α;
    $(X, y, S) := (X, y, S) + \alpha(\Delta X, \Delta y, \Delta S)$;
    $V := \frac{1}{\sqrt{\mu}}(PXSP^{-1})^{\frac{1}{2}}$;
  end while (inner iteration)
end while (outer iteration)

3.1. *Some technical results*

Let us define the new kernel function as

$$\psi(t) = \left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right)\frac{t^2-1}{2} + \coth^2(t) - \log t - \coth^2(1),\ \forall t > 0. \tag{9}$$

We have

$$\lim_{t\to 0^+} \psi(t) = \lim_{t\to +\infty} \psi(t) = +\infty,$$

and the first three derivatives of the function ψ on $]0, +\infty[$ are

$$\psi'(t) = \left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right)t - \frac{1}{t} - \frac{2\coth(t)}{\sinh^2(t)}, \tag{10}$$

$$\psi''(t) = \left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right) + 2\left(\frac{3\coth^2(t)-1}{\sinh^2(t)}\right) + \frac{1}{t^2} > 1 + \frac{2\coth(1)}{\sinh^2(1)}, \tag{11}$$

and

$$\psi'''(t) = -8\frac{\coth t}{\sinh^2(t)}\left(3\coth^2(t) - 2\right) - \frac{2}{t^3} < 0, \tag{12}$$

where

$$\coth(t) > 1,\ \forall t > 0.$$

The kernel function $\psi(t)$ is twice differentiable and one can easily verify that $\psi(1) = \psi'(1) = 0$. This implies that $\psi(t)$ is completely defined by its second derivative as follows

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)\, d\zeta d\xi. \tag{13}$$

In what follows, we develop some technical lemmas on the new kernel function.

Lemma 3.1. *Let $\psi(t)$ be the function defined in* (9)*. Then, we have*

$$2t\coth(t) - 1 > 0,\ \forall t > 0. \tag{14}$$

$$\psi'(t) = \left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right)t - \frac{1}{t} - \frac{2\coth(t)}{\sinh^2(t)} > 0,\ \forall t > 1. \tag{15}$$

Proof. (14) is proven in [24, Lemma 3.2].
For the inequality (15), $\psi'(t)$ can be written as follows

$$\begin{aligned}\psi'(t) &= 2\left(\frac{t\coth(1)}{\sinh^2(1)} - \frac{\coth(t)}{\sinh^2(t)}\right) + t - \tfrac{1}{t}\\ &> 2\left(\frac{t\coth(1)}{\sinh^2(1)} - \frac{\coth(t)}{\sinh^2(t)}\right),\ \forall t > 1.\end{aligned}$$

Thus, $\psi'(t) > 0,\ \forall t > 1$, since sinh (respectively coth) is positive and increasing (respectively decreasing) in $]0, +\infty[$ then in $]1, +\infty[$, this implies that

$$\frac{\coth t}{\sinh^2(t)} < \frac{\coth 1}{\sinh^2(t)} < \frac{\coth 1}{\sinh^2(1)} < \frac{t\coth 1}{\sinh^2(1)},\ \forall t > 1.$$

Thus, the inequality (15) is a direct consequence, which completes the proof. □

Lemma 3.2. *Let $\psi(t)$ be the function defined in* (9)*, then we have*

(i) *$\psi(t)$ is convex exponentially for all $t > 0$; that is*

$$\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)),\ \forall t_1, t_2 > 0.$$

(ii) *$\psi''(t)$ is monotonically decreasing, $\forall t > 0$.*

(iii) $t\psi''(t) - \psi'(t) > 0,\ \forall t > 0.$

Proof. For (*i*), by Lemma 2.1 in [9], it suffices to show that $t\psi''(t) + \psi'(t) \geq 0,\ \forall t > 0$. Indeed, using (10), (11) and (14) of Lemma 3.1, we have

$$\begin{aligned}t\psi''(t) + \psi'(t) &= 2t\left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right) + \frac{2}{\sinh^2(t)}\left(3\coth^2(t) - \coth t - 1\right)\\ &= 2t\left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right) + \frac{2t}{\sinh^4(t)} + \frac{2\coth t}{\sinh^2(t)}(2t\coth t - 1) > 0.\end{aligned}$$

(*ii*) follows immediately since $\psi'''(t) < 0, \forall t > 0$, from (12).
For the item (*iii*), using (10) and (11), we have

$$t\psi''(t) - \psi'(t) = 2t\left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right) + \frac{2}{\sinh^2(t)}\left(3\coth^2(t) + \coth t - 1\right) > 0.$$

This completes the proof. □

Lemma 3.3. *For $\psi(t)$, we have*

(i) $2(t-1)^2 \leq \psi(t) \leq \left(\frac{\psi'(t)}{2}\right)^2,\ \forall t > 0.$

(ii) $\psi(t) \leq \dfrac{1}{2}\psi''(1)(t-1)^2,\ \forall t \geq 1.$

Proof. Using (13) and (11), we have

$$\psi(t) \geq \left(1 + \frac{2\coth 1}{\sinh^2 1}\right)\int_1^t \int_1^\xi d\zeta d\xi = \left(1 + \frac{2\coth 1}{\sinh^2 1}\right)(t-1)^2,$$

and

$$\psi(t) \leq \frac{\sinh^2 1}{2\coth 1 + \sinh^2 1}\int_1^t \int_1^\xi \psi''(\xi)\psi''(\zeta)d\zeta d\xi = \frac{\sinh^2 1}{2\coth 1 + \sinh^2 1}\frac{\psi'(t)^2}{2},\ \forall t > 0.$$

This gives (i) by using the inequality $2\coth 1 > \sinh^2 1$.
For (ii), since $\psi(1) = \psi'(1) = 0$, and by using Taylor's theorem, we have

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

where $1 < \xi < t$. The result follows since $\psi'''(t) < 0$. □

Lemma 3.4. *Let* $\varrho : [0, +\infty[\ \to [1, +\infty[$ *be the inverse function of* $\psi(t)$ *for* $t \geq 1$ *and* $\rho : [0, +\infty[\ \to\ ]0, 1]$ *be the inverse function of* $-\frac{1}{2}\psi'(t)$ *for* $0 < t \leq 1$, *then*

(i) $1 + \sqrt{\dfrac{2z}{\psi''(1)}} \leq \varrho(z) \leq 1 + \sqrt{\dfrac{z}{2}},\ \forall z \in [0, +\infty[.$

(ii) $\coth t \leq \sqrt{2}(z+1)^{\frac{1}{3}},\ z = -\dfrac{1}{2}\psi'(t) \geq 0,\ \forall t \in]0, 1].$

Proof. For (i), let $z \geq 0$ and let $t \in [1, +\infty[$ such that $z = \psi(t)$, then $\varrho(z) = t$. By (ii) of Lemma 3.3, we have

$$z \leq \frac{1}{2}\psi''(1)(t-1)^2,$$

this implies that

$$t - 1 \geq \sqrt{\frac{2z}{\psi''(1)}}.$$

For the second hand, we use (i) of Lemma 3.3, we have $z = \psi(t) \geq 2(t-1)^2$, then we obtain

$$t \leq 1 + \sqrt{\frac{z}{2}}.$$

For (ii), let $z \geq 0$ and let $t \in]0, 1]$ such that $z = -\frac{1}{2}\psi'(t)$, for $t \in]0, 1]$, then $\rho(z) = t$. So, by (10), we have

$$2z = -\psi'(t) = -\left(1 + \frac{2\coth(1)}{\sinh^2(1)}\right)t + \frac{1}{t} + \frac{2\coth(t)}{\sinh^2(t)}.$$

Using the hyperbolic equation, we have

$$\begin{aligned} 2\coth t\left(\coth^2(t) - 1\right) &= \frac{2\coth t}{\sinh^2 t} = 2z + \frac{2\coth 1}{\sinh^2 1}t + t - \frac{1}{t} \\ &\leq 2\left(z + \frac{\coth 1}{\sinh^2 1}\right), \end{aligned}$$

this implies

$$\frac{\coth^3 t}{\cosh^2 t} \leq z + \frac{\coth 1}{\sinh^2 1}, \ \forall t \in]0,1].$$

Since cosh is positive and increasing in $]0,+\infty[$ then in $]0,1]$, we obtain

$$\coth^3 t \leq z \cosh^2 1 + \coth^3 1,$$

hence (ii) is obtained by observing that $\cosh^2 1 < 2\sqrt{2}$ and $\coth 1 < \sqrt{2}$. □

3.2. Growth behavior of the proximity function $\Psi(V)$

Note that, at the start of each outer iteration of the algorithm, just before the update of μ with the factor $(1-\theta)$, we have $\Psi(V) \leq \tau$. After updating μ in an outer iteration, all the eigenvalues of V are divided by the factor $\sqrt{1-\theta}$, which in general leads to an increase of the value of $\Psi(V)$. Then during the inner iteration, the value of the proximity function decreases until its value gets back to the situation in which $\Psi(V) \leq \tau$. In this section, we derive an estimate for the effect of updating the barrier parameter μ on the value of the proximity function during an iteration. We start with an important theorem which is valid for all kernel functions $\psi(t)$ that are strictly convex and satisfies (ii) and (iii) of Lemma 3.2.

Theorem 3.5. *([25]) Let ϱ be as defined in Lemma 3.4. Then, for any $V \in \mathcal{S}^n_{++}$ and $\beta > 1$, we have*

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

Corollary 3.6. *Let θ be such that $0 < \theta < 1$. If $\Psi(V) \leq \tau$, then*

$$\Psi(\beta V) \leq \frac{\psi''(1)}{4(1-\theta)}\left(\theta\sqrt{2n} + \sqrt{\tau}\right)^2, \beta = \frac{1}{\sqrt{1-\theta}} > 1.$$

Proof. Using (ii) of Lemma 3.3 for $t \geq 1$, we have

$$\psi(t) \leq \frac{1}{2}\psi''(1)(t-1)^2.$$

Consequently, by the above Theorem, we get

$$\Psi(\beta V) \leq \frac{n}{2}\psi''(1)\left(\beta\varrho\left(\frac{\Psi(V)}{n}\right) - 1\right)^2.$$

Hence, from (i) of Lemma 3.4, we have

$$\Psi(\beta V) \leq \frac{n}{2}\psi''(1)\left(\beta\left(1 + \sqrt{\frac{\frac{\Psi(V)}{n}}{2}}\right) - 1\right)^2.$$

Since $\beta = \frac{1}{\sqrt{1-\theta}}$, then

$$\begin{aligned}\Psi(\beta V) &\leq \frac{n\psi''(1)}{2(1-\theta)}\left(\sqrt{\frac{\Psi(V)}{2n}} + 1 - \sqrt{1-\theta}\right)^2 \\ &\leq \frac{\psi''(1)}{4(1-\theta)}\left(\theta\sqrt{2n} + \sqrt{\Psi(V)}\right)^2.\end{aligned}$$

The last inequality is obtained from the fact that $1 - \sqrt{1-\theta} = \dfrac{\theta}{1+\sqrt{1-\theta}} \leq \theta$.

By the assumption $\Psi(V) \leq \tau$ just before the μ−update and from (11), we have

$$\Psi(\beta V) \leq \frac{\psi''(1)}{4(1-\theta)}\left(\theta\sqrt{2n} + \sqrt{\tau}\right)^2 =: \Psi_0. \tag{16}$$

□

Obviously, Ψ_0 is an upper bound of $\Psi(\beta V)$, the value of $\Psi(V)$ after the $\mu-$update.

The following Lemma gives a lower bound on the norm-based proximity measure $\sigma(V)$, in terms of $\Psi(V)$.

Lemma 3.7. *Let $\sigma(V)$ be defined by* (8)*. Then, for any $V \in \mathcal{S}^n_{++}$, we have*

$$\sigma(V) \geq \sqrt{\Psi(V)}.$$

Proof. From (i) of Lemma 3.3, we have for all $t > 0$

$$\psi(t) \leq \frac{(\psi'(t))^2}{4},$$

using (8) and (5), we obtain

$$\sigma(V)^2 = \frac{1}{4}\mathrm{tr}\left(\psi'(V)^2\right) = \frac{1}{4}\sum_{i=1}^{n}\psi'(\lambda_i(V))^2 \geq \sum_{i=1}^{n}\psi(\lambda_i(V)).$$

□

Remark 3.8. *Throughout the paper, we assume that $\tau \geq 1$. Using the above lemma and the assumption that $\Psi(V) \geq \tau$, we have $\sigma(V) \geq 1$.*

4. Analysis and complexity of the algorithm

4.1. Computation of displacement step

The choice of the step size α is another crucial issue in the analysis of the algorithm. In this section, we compute a default step size α, such that the proximity function is decreasing during an inner iteration and give the complexity results of the algorithm.
Using (7), (4) and (2) for fixed μ, we get

$$X_+ = X + \alpha\Delta X = X + \alpha\sqrt{\mu}P^{-1}D_XP^{-1} = \sqrt{\mu}P^{-1}(V + \alpha D_X)P^{-1},$$

$$S_+ = S + \alpha\Delta S = S + \alpha\sqrt{\mu}PD_SP = \sqrt{\mu}P(V + \alpha D_S)P.$$

So, by (2), we have $V_+ = V + \alpha D_X = V + \alpha D_S$.
This implies that the eigenvalues of V_+ are the same as those of the matrix $((V+\alpha D_X)^{\frac{1}{2}}(V+\alpha D_S)(V+\alpha D_X)^{\frac{1}{2}})^{\frac{1}{2}}$.
Let $f(\alpha) = \Psi(V_+) - \Psi(V) = \Psi\left(\left((V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_S)(V + \alpha D_X)^{\frac{1}{2}}\right)^{\frac{1}{2}}\right) - \Psi(V)$.
Throughout the paper, we assume that the step size α satisfied

$$V + \alpha D_X \in \mathcal{S}^n_{++} \quad \text{and} \quad V + \alpha D_S \in \mathcal{S}^n_{++}.$$

Due to the following proposition which is a consequence of the exponentially convexity property of ψ

Proposition 4.1. *([9]) For any $X_1, X_2 \in \mathcal{S}^n_{++}$, we have*

$$\Psi\left(\left(X_1^{\frac{1}{2}}X_2X_1^{\frac{1}{2}}\right)^{\frac{1}{2}}\right) \leq \frac{1}{2}\left(\Psi(X_1) + \Psi(X_2)\right),$$

it follows that

$$\Psi(V_+) \leq \frac{1}{2}\left(\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)\right).$$

Therefore, $f(\alpha) \leq f_1(\alpha)$, where

$$f_1(\alpha) = \frac{1}{2}\left(\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)\right) - \Psi(V). \tag{17}$$

Obviously, we have $f(0) = f_1(0) = 0$. Taking the first two derivatives of $f_1(\alpha)$ with respect to α, we get

$$f_1'(\alpha) = \frac{1}{2}\mathrm{tr}\left(\psi'(V + \alpha D_X)D_X + \psi'(V + \alpha D_S)D_S\right),$$

$$f_1''(\alpha) = \frac{1}{2}\mathrm{tr}\left(\psi''(V + \alpha D_X)D_X^2 + \psi''(V + \alpha D_S)D_S^2\right).$$

This gives, using the last equality of (6) and from (8)

$$f_1'(0) = -\frac{1}{2}\mathrm{tr}\left((\psi'(V))^2\right) = -2\sigma(V)^2. \tag{18}$$

Below we use the following notation, $\sigma(V) := \sigma$.
From Lemmas 4.1-4.4 in [4], we have the following Lemmas 4.2-4.4, since $\psi(t)$ is a kernel function and $\psi''(t)$ is monotonically decreasing (see also [6, 9, 25]).

Lemma 4.2. *Let $f_1(\alpha)$ as defined in* (17), *one has*

$$f_1''(\alpha) \leq 2\sigma^2\psi''(\lambda_{\min}(V) - 2\alpha\sigma).$$

Lemma 4.3. *We have $f_1'(\alpha) \leq 0$ if the step size α satisfies the inequality*

$$\psi'(\lambda_{\min}(V)) - \psi'(\lambda_{\min}(V) - 2\alpha\sigma) \leq 2\sigma. \tag{19}$$

Lemma 4.4. *Let ρ be as defined in Lemma* 3.4. *Then the step size*

$$\alpha^* := \frac{\rho(\sigma) - \rho(2\sigma)}{\sigma},$$

is the largest possible solution of inequality (19). *And then,*

$$\alpha^* \geq \frac{1}{\psi''(\rho(2\sigma))}.$$

Now, we give a suitable step-size α for Algorithm 1.

Lemma 4.5. *Let ρ and α^* be as defined in Lemma* 4.4. *If $\Psi(V) \geq \tau \geq 1$, then we have*

$$\alpha^* \geq \frac{1}{8 + 24(2\sigma+1)^{\frac{4}{3}} + 8(2\sigma+1)^{\frac{2}{3}}}.$$

Proof. Using (11), (14) and (ii) of Lemma 3.4 with $t = \rho(2\sigma) \in]0, 1]$, we obtain

$$\begin{aligned}
\psi''(t) &= \left(1 + \frac{2\coth 1}{\sinh^2 1}\right) + 2\left(\frac{3\coth^2 t - 1}{\sinh^2 t}\right) + \frac{1}{t^2} \\
&= 1 + 2\coth 1(\coth^2 1 - 1) + 2(3\coth^2 t - 1)(\coth^2 t - 1) + \frac{1}{t^2} \\
&\leq 1 + 2\coth 1(\coth^2 1 - 1) + 2(3\coth^2 t - 1)(\coth^2 t - 1) + 4\coth^2(t) \\
&\leq 1 + 2\coth^3 1 + 6\coth^4 t + 4\coth^2 t \\
&\leq 1 + 4\sqrt{2} + 6\coth^4 t + 4\coth^2 t \\
&\leq 8 + 24(2\sigma+1)^{\frac{4}{3}} + 8(2\sigma+1)^{\frac{2}{3}},
\end{aligned}$$

application of Lemma 4.4 yields the desired inequality. □

Denoting

$$\bar{\alpha} = \frac{1}{8 + 24(2\sigma+1)^{\frac{4}{3}} + 8(2\sigma+1)^{\frac{2}{3}}}, \tag{20}$$

we have that $\bar{\alpha}$ is the default step size in the algorithm and that $\bar{\alpha} \leq \alpha^*$.

4.2. Decreasing of $\Psi(V)$ during an inner iteration

Lemma 4.6. *([25]) Let g be a twice differentiable convex function with $g(0) = 0$, $g'(0) < 0$, which attains its (global) minimum at $t^* > 0$. If g'' is increasing for $t \in [0, t^*]$, then*

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

As a direct consequence, we get the following lemma.

Lemma 4.7. *If the step size α is such that $\alpha \leq \alpha^*$, then*

$$f(\alpha) \leq -\alpha\sigma^2.$$

Proof. It's easy to verify that $f_1(\alpha)$ satisfies the conditions of Lemma 4.6, then

$$f(\alpha) \leq f_1(\alpha) \leq \frac{f_1'(0)}{2}\alpha, \quad \text{for} \quad \text{all} \quad 0 \leq \alpha \leq \alpha^*.$$

The result follows from (18). □

We can obtain the upper bound for the decreasing value of the proximity in the inner iteration by the following theorem.

Theorem 4.8. *Let $\bar{\alpha}$ as in* (20) *and $\Psi(V) \geq 1$. Then*

$$f(\bar{\alpha}) \leq -\frac{(\Psi(V))^{\frac{1}{3}}}{130}. \tag{21}$$

Proof. Since $\bar{\alpha} \in [0, \alpha^*]$, we obtain thanks to Lemma 4.7 and Remark 3.8

$$\begin{aligned} f(\bar{\alpha}) &\leq -\frac{\sigma^2}{8\left(1 + 3(2\sigma+1)^{\frac{4}{3}} + (2\sigma+1)^{\frac{2}{3}}\right)} \\ &= -\frac{\sigma^2}{8\left(\frac{1}{\sigma^{\frac{4}{3}}} + 3\left(2 + \frac{1}{\sigma}\right)^{\frac{4}{3}} + \left(\frac{2}{\sigma} + \frac{1}{\sigma^2}\right)^{\frac{2}{3}}\right)\sigma^{\frac{4}{3}}} \\ &\leq -\frac{\sigma^{\frac{2}{3}}}{8\left(\frac{1}{\sigma^{\frac{4}{3}}} + 3\left(2 + \frac{1}{\sigma}\right)^{\frac{4}{3}} + \left(\frac{2}{\sigma} + \frac{1}{\sigma^2}\right)^{\frac{2}{3}}\right)} \\ &\leq -\frac{\sigma^{\frac{2}{3}}}{8\left(1 + 3(3)^{\frac{4}{3}} + 3^{\frac{2}{3}}\right)} \leq -\frac{\sigma^{\frac{2}{3}}}{130}. \end{aligned}$$

Hence, from Lemma 3.7, we get

$$\begin{aligned} f(\bar{\alpha}) &\leq -\frac{\left(\sqrt{\Psi(V)}\right)^{\frac{2}{3}}}{130} \\ &= -\frac{(\Psi(V))^{\frac{1}{3}}}{130}, \end{aligned}$$

which completes the proof. □

4.3. Iteration bound of our algorithm

We need to count how many inner iterations are required to return to the situation where $\Psi(V) \leq \tau$ after $\mu-$update. We define the value of $\Psi(V)$ after $\mu-$update as Ψ_0, and the subsequent values in the same outer iteration are denoted as $\Psi_k, k = 1, ..., K$, where K denotes the total number of inner iterations in the outer iteration. By the definition of $f(\alpha)$ and according to (21), for $k = 0, ..., K-1$, we obtain

$$\Psi_{k+1} \leq \Psi_k - \frac{(\Psi_k(V))^{\frac{1}{3}}}{130}.$$

Lemma 4.9. *([9]) Suppose $t_0, t_1, ..., t_k$ be a sequence of positive numbers such that*

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

where $\beta > 0$ and $0 < \gamma \leq 1$. Then $K \leq \left[\frac{t_0^{\gamma}}{\beta\gamma}\right]$.

Letting $t_k = \Psi_k$, $\beta = \dfrac{1}{130}$ and $\gamma = \dfrac{2}{3}$, we can get that the total number of inner iterations K in the outer iteration satisfies

$$K \leq 195\ \Psi_0^{\frac{2}{3}},$$

where Ψ_0 is the value of $\Psi(V)$ after the $\mu-$update in an outer iteration.
Now, we derive the complexity bounds for large and small-update methods.

Theorem 4.10. *Let Ψ_0 be the value defined in* (16) *and let $\tau \geq 1$. Then, the total number of iterations to obtain an approximation solution with $n\mu \leq \epsilon$ is bounded by*

$$\frac{K}{\theta} \log \frac{n}{\epsilon} \leq \left[195\ \Psi_0^{\frac{2}{3}}\right]\left[\frac{1}{\theta} \log \frac{n}{\epsilon}\right].$$

Proof. Recall that Ψ_0 is the upper bound according to (16). An upper bound for the total number of iterations is obtained by multiplying the upper bound K by the number of outer iterations, which is bounded above by $\frac{1}{\theta}(\log \frac{n}{\epsilon})$ (see [9]), that gives the result thanks to the above lemma. □

Large-update methods use $\tau = O(n)$ and $\theta = \Theta(1)$. We then easily verify that the right-hand side expression is

$$O\left(n^{\frac{2}{3}} \log \frac{n}{\epsilon}\right).$$

For small-update methods one has $\tau = O(1)$ and $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$. The iteration bound then becomes

$$O\left(\sqrt{n} \log \frac{n}{\epsilon}\right).$$

5. Conclusions and remarks

In this paper, we have introduced a new kernel function, which is a combination of the classic kernel function and a hyperbolic barrier term, and presented various properties of this new kernel function. We have showed that the large-update IPM based on this kernel function has $O\left(n^{\frac{2}{3}} \log \frac{n}{\epsilon}\right)$ iteration bound in the worst-case. This bound improves the complexity of the IPMs based on hyperbolic as well as classic kernel functions significantly, but still needs to be improved so that it meets at least the so far best known complexity result for large-update IPMs, i.e., $O\left(\sqrt{n} \log n \log \frac{n}{\epsilon}\right)$. For small-update methods, we have obtained $O\left(\sqrt{n} \log \frac{n}{\epsilon}\right)$ iteration bound which matches the currently best known iteration bound.

Some interesting topics for forthcoming works are: The extension to linear programming (LP), second-order cone programming (SOCP), and Convex quadratic semidefinite programming (CQSDP). Find a parameterized hyperbolic kernel function to improve the complexity bound for large-update IPMs. Furthermore, numerical tests to investigate the behavior of the algorithm so as to be compared with other approaches.

References

[1] F. Alizadeh, Combinatorial optimization with interior point methods and semi-definite matrices, Ph. D. thesis, University of Minnesota, Minneapolis, MN, Oct. (1991).

[2] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, SIAM J. Optim. 5 (1995) 13-51.

[3] Y.Q. Bai, M. EL. Ghami, C. Roos, A new efficient large-update primal-dual interior-point method based on a finite barrier. SIAM J. Optim. 13 (2003) 766-782.

[4] Y.Q. Bai, M. EL. Ghami, C. Roos, A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. SIAM J. Optim. 15 (2004) 101-128.

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

[6] Y.Q. Bai, C. Roos, A primal-dual interior-point method based on a new kernel function with linear growth rate. Proceedings of Industrial Optimization Symposium and Optimization Day. Australia, November (2002).

[7] M. Bouafia, D. Benterki, A. Yassine, An Efficient Primal-Dual Interior Point Method for Linear Programming Problems Based on a New Kernel Function with a Trigonometric Barrier Term. J. Optim. Theory Appl. 170 (2016) 528-545.

[8] X.Z. Cai, G.Q. Wang, M. El Ghami, Y.J. Yue, Complexity analysis of primal-dual interior-point methods for linear optimization based on a new parametric kernel function with a trigonometric barrier term. Abstr. Appl. Anal. Art. ID 710158. 11 (2014).

[9] B.K. Choi, G.M. Lee, On complexity analysis of the primal-dual interior-point method for semidefinite optimization problem based on a new proximity function. Nonlinear Anal. 71 (2009) 2628-2640.

[10] M. El Ghami, I.D. Ivanov, C. Roos, T. Steihaug, A polynomial-time algorithm for LO based on generalized logarithmic barrier functions. Int. J. Appl. Math. 21 (2008) 99-115.

[11] M. El Ghami, Z.A. Guennoun, S. Bouali, T. Steihaug, Interior point methods for linear optimization based on a kernel function with a trigonometric barrier term. J. Comput. Appl. Math. 236 (2012) 3613-3623.

[12] B. Kheirfam, Primal-dual interior-point algorithm for semidefinite optimization based on a new kernel function with trigonometric barrier term, Numer. Algorithms. 61 (2012) 659-680.

[13] X. Li, M. Zhang, Interior-point algorithm for linear optimization based on a new trigonometric kernel function. Operations Research Letters. 43 (2015) 471-475.

[14] R.E. Nesterov, A.S. Nemirovskii, Interior-Point Ploynomial Methods in Convex Programming. Society for Industrial and Applied Mathematics (1994).

[15] J. Peng, C. Roos, T. Terlaky, A new and efficient large-update interior point method for linear optimization. J. Comput. Technol. 6 (2001) 61-80.

[16] J. Peng, C. Roos, T. Terlaky, Self-regular functions and new search directions for linear and semidefinite optimization. Math. Program. 93 (2002) 129-171.

[17] J. Peng, C. Roos, T. Terlaky, Self-regularity: A new Paradigm for Primal-Dual Interior-Point Algorithms. Princeton University Press, Princeton, NJ (2002).

[18] M.R. Peyghami, S.F. Hafshejani, L. Shirvani, Complexity of interior point methods for linear optimization based on a new trigonometric kernel function. J. Comput. Appl. Math. 255 (2014) 74-85.

[19] M.R. Peyghami, S.F. Hafshejani, S. Chen, A primal-dual interior-point method for semidefinite optimization based on a class of trigonometric barrier functions, Operations Research Letters. 44 (2016) 319-323.

[20] Z.G. Qian, Y.Q. Bai, G.Q. Wang, Complexity analysis of interior-point algorithm based on a new kernel function for semi-definite optimization. J. Shanghai Univ (Engl Ed). 12(5) (2008) 388-394.

[21] C. Roos, T. Terlaky, J.Ph. Vial, Theory and algorithms for linear optimization, in: An interior point Approach. John Wiley & Sons, Chichester, UK (1997).

[22] M.J. Todd, K.C. Toh, R. H. Tütüncü, On the Nestrov-Todd direction in semidefinite programming. SIAM J. Optim. 8 (1998) 769-796.

[23] I. Touil, D. Benterki, A. Yassine, A feasible primal-dual interior point method for linear semidefinite programming, J. Comput. Appl. Math. 312 (2017) 216-230.

[24] I. Touil, W. Chikouche, Novel kernel function with a hyperbolic barrier term to primal-dual interior point algorithm for SDP problems. Submitted for publication.

[25] G.Q. Wang, Y.Q. Bai, C. Roos, Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function. Journal Math. Model. Algorithms. 4 (2005) 409-433.

[26] H. Wolkowicz, R. Saigal, L. Vandenberghe, Handbook of Semidefinite Programming, Theory, Algorithms, and Applications, Kluwer Academic Publishers (2000).