Acta Mathematicae Applicatae Sinica, English Series
Vol. 36, No. 3 (2020) 753–770
https://doi.org/10.1007/s10255-020-0955-y
http://www.ApplMath.com.cn & www.SpringerLink.com

Acta Mathematicae Applicatae Sinica, English Series

An Efficient Parameterized Logarithmic Kernel Function for Semidefinite Optimization

Louiza DERBAL†, Zakia KEBBICHE

University of Ferhat Abbas Setif 1, Algeria (†E-mail: louiza.derbal@univ-setif.dz)

Abstract In this paper, we present a primal-dual interior point algorithm for semidefinite optimization problems based on a new class of kernel functions. These functions constitute a combination of the classic kernel function and a barrier term.

We derive the complexity bounds for large and small-update methods respectively. We show that the best result of iteration bounds for large and small-update methods can be achieved, namely $O(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon})$ for large-update methods and $O(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon})$ for small-update methods.

We test the efficiency and the validity of our algorithm by running some computational tests, then we compare our numerical results with results obtained by algorithms based on different kernel functions.

Keywords kernel function; interior-point algorithms; semidefinite optimization; complexity bound; primal-dual methods

2000 MR Subject Classification 90C22; 90C31; 90C51

1 Introduction

Primal-dual interior-point methods (IPMs) have been well known as the most effective methods for solving wide classes of optimization problems, for example, linear optimization (LO), quadratic optimization problem (QOP), semidefinite optimization (SDO), second-order cone optimization (SOCO) and symmetric optimization (SO).

SDO has wide applications in continuous and combinatorial optimization [1]. It has become a popular research area in mathematical programming when it is clear that the IPMs for LO can often be extended to the more general SDO case. Many researchers have studied SDO and achieved plentiful and beautiful results. For an overview of these results we refer to the book [27] and its references. Among them, IPMs based on kernel functions have been designed.

Recently, a new variant of primal-dual IPM for LO and SDO based on the so-called self-regular (SR) barrier functions was presented by Peng et al. [16] and aforementioned gap was narrowed. Each such barrier function is determined by its (univariate) self-regular kernel function. They obtained so far the best iteration bound, namely, $O(\sqrt{n}\log n\log\frac{n}{\epsilon})$, for large-update IPMs for LO and SDO. Subsequently, Bai et al. [3, 4] developed a class of primal-dual IPMs for LO based on non-self-regular barrier functions and obtained the same favorable iteration bounds for the algorithms with large-update strategy as [16]. Similar algorithms for SDO also were presented (see [7, 26]).

Manuscript received April 30, 2019. Accepted on January 31, 2020.
†Corresponding author.

Table 1. Complexity Results for the Eligible Kernel Functions

i	Kernel functions $\psi_i(t)$	Large-update	Small-update	Ref
1	$\frac{1-t^2}{2} - \log t$	$O(n \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[20]
2	$\frac{1}{2}(t - \frac{1}{t})^2$	$O(n^{\frac{2}{3}} \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[4]
3	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1}, q > 1$	$O(qn^{\frac{q+1}{2q}} \log \frac{n}{\varepsilon})$	$O(q^2\sqrt{n} \log \frac{n}{\varepsilon})$	[4]
4	$\frac{t^2-1}{2} + \frac{t^{1-q}-1}{q(q-1)} - \frac{q-1}{q}, q > 1$	$O(qn^{\frac{q+1}{2q}} \log \frac{n}{\varepsilon})$	$O(q^2\sqrt{n} \log \frac{n}{\varepsilon})$	[16]
5	$\frac{t^2-1}{2} + \frac{e^{\frac{1}{t}-1}-1}{e}$	$O(\sqrt{n}(\log n)^2 \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[4]
6	$\frac{t^2-1}{2} - \int_1^t e^{\frac{1}{\xi}-1} d\xi$	$O(\sqrt{n}(\log n)^2 \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[4]
7	$\frac{t^2-1}{2} - \frac{e^{q(\frac{1}{t}-1)}-q}{q}, q \geq 1$	$O(q\sqrt{n} \log \frac{n}{\varepsilon})$	$O(q\sqrt{qn} \log \frac{n}{\varepsilon})$	[2]
8	$\frac{t^2-1}{2} - \int_1^t e^{q(\frac{1}{\xi}-1)} d\xi, q \geq 1$	$O(q\sqrt{n} \log \frac{n}{\varepsilon})$	$O(q\sqrt{qn} \log \frac{n}{\varepsilon})$	[4]
9	$\frac{t^2-1}{2} + \frac{6}{8} \tan(\frac{\pi(1-t)}{4t+2})$	$O(n^{\frac{3}{4}} \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[11]
10	$\frac{t^2-1}{2} - \log t + \frac{1}{8} \tan^2(\frac{\pi(1-t)}{4t+2})$	$O(n^{\frac{2}{3}} \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[19]
11	$\frac{t^2-1}{2} - \int_1^t e^{3(\tan(\frac{\pi}{2\xi+2})-1)} d\xi$	$O(\sqrt{n}(\log n)^2 \log \frac{n}{\varepsilon})$	$O(\sqrt{n} \log \frac{n}{\varepsilon})$	[18]
12	$t - 1 + \frac{t^{1-q}-1}{q-1}, q > 1$	$O(qn \log(\frac{n}{\varepsilon}))$	$O(q^2\sqrt{n} \log \frac{n}{\varepsilon})$	[4]
13	$\begin{cases} \frac{t^{p+1}-1}{p+1} + \frac{t^{1-q}-1}{q-1}, & t > 0, p \in [0,1], q > 1 \\ \frac{t^{p+1}-1}{p+1} - \log t, & t > 0, p \in [0,1], q = 1 \end{cases}$	$O(qn^{\frac{(p+q)}{(q(1+p))}} \log(\frac{n}{\varepsilon}))$	$O(q^2\sqrt{n} \log \frac{n}{\varepsilon})$	[5]

In this paper, we define a new familly of kernel functions for the SDO defined by:

$$\psi(t) = \frac{t^2 - 1 - \log(t)}{2} + \frac{e^{\frac{1}{t^q}-1} - 1}{2q} \qquad \text{for} \qquad t > 0, q \geq 1. \tag{1}$$

This kernel function is a combination of the classic kernel function and a barrier term. This work constitute an extension of the work presented for LO by the same authors L. Derbal et al. in [9]. We show that the best result of iteration bounds for large and small-update methods based on this new kernel function can be achieved, namely $O(q\sqrt{n}(\log \sqrt{n})^{\frac{q+1}{q}} \log \frac{n}{\varepsilon})$ for large-update methods and $O(q^{\frac{3}{2}}(\log \sqrt{q})^{\frac{q+1}{q}} \sqrt{n} \log \frac{n}{\varepsilon})$ for small-update methods. Moreover, we report some numerical results to confirm the validity of this approach and to compare our obtained results with those obtained by methods based on different kernel functions.

The paper is organized as follows. In Section 2, the SDO problem is presented and the primal-dual IPM is described, we also define the new kernel function and its properties. In section 3, we establish the complexity analysis and the iteration complexity bounds for large and small-update methods. We provide our numerical resuls in Section 4. Finally, concluding remarks and future work are presented in Section 5.

Throughout this paper. $\mathbb{R}^n$, $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ denote the set of vectors with n components, the set of nonnegative vectors and the set of positive vectors, respectively. $\|.\|$ denotes the Frobenius norm for matrices, and the 2-norm for vectors. S^n, S^n_+ and S^n_{++} denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n \times n$ matrices, respectively. Furthermore, $A \succeq 0$ ($A \succ 0$) means that $A \in S^n_+$ ($A \in S^n_{++}$). We use the matrix inner product, i.e., $A \cdot B := tr(A^T B) = \sum_{i,j} A_{ij} B_{ij}$. For any $Q \in S^n_{++}$, the expression $Q^{\frac{1}{2}}$ denotes its symmetric square root. For any $V \in S^n_{++}$, we assume that the eigenvalues of V are arranged in non-increasing order, that is, $\lambda_1(V) \geq \lambda_2(V) \geq \cdots \geq \lambda_n(V)$. If $g(x) \geq 0$ is a real valued function of real nonnegative variable, the notation $g(x) = O(x)$ means that $g(x) \leq \bar{c}x$ for some positive constant $\bar{c}$ and $g(x) = \Theta(x)$ that $c_1 x \leq g(x) \leq c_2 x$ for two positive constants c_1 and c_2.

2 The SDO Problem

In this section, we first introduce some useful results regarding the properties of the symmetric matrices.Then some concepts of central path for SDO problems are briefly recalled. The structure of generic primal-dual IPMs based on kernel function is also given in this section. Now, we recall some well known facts from linear algebra which are essential in our analysis. Consider the SDO problem as:

$$(\text{SDO}) \begin{cases} \min \operatorname{tr}(CX) \\ \text{subject to } \operatorname{tr}(A_i X) = b_i, \quad i = 1, 2, \cdots, m, \quad X \succeq 0, \end{cases}$$

and its dual problem:

$$(\text{SDD}) \begin{cases} \max b^T y \\ \text{subject to } \sum_{i=1}^{m} y_i A_i + S = C, \qquad S \succeq 0, \end{cases}$$

where each $A_i \in S^n$, $b = (b_1, b_2, \cdots, b_m)^T \in R^m$ and $C \in S^n$. Moreover, the matrices A_i are linearly independent, with $\mathbf{y} \in R^m$ and $S \in S^n$. Also, we assume that every point (X, y, S) is strictly feasible, that is, SDO and SDD satisfy the interior-point condition (IPC). We have the following lemma which is well known:

Lemma 2.1 ([8])**.** *The following statements are equivalent:*
(i) $X \geq 0,\ S \geq 0$ *and* $tr(XS) = 0$;
(ii) $X \geq 0,\ S \geq 0$ *and* $\|X^{\frac{1}{2}} S^{\frac{1}{2}}\|^2 = 0$;
(iii) $X \geq 0,\ S \geq 0$ *and* $XS = 0$.

Using lemma 2.1, we can easily check that a pair of optimal solution of SDO and SDD is equivalent to solving the following Newton system

$$\begin{cases} \operatorname{tr}(A_i X) = b_i \ , & i = 1, 2, \cdots, m, \\ \sum_{i=1}^{m} y_i A_i + S = C, & \\ XS = 0, & X, S \succeq 0. \end{cases}$$

In a similar way to the LO case, we can rewrite the above system as

$$\begin{cases} \operatorname{tr}(\text{AiX}) = b_i \ , & i = 1, 2, \cdots, m, \\ \sum_{i=1}^{m} y_i A_i + S = C, & \\ XS = \mu E, & X, S \succeq 0. \end{cases} \tag{2}$$

where E denotes the $n \times n$ identity matrix and $\mu > 0$ is a parameter. If both SDO and SDD satisfy IPC, then for each μ the system (2) has a unique solution $(X(\mu), y(\mu), S(\mu))$. The set of μ centers gives an homotopy path, which is called the central path of (P) and (D)**.** The basic idea of IPMs is to follow this path, if $\mu \to 0$ then the limit of the central path exists (see [13, 15]) .

In general, IPMs for the SDO consist of two strategies: The first one, which is called the inner iteration scheme, is to keep the iterative sequence in a certain neighborhood of the central path or to keep the iterative sequence in a certain neighborhood of the center and the second one is called the outer iteration scheme, is to decrease the parameter μ to $\mu_+ = (1 - \theta)\mu$, for some $\theta \in (0, 1)$.

2.1 Proximity Functions and Search Directions

Newton's method is a well-known procedure to solve a system of nonlinear equations. Suppose that the point (X, y, S) is strictly feasible. Newton method amounts to linearizing the system (2), thus yielding the following system:

$$\begin{cases} \operatorname{tr}(A_i \Delta X) = 0, \qquad i = 1, 2, \cdots, m, \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta S = 0 \ , \\ X\Delta S + \Delta X S = \mu E - XS, \qquad\qquad X, S \succeq 0. \end{cases} \tag{3}$$

A decisive observation for SDO is that the above Newton system might have no symmetric solution ΔX. Several ways for symmetrizing the third equation in the Newton system are proposed, such that the resulting new system has a unique symmetric solution [22, 23]. Y. Zhang [29] proposed the symmetrization scheme

$$H_p(XS) = \frac{1}{2}(PXSP^{-1} + P^{-T}XSP^T)$$

where P is a general nonsingular matrix. We use the Nesterov-Tood [NT] direction [22], so $P := X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} = S^{-\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})S$, we can derive a new linearized Newton equation as follows

$$\begin{cases} \operatorname{tr}(A_i \Delta X) = 0, \qquad i = 1, 2, \cdots, m, \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta S = 0 \ , \\ \Delta X + P\Delta S P^T = \mu S^{-1} - X, \qquad\qquad X, S \succeq 0. \end{cases} \tag{4}$$

Now ΔX is automatically a symmetric matrix. Let $D = P^{\frac{1}{2}}$ witch can be used to scale X and S to the same matrix V, namely [8, 21]

$$V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} DSD. \tag{5}$$

Obviously the matrices D and V are symmetric and positive definite. Let us further define

$$\overline{A}_i := \frac{1}{\sqrt{\mu}} D A_i D, \qquad i = 1, 2, \cdots, m,$$

$$D_{\mathbf{X}} := \frac{1}{\sqrt{\mu}} D^{-1} \Delta X D^{-1}, \qquad D_{\mathbf{S}} := \frac{1}{\sqrt{\mu}} D \Delta S D. \tag{6}$$

Then it follows from (4) that the [NT] search direction $(D_{\mathbf{X}}, \Delta y, D_{\mathbf{S}})$ (See [17]) is obtained from the system

$$\begin{cases} tr(\overline{A}_i D_{\mathbf{X}}) = 0, \qquad\qquad i = 1, 2, \cdots, m \\ \sum_{i=1}^{n} \Delta y_i \overline{A}_i + D_{\mathbf{S}} = 0 \\ D_X + D_S \ = V^{-1} - V. \end{cases} \tag{7}$$

We can say that $\operatorname{tr}(D_X \ D_S) = 0$, which is coming from the first and second equations of (7) or from the orthogonality of ΔX and ΔS.

Let us recall some basics notions of linear algebra.

Theorem 2.2. *(Spectral theorem for symmetric matrices[27]). The real $n \times n$ matrix A is symmetric if and only if there exists a matrix $Q \in R^{n\times n}$ such that $Q^TQ = E$ and $Q^TAQ = \Lambda$ where Λ is a diagonal matrix.*

Now we are ready to show how a matrix function can be obtained from $\psi(t)$ as defined by (1).

Definition 2.3. *Let $V \in S^n$ and $V = Q_V^T \operatorname{diag}(\lambda(V))Q_V$ where Q_V is any orthonormal matrix that diagonalizes V. Let $\psi(t)$ be defined in (1). Then the matrix function $\psi(V) : S^n \to S^n$ is defined by*

$$\psi(V) = Q_V^T \operatorname{diag}(\psi(\lambda_1(V)), \psi(\lambda_2(V)), \cdots, \psi(\lambda_n(V)))Q_V. \tag{8}$$

Note that $\psi(V)$ depends only on the restriction of $\psi(t)$ to the set of eigenvalues of V. Since $\psi(t)$ is differentiable, the derivative $\psi(t)$ is well defined for $t > 0$. Hence, replacing $\psi(\lambda_i(V))$ in (8) by $\psi'(\lambda_i(V))$, we obtain that the matrix function $\psi'(V)$ is defined as well. Using ψ, we define the barrier function.

Definition 2.4. $\Psi(V) : S^n_+ \to \mathbb{R}$ *such that:*

$$\Psi(V) := \operatorname{tr}(\psi(\mathbf{V})) = \sum_{i=1}^{n} \psi(\lambda_i(V)), \tag{9}$$

As in the linear case, we can call $\psi(t)$ the kernel function for the matrix function $\psi(V)$ and $\Psi(V)$. When the function $\psi(t)$ is triple differentiable, the derivatives $\psi'(t)$,$\psi''(t)$, and $\psi'''(t)$ are well-defined, and we can define $\psi'(V)$,$\psi''(V)$, and $\psi'''(V)$ by replacing $\psi(\lambda_i(V))$ in (8) by $\psi'(\lambda_i(V))$, $\psi''(\lambda_i(V))$, and $\psi'''(\lambda_i(V))$, respectively.

Definition 2.5. *A matrix $X(t)$ is said to be a matrix of function (or a matrix-valued function) if each entry of $X(t)$ is a function of t, i.e., $X(t) = [X_{ij}(t)]$.*

The usual concepts of continuity, differentiability, integrability and some basic rules in calculus can be extended to matrix-valued functions by interpreting them entry-wise. Suppose that $X(t)$, $G(t)$, and $H(t)$ are all matrices of functions. And we denote by $X'(t)$ the derivative to t of the matrix of function $X(t)$. Then, we have

$$X'(t) = \frac{d}{dt}X(t) = \Big(\frac{d}{dt}X_{ij}(t)\Big), \tag{10}$$

$$\frac{d}{dt}\operatorname{tr}(X(t)) = \operatorname{tr}\Big(\frac{d}{dt}X(t)\Big) = \operatorname{tr}(X'(t)), \tag{11}$$

$$\frac{d}{dt}\operatorname{tr}(\psi(X(t))) = \operatorname{tr}(\psi'(X(t))X'), \tag{12}$$

$$\begin{aligned}\frac{d}{dt}(X(t)G(t)) &= \Big(\frac{d}{dt}X(t)\Big)G(t) + X(t)\Big(\frac{d}{dt}G(t)\Big),\\ &= X'(t)G(t) + X(t)G'(t).\end{aligned} \tag{13}$$

In fact, the right-hand side of the third equation in (7) is the negative gradient of the matrix function $\psi_c(V)$ with the classical kernel function $\psi_c(t) = \frac{t^2-1}{2} - \log(t)$. However, in this paper we consider the kernel function (1). As in the linear case we have the following lemma:

Lemma 2.6. $\Psi(V)$ *is strictly convex with respect to $V \succ 0$ and vanishes at its global minimal point $X = E$, i.e., $\psi(E) = \psi'(E) = 0_{n\times n}$.*

We replace the right-hand side of the third equation in (7) by $-\psi'(V)$. Thus the above system can be rewriten as

$$\begin{cases} \overline{A}_i \cdot D_X = 0, \quad i = 1, 2, \cdots, m \\ \sum_{i=1}^{n} \Delta y_i \overline{A}_i + D_S = 0, \\ D_X + D_S = -\psi'(\mathbf{V}). \end{cases} \tag{14}$$

Since D_X and D_S are orthogonal due to the orthogonality of ΔX and ΔS, it is trivial to verify that $tr(D_X D_S) = tr(D_X D_S) = 0$. Then we have

$$D_X = D_S = 0_{n\times n} \Leftrightarrow \psi'(V) = 0_{n\times n} \Leftrightarrow V = E \quad \Leftrightarrow \Psi(V) = 0,$$

i.e., if and only if $XS = \mu E$, that is, if and only if $X = X(\mu)$ and $S = S(\mu)$ as it should. Otherwise $\Psi(V) > 0$, hence, if $(X, y, S) \neq (X(\mu), y(\mu), S(\mu))$, then $(\Delta X, \Delta y, \Delta S)$ is nonzero. By taking a step along the search direction, with the step size α defined by some line search rules, one constructs a new triple (X_+, y_+, S_+) according to

$$X_+ = X + \alpha \Delta X, \qquad y_+ = y + \alpha \Delta y, \qquad S_+ = S + \alpha \Delta S. \tag{15}$$

We can now describe the algorithm in a more formal way. The generic form of this algorithm is shown below

Generic Algorithm

Input:
a kernel function $\psi(t)$;
a threshold parameter $\tau > 1$;
an accuracy parameter $\varepsilon > 0$;
a fixed barrier update parameter $\theta, 0 < \theta < 1$;
begin
  $X := E; S := E; \mu := 1; V := E$
  while $n\mu \geq \varepsilon$ do
   begin (outer iteration)
    $\mu := (1 - \theta)\mu$;
    $V := \frac{V}{\sqrt{1-\theta}}$
    while $\Psi(V) > \tau$ do
    begin (inner iteration)
      Find search directions by solving system (14);
      determine a step size α;
      $X := X + \alpha \Delta X$;
      $S := S + \alpha \Delta S$;
      $y := y + \alpha \Delta y$;
      $V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} DSD$;
    end (inner iteration)
  end (outer iteration)
end

2.2 Generic Primal-dual IPMs for SDO

It is clear from this description that closeness of (X, y, S) to $(X(\mu), y(\mu), S(\mu))$ is mesured by the value of $\Psi(V)$, with τ as threshold value: if $\Psi(V) \leq \tau$, then start a new outer iteration by performing a μ-update, otherwise we enter an inner iteration relative to the current value of μ and apply (15) to get new iterates. The parametrs τ, θ and the step size α should be chosen in such a way that the number of iterations required by the algorithm is as small as possible.

2.3 Properties of the Kernel Function

We start by recalling some properties of $\psi(t)$,

$$\begin{aligned}
&\psi(1) = 0, \psi'(1) = 0, \qquad \psi_c''(t) > 0, \qquad \lim_{t\to 0^+} \psi(t) = \lim_{t\to+\infty} \psi(t) = +\infty,\\
&\psi'(t) = t - \frac{1}{2t} - \frac{e^{\frac{1}{t^q}-1}}{2t^{q+1}},\\
&\psi''(t) = 1 + \frac{1}{2t^2} + \frac{1}{2}\Big(\frac{(q+1)t^q + q}{t^{2q+2}}\Big)e^{\frac{1}{t^q}-1} > 1, \qquad (16)\\
&\psi'''(t) = \frac{-1}{t^3} - \frac{1}{2}(q^2 t^{-(3q+3)} + 3q(q+1)t^{-(2q+3)} + (q+1)(q+2)t^{-(q+3)})e^{\frac{1}{t^q}-1} < 0.
\end{aligned}$$

Moreover, from (16) $\psi(t)$ is strictly convex and $\psi''(t)$ is monotonically decreasing in $t \in (0, \infty)$.

Lemma 2.7 ([9]). *For $\psi(t)$, we have the following*
$\psi(t)$ is exponentially convex for all $t > 0$; that is

$$\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)), \qquad \text{(a)}$$

$$\psi''(t) \text{ is monotonically decreasing for all } t > 0, \qquad \text{(b)}$$

$$t\psi''(t) - \psi'(t) > 0 \text{ for all } t > 0, \qquad \text{(c)}$$

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

Lemma 2.8 ([9]). *For $\psi(t)$, we have*

$$\frac{1}{2}(t-1)^2 \leq \psi(t) \leq \frac{1}{2}[\psi'(t)]^2, \qquad t > 0 \qquad (17)$$

$$\psi(t) \leq \frac{2+q}{2}(t-1)^2, \qquad t > 1. \qquad (18)$$

Let $\psi(t)$ be as defined in (1), one has

$$t\psi'(t) \geq \psi(t), \qquad t > 1.$$

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 all $t \in [0, 1]$. In the next lemma we use the so-called barrier term $\psi_b(t)$ of $\psi(t)$, which is defined by $\psi(t) = \frac{t^2-1}{2} + \psi_b(t), \quad t > 0$. $\underline{\rho} : [0, \infty) \to (0, 1]$ be the inverse function of the restriction of $-\psi_b'(t)$ in the interval $(0, 1]$ and $s_b = -\psi_b'(t)$. Then one has

Lemma 2.9. *for $\psi(t)$, we have*

$$1+\sqrt{\frac{2s}{q+2}} \leq \varrho(s) \leq 1+\sqrt{2s}, \tag{19}$$

$$\rho(s) \geq \underline{\rho}(1+2s), \tag{20}$$

$$\underline{\rho}(s_b) \geq \frac{1}{(\log(2s_b)+1)^{\frac{1}{q}}}.s_b > \frac{1}{2}.$$

At some place below we apply the function Ψ to a positive vector v. The interpretation of $\Psi(v)$ is compatible with definition 2.3, when identifying the vector v with its diagonal matrix $V = \operatorname{diag}(v)$. when applying Ψ to this matrix we obtain

$$\Psi(V) = \sum_{i=1}^{n} \psi(v_i), \qquad v \in \mathbb{R}_+^n.$$

2.4 Properties of $\Psi(V)$ and $\delta(V)$

In the analysis of the algorithm, we also use the norm- based proximity measure $\delta(V)$ defined by

$$\delta(V) := \frac{1}{2}\left\|\Psi'(V)\right\| = \frac{1}{2}\sqrt{\sum_{i=1}^{n}\psi'(\lambda_i(V))^2} = \frac{1}{2}\left\|D_X + D_S\right\|. \tag{21}$$

Now we will derive a bound of $\delta(V)$ in term of $\Psi(V)$. Since $\Psi(V)$ is strictly convex and attains its minimal value zero at $V = E$, we have $\Psi'(V) = 0 \Leftrightarrow \delta(V) = 0 \Leftrightarrow V = E$.

Theorem 2.10 (Theorem 5 in [10])**.** *Let σ be the inverse function of $\psi(t)$ for $t \geq 1$. Then*

$$\delta(V) \geq \frac{1}{2}\psi'(\sigma(\Psi(V))).$$

Lemma 2.11. *Let $\delta(V)$ be as defined in* (21)*. Then we have*

$$\delta(V) \geq \sqrt{\frac{1}{2}\Psi(V)}. \tag{22}$$

Proof. Using (17), we have

$$\Psi(V) = \sum_{i=1}^{n}\psi(\lambda_i(V)) \leq \sum_{i=1}^{n}\frac{1}{2}\psi'(\lambda_i(V))^2 = \frac{1}{2}\left\|\nabla\Psi(V)\right\|^2 = 2\delta(V)^2,$$

so $\delta(V) \geq \sqrt{\frac{1}{2}\Psi(V)}$.

This proves the lemma. □

Remark 2.12. Throughout the paper, we assume that $\tau \geq 1$. Using lemma 2.11 and the assumption that $\Psi(V) \geq \tau$, we have $\delta(V) \geq \sqrt{\frac{1}{2}}$.

3 Complexity Analysis and the Iteration Complexity Bounds

We derive an estimate for effect of μ-update on the value of $\Psi(V)$.

3.1 Three Technical Lemmas

The next lemma is cited from [12, Lemma 3.3.14 (c)].

Lemma 3.1. *Let $A,\ B \in S^n$ be two nonsingular matrices and $f(t)$ be a given real-valued function such that $f(e^t)$ is a convex function. One has*

$$\sum_{i=1}^{n} f(\eta_i(AB)) \leq \sum_{i=1}^{n} f(\eta_i(A)\eta_i(B)),$$

where $\eta_i(A)$, and $\eta_i(B)$, $i = 1, 2, \cdots, n$ denote the singular values of A and B respectively.

Lemma 3.2. *Let $A,\ A + B \in S^n_+$, then one has*

$$\lambda_i(A + B) \geq \lambda_n(A) - |\lambda_1(B)|, \qquad i = 1, 2, \cdots, n.$$

A consequence of lemma 2.7, is that any eligible kernel function is exponentially convex. This implies the following lemma, which is crucial for our purpose.

Lemma 3.3 (Lemma 9 in [10])**.** *Let V_1 and V_2 be two symmetric positive definite matrices, then*

$$\Psi((V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}})^{\frac{1}{2}}) \leq \frac{1}{2}(\Psi(V_1) + \Psi(V_2)), \qquad \forall\ V_1 > 0,\ V_2 > 0.$$

Proof. For any nonsingular matrix $U \in S^n$, we have

$$\eta_i(U) = (\lambda_i(U^T U))^{\frac{1}{2}} = (\lambda_i(U U^{\frac{1}{2}}))^{\frac{1}{2}}, \qquad i = 1, 2, \cdots, n.$$

Taking U=$V_1^{\frac{1}{2}} V_2^{\frac{1}{2}}$, we may write

$$\eta_i(V_1^{\frac{1}{2}} V_2^{\frac{1}{2}}) = (\lambda_i(V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}}))^{\frac{1}{2}} = (\lambda_i(V_2^{\frac{1}{2}} V_1 V_2^{\frac{1}{2}}))^{\frac{1}{2}}, \qquad i = 1, 2, \cdots, n.$$

Since V_1 and V_2 are symmetric positive definite, using lemma 3.1 one has

$$\Psi((V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}})^{\frac{1}{2}}) = \sum_{i=1}^{n} \psi(\eta_i(V_1^{\frac{1}{2}} V_2^{\frac{1}{2}})) \leq \sum_{i=1}^{n} \psi(\eta_i(V_1^{\frac{1}{2}})\eta_i(V_2^{\frac{1}{2}})).$$

Since $\eta_i(V_1^{\frac{1}{2}}), \eta_i(V_2^{\frac{1}{2}}) > 0$, using lemma 2.7(a), we obtain

$$\begin{aligned}\Psi((V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}})^{\frac{1}{2}}) \leq & \frac{1}{2} \sum_{i=1}^{n} (\psi(\eta_i^2(V_1^{\frac{1}{2}})) + \psi(\eta_i(V_2^{\frac{1}{2}}))) \\ = & \frac{1}{2} \sum_{i=1}^{n} (\psi(\lambda_i(V_1)) + \psi(\lambda_i(V_2))) = \frac{1}{2}(\Psi(V_1) + \Psi(V_2)).\end{aligned}$$

This completes the proof of lemma 3.3. □

3.2 Decrease the Value of $\Psi(V)$ in the Inner Iteration

In each inner iteration the search directions $\Delta X, \Delta y$ and ΔS are obtained by solving the system (14), and using (6). After a step with size α, the new iterate is given by

$$X_+ = X + \alpha\Delta X, \qquad y_+ = y + \alpha\Delta y, \qquad S_+ = S + \alpha\Delta S.$$

Due to (6), we may write

$$X_+ = X + \alpha\Delta X = X + \alpha\sqrt{\mu}DD_X D = \sqrt{\mu}D(V + \alpha D_X)D$$

and

$$S_+ = S + \alpha\Delta S = S + \alpha\sqrt{\mu}D^{-1}D_S D^{-1} = \sqrt{\mu}D^{-1}(V + \alpha D_S)D^{-1}.$$

Denoting the matrix V after the step as V_+, we have

$$V_+ = \frac{1}{\sqrt{\mu}}(D^{-1}X_+S_+D)^{\frac{1}{2}}.$$

We can verify that V_+^2 is unitarily similar to the matrix $\frac{1}{\mu}X_+^{\frac{1}{2}}S_+X_+^{\frac{1}{2}}$ and thus to

$$(V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_S)(V + \alpha D_X)^{\frac{1}{2}}.$$

Consequently, the eigenvalues of the matrix V_+ are precisely the same as those of

$$\widetilde{V}_+ = ((V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_S)(V + \alpha D_X)^{\frac{1}{2}})^{\frac{1}{2}}.$$

By the definition of $\Psi(V)$, we obtain

$$\Psi(V_+) = \Psi(\widetilde{V}_+).$$

Let us denote the difference between the proximity before and after one step by a function of the step size, that is,

$$f(\alpha) = \Psi(V_+) - \Psi(V) = \Psi(\widetilde{V}_+) - \Psi(V), \tag{23}$$

we find α such that $f(\alpha)$ is as small as possible. Hence, by Lemma 3.3,

$$\begin{aligned}\Psi(\widetilde{V}_+) =& \Psi(((V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_S)(V + \alpha D_X)^{\frac{1}{2}})^{\frac{1}{2}})\\ \leq& \frac{1}{2}(\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)).\end{aligned}$$

We have $f(\alpha) \leq f_1(\alpha)$, where

$$f_1(\alpha) = \frac{1}{2}(\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)) - \Psi(V)$$

which means that $f_1(\alpha)$ gives an upper bound for the decrease of the barrier function $\Psi(V)$. Furthermore, we can easily verify that

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

It follows from (11), (12) and (13) that

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

and

$$f_1''(\alpha) = \frac{1}{2}\mathrm{tr}\,(\psi''(V + \alpha D_X)D_X^2 + \psi''(V + \alpha D_S)D_S^2), \tag{24}$$

hence, using (21) and the last equation of (14), we obtain

$$f_1'(0) = \frac{1}{2}tr(\psi'(V)(D_X + D_S)) = \frac{1}{2}tr(-\psi'(V)^2) = -2\delta(V)^2.$$

In what follows, we use the short notation: $\delta := \delta(V)$, and state some important results.

Lemma 3.4 (Lemma 5.19 in [26]). *One has*

$$f_1''(\alpha) \leq 2\delta^2\psi''(\lambda_n(V) - 2\alpha\delta)$$

Proof. Since $D_X + D_S = -\psi'(V)$ and (21) imply that $\|D_X + D_S\|^2 = \|D_X\|^2 + \|D_S\|^2 = 4\lambda^2$. Thus we have $|\lambda_1(D_X)| \leq 2\delta$ and $|\lambda_1(D_S)| \leq 2\delta$. Using lemma 3.2 and $V + \alpha D_X \geq 0$, we have , for each i,

$$\lambda_i(V + \alpha D_X) \geq \lambda_n(V) - \alpha\,|\lambda_1(D_X)| \geq \lambda_n(V) - 2\alpha\delta,$$
$$\lambda_i(V + \alpha D_S) \geq \lambda_n(V) - \alpha\,|\lambda_1(D_S)| \geq \lambda_n(V) - 2\alpha\delta.$$

Since ψ'' is monotonically decreasing in $t \in (0, +\infty)$, we obtain

$$\psi''(\lambda_i(V + \alpha D_X)) \leq \psi''(\lambda_n - 2\alpha\delta), \qquad \psi''(\lambda_i(V + \alpha D_S)) \leq \psi''(\lambda_n - 2\alpha\delta).$$

Substitution into (24) gives

$$f''(\alpha) \leq \frac{1}{2}\psi''(\lambda_n - 2\alpha\delta)\mathrm{tr}(D_X^2 + D_S^2) = \frac{1}{2}\psi''(\lambda_n - 2\alpha\delta)(\|D_X\|^2 + \|D_S\|^2).$$

Now, using that $D_X \perp D_S$ and $\|D_X + D_S\| = 2\delta$, we obtain

$$f''(\alpha) \leq 2\psi''(\lambda_n(V) - 2\alpha\delta).$$

This proves the lemma. □

Next we will choose a suitable step size for the algorithm. This should be chosen such that X_+ and S_+ are feasible and another one is to make $\Psi(V_+) - \Psi(V)$ decreases sufficiently. Putting $v_i = \lambda_i(V), \quad 0 \leq i \leq n$, and $v_1 := \min(v)$, we have:

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

which is the same inequality as in Lemma 4.1 in [4]. From this stage on we can apply similar arguments as in [4] for the LO case to obtain the following results which require no further proof.

Lemma 3.5 (Lemma 4.2 in [4]). *One has $f_1'(\alpha) \leq 0$, certainly holds if α satisfies the inequality*

$$-\psi'(v_1 - 2\alpha\delta) + \psi'(v_1) \leq 2\delta. \tag{25}$$

Lemma 3.6 (Lemma 4.3 in [4]). *The largest step size α that satisfies* (25) *is given by*

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

Lemma 3.7 (Lemma 4.4 in [4])**.** *Let $\bar{\alpha}$ is defined in Lemma 3.6. Then*

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

As in the LO case, we use a default step size $\widetilde{\alpha}$ that is the lower bound of the $\bar{\alpha}$ and consists of δ.

$$\widetilde{\alpha} := \frac{1}{\psi''(\rho(2\delta))}. \tag{26}$$

Lemma 3.8. *Let ρ and $\bar{\alpha}$ be the same as defined in Lemma 3.7. Then if $\Psi(V) \geq \tau \geq 1$, then we have*

$$\bar{\alpha} \geq \frac{1}{1 + (2q+1)(1+4\delta)[\log(2+8\delta)+1]^{\frac{q+1}{q}}}.$$

Proof. Using (16), (20) and lemma 3.7, we have

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))} \geq \frac{1}{\psi''(\underline{\rho}(1+4\delta))}$$

by setting $t = \underline{\rho}(1+4\delta)$, $(0 < t \leq 1)$, it follows that

$$\begin{aligned}
\bar{\alpha} \geq & \frac{1}{\psi''(t)} = \frac{1}{1 + \frac{1}{2t^2} + [\frac{1}{2}(q+1)t^{-(q+2)} + \frac{1}{2}qt^{-(2q+2)}]e^{t^{-q}-1}} \\
> & \frac{1}{1 + (2q+1)t^{-(q+1)}(-\psi_b'(t))} \\
> & \frac{1}{1 + (2q+1)(1+4\delta)[\log(2+8\delta)+1]^{\frac{q+1}{q}}}, \qquad (\text{put } t = \underline{\rho}(1+\sqrt{\Psi(v)})).
\end{aligned}$$

This completes the proof. □

By lemma 3.8, we have

$$\widetilde{\alpha} := \frac{1}{1 + (2q+1)(1+4\delta)[\log(2+8\delta)+1]^{\frac{q+1}{q}}} \tag{27}$$

Lemma 3.9 (Lemma. 4.5 in [4])**.** *If the step size α satisfies $\alpha \leq \bar{\alpha}$, then*

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

Lemma 3.10. *Let $\widetilde{\alpha}$ be the default step size as defined in* (27) *and $\Psi(V) \geq 1$, then*

$$f(\widetilde{\alpha}) \leq -\frac{\sqrt{\Psi(v)}}{2 + (2q+1)(1+4\sqrt{2})[\log(2+4\sqrt{2\Psi_0})+1]^{\frac{q+1}{q}}}.$$

Proof. Since $\Psi(V) \geq 1$, then from (22), we have

$$\delta \geq \sqrt{\frac{1}{2}\Psi(V)} \geq \sqrt{\frac{1}{2}}.$$

Using lemma 3.9 with $\alpha = \widetilde{\alpha}$, we have

$$f(\widetilde{\alpha}) \leq -\widetilde{\alpha}\delta^2,$$
$$f(\widetilde{\alpha}) \leq -\frac{\delta}{\sqrt{2} + (2q+1)(\sqrt{2}+4)[\log(2+8\delta)+1]^{\frac{q+1}{q}}}.$$

Since the decrease depends monotonically on δ, subtitution yields

$$f(\widetilde{\alpha}) \leq -\frac{\sqrt{\Psi(v)}}{2 + (2q+1)(2+4\sqrt{2})[\log(2+4\sqrt{2\Psi_0})) + 1]^{\frac{q+1}{q}}}.$$

Where the last inequality follows from $\Psi_0 \geq \Psi \geq \tau \geq 1$.
This completes the proof. □

3.2.1 A Lower Bound of Ψ

The next theorem is an extension of Theorem 3.2 in [4] to positive definite matrices.

Theorem 3.11 (Theorem 3 in [24]). *Let ϱ be as defined in lemma 2.9. Then for any positive definite matrix V, and any $\beta > 1$ we have: $\Psi(\beta V) \leq n\psi(\beta\varrho(\frac{\Psi(V)}{n}))$.*

Corollary 3.12. *Let $0 \leq \theta < 1$ and $V_+ = \frac{V}{\sqrt{1-\theta}}$. If $\Psi(V) \leq \tau$, then*

$$\Psi(V_+) \leq n\psi\Big(\frac{\varrho(\dfrac{\tau}{n})}{\sqrt{1-\theta}}\Big) \leq \frac{n\theta + 2\tau + 2\sqrt{2\tau n}}{2(1-\theta)}.$$

Proof. Since $\frac{1}{\sqrt{1-\theta}} \geq 1$ and $\varrho(\frac{\Psi(V)}{n}) \geq 1$, we have $\frac{1}{\sqrt{1-\theta}}\varrho(\frac{\Psi(V)}{n}) \geq 1$. Using Theorem 3.11 with $\beta = \frac{1}{\sqrt{1-\theta}}$ and the function ϱ is monotonically increasing since $\psi(t)$ is that for $t \geq 1$, we have

$$\Psi(V_+) \leq n\psi\Big(\frac{\varrho(\dfrac{\Psi(V)}{n})}{\sqrt{1-\theta}}\Big) \leq n\psi\Big(\frac{\varrho(\dfrac{\tau}{n})}{\sqrt{1-\theta}}\Big).$$

This prove the first inequality. The second inequality follows from: for $t \geq 1$, we have $\psi(t) \leq \frac{t^2-1}{2}$.
Then

$$\begin{aligned}\Psi(V_+) \leq & n\psi\Big(\frac{\sigma(\frac{\tau}{n})}{\sqrt{1-\theta}}\Big) \leq \frac{n}{2}\Big(\frac{\sigma^2(\dfrac{\tau}{n})}{1-\theta} - 1\Big)\\ =&\frac{n}{2(1-\theta)}\Big(\sigma^2\Big(\frac{\tau}{n}\Big) - (1-\theta)\Big)\\ \leq&\frac{n}{2(1-\theta)}\Big(\Big(1+\sqrt{\frac{2\tau}{n}}\Big)^2 - (1-\theta)\Big)\\ =&\frac{(n\theta + 2\tau + 2\sqrt{2\tau n})}{2(1-\theta)}.\end{aligned}$$

□

3.3 Complexity of the Algorithm

We denote the value of $\Psi(V)$ after μ-update by Ψ_0

$$\Psi(V_+) \leq \Psi_0 = \frac{n\theta + 2\tau + 2\sqrt{2\tau n}}{2(1-\theta)} = L(n,\theta,\tau).$$

We need to count how many inner iterations are required to return to the situation where $\Psi(V) \leq \tau$. the subsequent values in the same outer iteration are denoted as Ψ_k , $k = 1, 2, \cdots, K$, where K denotes the total number of inner iterations in the outer iteration. According to decrease of $f(\widetilde{\alpha})$, for $k = 1, 2, \cdots, K-1$, we obtain $K \leq (4+(2q+1)(4+8\sqrt{2})[\log(2+4\sqrt{2\Psi_0})+1]^{\frac{q+1}{q}})\Psi_0^{\frac{1}{2}}$.

Lemma 3.13 (Lemma 14 in [24])**.** *Suppose $t_0, t_1, \cdots, t_k$ be a sequence of positive numbers such that $t_{k+1} \leq t_k - \beta t_k^{1-\gamma}$, $k = 0, 1, \cdots, K-1$, where $\beta > 0$ and $0 < \gamma \leq 1$. Then $K \leq [\frac{t_0^{\gamma}}{\beta\gamma}]$, letting $t_k = \Psi_k$, $\beta = \frac{1}{2+(2q+1)(2+4\sqrt{2})[\log(2+4\sqrt{2\Psi_0})+1]^{\frac{q+1}{q}}}$ and $\gamma = \frac{1}{2}$.*

Theorem 3.14. *Let K be the total number of inner iterations in the outer iterations. Then we have*

$$K \leq (4 + (2q+1)(4+8\sqrt{2})[\log(2+4\sqrt{2\Psi_0})+1]^{\frac{q+1}{q}})\Psi_0^{\frac{1}{2}}.$$

Proof. By lemma 3.13, we have

$K \leq [\frac{t_0^{\gamma}}{\beta\gamma}] = (4 + (2q+1)(4+8\sqrt{2})[\log(2+4\sqrt{2\Psi_0})+1]^{\frac{q+1}{q}})\Psi_0^{\frac{1}{2}}$.

This completes the proof. □

The number of outer iterations is bounded above by $\frac{1}{\theta}\log\frac{n}{\varepsilon}$ (cf. [20], II. 17, page 116). By multiplying the number of outer iterations and the number of inner iterations we get an upper bound for the total number of iterations, namely,

$$(4 + (2q+1)(4+8\sqrt{2})[\log(2+4\sqrt{2\Psi_0})+1]^{\frac{q+1}{q}})\Psi_0^{\frac{1}{2}}\frac{\log\frac{n}{\varepsilon}}{\theta}.$$

For large-update methods with $\tau = O(\sqrt{n})$ and $\theta = \Theta(1)$, we have $\Psi_0 = O(n)$ and $O(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon})$ iterations complexity.

In the case of small-update methods, $\tau = O(1)$, $\theta = \Theta(\frac{1}{\sqrt{n}})$ and $\psi(t) \leq \frac{2+q}{2}(t-1)^2$, $t > 1$. We then obtain

$$\begin{aligned}\Psi(V_+) \leq & n\psi\Big(\frac{1}{\sqrt{1-\theta}}\varrho(\frac{\Psi(V)}{n})\Big) \leq n\frac{q+2}{2}\Big(\frac{1}{\sqrt{1-\theta}}\varrho\Big(\frac{\Psi(V)}{n}\Big)-1\Big)^2 \\ = & \frac{n(q+2)}{2(1-\theta)}\Big(\varrho\Big(\frac{\Psi(V)}{n}\Big)-\sqrt{1-\theta}\Big)^2\end{aligned}$$

Using (19), we have

$$\begin{aligned}&\frac{n(q+2)}{2(1-\theta)}\Big(\varrho\Big(\frac{\Psi(V)}{n}\Big)-\sqrt{1-\theta}\Big)^2 \\ \leq & \frac{n(q+2)}{2(1-\theta)}\Big(1+\sqrt{2\frac{\Psi(V)}{n}}-\sqrt{1-\theta}\Big)^2 = \frac{n(q+2)}{2(1-\theta)}\Big((1-\sqrt{1-\theta})+\sqrt{2\frac{\Psi(V)}{n}}\Big)^2 \\ \leq & \frac{n(q+2)}{2(1-\theta)}\Big(\theta+\sqrt{2\frac{\tau}{n}}\Big)^2 = \frac{(q+2)}{2(1-\theta)}(\theta\sqrt{n}+\sqrt{2\tau})^2 = \Psi_0,\end{aligned}$$

where we also used that $1-\sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \leq \theta$ and $\Psi(V) \leq \tau$, using this upper for Ψ_0, we get in this case $\Psi_0 = O(q)$ and the iteration bound becomes $O(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon})$ iteration complexity.

4 Numerical Results

In this section, we present some numerical results, where the Algorithm is coded in MATLAB (R2014a) and our experiments were performed on PC with Processeur Genuine Intel (R) CPR T2080 @ 1,73GHZ installed memory (RAM) 2,00GO.

We consider the SDO problems in [26], whose primal-dual pair of (SDO) and (SDD) have the following data:

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 & -1 \\ 0 & -1 & 1 & -1 & -2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & -2 & 2 & 0 \\ 0 & 2 & 1 & 0 & 2 \\ -2 & 1 & -2 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 2 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 2 & 2 & -1 & -1 & 1 \\ 2 & 0 & 2 & 1 & 1 \\ -1 & 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & 0 & -2 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 3 & -3 & 1 & 1 \\ 3 & 5 & 3 & 1 & 2 \\ -3 & 3 & -1 & 1 & 2 \\ 1 & 1 & 1 & -3 & -1 \\ 1 & 2 & 2 & -1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}$$

One may easily verify that $X = E$ is feasible for the primal problem, and that $y = (1; 1; 1)$ and $S = E$ is feasible for the dual problem. An optimal solution of the primal problem is given by

$$X^* = \begin{pmatrix} 0.0714 & -0.0718 & 0.0169 & 0.0649 & -0.1583 \\ -0.0718 & 0.0724 & -0.0183 & -0.0602 & 0.1676 \\ 0.0169 & -0.0183 & 0.0103 & -0.0084 & -0.0772 \\ 0.0649 & -0.0602 & -0.0084 & 0.1481 & 0.0056 \\ -0.1583 & 0.1676 & -0.0772 & 0.0056 & 0.6022 \end{pmatrix}$$

and for the dual problem an optimal solution is given by

$$Z^* = \begin{pmatrix} 1.4338 & 0.5754 & -0.0295 & -0.4043 & 0.2169 \\ 0.5754 & 1.0956 & 0.3401 & 0.2169 & -0.1120 \\ -0.0295 & 0.3401 & 1.1874 & 0.2169 & 0.0478 \\ -0.4043 & 0.2169 & 0.2169 & 0.2831 & -0.1415 \\ 0.2169 & -0.1120 & 0.0478 & -0.1415 & 0.0957 \end{pmatrix}$$

$$y^* = \begin{pmatrix} 0.8585 \\ 1.0937 \\ 0.7831 \end{pmatrix}.$$

The optimal value of both problem is equal to -1.0957.

The main goal of this section is to compare iteration numbers and calculation time of the algorithm for the following four kernel functions:

$\psi(\mathbf{t}) = \frac{\mathbf{t^2-1-\log(t)}}{\mathbf{2}} + \frac{\mathbf{e^{\frac{1}{t^q}-1}-1}}{\mathbf{2q}}$ **for** $\mathbf{t > 0, q \geq 1}$**;** $\psi_1(\mathbf{t}) = \frac{t^2-1}{2} - \log(t)$ **;** $\psi_2(\mathbf{t}) = \frac{t^2-1}{2} - (t-1)e^{\frac{1}{t}-1}$; $\psi_3(t) = (m+1)t^2 - (m+2)t + \frac{1}{t^m}, t > 0$, where $m > 4$, where $\psi_1(\mathbf{t})$ is the classical logarithmic kernel function (see [20]), $\psi_2(\mathbf{t})$ is the non-self-regular kernel function (see

[28]), and $\psi_3(t)$ is the non-self-regular kernel function (see [14]). We take the parameter $\theta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, the step size $\alpha \in \{0.5, 0.6, 0.7, \alpha_{\max}\}$. A practical step size $\alpha_{\max}$ such that $\alpha_{\max} = \rho \min(\alpha_X, \alpha_Z)$ and $\rho \in (0, 1)$, where

$$\alpha_X = \begin{cases} -\dfrac{1}{\lambda_{\min}(X^{-1}\Delta X)}, & \text{if } \lambda_{\min}(X^{-1}\Delta X) < 0, \\ 1, & \text{if } \lambda_{\min}(X^{-1}\Delta X) \geq 0, \end{cases}$$

$$\alpha_Z = \begin{cases} -\dfrac{1}{\lambda_{\min}(Z^{-1}\Delta Z)}, & \text{if } \lambda_{\min}(Z^{-1}\Delta Z) < 0, \\ 1, & \text{if } \lambda_{\min}(Z^{-1}\Delta Z) \geq 0. \end{cases}$$

The threshold parameter $\tau = 3$, and the accuracy parameter $\varepsilon = 10^{-8}$ in all experiments. The iteration numbers and the calculation time of the algorithm based on the above kernel functions are given in the following tables.

Table 1. Numerical Results for $\psi(t)$ $(q = 1, q = 2$ Alternately)

	0.5		0.6		0.7		$\alpha_{\max}$	
θ/α	Int	Cpu	Int	Cpu	Int	Cpu	Int	Cpu
0.1	42	0.4532	33	0.3095	26	0.4330	15	0.3103
0.1	42	0.5413	33	0.3734	28	0.3685	15	0.2925
0.3	39	0.2695	31	0.2231	25	0.2083	15	0.1737
0.3	40	0.2651	32	0.2460	25	0.2147	15	0.2178
0.5	37	0.2804	26	0.1961	26	0.1286	15	0.1381
0.5	39	0.2509	27	0.2002	26	0.1441	15	0.1427
0.7	35	0.1456	28	0.2336	22	0.2672	16	0.1373
0.7	36	0.3283	31	0.2032	23	0.1719	16	0.1336
0.9	31	0.1133	23	0.1592	19	0.1446	15	0.1796
0.9	31	0.1835	23	0.1616	19	0.1577	15	0.1185

Table 2. Numerical Results for $\psi_1(t)$

	0.5		0.6		0.7		$\alpha_{\max}$	
θ/α	Int	Cpu	Int	Cpu	Int	Cpu	Int	Cpu
0.1	39	0.4480	31	0.3727	25	0.3483	15	0.2871
0.3	37	0.3298	30	0.2187	25	0.2083	15	0.2419
0.5	35	0.2066	26	0.2687	26	0.2462	15	0.1313
0.7	34	0.3439	27	0.0948	21	0.1918	16	0.1354
0.9	30	0.2553	23	0.2868	18	0.2298	15	0.1132

Table 3. Numerical Results for $\psi_2(t)$

	0.5		0.6		0.7		$\alpha_{\max}$	
θ/α	Int	Cpu	Int	Cpu	Int	Cpu	Int	Cpu
0.1	41	0.3990	32	0.3406	26	0.4055	15	0.2871
0.3	38	0.3449	30	0.3135	25	0.2774	15	0.1525
0.5	36	0.3012	26	0.9552	26	0.2743	15	0.2216
0.7	34	0.1812	28	0.2142	22	0.2342	16	0.1917
0.9	31	0.1718	23	0.2442	18	0.1382	15	0.1031

Table 4. Numerical results for $\psi_3(t)$

	0.5		0.6		0.7		$\alpha_{\max}$	
θ/α	Int	Cpu	Int	Cpu	Int	Cpu	Int	Cpu
0.1	98	0.5745	86	0.5279	69	0.5514	69	0.5514
0.3	77	0.3306	52	0.2903	52	0.5573	48	0.4156
0.5	67	0.4727	53	0.3989	40	0.3151	29	0.3111
0.7	55	0.2651	47	0.2471	32	0.3329	52	0.2719
0.9	40	0.2812	32	0.3669	24	0.1611	52	0.3636

The results in these four tables show that the algorithm based on our new kernel function $\psi(x)$ is efficient. The iteration numbers of the algorithm depend on the values of the parameter θ and step size α. In fact, for each θ that considered, larger values of α give better iteration numbers. However the step size α should have an upper bound in practical computation. For each α, larger θ gives better iteration numbers for $\psi(t)$, $\psi_1(t)$ and $\psi_2(t)$, while for $\psi_3(t)$, $\theta = 0.5$ gives better results. But not better than $\alpha = \alpha_{\max}$ which gives better iteration numbers in all cases.

5 Concluding Remarks

The method presented in this article shows an extension of a large-update primal–dual IPMs for LO to SDO. This method is based on a new class of parametric kernel functions. It is shown that the best result of iteration bounds of our algorithm for large-update and small-update methods based on this new kernel function can be achieved, namely $O(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon})$ for large-update methods and $O(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon})$ for small-update methods. Moreover, we have reported some numerical results to show the validity of our approach by made a comparaison with methods based on different kernel functions. It is shown that our algorithm is efficient for these preliminary.

Some interesting topics remain for further research. Firstly, the search directions used in this paper are all based on the NT-symmetrization scheme. It may be possible to design similar algorithms using other symmetrization schemes and to obtain polynomial-time complexity bound. Secondly, further research my extend this result to linear complementarity problems (LCP) over symetric cones (SCLCP) and the Cartesian $P_*(\kappa)$ SCLCP.

References

[1] Alizadeh, F. Interior-point methods in semi-definite programming with applications to combinatorial optimization. *SIAM J. Optim.*, 5: 13–51 (1995)

[2] Amini, K., Haseli, A. A new proximity function generating the best known iteration bounds for both large-update and small-update interior-point methods. *ANZIAM Journal*, 49(2): 259–270 (2007)

[3] Bai, Y.Q., El Ghami, M., Roos, C. A new efficient large update interior-point method based a finite barrier. *SIAM Journal on Optimization*, 13(3): 766–782 (2003)

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

[5] Bai, Y.Q., Lesaja, G., Roos, C., Wang, G. Q., El Ghami, M. 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 (2008)

[6] Cai, X.Z., Li, L., El Ghami, M., Steihaug, T., Wang, G.Q. A new parametric kernel function yielding the best known iteration bounds of interior-point methods for the Cartesian $P_*(\kappa)$-LCP over symmetric cones. *Pacific Journal of Optimization*, 13(4): 547–570 (2017)

[7] Choi, B.K., Kee, G.M. On complexity analysis of the primal-dual interior-point method for semidefinite optimization problem based on a new proximity function. *Nonlinear Anal.*, 71(12): 2540–2550 (2009)

[8] DeKlerk, E. Aspects of Semidefinite Programming. Applied Optimization, Vol. 65. Kluwer Academic, Dordrecht, 2002

[9] Derbal, L., Kebbiche, Z. Theoretical and Numerical Result for Linear Optimization. Journal of Siberian Federal University. *Mathematics & Physics*, 12(2): 160–172 (2019)

[10] EI Ghami, M. Primal-dual algorithms for semidefinite optimization problems based on generalized trigonometric barrier function. *International Journal of Pure and Applied Mathematics*, 114(4): 797-818 (2017)

[11] El Ghami, M., Guennoun, Z. A., Bouali, S., Steihaug, T. Interior-point methods for linear optimization based on a kernel function with a trigonometric barrier term. *Journal of Computational and Applied Mathematics*, 236(15): 3613–3623 (2012)

[12] Horn, R.A., Charles, R.J. Matrix Analysis. Cambridge University Press, UK, 1986

[13] Kojima, M., Shida, M., Hara, S. Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM J. Optim.*, 7: 86–125 (1997)

[14] Liu, L., Li, S. A new kind of simple kennel function yielding good iteration bounds for primal-dual interior-point methods. *Journal of Computational and Applied Mathematics*, ??, 2011

[15] Nesterov, Y.E., Nemirovskii, A.S. Interior point polynomial algorithms in convex programming. In: SIAM Stud. Appl. Math., Vol. 13. SIAM, Philadelphia, 1994

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

[17] Peng, J., Roos, C., Terlaky, T. Self-regularity: A New Paradigm for Primal-dual Interior-point Algorithms. Princeton University Press, New Jersey, 2002

[18] Peyghami, M.R., Hafshejani, S.F. Complexity analysis of an interior-point algorithm for linear optimization based on a new proximity function. Numerical Algorithms, 2014

[19] Peyghami, M.R., Hafshejani, S.F., Shirvani, L. Complexity of interior-point methods for linear optimization based on a new trigonometric kernel function. *Journal of Computational and Applied Mathematics*, 255(1): 74–85 (2014)

[20] Roos, C., Terlaky, T., Vial, J. Ph. Theory and Algorithms for Linear Optimization. An Interior-Point Approach. John Wiley & Sons, Chichester, UK, 1997

[21] Sturm, J.F., Zhang, S. Symmetric primal-dual path following algorithms for semidefinite programming. *Appl. Num. Math.*, 29: 301–315 (1999)

[22] Todd, M.J., Toh, K.C., T¨ut¨ unc¨u, R. H. The Nesterov-Todd directions in semidefinite programming. *SIAM Journal on Optimization*, 8(3): 769–796 (1998)

[23] Todd, M.J. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Optim. Methods Softw.*, 11, 1–46 (1999)

[24] Wang, G.Q., Bai, Y.Q. A class of polynomial primal-dual interior-point algorithms for semidefinite optimization. *Journal of Shanghai University (English Edition)*, 10(3): 198–207 (2006)

[25] Wang, G.Q., Lesaja, G. Full Nesterov-Todd step feasible interior-point method for the Cartesian $P_*(\kappa)$-SCLCP. *Optim. Methods Softw.*, 28(3): 600–618 (2013)

[26] Wang, G.Q., Bai, Y.Q., Roos, C. Primal-dual interior point algorithms for semidefinite optimization based on a simple kernel function. *Journal of Mathematical Modelling and Algorithms*, 4(4): 409–433 (2005)

[27] Wolkowicz, H., Saigal, R., Vandenberghe, L. Handbook of Semidefinite Programming: Theory, Algorithms, and Applications. Kluwer Academic Publishers, Dordrecht, 2000

[28] Zhang, M.W. A Large-update Interior-point Algorithm for Convex Quadratic Semi-definite Optimization Based on a New Kernel Function. *Acta Mathematica Sinica, English Series*, 28(11): 2313–2328 (2012)

[29] Zhang, Y. On extending some primal dual algorithms from linear programming to semidefinite programming. *SIAM J. Optim.*, 8: 365–386 (1998)