COMPLEXITY ANALYSIS OF PRIMAL-DUAL INTERIOR-POINT METHODS FOR SEMIDEFINITE OPTIMIZATION BASED ON A PARAMETRIC KERNEL FUNCTION WITH A TRIGONOMETRIC BARRIER TERM

GUOQIANG WANG, ZHONGCHEN WU, ZHONGTUAN ZHENG AND XINZHONG CAI

College of Fundamental Studies
Shanghai University of Engineering Science
Shanghai 201620, China

ABSTRACT. In this paper, we present a class of large- and small-update primal-dual interior-point methods for semidefinite optimization based on a parametric kernel function with a trigonometric barrier term. For both versions of the kernel-based interior-point methods, the worst case iteration bounds are established, namely, $O(n^{\frac{3}{2}} \log n)$ and $O(\sqrt{n} \log n)$, respectively. These results match the ones obtained in the linear optimization case.

1. Introduction. In this paper, we consider the primal problem of semidefinite optimization (SDO) in the standard form

$$\text{(SDOP)} \quad \min \{C \cdot X : A_i \cdot X = b_i, i = 1, \ldots, m, X \succeq 0\},$$

and its dual problem

$$\text{(SDOD)} \quad \max \left\{ b^T y : \sum_{i=1}^m y_i A_i + S = C, S \succeq 0 \right\}.$$  

Here, each $A_i \in \mathbb{S}^n$, $b \in \mathbb{R}^m$, and $C \in \mathbb{S}^n$. Throughout the paper, we assume that the matrices $A_i$ are linearly independent.

For years, SDO has been one of the most active research areas in mathematical programming. There are many solution approaches for SDO. Among them, the interior-point methods (IPMs) gain much more attention. Several IPMs designed for linear optimization (LO) (e.g., [2,3,5,7,14,16–18]) have been successfully extended to SDO, (e.g., [1,4,6,8,10–15,19–24]), due to their polynomial complexity and practical efficiency.

The IPM is based on the barrier functions that are defined by a large class of univariate functions called the eligible kernel functions [2] which have recently been successfully used to design new primal-dual IPMs for various optimization problems. It is well known that the use of certain eligible kernel functions lead to significant reduction of the complexity gap between large- and small-update methods comparing to the logarithmic kernel function. This was one of the main
motivations of considering eligible kernel functions as an alternative to classical logarithmic kernel function. For some other related kernel-based IPMs we refer to the references [1, 3, 4, 7, 8, 10, 15–17, 20, 21, 23].

El Ghami et al. [7] first introduced a trigonometric kernel function for primal-dual IPMs in LO. They established the worst case iteration bounds for large- and small-update methods, namely, $O(n^{\frac{3}{2}} \log \frac{n}{\epsilon})$ and $O(\sqrt{n} \log \frac{n}{\epsilon})$, respectively. Recently, Peyghami et al. [16, 17] and Cai et al. [3] constructed some new trigonometric kernel functions and presented a class of large- and small-update IPMs for LO. Some trigonometric kernel functions and the corresponding iteration bounds for large- and small-update methods are collected in Table 1.

The purpose of this paper is to present a class of large- and small-update primal-dual IPMs for SDO based on the following parametric kernel function

$$
\psi(t) = \frac{t^2 - 1}{2} - \log t + \lambda \tan^2 \left( \frac{\pi(1-t)}{2 + 3t} \right), \quad 0 < \lambda \leq \frac{8}{25\pi},
$$

which was first studied by Cai et al. [3] for LO. It should be noted that if $\lambda = 0$, then

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

which is the kernel function of the classic barrier function. Comparing to the existing ones [7, 16, 17], the proposed kernel function includes a class of kernel functions. By utilizing the feature of the parametric kernel function, we obtained the iteration bound for large-update methods, namely, $O(n^{\frac{3}{2}} \log \frac{n}{\epsilon})$, which improves the classical iteration complexity with a factor $n^{\frac{1}{2}}$. For small-update methods, we derived the iteration bound, namely, $O(\sqrt{n} \log \frac{n}{\epsilon})$, which coincides the currently best known iteration bound for these type methods.

The paper is organized as follows. In Section 2, we introduce the framework of kernel-based IPMs for SDO. In Section 3, we recall some useful properties of the parametric kernel function, as well as the corresponding barrier function. The complexity and analysis of the algorithms are presented in Section 4. Finally, Section 5 contains some conclusions and remarks.

Some of the notations used throughout the paper are as follows. $\mathbf{S}^n$, $\mathbf{S}^n_+$ and $\mathbf{S}^{n+}_+$ denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n \times n$ matrices, respectively. $A \cdot B = \text{Tr}(A^T B)$ denotes the matrix inner product of two matrices $A$ and $B$. $\| \cdot \|$ denotes the Frobenius norm for matrices. The Löwner partial order $\succeq$ (or $\succ$) on positive semidefinite (or positive definite) matrices means $A \succeq B$ (or $A \succ B$) if $A - B$ is positive semidefinite (or positive definite). For any $Q \in \mathbf{S}^n_{++}$, the expression $Q^\frac{1}{2}$ (or $\sqrt{Q}$) denotes its symmetric

| Eligible kernel functions $\psi(t)$ | Large-update methods | Small-update methods | Ref. |
|-----------------------------------|---------------------|---------------------|-----|
| $\frac{t^2 - 1}{2} + \frac{8}{\pi} \tan \left( \frac{\pi(1-t)}{2 + 3t} \right)$ | $O(n^{\frac{3}{2}} \log \frac{n}{\epsilon})$ | $O(\sqrt{n} \log \frac{n}{\epsilon})$ | [7] |
| $\frac{t^2 - 1}{2} + \frac{4}{\pi} \cot \left( \frac{\pi t}{2 + 3t} \right)$ | $O(n^{\frac{3}{2}} \log \frac{n}{\epsilon})$ | $O(\sqrt{n} \log \frac{n}{\epsilon})$ | [10] |
| $\frac{t^2 - 1}{2} - \log t + \frac{1}{8} \tan^2 \left( \frac{\pi(1-t)}{2 + 3t} \right)$ | $O(n^{\frac{3}{2}} \log \frac{n}{\epsilon})$ | $O(\sqrt{n} \log \frac{n}{\epsilon})$ | [17] |
| $\frac{t^2 - 1}{2} - \int_1^{t} e^{3(\tan (\pi/(2 + 2x)) - 1)}dx$ | $O(\sqrt{n} (\log n)^2 \log \frac{n}{\epsilon})$ | $O(\sqrt{n} \log \frac{n}{\epsilon})$ | [16] |

Table 1. Complexity results for the eligible kernel functions
square root. For any \( V \in S^n \), we define \( \lambda_{\min}(V) \) (or \( \lambda_{\max}(V) \)) to be the minimal (or maximal) eigenvalue of \( V \). Finally, if \( g(x) \geq 0 \) is a real valued function of a real nonnegative variable, the notation \( g(x) = O(x) \) means that \( g(x) \leq \bar{c}x \) for some positive constant \( \bar{c} \) and \( g(x) = \Theta(x) \) that \( c_1 x \leq g(x) \leq c_2 x \) for two positive constants \( c_1 \) and \( c_2 \).

2. The framework of kernel-based IPMs for SDO. Throughout the paper, we assume that (SDOP) and (SDOD) satisfy the interior-point condition (IPC), i.e., there exists \((X^0, y^0, S^0)\) such that

\[
A_i \cdot X^0 = b_i, \; i = 1, \ldots, m, \; X^0 \succ 0, \; \sum_{i=1}^m y^0_i A_i + S^0 = C^0, \; S^0 \succ 0.
\]

The perturbed Karush-Kuhn-Tucker (KKT) conditions for (SDOP) and (SDOD) are equivalent to the following system

\[
A_i \cdot X = b_i, \; i = 1, \ldots, m, \; X \succeq 0,
\]

\[
\sum_{i=1}^m y_i A_i + S = C, \; S \succeq 0,
\]

\[
XS = \mu E,
\]

where \( \mu > 0 \) and \( E \) is the \( n \times n \) unit matrix. This system (5) has a unique solution, denoted by \((X(\mu), y(\mu), S(\mu))\). The set of \( \mu \)-centers (with \( \mu \) running through positive real numbers) gives a homotopy path, which is called the central path of (SDOP) and (SDOD). If \( \mu \to 0 \), then the limit of the central path exists, and since the limit points satisfy the complementarity condition, i.e., \( XS = 0 \), it naturally yields an optimal solution for (SDOP) and (SDOD) (see, e.g., [6]).

It is well known that the search directions for SDO require some symmetrization scheme. Zhang [24] suggested to replace the nonlinear equation \( XS = \mu E \) by

\[
H_P(M) = \mu E,
\]

where \( H_P \) is the linear transformation given by

\[
H_P(M) := \frac{1}{2}(PMP^{-1} + (PMP^{-1})^T),
\]

for any symmetric matrix \( M \), and where the scaling matrix \( P \) determines the symmetrization strategy. For any given nonsingular matrix \( P \), the system (5) is equivalent to

\[
A_i \cdot X = b_i, \; i = 1, \ldots, m, \; X \succeq 0,
\]

\[
\sum_{i=1}^m y_i A_i + S = C, \; S \succeq 0,
\]

\[
H_P(XS) = \mu E.
\]

Applying Newton’s method to the system (7), we have

\[
A_i \cdot \Delta X = 0, \; i = 1, \ldots, m,
\]

\[
\sum_{i=1}^m \Delta y_i A_i + \Delta S = 0,
\]

\[
H_P(X \Delta S + \Delta XS) = \mu E - H_P(XS).
\]
The search direction obtained through the system (8) is called the Monteiro-Zhang (MZ) unified direction. Different choices of the matrix $P$ result in different search directions (see, e.g., [6,24]). In this paper, we consider the so-called NT-symmetrization scheme, from which the NT search direction is derived. Let 

$$P := X^\frac{1}{2} (X^\frac{1}{2} S X^\frac{1}{2})^{-\frac{1}{2}} X^\frac{1}{2} = S^{-\frac{1}{2}} (S^\frac{1}{2} X S^\frac{1}{2})^\frac{1}{2} S^{-\frac{1}{2}},$$

and $D = P^\frac{1}{2}$. The matrix $D$ can be used to rescale $X$ and $S$ to the same matrix $V$, defined by

$$V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} D S D.$$  \hfill (9)

Furthermore, we define

$$\bar{A}_i := \frac{1}{\sqrt{\mu}} DA_i D, \quad i = 1, \ldots, m, \quad D_X := \frac{1}{\sqrt{\mu}} D^{-1} \Delta X D^{-1}, \text{ and } D_S := \frac{1}{\sqrt{\mu}} D \Delta S D.$$  \hfill (10)

From (9) and (10), after some elementary reductions, we have

$$\sum_{i=1}^m \Delta y_i \bar{A}_i + D_S = 0,$$

$$D_X + D_S = V^{-1} - V.$$  \hfill (11)

The scaled NT search direction $(D_X, \Delta y, D_S)$ is computed by solving (11) so that $\Delta X$ and $\Delta S$ are obtained through (10).

So far we described the scheme that defines the classical NT search direction. Now, following [19] we turn to the new approach of this paper. Given the kernel function $\psi(t)$ as given by (3), we briefly recall the definition of matrix function (see, e.g., [9,14,19]).

Let $V \in S_{++}^n$ and

$$V = Q^T \text{diag}(\lambda_1(V), \lambda_2(V), \ldots, \lambda_n(V)) Q$$  \hfill (12)

be an eigenvalue decomposition of $V$, where $\lambda_i(V)$ with $i = 1, \ldots, n$ denote the eigenvalues of $V$ and $Q$ is any orthonormal matrix that diagonalizes $V$. Then, the (matrix valued) matrix function $\psi(V) : S_{++}^n \rightarrow S^n$ is defined by

$$\psi(V) = Q^T \text{diag}(\psi(\lambda_1(V)), \psi(\lambda_2(V)), \ldots, \psi(\lambda_n(V))) Q.$$  \hfill (13)

Note that $\psi(V)$ depends only on the restriction of $\psi(t)$ to the set of eigenvalues of the symmetric matrix $V$. Since $\psi(t)$ is differentiable, the derivative $\psi'(t)$ is well-defined for $t > 0$. Hence, replacing $\psi(\lambda_i(V))$ in (13) by $\psi'(\lambda_i(V))$, we obtain that the matrix function $\psi'(V)$ is defined as well.

Furthermore, we define the real valued matrix function $\Psi(V) : S_{+++}^n \rightarrow \mathbb{R}_+$ as follows

$$\Psi(V) := \Psi(X,S;\mu) = \text{Tr}(\psi(V)) = \sum_{i=1}^n \psi(\lambda_i(V)),$$  \hfill (14)

where $\psi(V)$ is given by (13). It should be noted that the gradient of $\Psi(V)$ is $\psi'(V)$.

A crucial observation is that the right-hand side $V^{-1} - V$ in the third equation of the system (11) equals minus the gradient of the classical logarithmic barrier function

$$\Psi_c(V) := \text{Tr}(\psi_c(V)) = \sum_{i=1}^n (\psi_c(\lambda_i(V)),$$  \hfill (15)
where \( \psi_c(t) \) is the kernel function of the classic barrier function given by (4). That is to say

\[
V^{-1} - V = -\nabla \Psi_c(V),
\]

where \( \nabla \Psi_c(V) \) denotes the gradient of \( \Psi_c(V) \), i.e., \( \psi'_c(V) \). Hence, the system (11) is equivalent to

\[
\begin{align*}
\bar{A}_i \cdot D_X &= 0, \ i = 1, \ldots, m, \\
\sum_{i=1}^m \Delta y_i \bar{A}_i + D_S &= 0, \\
D_X + D_S &= -\nabla \Psi_c(V).
\end{align*}
\]

This means that the logarithmic barrier function essentially determines the classical NT search direction.

Now, we replace the right-hand side \(-\nabla \Psi_c(V)\) in the third equation in (16) by \(-\nabla \Psi(V)\), i.e., \(-\psi'(V)\). This yields

\[
\begin{align*}
\bar{A}_i \cdot D_X &= 0, \ i = 1, 2, \ldots, m, \\
\sum_{i=1}^m \Delta y_i \bar{A}_i + D_S &= 0, \\
D_X + D_S &= -\nabla \Psi(V).
\end{align*}
\]

The scaled new search direction \((D_X, \Delta y, D_S)\) is computed by solving (17) so that \(\Delta X\) and \(\Delta S\) are obtained through (10). If \((X, y, S) \neq (X(\mu), y(\mu), S(\mu))\), then \((\Delta X, \Delta y, \Delta S)\) is nonzero. By taking a default step size along the search directions, we get the new iterate \((X_+, y_+, S_+)\) according to

\[
X_+ := X + \alpha \Delta X, \ y_+ := y + \alpha \Delta y, \ S_+ := S + \alpha \Delta S.
\]

One can easily verify that

\[
XS = \mu E \iff V = E \iff \psi'(V) = 0 \iff \Psi(V) = 0.
\]

Hence, the value of \(\Psi(V)\) can be considered as a measure for the distance between the given iterate \((X, y, S)\) and the \(\mu\)-center \((X(\mu), y(\mu), S(\mu))\).

The above discussion can be summarized in the form of the generic kernel-based primal-dual IPMs for SDO presented in Figure 1.

3. **Properties of the parametric kernel (barrier) function.** Firstly, we recall some important properties of the parametric kernel function \(\psi(t)\). The detailed can be found in [3].

**Lemma 3.1** (Lemma 3 in [3]). Let \( t > 0 \). Then

\[
\begin{align*}
\psi''(t) > 1, \\
t \psi''(t) + \psi'(t) > 0, \\
t \psi''(t) - \psi'(t) > 0, \\
\psi'''(t) < 0.
\end{align*}
\]

The condition (19) of Lemma 3.1, i.e., \( \psi''(t) > 1 \), implies that the parametric kernel function \(\psi(t)\) is strongly convex. The following lemma provides an important consequence of this property.
Input:
- A threshold parameter $\tau \geq 1$;
- an accuracy parameter $\varepsilon > 0$;
- a fixed barrier update parameter $\theta$, $0 < \theta < 1$;
- a strictly feasible pair $(X^0, y^0, S^0)$ and $\mu^0 = 1$ such that $\Psi(X^0, S^0; \mu^0) \leq \tau$.

begin
  $X := X^0$; $y := y^0$; $S := S^0$; $\mu := \mu^0$;
  while $n\mu \geq \varepsilon$ do
    $\mu := (1 - \theta)\mu$;
    while $\Psi(X, S; \mu) > \tau$ do
      calculate the search direction $(\Delta X, \Delta y, \Delta S)$;
      determine the default step size $\alpha$;
      update $(X, y, S) := (X, y, S) + \alpha(\Delta X, \Delta y, \Delta S)$.
  end
end

Figure 1. Generic Primal-Dual IPM for SDO

Lemma 3.2 (Lemma 5 in [3]). Let $t > 0$. Then
$$\frac{1}{2}(t - 1)^2 \leq \psi(t) \leq \frac{1}{2}\psi'(t)^2.$$  

Lemma 3.3. Let $\varrho : [0, \infty) \to [1, \infty)$ be the inverse function of the parametric kernel function $\psi(t)$ for $t \geq 1$. Then
$$\varrho(s) \leq 1 + \sqrt{2}s.$$  

Proof. It follows immediately from the first inequality of Lemma 3.2 that the result is obvious. This completes the proof of the lemma.

Lemma 3.4 (Lemma 2.6 in [2]). If the kernel function $\psi(t)$ satisfies (22), then
$$\frac{1}{2}\psi''(t)(t - 1)^2 < \psi(t) < \frac{1}{2}\psi''(1)(t - 1)^2, \ t < 1,$$
$$\frac{1}{2}\psi''(1)(t - 1)^2 < \psi(t) < \frac{1}{2}\psi''(1)(t - 1)^2, \ t > 1.$$  

The condition (20) of Lemma 3.1, i.e., $t\psi''(t) + \psi'(t) > 0$, implies that the parametric kernel function has the exponential convexity.

Lemma 3.5 (Lemma 4 in [3]). Let $t_1, t_2 \geq 0$. Then
$$\psi(\sqrt{t_1t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)).$$  

The exponential convexity of the parametric kernel function implies the exponential convexity of the associated barrier function.

Lemma 3.6 (Proposition 3 in [14]). Let $V_1, V_2 \in S_+^n$. Then
$$\Psi((V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}})^\frac{1}{2}) \leq \frac{1}{2}(\Psi(V_1) + \Psi(V_2)).$$
The norm-based proximity measure \( \delta(V) : S^n_{++} \to \mathbb{R}_+ \) is given by

\[
\delta(V) := \frac{1}{2}\|\nabla \Psi(V)\| = \frac{1}{2} \sqrt{\sum_{i=1}^{n} \psi'(\lambda_i(V))^2}.
\]

(23)

As a consequence of Lemma 3.2, the following result obtained in the LO case.

**Corollary 3.1** (Corollary 6 in [3]). Let \( v \in \mathbb{R}^n_{++} \) and \( \Psi(v) \geq 1 \). Then

\[
\delta(v) \geq \sqrt{\frac{\Psi(v)}{2}}.
\]

Note that \( \Psi(V) \) and \( \delta(V) \) depend only on the eigenvalues \( \lambda_i(V) \) of the symmetric matrix \( V \). From Corollary 3.1, we have

**Theorem 3.7.** Let \( V \in S^n_{++} \) and \( \Psi(V) \geq 1 \). Then

\[
\delta(V) \geq \sqrt{\frac{\Psi(V)}{2}}.
\]

It should be mentioned that during the course of the algorithm the largest value of \( \Psi(V) \) occur just after the update of \( \mu \). So next we need derive an estimate for the effect of a \( \mu \)-update on the value of \( \Psi(V) \). For this purpose, we recall the corresponding result in the LO case.

**Theorem 3.8** (Theorem 9 in [3]). Let \( v \in \mathbb{R}^n_{++} \) and \( v_+ = \frac{v}{\sqrt{1+\theta}} \) with \( 0 < \theta < 1 \). Then

\[
\Psi(v_+) \leq \Psi(v) + \frac{\theta}{2(1-\theta)} \left( 2\Psi(v) + 2\sqrt{2n\Psi(v)} + n \right).
\]

By applying Theorem 3.8, with \( v \) being the vector in \( \mathbb{R}^n \) consisting of all the eigenvalues \( \lambda_i(V) \) of the symmetric matrix \( V \), the theorem below immediately follows.

**Theorem 3.9.** Let \( V \in S^n_{++} \) and \( v_+ = \frac{V}{\sqrt{1+\theta}} \) with \( 0 < \theta < 1 \). Then

\[
\Psi(V_+) \leq \Psi(V) + \frac{\theta}{2(1-\theta)} \left( 2\Psi(V) + 2\sqrt{2n\Psi(V)} + n \right).
\]

For any function \( \psi(t) \), let us denote by \( \Delta \psi \) the divided difference of \( \psi(t) \):

\[
\Delta \psi(t_1, t_2) = \frac{\psi(t_1) - \psi(t_2)}{t_1 - t_2}, \forall t_1, t_2 \in \mathbb{R}.
\]

If \( t_1 = t_2 = t \), we simply write \( \Delta \psi(t, t) = \psi'(t) \). Now we have

**Theorem 3.10** (Lemma 16 in [14]). Suppose that \( H(t) \) is a matrix of functions such that the matrix \( H(t) \) is positive definite with eigenvalues \( \lambda_1(t) \geq \lambda_2(t) \geq \cdots \geq \lambda_n(t) > 0 \). If \( H(t) \) is twice differentiable with respect to \( t \in (l_t, u_t) \) and \( \psi(t) \) is twice continuously differentiable function in a domain that contains all the eigenvalues of \( H(t) \), then

\[
\frac{d}{dt} \text{Tr}(\psi(H(t))) = \text{Tr}(\psi'(H(t))H'(t)),
\]

\[
\frac{d^2}{dt^2} \text{Tr}(\psi(H(t))) \leq \omega \|H'(t)\|^2 + \text{Tr}(\psi'(H(t))H''(t)),
\]

where

\[
\omega = \max\{|\Delta \psi'(\lambda_j(t), \lambda_k(t))| : t \in (l_t, u_t), j, k = 1, 2, \ldots, n\}.
\]
4. Complexity and analysis of the algorithms. From (18) and (10), we have
\[ X_+ = X + \alpha \Delta X = X + \alpha \sqrt{\mu} DD_X D = \sqrt{\mu} D(V + \alpha D_X) D \]
and
\[ S_+ = S + \alpha \Delta S = S + \alpha \sqrt{\mu} D^{-1} D S D^{-1} = \sqrt{\mu} D^{-1} (V + \alpha D_S) D^{-1}. \]
It follows from (9) that
\[ V_+ = \frac{1}{\sqrt{\mu}} (D^{-1} X_+ S_+ D)^{\frac{1}{2}}. \]
We can verify that \( V_+^2 \) is unitarily similar to the matrix \( X_+^{\frac{1}{2}} S_+ X_+^{\frac{1}{2}} \) and thus to
\( (V + \alpha D_X)^{\frac{1}{2}} (V + \alpha D_S)(V + \alpha D_X)^{\frac{1}{2}} \). This implies that the eigenvalues of \( V_+ \) are precisely the same as those of the matrix
\[ \mathbf{V}_+ := \left( (V + \alpha D_X)^{\frac{1}{2}} (V + \alpha D_S)(V + \alpha D_X)^{\frac{1}{2}} \right)^{\frac{1}{2}}. \]
From the definition of \( \Psi(V) \), we obtain \( \Psi(V_+) = \Psi(\mathbf{V}_+) \). Theorem 3.6 implies that
\[ \Psi(V_+) = \Psi(\mathbf{V}_+) \leq \frac{1}{2} (\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)). \]
Now, we consider the decrease in \( \Psi(V) \) as a function of \( \alpha \) and define
\[ f(\alpha) := \Psi(V_+) - \Psi(V) = \Psi(\mathbf{V}_+) - \Psi(V). \]
Furthermore, we define
\[ f_1(\alpha) := \frac{1}{2} (\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S) - \Psi(V). \]
It follows that \( f(\alpha) \leq f_1(\alpha) \), which means that \( f_1(\alpha) \) gives an upper bound for the
decrease of the barrier function \( \Psi(V) \). It is worth pointing out that \( f_1(\alpha) \) is convex
and in general \( f(\alpha) \) is not convex. That is an important advantage of using the
function \( f_1(\alpha) \) instead of using the original decrease function \( f(\alpha) \). Moreover, we have
\( f(0) = f_1(0) = 0 \).
From Lemma 3.10, we have
\[ f_1'(\alpha) = \frac{1}{2} \text{Tr} (\psi'(V + \alpha D_X) D_X + \psi'(V + \alpha D_S) D_S), \quad (24) \]
and
\[ f_1''(\alpha) = \frac{1}{2} \frac{d^2}{d\alpha^2} \text{Tr} (\psi(V + \alpha D_X) + \psi(V + \alpha D_S)) \leq \frac{1}{2} (\omega_1 \| D_X \|^2 + \omega_2 \| D_S \|^2), \quad (25) \]
where
\[ \omega_1 = \max \{|\Delta \psi'(\lambda_j(V + \alpha D_X), \lambda_k(V + \alpha D_X))| : j, k = 1, \ldots, n\}, \]
and
\[ \omega_2 = \max \{|\Delta \psi'(\lambda_j(V + \alpha D_S), \lambda_k(V + \alpha D_S))| : j, k = 1, \ldots, n\}. \]
Hence, using the third equation of the system (16), we have
\[ f_1'(0) = \frac{1}{2} \text{Tr} (\psi'(V)(D_X + D_S)) = \frac{1}{2} \text{Tr} (-\psi'(V)^2) = -2\delta(V)^2. \quad (26) \]
In order to facilitate discussion, we denote \( \delta := \delta(V) \), and we have the main result, which provides an upper bound of \( f_1''(\alpha) \).

**Theorem 4.1** (Lemma 5.19 in [19]). One has
\[ f_1''(\alpha) \leq 2\delta^2 \psi''(\lambda_{\min}(V) - 2\alpha \delta). \]
The default step size for the algorithm should be chosen such that $X_+$ and $S_+$ are feasible and $\Psi(V_+) - \Psi(V)$ decreases sufficiently. For the details we leave it for the interested readers (see, e.g., [2, 3, 14]. Following the strategy considered in [3], we briefly recall how to choose the default step size. Suppose that the step size $\alpha$ satisfies

$$-\psi'(\lambda_{\min}(V) - 2\alpha \delta) + \psi'(\lambda_{\min}(V)) \leq 2\delta.$$  \hspace{1cm} (27)

Then $f_1(\alpha) \leq 0$. The largest possible value of the step size of $\alpha$ satisfying (27) is given by

$$\bar{\alpha} := \frac{1}{2\delta} (\rho(\delta) - \rho(2\delta)).$$  \hspace{1cm} (28)

where $\rho(s) : [0, \infty) \to (0, 1]$ is the inverse function of $-\frac{1}{2}\psi'(t)$ for $t \in (0, 1]$. Furthermore, we can conclude that

$$\frac{1}{\psi''(\rho(2\delta))} \leq \bar{\alpha} \leq \frac{1}{\psi''(\rho(\delta))}.$$  \hspace{1cm} (29)

After some elementary reductions, we can conclude that

$$\bar{\alpha} \geq \frac{1}{C(\lambda)\delta^{\frac{3}{4}}}.$$  \hspace{1cm} (30)

where

$$C(\lambda) := 2^{\frac{3}{4}} + 9\pi^2(2^{\frac{3}{4}} + \left(\frac{20}{\lambda \pi}\right)^{\frac{1}{4}})^2 + 75\sqrt{2} + \frac{25\lambda \pi^2}{8}(2^{\frac{1}{4}} + \left(\frac{20}{\lambda \pi}\right)^{\frac{1}{4}} + 3\left(\frac{10}{\lambda \pi}\right)^{\frac{1}{4}}).$$  \hspace{1cm} (31)

In the sequel, we use

$$\tilde{\alpha} := \frac{1}{C(\lambda)\delta^{\frac{3}{4}}}$$  \hspace{1cm} (32)

as the default step size, which essentially depends only on the norm $\delta$ and $C(\lambda)$. It is clear that $\bar{\alpha} \geq \tilde{\alpha}$.

In what follows, we will show that the barrier function $\Psi(V)$ in each inner iteration with the default step size $\alpha$, as defined by (32), is decreasing. For this, we need the following technical result.

**Lemma 4.2** (Lemma 12 in [14]). Let $h(t)$ be a twice differentiable convex function with $h(0) = 0$, $h'(0) < 0$ and let $h(t)$ attain its (global) minimum at $t^* > 0$. If $h''(t)$ is increasing for $t \in [0, t^*)$, then

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

As a consequence of Lemma 4.2 and the fact that $f(\alpha) \leq f_1(\alpha)$, which is a twice differentiable convex function with $f_1(0) = 0$, and $f_1'(0) = -2\delta^2 < 0$, the following lemma is obvious.

**Lemma 4.3.** Let the step size $\alpha$ be such that $\alpha \leq \tilde{\alpha}$. Then

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

The following lemma shows that the default step size (32) yields sufficient decrease of the barrier function during each inner iteration.

**Lemma 4.4.** Let $\tilde{\alpha}$ be the default step size as given by (32). Then

$$f(\tilde{\alpha}) \leq -\left(\sqrt[3]{2}C(\lambda)\right)^{-1}\Psi(V)^{\frac{1}{3}}.$$  


After updating the parameter $\mu$ to $(1-\theta)\mu$ with $0 < \theta < 1$, we have, by Theorem 3.9,
\[
\Psi(V_+) \leq \Psi(V) + \frac{\theta}{2(1-\theta)} \left( 2\Psi(V) + 2\sqrt{2n\Psi(V) + n} \right). 
\tag{33}
\]
At the start of an outer iteration and just before updating the parameter $\mu$, we have $\Psi(V) \leq \tau$. It follows from (33) that the value of $\Psi(V)$ exceeds the threshold $\tau$ after updating of $\mu$. Therefore, we need to count how many inner iterations are required to return to the situation where $\Psi(V) \leq \tau$. We denote the value of $\Psi(V)$ after the $\mu$-update as $\Psi_0$, the subsequent values in the same outer iteration are denoted as $\Psi_k$, $k = 1, \ldots, K$, where $K$ denotes the total number of inner iterations in the outer iteration. Hence, we have
\[
\Psi_0 \leq \tau + \frac{\theta}{2(1-\theta)} \left( 2\tau + 2\sqrt{2n\tau + n} \right). 
\tag{34}
\]
According to the decrease of $f(\tilde{\alpha})$ in Lemma 4.4, we have
\[
\Psi_{k+1} \leq \Psi_k - \beta(\Psi_k)^{1-\gamma}, \quad k = 0, 1, \ldots, K - 1, 
\tag{35}
\]
where $\beta = (\sqrt{2}C(\lambda))^{-1}$, and $\gamma = \frac{2}{3}$.

**Lemma 4.5** (Lemma 14 in [14]). Let $t_0, t_1, \ldots, t_K$ be a sequence of positive numbers such that
\[
t_{k+1} \leq t_k - \beta t_k^{1-\gamma}, \quad k = 0, 1, \ldots, K - 1,
\]
where $\beta > 0$ and $0 < \gamma \leq 1$. Then $K \leq \lceil \frac{t_0}{\beta\gamma} \rceil$.

The following lemma provides an estimate for the number of inner iterations between two successive barrier parameter updates, in terms of $\Psi_0$ and the constant $C(\lambda)$.

**Lemma 4.6.** One has
\[
K \leq \frac{3\sqrt{2}C(\lambda)}{2} (\Psi_0)^{\frac{\gamma}{2}}. 
\]

**Proof.** From Lemma 4.5 and (35), the result of the lemma follows.

The following lemma provides an upper bound of the number of outer iterations.

**Lemma 4.7** (Lemma II.17 in [18]). If the barrier parameter $\mu$ has the initial value $\mu^0$ and is repeatedly multiplied by $1 - \theta$ with $0 < \theta < 1$, then after at most
\[
\left\lceil \frac{1}{\theta} \log \frac{n\mu^0}{\varepsilon} \right\rceil
\]
steps we have $n\mu \leq \varepsilon$.

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,
\[
\frac{3\sqrt{2}C(\lambda)}{2\theta} \left( \tau + \frac{\theta}{2(1-\theta)} \left( 2\tau + 2\sqrt{2n\tau + n} \right) \right)^{\frac{3}{2}} \log \frac{n\mu^0}{\varepsilon}. 
\tag{36}
\]

**Theorem 4.8.** For large-update methods, one takes for $\theta$ a constant (independent on $n$), namely $\theta = \Theta(1)$, and $\tau = O(n)$. The iteration bound then becomes
\[
O \left( n^{\frac{3}{2}} \log \frac{n}{\varepsilon} \right),
\]
which improves the classical iteration bound with a factor $n^{\frac{3}{2}}$ for large-update methods.
For the analysis of the iteration bound of small-update methods, we need to estimate the upper bound of $\Psi_0$ more accurately. This is due to the following lemma.

**Lemma 4.9** (Corollary 4.17 in [20]). Let $V \in S_{++}^n$ and $V_+ = \frac{V}{\sqrt{1-\theta}}$ with $0 < \theta < 1$. If $\Psi(V) \leq \tau$, then

$$
\Psi(V_+) \leq n\psi\left(\frac{\rho(\frac{\sqrt{n}}{\tau})}{\sqrt{1-\theta}}\right).
$$

From Lemma 4.9, Lemma 3.4 with $\psi''(t) > 1$, Lemma 3.3, and the fact that $1 - \sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \leq \theta$, we have

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

(37)

Applying Theorem 4.6 again, the total number of iterations is bounded above by

$$
\frac{3\sqrt{2}C(\lambda)}{2\theta} \left(\frac{1}{1-\theta} \left(\theta\sqrt{n} + \sqrt{2\tau}\right)^2\right)^{\frac{3}{2}} \log \frac{n\mu^0}{\epsilon}.
$$

(38)

After some elementary reductions, we have the following theorem, which gives the currently best known iteration bound for small-update methods.

**Theorem 4.10.** For small-update methods, one takes $\theta = \Theta\left(\frac{1}{\sqrt{n}}\right)$ and $\tau = O(1)$. The iteration bound then becomes

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

which matches the currently best known iteration bound for small-update methods.

5. Conclusions and remarks. In this paper, we have shown that a class of large- and small-update primal-dual IPMs for LO based on the parametric kernel function with a trigonometric barrier term presented in [3] can be extended to the context of SDO. By utilizing the feature of the parametric kernel function, we obtained the same iteration bounds as in the LO case, both for large- and small-update methods. Although expected, these results were not obvious and, at certain steps of the analysis, they were not trivial and/or straightforward generalization of the LO case.

Some interesting topics for further research remain. One interesting topic is to investigate whether it is possible to replace the NT-scaling scheme by some other scaling schemes and still obtain the polynomial-time iteration bounds. Furthermore, numerical results may help us to compare the behavior of the algorithms of the paper with the existing methods.

**Acknowledgments.** The authors would like to thank the anonymous referees for their useful comments and suggestions, which helped to improve the presentation of this paper. This work is supported by Shanghai Natural Science Fund Project (No. 14ZR1418900), National Natural Science Foundation of China (No. 11471211), Shanghai Municipal Education Commission Research Foundation (No. ZZGJD12061).
REFERENCES

[1] M. Achache and L. Guerra, A full Nesterov-Todd step feasible primal-dual interior-point algorithm for convex quadratic semidefinite optimization, *Applied Mathematics and Computation*, **231** (2014), 581–590.

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

[3] X. Z. Cai, G. Q. Wang, M. El Ghami and Y. J. Yue, Complexity analysis of primal-dual interior-point methods for linear optimization based on a parametric kernel function with a trigonometric barrier term, *Abstract and Applied Analysis*, **2014** (2014), 710158.

[4] B. K. Choi and G. M. Lee, On complexity analysis of the primal-dual interior-point method for semidefinite optimization problem based on a new proximity function, *Nonlinear Analysis, Theory, Methods & Applications*, **71** (2009), 2628–2640.

[5] Zs. Darvay, New interior point algorithms in linear programming, *Advanced Modeling and Optimization*, **5** (2003), 51–92.

[6] E. De Klerk, *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

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

[8] M. El Ghami, C. Roos and T. Steihaug, A generic primal-dual interior-point method for semidefinite optimization based on a new class of kernel functions, *Optimization Methods & Software*, **25** (2010), 387–403.

[9] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, UK, 1991.

[10] B. Kheirfam, Simplified infeasible interior-point algorithm for SDO using full Nesterov-Todd step, *Numerical Algorithms*, **59** (2012), 589–606.

[11] M. Kojima, M. Shida and S. Shindoh, Local convergence of predictor-corrector infeasible-interior-point method for SDPs and SDLCPs, *Mathematical Programming*, **80** (1998), 129–160.

[12] H. W. Liu, C. H. Liu and X. M. Yang, New complexity analysis of a Mehrotra-type predictor-corrector algorithm for semidefinite programming, *Optimization Methods & Software*, **28** (2013), 1179–1194.

[13] H. Mansouri and C. Roos, A new full-Newton step $O(n)$ infeasible interior-point algorithm for semidefinite optimization, *Numerical Algorithms*, **52** (2009), 225–255.

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

[15] M. R. Peyghami, An interior-point approach for semidefinite optimization using new proximity functions, *Asia-Pacific Journal of Operational Research*, **26** (2009), 365–382.

[16] M. R. Peyghami and S. F. Hafshejani, Complexity analysis of an interior-point algorithm for linear optimization based on a new proximity function, *Numerical Algorithms*, **67** (2014), 33–48.

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

[18] C. Roos, T. Terlaky and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization*, Springer, New York, 2005.

[19] G. Q. Wang and Y. Q. Bai, A new primal-dual path-following interior-point algorithm for semidefinite optimization, *Journal of Mathematical Analysis and Applications*, **353** (2009), 339–349.

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

[21] G. Q. Wang and D. T. Zhu, A unified kernel function approach to primal-dual interior-point algorithms for convex quadratic SDO, *Numerical Algorithms*, **57** (2011), 537–558.

[22] L. P. Zhang, Y. H. Xu and Z. J. Jin, An efficient algorithm for convex quadratic semidefinite optimization, *Numerical Algebra, Control and Optimization*, **2** (2012), 129–144.
[23] M. W. Zhang, A large-update interior-point algorithm for convex quadratic semidefinite optimization based on a new kernel function, *Acta Mathematica Sinica (English Series)*, 28 (2012), 2313–2328.

[24] Y. Zhang, On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming, *SIAM Journal on Optimization*, 8 (1998), 365–386.

Received October 2014; 1st revision January 2015; final revision May 2015.

E-mail address: guoq_wang@hotmail.com
E-mail address: wzcz@sues.edu.cn
E-mail address: zhongtuanzheng@163.com
E-mail address: xzhcai@163.com