Journal of Mathematical Modeling
Vol. 4, No. 1, 2016, pp. 35-58

JMM

A path following interior-point algorithm for semidefinite optimization problem based on new kernel function

El Amir Djeffal[a*] **and Lakhdar Djeffal**[b]

[a] *Department of Mathematics, University of Batna 2, Batna, Algeria*
email: djeffal_elamir@yahoo.fr

[b] *Department of Mathematics, University of Batna 2, Batna, Algeria*
email: lakdar_djeffal@yahoo.fr

Abstract. In this paper, we deal to obtain some new complexity results for solving semidefinite optimization (SDO) problem by interior-point methods (IPMs). We define a new proximity function for the SDO by a new kernel function. Furthermore we formulate an algorithm for a primal dual interior-point method (IPM) for the SDO by using the proximity function and give its complexity analysis, and then we show that the worst-case iteration bound for our IPM is $O(6(m+1)^{\frac{3m+4}{2(m+1)}}\Psi_0^{\frac{m+2}{2(m+1)}}\frac{1}{\theta}\log\frac{n\mu^0}{\varepsilon})$, where $m>4$.

Keywords: quadratic programming, convex nonlinear programming, interior point methods.
AMS Subject Classification: 90C25, 90C30, 90C51, 90C20.

1 Introduction

Infeasible interior-point methods (IIPMs) start with an arbitrary positive point and feasibility is reached as optimality is approached. The choice of the starting point in IPMs is crucial for the performance. Lustig [8] and Tanabe [16] were the first to present IPMs for LO. Kojima et al. [6] were the first that proved the global convergence of a primal–dual IPM for LO.

*Corresponding author.
Received: 3 December 2015 / Revised: 28 May 2016 / Accepted: 29 May 2016.

 http://jmm.guilan.ac.ir

Zhang [18] was the first who presented a primal–dual IPM with polynomial iteration complexity for LO. Primal–dual interior-point methods (IPMs) for semidefinite optimization have been widely studied, the reader is referred to Klerk [4]. Recently a full-Newton step infeasible interior-point algorithm for linear programming (LP) was presented by Roos [14]. The result of polynomial complexity coincides with the best known one for IIPMs. Mansouri and Roos [9, 10] extended this algorithm to semidefinite optimization by using a specific feasibility step. The barrier function is determined by a simple univariate function, called its kernel function. Bai et al. [2] introduced a new barrier function which is not a barrier function in the usual sense.

In this paper, we define a new proximity function for the SDO by a new kernel function. Also, we formulate an interior-point algorithm for SDO by using a new proximity function and give its complexity analysis, and then we show that the worst-case iteration bound for our IPM is $O(6(m+1)^{\frac{3m+4}{2(m+1)}}\Psi_0^{\frac{m+2}{2(m+1)}}\frac{1}{\theta}\log\frac{n\mu^0}{\varepsilon})$. Furthermore, the complexity bounds obtained by the algorithm are $O(m^{\frac{3m+1}{2m}}n^{\frac{m+1}{2m}}\log\frac{Tr(X^0S^0)}{\varepsilon})$ and $O(m^{\frac{3m+1}{2m}}\sqrt{n}\log\frac{Tr(X^0S^0)}{\varepsilon})$, for large and small-update methods, respectively.

The paper is organized as follows. In Section 2, we recall the preliminaries. In Section 3, we define a new kernel function and give its properties which are essential for the complexity analysis. In Section 4, we derive the complexity result for both large-update and small-update methods. Finally, concluding remarks are given in Section 5.

Some of the notations used throughout the paper are as follows: $\Re^n$, $\Re^n_+$ and $\Re^n_{++}$ denote the set of vectors with n components, the set of nonnegative vectors and the set of positive vectors, respectively. $\Re^{n\times n}$ denotes the set of $n\times n$ real matrices. $\|.\|_F$ and $\|.\|_2$ denote the Frobenius norm and the spectral norm for matrices, respectively. S_n, S_n^+ and S_n^{++} denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n\times n$ matrices, respectively. E denotes $n\times n$ identity matrix. 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). We use the matrix inner product $A\bullet B = Tr(A^TB)$. For any $Q\in S_n^{++}$, the expression $Q^{\frac{1}{2}}$ (or $\sqrt{Q}$) denotes its symmetric square root. For any $V\in S_n^+$, we define $\lambda_{\min}(V)$ to be the minimal eigenvalue of V.

2 Preliminaries

2.1 The central path

We consider the semidefinite optimization problem in the following form:

$$\min \{Tr(CX) : \ Tr(A_iX) = b_i, i = 1, ..., m, \ X \succeq 0\}, \tag{SDO}$$

and its dual

$$\max \left\{ b^t y : \ \sum_{i=1}^{m} y_i A_i + S = C, \ S \succeq 0 \right\}, \tag{SDD}$$

where $A_i \in \mathbb{R}^{m\times n}, i = 1, ..., m$ and C are symmetric $n \times n$ matrices, and $b, y \in \Re^m$.
Throughout the paper, we make the following assumptions:

Assumption 1: The matrices $A_i \in \mathbb{R}^{m\times n}, i = 1, ..., m$ are linearly independent.

Assumption 2: The initial iterate (X^0, y^0, S^0) is strictly feasible:

$$Tr(A_iX^0) = b_i, i = 1, ..., m, \ X^0 \succ 0, \quad \sum_{i=1}^{m} y_i^0 A_i + S^0 = C, \ S \succ 0.$$

We have the following well-known lemma:

Lemma 1. *[13] The following statements are equivalent:*
1) $X \succeq 0, \ S \succeq 0$ and $Tr(XS) = 0$,
2) $X \succeq 0, \ S \succeq 0$ and $\left\| X^{\frac{1}{2}} S^{\frac{1}{2}} \right\|^2 = 0$,
3) $X \succeq 0, \ S \succeq 0$ and $XS = 0$.

It is well known that finding an optimal solution (X^*, y^*, S^*) of SDO and SDD is equivalent to solve the following system:

$$\begin{cases} Tr(A_iX) = b_i, i = 1, ..., m, \ X \succeq 0, \\ \sum\limits_{i=1}^{m} y_i A_i + S = C, \ S \succeq 0, \\ XS = 0. \end{cases} \tag{1}$$

The basic idea of primal dual IPMs is to replace the third equation in (1), the so-called complementarity condition for SDO and SDD, by the parameterized equation $XS = \mu E$ with $\mu > 0$, where E denotes the $n \times n$ identity matrix. Thus we consider the system:

$$\begin{cases} Tr(A_i X) = b_i, i = 1, ..., m, \;\; X \succeq 0, \\ \sum\limits_{i=1}^{m} y_i A_i + S = C, \;\; S \succeq 0, \\ XS = \mu E. \end{cases} \tag{2}$$

If both SDO and SDD satisfy IPC, then for each $\mu > 0$ the parameterized system (2) has a unique solution $(X(\mu), y(\mu), S(\mu))$ (see [7, 17]), which is called μ-center of SDO and SDD. The set of μ-centers, that is, $\Lambda = \{(X(\mu), y(\mu), S(\mu)) \,/ \mu > 0\}$, is called the central path of SDO and SDD. The central path converges to the solution pair of SDO and SDD as reduces to zero [11, 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 μ-center and the second one is called the outer iteration scheme, is to decrease the parameter μ to $\mu^+ = (1 - \theta)\mu$, for some $\theta \in (0, 1)$.

2.2 The search directions

IPMs follow the central path approximately. We briefly describe the usual approach. Without loss of generality, we assume that $(X(\mu), y(\mu), S(\mu))$ is known for some positive μ. For example, due to the above assumption, we may assume this for $\mu^0 = \frac{Tr(X^0 S^0)}{n}$, with $X^0 \succ 0$ and $S^0 \succ 0$. We then decrease μ to $\mu^+ = (1 - \theta)\mu$, for some $\theta \in (0, 1)$, and solve the following Newton system:

$$\begin{cases} Tr(A_i \triangle X) = 0, \;\; i = 1, ..., m, \\ \sum\limits_{i=1}^{m} \triangle y_i A_i + \triangle S = 0, \\ X \triangle S + \triangle X S = \mu E - XS. \end{cases} \tag{3}$$

We consider the symmetrization scheme that yields the Nesterov-Todd (NT) direction, let us define the matrix:

$$P = X^{\frac{1}{2}} (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{-\frac{1}{2}} X^{\frac{1}{2}} = S^{-\frac{1}{2}} (S^{\frac{1}{2}} X S^{\frac{1}{2}})^{\frac{1}{2}} S^{-\frac{1}{2}}, \tag{4}$$

and also define $D = P^{\frac{1}{2}}$, the matrix D can be used to scale X and S to same matrix V because

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

Note that the matrices D and V are symmetric and positive semidefinte, similary to the SDO [10]. We can conclude that the system (1) has a unique symmetric solution, let use define

$$\begin{array}{l} \overline{A}_i := \frac{1}{\sqrt{\mu}} D A_i D, \;\; i = 1, \dots, m, \\ D_X := \frac{1}{\sqrt{\mu}} D^{-1} \triangle X D^{-1}, \\ D_S := \frac{1}{\sqrt{\mu}} D \triangle S D. \end{array} \tag{6}$$

The $NT-$ search direction are obtained from the system:

$$\begin{cases} Tr(\overline{A}_i D_X) = 0, \;\; i = 1, ..., m, \\ \sum\limits_{i=1}^{m} \triangle y_i \overline{A}_i + D_S = 0, \\ D_X + D_S = V^{-1} - V = -\nabla \Psi_l(V). \end{cases} \tag{7}$$

Clearly, $Tr(D_X D_S) = 0$ which is concluded from the first and second equations of (7) or from the orthogonality of $\triangle X$ and $\triangle S$. The classical kernel function defined as follows:

$$\psi_l(t) = \frac{1}{2}(t^2 - 1) - \log t.$$

2.3 The generic interior-point algorithm for SDO

We call $\psi_l(t)$ the kernel function of the logarithmic barrier function $\Psi_l(V)$. In this paper, we replace $\psi_l(t)$ with a new kernel function $\psi(t)$ which is defined in the next Section and assume that $\tau \geq 1$.

The new interior-point algorithm works as follows. Assume that we are given a strictly feasible point (X, y, S) which is in a τ-neighborhood of the given μ-center. Then we decrease μ to $\mu^+ = (1-\theta)\mu$, for some fixed $\theta \in (0, 1)$ and then we solve the Newton system (3) to obtain the unique search direction. The positivity condition of a new iterate is ensured with the right choice of the step size α which is defined by some line search rule. This procedure is repeated until we find a new iterate (X^+, y^+, S^+) that is in a τ-neighborhood of the μ^+-center. 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, i.e., $n\mu \leq \varepsilon$.

The parameters τ, θ and the step size α should be chosen in such a way that the algorithm is optimized in the sense that the number of iterations required by the algorithm is as small as possible. The choice of the so-called barrier update parameter θ plays an important role both in theory and practice of IPMs. Usually, if θ is a constant independent of the dimension n of the problem, for instance $\theta = \frac{1}{2}$, then we call the algorithm a large-update (or long-step) method. If θ depends on the dimension of the problem, such as $\theta = \frac{1}{\sqrt{n}}$, then the algorithm is named a small-update (or short-step) method.

The algorithm for our primal-dual IPM for the SDO is given as follows:

Primal-Dual IPM **for the** SDO

Begin algorithm
    Input:
    An accuracy parameter $\varepsilon > 0$,
    An update parameter θ, $0 < \theta < 1$,
    A threshold parameter $\tau, 0 < \tau < 1$,
    A strictly feasible point (X^0, y^0, S^0) and $\mu^0 = \frac{Tr(X^0 S^0)}{n}$
      such that $\Psi(X^0 S^0, \mu^0) \leq \tau$.
    begin
        $X := X^0, S := S^0, \mu := \mu^0$,
        While $(n\mu) \geq \varepsilon$ do
            begin
                $mu = (1 - \theta)\mu$
                While $(\Psi(V) > \tau)$ do
                    begin
                        Solve system (3) to obtain $(\triangle X, \triangle y, \triangle S)$,
                        Determine a step size α
                        $X := X + \alpha \triangle X$
                        $y := y + \alpha \triangle y$
                        $S := S + \alpha \triangle S$
                End While.
        End While
End algorithm.

In the next section, we define a new kernel function and give its properties which are essential to our complexity analysis.

3 The new kernel function

We call $\psi : \Re_{++} \to \Re_{+}$ a kernel function if ψ is twice differentiable and satisfies the following conditions [1]:

$$\begin{array}{l} \psi^{'}(1) = \psi(1) = 0, \\ \psi^{''}(t) > 0, \\ \lim\limits_{t \to 0^{+}} \psi(t) = \lim\limits_{t \to \infty} \psi(t) = 0. \end{array} \tag{8}$$

For our IPM, we use the following new kernel function:

$$\psi(t) = (m+1)t^2 - (m+2)t + \frac{1}{t^m}, \text{ for all } t > 0, \tag{9}$$

where $m > 4$. Then we have the following:

$$\begin{array}{l} \psi^{'}(t) = 2(m+1)t - (m+2) - mt^{-m-1}, \\ \psi^{''}(t) = 2(m+1) - m(-m-1)t^{-m-2}, \\ \psi^{'''}(t) = -m(-m-1)(-m-2)t^{-m-3}. \end{array} \tag{10}$$

From (8), $\psi(t)$ is clearly a kernel function and

$$\psi^{''}(t) > 2(m+1), \text{ for all } t > 0. \tag{11}$$

In this paper, we replace the function $\Psi_l(V)$ in (7) with the function $\Psi(V)$ as follows:

$$D_X + D_S = -\nabla\Psi(V), \tag{12}$$

where $\Psi(V) = Tr(\psi(V)) = \sum\limits_{i=1}^{n} \psi(\lambda_i(V))$, $\psi(t)$ is defined in (9). Hence, the new search direction $(\triangle X, \triangle y, \triangle S)$ is obtained by solving the following modified Newton system:

$$\begin{cases} Tr(A_i \triangle X) = 0, i = 1, ..., m, \\ \sum\limits_{i=1}^{m} \triangle y_i A_i + \triangle S = 0, \\ X \triangle S + \triangle X S = -\mu V \nabla\Psi(V). \end{cases} \tag{13}$$

Note that D_X and D_S are orthogonal because matrix D_X belongs to null space and matrix D_S to the row space of the matrix $\overline{A_i}, i = 1, ..., m$. Since D_X and D_S are orthogonal, we have

$$\begin{aligned} D_X = D_S = 0 &\Leftrightarrow \nabla\Psi(V) = 0 \\ &\Leftrightarrow V = E \\ &\Leftrightarrow X = X(\mu), S = S(\mu). \end{aligned} \tag{14}$$

We use $\Psi(V)$ as the proximity function to measure the distance between the current iterate and the μ-center for given $\mu > 0$. We also define the norm-based proximity measure, $\delta(V) : \Re_{++} \to \Re_{+}$, as follows:

$$\delta(XS, \mu) = \delta(V) = \frac{1}{2} \|\nabla \Psi(V)\| = \frac{1}{2} \|D_X + D_S\|. \tag{15}$$

Lemma 2. *For $\psi(t)$, we have the following.*
(i) $\psi(t)$ is exponentially convex for all $t > 0$,
(ii) $\psi^{''}(t)$ is monotonically decreasing for all $t > 0$,
(iii) $t\psi^{''}(t) - \psi^{'}(t) > 0,$ for all $t > 0$.

Proof. For (i), by lemma 1 in [11], it suffices to show that the function $\psi(t)$ satisfies $t\psi^{''}(t) + \psi^{'}(t) \geq 0$, for all $t > 0$. Using (10), we have

$$\begin{aligned} t\psi''(t) + \psi'''(t) &= t\left(2(m+1) - m(-m-1)t^{-m-2}\right) \\ &\quad + \left(2(m+1)t - (m+2) - mt^{-m-1}\right) \\ &= 4(m+1)t + m^2 t^{-m-1} - (m+2). \end{aligned}$$

Let

$$g(t) = 4(m+1)t + m^2 t^{-m-1} - (m+2).$$

Then

$$g^{'}(t) = 4(m+1) - m^2(m+1)t^{-m-2},$$

$$g^{''}(t) = m^2(m+1)(m+2)t^{-m-3} > 0, \text{ for all } t > 0.$$

Let $g^{'}(t) = 0$, we get $t = (\frac{m^2}{4})^{\frac{1}{m+2}}$. Since $g(t)$ is strictly convex and has a global minimum, $g((\frac{m^2}{4})^{\frac{1}{m+2}}) > 0$. And by lemma 1 in [13], we have the result.
For (ii), using (10), we have $\psi^{'''}(t) > 0$, so we have the result.
For (iii), using (10), we have

$$t\psi^{''}(t) - \psi^{'}(t) = m(m+2)t^{-m-1} + (m+2) > 0, \text{ for all } t > 0.$$

□

Lemma 3. *For $\psi(t)$, we have the following.*

$$(m+1)(t-1)^2 \leq \psi(t) \leq \frac{1}{4(m+1)} \psi^{'}(t)^2, \textit{ for all } t > 0, \tag{16}$$

$$\psi(t) \leq \frac{(m+1)(m+2)}{2}(t-1)^2, \textit{ for all } t \geq 1. \tag{17}$$

Proof. For (16), using (8) and (11), we have

$$\psi(t)=\int_1^t\int_1^\xi \psi^{''}(\zeta)d\zeta d\xi \geq 2(m+1)\int_1^t\int_1^\xi d\zeta d\xi=(m+1)(t-1)^2,$$

also,

$$\begin{aligned}
\psi(t) &= \int_1^t\int_1^\xi \psi^{''}(\zeta)d\zeta d\xi \\
&\leq \frac{1}{2(m+1)}\int_1^t\int_1^\xi \psi^{''}(\xi)\psi^{''}(\zeta)d\zeta d\xi \\
&= \frac{1}{2(m+1)}\int_1^t \psi^{''}(\xi)\psi^{'}(\xi)\, d\xi \\
&= \frac{1}{2(m+1)}\int_1^t \psi^{'}(\xi)\, d(\psi^{'}(\xi)) \\
&= \frac{1}{4(m+1)}\psi^{'}(t)^2.
\end{aligned}$$

For (17), using Taylor's Theorem, we have

$$\begin{aligned}
\psi(t) &= \psi(1)+\psi^{'}(1)(t-1)+\frac{1}{2}\psi^{''}(1)(t-1)^2+\frac{1}{6}\psi^{'''}(\xi)(\xi-1)^3 \\
&= \frac{1}{2}\psi^{''}(1)(t-1)^2+\frac{1}{6}\psi^{'''}(\xi)(\xi-1)^3 \\
&\leq \frac{1}{2}\psi^{''}(1)(t-1)^2 \\
&= \frac{(m+1)(m+2)}{2}(t-1)^2.
\end{aligned}$$

□

Now, we define $\gamma:(0,\infty)\to(1,\infty)$, be the inverse function of $\psi(t)$ for all $t\geq 1$, and $\rho:(0,\infty)\to(0,1)$, be the inverse function of $-\frac{1}{2}\psi^{'}(t)$ for all $t\in(0,1)$. Then we have the following lemma.

Lemma 4. *For $\psi(t)$, we have the following*

$$\sqrt{\frac{s}{m+1}+1}\leq\gamma(s)\leq 1+\sqrt{\frac{s}{m+1}},\; s\geq 0, \tag{18}$$

and

$$\rho(s) \geq (\frac{m}{2s+m})^{\frac{1}{m+1}}, \;\; s \geq 0. \tag{19}$$

Proof. For (18), let $s = \psi(t),\ t \geq 1,$ i.e., $\gamma(s) = t, t \geq 1$, then we have

$$(m+1)t^2 = s + (m+2)t - t^{-m}.$$

Because $(m+2)t - t^{-m}$ is monotone increasing with respect to $t \geq 1$, we have

$$(m+1)t^2 \geq s + m + 1,$$

which implies that

$$t = \gamma(s) \geq \sqrt{\frac{s}{m+1} + 1}.$$

By (16), we have

$$s = \psi(t) \geq (m+1)(t-1)^2,$$

so

$$t = \gamma(s) \leq 1 + \sqrt{\frac{s}{m+1}}.$$

For (19), let $z = -\frac{1}{2}\psi^{'}(t)$ for all $t \in (0,1)$. By the definition of ρ, we have $\rho(z) = t$ and $2z = -\psi^{'}(t)$. Then

$$mt^{-m-1} = 2z + 2(m+1)t - (m+2).$$

Because $2(m+1)t - (m+2)$ is monotone increasing with respect to $t \in (0,1)$, we have

$$mt^{-m-1} \leq 2z + m,$$

which implies that

$$\rho(z) = t \geq (\frac{m}{2z+m})^{\frac{1}{m+1}}.$$

□

Lemma 5. *Let* $\gamma : (0,\infty) \to (1,\infty)$,*be the inverse function of* $\psi(t)$ *for all* $t \geq 1$. *Then we have*

$$\Psi(\beta V) \leq n\psi\left(\beta\gamma\left(\frac{\Psi(V)}{n}\right)\right), \;\; V \in, \;\; \beta \geq 1. \tag{20}$$

Proof. Using Theorem 3.2 in [1], we get the result. This completes the proof. □

Lemma 6. *Let* $0 \leq \theta \leq 1,\ V^+ = \frac{1}{\sqrt{1-\theta}}V$. *If* $\Psi(V) \leq \tau$, *then we have*

$$\Psi(V^+) \leq \frac{(m+1)(m+2)}{2(1-\theta)} \left(\sqrt{n}\theta + \sqrt{\frac{\tau}{m+1}}\right)^2. \tag{21}$$

Proof. Since $\frac{1}{\sqrt{1-\theta}} \geq 1$ and $\gamma\left(\frac{\Psi(V)}{n}\right) \geq 1$, we have $\frac{\gamma\left(\frac{\Psi(V)}{n}\right)}{\sqrt{1-\theta}} \geq 1$. Using lemma 5, with $\beta = \sqrt{1-\theta}$, (17), (18) and $\Psi(V) \leq \tau$, we have

$$\begin{aligned}
\Psi(V^+) &\leq n\psi\left(\frac{1}{\sqrt{1-\theta}}\gamma\left(\frac{\Psi(V)}{n}\right)\right) \\
&\leq n\frac{(m+1)(m+2)}{2}\left(\frac{1}{\sqrt{1-\theta}}\gamma\left(\frac{\Psi(V)}{n}\right) - 1\right)^2 \\
&= n\frac{(m+1)(m+2)}{2(1-\theta)}\left(\gamma\left(\frac{\Psi(V)}{n}\right) - \sqrt{1-\theta}\right)^2 \\
&\leq n\frac{(m+1)(m+2)}{2(1-\theta)}\left(1 + \sqrt{\frac{\Psi(V)}{(m+1)n}} - \sqrt{1-\theta}\right)^2 \\
&\leq n\frac{(m+1)(m+2)}{2(1-\theta)}\left(\theta + \sqrt{\frac{\tau}{(m+1)n}}\right)^2 \\
&\leq n\frac{(m+1)(m+2)}{2(1-\theta)}\left(\sqrt{n}\theta + \sqrt{\frac{\tau}{(m+1)n}}\right)^2.
\end{aligned}$$

□

Denote

$$\Psi_0 = L(n,\theta,\tau) = n\frac{(m+1)(m+2)}{2(1-\theta)}\left(\sqrt{n}\theta + \sqrt{\frac{\tau}{(m+1)n}}\right)^2, \tag{22}$$

then Ψ_0 is an upper bound for $\Psi(V)$ during the process of the algorithm.

4 Analysis of algorithm

The aim of this paper is to define a new kernel function and to obtain new complexity results for an SDO problem using the proximity function defined by the kernel function and following the approach of Bai et al. [1]. Using the concept of a matrix function [3], the definition of kernel function ψ can be extended to any diagonalizable matrix with positive eigenvalues. In particular, for a given eigen-decomposition

$$V = Q_V^{-1}\ diag\left(\lambda_1(V), \lambda_2(V), ..., \lambda_n(V)\right) Q_V,$$

of V with a nonsingular matrix Q_V, the matrix function $\psi(V)$ is defined by

$$\psi(V) = Q_V^{-1}\ diag\left(\psi(\lambda_1(V)), \psi(\lambda_2(V)), ..., \psi(\lambda_n(V))\right) Q_V. \tag{23}$$

In this section, we compute a proper step size α and the decrease of the proximity function during an inner iteration and give the complexity results of the algorithm.

4.1 Determining a default step size

Taking a step size α, we have new iterates

$$X^+ = X + \alpha \triangle X,\ y^+ = y + \alpha \triangle y \text{ and } S^+ = S + \alpha \triangle S.$$

Let

$$X^+ = X\left(I + \alpha \frac{\triangle X}{X}\right) = X\left(I + \alpha \frac{D_X}{V}\right) = \frac{X}{V}\left(V + \alpha D_X\right),$$

$$S^+ = S\left(I + \alpha \frac{\triangle S}{S}\right) = S\left(I + \alpha \frac{D_S}{V}\right) = \frac{S}{V}\left(V + \alpha D_S\right).$$

So, we have

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

Since the proximity after one step is defined by

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

By (i) in Lemma 2, we have

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

Define, for $\alpha > 0$

$$f(\alpha) = \Psi(V^+) - \Psi(V).$$

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

$$f_1(\alpha) = \frac{1}{2}\left(\Psi\left(V + \alpha D_X\right) + \Psi(V + \alpha D)\right) - \Psi(V). \tag{24}$$

Obviously,

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

Now, we deal with another concepts relevant to matrix functions in matrix theory [5].

Definition 1. [5] A matrix $X(t)$ is said to be a matrix of functions if each entry of $X(t)$ is a function of t, that is, $X(t) = \left[X_{ij}(t)\right].$

The concepts of continuity, differentiability and integrability naturally extended to matrix-valued functions of a scalar by interpreting them componentwise. Thus we can say that

$$\frac{d}{dt}X(t) = \frac{d}{dt}\left[X_{ij}(t)\right] = X^{'}(t).$$

Suppose that the matrix-valued functions $H(t)$ and $G(t)$ are differentiable with respect to t. Then we have

$$\frac{d}{dt}\left(Tr(G(t)\right) = Tr\left(\frac{d}{dt}G(t)\right) = Tr\left(G^{'}(t)\right),$$

$$\frac{d}{dt}\left((G(t)H(t)\right) = G^{'}(t)H(t) + G(t)H^{'}(t).$$

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

$$\triangle\psi(t_1, t_2) = \frac{\psi(t_1) - \psi(t_2)}{t_1 - t_2},\ \forall t_1 \neq t_2 \in \Re^*.$$

If $t_1 = t_2$, we simply write $\triangle\psi(t_1, t_2) = \psi^{'}(t)$.

Let us define that Q_α is orthogonal matrix such that

$$V + \alpha D_X = Q_\alpha^T\ diag\left(\lambda_1(V + \alpha D_X), \lambda_2(V + \alpha D_X), ..., \lambda_n(V + \alpha D_X)\right) Q_\alpha,$$

and let us denote by D_i the diagonal matrix that has one in its (i, i) position and all other components of D_i equal to zero. It follows from [5,12,13] that

$$\frac{d}{d\alpha}\left(\psi(V+\alpha D_X)\right) = Q_\alpha^T \left(\sum_{j,k=1}^{n} \triangle\psi\left(\lambda_j(V+\alpha D_X), \lambda_k\left(V+\alpha D_X\right)\right) \right.$$
$$\left. D_j\left(Q_\alpha (V+\alpha D_X)' Q_\alpha^T\right) D_k \right) Q_\alpha. \tag{25}$$

Now by the choice of D_i, it holds $Tr\left(D_j\left(Q_\alpha (V+\alpha D_X)' Q_\alpha^T\right) D_k\right) = 0$, for $j \neq k$. Thus it follows that

$$\begin{aligned}
&\frac{d}{d\alpha}\left(\psi(V+\alpha D_X)\right)\\
&\quad = \sum_{i=1}^{n} \psi'\left(\lambda_i(V+\alpha D_X)\right) D_i \left(Q_\alpha (V+\alpha D_X)' Q_\alpha^T\right) D_i\\
&\quad = Tr\left(Q_\alpha^T \left(\sum_{i=1}^{n} D_i \psi'\left(\lambda_i(V+\alpha D_X)\right) D_i\right) Q_\alpha (V+\alpha D_X)'\right)\\
&\quad = Tr\left(\psi'(V+\alpha D_X) D_X\right),
\end{aligned}$$

and

$$\frac{d}{d\alpha}\left(\psi(V+\alpha D_S)\right) = Q_\alpha^T \left(\sum_{j,k=1}^{n} \triangle\psi(\lambda_j(V+\alpha D_S), \lambda_k(V+\alpha D_S)) \right.$$
$$\left. D_j\left(Q_\alpha (V+\alpha D_S)' Q_\alpha^T\right) D_k \right) Q_\alpha. \tag{26}$$

Now by the choice of D_i, it holds $Tr\left(D_j\left(Q_\alpha (V+\alpha D_S)' Q_\alpha^T\right) D_k\right) = 0$, for $j \neq k$. Thus it follows that

$$\begin{aligned}
&\frac{d}{d\alpha}\left(\psi(V+\alpha D_S)\right)\\
&\quad = Tr\left(\sum_{i=1}^{n} \psi'(\lambda_i(V+\alpha D_S)) D_i \left(Q_\alpha (V+\alpha D_S)' Q_\alpha^T\right) D_i\right)\\
&\quad = Tr\left(Q_\alpha^T \left(\sum_{i=1}^{n} D_i \psi'(\lambda_i(V+\alpha D_S)) D_i\right) Q_\alpha (V+\alpha D_S)'\right)\\
&\quad = Tr\left(\psi'(V+\alpha D_S) D_S\right).
\end{aligned}$$

Now, we can write

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

This gives

$$f_1'(0) = -2\delta^2.$$

Furthermore,

$$\begin{aligned}
&\tfrac{d^2}{d\alpha^2}\left(Tr\left(\psi\left(V+\alpha D_X\right)\right)\right)\\
&\quad = Tr\left(\tfrac{d}{d\alpha}\left(\psi'\left(V+\alpha D_X\right)D_X\right)\right)\\
&\quad = Tr\left(Q_\alpha^T\left(\sum_{j,k=1}^{n}\triangle\psi'(\lambda_j(V+\alpha D_X),\lambda_k(V+\alpha D_X))D_j\right.\right.\\
&\qquad\qquad\left.\left.\left(Q_\alpha(V+\alpha D_X)'Q_\alpha^T\right)D_k\right)Q_\alpha D_X\right)\\
&\quad = Tr\left(\sum_{j,k=1}^{n}\triangle\psi'(\lambda_j(V+\alpha D_X),\lambda_k(V+\alpha D_X))D_j(Q_\alpha D_X Q_\alpha^T)\right.\\
&\qquad\qquad\left.D_k(Q_\alpha D_X Q_\alpha^T)\right)\\
&\quad = Tr\left(\sum_{j,k=1}^{n}\triangle\psi'(\lambda_j(V+\alpha D_X),\lambda_k(V+\alpha D_X))(Q_\alpha D_X Q_\alpha^T)_{jk}^2\right)\\
&\quad \le \max\left\{\left|\triangle\psi'(\lambda_j(V+\alpha D_X),\lambda_k(V+\alpha D_X))\right|,\ j,k=1,...,n\right\}\|D_X\|^2.
\end{aligned}$$

Also, we have

$$\begin{aligned}
&\tfrac{d^2}{d\alpha^2}\left(Tr\left(\psi\left(V+\alpha D_S\right)\right)\right)\\
&\quad = Tr\left(\tfrac{d}{d\alpha}\left(\psi'\left(V+\alpha D_S\right)D_S\right)\right)\\
&\quad = Tr\left(Q_\alpha^T\left(\sum_{j,k=1}^{n}\triangle\psi'(\lambda_j(V+\alpha D_S),\lambda_k(V+\alpha D_S))D_j\right.\right.\\
&\qquad\qquad\left.\left.\left(Q_\alpha(V+\alpha D_S)'Q_\alpha^T\right)D_k\right)Q_\alpha D_S\right)\\
&\quad = Tr\left(\sum_{j,k=1}^{n}\triangle\psi'(\lambda_j(V+\alpha D_S),\lambda_k(V+\alpha D_S))D_j(Q_\alpha D_S Q_\alpha^T)\right.\\
&\qquad\qquad\left.D_k(Q_\alpha D_S Q_\alpha^T)\right)\\
&\quad = Tr\left(\sum_{j,k=1}^{n}\triangle\psi'(\lambda_j(V+\alpha D_S),\lambda_k(V+\alpha D_S))(Q_\alpha D_S Q_\alpha^T)_{jk}^2\right)\\
&\quad \le \max\left\{\left|\triangle\psi'(\lambda_j(V+\alpha D_S),\lambda_k(V+\alpha D_S))\right|,\ j,k=1,...,n\right\}\|D_S\|^2.
\end{aligned}$$

We denote by

$$\omega_1=\max\left\{\left|\triangle\psi'(\lambda_j(V+\alpha D_X),\lambda_k(V+\alpha D_X))\right|,\ j,k=1,...,n\right\},$$

and

$$\omega_2 = \max\left\{\left|\triangle\psi^{'}(\lambda_j(V+\alpha D_S),\lambda_k(V+\alpha D_S))\right|,\ j,k=1,...,n\right\}.$$

Therefore,

$$\begin{aligned} f^{''}(\alpha) &= \tfrac{1}{2}\tfrac{d^2}{d\alpha^2}Tr\left(\psi\left(V+\alpha D_X\right)+\psi\left(V+\alpha D_S\right)\right) \\ &\leq \tfrac{1}{2}\left(\omega_1\left\|D_X\right\|^2+\omega_2\left\|D_S\right\|^2\right). \end{aligned} \tag{27}$$

Lemma 7. *Let $\delta(V)$ be as defined in (15). Then we have*

$$\delta(V) \geq \sqrt{(m+1)\Psi(V)}.$$

Proof. Using (16), we have

$$\begin{aligned} \Psi(V) &= Tr(\psi(V)) = \sum_{i=1}^{n}\psi(\lambda_i(V)) \leq \sum_{i=1}^{n}\frac{1}{4(m+1)}\psi^{'}(\lambda_i(V))^2 \\ &= \frac{1}{4(m+1)}\left\|\nabla\Psi(V)\right\|^2 = \frac{1}{m+1}\delta(V)^2. \end{aligned}$$

This gives

$$\delta(V) \geq \sqrt{(m+1)\Psi(V)}.$$

□

Throughout the paper, we assume that $\tau \geq 1$. Using lemma 7 and the assumption that $\Psi(V) \geq \tau$, we have $\delta(V) \geq \sqrt{(m+1)}$. We have the following lemma.

Lemma 8. *Let $f_1(\alpha)$ be as defined in (24), $\delta(V)$ be as defined in (15) and ψ the kernel function be as defined in (9). Then we have*

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

Proof. We have

$$f_1^{''}(\alpha) \leq 2\max\left\{\omega_1,\omega_2\right\}\delta^2.$$

It suffices to prove the following inequality:

$$\max\left\{\omega_1,\omega_2\right\} \leq \psi^{''}(\lambda_{\min}(V)-2\alpha\delta).$$

We can choose $j^*, k^* \in \{1,2,...,n\}$ such that

$$\omega_1 = \left|\frac{\psi^{'}(\lambda_{j^*}(V+\alpha D_X)-\psi^{'}(\lambda_{k^*}(V+\alpha D_X)}{\lambda_{j^*}(V+\alpha D_X)-\lambda_{k^*}(V+\alpha D_X)}\right|.$$

Using the Mean value theorem, there exists

$$\eta \in \Big[\min\left(\lambda_{j^*}(V+\alpha D_X), \lambda_{k^*}(V+\alpha D_X)\right), \\ , \max\left(\lambda_{j^*}(V+\alpha D_X), \lambda_{k^*}(V+\alpha D_X)\right) \Big],$$

such that

$$\psi^{''}(\eta) = \triangle\psi^{'}(\lambda_{j^*}(V+\alpha D_X), \lambda_{k^*}(V+\alpha D_X)).$$

Since D_X are symmetric matrices, from the definition of δ and Frobenius norm, we have

$$\begin{aligned} \eta &\geq \min\{\lambda_{j^*}(V+\alpha D_X), \lambda_{k^*}(V+\alpha D_X)\} \\ &\geq \lambda_{\min}(V) - 2\alpha\delta, \end{aligned}$$

because, $\psi^{''}$ is monotonically decreasing, then we obtain

$$\omega_1 \leq \psi^{''}(\lambda_{\min}(V) - 2\alpha\delta).$$

By the same method to get ω_1, we obtain

$$\omega_2 = \left| \frac{\psi^{'}(\lambda_{j^*}(V+\alpha D_S) - \psi^{'}(\lambda_{k^*}(V+\alpha D_S)}{\lambda_{j^*}(V+\alpha D_S) - \lambda_{k^*}(V+\alpha D_S)} \right|.$$

Using the Mean value theorem, there exists

$$\eta \in \Big[\min\left(\lambda_{j^*}(V+\alpha D_S), \lambda_{k^*}(V+\alpha D_S)\right), \\ \max\left(\lambda_{j^*}(V+\alpha D_S), \lambda_{k^*}(V+\alpha D_S)\right) \Big],$$

such that

$$\psi^{''}(\eta) = \triangle\psi^{'}(\lambda_{j^*}(V+\alpha D_S), \lambda_{k^*}(V+\alpha D_S)).$$

Since D_S are symmetric matrices, from the definition of δ and Frobenius norm, we have

$$\begin{aligned} \eta &\geq \min\{\lambda_{j^*}(V+\alpha D_S), \lambda_{k^*}(V+\alpha D_S)\} \\ &\geq \lambda_{\min}(V) - 2\alpha\delta, \end{aligned}$$

because, $\psi^{''}$ is monotonically decreasing, then we obtain

$$\omega_2 \leq \psi^{''}(\lambda_{\min}(V) - 2\alpha\delta).$$

Therefore,

$$\max\{\omega_1, \omega_2\} \leq \psi^{''}(\lambda_{\min}(V) - 2\alpha\delta).$$

This gives

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

$\square$

Since $f_1(0) = 0$ and $f_1^{'}(0\ ) = -2\delta(V)^2$, we have

$$\begin{aligned} f(\alpha) &\leq f_1(\alpha) := f_1(0) + f_1^{'}(0)\alpha + \int_0^{\alpha}\int_0^{\xi} f_1^{''}(\zeta) d\zeta d\xi \\ &\leq f_2(\alpha) := f_1(0) + f_1^{'}(0)\alpha + 2\delta^2 \int_0^{\alpha}\int_0^{\xi} \psi^{''}(\lambda_{\min}(V) - 2\zeta\delta) d\zeta d\xi. \end{aligned}$$

Note that $f_2(0) = 0$. Furthermore, since

$$f_2^{'}(\alpha) = -2\delta^2 + \delta\left(\psi^{'}(\lambda_{\min}(V)) - \psi^{'}(\lambda_{\min}(V) - 2\alpha\delta)\right),$$

then, we have $f_2^{'}(0) = -2\delta^2$, which is the same value of the $f_1^{'}(0)$ and $f_2^{''}(\alpha) = 2\delta^2\psi^{''}(\lambda_{\min}(V) - 2\alpha\delta)$, which is increasing on $\alpha \in \left[0, \frac{\lambda_{\min}(V)}{2\delta}\right]$. So we can rewrite $f_2(\alpha)$ as follows:

$$f_2(\alpha) = f_2(0) + f_2^{'}(0)\alpha + 2\delta^2 \int_0^{\alpha}\int_0^{\xi} \psi^{''}(\lambda_{\min}(V) - 2\zeta\delta) d\zeta d\xi.$$

Now, using $f_1^{'}(0) = f_2^{'}(0)$ and $f_1^{''}(\alpha) \leq f_2^{''}(\alpha)$, we can easily check that

$$f_1^{'}(\alpha) = f_1^{'}(0) + \int_0^{\alpha} f^{''}(\xi) d\xi \leq f_2^{'}(\alpha).$$

This gives that

$$f_1^{'}(\alpha) \leq 0, \text{ if } f_2^{'}(\alpha) \leq 0.$$

For each $\mu > 0$, we compute a feasible iterate such that the proximity measure is decreasing. We want to compute the step size α which satisfies that $f_2^{'}(\alpha) \leq 0$ holds with α as large as possible. Since $f_2^{''}(\alpha) > 0$, that is, $f_2^{'}(\alpha)$ is monotonically increasing at α, the largest possible value at α satisfying $f_2^{'}(\alpha) \leq 0$ occurs when $f_2^{'}(\alpha) = 0$, that is,

$$\psi^{'}(\lambda_{\min}(V)) - \psi^{'}(\lambda_{\min}(V) - 2\alpha\delta) = 2\delta. \tag{28}$$

Since $\psi^{''}(t)$ is monotonically decreasing, the derivative of the left-hand side in (28) with respect to $\lambda_{\min}(V)$ is

$$\psi^{''}(\lambda_{\min}(V)) - \psi^{''}(\lambda_{\min}(V) - 2\alpha\delta) \leq 0.$$

So, the left-hand side in (28) is decreasing at $\lambda_{\min}(V)$. This implies that if $\lambda_{\min}(V)$ gets smaller, then α gets smaller with fixed δ. Note that

$$\delta = \sqrt{\sum_{i=1}^{n} (\psi'(\lambda_i(V)))^2} \geq \left|\psi^{'}(\lambda_{\min}(V))\right| \geq -\psi^{'}(\lambda_{\min}(V)),$$

and the equality is true if and only if $\lambda_{\min}(V)$ is only the cordinate in $(\lambda_1(V), \lambda_2(V), ..., \lambda_n(V))$ which is different from 1 and $\lambda_{\min}(V) < 1$, that is, $\psi^{'}(\lambda_{\min}(V)) < 0$. Hence, the worst situation for the largest step size α occurs when $\lambda_{\min}(V)$ satisfies

$$-\psi^{'}(\lambda_{\min}(V)) = \delta. \tag{29}$$

In that case, the largest satisfying (28) is minimal. For our purpose, we need to deal with the worst case and so we assume that (29) holds. This implies

$$\lambda_{\min}(V) = \rho(\delta). \tag{30}$$

By using (28) and (29) we immediately obtain

$$-\psi^{'}(\lambda_{\min}(V) - 2\alpha\delta) = 4\delta.$$

By the definition of ρ and using (30), the largest step size of the worst case is given as follows:

$$\alpha^* = \frac{\rho(\delta) - \rho(2\delta)}{2\delta}. \tag{31}$$

Lemma 9. *Let the definition of ρ and α^* be as defined in (31), then we have*

$$\alpha^* \geq \frac{1}{(m+1)(m+2)^{\frac{m+2}{m+1}}}.$$

Proof. Using lemma 4.4 in [1], the definition of $\psi''(t)$ and (19), we have

$$\begin{aligned}
\alpha^* &\geq \frac{1}{\psi''(\rho(2\delta))} = \frac{1}{2(m+1) + \frac{m(m+1)}{\rho(2\delta)^{m+2}}} \\
&\geq \frac{1}{2(m+1) + m(m+1)(\frac{4\delta+m}{m})^{\frac{m+2}{m+1}}} \geq \frac{1}{2(m+1)\delta + 3m(m+2)\delta^{\frac{m+2}{m+1}}} \\
&\geq \frac{1}{3(m+1)(m+2)\delta^{\frac{m+2}{m+1}}},
\end{aligned}$$

which completes the proof. □

For using $\overline{\alpha}$ as the default step size in the algorithm, define the $\overline{\alpha}$ as follows

$$\overline{\alpha} = \frac{1}{3(m+1)(m+2)\delta^{\frac{m+2}{m+1}}}. \tag{32}$$

4.2 Decrease of the proximity function during an inner iteration

Now, we show that our proximity function Ψ with our default step size $\overline{\alpha}$ is decreasing. It can be easily established by using the following result:

Lemma 10. *[12] 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*

$$h(t) = \frac{th^{'}(0)}{2}.$$

Let the univariate function h be such that

$$h(0) = f_1(0) = 0,\ h^{'}(0) = f_1^{'}(0) = -2\delta^2,\ h^{''}(\alpha) = 2\delta^2\psi^{''}(\lambda_{\min}(V) - 2\alpha\delta).$$

Since $f_2(\alpha)$ satisfies the condition of the above lemma,

$$f(\alpha) \leq f_1(\alpha) \leq f_2(\alpha) \leq \frac{f_2^{'}(0)}{2}\alpha,\ \text{for all } 0 \leq \alpha \leq \alpha^*.$$

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

Theorem 1. *Let $\overline{\alpha}$ be a step size as defined in (32) and $\delta = \Psi(V) \geq \tau = 1$. Then we have*

$$f(\overline{\alpha}) \leq -\frac{(m+1)^{\frac{-m-2}{2(m+1)}}}{3(m+2)}\Psi(V)^{\frac{m}{2(m+1)}}.$$

Proof. For all $\overline{\alpha} \leq \alpha^*$, we have

$$\begin{aligned} f(\overline{\alpha}) &\leq -\overline{\alpha}\delta^2 = -\frac{1}{3(m+1)(m+2)\delta^{\frac{m+2}{m+1}}}\delta^2 \\ &= -\frac{1}{3(m+1)(m+2)}\delta^{\frac{m}{m+1}} \\ &\leq -\frac{1}{3(m+1)(m+2)}\left(\sqrt{(m+1)\Psi(V)}\right)^{\frac{m}{m+1}} \\ &\leq -\frac{(m+1)^{\frac{m}{2(m+1)}}}{3(m+1)(m+2)}\Psi(V)^{\frac{m}{2(m+1)}} \leq \frac{(m+1)^{\frac{-m-2}{2(m+1)}}}{3(m+2)}\Psi(V)^{\frac{m}{2(m+1)}}. \end{aligned}$$

□

4.3 Iteration bound

We need to count how many inner iterations are required to return to the situation where $\Psi(V) \leq \tau$ after a $\mu-$update. We denote the value of $\Psi(V)$ after $\mu-$update as Ψ_0 the subsequent values in the same outer iteration are denoted as $\Psi_k,\ k = 1, 2, ...$ If K denotes the total number of inner iterations in the outer iteration, then we have

$$\Psi_0 \leq L = O(n, \theta, \tau),\ \Psi_{K-1} > \tau\,,\ 0 \leq \Psi_K \leq \tau.$$

and according to (20),

$$\Psi_{k+1} \leq \Psi_k - \frac{(m+1)^{\frac{-m-2}{2(m+1)}}}{3(m+2)}\Psi_k^{\frac{m}{2(m+1)}}.$$

At this stage we invoke the following lemma from [12]

Lemma 11. *[12] Let $t_0, t_1, ..., t_k$ be a sequence of positive numbers such that*

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

where $\beta > 0,\ 0 < \nu \leq 1$. Then

$$K \leq \frac{t_0^\nu}{\beta\nu}.$$

Letting

$$t_k = \Psi_k,\ \beta = \frac{(m+1)^{\frac{-m-2}{2(m+1)}}}{3(m+2)} \text{ and } \nu = \frac{m+2}{2(m+1)},$$

we can get the following lemma.

Lemma 12. *Let K be the total number of inner iterations in the outer iteration. Then we have*

$$K \leq 6(m+1)^{\frac{3m+4}{2(m+1)}} \Psi_0^{\frac{m+2}{2(m+1)}}.$$

Proof. Using lemma 11, we have

$$K \leq \frac{\Psi_0^{\nu}}{\beta\nu} = 6(m+1)^{\frac{3m+4}{2(m+1)}} \Psi_0^{\frac{m+2}{2(m+1)}}.$$

□

Now we estimate the total number of iterations of our algorithm.

Theorem 2. *If $\tau \geq 1$, the total number of iterations is not more than*

$$6(m+1)^{\frac{3m+4}{2(m+1)}} \Psi_0^{\frac{m+2}{2(m+1)}} \frac{1}{\theta} \log \frac{n\mu^0}{\varepsilon}.$$

Proof. In the algorithm, $n\mu \leq \varepsilon,\ \mu^k = (1-\theta)^k \mu^0$ and $\mu^0 = \frac{Tr(X^0S^0)}{n}$. By simple computation, we have

$$K \leq \frac{1}{\theta} \log \frac{n\mu^0}{\varepsilon}.$$

Therefore, the number of outer iterations is bounded above by $\frac{1}{\theta} \log \frac{n\mu^0}{\varepsilon}$. Multiplying the number of outer iterations by the number of inner iterations, we get an upper bound for the total number of iterations, namely,

$$6(m+1)^{\frac{3m+4}{2(m+1)}} \Psi_0^{\frac{m+2}{2(m+1)}} \frac{1}{\theta} \log \frac{n\mu^0}{\varepsilon}.$$

□

5 Conclusion

We propose a new barrier function and primal–dual interior point algorithms for SDO problems and analyze the iteration complexity of the algorithm based on the kernel function. We have $O(m^{\frac{3m+1}{2m}} n^{\frac{m+1}{2m}} \log \frac{Tr(X^0S^0)}{\varepsilon})$ for large-update methods and $O(m^{\frac{3m+1}{2m}} \sqrt{n} \log \frac{Tr(X^0S^0)}{\varepsilon})$ for small-update methods which are the best known iteration bounds for such methods. Future research might focus on the extension to symmetric cone optimization. Finally, the numerical test is an interesting work for investigating the behavior of the algorithms so as to be compared with other existing approaches.

Acknowledgements

The authors thank the referees for their careful reading and their precious comments. Their help is much appreciated.

References

[1] 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 J. Optim. **15** (2004) 101–128.

[2] Y.Q. Bai, M. El Ghami and C. Roos, *A new efficient large-update primal–dual interior–point method based a finite barrier*, SIAM J. Optim. **13** (2002) 766–782.

[3] R. Bellman, *Introduction to Matrix Analysis, in: Classics in Applied Mathematics*, SIAM, Philadelphia, 1995.

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

[5] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

[6] M. Kojima, N. Megiddo and S. Mizuno, *A primal–dual infeasible-interior-point algorithm for linear programming*, Math. Program. **61** (1993) 263–280.

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

[8] I.J. Lustig, *Feasible issues in a primal–dual interior point method for linear programming*, Math. Program. **49** (1990-1991) 145–162.

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

[10] H. Mansouri and C. Roos, *Simplified $O(nL)$ infeasible interior-point algorithm for linear optimization using full Newton steps*, Optim. Methods Softw. **22** (2007) 519–530.

[11] Y.E. Nesterov and A.S. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Stud. Appl. Math. SIAM, Philadelphia, 1994.

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

[13] J. Peng, C. Roos and T. Terlaky, *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*, Princeton University Press, Princeton, NJ, 2002.

[14] C. Roos, *A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization*, SIAM J. Optim. **16** (2006) 1110–1136.

[15] J. Sturm, *Theory and algorithms of semidefinite programming, in: H. Frenk, C. Roos, T. Terlaky, S. Zhang (Eds.), High Performance Optimization*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[16] K. Tanabe, *Centered Newton method for linear programming: interior and exterior point method (Japanese), in: New Methods for Linear Programming 3, ed. K. Tone*, Inst. Statist. Math. Tokyo, Japan, 1990.

[17] L. Tuncel, *Potential reduction and primal dual methods, in: H. Wolkowicz, R. Saigal, L. Vandenberghe (Eds.), Theory, Algorithms and Applications, in: Handbook of Semidefinite Programming*, Kluwer Academic Publishers, Boston, MA, 2000.

[18] Y. Zhang, *On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problem*, SIAM J. Optim. **4** (1994) 208–227.