Acta Mathematica Sinica, English Series
Mar., 2015, Vol. 31, No. 3, pp. 543–556
Published online: February 15, 2015
DOI: 10.1007/s10114-015-1314-4
Http://www.ActaMath.com

Acta Mathematica Sinica, English Series

Complexity Analysis of an Interior Point Algorithm for the Semidefinite Optimization Based on a Kernel Function with a Double Barrier Term

Mohamed ACHACHE
Laboratoire de Mathématiques Fondamentales et Numériques,
Faculté des Sciences, Université Ferhat Abbas Sétif 1*, Algérie*
E-mail: *achache_m@yahoo.fr*

Abstract In this paper, we establish the polynomial complexity of a primal-dual path-following interior point algorithm for solving semidefinite optimization (SDO) problems. The proposed algorithm is based on a new kernel function which differs from the existing kernel functions in which it has a double barrier term. With this function we define a new search direction and also a new proximity function for analyzing its complexity. We show that if $q_1 > q_2 > 1$, the algorithm has $O((q_1+1)\,n^{\frac{q_1+1}{2(q_1-q_2)}}\log\frac{n}{\epsilon})$ and $O((q_1+1)^{\frac{3q_1-2q_2+1}{2(q_1-q_2)}}\sqrt{n}\log\frac{n}{\epsilon})$ complexity results for large- and small-update methods, respectively.

Keywords Semidefinite optimization, kernel functions, primal-dual interior point methods, large and small-update algorithms, complexity of algorithms

MR(2010) Subject Classification 65K05, 90C22, 90C25, 90C51

1 Introduction

We consider the semidefinite optimization problem (SDO) in its primal format:

$$(P)\quad \min_X C \bullet X \ \text{ s.t. } A_i \bullet X = b_i, \quad 1 \le i \le m, X \succeq 0$$

and its dual problem

$$(D)\quad \max_{(y,Z)} b^T y \ \text{ s.t. } \sum_{i=1}^m y_i A_i + Z = C, \quad Z \succeq 0,$$

where $C, A_i \in \mathbf{S}^n$, $b \in \mathbb{R}^m$ and (X, y, Z) are unknown variables. Here $\mathbf{S}^n$ denotes the space of $n \times n$ real symmetric matrices. In addition, $X \succeq 0$ indicates that X is a symmetric positive semidefinite matrix and the expression $A \bullet B := \sum_{j=1}^n \sum_{i=1}^n A_{ij}B_{ij}$ denotes the inner-product between two symmetric matrices. Without loss of generality we assume that the matrices A_i are linearly independent.

The semidefinite optimization is an exciting problem in mathematical programming that has many applications in both engineering and scientific fields. For more details concerning this subject the reader consults [2, 15, 16]. Recently, generalized barrier methods play an important role to designing new and efficient primal-dual path-following interior point (IP) algorithms to solve the linear optimization (LO) and SDO [3–10]. These methods are based on the so-called

Received June 14, 2011, accepted June 3, 2014

kernel functions for determining new search directions and new proximity functions for analyzing the complexity of these algorithms. With this new paradigm the polynomial complexity of large-update primal-dual IP algorithms is improved in contrast with the classical complexity given by the logarithmic barrier functions (see [3–14]).

In this paper, we establish the polynomial complexity of a primal-dual path-following IP algorithm for SDO by introducing the following new kernel function

$$\psi(t) = t^2 - 1 + \frac{t^{1-q_1} - 1}{q_1 - 1} + \frac{t^{1-q_2} - 1}{q_2 - 1}, \quad q_i > 1,\, i = 1, 2.$$

In contrast with the existing kernel functions, $\psi(t)$ has a double barrier term defined by two parameters q_1 and q_2. With this function a new search direction and a proximity are determined for analyzing the proposed algorithm. We derive complexity results for this algorithm with large- and small-update methods. We note that if $q_1 = q_2$, then $\psi(t)$ reduces to a well-known kernel function with a single barrier term introduced by Bai et al. [3] and EL Ghami [8] for LO and SDO and extended by other authors for different problems in mathematical programming.

This paper is organized as follows. In Section 2, the central path and the classical Nesterov–Todd direction are given. In Section 3, the new kernel-function-based Nesterov–Todd direction and the generic primal-dual algorithm are presented. In Section 4, a new eligible kernel function and its growth properties for SDO are studied. In Section 5, the iteration bounds for the algorithm are computed. Finally, a conclusion ends Section 6.

The following notations are used throughout the paper. $\mathbf{S}^n_+$ and $\mathbf{S}^n_{++}$ denote the cone of symmetric positive semidefinite and symmetric definite positive matrices, respectively. Furthermore, $X \succeq 0\,(X \succ 0)$ means that $X \in \mathbf{S}^n_+$ $(X \in \mathbf{S}^n_{++})$. For any $X \succ 0$, $\lambda_i(X), 1 \le i \le n$, denote its eigenvalues. The trace of an $n \times n$ matrix X is denoted by $\mathrm{Tr}(X) = \sum_{i=1}^n X_{ii}$. $\|\cdot\|$ denotes the Frobenius norm and I is the identity matrix. For two real valued functions $f(x)$, $g(x) : \mathbb{R}^n_+ \to \mathbb{R}^n_{++}$, $f(x) = O(g(x))$ if $f(x) \le kg(x)$ for some positive constant k and $f(x) = \Theta g(x)$ if $k_1 g(x) \le f(x) \le k_2 g(x)$ for some positive constants k_1 and k_2.

2 The Central Path and the Classical Nesterov–Todd Search Direction for SDO

In this section, we recall the notion of the central path with its properties and we derive the classical Nesterov–Todd search direction for SDO.

Throughout the paper, we assume that (P) and (D) satisfy the interior point condition, i.e., there exists (X^0, y^0, Z^0) such that

$$A_i \bullet X^0 = b_i, \quad 1 \le i \le m, \quad \sum_{i=1}^m y_i^0 A_i + Z^0 = C, \quad X^0 \succ 0,\, Z^0 \succ 0.$$

It is well known that finding a solution of (P) and (D) is equivalent to find a solution of the following system:

$$\begin{cases} A_i \bullet X = b_i, & 1 \le i \le m,\, X \succeq 0, \\ \sum_{i=1}^m y_i A_i + Z = C, & Z \succeq 0, \\ XZ = 0. \end{cases} \tag{2.1}$$

The basic idea of primal-dual interior point methods is to replace the third equation in (2.1) by the parameterized equation $XZ = \mu I$ where $\mu > 0$. Then one considers the following system:

$$\begin{cases} A_i \bullet X = b_i, & 1 \le i \le m,\ X \succeq 0, \\ \sum_{i=1}^{m} y_i A_i + Z = C, & Z \succeq 0, \\ XZ = \mu I. \end{cases} \quad (2.2)$$

It is also well known that the system (2.2) has a unique solution for each $\mu > 0$. It is denoted by $(X(\mu), y(\mu), Z(\mu))$ and we call it the μ-center of both problems (P) and (D). The set of μ-centers defines a homotopy which is called the central path of (P) and (D). If $\mu \to 0$, then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D).

Now to obtain the search direction for SDO, we apply Newton's method for the system (2.2) for a given strictly feasible primal-dual point $(X,\ y,\ Z)$, which yields the following linear system of equations:

$$\begin{cases} A_i \bullet \Delta_X = 0, \quad 1 \le i \le m, \\ \sum_{i=1}^{m} \Delta_{y_i} A_i + \Delta Z = 0, \\ X\Delta_Z + \Delta_X Z = \mu I - XZ. \end{cases} \quad (2.3)$$

Since A_i are linearly independent and $X \succ 0$, $Z \succ 0$, the system (2.3) has a unique solution $(\Delta_X, \Delta_y, \Delta_Z)$. Note that Δ_Z is symmetric due to the second equation in (2.3) but Δ_X may be not symmetric. Therefore the proposal made by researchers is to symmetrize the third equation in (2.3) by using a symmetric nonsingular matrix P and by replacing (2.3) by the system:

$$\begin{cases} A_i \bullet \Delta_X = 0, \quad 1 \le i \le m, \\ \sum_{i=1}^{m} \Delta_{y_i} A_i + \Delta_Z = 0, \\ \Delta_X + P\Delta_Z P^T = \mu Z^{-1} - X. \end{cases} \quad (2.4)$$

In [15], Todd studied several symmetrization schema. Among them, we consider the Nesterov–Todd (NT) symmetrization scheme where P is defined as

$$P = X^{\frac{1}{2}}(X^{\frac{1}{2}}ZX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} = Z^{-\frac{1}{2}}(Z^{\frac{1}{2}}XZ^{\frac{1}{2}})^{\frac{1}{2}}Z^{-\frac{1}{2}}.$$

Let $D = P^{\frac{1}{2}}$ where $P^{\frac{1}{2}}$ denotes the symmetric square root of P. The matrix D is used to scale both matrices X and Z to the same matrix V defined by

$$V := \frac{1}{\sqrt{\mu}} D^{-1}XD^{-1} = \frac{1}{\sqrt{\mu}} DZD = \frac{1}{\sqrt{\mu}} (D^{-1}XZD)^{\frac{1}{2}}. \quad (2.5)$$

Note that both matrices D and V are symmetric and positive definite. So by using (2.5), the linear system (2.4) becomes

$$\begin{cases} \bar{A}_i \bullet D_X = 0, \quad 1 \le i \le m, \\ \sum_{i=1}^{m} \Delta_{y_i} \bar{A}_i + D_Z = 0, \\ D_X + D_Z = V^{-1} - V, \end{cases} \quad (2.6)$$

with

$$\bar{A}_i = DA_iD, \quad 1 \le i \le m, \quad D_X := \frac{1}{\mu}D^{-1}\Delta_X D^{-1} \quad \text{and} \quad D_Z := \frac{1}{\mu}D\Delta_Z D. \tag{2.7}$$

The system (2.6) determines a uniquely symmetric NT direction with the matrices D_X and D_Z be orthogonal, since $D_X \bullet D_Z = 0$, and it is evident to see

$$\mathrm{Tr}(D_X D_Z) = \mathrm{Tr}(D_Z D_X) = 0.$$

The above Nesterov–Todd direction leads to the classical primal dual IP algorithms for SDO.

3 The New Search Direction and the Generic Primal-dual IP Algorithm for SDO

In this section, we recall the definition of a matrix function and we derive the new kernel-function-based Nesterov–Todd direction and then we describe our generic primal-dual (IP) algorithm to SDO.

Definition 3.1 ([8, Definition 3.2.1]) *Let X be a symmetric matrix, and let*

$$X = Q_X^{-1}\mathrm{diag}(\lambda_1(X), \lambda_2(X), \dots, \lambda_n(X))Q_X$$

be an eigenvalue decomposition of X, where $\lambda_i(X), 1 \le i \le n$, denote the eigenvalues of X, and Q_X is orthogonal. If $\psi(t)$ is any univariate function whose domain contains $\{\lambda_i(X); 1 \le i \le n\}$, then the matrix function $\psi(X)$ is defined by

$$\psi(X) = Q_X^{-1}\mathrm{diag}(\psi(\lambda_1(X)), \psi(\lambda_2(X)), \dots, \psi(\lambda_n(X)))Q_X.$$

Definition 3.1 is called the spectral decomposition theorem of symmetric matrices and its importance enables us to extend the definition of any function $\psi : \mathbb{R} \to \mathbb{R}$ to a function from $\mathbf{S}^n$ to $\mathbf{S}^n$.

Now in a similar way to [12], the barrier function $\Psi(X)$ is defined as follows:

$$\Psi(X) := \sum_{i=1}^{n} \psi(\lambda_i(X)) = \mathrm{Tr}(\psi(X)). \tag{3.1}$$

When we use the function $\psi(\cdot)$ and its first three derivatives $\psi'(\cdot), \psi''(\cdot)$ and $\psi'''(\cdot)$ without any specification, it denotes a matrix function if the argument is a matrix and a univariate function (from $\mathbb{R}$ to $\mathbb{R}$) if the argument is in $\mathbb{R}$.

Now following [6–8], [11] and [12], the kernel-function-based Nesterov–Todd direction for SDO is based in replacing the right hand side $V^{-1} - V$ in the third equation in (2.6) by $-\psi'(V)$. Thus we have the linear system:

$$\begin{cases} \bar{A}_i \bullet D_X = 0, \quad 1 \le i \le m, \\ \sum_{i=1}^{m} \Delta y_i \bar{A}_i + D_Z = 0, \\ D_X + D_Z = -\psi'(V), \end{cases} \tag{3.2}$$

where $\psi(t)$ is a given kernel function, and $\psi(V), \psi'(V)$ are the associated matrix functions. The system (3.2) has a unique symmetric solution.

In the algorithm, we use the barrier function $\Psi(V)$ defined in (3.1) as a measure function and also we introduce the norm-based proximity measure $\delta(V)$ as follows:

$$\delta(V) := \frac{1}{2}\|D_X + D_Z\| = \frac{1}{2}\|\psi'(V)\| = \frac{1}{2}\sqrt{\mathrm{Tr}(\psi'(V))^2}. \tag{3.3}$$

For the notational convenience we denote $\delta(V)$ by δ. Hence, our generic primal-dual IP algorithm for SDO is described in Figure 1 as follows.

Generic primal-dual IP algorithm for SDO

Input:
An accuracy parameter ϵ;
an update parameter $\theta, 0 < \theta < 1$;
a proximity parameter τ, $\tau > 0$;
a strictly feasible pair (X^0, Z^0) and $\mu^0 > 0$ s.t. $\Psi(X^0, Z^0, \mu^0) \le \tau$;
begin
$X := X^0;\ Z := Z^0;\ \mu := \mu^0;$
while $n\mu > \epsilon$ **do**
begin
$\mu := (1-\theta)\mu;$
while $\Psi(X, Z, \mu) > \tau$ **do**
begin
solve the system (3.2) and use (2.7) to obtain $(\Delta_X, \Delta_y, \Delta_Z)$;
determine a default step size α;
update $X := X + \alpha\Delta_X;\ y := y + \alpha\Delta_y;\ Z := Z + \alpha\Delta_Z;$
end
end
end.

Figure 1 Algorithm

4 The Barrier Kernel Function and Its Properties

In this section, we present the eligible kernel function and its growth properties for SDO.

Definition 4.1 *We call $\psi : \mathbb{R}_{++} \to \mathbb{R}_+$ a barrier kernel function if it is twice differentiable and the following conditions are satisfied*:

(1) $\psi(1) = \psi'(1) = 0$;
(2) $\psi''(t) > 0$ *for all* $t > 0$;
(3) $\lim_{t\mapsto 0} \psi(t) = \lim_{t\mapsto +\infty} \psi(t) = +\infty$.

From the conditions (1) and (2), it follows that $\psi(t)$ is strictly convex and minimal at $t = 1$, and $\psi(t)$ is expressed in term of its second derivative as follows:

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi. \tag{4.1}$$

However, the condition (3) indicates the barrier property of $\psi(t)$.

Let

$$\psi(t) = (t^2 - 1) + \frac{t^{1-q_1} - 1}{q_1 - 1} + \frac{t^{1-q_2} - 1}{q_2 - 1}, \quad q_i > 1,\ i = 1, 2,\ t > 0. \tag{4.2}$$

It is easy to check that $\psi(t)$ is indeed a barrier kernel function and its three first derivatives are

$$\psi'(t) = 2t - t^{-q_1} - t^{-q_2}, \tag{4.3}$$

$$\psi''(t) = 2 + q_1 t^{-q_1-1} + q_2 t^{-q_2-1} > 0 \quad \text{for } t > 0,\ q_1, q_2 > 1, \tag{4.4}$$

$$\psi'''(t) = -q_1(q_1+1)t^{-q_1-2} - q_2(q_2+1)t^{-q_2-2}. \tag{4.5}$$

Lemma 4.2 *Let $\psi(t)$ be as defined in* (4.2)*. Then*

$$t\psi''(t) + \psi'(t) > 0, \quad t < 1, \tag{4.6}$$

$$\psi'''(t) < 0, \quad t > 0, \tag{4.7}$$

$$t\psi''(t) - \psi'(t) > 0, \quad t > 1, \tag{4.8}$$

$$2\psi''(t)^2 - \psi'''(t)\psi'(t) > 0, \quad t < 1, \tag{4.9}$$

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, \quad t > 1, \beta > 1. \tag{4.10}$$

Proof For (4.6), by using (4.3) and (4.4), it follows that $t\psi''(t) + \psi'(t) = 4t + (q_1-1)t^{-q_1} + (q_2-1)t^{-q_2} > 0$ for all q_1, $q_2 > 1$ and $t > 0$. For (4.7), it is clear from (4.5), $\psi'''(t) < 0$ for $t > 0$. For (4.8), we have $t\psi''(t) - \psi'(t) = (q_1+1)t^{-q_1} + (q_2+1)t^{-q_2} > 0$ for all $q_1, q_2 > 1$ and $t > 0$. For (4.9), we have $2\psi''(t)^2 - \psi'''(t)\psi'(t) = 8 + (8q_1 + 2q_1(q_1+1))t^{-q_1-1} + (8q_2 + 2q_2(q_2+1))t^{-q_2-1} + +(q_1(q_1-1))t^{-2(q_1+1)} + (q_2(q_2-1))t^{-2(q_2+1)} + g(t)$, where $g(t) = (4q_1q_2 - q_1(q_1+1) - q_2(q_2+1))t^{-q_1-q_2-2}$. Then $2(\psi''(t))^2 - \psi'''(t)\psi'(t) > 0$ if $g(t) > 0$ for all $q_1, q_2 > 1$ and $t < 1$. Indeed, $g(t) > 0$ since $g(t) > [(q_1-1) + (q_2-1)]\, t^{-q_1-q_2-2} > 0$ for all $q_1, q_2 > 1$ and $t > 0$. For (4.10), by [3, Lemma 2.4], it suffices to show that $\psi(t)$ satisfies (4.6) and (4.8). These properties show the eligibility of $\psi(t)$. This completes the proof. □

Lemma 4.3 *One has*

(i) $\psi(\sqrt{t_1t_2}) \le \frac{1}{2}(\psi(t_1) + \psi(t_2))$ *for all* $t_1, t_2 > 0$,

(ii) $\frac{1}{2}(t-1)^2 \le \psi(t) \le \frac{1}{2}(\psi'(t))^2$, $t > 0$,

(iii) $\psi(t) \le (\frac{2+q_1+q_2}{4})(t-1)^2$, $t \ge 1$.

Proof For (i), by [12, Lemma 2.1.2], it suffices to show that $\psi(t)$ satisfies (4.8). This statement indicates the exponential property of $\psi(t)$. For (ii), it obtained by using the definition of $\psi(t)$ in (4.1), and the fact that $\psi''(t) > 1$ for all $t > 0$, then we have

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)d\zeta dt \ge \int_1^t \int_1^\xi d\zeta dt = \frac{1}{2}(t-1)^2.$$

The second inequality is obtained as follows:

$$\begin{aligned}\int_1^t \int_1^\xi \psi''(\zeta)d\zeta d\xi &\le \int_1^t \int_1^\xi \psi''(\xi)\psi''(\zeta)d\zeta d\xi \\ &= \int_1^t \psi''(\xi)\psi'(\xi)d\xi \\ &= \int_1^t \psi'(\xi)d\psi'(\xi)d\xi = \frac{1}{2}(\psi'(t))^2.\end{aligned}$$

For (iii), by using Taylor's theorem with $\psi(1) = \psi'(1) = 0$, $\psi''(1) = \frac{2+q_1+q_2}{2}$ and $\psi'''(t) < 0$, we obtain $\psi(t) = \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(c)(c-1)^3 = \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(c)(c-1)^3 < \frac{1}{2}\psi''(1)(t-1)^2 = (\frac{2+q_1+q_2}{4})(t-1)^2$ for some c, $1 \le c \le t$. This completes the proof. □

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 $t \in (0,1]$. Then we have the following lemma.

Lemma 4.4 *For $\psi(t)$, we have*

(i) $\sqrt{1+s} \leq \varrho(s) \leq 1+\sqrt{2s},\ s \geq 0,$

(ii) $\rho(s) \geq (\frac{1}{1+2a})^{\frac{1}{q_1-q_2}}$, *if* $q_1 > q_2$ *and* $s \geq 0$.

Proof For (i), let $s = \psi(t)$ for $t \geq 1$, i.e., $\varrho(s) = t,\ t \geq 1$. By the definition of $\psi(t)$, $s = t^2 - 1 + \psi_b(t),\ t > 0$, where $\psi_b(t) = \frac{t^{1-q_1}-1}{q_1-1} + \frac{t^{1-q_2}-1}{q_2-1}$ denotes the barrier term of $\psi(t)$. It follows that

$$t^2 - 1 + \psi_b(t) \leq t^2 - 1. \tag{4.11}$$

This inequality is due to the fact that $\psi_b(1) = 0$ and $\psi_b(t)$ is monotonically decreasing. This implies that $t = \varrho(s) \geq \sqrt{s+1}$. By Lemma 4.3 (ii), we have $s = \psi(t) \geq \frac{1}{2}(t-1)^2,\ t \geq 1$. Then we have $t = \varrho(s) \leq 1+\sqrt{2s},\ s \geq 0$. For (ii), let $a = -\frac{1}{2}\psi'(t)$ for $t \in (0,1]$. Then by definition of ρ, $\rho(a) = t \Leftrightarrow a = -\frac{1}{2}\psi'(t)$ for all $t \in (0,\ 1]$. So we have $a = -t + \frac{t^{-q_1}+t^{-q_2}}{2} \Leftrightarrow \frac{t^{-q_1}+t^{-q_2}}{2} = a+t \Leftrightarrow \frac{t^{-q_1+q_2}+1}{2t^{q_2}} = a+t \Leftrightarrow t^{-q_1+q_2}+1 = 2t^{q_2}(a+t)$. It follows that $t^{q_2-q_1} \leq 2(a+1)-1 = 2a+1$. Hence, we obtain $t \geq (\frac{1}{1+2a})^{\frac{1}{q_1-q_2}}$ for all $t \in (0,\ 1]$ and $q_1 > q_2$. This completes the proof. □

Theorem 4.5 ([8, Theorem 3.3.2]) *Let $\varrho : [0,\infty) \to [1,\infty)$ be the inverse function of $\psi(t),\ t \geq 1$. Then we have*

$$\Psi(\beta V) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(V)}{n}\right)\right), \quad \beta \geq 1 \text{ for } V \in \mathbf{S}^n_{++}.$$

In the next theorem, we obtain an estimate for the effect of a μ-update on the value of $\Psi(V)$.

Theorem 4.6 *Let $0 \leq \theta < 1$ and $V_+ = \frac{V}{\sqrt{1-\theta}}$. If $\Psi(V) \leq \tau$, then we have*

$$\Psi(V_+) \leq \frac{2+q_1+q_2}{4(1-\theta)}(\sqrt{n}\theta + \sqrt{2\tau})^2.$$

Proof Since $\frac{1}{\sqrt{1-\theta}} \geq 1$ and $\varrho(\frac{\Psi(V)}{n}) \geq 1$, we have $\frac{\varrho(\frac{\Psi(V)}{n})}{\sqrt{1-\theta}} \geq 1$. Using Theorem 4.5 with $\beta = \frac{1}{\sqrt{1-\theta}}$, Lemmas 4.3–4.4 and $\Psi(V) \leq \tau$, we have

$$\begin{aligned}
\Psi(V_+) &\leq n\psi\left(\frac{1}{\sqrt{1-\theta}}\varrho\left(\frac{\Psi(V)}{n}\right)\right)\\
&\leq \frac{(2+q_1+q_2)n}{4}\left(\frac{\varrho(\frac{\Psi(V)}{n}) - \sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2\\
&\leq \frac{(2+q_1+q_2)n}{4}\left(\frac{\varrho(\frac{\Psi(V)}{n}) - \sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2\\
&\leq \frac{(2+q_1+q_2)n}{4}\left(\frac{1+\sqrt{\frac{2\tau}{n}} - \sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2\\
&\leq \frac{(2+q_1+q_2)n}{4}\left(\frac{\theta+\sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}}\right)^2\\
&= \frac{2+q_1+q_2}{4(1-\theta)}(\sqrt{n}\theta + \sqrt{2\tau})^2,
\end{aligned}$$

where the last inequality holds since $1-\sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \leq \theta$ for all $0 \leq \theta < 1$. This completes the proof. □

Denote

$$\tilde{\Psi}_0 = \frac{2+q_1+q_2}{4(1-\theta)}(\sqrt{n}\theta + \sqrt{2\tau})^2. \tag{4.12}$$

Then $\tilde{\Psi}_0$ is an upper bound for $\Psi(V)$ during the Newton process of the algorithm.

In the next proposition, we give a lower bound for the proximity δ in term of $\Psi(V)$.

4.1 A Lower Bound for δ in Term of $\Psi(V)$

Proposition 4.7 *For any* $V \succ 0$,

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

Proof By Lemma 4.3(ii) and (3.3), we have

$$\delta^2 = \frac{1}{4}\mathrm{Tr}(\psi'(V)^2) \geq \frac{2}{4}\sum_{i=1}^{n}\psi(\lambda_i(V)) \geq \frac{1}{2}\Psi(V).$$

Hence, we have $\delta \geq \sqrt{\frac{\Psi(V)}{2}}$. This completes the proof. □

Remark 4.8 During the algorithm we assume that $\tau \geq 1$. Using Proposition 4.7 and the assumption $\Psi(V) \geq \tau$, we have $\delta \geq \sqrt{\frac{\Psi(V)}{2}} \geq \frac{1}{\sqrt{2}}$.

5 Complexity Analysis

In this subsection, we compute a default step size α and the decrease of the proximity function during an inner iteration. After a step size α, new iterates are denoted by

$$X_+ = X + \alpha\Delta X, \quad y_+ = y + \alpha\Delta y, \quad Z_+ = Z + \alpha\Delta Z.$$

5.1 Determining a Default Step Size

Define for $\alpha > 0$,

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

Then $f(\alpha)$ is the difference of the proximity between a new iterate and a current iterate for a fixed $\mu > 0$.

Due to (2.5) and (2.7), we may write

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

and

$$Z_+ = Z + \alpha\Delta Z = Z + \alpha\sqrt{\mu}D^{-1}D_ZD^{-1} = \sqrt{\mu}D^{-1}(V + \alpha D_Z)D^{-1}.$$

Thus we have

$$V_+^2 := (V + \alpha D_X)(V + \alpha D_Z),$$

and

$$V_+^2 = (V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_Z)(V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_X)^{-\frac{1}{2}},$$

by assuming that $V + \alpha D_X \succ 0$ and $V + \alpha D_Z \succ 0$ for such feasible step size α. Then it is clear that V_+^2 is similar to the matrix

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

As a result we deduce that V_+ has the same eigenvalues as the matrix

$$[(V+\alpha D_X)^{\frac{1}{2}}(V+\alpha D_Z)(V+\alpha D_X)^{\frac{1}{2}}]^{\frac{1}{2}},$$

and then we have

$$\Psi(V_+)=\Psi([(V+\alpha D_X)^{\frac{1}{2}}(V+\alpha D_Z)(V+\alpha D_X)^{\frac{1}{2}}]^{\frac{1}{2}}).$$

Next we show that our proximity $\Psi(V)$ satisfy the following proposition.

Proposition 5.1 ([11, Proposition 5.2.6]) *For any* $V_1, V_2 \succ 0$,

$$\Psi((V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}})^{\frac{1}{2}}) \le \frac{1}{2}(\Psi(V_1)+\Psi(V_2)).$$

Using Proposition 5.1, it follows that

$$f(\alpha)\le f_1(\alpha),$$

where

$$f_1(\alpha)=\frac{1}{2}\left[\Psi(V+\alpha D_X)+\Psi(V+\alpha D_Z)\right]-\Psi(V).$$

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

Now to estimate the decrease of the proximity during one step, we need the two successive derivatives of $f_1(\alpha)$ with respect to α. By using the rule of differentiability involving matrix functions, we obtain

$$f_1'(\alpha)=\frac{1}{2}\mathrm{Tr}\left[\psi'(V+\alpha D_X)D_X+\psi'(V+\alpha D_Z)D_Z\right]$$

and

$$f_1''(\alpha)=\frac{1}{2}\mathrm{Tr}\left[\psi''(V+\alpha D_X)D_X^2+\psi''(V+\alpha D_Z)D_Z^2\right].$$

It is obvious that $f_1''(\alpha)>0$, unless $D_X=D_Z=0$.

Using the definition of δ, we get

$$\begin{aligned} f_1'(0) &= \frac{1}{2}\mathrm{Tr}\left[\psi'(V)D_X+\psi'(V)D_Z\right], \\ &= \frac{1}{2}\mathrm{Tr}\psi'(V)(D_X+D_Z), \\ &= -2\delta^2. \end{aligned}$$

Lemma 5.2 ([3, Lemma 4.1]) *One has*

$$f_1''(\alpha)\le 2\delta^2\psi''(\lambda_1(V)-2\alpha\delta),$$

where $\lambda_1(V)$ *is the smallest eigenvalue of* V.

Lemma 5.3 ([3, Lemma 4.2]) *One has* $f_1'(\alpha)\le 0$ *if* α *satisfies the inequality*

$$-\psi'(\lambda_1(V)-2\alpha\delta)+\psi'(\lambda_1(V))\le 2\delta. \tag{5.1}$$

Lemma 5.4 ([3, Lemma 4.4]) *Let* ρ *and* $\tilde{\alpha}$ *be defined as* (5.1)*. Then*

$$\tilde{\alpha}\ge\frac{1}{\psi''(\rho(2\delta))}.$$

Lemma 5.5 *One has*

$$\tilde{\alpha}\ge\frac{1}{2+2q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}}. \tag{5.2}$$

Proof Using the definition of $\psi''(t)$, and Lemma 4.4 (ii), we have

$$\begin{aligned}\tilde{\alpha} &\geq \frac{1}{\psi''(\rho(2\delta))}\\ &\geq \frac{1}{2+q_1(\rho(2\delta))^{q_1+1}+q_2(\rho(2\delta))^{q_2+1}}\\ &= \frac{1}{2+q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}+q_2(1+4\delta)^{\frac{q_2+1}{q_1-q_2}}}.\end{aligned}$$

Assuming that $q_1 \geq q_2$, then it follows that

$$q_2(1+4\delta)^{\frac{q_2+1}{q_1-q_2}} \leq q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}$$

and we deduce that

$$\tilde{\alpha} \geq \frac{1}{2+2q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}}.$$

This completes the proof. □

As a default step size, we take

$$\bar{\alpha} = \frac{1}{2+2q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}}, \tag{5.3}$$

with $\tilde{\alpha} \geq \bar{\alpha}$.

Next lemma shows that our proximity function Ψ with the default step size $\bar{\alpha}$ is decreasing.

Lemma 5.6 ([11, Lemma 3.12]) *Suppose that $h(t)$ is a twice differentiable convex function with*

$$h(0) = 0, \quad h'(0) < 0$$

and attains its global minimum at $t^* > 0$, and $h''(t)$ is increasing with respect to t. Then for any $t \in [0, t^*]$, we have

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

Since $f_1(\alpha)$ holds the condition of the above lemma,

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

then we have the following lemma to obtain the upper bound for the decreasing value of the proximity in the inner iteration.

Lemma 5.7 ([3, Lemma 4.5]) *If the step size α is such that $\alpha \leq \tilde{\alpha}$, then*

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

Lemma 5.8 *Let $\bar{\alpha}$ be a step size as defined in (5.3), and $\delta \geq 1$. Then we have*

$$f(\bar{\alpha}) \leq -\frac{\delta^2}{2+2q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}}.$$

Proof By Lemma 5.7 and (5.3), we have

$$f(\bar{\alpha}) \leq -\frac{\delta^2}{2+2q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}}. \tag{5.4}$$

This completes the proof. □

Now, we express the decrease in one inner iteration in terms of $\Psi(V)$ by using Proposition 4.7 as follows.

Lemma 5.9 *One has*

$$f(\bar{\alpha}) \leq -\frac{1}{24q_1}\Psi^{\frac{q_1-2q_2-1}{2(q_1-q_2)}}.$$

Proof Using Proposition 4.7, Remark 4.8, and substituting the value of $\bar{\alpha}$, we have

$$\begin{aligned} f(\bar{\alpha}) &\leq -\bar{\alpha}\delta^2 \\ &= -\frac{\delta^2}{2+2q_1(1+4\delta)^{\frac{q_1+1}{q_1-q_2}}} \\ &\leq -\frac{\delta^2}{2+2q_1(\sqrt{2}\delta+4\delta)^{\frac{q_1+1}{q_1-q_2}}} \\ &= -\frac{\delta^2}{4q_1(\sqrt{2}\delta+4\delta)^{\frac{q_1+1}{q_1-q_2}}} \\ &\leq -\frac{\delta^2}{22q_1\delta^{\frac{q_1+1}{q_1-q_2}}} \\ &= -\frac{\Psi}{\frac{44}{\sqrt{2}}\Psi^{\frac{q+1}{2(q_1-q_2)}}} = \frac{1}{31q_1}\Psi^{\frac{q_1-q-2q_2-1}{2(q_1-q_2)}}, \quad \text{since} \frac{44}{\sqrt{2}} \simeq 31. \end{aligned}$$

This completes the proof. □

Now we are ready to estimate the total number of iteration bound of the algorithm.

5.2 Iteration Bound

At this stage, we need the following proposition from Proposition 1.3.2 in [12] without proof.

Proposition 5.10 *Let $t_0, t_1, \ldots, t_K$ be a sequence of positive numbers such that*

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

where $\gamma > 0$ and $0 < \bar{\beta} \leq 1$. Then

$$K \leq \frac{t_0^{\bar{\beta}}}{\gamma\bar{\beta}}.$$

Denote the value of $\Psi(v)$ after the μ-update as Ψ_0 and the subsequent values in the same manner as Ψ_k, $k = 1, 2, \ldots$. Then we have

$$\Psi_0 \leq \tilde{\Psi}_0$$

with $\tilde{\Psi}_0$ being defined in (4.12). Let K be the total number of inner iterations per the outer iteration. Then we have

$$\Psi_{K-1} > \tau, \quad 0 \leq \Psi_k \leq \tau.$$

Lemma 5.11 *Let $\tilde{\Psi}_0$ be defined as in* (4.12) *and K be the total number of inner iterations in the outer iteration. Then we have*

$$K \leq \frac{62q_1(q_1-q_2)}{q_1+1}\tilde{\Psi}_0^{\frac{q_1+1}{2(q_1-q_2)}}.$$

Proof Using Proposition 5.10 and Lemma 5.9 with $\gamma = \frac{1}{31q_1}$ and $\bar{\beta} = \frac{q_1+1}{2(q_1-q_2)}$, we get

$$K \leq \frac{62q_1(q_1-q_2)}{q_1+1} \tilde{\Psi}_0^{\frac{q_1+1}{2(q_1-q_2)}}.$$

This completes the proof. □

Theorem 5.12 *If $\tau \geq 1$, the total number of iterations to get a primal-dual optimal solution for* SDO *is no more than*

$$\frac{K}{\theta}\log\frac{n}{\epsilon} \leq \frac{62(q_1+1)^{\frac{3q_1-2q_2+1}{2(q_1-q_2)}}}{\theta}\left(\frac{(\sqrt{n}\theta+\sqrt{2\tau})^2}{(1-\theta)}\right)^{\frac{q_1+1}{2(q_1-q_2)}}\log\frac{n}{\epsilon}.$$

Proof Since $n\mu \leq \epsilon$, $\mu_k := (1-\theta)^k\mu_0$ and $\mu_0 = 1$, by a simple computation, we obtain

$$k \leq \frac{1}{\theta}\log\frac{n}{\epsilon}.$$

Therefore the number of outer iterations is bounded above by

$$\frac{1}{\theta}\log\frac{n}{\epsilon}.$$

Multiplication of this result by the number in Lemma 5.11 the theorem holds. Hence, we have

$$\begin{aligned}\frac{K}{\theta}\log\frac{n}{\epsilon} &\leq \eta\left(\frac{(2+q_1+q_2)(\sqrt{n}\theta+\sqrt{2\tau})^2}{4(1-\theta)}\right)^{\frac{q_1+1}{2(q_1-q_2)}}\log\frac{n}{\epsilon}\\ &= \eta\left(\frac{(\sqrt{n}\theta+\sqrt{2\tau})^2}{1-\theta}\right)^{\frac{q_1+1}{2(q_1-q_2)}}\log\frac{n}{\epsilon},\end{aligned}$$

with $\eta = \frac{62q_1(q_1-q_2)}{\theta(q_1+1)}\left(\frac{2+q_1+q_2}{4}\right)^{\frac{q_1+1}{2(q_1-q_2)}}$.

Now since $\left(\frac{2+q_1+q_2}{4}\right)^{\frac{q_1+1}{2(q_1-q_2)}} \leq \frac{1}{2}(q_1+1)^{\frac{q_1+1}{2(q_1-q_2)}}$ and $\frac{62q_1(q_1-q_2)}{(q_1+1)} \leq 62(q_1+1)$ for all $q_1 > q_2 > 1$, it follows that

$$\frac{K}{\theta}\log\frac{n}{\epsilon} \leq \frac{62(q_1+1)^{\frac{3q_1-2q_2+1}{2(q_1-q_2)}}}{\theta}\left(\frac{(\sqrt{n}\theta+\sqrt{2\tau})^2}{(1-\theta)}\right)^{\frac{q_1+1}{2(q_1-q_2)}}\log\frac{n}{\epsilon}.$$

This completes the proof. □

Remark 5.13 If $\tau = O(1)$ and $\theta = \Theta(n)$, then the iteration bound for the small-update algorithm is

$$O\left((q_1+1)^{\frac{3q_1-2q_2+1}{2(q_1-q_2)}}\sqrt{n}\log\frac{n}{\epsilon}\right).$$

In particular if $q_1 = lq_2$, with $l > 1$, then the total iteration bound is

$$O\left((q_1+1)^{\frac{(3l-2)q_1+1}{2(l-1)q_1}}\sqrt{n}\log\frac{n}{\epsilon}\right).$$

Therefore, the larger l, the better the total iteration bound. In addition if $l \mapsto \infty$, then the iteration bound is

$$O\left((q_1+1)^{\frac{3q_1+1}{2q_1}}\sqrt{n}\log\frac{n}{\epsilon}\right).$$

Remark 5.14 Since

$$\Psi_0 \leq n\psi\left(\frac{1}{\sqrt{1-\theta}}\varrho\left(\frac{\Psi(V)}{n}\right)\right) \leq n\psi\left(\frac{1+\sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}}\right)$$

and by (4.11), if $t \geq 1$, then $\psi(t) \leq (t^2 - 1)$. Using this we obtain

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

With this estimate the total number of iterations is bounded above by

$$\begin{aligned}\Psi_0 &\leq \frac{62q_1(q_1-q_2)}{\theta(q_1+1)}\left(\frac{\theta n + 2\sqrt{2n\tau} + 2\tau}{(1-\theta)}\right)^{\frac{q_1+1}{2(q_1-q_2)}} \log\frac{n}{\epsilon} \\ &\leq \frac{62(q_1+1)}{\theta}\left(\frac{\theta n + 2\sqrt{2n\tau} + 2\tau}{(1-\theta)}\right)^{\frac{q_1+1}{2(q_1-q_2)}} \log\frac{n}{\epsilon}\end{aligned}$$

since

$$\frac{62q_1(q_1-q_2)}{q_1+1} \leq \frac{62q_1^2}{q_1+1} \leq 62(q_1+1) \quad \text{for all } q_1 > q_2 > 1.$$

Therefore, if $\tau = O(n)$ and $\theta = \Theta(1)$, the iteration bound for the large-update algorithm is

$$O\big((q_1+1)n^{\frac{q_1+1}{2(q_1-q_2)}}\big)\log\frac{n}{\epsilon}.$$

In particular, if $q_1 = lq_2$ with $l > 1$, then the total iteration bound is

$$O\left((q_1+1)n^{\frac{l(q_1+1)}{2(l-1)q_1}}\log\frac{n}{\epsilon}\right).$$

In addition, if $l \mapsto \infty$, then the total iteration bound is

$$O\left((q_1+1)n^{\frac{q_1+1}{2q_1}}\log\frac{n}{\epsilon}\right).$$

6 Conclusion

In this paper, we proposed a new primal-dual path-following interior point algorithm for solving semidefinite optimization problems based on a new kernel function which has a double barrier term. The algorithm yields the iteration bounds $O((q_1+1)\,n^{\frac{q_1+1}{2(q_1-q_2)}}\log\frac{n}{\epsilon})$, and $O((q_1+1)^{\frac{3q_1-2q_2+1}{2(q_1-q_2)}}\sqrt{n}\log\frac{n}{\epsilon})$ for large and small-update algorithms, respectively provided that $q_1 > q_2 > 1$. Future researches might extend this analysis for linear and convex quadratic optimization problems, complementarity and conic problems. Finally, numerical tests will be an important topic of research in the future.

References

[1] Achache, M.: A new primal-dual path-following method for convex quadratic programming. *Comput. Appl. Math.*, **25**(1), 97–110 (2006)

[2] Alizadeh, F.: Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal. Optim.*, **5**(1), 13–51 (2006)

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

[4] Bai, Y. Q., Wang, G. Q., Roos, C.: Primal-dual interior point interior point algorithms for second-order cone optimization based on kernel functions. *Nonlinear Anal.*, **70**, 3548–3602 (2009)

[5] Bai, Y. Q., Roos, C., El Ghami, M.: A primal-dual interior point method for linear optimization based on a new proximity function. *Optimiz. Methods Software*, **17**(6), 985–1008 (2002)

[6] Choi, B. K., Lee, 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)

[7] El Ghami, M., Steihaug, T., Roos, C.: Primal-dual IPMs for semidefinite optimization based on finite barrier functions, Report $N^0$341, Department of Informatics, University of Bergen, Bergen, Norway, 2006

[8] El Ghami, M.: New primal-dual interior-point methods based on kernel functions. Ph.D. thesis, Delft University, Netherland, 2005

[9] Liu, L., Li, S.: A new kind of simple kernel function yielding good iteration bounds for for primal-dual interior point methods. *J. Comput. Appl. Math.*, **235**(9), 2944–2955 (2011)

[10] Cho, G. M.: A new primal-dual interior point method for linear optimization. *J. KSIAM*, **13**(1), 41–53 (2009)

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

[12] Peng, J., Roos, C., Terlaky, T.: Self-regular functions and new search directions for linear and semidefinite optimization. *Math. Programming*, **93**, 129–117 (2002)

[13] Peng, J., Roos, C., Terlaky, T.: Self-regularity. A New Paradigm for Primal-Dual Interior Point Algorithm, Princeton University Press, Princeton, 2002

[14] Roos, C., Terlaky, T., Vial, J. Ph.: Theory and Algorithms for Linear Optimization. An Interior Point Approach. John Wiley and Sons, Chister, U.K., 1997

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

[16] Vanderberghe, L., Boyd, S.: Semidefinite programming. *SIAM Review*, **38**(1), 49–95 (1996)

[17] Horn, R. A., Johnson, C. R.: Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991