J. Math. Anal. Appl. 353 (2009) 339–349

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

# A new primal-dual path-following interior-point algorithm for semidefinite optimization

G.Q. Wang [a,b,*,1], Y.Q. Bai [b,2]

[a] College of Vocational Technology, Shanghai University of Engineering Science, Shanghai 200437, PR China
[b] Department of Mathematics, College of Sciences, Shanghai University, Shanghai 200444, PR China

ARTICLE INFO

*Article history:*
Received 5 April 2008
Available online 13 December 2008
Submitted by M. Laurent

*Keywords:*
Semidefinite optimization
Interior-point algorithm
Small-update method
Iteration bound

ABSTRACT

In this paper we present a new primal-dual path-following interior-point algorithm for semidefinite optimization. The algorithm is based on a new technique for finding the search direction and the strategy of the central path. At each iteration, we use only full Nesterov–Todd step. Moreover, we obtain the currently best known iteration bound for the algorithm with small-update method, namely, $O(\sqrt{n}\log\frac{n}{\epsilon})$, which is as good as the linear analogue.



## 1. Introduction

Semidefinite optimization (SDO) problems are convex optimization problems over the intersection of an affine set and the cone of positive semidefinite matrices. SDO has wide applications in continuous and combinatorial optimization [2,21]. In the past decade, SDO has become a popular research area in mathematical programming when it became clear that the algorithm for linear optimization (LO) can often be extended to the more general SDO case. Several interior-point methods (IPMs) designed for LO have been successfully extended to SDO [12,16,20,22] and second-order cone optimization (SOCO) [7]. An important contribution to this field was made by Nesterov and Todd [13,14] who showed that the primal-dual algorithm maintains its theoretical efficiency when the nonnegativity constrains in LO are replaced by a convex cone, as long as the cone is homogeneous and self-dual. An interesting fact is that almost all known polynomial-time variants of IPMs use the so-called central path as a guideline to the optimal set, and some variants of Newton's method follow the central path approximately. For an overview of these related results we refer to [10,17,19,21] and their references. In particular, primal-dual interior-point algorithms are of high efficiency both in theory and in practice. The convergence rates, the stability and the numerical results of primal-dual interior-point methods for SDO have been presented by Alizadeh in [3]. Recently, Darvay [8] proposed a new technique for finding a class of search directions. Based on this technique, the author designed a new primal-dual path-following interior-point algorithm for LO with iteration bound $O(\sqrt{n}\log\frac{n}{\epsilon})$. Later on Achache [1] extended it to convex quadratic optimization (CQO).

Motivated by their work, we propose a new primal-dual path-following interior-point algorithm for SDO. We adopt the basic analysis used in [8] to the SDO case. The favorable iteration bound for the algorithm with small-update method,

* Corresponding author at: College of Vocational Technology, Shanghai University of Engineering Science, Shanghai 200437, PR China.
*E-mail addresses:* guoq_wang@hotmail.com (G.Q. Wang), yqbai@shu.edu.cn (Y.Q. Bai).

1 The research is supported by the Foundation of Scientific Research for Selecting and Cultivating Young Excellent University Teachers in Shanghai (No. 06XPYQ52) and Shanghai Educational Committee Foundation (No. 06NS031).
2 The research is supported by National Natural Science Foundation of China (No. 10117733) and Shanghai Leading Academic Discipline Project (No. J50101).

0022-247X/$ – see front matter 
doi:10.1016/j.jmaa.2008.12.016

namely, $O(\sqrt{n}\log\frac{n}{\varepsilon})$ is obtained. The iteration bound is as good as the bound for the LO case. Moreover, our analysis is relatively simple and straightforward to the LO analogue.

The outline of the paper is as follows. In Section 2, we recall some well-known results on matrices and matrix functions. In Section 3, we briefly introduce the central path for SDO and its properties. In Section 4, we extend Darvay's new technique for LO to SDO and obtain the new search direction for SDO. In Section 5, we present the generic primal-dual path-following algorithm for SDO. In Section 6, we analyze the algorithm and derive the iteration bound with small-update method. Finally, some conclusions and remarks follow in Section 7.

Some notations used throughout the paper are as follows. $\mathbf{R}^n$, $\mathbf{R}^n_+$ and $\mathbf{R}^n_{++}$ denote the set of vectors with $n$ components, the set of nonnegative vectors and the set of positive vectors, respectively. $\mathbf{R}^{n\times n}$ denotes the set of $n\times n$ real matrices. $\|\cdot\|_F$ and $\|\cdot\|_2$ denote the Frobenius norm and the spectral norm for matrices, respectively. $\mathbf{S}^n$, $\mathbf{S}^n_+$ and $\mathbf{S}^n_{++}$ denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n\times n$ matrices, respectively. The Löwner partial order "$\succcurlyeq$" (or "$\succ$") on positive semidefinite (or positive) matrices means $A\succcurlyeq B$ (or $A\succ B$) if $A-B$ is positive semidefinite (or positive). We use the matrix inner product $A\bullet B=\mathbf{Tr}(A^TB)$ (i.e., the trace of the matrix $A^TB$). For any $A\in\mathbf{R}^{n\times n}$, $\det(A)$ denotes the determinant of $A$. For any $Q\in\mathbf{S}^n_{++}$, the expression $Q^{\frac{1}{2}}$ (or $\sqrt{Q}$) denotes its symmetric square root. When $\lambda$ is a vector we denote the diagonal matrix $\Lambda$ with entries $\lambda_i$ by $\operatorname{diag}(\lambda)$. For any $V\in\mathbf{S}^n$, $\lambda_{\max}(V)$ and $\lambda_{\min}(V)$ denote the largest eigenvalue and the smallest eigenvalue of $V$. Furthermore, we assume that the eigenvalues of $V$ are listed according to the order of their absolute values such that $|\lambda_1(V)|\geqslant|\lambda_2(V)|\geqslant\cdots\geqslant|\lambda_n(V)|$. If $V$ is positive semidefinite, then we have $\lambda_{\min}(V)=\lambda_n(V)\geqslant 0$, $\lambda_{\max}(V)=\lambda_1(V)$. Finally, if $g(x)\geqslant 0$ is a real valued function of a real nonnegative variable, the notation $g(x)=O(x)$ means that $g(x)\leqslant\bar{c}x$ for some positive constant $\bar{c}$ and $g(x)=\Theta(x)$ that $c_1x\leqslant g(x)\leqslant c_2x$ for two positive constants $c_1$ and $c_2$.

## 2. Preliminaries on matrices and matrix functions

Firstly, we briefly recall some known facts from linear algebra. For the details we refer to the book [9].

**Definition 2.1.** Let $A,B\in\mathbf{R}^{n\times n}$, $c\in\mathbf{R}$, we call a function $\|\cdot\|:\mathbf{R}^{n\times n}\to\mathbf{R}$ a matrix norm if it satisfies the following axioms

(1) $\|A\|\geqslant 0$;
(2) $\|cA\|=|c|\|A\|$;
(3) $\|A+B\|\leqslant\|A\|+\|B\|$;
(4) $\|AB\|\leqslant\|A\|\|B\|$.

It is well known that for any matrix $A\in\mathbf{R}^{n\times n}$, the Frobenius norm

$$\|A\|_F=\sqrt{\mathbf{Tr}\big(A^TA\big)}=\sqrt{\sum_{i=1}^n\sum_{j=1}^n a_{ij}^2}\quad\left(=\sqrt{\sum_{i=1}^n\lambda_i^2(A)}\ \text{if}\ A\in\mathbf{S}^n\right),$$

and the spectral norm

$$\|A\|_2=\sqrt{\lambda_{\max}\big(A^TA\big)}\quad\big(=\lambda_{\max}(A)\ \text{if}\ A\succcurlyeq 0\big),$$

for matrices are norms in this sense.

The trace of the matrix $A$ is the sum of the diagonal elements of the matrix $A$ and often denoted by $\mathbf{Tr}(A)$. It has the following property.

**Property 2.2.** *Let $A,B\in\mathbf{R}^{n\times n}$, then*

(1) $\mathbf{Tr}(A)=\sum_{i=1}^n\lambda_i(A)$, *where $\lambda_i(A)$ is the ith eigenvalue of matrix $A$*;
(2) $\mathbf{Tr}(A)=\mathbf{Tr}(A^T)$;
(3) $\mathbf{Tr}(AB)=\mathbf{Tr}(BA)$;
(4) $\mathbf{Tr}(A+B)=\mathbf{Tr}(A)+\mathbf{Tr}(B)$.

**Theorem 2.3** *(Spectral theorem for symmetric matrices). (See [21].) The real $n\times n$ matrix $A$ is symmetric if and only if there exists a matrix $Q\in\mathbf{R}^{n\times n}$ such that $Q^TQ=E$ and $Q^TAQ=\Lambda$ where $\Lambda$ is a diagonal matrix.*

It readily follows from Theorem 2.3 that $\lambda_i(A^2)=\lambda_i^2(A)$, $i=1,2,\ldots,n$, whence $A^2\in\mathbf{S}^n_+$.

Throughout the paper, we assume that $\psi(t)$ is a real valued function on $[0,+\infty)$ and differentiable on $(0,+\infty)$ such that $\psi'(t)>0$ for all $t>0$. Now we are ready to show how a matrix function can be obtained from $\psi(t)$.

**Definition 2.4.** Let $V \in \mathbf{S}^n_+$ and

$$V = Q^T \operatorname{diag}\big(\lambda_1(V), \lambda_2(V), \ldots, \lambda_n(V)\big) Q,$$

where $Q$ is any orthonormal matrix ($Q^T = Q^{-1}$) that diagonalizes $V$. The matrix valued function $\psi(V)$ is defined by

$$\psi(V) = Q^T \operatorname{diag}\big(\psi\big(\lambda_1(V)\big), \psi\big(\lambda_2(V)\big), \ldots, \psi\big(\lambda_n(V)\big)\big) Q. \tag{1}$$

It should be noted that the matrix $Q$ is not unique, but $\psi(V)$ is well defined whenever $\psi(t)$ is well defined on the eigenvalues of $V$ [16].

Furthermore, replacing $\psi(\lambda_i(V))$ in (1) by $\psi'(\lambda_i(V))$, we can conclude that the matrix functions $\psi'(V)$

$$\psi'(V) = Q^T \operatorname{diag}\big(\psi'\big(\lambda_1(V)\big), \psi'\big(\lambda_2(V)\big), \ldots, \psi'\big(\lambda_n(V)\big)\big) Q \tag{2}$$

is defined as well.

It is well known that two matrixes $A$ and $B$ are called similar (abbreviated $A \sim B$) if $A = PBP^{-1}$ for some invertible matrix $P$ and, moreover, if $A$ and $B$ are symmetric then this happens if and only if $A$ and $B$ have the same eigenvalues [9].

**Lemma 2.5.** *Let $A, B \in \mathbf{S}^n$, and $AB = BA$, then*

$$\lambda_i(A+B) = \lambda_i(A) + \lambda_i(B), \quad i = 1, 2, \ldots, n.$$

*Furthermore, if $|\lambda_i(B)|$ is small enough, we have*

$$\psi(A+B) \doteq \psi(A) + \psi'(A)B.$$

**Proof.** Theorem 2.3 implies that the first part of the lemma. On the other hand, since $\psi(V)$ depends only on the eigenvalues of the matrix $V$, from Taylor's theorem, the second part of the lemma is trivial. □

**Lemma 2.6.** *Let $t > 0$ and $V \in \mathbf{S}^n_+$, then*

$$\big\|\big(tE - V^2\big)(tE + V)^{-1}\big\|_F \leqslant \frac{1}{t + \lambda_{\min}(V)} \big\|tE - V^2\big\|_F.$$

**Proof.** Since $V \in \mathbf{S}^n_+$, from Theorem 2.3 we have

$$V = Q^T \operatorname{diag}\big(\lambda_1(V), \lambda_2(V), \ldots, \lambda_n(V)\big) Q,$$

and

$$V^2 = Q^T \operatorname{diag}\big(\lambda_1^2(V), \lambda_2^2(V), \ldots, \lambda_n^2(V)\big) Q.$$

Then

$$tE - V^2 = Q^T \operatorname{diag}\big(t - \lambda_1^2(V), t - \lambda_2^2(V), \ldots, t - \lambda_n^2(V)\big) Q,$$

and

$$(tE + V)^{-1} = Q^T \operatorname{diag}\left(\frac{1}{t + \lambda_1(V)}, \frac{1}{t + \lambda_2(V)}, \ldots, \frac{1}{t + \lambda_n(V)}\right) Q.$$

Thus

$$\big\|\big(tE - V^2\big)(tE + V)^{-1}\big\|_F = \sqrt{\sum_{i=1}^n \left(\frac{t - \lambda_i^2(V)}{t + \lambda_i(V)}\right)^2} \leqslant \frac{1}{t + \lambda_{\min}(V)} \sqrt{\sum_{i=1}^n \big(t - \lambda_i^2(V)\big)^2} \leqslant \frac{1}{t + \lambda_{\min}(V)} \big\|tE - V^2\big\|_F.$$

This completes the proof of the lemma. □

**Remark 2.7.** In the rest of the section, when we use the function $\psi(\cdot)$ and its derivatives $\psi'(\cdot)$, they denote matrix function if the argument is a matrix and a univariate function if the argument is in $\mathbf{R}$.

## 3. The central path

We consider the SDO problem in standard form

$$(P) \quad \text{minimize } C \bullet X \quad \text{subject to} \quad A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m, \; X \succcurlyeq 0 \tag{3}$$

and its dual problem

$$(D) \quad \text{maximize } b^T y \quad \text{subject to} \quad \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succcurlyeq 0, \tag{4}$$

where each $A_i \in \mathbf{S}^n$, $b \in \mathbf{R}^m$, and $C \in \mathbf{S}^n$. Moreover, the matrices $A_i$ are linearly independent.

Throughout the paper, we assume that $(P)$ and $(D)$ satisfy the interior-point condition (IPC), i.e., there exists $(X^0 \succ 0, y^0, S^0 \succ 0)$ such that

$$A_i \bullet X^0 = b_i, \quad X^0 \succ 0, \; i = 1, 2, \ldots, m, \qquad \sum_{i=1}^{m} y_i^0 A_i + S^0 = C, \quad S^0 \succ 0.$$

It is well known that the IPC can be assumed without loss of generality. In fact we may choose $X^0 = S^0 = E$ as the initial start point, where $E$ is the $n \times n$ unit matrix. The detailed analysis can be found in [10,21]. The optimality conditions for $(P)$ and $(D)$ are given by the following system

$$\begin{aligned} & A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m, \; X \succ 0, \\ & \sum_{i=1}^{m} y_i A_i + S = C, \quad S \succ 0, \\ & XS = 0. \end{aligned} \tag{5}$$

If the IPC holds, the $\mu$-central of $(P)$ and $(D)$ is defined by the solution $(X(\mu), y(\mu), S(\mu))$ of the following system

$$\begin{aligned} & A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m, \; X \succ 0, \\ & \sum_{i=1}^{n} y_i A_i + S = C, \quad S \succ 0, \\ & XS = \mu E, \end{aligned} \tag{6}$$

with $\mu > 0$. The set of $\mu$-centers (with $\mu$ running through all positive real numbers) gives a homotopy path, 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 an $\varepsilon$-approximate solution for $(P)$ and $(D)$ [10,21].

## 4. The new search directions

In [8], Darvay presented a new technique for finding a class of search directions for LO. He replaces the standard centering equation $xs = \mu e$ by $\psi(\frac{xs}{\mu}) = \psi(e)$, where $\psi(\cdot)$ is the vector function induced by function $\psi(t)$, and then applies Newton's method to obtain the new search directions. Similar to the LO case, we replace the standard centering equation $XS = \mu E$ by $\psi(\frac{XS}{\mu}) = \psi(E)$, then the system (6) can be written as

$$\begin{aligned} & A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m, \; X \succ 0, \\ & \sum_{i=1}^{n} y_i A_i + S = C, \quad S \succ 0, \\ & \psi\left(\frac{XS}{\mu}\right) = \psi(E). \end{aligned} \tag{7}$$

Applying Newton's method to system (7) produces the following equations for the search direction $\Delta X$, $\Delta y$ and $\Delta S$

$$\begin{aligned} & A_i \bullet (X + \Delta X) = b_i, \quad i = 1, 2, \ldots, m, \\ & \sum_{i=1}^{n} (y_i + \Delta y_i) A_i + (S + \Delta S) = C, \\ & \psi\left(\frac{(X + \Delta X)(S + \Delta S)}{\mu}\right) = \psi(E). \end{aligned} \tag{8}$$

The third equation of the system (8) is equivalent to

$$\psi\left(\frac{XS}{\mu}+\frac{X\Delta S+\Delta XS+\Delta X\Delta S}{\mu}\right)=\psi(E). \tag{9}$$

Applying Lemma 2.5 and neglecting the term $\Delta X\Delta S$, Eq. (9) can be written as

$$\psi\left(\frac{XS}{\mu}\right)+\psi'\left(\frac{XS}{\mu}\right)\left(\frac{X\Delta S+\Delta XS}{\mu}\right)=\psi(E). \tag{10}$$

Then we consider the following system

$$\begin{aligned}
&A_i\bullet\Delta X=0,\quad i=1,2,\ldots,m,\\
&\sum_{i=1}^{n}\Delta y_iA_i+\Delta S=0,\\
&\Delta X+X\Delta SS^{-1}=\mu\left(\psi'\left(\frac{XS}{\mu}\right)\right)^{-1}\left(\psi(E)-\psi\left(\frac{XS}{\mu}\right)\right)S^{-1}.
\end{aligned} \tag{11}$$

It is obvious that $\Delta S$ is symmetric due to the second equation in (11). However, a crucial observation is that $\Delta X$ is not necessarily symmetric because $X\Delta SS^{-1}$ may not be symmetric [20]. Several ways exist for symmetrizing the third equation in the Newton system such that the resulting new system has a unique symmetric solution [10,13,14,16–18,21].

In this paper we consider the Nesterov–Todd (NT)-symmetrization scheme in [13,14]. Let us define

$$P:=X^{\frac12}\left(X^{\frac12}SX^{\frac12}\right)^{-\frac12}X^{\frac12}=S^{-\frac12}\left(S^{\frac12}XS^{\frac12}\right)^{\frac12}S^{-\frac12}. \tag{12}$$

We replace the term $X\Delta SS^{-1}$ in the third equation of (11) by $P\Delta SP^T$. The system (11) becomes

$$\begin{aligned}
&A_i\bullet\Delta X=0,\quad i=1,2,\ldots,m,\\
&\sum_{i=1}^{n}\Delta y_iA_i+\Delta S=0,\\
&\Delta X+P\Delta SP^T=\mu\left(\psi'\left(\frac{XS}{\mu}\right)\right)^{-1}\left(\psi(E)-\psi\left(\frac{XS}{\mu}\right)\right)S^{-1}.
\end{aligned} \tag{13}$$

Furthermore, we define $D=P^{\frac12}$. The matrix $D$ can be used to scale $X$ and $S$ to the same matrix $V$ because

$$V:=\frac{1}{\sqrt{\mu}}D^{-1}XD^{-1}=\frac{1}{\sqrt{\mu}}DSD. \tag{14}$$

Note that the matrices $D$ and $V$ are symmetric and positive definite. Furthermore, we have

$$V^2=\left(\frac{1}{\sqrt{\mu}}D^{-1}XD^{-1}\right)\left(\frac{1}{\sqrt{\mu}}DSD\right)=D^{-1}\frac{XS}{\mu}D.$$

From Definition 2.4, we obtain

$$\psi\left(\frac{XS}{\mu}\right)=D\psi\left(V^2\right)D^{-1}\quad\text{and}\quad\psi'\left(\frac{XS}{\mu}\right)=D\psi'\left(V^2\right)D^{-1}. \tag{15}$$

Let us further define

$$\bar A_i:=\frac{1}{\sqrt{\mu}}DA_iD,\quad i=1,2,\ldots,m;\qquad D_X:=\frac{1}{\sqrt{\mu}}D^{-1}\Delta XD^{-1};\qquad D_S:=\frac{1}{\sqrt{\mu}}D\Delta SD. \tag{16}$$

Then it follows from (13) that the scaled NT search directions $(D_X,\Delta y,D_S)$ are defined by the following system

$$\begin{aligned}
&\bar A_i\bullet D_X=0,\quad i=1,2,\ldots,m,\\
&\sum_{i=1}^{n}\Delta y_i\bar A_i+D_S=0,\\
&D_X+D_S=P_V,
\end{aligned} \tag{17}$$

where

$$P_V=\sqrt{\mu}D^{-1}\left(D\psi'\left(V^2\right)D^{-1}\right)^{-1}\left(\psi(E)-D\psi\left(V^2\right)D^{-1}\right)S^{-1}D^{-1}.$$

Recently, Peng et al. [16] introduced a class of search directions based on self-regular kernel functions and Bai et al. [4–6] also defined a class of new search directions by using the so-called eligible kernel functions. The general approach in this paper can be particularized in such a way as to obtain, the directions defined in [5,6,16] only by a constant multiplier. Such as

- $\psi(t)=t$ yields $P_V=V^{-1}-V$ which gives the classical search direction. The classical search direction has been studied by many researchers (e.g., [5,6,10,16]);
- $\psi(t)=t^2$ yields $P_V=\frac{1}{2}(V^{-3}-V)$, see [5,6,16];
- $\psi(t)=t^{\frac{q+1}{2}}, q\geqslant 0$ yields $P_V=\frac{2}{q+1}(V^{-q}-V)$, see [5,6,16].

For a detailed discussion we refer to [1,8]. Related discussions can be found in [11,15] for LO and linear complementarity problems (LCP).

Following [1,8], in this paper we restrict the analysis to the case where $\psi(t)=\sqrt{t}$, this yields

$$P_V=2(E-V). \tag{18}$$

As an anonymous referee pointed out, one interesting question is that whether the new search direction fits into the framework considered in [6]. The answer is positive! Similarly to the strategy in [20], we can consider the kernel function

$$\phi(t)=(t-1)^2. \tag{19}$$

For this one has $\phi'(t)=2(t-1)$, whence the induced barrier function $\Phi(V)=\sum_{i=1}^{n}\phi(\lambda_i(V))$ satisfies $\nabla\Phi(V)=2(V-E)$, and hence, due to (18), $P_V=-\nabla\Phi(V)$. It should be noted that except for the kernel function considered in [4], all kernel functions considered so far are coercive, i.e., have the properties

$$\lim_{t\to 0}\phi(t)=\infty \quad \text{and} \quad \lim_{t\to\infty}\phi(t)=\infty. \tag{20}$$

The present kernel function $\phi(t)$ as defined in (19) has the second property, but it fails to have the first property, because $\lim_{t\to 0}\phi(t)=1$. This situation has also appeared in [4].

Furthermore, we have

$$V^2+VP_V=V^2+2V(E-V)=E-(E-V)^2=E-\frac{P_V^2}{4}. \tag{21}$$

For the analysis of the algorithm, we define a norm-based proximity measure $\delta(X,S;\mu)$ as follows

$$\delta(V):=\delta(X,S;\mu):=\frac{\|P_V\|_F}{2}=\|E-V\|_F. \tag{22}$$

Due to the first two equations of the system (13), $D_X$ and $D_S$ are orthogonal. Thus

$$D_X\bullet D_S=D_S\bullet D_X=0. \tag{23}$$

One can easily verify that

$$\delta(V)=0 \quad\Leftrightarrow\quad V=E \quad\Leftrightarrow\quad D_X=D_S=0 \quad\Leftrightarrow\quad XS=\mu E. \tag{24}$$

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

The new search directions $D_X$ and $D_S$ are obtained by solving (17) with $P_V=2(E-V)$ so that $\Delta X$ and $\Delta S$ are computed via (16). If $(X,y,S)\neq(X(\mu),y(\mu),S(\mu))$ then $(\Delta X,\Delta y,\Delta S)$ is nonzero. One can construct a new full-Newton triple according to

$$X_+=X+\Delta X, \qquad y_+=y+\Delta y, \qquad S_+=S+\Delta S. \tag{25}$$

## 5. The generic interior-point algorithm

The generic form of the algorithm is shown in Fig. 1. We will prove that the algorithm is well defined in Section 6.

## 6. Analysis of the algorithm

In this section we will show that the algorithm can solve the SDO problem in polynomial time and prove the local quadratic convergence of the algorithm.

**Primal-Dual Path-Following Interior-Point Algorithm for SDO**

**Input:**
A threshold parameter $0 < \tau < 1$ (default $\tau = \frac{1}{2}$);
an accuracy parameter $\varepsilon > 0$;
a fixed barrier update parameter $0 < \theta < 1$ (default $\theta = \frac{1}{2\sqrt{n}}$);
a strictly feasible $(X^0, y^0, S^0)$ and $\mu^0 = 1$ such that $\delta(X^0, S^0; \mu^0) < \tau$.
**begin**
$X := X^0;\ y := y^0;\ S := S^0;\ \mu := \mu^0;$
**while** $n\mu \geqslant \varepsilon$ **do**
**begin**
solve system (17) and via (16) to obtain $(\Delta X, \Delta y, \Delta S)$;
update $(X, y, S) := (X, y, S) + (\Delta X, \Delta y, \Delta S)$;
$\mu := (1-\theta)\mu$;
**end**
**end**

**Fig. 1.** Algorithm.

For the analysis of the algorithm we introduce the notation

$$Q_V = D_X - D_S. \tag{26}$$

Thus we have

$$D_X = \frac{P_V + Q_V}{2}, \qquad D_S = \frac{P_V - Q_V}{2}, \qquad D_X D_S + D_S D_X = \frac{P_V^2 - Q_V^2}{2}. \tag{27}$$

Note that $D_X$ and $D_S$ are orthogonal, therefore $\|P_V\|_F = \|Q_V\|_F = 2\delta(V)$.

Let $0 \leqslant \alpha \leqslant 1$, we define

$$X(\alpha) = X + \alpha\Delta X, \qquad S(\alpha) = S + \alpha\Delta S. \tag{28}$$

We cite two useful lemmas in [10], which will be used in the proof of Lemma 6.3.

**Lemma 6.1.** *(See [10, Lemma 6.1].) Suppose that $X \succ 0$ and $S \succ 0$. If one has*

$$\det\big(X(\alpha)S(\alpha)\big) > 0, \quad \forall 0 \leqslant \alpha \leqslant \bar{\alpha},$$

*then $X(\bar{\alpha}) \succ 0$ and $S(\bar{\alpha}) \succ 0$.*

**Lemma 6.2.** *(See [10, Lemma 6.3].) Suppose that $Q \in \mathbf{S}^n_{++}$, and $M \in \mathbf{R}^{n\times n}$ be skew-symmetric, i.e. $M = -M^T$. One has $\det(Q+M) > 0$. Moreover, if $\lambda_i(Q+M) \in \mathbf{R}$ $(i = 1, 2, \dots, n)$, then*

$$0 < \lambda_{\min}(Q) \leqslant \lambda_{\min}(Q+M) \leqslant \lambda_{\max}(Q+M) \leqslant \lambda_{\max}(Q),$$

*which implies $(Q+M) \succ 0$.*

The following lemma shows the strict feasibility of the full NT-step under the condition $\delta(X, S; \mu) < 1$.

**Lemma 6.3.** *Let $\delta := \delta(X, S; \mu) < 1$, then the full NT-step is strictly feasible.*

**Proof.** By applying (28) and (16), we have

$$\begin{aligned} X(\alpha)S(\alpha) &= XS + \alpha(X\Delta S + \Delta X S) + \alpha^2 \Delta X \Delta S \\ &= \mu D\big(V^2 + \alpha(D_X V + V D_S) + \alpha^2 D_X D_S\big)D^{-1} \\ &\sim \mu\big(V^2 + \alpha(D_X V + V D_S) + \alpha^2 D_X D_S\big) \\ &= Q(\alpha) + M(\alpha), \end{aligned} \tag{29}$$

where

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

and

$$M(\alpha)=\mu\left(\frac{1}{2}\alpha(D_XV+VD_S-VD_X-D_SV)+\frac{1}{2}\alpha^2(D_XD_S-D_SD_X)\right).$$

One can easily verify that the matrix $M(\alpha)$ is skew-symmetric. Lemma 6.2 implies that $\det(X(\alpha)S(\alpha))>0$ if the matrix $Q(\alpha)\succ 0$. To this end, using (18), (21) and (27), we have

$$\begin{aligned}Q(\alpha)&=\mu\left(V^2+\frac{1}{2}\alpha(VP_V+P_VV)+\alpha^2\frac{P_V^2-Q_V^2}{4}\right)=\mu\left(V^2+\alpha VP_V+\alpha^2\frac{P_V^2-Q_V^2}{4}\right)\\&=\mu\left((1-\alpha)V^2+\alpha\big(V^2+VP_V\big)+\alpha^2\frac{P_V^2-Q_V^2}{4}\right)=\mu\left((1-\alpha)V^2+\alpha\left(E-\frac{P_V^2}{4}\right)+\alpha^2\frac{P_V^2-Q_V^2}{4}\right)\\&=\mu\left((1-\alpha)V^2+\alpha\left(E-(1-\alpha)\frac{P_V^2}{4}-\alpha\frac{Q_V^2}{4}\right)\right).\end{aligned}$$

Furthermore, since $0\leqslant\alpha\leqslant 1$, we have

$$\left\|(1-\alpha)\frac{P_V^2}{4}-\alpha\frac{Q_V^2}{4}\right\|_F\leqslant(1-\alpha)\left\|\frac{P_V^2}{4}\right\|_F+\alpha\left\|\frac{Q_V^2}{4}\right\|_F\leqslant(1-\alpha)\frac{\|P_V\|_F^2}{4}+\alpha\frac{\|Q_V\|_F^2}{4}=\delta^2<1.$$

Since the set of the positive matrices is cone, we can conclude that

$$\left((1-\alpha)V^2+\alpha\left(E-(1-\alpha)\frac{P_V^2}{4}-\alpha\frac{Q_V^2}{4}\right)\right)\succ 0,$$

i.e., $Q(\alpha)\succ 0$. Thus $\det(X(\alpha)S(\alpha))>0$. In addition, since $X(0)=X\succ 0$ and $S(0)=S\succ 0$, Lemma 6.1 implies that $X(1)=X\succ 0$ and $S(1)=S\succ 0$ for $\bar\alpha=1$. This complete the proof of the lemma. □

In the next lemma, we proceed to prove the local quadratic convergence of full NT-step to the target point $(X(\mu),y(\mu),S(\mu))$.

**Lemma 6.4.** *Let $\delta=\delta(X,S;\mu)<1$, then*

$$\delta(X_+,S_+;\mu)\leqslant\frac{\delta^2}{1+\sqrt{1-\delta^2}}.$$

*Thus $\delta(X_+,S_+;\mu)\leqslant\delta^2$, which shows the quadratical convergence of the algorithm.*

**Proof.** From (29) in the proof of Lemma 6.3, letting $\alpha=1$, we derive that

$$\frac{X_+S_+}{\mu}\sim E-\frac{Q_V^2}{4}+M,$$

where

$$M=\frac{1}{2}(D_XV+VD_S-VD_X-D_SV+D_XD_S-D_SD_X)$$

is a skew-symmetric matrix. Furthermore, $E-\frac{Q_V^2}{4}\succ 0$. Therefore

$$V_+^2\sim\frac{X_+S_+}{\mu}\sim\left(E-\frac{Q_V^2}{4}+M\right). \tag{30}$$

Since $M$ a skew-symmetric matrix, using Lemma 6.2, we have

$$\lambda_{\min}\big(V_+^2\big)=\lambda_{\min}\left(E-\frac{Q_V^2}{4}+M\right)\geqslant\lambda_{\min}\left(E-\frac{Q_V^2}{4}\right).$$

Thus

$$\lambda_{\min}\big(V_+^2\big)\geqslant\lambda_{\min}\left(E-\frac{Q_V^2}{4}\right)\geqslant 1-\lambda_{\max}\left(\frac{Q_V^2}{4}\right)\geqslant 1-\left\|\frac{Q_V^2}{4}\right\|_F\geqslant 1-\frac{\|Q_V\|_F^2}{4}=1-\delta^2.$$

This implies that

$$\lambda_{\min}(V_+)\geqslant\sqrt{1-\delta^2}. \tag{31}$$

On the other hand, using (30) and Lemma 2.6 (with $t=1$), we have

$$\begin{aligned}\delta(X_+, S_+; \mu) &= \|E - V_+\|_F = \left\|(E - V_+)(E + V_+)(E + V_+)^{-1}\right\|_F = \left\|\left(E - V_+^2\right)(E + V_+)^{-1}\right\|_F \\ &\leqslant \frac{1}{1 + \lambda_{\min}(V_+)}\left\|E - V_+^2\right\|_F = \frac{1}{1 + \lambda_{\min}(V_+)}\sqrt{\mathbf{Tr}\left(\frac{Q_V^2}{4} - M\right)^2}.\end{aligned} \tag{32}$$

Since $M$ is a skew-symmetric matrix, after some elementary reductions, we have

$$\begin{aligned}\delta(X_+, S_+; \mu) &\leqslant \frac{1}{1 + \lambda_{\min}(V_+)}\mathbf{Tr}\left(\frac{Q_V^2}{4}\right) \leqslant \frac{1}{1 + \sqrt{1 - \delta^2}}\left\|\frac{Q_V^2}{4}\right\|_F \\ &= \frac{1}{1 + \sqrt{1 - \delta^2}}\frac{\|Q_V\|_F^2}{4} = \frac{\delta^2}{1 + \sqrt{1 - \delta^2}}.\end{aligned}$$

This proves the lemma. □

The following lemma gives an upper bound of the duality gap after a full NT-step.

**Lemma 6.5.** *After a full NT-step, then*

$$X_+ \bullet S_+ \leqslant n\mu.$$

**Proof.** Since $E - \frac{Q_V^2}{4} \succcurlyeq 0$ and $M$ is a skew-symmetric matrix, using Property 2.2, we have

$$X_+ \bullet S_+ = \mu\,\mathbf{Tr}\left(V_+^2\right) = \mu\,\mathbf{Tr}\left(E - \frac{Q_V^2}{4} + M\right) = \mu\,\mathbf{Tr}\left(E - \frac{Q_V^2}{4}\right) \leqslant n\mu.$$

The proof is completed. □

In the following lemma, we investigate the effect on the proximity measure of a full NT-step followed by an update of the parameter $\mu$.

**Lemma 6.6.** *Let $\delta = \delta(X, S; \mu) < 1$ and $\mu_+ = (1 - \theta)\mu$, where $0 < \theta < 1$. Then*

$$\delta(X_+, S_+; \mu_+) \leqslant \frac{\theta\sqrt{n} + \delta^2}{1 - \theta + \sqrt{(1 - \theta)(1 - \delta^2)}}.$$

*Furthermore, if $\delta \leqslant \frac{1}{2}$, $\theta = \frac{1}{2\sqrt{n}}$ and $n \geqslant 4$, then we have*

$$\delta(X_+, S_+; \mu_+) \leqslant \frac{1}{2}.$$

**Proof.** By applying (30), (31) and Lemma 2.6 (with $t = \sqrt{1 - \theta}$), we have

$$\begin{aligned}\delta(X_+, S_+; \mu_+) &= \left\|E - \sqrt{\frac{X_+S_+}{\mu_+}}\right\|_F = \frac{1}{\sqrt{1 - \theta}}\|\sqrt{1 - \theta}E - V_+\|_F \\ &= \frac{1}{\sqrt{1 - \theta}}\left\|(\sqrt{1 - \theta}E - V_+)(\sqrt{1 - \theta}E + V_+)(\sqrt{1 - \theta}E + V_+)^{-1}\right\|_F \\ &= \frac{1}{\sqrt{1 - \theta}}\left\|\left((1 - \theta)E - V_+^2\right)(\sqrt{1 - \theta}E + V_+)^{-1}\right\|_F \\ &\leqslant \frac{1}{\sqrt{1 - \theta}(\sqrt{1 - \theta} + \lambda_{\min}(V_+))}\left\|(1 - \theta)E - V_+^2\right\|_F \\ &= \frac{1}{\sqrt{1 - \theta}(\sqrt{1 - \theta} + \lambda_{\min}(V_+))}\sqrt{\mathbf{Tr}\left(-\theta E + \frac{Q_V^2}{4} - M\right)^2}.\end{aligned}$$

Since $M$ is a skew-symmetric matrix, after some elementary reductions, we have

$$\begin{aligned}\delta(X_+, S_+; \mu_+) &\leqslant \frac{1}{\sqrt{1 - \theta}(\sqrt{1 - \theta} + \lambda_{\min}(V_+))}\mathbf{Tr}\left(-\theta E + \frac{Q_V^2}{4}\right) \\ &\leqslant \frac{1}{1 - \theta + \sqrt{(1 - \theta)(1 - \delta^2)}}\left\|-\theta E + \frac{Q_V^2}{4}\right\|_F\end{aligned}$$

$$\leqslant \frac{1}{1-\theta+\sqrt{(1-\theta)(1-\delta^2)}}\left(\theta\sqrt{n}+\left\|\frac{Q_V^2}{4}\right\|_F\right)$$
$$=\frac{1}{1-\theta+\sqrt{(1-\theta)(1-\delta^2)}}\left(\theta\sqrt{n}+\delta^2\right).$$

This completes the proof of the first part of the lemma. On the other hand, since $n \geqslant 4$, we have

$$1-\theta = 1-\frac{1}{2\sqrt{n}} \geqslant \frac{3}{4}.$$

Thus from $\delta \leqslant \frac{1}{2}$, we can conclude that

$$\delta(X_+, S_+; \mu_+) \leqslant \frac{1}{2}.$$

This proves the lemma. □

**Remark 6.7.** At the start of the algorithm we choose a strictly feasible pair $(X^0, y^0, S^0)$ and $\mu^0 = \frac{X^0 \bullet S^0}{n}$ such that $\delta(X^0, S^0; \mu^0) < \tau = \frac{1}{2}$. From Lemmas 6.3 and 6.5, we have $X_+ \succ 0$, $S_+ \succ 0$, and $X_+ \bullet S_+ \leqslant n\mu$. After the update of the barrier parameter to $\mu_+ = (1-\theta)\mu$, with $\theta = \frac{1}{2\sqrt{n}}$, we have, by Lemma 6.6,

$$\delta(X_+, S_+; \mu_+) \leqslant \frac{1}{2}.$$

Thus, after each iteration of the algorithm, the new iteration is strictly feasible, and the properties

$$X_+ \bullet S_+ \leqslant n\mu, \delta(X_+, S_+; \mu_+) \leqslant \frac{1}{2}$$

are maintained. Hence the algorithm is well defined.

The following lemma gives an upper bound for the total number of iterations produced by our algorithm.

**Lemma 6.8.** *Suppose that $X^0$ and $S^0$ are strictly feasible, $\mu^0 = \frac{X^0 \bullet S^0}{n}$ and $\delta(X^0, S^0; \mu^0) \leqslant \frac{1}{2}$. Moreover, let $X^k$ and $S^k$ be the matrices obtained after k iterations. Then the inequality $X^k \bullet S^k \leqslant \varepsilon$ is satisfied for*

$$k \geqslant \frac{1}{\theta}\log\frac{X^0 \bullet S^0}{\varepsilon}.$$

**Proof.** Lemma 6.5 implies that

$$X^k \bullet S^k \leqslant n\mu^k = n(1-\theta)^k\mu^0 = (1-\theta)^k X^0 \bullet S^0.$$

Then the inequality $X^k \bullet S^k \leqslant \varepsilon$ holds if

$$(1-\theta)^k X^0 \bullet S^0 \leqslant \varepsilon.$$

Taking logarithms, we obtain

$$k\log(1-\theta) + \log\left(X^0 \bullet S^0\right) \leqslant \log\varepsilon,$$

and using $-\log(1-\theta) \geqslant \theta$ we observe that the above inequality holds if

$$k\theta \geqslant \log\left(X^0 \bullet S^0\right) - \log\varepsilon = \log\frac{X^0 \bullet S^0}{\varepsilon}.$$

This implies the lemma. □

**Theorem 6.9.** *Let $\theta = \frac{1}{2\sqrt{n}}$, then the algorithm requires at most*

$$O\left(\sqrt{n}\log\frac{X^0 \bullet S^0}{\varepsilon}\right)$$

*iterations. The output is a primal-dual pair $(X, S)$ satisfying $X \bullet S \leqslant \varepsilon$.*

**Proof.** Let $\theta = \frac{1}{2\sqrt{n}}$, by using Lemma 6.8, the proof is straightforward. □

**Corollary 6.10.** *If one takes $X^0 = S^0 = E$, the iteration bound becomes*

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

*which is the currently best known iteration bound for the algorithm with small-update method.*

## 7. Conclusions and remarks

We have extended a primal-dual path-following interior-point algorithm for LO to SDO with full NT-step and derived the currently best known iteration bound for the algorithm with small-update method, namely, $O(\sqrt{n}\log\frac{n}{\varepsilon})$, which is the same iteration bound as in the LO case. Moreover, the resulting analysis is relatively simple and similar to the LO analogue in [8].

Some interesting topics remain for further research. Firstly, the search directions used in this paper are all based on the NT-symmetrization scheme. It may be possible to design similar algorithms using other symmetrization schemes and to obtain polynomial-time iteration bounds. Secondly, the extensions to SOCO and the general convex optimization deserve to be investigated. Furthermore, numerical test is an interesting topic for investigating the behavior of the algorithm so as to be compared with other approaches.

## Acknowledgments

The authors would like to thank the referees for their useful suggestions which improve the presentation of this paper.

## References

[1] M. Achache, A new primal-dual path-following method for convex quadratic programming, Appl. Comput. Math. 25 (1) (2006) 97–110.
[2] F. Alizadeh, Interior-point methods in semidefinite programming with applications to combinatorial optimization, SIAM J. Optim. 5 (1995) 13–51.
[3] F. Alizadeh, J.P.A. Heaberly, M.L. Overton, Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results, SIAM J. Optim. 8 (3) (1998) 746–768.
[4] Y.Q. Bai, M. El Ghami, C. Roos, A new efficient large-update primal-dual interior-point method based on a finite barrier, SIAM J. Optim. 13 (3) (2003) 766–782.
[5] Y.Q. Bai, G. Lesaja, C. Roos, G.Q. Wang, M. El Ghami, A class of large-update and small-update primal-dual interior-point algorithms for linear optimization, J. Optim. Theory Appl. 138 (3) (2008) 341–359.
[6] Y.Q. Bai, C. Roos, M. El Ghami, A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization, SIAM J. Optim. 15 (1) (2004) 101–128.
[7] Y.Q. Bai, G.Q. Wang, A new primal-dual interior-point algorithm for second-order cone optimization based on kernel function, Acta Math. Sin. (Engl. Ser.) 23 (11) (2007) 2027–2042.
[8] Z. Darvay, New interior-point algorithms in linear optimization, Adv. Model. Optim. 5 (1) (2003) 51–92.
[9] R.A. Horn, R.J. Charles, Topics in Matrix Analysis, Cambridge University Press, UK, 1991.
[10] E. de Klerk, Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
[11] X.S. Li, S.Y. He, S.Y. Pan, H.W. Zhang, A path-following algorithm for linear complementarity problems based on algebraically equivalent transformation, J. Dalian Univ. Technol. 44 (3) (2004) 337–341.
[12] Z.Q. Luo, J. Sturm, S.Z. Zhang, Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming, SIAM J. Optim. 8 (1998) 59–81.
[13] Y.E. Nesterov, M.J. Todd, Self-scaled barriers and interior-point methods for convex programming, Math. Oper. Res. 22 (1) (1997) 1–42.
[14] Y.E. Nesterov, M.J. Todd, Primal-dual interior-point methods for self-scaled cones, SIAM J. Optim. 8 (2) (1998) 324–364.
[15] S.Y. Pan, X.S. Li, S.Y. He, An infeasible primal-dual interior-point algorithm for linear programs based on logarithmic equivalent transformation, J. Math. Anal. Appl. 314 (2006) 644–660.
[16] J. Peng, C. Roos, T. Terlaky, Self-regular functions and new search directions for linear and semidefinite optimization, Math. Program. 93 (2002) 129–171.
[17] M.J. Todd, A study of search directions in primal-dual interior-point methods for semidefinite programming, Optim. Methods Softw. 11 (1999) 1–46.
[18] M.J. Todd, K.C. Toh, R.H. Tütüncü, On the Nesterov–Todd direction in semidefinite programming, SIAM J. Optim. 8 (3) (1998) 769–796.
[19] L. Vandenberghe, S. Boyd, Semidefinite programming, SIAM Rev. 38 (1996) 49–95.
[20] G.Q. Wang, Y.Q. Bai, C. Roos, Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function, J. Math. Model. Algorithms 4 (2005) 409–433.
[21] H. Wolkowicz, R. Saigal, L. Vandenberghe, Handbook of Semidefinite Programming, Theory, Algorithms, and Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
[22] Y. Zhang, On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming, SIAM J. Optim. 8 (1998) 365–386.