ON THE EXTENSION OF AN ARC-SEARCH INTERIOR-POINT ALGORITHM FOR SEMIDEFINITE OPTIMIZATION

Behrouz Kheirfam*1, Morteza Moslemi1

1Department of Applied Mathematics
Azarbaijan Shahid Madani University
Tabriz, I. R. Iran

(Communicated by Tatiana Tchemisova)

Abstract. This paper concerns an extension of the arc-search strategy that was proposed by Yang [26] for linear optimization to semidefinite optimization case. Based on the Nesterov-Todd direction as Newton search direction it is shown that the complexity bound of the proposed algorithm is of the same order as that of the corresponding algorithm for linear optimization. Some preliminary numerical results indicate that our primal-dual arc-search path-following method is promising for solving the semidefinite optimization problems.

1. Introduction. Semidefinite optimization (SDO) problem is a generalization of linear optimization (LO) problem over an affine subspace of symmetric matrices. SDO has wide applications in diverse areas, such as continuous and combinatorial optimization [1, 22], system and control theory [4] or mechanical and electrical engineering [21].

Interior-point methods (IPMs) are now among the most effective methods for solving SDO problems. Due to their polynomial complexity and practical efficiency, many IPMs for LO problems are successfully extended to the SDO case [9, 11, 19, 29]. Pioneering researches for this extension were independently done by Alizadeh [2] and Nesterov and Nemirovskii [16] in which the SDO problems were solved by the primal-dual interior-point algorithms with polynomial time complexity. Kojima et al. [9] and Potra and Sheng [19] independently discussed a generalization to SDO of the Mizuno-Todd-Ye (MTY) predictor-corrector algorithm [12] for infeasible starting points and derived the best known iteration bound under some mild assumptions. The convergence analysis of IPMs for SDO problems is more complicated than for LO case. A large part of the theoretical difficulty is due to the issue of maintaining symmetry in the linearized complementarity condition. A strategy to overcome this problem was proposed by Alizadeh et al. [3], but one of the most common strategies was developed by Monteiro [14]. Such formulation reveals some deep similarity between LO and SDO cases and allows insightful complexity analysis for SDO problems (see [29]). It turned out that the resulting directions by this approach could be seen as two special cases of the class of directions introduced earlier by Kojima et al. [8]. Nesterov and Todd [17, 18] introduced the
so-called Nesterov-Todd (NT) direction in their attempt to generalize primal-dual IPM beyond SDO. In [29], based on Monteiro’s idea, Zhang generalized all these aforementioned approaches to a unified scheme parameterized by a nonsingular scaling matrix. This family of search directions is referred to as the Monteiro-Zhang (MZ) family of search directions.

Yang [25, 26] proposed the arc-search approach for LO and convex quadratic optimization problems, respectively, which searches optimizers along an ellipse that approximates the entire central path. Based on the arc-search strategy, the author proved that the complexity bound can be improved for a higher-order feasible IPM. Yang et al. [27] used the same idea and proposed an arc-search infeasible IPM with a complexity bound $O(n^{5/4} \log \varepsilon^{-1})$ for LO problems, where $n$ is the larger dimension of a standard LO and $\varepsilon$ is the required precision. Later, Yang et al. [28] extended the algorithm to symmetric optimization (SO). Very recently, Kheirfam [7] proposed an arc-search infeasible-interior-point algorithm for SO in a wide neighborhood $N^\infty_\varepsilon$ of the central path and derived the complexity bound $O(r^{3/2} \log \varepsilon^{-1})$, where $r$ is the rank of the associated Euclidean Jordan algebra.

The purpose of this paper is to establish iteration-complexity bound for an extension of the arc-search strategy introduced by Yang in [26] for LO problems to the SDO case. Based on the NT direction, our derived complexity bound for the algorithm is $O(\sqrt{n} \log \mu_0^{-1})$, which is analogous to the LO case. Numerical results show that our algorithm performs significantly better than the MTY predictor-corrector algorithm.

This paper is structured as follows. In Section 2, we first introduce the SDO problems and discuss the arc-search approach for SDO problems. Then we present a primal-dual arc-search algorithm. In Section 3, we provide some technical lemmas, which will be needed in Section 4. Section 4 is devoted to complexity analysis of the proposed algorithm. Numerical results are given in Section 5. Finally, we end the paper with some concluding remarks in Section 6.

The following notations are used throughout the paper. Given $m, n \in \mathbb{N}$, $\mathbb{R}^n$ and $\mathbb{R}^{m \times n}$ denote the $n$-dimensional Euclidean space and the space of all real $m \times n$ matrices, respectively. The set of all $n \times n$ symmetric matrices is denoted by $\mathbb{S}^n$. $\mathbb{S}_+^n$ and $\mathbb{S}_{++}^n$ denote the set of all matrices in $\mathbb{S}^n$ which are positive semidefinite and positive definite respectively. For $Q \in \mathbb{S}^n$, $Q \succeq 0 (Q \succ 0)$ means that $Q$ is positive semidefinite (positive definite). For a matrix $Q \in \mathbb{S}^n$, with all real eigenvalues, we denote its eigenvalues by $\lambda_i(Q), i = 1, \ldots, n$ and its smallest and largest eigenvalues by $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$, respectively. $\|Q\|_F$ denotes the Frobenius norm of $Q \in \mathbb{R}^{n \times n}$, i.e., $\|Q\|_F = \sqrt{\text{Tr}(Q^T Q)}$. The Kronecker product of any two matrices of arbitrary size $A$ and $B$ is denoted by $A \otimes B$. Finally, for $A \in \mathbb{R}^{m \times n}$, $\text{vec}(A)$ denotes the $mn$-vector obtained by stacking the columns of $A$ one by one from the first to the last column.

2. SDO problem and search direction. In this paper, we deal with the primal SDO problem in the form

$$\min \left\{ \text{Tr}(CX) : \text{Tr}(A_i X) = b_i, \; i = 1, \ldots, m, \; X \succeq 0 \right\}, \tag{P}$$

along with its dual problem, which can be given as follows

$$\max \left\{ b^T y : \sum_{i=1}^m y_i A_i + S = C, \; S \succeq 0 \right\}, \tag{D}$$
where $C, X, S \in \mathbb{S}^n$, $b, y \in \mathbb{R}^m$ and $A_i \in \mathbb{S}^n$, $i = 1, \ldots, m$, are, without loss of generality, linearly independent. Throughout the paper, we assume that both problems (P) and (D) satisfy the interior-point condition (IPC), that is,

$$\mathcal{F}^0 := \{(X, y, S) : \text{Tr}(A_i X) = b_i, i = 1, \ldots, m, \text{Tr}(X) = 0, \sum_{i=1}^{m} y_i A_i + S = C, S > 0\} \neq \emptyset.$$  

It is well known [5] that under the assumptions that $\mathcal{F}^0 \neq \emptyset$ and the matrices $A_i, i = 1, \ldots, m$, are linearly independent, the set of all primal and dual optimal solutions consists of all the solutions $(X, y, S) \in \mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$ of the following optimality system:

$$\text{Tr}(A_i X) = b_i, \quad i = 1, \ldots, m, \quad X \succeq 0,$$

$$\sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0,$$

$$XS = 0. \quad (1)$$

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

$$\text{Tr}(A_i X) = b_i, \quad i = 1, \ldots, m, \quad X \succeq 0,$$

$$\sum_{i=1}^{m} y_i A_i + S = C, \quad S \succeq 0,$$

$$XS = \mu I. \quad (2)$$

It is proven in [8, 16] that the system (2) has a unique solution $(X(\mu), y(\mu), S(\mu))$ for any $\mu > 0$ as a $\mu$-center of (P) and (D) under assumption that $\mathcal{F}^0 \neq \emptyset$ and $A_i, i = 1, \ldots, m$, are linearly independent. The set of $\mu$-centers with $\mu > 0$ is said to be the central path of (P) and (D), i.e., an arc that is parameterized as a function of $\mu$ and is denoted as

$$C = \{(X(\mu), y(\mu), S(\mu)) : \mu > 0\}.$$ 

As $\mu \to 0$, the limit $(X(\mu), y(\mu), S(\mu))$ is a primal-dual optimal solution of the corresponding SDO problem.

We consider a set of neighborhoods of the central path as

$$\mathcal{N}_F(\theta) := \{(X, y, S) \in \mathcal{F}^0 : \|X^2 S X^2 - \mu I\|_F \leq \theta \mu\},$$

where $\theta \in (0, 1)$ and $\mu := \frac{\text{Tr}(XS)}{n}$ is associated with the actual duality gap. As we reduce the duality gap to zero, $\mathcal{N}_F(\theta)$ will be a neighborhood of the central path that approaches the optimizer(s) of the SDO problem. Thus, all points inside $\mathcal{N}_F(\theta)$ will approach the optimizer(s) of the SDO problem.

A promising way to obtain a search direction is to apply Newton’s method to system (2). Observe that for $X, S \in \mathbb{S}^n$, the product $XS$ is generally not in $\mathbb{S}^n$. Hence, the left-hand side of (2) is a map from $\mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$ to $\mathbb{S}^n \times \mathbb{S}^n \times \mathbb{R}^n \times \mathbb{n}$. Thus, the system (2) is not a square system when $X$ and $S$ are restricted to $\mathbb{S}^n$, which it is needed for applying Newton-like methods. To remedy, we use the so-called similar symmetrization operator $H_P : \mathbb{R}^{n \times n} \to \mathbb{S}^n$ introduced by Zhang [29]
defined as follows

\[ H_P(M) = \frac{1}{2} (PMP^{-1} + (PMP^{-1})^T), \quad \forall M \in \mathbb{R}^{n \times n}, \]  

(3)

where \( P \in \mathbb{R}^{n \times n} \) is some nonsingular matrix. Zhang [29] also observed that

\[ H_P(M) = \mu I \iff M = \mu I, \]

for any nonsingular matrix \( P \), any matrix \( M \) similar to a (symmetric) positive definite matrix (e.g. \( M = XS \) with \( X, S \in S_+^n \)) and any \( \mu \in \mathbb{R} \). Consequently, for any given nonsingular matrix \( P \), system (2) is equivalent to the following square system from \( \mathbb{R}^m \times S^n \times S^n \) to itself:

\[ \text{Tr}(A_iX) = b_i, \quad i = 1, \ldots, m, \]

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

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

(4)

A Newton-like method applied to system (4) leads to the following linear system for direction \((\Delta X, \Delta y, \Delta S)\) \(\in S^n \times \mathbb{R}^m \times S^n\):

\[ \text{Tr}(A_i \Delta X) = 0, \quad i = 1, \ldots, m, \]

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

\[ H_P(X \Delta S + \Delta XS) = \mu I - H_P(XS). \]  

(5)

Todd et al. [20] showed that the system (5) has a unique solution for any \((X, y, S)\) \(\in S^n_+ \times \mathbb{R}^m \times S^n_+\) and the scaling matrix \( P \) satisfies \( PXSP^{-1} \in S^n \). The choice of \( P = W \frac{1}{2} \) in (5), where

\[ W := X^{-\frac{1}{2}} (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{\frac{1}{2}} X^{-\frac{1}{2}} = S^{\frac{1}{2}} (S^{\frac{1}{2}} X S^{\frac{1}{2}})^{-\frac{1}{2}} S^{\frac{1}{2}} \]  

(6)

leads to the NT-direction [18]. For the NT scaling matrix \( P = W \frac{1}{2} \), we have \( PXP = P^{-1} SP^{-1} \) and hence \( PXSP^{-1} \in S^n \).

Recently, Yang [24, 25, 26] proposed an interesting approach based on approximating the central path by an ellipse and presented some polynomial feasible interior-point algorithms for LO problems that search optimizers along an ellipsoidal estimation of the central path. Yang [26] showed that arc-search along ellipse may be a better method than line-search methods because the algorithm is proved to be polynomial with a better bound than the bounds of all existing higher-order algorithms. Based on the NT direction as Newton search direction, we extend the arc-search strategy proposed by Yang [26] for LO problems to SDO case. For this purpose, we define an ellipse \( \mathcal{E} \in S^{2n+m} \) to estimate the central path \( C \) as follows

\[ \mathcal{E} = \left\{ Z(\alpha) := \begin{bmatrix} X(\alpha) & 0 & 0 \\ 0 & y(\alpha) & 0 \\ 0 & 0 & S(\alpha) \end{bmatrix} : Z(\alpha) = \cos(\alpha)B_1 + \sin(\alpha)B_2 + B_3 \right\}, \]  

(7)

where \( B_1 \) and \( B_2 \) are the axes of the ellipse, and \( B_3 \) is the center of the ellipse. Let \((\dot{X}, \dot{y}, \dot{S})\) and \((\ddot{X}, \ddot{y}, \ddot{S})\) respectively denote the first and second derivatives of \((X(\mu), y(\mu), S(\mu))\) and \( Z = Z(\alpha_0) \in \mathcal{E} \) be close to or on the central path. We
proceed to determine the matrices $B_1, B_2$ and $B_3$ such that the first and second derivatives of $(X(\mu), y(\mu), S(\mu))$ satisfy

\[
\text{Tr}(A_i \dot{X}) = 0, \ i = 1, \ldots, m, \\
\sum_{i=1}^{m} \dot{y}_i A_i + \dot{S} = 0, \\
H_P(X \dot{S} + \dot{X} S) = H_P(X S),
\]

and

\[
\text{Tr}(A_i \ddot{X}) = 0, \ i = 1, \ldots, m, \\
\sum_{i=1}^{m} \ddot{y}_i A_i + \ddot{S} = 0, \\
H_P(X \ddot{S} + \ddot{X} S) = H_P(-2\dot{X} \dot{S}).
\]

For the scaling matrix $P$, we define

\[
\hat{X} := PXP, \ \hat{S} := P^{-1}SP^{-1}, \ \hat{A}_i := P^{-1}A_iP^{-1}, \ i = 1, \ldots, m.
\]

Now, we scale systems (8) and (9) and obtain the following scaled systems

\[
\text{Tr}(\hat{A}_i \hat{X}) = 0, \ i = 1, \ldots, m, \\
\sum_{i=1}^{m} \dot{y}_i \hat{A}_i + \hat{S} = 0, \\
H(\hat{X} \dot{S} + \dot{X} \hat{S}) = H(\hat{X} \hat{S}),
\]

and

\[
\text{Tr}(\hat{A}_i \ddot{X}) = 0, \ i = 1, \ldots, m, \\
\sum_{i=1}^{m} \ddot{y}_i A_i + \ddot{S} = 0, \\
H(\hat{X} \ddot{S} + \ddot{X} \hat{S}) = -2H(\dot{X} \dot{S}),
\]

where $H(\cdot) = H_I(\cdot), \ X := PXP, \ S := P^{-1}SP^{-1}, \ \dot{X} := P\dot{X}P$ and $\ddot{S} := P^{-1}\ddot{S}P^{-1}$. Furthermore, similar to [10, 15], in terms of the Kronecker product, the third equations in (11) and (12) can be expressed as

\[
\dot{E} \text{vec}(\hat{X}) + \dot{F} \text{vec}(\hat{S}) = \text{vec}(H(\hat{X} \hat{S})), \quad (13)
\]

\[
\dot{E} \text{vec}(\hat{X}) + \dot{F} \text{vec}(\hat{S}) = -2\text{vec}(H(\hat{X} \hat{S})), \quad (14)
\]

where

\[
\dot{E} \equiv \frac{1}{2}(\dot{S} \otimes I + I \otimes \dot{S}), \quad \dot{F} \equiv \frac{1}{2}(\dot{X} \otimes I + I \otimes \dot{X}).
\]
By taking the first and second derivatives of $E$ and using (11) and (12), after some calculations, we obtain

$$
B_1 = \begin{bmatrix}
\sin(\alpha)\dot{X} + \cos(\alpha)\ddot{X} & 0 & 0 \\
0 & \sin(\alpha)\dot{y} + \cos(\alpha)\ddot{y} & 0 \\
0 & 0 & \sin(\alpha)\dot{S} + \cos(\alpha)\ddot{S}
\end{bmatrix},
$$

$$
B_2 = \begin{bmatrix}
\cos(\alpha)\dot{X} - \sin(\alpha)\ddot{X} & 0 & 0 \\
0 & \cos(\alpha)\dot{y} - \sin(\alpha)\ddot{y} & 0 \\
0 & 0 & \cos(\alpha)\dot{S} - \sin(\alpha)\ddot{S}
\end{bmatrix},
$$

$$
B_3 = \begin{bmatrix}
\dot{X} + \dot{X} & 0 & 0 \\
0 & \ddot{y} + \ddot{y} & 0 \\
0 & 0 & \dot{S} + \dot{S}
\end{bmatrix}.
$$

Let $\alpha \in (0, \pi/2]$. The following lemma denotes the mathematical formulation of the ellipsoidal approximation of the central path $C$, which substituting $B_1, B_2$ and $B_3$ into (7), the lemma follows.

**Lemma 2.1.** Let $(\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha))$ be the defined ellipse in (7) which passes a point $(\hat{X}, y, \hat{S})$. Moreover, assume that the first and second derivatives $(\hat{X}, \hat{y}, \hat{S})$ and $(\ddot{X}, \ddot{y}, \ddot{S})$ satisfy (11) and (12). Then, after searching along the ellipse $\mathcal{E}$, the new generated point $(\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha))$ is given by

$$
\dot{X}(\alpha) = \dot{X} - \sin(\alpha)\dot{X} + (1 - \cos(\alpha))\ddot{X},
$$

$$
y(\alpha) = y - \sin(\alpha)\dot{y} + (1 - \cos(\alpha))\ddot{y},
$$

$$
\dot{S}(\alpha) = \dot{S} - \sin(\alpha)\dot{S} + (1 - \cos(\alpha))\ddot{S}.
$$

Due to Lemma 3.1, the ellipsoidal approximation of the central path $C$ is defined as follows:

$$
\mathcal{E} := \left\{ \hat{Z}(\alpha) : \hat{Z}(\alpha) \text{ satisfies (15), (16) and (17) with } \alpha \in (0, \pi/2] \right\}
$$

Moreover, in terms of the Kronecker product and using (10), the system (5) can be rewritten as follows:

$$
\hat{A}\text{vec}(\Delta \hat{X}) = 0, \ i = 1, \ldots, m,
$$

$$
\hat{A}^T\Delta y + \text{vec}(\Delta \hat{S}) = 0,
$$

$$
\hat{E}\text{vec}(\Delta \hat{X}) + \hat{F}\text{vec}(\Delta \hat{S}) = \text{vec}(\mu I - H(\hat{X}, \hat{S})),
$$

where $\hat{A}^T = [\text{vec}(\hat{A}_1), \ldots, \text{vec}(\hat{A}_m)], \Delta y = [\Delta y_1, \ldots, \Delta y_m]^T, \Delta \hat{X} = P\Delta XP$ and $\Delta \hat{S} = P^{-1}\Delta SP^{-1}$.

Now, we are ready to present a generic frame for our primal-dual arc-search algorithm for SDO problem.
Algorithm 1: Primal–dual arc–search interior–point algorithm for SDO

Let neighborhood parameter $\theta = 0.25$, required precision $\varepsilon > 0$, and the initial solution $(X^0, y^0, S^0) \in \mathcal{N}_F(\theta)$ be given.

Set $k = 0$, and $\mu^0 = \frac{1}{n} \text{Tr}(X^0 S^0)$.

Repeat until $\mu^k \leq \varepsilon$, do

Step 1: Compute the NT scaling matrix $P = (W^k)^{\frac{1}{2}}$ by (6).

Step 2: Solve $(\hat{X}, \hat{y}, \hat{S})$ and $(\check{X}, \check{y}, \check{S})$ based on (11) and (12).

Step 3: Find the smallest positive $\sin(\alpha)$ that satisfies the quartic polynomial in terms of $\sin(\alpha)$

$$q(\alpha) = \left(\|\hat{X}\|_F^2 + \|\check{X}\|_F^2\right) \sin^4(\alpha) + \left(\|\hat{X}\|_F^2 + \|\check{X}\|_F^2\right) \sin^2(\alpha) + \theta \mu^k \sin(\alpha) - \theta \mu^k = 0;$$

Step 4: Compute $(\hat{X}(\alpha), y(\alpha), S(\alpha))$ by (15), (16) and (17) and set $\mu(\alpha) = \frac{1}{n} \text{Tr}(\hat{X}(\alpha) S(\alpha))$.

Step 5: Calculate $(\Delta X, \Delta y, \Delta S)$ by solving

$$\begin{bmatrix} \hat{A} & 0 & 0 \\ 0 & \hat{A}^T & I \end{bmatrix} \begin{bmatrix} \text{vec}(\Delta X) \\ \text{vec}(\Delta y) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{vec}(\mu(\alpha) I - H(\hat{X}(\alpha) S(\alpha)))$$

where $\hat{E}(\alpha) = \frac{1}{2} (S(\alpha) \otimes I + I \otimes S(\alpha))$, $\check{F}(\alpha) = \frac{1}{2} (X(\alpha) \otimes I + I \otimes X(\alpha))$.

Step 6: Set $(X^{k+1}, y^{k+1}, S^{k+1}) = (X(\alpha), y(\alpha), S(\alpha)) + (\Delta X, \Delta y, \Delta S)$.

Step 7: Compute $(X^{k+1}, y^{k+1}, S^{k+1}) = (P^{-1} X^{k+1} P^{-1}, y^{k+1}, P S^{k+1} P)$.

Step 8: Set $\mu^{k+1} := \frac{1}{n} \text{Tr}(X^{k+1} S^{k+1})$, and $k := k + 1$.

Step 3 of the algorithm finds a step size $\sin(\alpha)$ such that $(X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}_F(\theta)$. Lemma 4.1 will show that the step size $\sin(\alpha)$ is determined by the smallest positive root of $q(\alpha)$. Since $q$ is a monotonic increasing function of $\alpha \in [0, \frac{\pi}{4}]$, for $(X^k, S^k) \succ 0$, (which will be shown in Lemma 4.2), $q(0) < 0$ and $q(\frac{\pi}{4}) \geq 0$. Therefore $q(\alpha)$ has exactly one positive root in $[0, \frac{\pi}{4}]$. Moreover, all the solutions are analytical and the computational cost is independent of the size of $A$ ($n$ and $m$) and is negligible [6]. We will show that the correction step brings the iterate from $\mathcal{N}_F(\theta)$ back to $\mathcal{N}_F(\theta)$, and will also estimate the step size.

3. Technical results. The following lemma will be frequently used in the analysis of Algorithm 1.

Lemma 3.1. Let $(\hat{X}, \hat{y}, \hat{S})$ and $(\check{X}, \check{y}, \check{S})$ be the solutions of (11) and (12), respectively. Then

(i) $\text{Tr}(\hat{X} \hat{S}) = \text{Tr}(\check{X} \check{S}) = \text{Tr}(\hat{X} \hat{S}) = \text{Tr}(\check{X} \check{S}) = 0$.

(ii) $\text{Tr}(\hat{A}_i \hat{X}(\alpha)) = \hat{b}_i$, $i = 1, \ldots, m$, and $\sum_{i=1}^m y_i(\alpha) \hat{A}_i + \check{S}(\alpha) = \check{C}$, where $\check{C} = P^{-1} C P^{-1}$.

Proof. Multiplying the second equations of systems (11) and (12) by $\hat{X}$ or $\check{X}$, taking the trace of the resulting equations, and using the first equations of (11) and (12), we obtain the results in part (i). To prove (ii), we have

$$\text{Tr}(\hat{A}_i \hat{X}(\alpha)) = \text{Tr}(\hat{A}_i \hat{X} - \sin(\alpha) \hat{X} + (1 - \cos(\alpha)) \hat{X})$$

$$= \text{Tr}(\hat{A}_i \hat{X}) - \sin(\alpha) \text{Tr}(\hat{A}_i \hat{X}) + (1 - \cos(\alpha)) \text{Tr}(\hat{A}_i \hat{X})$$

$$= \hat{b}_i.$$
where the last equality become from the first equations of (11) and (12). Similarly, we obtain
\[ \sum_{i=1}^{m} y_i(\alpha)A_i + \hat{S}(\alpha) = \hat{C}. \]
Therefore, the proof is complete. \qed

The next lemma shows that searching along the ellipse reduces the duality gap.

**Lemma 3.2.** Let \((X, y, S) \in \mathcal{F}^0\). Then for every \(\alpha \in [0, \frac{\pi}{2}]\), we have
(i) \(\hat{\mu}(\alpha) = \mu(\alpha)\), \(\hat{\mu} = \mu\).
(ii) \(\mu(\alpha) = (1 - \sin(\alpha))\mu\).

**Proof.** We have
\[
\hat{\mu}(\alpha) = \frac{1}{n} \text{Tr}(\hat{X}(\alpha)\hat{S}(\alpha)) = \frac{1}{n} \text{Tr}(P^{-1}X(\alpha)S(\alpha)P) = \frac{1}{n} \text{Tr}(X(\alpha)S(\alpha)) = \mu(\alpha),
\]
and similarly \(\hat{\mu} = \mu\). This proves part (i). For the proof of part (ii), using (15), (17) and Lemma 3.1(i), we obtain
\[
\mu(\alpha) = \frac{1}{n} \text{Tr}(\hat{X}(\alpha)\hat{S}(\alpha)) = \frac{1}{n} \text{Tr}((\hat{X} - \sin(\alpha)\hat{X} + (1 - \cos(\alpha))\hat{X}((\hat{S} - \sin(\alpha)\hat{S} + (1 - \cos(\alpha))\hat{S}))
\]
\[
= \frac{1}{n} \text{Tr}(\hat{X}\hat{S} - \sin(\alpha)(\hat{X}\hat{S} + \hat{X}\hat{S}) + (1 - \cos(\alpha))(\hat{X}\hat{S} + \hat{X}\hat{S}))
\]
\[
= \frac{1}{n} \text{Tr}(\hat{X}\hat{S}) - \frac{\sin(\alpha)}{n} \text{Tr}(\hat{X}\hat{S}) + \frac{1 - \cos(\alpha)}{n} \text{Tr}(-2\hat{X}\hat{S})
\]
\[
= \frac{1 - \sin(\alpha)}{n} \text{Tr}(\hat{X}\hat{S}) = (1 - \sin(\alpha))\mu,
\]
where the fourth equality is due to the identity \(\text{Tr}(H_P(M)) = \text{Tr}(M)\) and the third equations of (11) and (12). The fifth equality is due to Lemma 3.1(i) and the last equality follows from part (i). This proves part (ii). \qed

Clearly, part (ii) of Lemma 3.2 says that \(\mu(\alpha) < \mu\) holds for any choice of \(\alpha \in (0, \frac{\pi}{2}]\). Furthermore, if \(X(\alpha) > 0\) and \(S(\alpha) > 0\) hold in all iterations, reducing the duality gap to zero means approaching the solution of SDO problem.

Here, we cite the following two lemmas that will be frequently used in the convergence analysis.

**Lemma 3.3.** ([24, Lemma 3.3]) Let \(p > 0\), \(q > 0\), and \(r > 0\). If \(p + q \leq r\), then \(pq \leq \frac{r^2}{4}\).

**Lemma 3.4.** ([13, Lemma 4.1]) Let \(u, v\) and \(w\) be real vectors of same size satisfying \(u + v = w\) and \(u^Tv \geq 0\). Then, we have
\[ 2||u||^2||v||^2 \leq ||u||^2 + ||v||^2 \leq \frac{r^2}{4} + 2u^Tv = ||u + v||^2 = ||w||^2. \]

The following lemma will be useful in the next sections.

**Lemma 3.5.** Let \((\hat{X}, \hat{y}, \hat{S})\) and \((\hat{X}, \hat{y}, \hat{S})\) be calculated by (11) and (12), respectively. Assuming that \((X, y, S) \in \mathcal{N}_P(\theta)\), then
(i) \(\|\hat{X}\hat{S}\|_F \leq \|\hat{X}\hat{S}\|_F \leq \frac{\mu}{2} + \frac{n^2\mu}{2(1 - \theta)}\).
(ii) \( \| \hat{X} \hat{S} \|_F + \| \hat{X} \hat{S} \|_F \leq \frac{n^2 \mu}{(1 - \theta)^2} \).

**Proof.** By \( \hat{E} = \hat{F} \), (13) and (14) can be rewritten as

\[
\text{vec}(\hat{X}) + \text{vec}(\hat{S}) = \hat{E}^{-1} \text{vec}(H(\hat{X} \hat{S})) ,
\]

and

\[
\text{vec}(\hat{X}) + \text{vec}(\hat{S}) = -2\hat{E}^{-1} \text{vec}(H(\hat{X} \hat{S})) .
\]

Taking norm squared of both sides of (20) and using \((\text{vec}(\hat{X}))^T (\text{vec}(\hat{S})) = \text{Tr}(\hat{X} \hat{S}) = 0\), we get

\[
\| \text{vec}(\hat{X}) \|^2 + \| \text{vec}(\hat{S}) \|^2 = \| \hat{E}^{-1} \text{vec}(H(\hat{X} \hat{S})) \|^2
\]

\[
= \| \hat{E}^{-1} \text{vec}(\hat{X} \hat{S}) \|^2 = \sum_{i=1}^{n} \frac{\lambda_i^2(\hat{X} \hat{S})}{\lambda_i(X \hat{S})}
\]

\[
= \sum_{i=1}^{n} \lambda_i(\hat{X} \hat{S}) = \text{Tr}(\hat{X} \hat{S}) = n\mu.
\]

Thus, we have

\[
\| \hat{X} \hat{S} \|_F \leq \| \hat{X} \|_F \| \hat{S} \|_F = \| \text{vec}(\hat{X}) \| \| \text{vec}(\hat{S}) \|
\]

\[
\leq \frac{1}{2} (\| \text{vec}(\hat{X}) \|^2 + \| \text{vec}(\hat{S}) \|^2) = \frac{1}{2} n\mu.
\]

Using \((X, y, S) \in \mathcal{N}_F(\theta)\) and proceeding exactly as in (22) and (23) on (21), we have

\[
\| \text{vec}(\hat{X}) \|^2 + \| \text{vec}(\hat{S}) \|^2 = 4\| \hat{E}^{-1} \text{vec}(\hat{X} \hat{S}) \|^2 \leq 4\rho(\hat{E}^{-1})^2 \| \hat{X} \hat{S} \|_F^2
\]

\[
\leq 4\| \hat{X} \hat{S} \|_F^2 \leq \frac{(n\mu)^2}{\lambda_{\min}(X \hat{S})} \leq \frac{n^2 \mu}{(1 - \theta)^2},
\]

and

\[
\| \hat{X} \hat{S} \|_F \leq \| \hat{X} \|_F \| \hat{S} \|_F = \| \text{vec}(\hat{X}) \| \| \text{vec}(\hat{S}) \|
\]

\[
= \frac{1}{2} (\| \text{vec}(\hat{X}) \|^2 + \| \text{vec}(\hat{S}) \|^2) \leq \frac{n^2 \mu}{2(1 - \theta)}.
\]

The inequalities (23) and (25) imply part (i). To prove (ii), we have

\[
(\| \hat{X} \hat{S} \|^2 + \| \hat{X} \hat{S} \|^2) \leq \| \hat{X} \|^2 \| \hat{S} \|^2 + \| \hat{X} \|^2 \| \hat{S} \|^2 + 2\| \hat{X} \| \| \hat{S} \| \| \hat{S} \| \| \hat{S} \|
\]

\[
\leq \| \text{vec}(\hat{X}) \|^2 \| \text{vec}(\hat{S}) \|^2 + \| \text{vec}(\hat{X}) \|^2 \| \text{vec}(\hat{S}) \|^2
\]

\[
+ \| \text{vec}(\hat{X}) \|^2 \| \text{vec}(\hat{S}) \|^2 + \| \text{vec}(\hat{S}) \|^2 \| \text{vec}(\hat{S}) \|^2
\]

\[
= (\| \text{vec}(\hat{X}) \|^2 + \| \text{vec}(\hat{S}) \|^2) (\| \text{vec}(\hat{X}) \|^2 + \| \text{vec}(\hat{S}) \|^2)
\]

\[
\leq n\mu \frac{n^2 \mu}{1 - \theta} = \frac{n^3 \mu^2}{1 - \theta},
\]

where the last inequality is due to (23) and (25). Taking the square root of both sides of the above inequality implies part (ii). Therefore, the proof is complete. \(\square\)
4. Complexity analysis.

Lemma 4.1. Let \((\hat{X}^k, y^k, \hat{S}^k) \in \mathcal{N}_F(\theta)\), and \((\hat{X}, \hat{y}, \hat{S})\) and \((\hat{X}, \hat{y}, \hat{S})\) be computed by (11) and (12), \(\sin(\alpha) > 0\) be the solution of \(q(\alpha) = 0\) in Step 3 of the algorithm. Then \((\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha)) \in \mathcal{N}_F(2\theta)\).

Proof. Using (15), (17) and the third equations of (8) and (9) and Lemma 3.2 we obtain

\[
\left\| \hat{X}(\alpha) \hat{S}(\alpha) - \hat{\mu}(\alpha) I \right\|_F \\
= \left\| (\hat{X} - \sin(\alpha) \hat{X} + (1 - \cos(\alpha)) \hat{X}) (\hat{S}^k - \sin(\alpha) \hat{S} + (1 - \cos(\alpha)) \hat{S}) - \hat{\mu}(\alpha) I \right\|_F \\
= \left\| \hat{X} \hat{S}^k - \sin(\alpha)(\hat{X} \hat{S}^k + \hat{X} \hat{S}^k) + (1 - \cos(\alpha))(\hat{X} \hat{S}^k + \hat{X} \hat{S}) + \sin^2(\alpha) \hat{X} \hat{S} \\
- \sin(\alpha)(1 - \cos(\alpha))(\hat{X} \hat{S} + \hat{X} \hat{S}) + (1 - \cos(\alpha))^2 \hat{X} \hat{S} - (1 - \sin(\alpha)) \hat{\mu} I \right\|_F \\
= \left\| (1 - \sin(\alpha))(\hat{X} \hat{S}^k - \hat{\mu}^k I) + (1 - \cos(\alpha))^2 (\hat{X} \hat{S} - \hat{X} \hat{S}) \\
- \sin(\alpha)(1 - \cos(\alpha))(\hat{X} \hat{S} + \hat{X} \hat{S}) \right\|_F \\
\leq (1 - \sin(\alpha)) \left\| \hat{X} \hat{S}^k - \hat{\mu}^k I \right\|_F + (1 - \cos(\alpha))^2 \left( \left\| \hat{X} \hat{S} \right\|_F + \left\| \hat{X} \hat{S} \right\|_F \right) \\
+ \sin(\alpha)(1 - \cos(\alpha)) \left( \left\| \hat{X} \hat{S} \right\|_F + \left\| \hat{X} \hat{S} \right\|_F \right) \\
\leq (1 - \sin(\alpha)) \theta \hat{\mu}^k + \sin^2(\alpha) \left( \left\| \hat{X} \hat{S} \right\|_F + \left\| \hat{X} \hat{S} \right\|_F \right) + \sin^3(\alpha) \left( \left\| \hat{X} \hat{S} \right\|_F + \left\| \hat{X} \hat{S} \right\|_F \right) \\
= \theta \hat{\mu}(\alpha) + \theta \hat{\mu}(\alpha) = (2\theta) \hat{\mu}(\alpha),
\]

where the first inequality is due to the triangle inequality, the second inequality follows from \((\hat{X}^k, y^k, \hat{S}^k) \in \mathcal{N}_F(\theta)\) and \(\sin^2(\alpha) \geq (1 - \cos(\alpha))\) and the last two inequalities hold by Step 3 of the algorithm. Moreover, from the above inequality it follows that

\[
\lambda_{\min}(\hat{X}(\alpha) \hat{S}(\alpha)) \geq (1 - 2\theta) \hat{\mu}(\alpha) = (1 - 2\theta)(1 - \sin(\alpha)) \hat{\mu}^k > 0,
\]

for all \(\alpha \in [0, \frac{\pi}{2}]\) and \(\theta < \frac{1}{2}\). This means that \(\hat{X}(\alpha) \hat{S}(\alpha)\) is a nonsingular matrix and further implies that \(\hat{X}(\alpha)\) and \(\hat{S}(\alpha)\) are non-singular as well. Using continuity of the eigenvalues of a symmetric matrix, it follows that \(\hat{X}(\alpha)\) and \(\hat{S}(\alpha)\) are positive definite matrices for all \(\alpha \in [0, \frac{\pi}{2}]\), since \(\hat{X}\) and \(\hat{S}\) are positive definite matrices. Therefore, \((\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha)) \in \mathcal{N}_F(2\theta)\). This completes the proof.

The following lemma shows that the corrector steps return any point in \(\mathcal{N}_F(2\theta)\) to the inner \(\mathcal{N}_F(\theta)\) after the \(\hat{\mu}\)-update.

Lemma 4.2. Assuming \((\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha)) \in \mathcal{N}_F(2\theta)\) and \(\theta \in [0, 0.25]\). Then

\[(i) \quad \hat{\mu}^{k+1} = \hat{\mu}(\alpha).
\]

\[(ii) \quad (\hat{X}^{k+1}, y^{k+1}, \hat{S}^{k+1}) \in \mathcal{N}_F(\theta).
\]

Proof. Introduce a step length \(t\) with \(0 \leq t \leq 1\) and define

\[(\hat{X}^{k+1}(t), \hat{S}^{k+1}(t)) = (\hat{X}(\alpha) + t\Delta \hat{X}, \hat{S}(\alpha) + t\Delta \hat{S}).\]
From the third equation of (5) at $(\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha))$, it follows that

$$
\text{Tr}(H(\hat{X}(\alpha)\Delta \hat{S} + \Delta \hat{X}\hat{S}(\alpha))) = \text{Tr}(\hat{X}(\alpha)\Delta \hat{S} + \Delta \hat{X}\hat{S}(\alpha)) = \text{Tr}(\hat{\mu}(\alpha)I - (\hat{X}(\alpha)\hat{S}(\alpha))) = 0,
$$

and from the first two equations of (5) we get $\text{Tr}(\Delta \hat{X}\Delta \hat{S}) = 0$. Invoking (26) and the last equality, we have

$$
\hat{\mu}^{k+1}(t) = \frac{1}{n} \text{Tr}\left( (\hat{X}(\alpha) + t\Delta \hat{X}) (\hat{S}(\alpha) + t\Delta \hat{S}) \right) = \frac{1}{n} \text{Tr}(\hat{X}(\alpha)\hat{S}(\alpha) + t(\hat{X}(\alpha)\Delta \hat{S} + \Delta \hat{X}\hat{S}(\alpha)) + t^2 \Delta \hat{X}\Delta \hat{S})
= \frac{1}{n} \text{Tr}(\hat{X}(\alpha)\hat{S}(\alpha)) = \hat{\mu}(\alpha).
$$

This proves part (i). To prove part (ii), using the third equation of the given system in Step 5 of the algorithm, we have

$$
\text{vec}(\Delta \hat{X}) + \text{vec}(\Delta \hat{S}) = \bar{E}^{-1}(\alpha)\text{vec}(\hat{\mu}(\alpha)I - H(\hat{X}(\alpha)\hat{S}(\alpha))).
$$

Applying Lemma 3.4 to (27) and the orthogonality between $\text{vec}(\Delta \hat{X})$ and $\text{vec}(\Delta \hat{S})$, we can get

$$
\|\text{vec}(\Delta \hat{X})\|^2 + \|\text{vec}(\Delta \hat{S})\|^2 = \|\bar{E}^{-1}(\alpha)\text{vec}(\hat{\mu}(\alpha)I - H(\hat{X}(\alpha)\hat{S}(\alpha)))\|^2.
$$

By equation (4.5) in [14] and the fact that $\|H_F(M)\|_F \leq \|M\|_F$, we have

$$
\|\bar{E}^{-1}(\alpha)\text{vec}(H(\hat{\mu}(\alpha)I - \hat{X}(\alpha)\hat{S}(\alpha)))\|^2 = \sum_{i+1}^n \frac{\lambda_i^2(H(\hat{\mu}(\alpha)I - \hat{X}(\alpha)\hat{S}(\alpha)))}{\lambda_i(\hat{X}(\alpha)\hat{S}(\alpha))} \leq \frac{\|H(\hat{\mu}(\alpha)I - \hat{X}(\alpha)\hat{S}(\alpha))\|^2}{\lambda_{\min}(\hat{X}(\alpha)\hat{S}(\alpha))} \leq \frac{\|\hat{\mu}(\alpha)I - \hat{X}(\alpha)\hat{S}(\alpha)\|^2}{(1 - 2\theta)\hat{\mu}(\alpha)} \leq \frac{(2\theta)^2\hat{\mu}(\alpha)}{1 - 2\theta} = \frac{4\theta^2}{1 - 2\theta}\hat{\mu}^{k+1},
$$

where the second and third inequalities are due to $(\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha)) \in \mathcal{N}_F(2\theta)$ and the last equality follows from part (i). Therefore, using (28) and the above inequality, we obtain

$$
\|H(\Delta \hat{X}\Delta \hat{S})\|_F \leq \|\Delta \hat{X}\|_F\|\Delta \hat{S}\|_F = \|\text{vec}(\Delta \hat{X})\|\|\text{vec}(\Delta \hat{S})\| \leq \frac{1}{2} \left( \|\text{vec}(\Delta \hat{X})\|^2 + \|\text{vec}(\Delta \hat{S})\|^2 \right) \leq \frac{2\theta^2}{1 - 2\theta}\hat{\mu}^{k+1}.
$$
Hence, by part (i), we have
\[
\| (X^{k+1})^\frac{1}{2} (t) S^{k+1}(t) (X^{k+1})^\frac{1}{2}(t) - \mu^{k+1}(t) I \|_F
\leq \| H_p (X^{k+1}(t) S^{k+1}(t) - \mu^{k+1}(t) I) \|_F
= \| H (\hat{X}^{k+1}(t) \hat{S}^{k+1}(t)) - \hat{\mu}^{k+1}(t) I \|_F
= \| H ((\hat{X}(\alpha) + t \Delta \hat{X}) (\hat{S}(\alpha) + t \Delta \hat{S})) - \hat{\mu}(\alpha) I \|_F
= \| (1 - t) (H (\hat{X}(\alpha) \hat{S}(\alpha)) - \hat{\mu}(\alpha) I) + t^2 H (\Delta \hat{X} \Delta \hat{S}) \|_F
\leq (1 - t) \| \hat{X}^\frac{1}{2}(\alpha) \hat{S}(\alpha) \hat{X}^\frac{1}{2}(\alpha) - \hat{\mu}(\alpha) I \|_F + t^2 \| H (\Delta \hat{X} \Delta \hat{S}) \|_F
\leq \left( (1 - t) 2\theta + \frac{2t^2\theta^2}{1 - 2\theta} \right) \hat{\mu}^{k+1} := h(t, \theta) \hat{\mu}^{k+1},
\]
where the third equality is obtained by the third equation of (5) at \((\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha))\), the second inequality is due to the triangle inequality and the last inequality follows from (29) and \((\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha)) \in \mathcal{N}_F(2\theta)\). Taking \(t = 1\) gives
\[
\| (X^{k+1})^\frac{1}{2} S^{k+1}(X^{k+1})^\frac{1}{2} - \mu^{k+1} I \|_F \leq \frac{2\theta^2}{1 - 2\theta} \mu^{k+1}.
\]
It is easy to see that
\[
\frac{2\theta^2}{1 - 2\theta} \leq \theta
\]
if \(\theta \in [0, 0.25]\). Since, for \(\theta \in [0, 0.25]\) and \(t \in [0, 1]\), \(0 < h(t, \theta) \leq h(t, 0.25) < 1\), we have, for \(\alpha \in [0, \frac{\pi}{2}]\),
\[
\lambda_i (X^{k+1}(t) S^{k+1}(t)) \geq \mu^{k+1} - h(t, \theta) \hat{\mu}^{k+1} = (1 - h(t, \theta)) \hat{\mu}^{k+1}
= (1 - h(t, \theta)) \hat{\mu}(\alpha) = (1 - h(t, \theta))(1 - \sin(\alpha)) \hat{\mu}^k > 0,
\]
where the first inequality is due to Lemma 3.2 (i), the second equality is due to part (i) and the last equality follows from Lemma 3.2. This implies that \(X^{k+1} > 0\) and \(S^{k+1} > 0\). Therefore \((X^{k+1}, y^{k+1}, S^{k+1}) \in \mathcal{N}_F(\theta)\). This completes the proof. \(\square\)

The following lemma finds an upper bound on \(\alpha\).

**Lemma 4.3.** Let \(\theta \in [0, 0.25]\) and \(\sin(\alpha)\) for \(\alpha \in [0, \frac{\pi}{2}]\) be the positive real solution of \(q(\alpha) = 0\). Then, \(\sin(\alpha) \geq \frac{\vartheta}{\sqrt{n}}\).

**Proof.** Since \(q(\sin(\alpha))\) is a monotonic increasing function of \(\sin(\alpha) \in [0, 1]\), \(q(\sin(0)) < 0\) and \(q(\sin(\frac{\pi}{2})) \geq 0\), we need only to show that \(q(\frac{\vartheta}{\sqrt{n}}) < 0\). Using Lemma 3.5, for \(\sin(\alpha) \leq \frac{\vartheta}{\sqrt{n}}\), we have
\[
q(\alpha) = \left( \| \hat{X} \hat{S} \|_F + \| \hat{X} \hat{S} \|_F \right) \sin^4(\alpha) + \left( \| \hat{X} \hat{S} \|_F + \| \hat{X} \hat{S} \|_F \right) \sin^3(\alpha)
+ \theta \mu^k \sin(\alpha) - \theta \mu^k
\leq \left( \frac{n \mu^k}{2} + \frac{n^2 \mu^k}{2(1 - \theta)} \right) \left( \frac{\theta^4}{n^2} + \left( \frac{n \sqrt{n} \mu^k}{\sqrt{1 - \theta}} \right) \left( \frac{\theta^3}{\sqrt{n}} \right) + \theta \mu^k \left( \frac{\theta^2}{\sqrt{1 - \theta}} \right) - \theta \mu^k \right)
= \frac{\theta \mu^k}{2} \left( \theta^4 + \frac{\theta^3}{2} \frac{\theta^2}{\sqrt{1 - \theta}} + \frac{\theta^2}{\sqrt{1 - \theta}} + \frac{\theta^2}{\sqrt{n}} - 1 \right) := \theta \mu^k f(\theta, n).
\]
It is easy to verify that \(f(\theta, n)\) is a monotonic increasing function of \(\theta \in (0, 1)\), and a monotonic decreasing function of \(n = 1, 2, \ldots\). For \(\theta \leq 0.25\) and \(n > 1\),
\[ f(\theta, n) = -0.6596 < 0. \] Hence for \( \sin(\alpha) \leq \frac{\theta}{\sqrt{n}} \), we have \( q(\sin(\alpha)) < 0 \). This proves the lemma.

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

**Lemma 4.4.** Let \((X^0, y^0, S^0) \in F^0, \mu^0 = 1/n \text{Tr}(X^0S^0)\) and \((X^k, y^k, S^k)\) be the iterates obtained after \(k\) iterations. Then \(\mu^k \leq \varepsilon\), for \(k \geq \frac{\sqrt{n}}{\theta} \log \frac{\mu^0}{\varepsilon}\).

**Proof.** Using Lemma 3.2(ii) and Lemma 4.3 we have

\[
\mu^k \leq \left(1 - \frac{\theta}{\sqrt{n}}\right) \mu^{k-1} \leq \left(1 - \frac{\theta}{\sqrt{n}}\right)^k \mu^0.
\]

The inequality \(\mu^k \leq \varepsilon\) holds if

\[
\left(1 - \frac{\theta}{\sqrt{n}}\right)^k \mu^0 \leq \varepsilon.
\]

Taking logarithms and using \(\log(1 + z) \leq z, z > -1\) ([23, Lemma 4.1]), we conclude that the above inequality holds if

\[
k \geq \frac{\sqrt{n}}{\theta} \log \frac{\mu^0}{\varepsilon}.
\]

This completes the proof.

We are ready to state our main result.

**Theorem 4.5.** Algorithm 1 will terminate in \(O(\sqrt{n} \log \frac{\mu^0}{\varepsilon})\) iterations.

5. **Numerical results.** In this section, we compare the proposed primal-dual arc-search algorithm with MTY predictor-corrector algorithm [5]. These two algorithms will be denoted by NewAlgor and MTYAlgor, respectively. Numerical results were obtained by using MATLAB R2014b on an Intel Core i3 (3.40 Ghz) with 4 GB RAM.

We stop the iteration of each algorithm if \(\mu = \frac{\text{Tr}(XS)}{n} \leq \varepsilon\), where \(\varepsilon = 10^{-6}, 10^{-8}\) for Example 1, and \(\varepsilon = 10^{-5}\) for Example 2.

**Example 1.** Consider the primal-dual pair of \((P)\) and \((D)\) with the following data

\[
A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 & -1 \\ 0 & -1 & 1 & -1 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & -2 & 2 & 0 \\ 0 & 2 & 1 & 0 & 2 \\ -2 & 1 & -2 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 2 \end{bmatrix},
\]

\[
A_3 = \begin{bmatrix} 2 & 2 & -1 & -1 & 1 \\ 2 & 0 & 2 & 1 & 1 \\ -1 & 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & 0 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 3 & -3 & 1 & 1 \\ 3 & 5 & 3 & 1 & 2 \\ -3 & 3 & -1 & 1 & 2 \\ 1 & 1 & 1 & -3 & -1 \\ 1 & 2 & 2 & -1 & -1 \end{bmatrix},
\]

and \(b = [-2, 2, -2]^T\). We choose \((X^0, y^0, S^0) = (I_{5 \times 5}, 1_3, I_{5 \times 5})\) as starting point.

The numerical results are summarized in Table 1 where we present the duality gap (DGAP), the CPU time (in seconds) and the required iteration numbers (Iter.)
Table 1. Numerical results for Example 1.

| ε     | MTYAlgor | NewAlgor |
|-------|----------|----------|
|       | Iter. | CPU | DGAP | Iter. | CPU | DGAP |
| 10^{-6} | 55 | 0.1882 | 4.4981e^{-6} | 11 | 0.0729 | 2.1703e^{-6} |
| 10^{-8} | 73 | 0.2476 | 4.7261e^{-8} | 13 | 0.0679 | 9.6585e^{-9} |

Table 2. Numerical results for Example 2.

*Example 2.* In this example, we test some various dimensions primal-dual pair of (P) and (D) in which the coefficient matrices $A_i$ are generated randomly. The Table 2 list the number of the constraints ($m$), the dimension of the blocks ($n$), the average number of iterations and the average of CPU times of ten randomly generated SDO problems with the same $m$ and $n$.

| $n \times m$ | MTYAlgor | NewAlgor |
|--------------|----------|----------|
|              | Iter. | CPU | Iter. | CPU |
| 10 \times 20 | 97.2 | 3.4915 | 28.8 | 1.1028 |
| 30 \times 30 | 117.3 | 18.7385 | 32.3 | 5.9383 |
| 25 \times 40 | 116.2 | 67.5448 | 33.7 | 23.2533 |
| 40 \times 20 | 155.9 | 70.7626 | 33.1 | 24.0481 |
| 50 \times 50 | 136.7 | 179.9818 | 35.4 | 87.9356 |

From the results in Tables 1 and 2, we can see that our primal-dual arc-search algorithm for the SDO is robust and promising.

6. **Conclusion.** In this paper, we extended an arc-search strategy introduced by Yang [26] for LO problems to the SDO case. The proposed algorithm searches the optimizers along the ellipses that approximate the entire central path. The derived complexity bound is of the same order as the one for LO case. Moreover, some preliminary numerical results illustrate the effectiveness of our primal-dual arc-search path-following method for solving the SDO. The used problems are small-size. Therefore, as a further research the efficiency of our algorithm could be analyzed for sparse problems as well.

**Acknowledgments.** The authors would like to thank the editor and the two anonymous referees for their useful comments and suggestions, which helped to improve the presentation of this paper.

**REFERENCES**

[1] F. Alizadeh, Combinatorial Optimization with Interior-Point Methods and Semi-definite Matrices, Ph.D. thesis, Computer Science Department, University of Minnesota, Minneapolis, MN, 1991.

[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.A. Haeberly and M. L. Overton, Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results, *SIAM J. Optim.*, 8 (1998), 746–768.

[4] S. Boyd, L. Ghoui, E. Feron and V. Balakrishnan, Linear Matrix Inequalities in System and Control Theory, SIAM, Philadelphia, PA, 1994.

[5] E. De Klerk, Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications, Kluwer Academic Publishers, Dordrecht, Netherlands, 2002.
ON THE EXTENSION OF AN ARC-SEARCH ···

[6] D. Herbison-Evans, Solving quartics and cubics for graphics, Technical Report, R94-487, Basser Department of Computer Science, University of Sydney, Sydney, Australia, 1994.

[7] B. Kheirfam, An arc-search interior point method in the $N_{\infty}$ neighborhood for symmetric optimization, *Fundam. Inform.*, 146 (2016), 255–269.

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

[9] M. Kojima, M. Shida and S. Shindoh, Local convergence of predictor-corrector infeasible interior-point algorithm for SDPs and SDLCPs, *Math. Program.*, 80 (1998), 129–160.

[10] Y. Li and T. Terlaky, A new class of large neighborhood path-following interior point algorithms for semidefinite optimization with $O(\sqrt{n}\log(\frac{\text{tr}(X^0S^0)}{\epsilon}))$ iteration complexity, *SIAM J. Optim.*, 8 (2010), 2853–2875.

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

[12] S. Mizuno, M. J. Todd and Y. Ye, On adaptive-step primal-dual interior-point algorithms for linear programming, *Math. Oper. Res.*, 18 (1993), 964–981.

[13] R. D. C. Monteiro and I. Adler, Interior path following primal-dual algorithm. Part I: linear programming, *Math. Program.*, 44 (1989), 27–41.

[14] R. D. C. Monteiro, Polynomial convergence of primal-dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions, *SIAM J. Optim.*, 8 (1998), 797–812.

[15] R. D. C. Monteiro and Y. Zhang, A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming, *Math. Program.*, 81 (1998), 281–299.

[16] Y. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics, Vol. 13 SIAM, Philadelphia, USA, 1994.

[17] Y. E. Nesterov and M. J. Todd, Self-scaled barriers and interior-point methods for convex programming, *Math. Oper. Res.*, 22 (1997), 1–42.

[18] Y. E. Nesterov and M. J. Todd, Primal-dual interior point methods for self-scaled cones, *SIAM J. Optim.*, 8 (1998), 324–364.

[19] F. A. Potra and R. Sheng, A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming, *SIAM J. Optim.*, 8 (1998), 1007–1028.

[20] M. J. Todd, K. C. Toh and R. H. Tütüncü, On the Nesterov-Todd direction in semidefinite programming, *SIAM J. Optim.*, 8 (1998), 769–796.

[21] S. Veeraraghavan and D. A. Mazzotti, Semidefinite programming formulation of linear-scaling electronic structure theories Artic, *Phys. Rev. A*, 92 (2015), 215–220.

[22] H. Wolkowicz, R. Saigal and L. Vandenberghe, *Handbook of Semidefinite Programming*, International Series in Operations Research and Management Science (27) Kluwer Academic Publishers, Boston, MA, 2000.

[23] S. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.

[24] Y. Yang, Arc-search path-following interior-point algorithm for linear programming, Optimization online, 2009.

[25] Y. Yang, A polynomial arc-search interior-point algorithm for convex quadratic programming, *Eur. J. Oper. Res.*, 215 (2011), 25–38.

[26] Y. Yang, A polynomial arc-search interior-point algorithm for linear programming, *J. Optim. Theory Appl.*, 158 (2013), 859–873.

[27] X. Yang, Y. Zhang and H. Liu, A wide neighborhood infeasible-interior-point method with arc-search for linear programming, *J. Appl. Math. Comput.*, 51 (2016), 209–225.

[28] X. Yang, H. Liu and Y. Zhang, An arc-search infeasible-interior-point method for symmetric optimization in a wide neighborhood of the central path, *Optim. Lett.*, 11 (2017), 135–152.

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

Received June 2017; 1st revision February 2018; final revision March 2018.

E-mail address: b.kheirfam@azaruniv.edu
E-mail address: m.moslemi83@gmail.com