

FULL-NEWTON STEP INFEASIBLE INTERIOR-POINT ALGORITHM FOR SDO PROBLEMS

HOSSEIN MANSOURI

In this paper we propose a primal-dual path-following interior-point algorithm for semidefinite optimization. The algorithm constructs strictly feasible iterates for a sequence of perturbations of the given problem and its dual problem. Each main step of the algorithm consists of a feasibility step and several centering steps. At each iteration, we use only full-Newton step. Moreover, we use a more natural feasibility step, which targets at the μ^+ -center. The iteration bound of the algorithm coincides with the currently best iteration bound for semidefinite optimization problems.

Keywords: semidefinite optimization, infeasible interior-point method, primal-dual method, polynomial complexity, Newton-step, optimal solutions

Classification: 90C05, 90C51

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 [1, 20]. 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 [7, 11, 12, 16, 21]. 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 a comprehensive learning about IPMs, we refer to Roos et al. [19] and Klerk [4]. Infeasible IPMs (IIPMs) start with an arbitrary positive point and feasibility is reached as optimality is approached. The choice of the starting point in IIPMs is crucial for the performance. Lustig [8] was the first to present IIPMs for LO. Zhang [21] was the first to present a primal-dual IIPMs for SDO with polynomial iteration complexity. In [18] an IIPM for LO was proposed by Roos and later this algorithm extended to semidefinite optimization by Mansouri et al. [10, 11, 12, 13]. It differs from the classical IIPMs [6], since the new method uses only full steps which has the advantage that no line searches are needed. Our motivation for the use of full-Newton steps is that, we use another definition for feasibility step and show that

the complexity coincides with the best known complexity of IIPMs. In this paper, we further consider full-Newton infeasible interior-point algorithm for SDO problems. The main difference with the aforementioned algorithm is the way the search directions are generated. In our algorithm, we change the definition of the feasibility step by replacing equation

$$\Delta^f X + P\Delta^f S P^T = (1 - \theta) \mu S^{-1} - X$$

instead of classical directions to calculate the search directions. A special feature of the direction is that, if the iterate is strictly feasible for SDO, the full-Newton step targeting to central path has local quadratic convergence property according to the proximity measure. The complexity result shows that the full-Newton step IIPM for SDO based on new directions enjoys the best-known iteration complexity for SDO.

This paper is organized as follows. First, we review some results which are due to [11], and then, apply them to analyze the feasibility and the centering steps of our algorithm. Then we present our algorithm. Each main step of the algorithm consists of a feasibility step and several centering steps. Recall that in [11] the feasibility step targets at the μ -center of the next pair of perturbed problems. Since the aim of each main iteration is to get a good approximation of the μ^+ -center of the next pair of perturbed problems, we take a more natural approach to let the feasibility step target at the μ^+ -center of the next pair of perturbed problems. Finally, we give some concluding remarks.

NOTATIONS

Some notations used throughout the paper are as follows. The superscript T denotes transpose. \mathbf{R}^n , \mathbf{R}_+^n and \mathbf{R}_{++}^n denote the set of vectors with n components, the set of nonnegative vector and the set of positive vectors, respectively. $\mathbf{R}^{m \times n}$ is the space of all $m \times n$ matrices. \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. \mathcal{P} and \mathcal{D} denote the feasible sets of primal and dual problems respectively. $\text{ri}(\mathcal{C})$ denotes the relative interior of a convex set \mathcal{C} . I denotes $n \times n$ identity matrix. We use the classical Löwner partial order \succeq for symmetric matrices. So $A \succeq B$ ($A \succ B$) means that $A - B$ is positive semidefinite (positive definite). The sign \sim denotes similarity of two matrices. The matrix inner product is defined by $A \bullet B = \text{Tr}(A^T B)$. For any symmetric positive definite matrix $Q \in \mathbf{S}_{++}^n$, the expression $Q^{\frac{1}{2}}$ denotes the symmetric square root of Q . For any $x = (x_1; x_2; \dots; x_n) \in \mathbf{R}^n$, $x_{\min} = \min(x_1; x_2; \dots; x_n)$ and $x_{\max} = \max(x_1; x_2; \dots; x_n)$. For any symmetric matrix G , $\lambda_{\min}(G)$ ($\lambda_{\max}(G)$) denotes the minimal (maximal) eigenvalue of G . When λ is vector we denote the diagonal matrix $\text{diag}(\lambda)$ with entries λ_i by Λ . For any $V \in \mathbf{S}_{++}^n$, we denote by $\lambda(V)$ the vector of eigenvalues of V arranged in non-increasing order, that is, $\lambda_{\max}(V) = \lambda_1(V) \geq \lambda_2(V) \geq \dots \geq \lambda_n(V) = \lambda_{\min}(V)$. The Frobenius and infinity matrix norm are given by $\|U\|^2 := \sum_{i=1}^m \sum_{j=1}^n U_{ij}^2 = \text{Tr}(U^T U)$ and $\|U\|_{\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^n |u_{ij}|$, respectively. For any $p \times q$ matrix A , $\text{vec}(A)$ denotes the pq -vector obtained by stacking the columns of A . The Kronecker product of two matrices A and B is denoted by $A \otimes B$ (we refer to [3] for a comprehensive treatment on Kronecker products and related topics). \mathcal{F}^* denotes the set of optimal solutions with zero duality gap, i. e.,

$$\mathcal{F}^* := \{(X, y, S) \in \mathcal{P} \times \mathcal{D} : \text{Tr}(XS) = 0\}.$$

2. PRELIMINARIES

We consider the semidefinite optimization (SDO) problem given in the following standard form:

$$(P) \quad \min \quad \mathbf{Tr}(CX) \\ \text{s.t.} \quad \mathbf{Tr}(A_i X) = b_i, \quad i = 1, 2, \dots, m, \quad X \succeq 0,$$

and its dual:

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

where each $A_i \in \mathbf{S}^n$, $b \in \mathbf{R}^m$ and $C \in \mathbf{S}^n$. Without loss of generality we assume that the matrices A_i are linearly independent. As usual for infeasible interior-point methods (IIPMs), we use the starting point as in [11] that one knows a positive scalar ζ such that $X^* + S^* \preceq \zeta I$ for some optimal solution (X^*, y^*, S^*) of (P) and (D) such that $\mathbf{Tr}(X^* S^*) = 0$ and the initial iterates are $(X^0, y^0, S^0) = \zeta (I, 0, I)$, where I denotes the identity matrix of size $n \times n$. Using $\mathbf{Tr}(X^0 S^0) = n\zeta^2$, the total number of iterations in the algorithm of [11] is bounded above by

$$20n \log \frac{\max \{n\zeta^2, \|r_b^0\|, \|R_c^0\|\}}{\varepsilon}, \tag{1}$$

where r_b^0 and R_c^0 are the initial values of the primal and dual residuals:

$$(r_b^0)_i = b_i - A_i \bullet X^0, \quad i = 1, \dots, m, \tag{2}$$

$$R_c^0 = C - \sum_{i=1}^m y_i^0 A_i - S^0. \tag{3}$$

Up to a constant factor, the iteration bound (1) was first obtained by Kojima et al. [5] and Potra and Sheng [17], and it is still the best-known iteration bound for IIPMs.

To describe the aim of this article, we need to recall the main ideas underlying the algorithm in [11]. For any ν with $0 < \nu \leq 1$, we consider the perturbed problem (P_ν) , defined by

$$(P_\nu) \quad \min \left\{ (C - \nu(C - \sum_{i=1}^m y_i^0 A_i - S^0)) \bullet X : \right. \\ \left. A_i \bullet X = b_i - \nu(b_i - A_i \bullet X^0), X \succeq 0 \right\}$$

and its dual problem (D_ν) , which is given by

$$(D_\nu) \quad \max \left\{ \sum_{i=1}^m (b_i - \nu(b_i - A_i \bullet X^0)) y_i : \right. \\ \left. \sum_{i=1}^m y_i A_i + S = C - \nu(C - \sum_{i=1}^m y_i^0 A_i - S^0), S \succeq 0 \right\}.$$

Note that if $\nu = 1$ then $X = X^0$ yields a strictly feasible solution of (P_ν) , and $(y, S) = (y^0, S^0)$ a strictly feasible solution of (D_ν) . We conclude that if $\nu = 1$ then (P_ν) and (D_ν) are strictly feasible, which means that both perturbed problems then satisfy the well-known interior-point condition (IPC) [19]. More generally one has the following lemma [11, lemma 4.1].

Lemma 2.1. Let the original problems, (P) and (D) , be feasible. Then for each ν satisfying $0 < \nu \leq 1$ the perturbed problems (P_ν) and (D_ν) are strictly feasible.

Assuming that (P) and (D) are feasible, it follows from Lemma 2.1 that the problems (P_ν) and (D_ν) satisfy the IPC for each $\nu \in (0, 1]$. Therefore, their central paths exist. This means that the system

$$b_i - A_i \bullet X = \nu (r_b^0)_i, \quad i = 1, 2, \dots, m, \quad X \succeq 0 \tag{4}$$

$$C - \sum_{i=1}^m y_i A_i - S = \nu R_c^0, \quad S \succeq 0 \tag{5}$$

$$XS = \mu I$$

has a unique solution, for every $\mu > 0$. In the sequel this unique solution is denoted as $(X(\mu, \nu), y(\mu, \nu), S(\mu, \nu))$ for $\nu \in (0, 1)$. These are the μ -centers of the perturbed problems (P_ν) and (D_ν) . Note that since $X^0 S^0 = \mu^0 I$, X^0 is the μ^0 -center of the perturbed problem (P_1) and (y^0, S^0) the μ^0 -center of (D_1) . In other words, $(X(\mu^0, 1), y(\mu^0, 1), S(\mu^0, 1)) = (X^0, y^0, S^0)$. In the sequel we will always assume $\mu = \nu \mu^0$, and we will accordingly denote $(X(\mu, \nu), y(\mu, \nu), S(\mu, \nu))$ simply as $(X(\nu), y(\nu), S(\nu))$.

We measure proximity of iterates (X, y, S) to the μ -center of the perturbed problems (P_ν) and (D_ν) by the quantity $\delta(X, S; \mu)$, which is defined as follows:

$$\delta(X, S, \mu) := \delta(V) := \frac{1}{2} \|V^{-1} - V\|, \text{ where } V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} D S D. \tag{6}$$

Here $D = P^{\frac{1}{2}}$ with

$$P := X^{\frac{1}{2}} \left(X^{\frac{1}{2}} S X^{\frac{1}{2}} \right)^{-\frac{1}{2}} X^{\frac{1}{2}} = S^{-\frac{1}{2}} \left(S^{\frac{1}{2}} X S^{\frac{1}{2}} \right)^{\frac{1}{2}} S^{-\frac{1}{2}}, \tag{7}$$

which is a symmetric nonsingular matrix. For more details see [14, 15].

Initially, we have $X^0 = S^0 = \zeta I$ and $\mu^0 = \zeta^2$, where $V = I$ and $\delta(X^0, S^0; \mu^0) = 0$. In the sequel, we assume that at the start of each iteration, $\delta(X^0, S^0; \mu^0)$ is smaller than or equal to a (small) threshold $\tau > 0$. So, this is certainly true at the start of the first iteration.

3. AN ITERATION OF THE ALGORITHM

In this section we describe one iteration of our algorithm. As we established above, if $\nu = 1$ and $\mu = \mu^0$, then $X = X^0$ is the μ -center of the perturbed problem (P_ν) and $(y, S) = (y^0, S^0)$ the μ -center of (D_ν) . These are our initial iterates. We measure

proximity to the μ -center of the perturbed problems by the quantity $\delta(X, S; \mu)$ as defined in (6). Initially we thus have $\delta(X, S; \mu) = 0$. In what follows we assume that at the start of each iteration, just before the feasibility step, $\delta(X, S; \mu)$ is smaller than or equal to a small threshold value $\tau > 0$. So this is certainly true at the start of the the first iteration. Suppose that for some $\nu \in (0, 1]$, we have X, y and S , which are strictly feasible and satisfying the feasibility conditions (4) and (5) and such that

$$\mathbf{Tr}(XS) = n\mu, \quad \text{and} \quad \delta(X, S; \mu) \leq \tau, \tag{8}$$

where $\mu = \nu\zeta^2$. Each main iteration consists of one so-called feasibility step, a μ -update, and a few centering steps, respectively. First, we find new iterates X^f, y^f and S^f that satisfy equations (4) and (5), with ν replaced by ν^+ . As we will see, by taking θ small enough, this can be realized by one feasibility step, as discussed subsequently. Therefore, as a result of the feasibility step, we obtain iterates that are feasible for (P_{ν^+}) and (D_{ν^+}) . Then we reduce ν to $\nu^+ = (1 - \theta)\nu$, with $\theta \in (0, 1)$, and apply a limited number of centering steps with respect to the μ^+ -centers of (P_{ν^+}) and (D_{ν^+}) . The centering steps keep the iterates feasible for (P_{ν^+}) and (D_{ν^+}) , their purpose is to get iterates X^+, y^+ and S^+ such that $\mathbf{Tr}(X^+S^+) = n\mu^+$, where $\mu^+ = \nu^+\zeta^2$ and $\delta(X^+, S^+; \mu^+) \leq \tau$. This process is repeated until the duality gap and the norms of residual vectors are less than some prescribed accuracy parameter ε . Before describing the search directions used in the feasibility step and the centering step, we give a more formal description of the algorithm in Algorithm 1.

Algorithm 1 (Primal-dual Algorithm with full-Newton steps)

Input:

- Accuracy parameter $\varepsilon > 0$;
- barrier update parameter $\theta, 0 < \theta < 1$;
- threshold parameter $\tau, 0 < \tau \leq \frac{1}{\sqrt{2}}$.

begin

$X := X^0 \succ 0, S := S^0 \succ 0, y := y^0, X^0 S^0 = \mu^0 I, \mu = \mu^0, \nu = 1$;

while $\max(\mathbf{Tr}(XS), \|r_b\|, \|R_c\|) \geq \varepsilon$ **do**

begin

feasibility step:

$$(X, y, S) := (X, y, S) + (\Delta^f X, \Delta^f y, \Delta^f S);$$

μ -update:

$$\mu := (1 - \theta)\mu;$$

centering steps:

while $\delta(X, S, \mu) \geq \tau$ **do**

begin

$$(X, y, S) := (X, y, S) + (\Delta X, \Delta y, \Delta S);$$

end

end

end

In this paper, for the feasibility step, we use the search directions $\Delta^f X$, $\Delta^f y$ and $\Delta^f S$

$$\mathbf{Tr} (A_i \Delta^f X) = \theta \nu (r_b^0)_i, \quad i = 1, \dots, m, \tag{9}$$

$$\sum_{i=1}^m \Delta^f y_i A_i + \Delta^f S = \theta \nu R_c^0, \tag{10}$$

$$\Delta^f X + P \Delta^f S P^T = (1 - \theta) \mu S^{-1} - X, \tag{11}$$

where we used the Nesterov-Todd-‘trick’ to symmetrize $\Delta^f X$ with P as defined in (7). It is easy to see that if (X, y, S) is feasible for the perturbed problems (P_ν) and (D_ν) , then after the feasibility step the iterates satisfy the feasibility conditions for (P_{ν^+}) and (D_{ν^+}) , provided that they satisfy the positive semidefinite conditions. Assuming that before the step $\delta(X, S; \mu) \leq \tau$ holds, and by taking θ small enough, it can be guaranteed that after the step, the iterates

$$X^f = X + \Delta^f X, \tag{12}$$

$$y^f = y + \Delta^f y,$$

$$S^f = S + \Delta^f S, \tag{13}$$

are semidefinite and moreover $\delta(X^f, S^f; \mu^+) \leq \frac{1}{\sqrt{2}}$, where $\mu^+ = (1 - \theta) \mu$. So, after the μ -update, the iterates are feasible for (P_{ν^+}) and (D_{ν^+}) , and μ is such that $\delta(X^f, S^f; \mu^+) \leq \frac{1}{\sqrt{2}}$.

Remark 3.1. For (11), we use the linearization of $X^f S^f = (1 - \theta) \mu I$, which means that we are targeting at the μ^+ -center of (P_{ν^+}) and (D_{ν^+}) . While in [11], the linearization of $X^f S^f = \mu I$ is used (targeting at the μ -center). As our aim is to calculate a feasible solution to (P_{ν^+}) and (D_{ν^+}) , which should also lie in the quadratic convergence neighborhood to its μ^+ -center, the direction used here is more natural and intuitively better.

In the centering steps, starting at iterates $(X, y, S) = (X^f, y^f, S^f)$ and targeting at the μ -centers, the search directions ΔX , Δy and ΔS are the usual primal-dual Newton directions, (uniquely) defined by

$$A_i \bullet \Delta X = 0, \quad i = 1, 2, \dots, m,$$

$$\sum_{i=1}^m \Delta y_i A_i + \Delta S = 0, \tag{14}$$

$$\Delta X + P \Delta S P^T = \mu S^{-1} - X,$$

where matrix P is defined as in (7). Denoting the iterates after a centering step as X^+ , y^+ and S^+ , we recall the following from [4].

Lemma 3.2. If $\delta := \delta(X, S; \mu) \leq 1$, then the primal-dual Newton-step is feasible, i. e. X^+ and S^+ are nonnegative, and $\mathbf{Tr}(X^+ S^+) = n\mu$. Moreover, if $\delta = \delta(X, S; \mu) \leq \frac{1}{\sqrt{2}}$, then $\delta = \delta(X, S; \mu) \leq \delta^2$.

The centering steps serve to get iterates that satisfy $\text{Tr}(XS) = n\mu^+$ and $\delta(X, S; \mu^+) \leq \tau$, where τ is much smaller than $\frac{1}{\sqrt{2}}$. By using Lemma 3.2, the required number of centering steps can easily be obtained. This goes as follows. After μ -update, we have $\delta(X^f, S^f; \mu^+) \leq \frac{1}{\sqrt{2}}$, and hence after k centering steps, the iterates (X, y, S) satisfy

$$\delta(X, S; \mu^+) \leq \left(\frac{1}{\sqrt{2}}\right)^{2^k}.$$

Just as in [11] this implies that no more than

$$\log_2 \left(\log_2 \frac{1}{\tau^2} \right) \tag{15}$$

centering steps are needed.

4. AN ANALYSIS OF THE ALGORITHM

Let X, y and S denote the iterates at the start of an iteration with $\text{Tr}(XS) = n\mu$ and $\delta(X, S; \mu) \leq \tau$. Recall that at the start of first iteration this is certainly true, because $\text{Tr}(X^0S^0) = n\mu^0$ and $\delta(X^0, S^0; \mu^0) = 0$.

Before dealing with the analysis of the algorithm we recall some lemmas which we use several times in this paper.

Lemma 4.1. (Lemma A.1 in [4]) Let $Q \in \mathbf{S}_{++}^n$, and let $M \in \mathbf{R}^{n \times n}$ be skew-symmetric ($M = -M^T$). One has $\det(Q + M) > 0$. Moreover, if $\lambda_i(Q + M) \in \mathbf{R}$, ($i = 1, \dots, n$), then

$$0 < \lambda_{\min}(Q) \leq \lambda_{\min}(Q + M) \leq \lambda_{\max}(Q + M) \leq \lambda_{\max}(Q).$$

Lemma 4.2. (Lemma 1.2.4 in [2]) Let $A, B \in \mathbf{S}_+^n$. Then we have following inequalities

$$\lambda_{\min}(A) \lambda_{\max}(B) \leq \lambda_{\min}(A) \text{Tr}(B) \leq \text{Tr}(AB) \leq \lambda_{\max}(A) \text{Tr}(B) \leq n \lambda_{\max}(A) \lambda_{\max}(B).$$

Lemma 4.3. (Theorem A.4 in [4]) Let $A \in \mathbf{S}_{++}^n$ and $B \in \mathbf{S}_{++}^n$. Then all the eigenvalues of AB are real and positive.

4.1. The effect of the feasibility step and the choice of θ

As we established in Section 3, the feasibility step generates new iterates X^f, y^f and S^f that satisfy the feasibility equations for (P_{ν^+}) and (D_{ν^+}) . A crucial element in the analysis is to show that after the feasibility step $\delta(X^f, S^f; \mu^+) \leq \frac{1}{\sqrt{2}}$, i.e., that the new iterates are within the region where the Newton process targeting at the μ^+ -centers of (P_{ν^+}) and (D_{ν^+}) is quadratically convergent. We define

$$D_X^f := \frac{1}{\sqrt{\mu}} D^{-1} \Delta^f X D^{-1}, \quad D_S^f := \frac{1}{\sqrt{\mu}} D \Delta^f S D, \quad (V^f)^2 := \frac{1}{\mu^+} D^{-1} X^f S^f D, \tag{16}$$

with D as defined in Section 2. We can now rewrite (9),(10) and (11) as follows.

$$\begin{aligned}
 DA_iD \bullet D_X^f &= \frac{1}{\sqrt{\mu}}\theta\nu (r_b^0)_i, \quad i = 1, \dots, m, \\
 \sum_{i=1}^m \frac{\Delta y_i}{\sqrt{\mu}} DA_iD + D_S^f &= \frac{1}{\sqrt{\mu}}\theta\nu DR_c^0D, \\
 D_X^f + D_S^f &= (1 - \theta)V^{-1} - V.
 \end{aligned}
 \tag{17}$$

From the third equation in (17) we obtain, by multiplying both side from the left by V ,

$$VD_X^f + VD_S^f = (1 - \theta)I - V^2.
 \tag{18}$$

Using (6), (12), (13) and (16), we obtain

$$\begin{aligned}
 X^f &= X + \Delta^f X = \sqrt{\mu}D (V + D_X^f) D, \\
 S^f &= S + \Delta^f S = \sqrt{\mu}D^{-1} (V + D_S^f) D^{-1}.
 \end{aligned}$$

Therefore

$$X^f S^f = \mu D (V + D_X^f) (V + D_S^f) D^{-1}.$$

The last matrix is similar to $\mu (V + D_X^f) (V + D_S^f)$. Thus we have

$$X^f S^f \sim \mu (V + D_X^f) (V + D_S^f).$$

To simplify the notation in the sequel we introduce

$$D_{XS}^f := \frac{1}{2} (D_X^f D_S^f + D_S^f D_X^f),
 \tag{19}$$

and

$$M := (D_X^f V - VD_S^f) + \frac{1}{2} (D_X^f D_S^f - D_S^f D_X^f).
 \tag{20}$$

Note that D_{XS}^f is symmetric and M is skew-symmetric. Now we may write, using (18),

$$\begin{aligned}
 (V + D_X^f) (V + D_S^f) &= V^2 + VD_S^f + D_X^f V + D_X^f D_S^f \\
 &= (1 - \theta)I - VD_X^f + D_X^f V + D_X^f D_S^f.
 \end{aligned}$$

By subtracting and adding $\frac{1}{2}D_S^f D_X^f$ to the last expression we get

$$\begin{aligned}
 (V + D_X^f) (V + D_S^f) &= (1 - \theta)I + \frac{1}{2} (D_X^f D_S^f + D_S^f D_X^f) + (D_X^f V - VD_X^f) \\
 &\quad + \frac{1}{2} (D_X^f D_S^f - D_S^f D_X^f) \\
 &= (1 - \theta)I + D_{XS}^f + M.
 \end{aligned}$$

Using (19) and (20) we obtain

$$X^f S^f \sim \mu \left((1 - \theta)I + D_{XS}^f + M \right). \tag{21}$$

Lemma 4.4. Let $X \succ 0$ and $S \succ 0$. Then the iterates (X^f, y^f, S^f) are strictly feasible if

$$(1 - \theta)I + D_{XS}^f \succ 0.$$

Proof. The proof is similar to the proof of Lemma 5.4 in [11], and is therefore omitted. □

Corollary 4.5. (Corollary 5.5 in [11]) The iterates (X^f, y^f, S^f) are certainly strictly feasible if

$$\left\| D_{XS}^f \right\|_\infty < 1 - \theta.$$

Assuming $\left\| D_{XS}^f \right\|_\infty < 1 - \theta$, which guarantees strict feasibility of the iterates (X^f, y^f, S^f) , we proceed by deriving an upper bound for $\delta(X^f, S^f; \mu^+)$. Recall from definition (6) that

$$\delta(X^f, S^f; \mu^+) = \frac{1}{2} \left\| (V^f)^{-1} - V^f \right\|, \tag{22}$$

with $(V^f)^2$ as defined in (16). In the sequel we denote $\delta(X^f, S^f; \mu^+)$ by $\delta(V^f)$.

We proceed to find an upper bound for $\delta(V^f)$ in terms of $\left\| D_{XS}^f \right\|$. To this end we need some technical results which give information on the eigenvalues and the norm of V^f .

Lemma 4.6. One has

$$\lambda_{\min} \left((V^f)^2 \right) \geq 1 - \left\| \frac{D_{XS}^f}{1 - \theta} \right\|_\infty.$$

Proof. The proof is similar to the proof of Lemma 5.7 in [11], and is therefore omitted. □

Lemma 4.7. One has

$$\left\| I - (V^f)^2 \right\| \leq \left\| \frac{D_{XS}^f}{1 - \theta} \right\|.$$

Proof. Using (21), after division of both sides by $\mu^+ = (1 - \theta)\mu$ we get

$$(V^f)^2 \sim \frac{\mu \left((1 - \theta)I + D_{XS}^f + M \right)}{\mu^+} = \frac{(1 - \theta)I + D_{XS}^f + M}{1 - \theta}.$$

By using properties of the Frobenius norm we have

$$\begin{aligned} \|I - (V^f)^2\|^2 &= \sum_{i=1}^n \left(\frac{\lambda_i((1-\theta)I + D_{XS}^f + M)}{1-\theta} - 1 \right)^2 \\ &= \frac{1}{(1-\theta)^2} \sum_{i=1}^n (\lambda_i((1-\theta)I + D_{XS}^f + M) - 1 + \theta)^2 \\ &= \frac{1}{(1-\theta)^2} \sum_{i=1}^n (\lambda_i(D_{XS}^f + M))^2. \end{aligned}$$

Since $(\lambda_i(D_{XS}^f + M))^2 = \lambda_i((D_{XS}^f + M)^2)$, for each i , we obtain

$$\|I - (V^f)^2\|^2 = \frac{1}{(1-\theta)^2} \text{Tr} \left((D_{XS}^f + M)^2 \right). \tag{23}$$

Using the skew-symmetry of M we obtain

$$\text{Tr} \left((D_{XS}^f + M)^2 \right) = \text{Tr} \left((D_{XS}^f)^2 + MD_{XS}^f + D_{XS}^fM - MM^T \right).$$

Since $MD_{XS}^f + D_{XS}^fM$ is skew-symmetric we obtain

$$\text{Tr} \left((D_{XS}^f + M)^2 \right) = \text{Tr} \left((D_{XS}^f)^2 - MM^T \right) \leq \text{Tr} \left((D_{XS}^f)^2 \right) = \|D_{XS}^f\|^2,$$

where the inequality follows since the matrix MM^T is positive semidefinite. Substitution in (23) gives

$$\|I - (V^f)^2\|^2 \leq \left\| \frac{D_{XS}^f}{1-\theta} \right\|^2.$$

This completes the proof. □

Lemma 4.8. Let $\|D_{XS}^f\|_\infty < 1 - \theta$. Then one has

$$4\delta(V^f)^2 \leq \frac{\left\| \frac{D_{XS}^f}{1-\theta} \right\|^2}{1 - \left\| \frac{D_{XS}^f}{1-\theta} \right\|_\infty}.$$

Proof. The proof is similar to that of Lemma 5.9 in [11] and is omitted. □

Recall from Section 3 that we need to have $\delta(V^f) \leq \frac{1}{\sqrt{2}}$. By Lemma 4.8 it suffices for this that

$$\frac{\left\| \frac{D_{XS}^f}{1-\theta} \right\|^2}{1 - \left\| \frac{D_{XS}^f}{1-\theta} \right\|_\infty} \leq 2. \tag{24}$$

As we may easily verify that

$$\|D_{XS}^f\|^2 \leq \left(\|D_X^f\| \|D_S^f\|\right)^2 \leq \frac{1}{4} \left(\|D_X^f\|^2 + \|D_S^f\|^2\right)^2 \quad (25)$$

$$\|D_{XS}^f\|_\infty \leq \frac{1}{2} \left(\|D_X^f\|_\infty^2 + \|D_S^f\|_\infty^2\right) \leq \frac{1}{2} \left(\|D_X^f\|^2 + \|D_S^f\|^2\right). \quad (26)$$

Substituting (25) and (26) in (24) we obtain

$$\frac{\frac{1}{4} \left(\frac{\|D_X^f\|^2 + \|D_S^f\|^2}{1-\theta}\right)^2}{1 - \frac{1}{2} \left(\frac{\|D_X^f\|^2 + \|D_S^f\|^2}{1-\theta}\right)} \leq 2.$$

Considering $\frac{1}{2} \left(\frac{\|D_X^f\|^2 + \|D_S^f\|^2}{1-\theta}\right)$ as a single term, and by some elementary calculation, we obtain that (24) holds if

$$\frac{\|D_X^f\|^2 + \|D_S^f\|^2}{1-\theta} \leq 1.464. \quad (27)$$

The inequality (27) implies that after the feasibility step (X^f, y^f, S^f) is strictly feasible and lies in the quadratic convergence neighborhood with respect to the μ^+ -center of (P_{ν^+}) and (D_{ν^+}) .

4.2. An upper bound for $\|D_X^f\|^2 + \|D_S^f\|^2$

As became clear in (17), the system (9)–(11), which defines the search directions $\Delta^f X$, $\Delta^f y$ and $\Delta^f S$, can be expressed in terms of scaled search directions D_X^f and D_S^f as follows.

$$\begin{aligned} DA_i D \bullet D_X^f &= \frac{1}{\sqrt{\mu}} \theta \nu (r_b^0)_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^m \frac{\Delta y_i}{\sqrt{\mu}} DA_i D + D_S^f &= \frac{1}{\sqrt{\mu}} \theta \nu DR_c^0 D, \\ D_X^f + D_S^f &= (1-\theta)V^{-1} - V. \end{aligned} \quad (28)$$

We proceed to finding an upper bound for $\|D_X^f\|^2 + \|D_S^f\|^2$. We define the linear space \mathcal{L} as follows:

$$\mathcal{L} := \{\xi \in \mathbf{S}^n \quad : \quad DA_i D \bullet \xi = 0, \quad i = 1, \dots, m\}.$$

Using the linear space \mathcal{L} , it is clear from the first equation in (28) that the affine space

$$\left\{ \xi \in \mathbf{S}^n \quad : \quad DA_i D \bullet \xi = \frac{1}{\sqrt{\mu}} \theta \nu (r_b^0)_i, \quad i = 1, \dots, m \right\}$$

equals $D_X^f + \mathcal{L}$. By the second equation in system (28), we have $D_S^f \in \frac{1}{\sqrt{\mu}}\theta\nu DR_c^0 D + \mathcal{L}^\perp$. Since $\mathcal{L} \cap \mathcal{L}^\perp = \{0\}$, the spaces $D_X^f + \mathcal{L}$ and $D_S^f + \mathcal{L}^\perp$ meet in a unique matrix. This matrix is denoted below by Q .

Lemma 4.9. Let Q be the (unique) matrix in the intersection of the affine spaces $D_X + \mathcal{L}$ and $D_S + \mathcal{L}^\perp$. Then

$$\|D_X^f\|^2 + \|D_S^f\|^2 \leq \|Q\|^2 + \left(\|Q\| + \sqrt{4(1-\theta)^2\delta^2 + \theta^2n}\right)^2.$$

Proof. The proof is similar to the proof of Lemma 5.6 in [18], and is therefore omitted. □

From (27) we know that we want to have $\|D_X^f\|^2 + \|D_S^f\|^2 \leq 1.464(1-\theta)$ because then $\delta(V^f) \leq \frac{1}{\sqrt{2}}$. Due to Lemma 4.9 this will hold if $\|Q\|$ satisfies

$$\|Q\|^2 + \left(\|Q\| + \sqrt{4(1-\theta)^2\delta^2 + \theta^2n}\right)^2 \leq 1.464(1-\theta). \tag{29}$$

Before doing this as we mentioned in Section 2 we choose the initial iterates (X^0, y^0, S^0) as follows:

$$X^0 = S^0 = \zeta I, \quad y^0 = 0, \quad \mu^0 = \zeta^2, \tag{30}$$

where $\zeta > 0$ is such that

$$X^* + S^* \preceq \zeta I, \tag{31}$$

for some $(X^*, y^*, S^*) \in \mathcal{F}^*$. For the moment, let us write

$$(r_b)_i = \theta\nu (r_b^0)_i, \quad i = 1, 2, \dots, m, \quad R_c = \theta\nu R_c^0,$$

and let r_b be the vector $((r_b)_1; (r_b)_2; \dots; (r_b)_m)$. For any two matrices $E (m \times n)$ and $F (p \times q)$ the Kronecker product $E \otimes F$ is the $mp \times nq$ block matrix

$$E \otimes F = \begin{bmatrix} E_{11}F & \dots & E_{1n}F \\ \vdots & \ddots & \vdots \\ E_{m1}F & \dots & E_{mn}F \end{bmatrix}.$$

We recall from [3, 9] some properties of Kronecker product and the operator $\mathbf{vec}(\cdot)$ that are useful for our purpose. These properties are

- (a) $(E \otimes F)^T = E^T \otimes F^T$.
- (b) If E and F are square and nonsingular, then

$$(E \otimes F)^{-1} = E^{-1} \otimes F^{-1}.$$

(c) For any $E (m \times n)$, $F (n \times r)$ and $H (r \times s)$, we have

$$\mathbf{vec}(EHF) = (F^T \otimes E) \mathbf{vec}(H).$$

By using these properties and the definition of the inner product of two matrices, the matrix Q introduced in Lemma 4.9 is the solution of the following system:

$$\begin{aligned} \mathbf{vec}(A_i)^T (D \otimes D) \mathbf{vec}(Q) &= \frac{1}{\sqrt{\mu}} (r_b)_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^m \frac{\xi_i}{\sqrt{\mu}} (D \otimes D) \mathbf{vec}(A_i) + \mathbf{vec}(Q) &= \frac{1}{\sqrt{\mu}} (D \otimes D) \mathbf{vec}(R_c). \end{aligned} \tag{32}$$

Let $\mathcal{A}^T = [\mathbf{vec}(A_1) \quad \mathbf{vec}(A_2) \quad \dots \quad \mathbf{vec}(A_m)]$ and $\xi = (\xi_1; \xi_2; \dots; \xi_m)$. One may easily verify that we can rewrite the system (32) as follows:

$$\begin{aligned} \mathcal{A} (D \otimes D) \mathbf{vec}(Q) &= \frac{1}{\sqrt{\mu}} r_b, \\ (D \otimes D) \mathcal{A}^T \frac{\xi}{\sqrt{\mu}} + \mathbf{vec}(Q) &= \frac{1}{\sqrt{\mu}} (D \otimes D) \mathbf{vec}(R_c). \end{aligned} \tag{33}$$

Lemma 4.10. (Lemma 5.12 in [11]) With (X^0, y^0, S^0) as defined in (30) and (31), we have

$$\|Q\| \leq \theta \sqrt{\nu \mathbf{Tr}(P^2 + P^{-2})}. \tag{34}$$

Lemma 4.11. (Lemma 5.13 in [11]) With (X^0, y^0, S^0) as defined in (30) and (31), we have

$$\|Q\| \leq \frac{\theta}{\zeta \lambda_{\min}(V)} \mathbf{Tr}(X + S). \tag{35}$$

4.3. Some bounds for $\mathbf{Tr}(X + S)$ and $\lambda_{\min}(V)$. The choice of τ and α

Let X be feasible for (P_ν) and (y, S) for (D_ν) . In the same way as in [11], we can rewrite $\delta(V)$ in (6) as follows:

$$4\delta(V)^2 = \|V - V^{-1}\|^2 = \sum_{i=1}^n \left(\lambda_i(V) - \frac{1}{\lambda_i(V)} \right)^2. \tag{36}$$

Using this one easily derives the following result, which we state without further proof.

Lemma 4.12. (Cf. Lemma II.60 in [19]) Let $\delta = \delta(V)$ be given by (36). Then

$$\frac{1}{\rho(\delta)} \leq \lambda_i(V) \leq \rho(\delta), \tag{37}$$

where

$$\rho(\delta) := \delta + \sqrt{1 + \delta^2}. \tag{38}$$

Lemma 4.13. (Lemma 5.15 in [11]) Let X and (y, S) be feasible for the perturbed problems (P_ν) and (D_ν) respectively and let (X^0, y^0, S^0) and $(X^*, y^*, S^*) \in \mathcal{F}^*$ be as defined in (30) and (31). Then we have

$$\nu\zeta \text{Tr}(X + S) \leq S \bullet X + \nu n\zeta^2.$$

Lemma 4.14. (Lemma 5.16 in [11]) Using the same notations as in Lemma 4.13, one has

$$\text{Tr}(X + S) \leq (\rho(\delta)^2 + 1) n\zeta, \tag{39}$$

where $\rho(\delta)$ as defined in (38).

By substituting (37) and (39) into (35) we get

$$\|Q\| \leq n\theta\rho(\delta) (1 + \rho(\delta)^2).$$

At this stage we choose

$$\tau = \frac{1}{8}. \tag{40}$$

Since $\delta \leq \tau = \frac{1}{8}$ and $\rho(\delta)$ is monotonically increasing in δ , we have

$$\|Q\| \leq n\theta\rho(\delta) (1 + \rho(\delta)^2) \leq n\theta\rho\left(\frac{1}{8}\right) \left(1 + \rho\left(\frac{1}{8}\right)^2\right) = 2.586 n\theta. \tag{41}$$

By substituting (37) and (41) into (29), we obtain

$$\begin{aligned} \|Q\|^2 + \left(\|Q\| + \sqrt{4(1-\theta)^2\delta^2 + \theta^2n}\right)^2 &\leq (2.586 n\theta)^2 \\ &+ \left(2.586 n\theta + \sqrt{4(1-\theta)^2\delta^2 + \theta^2n}\right)^2. \end{aligned} \tag{42}$$

We have found that $\delta(v^f) \leq \frac{1}{\sqrt{2}}$ certainly holds if the inequality (29) is satisfied. Then by (42), inequality (29) holds if

$$(2.586 n\theta)^2 + \left(2.586 n\theta + \sqrt{4(1-\theta)^2\delta^2 + \theta^2n}\right)^2 \leq 1.464 (1-\theta).$$

Obviously, the left-hand side of the above inequality is increasing in δ , due to the definition $\rho(\delta) = \delta + \sqrt{1 + \delta^2}$. Using this one may easily verify that if τ is chosen as in (40) and

$$\theta = \frac{1}{4n}, \tag{43}$$

then the above inequality is satisfied for all $n \geq 2$. Then, according to (15), with τ as given, after the feasibility step at most 3 centering steps suffices to get iterates (X^+, y^+, S^+) that satisfy $\delta(X^+, S^+; \mu^+) \leq \tau$.

5. COMPLEXITY ANALYSIS

In the previous sections we have found that if at the start of an iteration the iterates satisfy $\delta(X, S; \mu) \leq \tau$, with τ as defined in (40), then after the feasibility step, with θ as defined in (43), the iterates satisfy $\delta(X^f, S^f; \mu) \leq \frac{1}{\sqrt{2}}$.

According to (15), at most 3 centering steps then suffice to get iterates (X^+, y^+, S^+) that satisfy $\delta(X^+, S^+; \mu^+) \leq \tau$ again. So each main iteration consists of at most 4 so-called inner iterations, in each of which we need to compute a search direction (for either a feasibility step or a centering step). It has become a custom to measure the complexity of an IPM by the required number of inner iterations. In each main iteration both the duality gap and the norms of the residuals are reduced by the factor $1 - \theta$. Hence, $\text{Tr}(X^0 S^0) = n\zeta^2$, the total number of main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max \{n\zeta^2, \|r_b^0\|, \|R_c^0\|\}}{\varepsilon}.$$

Due to (43) we may take

$$\theta = \frac{1}{4n}.$$

Hence the total number of inner iterations is bounded above by

$$16n \log \frac{\max \{n\zeta^2, \|r_b^0\|, \|R_c^0\|\}}{\varepsilon}.$$

Thus we may state without further proof the main result of the paper.

Theorem 5.1. If (P) and (D) have optimal solutions $(X^*, y^*, S^*) \in \mathcal{F}^*$ such that $X^* + S^* \preceq \zeta I$, then after at most

$$16n \log \frac{\max \{n\zeta^2, \|r_b^0\|, \|R_c^0\|\}}{\varepsilon}$$

iterations the algorithm finds an ε -solution of (P) and (D) .

Remark 5.2. The above iteration bound is derived under the assumption that there exists some optimal solutions of (P) and (D) with $(X^*, y^*, S^*) \in \mathcal{F}^*$ such that $X^* + S^* \preceq \zeta I$. One might ask what happens if this is not satisfied. In that case, during the course of the algorithm it may happen that after some main steps the proximity measure δ (after the feasibility step) exceeds $\frac{1}{\sqrt{2}}$, because otherwise there is no reason why the algorithm would not generate an ε -solution. So if this happens it tell us that either the problem (P) and (D) do not have optimal solutions (with zero duality gap) or the value of ζ has been too small. In the latter case one might run the algorithm once more with a larger ζ .

6. CONCLUDING REMARKS

We presented a new IIPM for SDO; each main iteration consists of a feasibility step and three centering steps. Our new feasibility step is more natural, as it targets at μ^+ -center, which results a better iteration bound in compare with [11]. The ideas underling this article can be used to extend the algorithm to second-order cone optimization and also to the symmetric cone optimization.

ACKNOWLEDGEMENTS

The author is indebted to referees for careful reading of the manuscript and for their suggestions which helped to improve the paper. The author also wish to thank Shahrekord University for financial support.

(Received November 7, 2011)

REFERENCES

- [1] F. Alizadeh: Interior-point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.* 5 (2006), 13–51.
- [2] C. Helmberg: Semidefinite Programming for Combinatorial Optimization. Technical Report 00-34, Konrad-Zuse-Zentrum für Informationstechnik Berlin 2000.
- [3] R. A. Horn and C. R. Johnson: Topics in Matrix Analysis. Cambridge University Press, Cambridge 1991.
- [4] E. de Klerk: Aspects of Semidefinite Programming. Kluwer Academic Publishers, Dordrecht 2002.
- [5] 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.
- [6] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise: A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems. *Lecture Notes in Comput. Sci.* 538, Springer, Berlin 1991.
- [7] Z. Q. Luo, J. Sturm, and S. Z. Zhang: Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming. *SIAM J. Optim.* 8 (1998), 59–81.
- [8] I. J. Lustig: Feasible issues in a primal-dual interior point method for linear programming. *Math. Program.* 49 (1990/1991), 145–162.
- [9] H. Lütkepohl: Handbook of Matrices. John Wiley & Sons, Chichester 1996.
- [10] H. Mansouri: Full-Newton Step Interior-Point Methods for Conic Optimization. Ph.D. Thesis, Faculty of Mathematics and Computer Science, TU Delft, 2008.
- [11] H. Mansouri and C. Roos: A new full-Newton step $o(n)$ infeasible interior-point algorithm for semidefinite optimization. *J. Numer. Algorithms* 52 (2009), 2, 225–255.
- [12] H. Mansouri and M. Zangiabadi: New complexity analysis of full Nesterov-Todd steps for semidefinite optimization. *Bull. Iranian Math. Soc.* 1 (2011), 269–286.
- [13] H. Mansouri, T. Siyavash, and M. Zangiabadi: A path-following infeasible interior-point algorithm for semidefinite programming. *Iranian J. Oper. Res.* 3 (2012), 1, 11–30.

- [14] Y. E. Nesterov and M. J. Todd: Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.* 22 (1997), 1, 1–42.
- [15] Y. E. Nesterov and M. J. Todd: Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.* 8 (1998), 2, 324–364.
- [16] J. Peng, C. Roos, and T. Terlaky: Self-regular functions and new search directions for linear and semidefinite optimization. *Math. Program.* 93 (2002), 129–171.
- [17] 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.
- [18] C. Roos: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. *SIAM J. Optim.* 16 (2006), 4, 1110–1136.
- [19] C. Roos, T. Terlaky, and J.-Ph. Vial: *Theory and Algorithms for Linear Optimization. An Interior-Point Approach.* John Wiley and Sons, Chichester 1997. (Second Edition Springer 2005.)
- [20] H. Wolkowicz, R. Saigal, and L. Vandenberghe: *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications.* Kluwer Academic Publishers, Dordrecht 2000.
- [21] Y. Zhang: On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM J. Optim.* 8 (1998), 365–386.

Hossein Mansouri, Department of Applied Mathematics, Faculty of Mathematical Sciences, Shahrekord University, P.O. Box 115, Shahrekord. Iran.

e-mail: Mansouri@sci.sku.ac.ir