Bulletin of the Iranian Mathematical Society Vol. 37 *No. 1* (*2011*), *pp 269-286.*

NEW COMPLEXITY ANALYSIS OF A FULL NESTEROV-TODD STEPS IIPM FOR SEMIDEFINITE OPTIMIZATION

H. MANSOURI* AND M. ZANGIABADI

Communicated by Nezam Mahdavi-Amiri

Abstract. In [H. Mansouri and C. Roos, *Numer. Algorithms* **52** (2009) 225-255.], Mansouri and Ross presented a primal-dual infeasible interior-point algorithm with full-Newton steps whose iteration bound coincides with the best known bound for infeasible interior-point methods. Here, we introduce a slightly different algorithm with a different search direction and show that the same complexity result is obtained using a simpler analysis.

1. **Introduction**

For a comprehensive study of interior-point methods (IPMs), we refer to Roos et al. [9] and deKlerk [1]. Ross in [8], presented a full-Newton step infeasible interior-point algorithm for linear optimization (LO); Later, this algorithm was extended to semidefinite optimization by Mansouri and Ross [5]. Here, we present a slightly different algorithm obtained by changing the definition of the search direction in the algorithm given in [5]. We show that the analysis of the new algorithm is easier than the one for the algorithm in [5], whereas the iteration bound

MSC(2010): Primary: 90C22; Secondary: 90C51.
Keywords: Semidefinite optimization, infeasible interior-point method, primal-dual method, polynomial complexity.
Received: 3 November 2009, Accepted: 29 February 2010.
∗Corresponding author

essentially remains the same.

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

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

and its dual,

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

where, each $A_i, \quad i = 1, \cdots, m$, and C are symmetric, i.e., $A_i,\, C \in \mathbf{S}^n$ and $b \in \mathbf{R}^m$. Furthermore, $X \succeq 0\,(X \succ 0)$ means that X is symmetric and positive semidefinite (symmetric and positive definite). 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 [4, 5] 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 $\mathbf{Tr}\,(XS) = 0$ and the initial iterates are $\left(x^0, y^0, S^0\right) = \zeta\,(I, 0, I)$, where I denotes the identity matrix of size $n \times n$. Using $\mathbf{Tr}\left(X^0 S^0\right) = n\zeta^2$, the total number of iterations for the algorithm in [5] is bounded above by

$$20\, n \log \frac{\max\left\{n\zeta^2, \left\|r_b^0\right\|, \left\|R_c^0\right\|\right\}}{\varepsilon}, \tag{1.1}$$

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

$$\left(r_b^0\right)_i = b_i - A_i \bullet X^0, \quad i = 1, \ldots, m, \tag{1.2}$$

$$R_c^0 = C - \sum_{i=1}^{m} y_i^0 A_i - S^0. \tag{1.3}$$

Up to a constant factor, the iteration bound (1.1) was first obtained by Kojima et al. [3] and Potra and Sheng [7], and it is still the best-known iteration bound for IIPMs.

To describe our aim here, we need to recall the main ideas underlying the algorithm in [5]. For any ν with $0 < \nu \leq 1$, we consider the perturbed problem (P_ν) to be defined by

$$(P_\nu) \qquad \min \qquad \left(C - \nu\left(C - \sum_{i=1}^{m} y_i^0 A_i - S^0\right)\right) \bullet X$$

$$\text{s.t.} \qquad A_i \bullet X = b_i - \nu\left(b_i - A_i \bullet X^0\right), \qquad X \succeq 0,$$

and its dual problem (D_ν) given be

$$(D_\nu) \quad \max \qquad \sum_{i=1}^{m}\left(b_i - \nu\left(b_i - A_i \bullet X^0\right)\right) y_i$$

$$\text{s.t.} \qquad \sum_{i=1}^{m} y_i A_i + S = C - \nu\left(C - \sum_{i=1}^{m} y_i^0 A_i - S^0\right), \ S \succeq 0.$$

Note that if $\nu = 1$, then $X = X^0$ yields a strictly feasible solution of (P_ν), and $(y, S) = (y^0, S^0)$ gives 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 satisfy the well-known interior-point condition (IPC). More generally, one has the following lemma [5, Lemma 4.1].

Lemma 1.1. *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 1.1 that the problems (P_ν) and (D_ν) satisfy the IPC for each $\nu \in (0, 1]$. But then their central paths exist. This means that the system

$$b_i - A_i \bullet X = \nu\left(r_b^0\right)_i, \quad i = 1, 2, \ldots, m, \qquad X \succeq 0 \tag{1.4}$$

$$C - \sum_{i=1}^{m} y_i A_i - S = \nu R_c^0, \quad S \succeq 0 \tag{1.5}$$

$$XS = \mu I,$$

has a unique solution, for every $\mu > 0$. In the sequel, this unique solution is denoted by $(X(\mu, \nu), y(\mu, \nu), S(\mu, \nu))$. 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) is the μ^0-center of (D_1). In other words,

$$\left(X\left(\mu^0, 1\right), y\left(\mu^0, 1\right), S\left(\mu^0, 1\right)\right) = \left(X^0, y^0, S^0\right).$$

In the sequel, we will always have $\mu = \nu\,\mu^0$, and we will accordingly denote $(X(\mu,\,\nu),\,y(\mu,\,\nu),\,S(\mu,\,\nu))$, simply by $(X(\nu),\,y(\nu),\,S(\nu))$.

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:

$$\delta\,(X,\,S,\,\mu) := \delta\,(V) := \tfrac{1}{2}\left\|V^{-1} - V\right\|, \tag{1.6}$$
$$\text{where,}\quad V := \tfrac{1}{\sqrt{\mu}}D^{-1}XD^{-1} = \tfrac{1}{\sqrt{\mu}}DSD.$$

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

$$P := X^{\frac{1}{2}}\left(X^{\frac{1}{2}}SX^{\frac{1}{2}}\right)^{\frac{-1}{2}}X^{\frac{1}{2}} = S^{\frac{-1}{2}}\left(S^{\frac{1}{2}}XS^{\frac{1}{2}}\right)^{\frac{1}{2}}S^{\frac{-1}{2}}, \tag{1.7}$$

which is a symmetric nonsingular matrix. For more details, see [6].

Initially, we have $X = S = \zeta I$ and $\mu = \zeta^2$, where, $V = I$ and

$$\delta\,(X,\,S;\,\mu) = 0.$$

In the sequel, we assume that at the start of each iteration, $\delta\,(X,\,S;\,\mu)$ is smaller than or equal to a (small) threshold $\tau > 0$. So, this is certainly true at the start of the first iteration.

We now describe one iteration of our algorithm. Suppose that for some $\nu \in (0,\,1]$, we have X, y and S satisfying the feasibility conditions (1.4) and (1.5) and such that

$$\mathbf{Tr}\,(XS) = n\mu, \quad\text{and}\quad \delta\,(X,\,S;\,\mu) \le \tau, \tag{1.8}$$

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 (1.4) and (1.5), 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^+}), and their purpose is to get the iterates X^+, y^+ and S^+ such that $\mathbf{Tr}\,(X^+S^+) = n\mu^+$, where, $\mu^+ = \nu^+\zeta^2$ and $\delta\,(X^+,\,S^+;\,\mu^+) \le \tau$. 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 Fig. 1.

Primal-Dual Infeasible IPM

Input:
Accuracy parameter $\varepsilon > 0$;
barrier update parameter θ, $0 < \theta < 1$;
threshold parameter τ, $0 < \tau \leq \frac{1}{\sqrt{2}}$;
$X^0 \succ 0$, $S^0 \succ 0$, $y^0 = 0$ and $\mu^0 > 0$ such that $X^0 S^0 = \mu^0 I$.

begin
$X := X^0$, $S := S^0$, $y := y^0$; $\mu := \mu^0$;
while $\max\left(\mathbf{Tr}\,(XS)\,,\, \|r_b\|\,,\, \|R_c\|\right) \geq \varepsilon$ **do**
begin
feasibility step:
$(X, y, S) := (X, y, S) + \left(\Delta^f X, \Delta^f y, \Delta^f S\right)$;
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

Figure 1. Infeasible Full-Newton-Step Algorithm.

For the feasibility step, in [5] we used the search directions $\Delta^f X$, $\Delta^f y$ and $\Delta^f S$,

$$\mathbf{Tr}\left(A_i \Delta^f X\right) = \theta\nu \left(r_b^0\right)_i, \quad i = 1, \ldots, m, \tag{1.9}$$

$$\sum_{i=1}^{m} \Delta^f y_i A_i + \Delta^f S = \theta\nu R_c^0, \tag{1.10}$$

$$\Delta^f X + P\Delta^f S P^T = \mu S^{-1} - X, \tag{1.11}$$

where, we used the NT-'trick' to symmetrize $\Delta^f X$ with P as defined in (1.7). 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 semidfinite conditions. Assuming that the step $\delta(X, S; \mu) \le \tau$ holds, before hand and by taking θ small enough, it can be guaranteed that after the step, the iterates

$$X^f = X + \Delta^f X, \tag{1.12}$$

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

$$S^f = S + \Delta^f S, \tag{1.13}$$

are semidefinite and moreover $\delta\left(X^f, S^f; \mu^+\right) \le \frac{1}{\sqrt{2}}$, 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 $\delta\left(X^f, S^f; \mu^+\right) \le \frac{1}{\sqrt{2}}$.

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

$$\begin{aligned} A_i \bullet \Delta X &= 0, \quad i = 1, 2, \ldots, m, \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta S &= 0, \\ \Delta X + P\Delta S P^T &= \mu S^{-1} - X, \end{aligned} \tag{1.14}$$

where, matrix P is defined as in (1.7). Denoting the iterates after a centering step as X^+, y^+ and S^+, we recall the following from [1].

Lemma 1.2. *If* $\delta := \delta(X, S; \mu) \leq 1$*, then the primal-dual NT 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 centerting steps serve to get iterates that satisfy $\mathbf{Tr}(XS) = n\mu^+$ and $\delta(X, S; \mu^+) \leq \tau$, where τ is much smaller than $\frac{1}{\sqrt{2}}$. By using Lemma 1.2, the required number of centering steps can easily be obtained. This goes as follows. After μ-update, we have $\delta\left(X^f, S^f; \mu^+\right) \leq \frac{1}{\sqrt{2}}$, and hence after k centering steps, the iterates (X, y, S) satisfy

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

Just as in [5], this implies that no more than

$$\log_2\left(\log_2 \frac{1}{\tau^2}\right) \tag{1.15}$$

centering steps are needed.

Having described the approach taken in [5], we are now able to explain our aim here. We present a slightly different algorithm which is obtained by changing the definition of the feasibility step, replacing (1.11) by

$$\Delta^f X + P\Delta^f SP^T = 0. \tag{1.16}$$

As we will see, this simplifies the analysis of our algorithm, whereas the iteration bound essentially remains the same.

The rest of the paper is organized as follows. Section 2 is devoted to the analysis of the new feasibility step, which is the main part of our work. We will see that the new search direction requires a different analysis, but at some places we can use the results obtained in [5]. In such cases, we will cite these results without repeating their proofs. The final iteration bound is derived in Section 3. The concluding remarks are given in Section 4.

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 vectors and the set of positive vectors, respectively. For any $x = (x_1; x_2; \ldots; x_n) \in \mathbf{R}^n$, $x_{\min} = \min(x_1; x_2; \ldots; x_n)$ and $x_{\max} = \max(x_1; x_2; \ldots; x_n)$.

$\mathbf{R}^{m\times n}$ 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 the primal and dual problem respectively. The relative interior of a convex set $\mathcal{C}$ is denoted as $\mathrm{ri}\,(\mathcal{C})$. I denotes the $n \times n$ 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 $A \bullet B = \mathbf{Tr}\left(A^T B\right)$. 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 symmetric matrix G, $\lambda_{\min}(G)$ $(\lambda_{\max}(G))$ denotes the minimal (maximal) eigenvalue of G. When λ is a vector, we denote the diagonal matrix $\mathrm{diag}\,(\lambda)$ 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, $\lambda_{\max}(V) = \lambda_1(V) \geq \lambda_2(V) \geq \ldots \geq \lambda_n(V) = \lambda_{\min}(V)$. The Frobenius matrix norm is given by $\|U\|^2 := \sum_{i=1}^m \sum_{j=1}^n U_{ij}^2 = \mathbf{Tr}\left(U^T U\right)$.

For any $p \times q$ matrix A, $\mathbf{vec}\,(A)$ 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 [2] for a comprehensive treatment of Kronecker products and related topics).

2. Analysis of the Feasibility Step

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

$$\delta\left(X^0, S^0, \mu^0\right) = 0.$$

Before dealing with the analysis of the algorithm, we recall a lemma that will be needed.

Lemma 2.1 (Lemma A.1 in [1]). *Let $Q \in \mathbf{S}^n_{++}$, and let $M \in \mathbf{R}^{n\times n}$ be skew-symmetric (i.e., $M = -M^T$). Then, $\det(Q+M) > 0$. Moreover, if the eigenvalues of $Q + M$ are real, then*

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

2.1. **The feasibility step and the choice of τ and θ.**

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

We introduce scaled versions of the search directions $\Delta^f X$ and $\Delta^f S$ as follows:

$$D_X^f := \tfrac{1}{\sqrt{\mu}} D^{-1}\Delta^f X D^{-1},\; D_S^f := \tfrac{1}{\sqrt{\mu}} D\Delta^f S D,$$
$$\left(V^f\right)^2 := \tfrac{1}{\mu^+} D^{-1} X^f S^f D, \tag{2.1}$$

with D as defined in Section 1. By using (2.1) and replacing (1.11) by (4.1), we can rewrite (1.9)-(1.11) as follows:

$$\begin{aligned} \mathbf{Tr}\left(DA_iDD_X^f\right) &= \frac{1}{\sqrt{\mu}}\theta\nu\left(r_b^0\right)_i, \quad i = 1,\ldots,m, \\ \sum_{i=1}^m \frac{\Delta^f y_i}{\sqrt{\mu}} DA_iD + D_S^f &= \frac{1}{\sqrt{\mu}}\theta\nu D R_c^0 D, \\ D_X^f + D_S^f &= 0. \end{aligned} \tag{2.2}$$

From the third equation in (2.2), we obtain, by multiplying both sides on the left with V,

$$VD_X^f + VD_S^f = 0. \tag{2.3}$$

Using (1.6), (1.12), (1.13) and (2.1), we obtain

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

Therefore,

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

The last equality shows that the matrix $X^f S^f$ is similar to

$$\mu\left(V + D_X^f\right)\left(V + D_S^f\right).$$

This means that we have

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

To simplify the notation, in the sequel we use

$$D_{XS}^f := \frac{1}{2}\left(D_X^f D_S^f + D_S^f D_X^f\right), \tag{2.4}$$

and

$$M := \left(D_X^f V - V D_X^f\right) + \frac{1}{2}\left(D_X^f D_S^f - D_S^f D_X^f\right). \tag{2.5}$$

Note that D_{XS}^f is symmetric and M is skew-symmetric. Now, we may write, using (2.3),

$$\begin{aligned}\left(V + D_X^f\right)\left(V + D_S^f\right) &= V^2 + V D_S^f + D_X^f V + D_X^f D_S^f \\ &= V^2 - V D_X^f + D_X^f V + D_X^f D_S^f.\end{aligned}$$

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

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

Using (2.4) and (2.5), we obtain

$$X^f S^f \sim \mu\left(V^2 + D_{XS}^f + M\right). \tag{2.6}$$

Lemma 2.2. *Let $X \succ 0$ and $S \succ 0$. Then, the iterates $\left(X^f, y^f, S^f\right)$ are strictly feasible if*

$$V^2 + D_{XS}^f \succ 0.$$

Proof. The proof is similar to the proof of Lemma 4.4 in [5]. □

By the Rayleigh-Ritz theorem (see [2]), we easily have the following result.

Lemma 2.3. *Let A, $B \in \mathbf{S}^n$. One has*

$$\lambda_i(A + B) \geq \lambda_{\min}(A) - |\lambda_{\max}(B)|, \quad i = 1, \cdots, n,$$

where, $\lambda_{\min}(A)$ denotes the smallest eigenvalue of A and $\lambda_{\max}(B)$ is the largest eigenvalue of B.

Proof. By the Rayleigh-Ritz theorem, there exists $x_0 \in \mathbf{R}^n$ such that

$$\begin{aligned}\lambda_i(A+B) &\geq \lambda_{\min}(A+B) = \frac{x_0^T(A+B)x_0}{x_0^Tx_0} = \frac{x_0^TAx_0}{x_0^Tx_0} + \frac{x_0^TBx_0}{x_0^Tx_0} \\ &\geq \frac{x_0^TAx_0}{x_0^Tx_0} - \left|\frac{x_0^TBx_0}{x_0^Tx_0}\right| \geq \min_{x\neq 0}\frac{x^TAx}{x^Tx} - \max_{x\neq 0}\left|\frac{x^TBx}{x^Tx}\right| \\ &= \lambda_{\min}(A) - |\lambda_{\max}(B)|.\end{aligned}$$

This completes the proof. □

From the third equation in (2.2), we have $D_S^f = -D_X^f$, and therefore by replacing in (2.4), we have

$$D_{XS}^f := \frac{1}{2}\left(D_X^fD_S^f + D_S^fD_X^f\right) = -\left(D_X^f\right)^2. \tag{2.7}$$

The equation (2.7) and lemmas 2.2 and 2.3 imply the following result.

Lemma 2.4. *The iterates* $(X^f,\, y^f,\, S^f)$ *are strictly feasible if*

$$\left|\lambda_i\left(D_X^f\right)\right| \leq \lambda_{\min}(V), \quad i = 1, \cdots, n.$$

We continue this section by recalling a lemma from [9] that is crucial in the analysis of the algorithm.

Lemma 2.5 (cf. Lemma II.60 in [9])**.** *Let* $\delta = \delta(V)$ *be given by (1.6). Then,*

$$\frac{1}{\rho(\delta)} \leq \lambda_i(V) \leq \rho(\delta), \tag{2.8}$$

where,

$$\rho(\delta) := \delta + \sqrt{1+\delta^2}. \tag{2.9}$$

Assuming $V^2 + D_{XS}^f \succ 0$, which according to Lemma 2.4 implies that the iterates $(X^f,\, y^f,\, S^f)$ are strictly feasible, we proceed by deriving an upper bound for $\delta(X^f,\, S^f;\, \mu^+)$. According to Definition 1.6, we have

$$\begin{aligned}&\delta\left(X^f,\, S^f;\, \mu^+\right) := \tfrac{1}{2}\left\|V^f - \left(V^f\right)^{-1}\right\|, \\ &\text{where,}\quad V^f = \tfrac{1}{\sqrt{\mu^+}}D^{-1}X^fD^{-1} = \tfrac{1}{\sqrt{\mu^+}}DS^fD.\end{aligned} \tag{2.10}$$

In the sequel, we also denote $\delta(X^f,\, S^f;\, \mu^+)$ by $\delta(V^f)$. We need some technical results which give information on the eigenvalues and the norm of V^f.

Lemma 2.6. *One has*

$$\lambda_{\min}\left(\left(V^f\right)^2\right) \geq \frac{1}{1-\theta}\left(\frac{1}{\rho(\delta)^2} - \left\|V^2 + \left(D_X^f\right)^2\right\|\right).$$

Proof. Using (2.6), after division of both sides into $\mu^+ = (1-\theta)\mu$, we get

$$\left(V^f\right)^2 \sim \frac{\mu\left(V^2 + D_{XS}^f + M\right)}{\mu^+} = \frac{V^2 + D_{XS}^f + M}{1-\theta}. \tag{2.11}$$

It follows that

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

Since M is skew-symmetric, lemmas 2.1 and 2.3 and (2.7) imply

$$\begin{aligned}\lambda_{\min}\left(\left(V^f\right)^2\right) &\geq \frac{1}{1-\theta}\lambda_{\min}\left(V^2 + D_{XS}^f\right) \\ &\geq \frac{1}{1-\theta}\left(\lambda_{\min}\left(V^2\right) - \left|\lambda_{\max}\left(\left(D_X^f\right)^2\right)\right|\right).\end{aligned}$$

By using Lemma 2.5, we easily obtain

$$\lambda_{\min}\left(\left(V^f\right)^2\right) \geq \frac{1}{1-\theta}\left(\frac{1}{\rho(\delta)^2} - \left\|\left(D_X^f\right)^2\right\|\right).$$

Since V^2 is a positive semidefinte matrix, we have

$$\lambda_{\min}\left(\left(V^f\right)^2\right) \geq \frac{1}{1-\theta}\left(\frac{1}{\rho(\delta)^2} - \left\|V^2 + \left(D_X^f\right)^2\right\|\right).$$

This completes the proof. □

Lemma 2.7. *One has*

$$\left\|I - \left(V^f\right)^2\right\| \leq \frac{1}{1-\theta}\left(\sqrt{n}(\theta - 1) + \left\|V^2 + \left(D_X^f\right)^2\right\|\right).$$

Proof. The proof is the same as the proof of Lemma 4.8 in [5]. □

Using (2.10) and lemmas 2.6 and 2.7, we have the following result.

Lemma 2.8. *One has*

$$2\delta\left(V^f\right) \leq \frac{\rho(\delta)\left(\sqrt{n}\,(\theta-1)+\left\|V^2+\left(D_X^f\right)^2\right\|\right)}{\sqrt{(1-\theta)\left(1-\rho(\delta)^2\left\|V^2+\left(D_X^f\right)^2\right\|\right)}}.$$

We conclude this section by presenting a value that we do not allow $\left\|V^2+\left(D_X^f\right)^2\right\|$ to exceed. Since we need to have $\delta\left(V^f\right)\leq\frac{1}{\sqrt{2}}$, it follows from Lemma 2.8 that it suffices to have

$$\frac{\rho(\delta)\left(\sqrt{n}\,(\theta-1)+\left\|V^2+\left(D_X^f\right)^2\right\|\right)}{\sqrt{(1-\theta)\left(1-\rho(\delta)^2\left\|V^2+\left(D_X^f\right)^2\right\|\right)}} \leq \sqrt{2}.$$

At this stage, we decide to choose

$$\tau=\frac{1}{8},\quad \theta=\frac{\alpha}{2\sqrt{n}},\quad \alpha\leq 1. \tag{2.12}$$

Then, for $n\geq 1$ and $\delta\leq\tau$, one may verify that

$$\left\|V^2+\left(D_X^f\right)^2\right\|\leq 1 \Rightarrow \delta\left(V^f\right)\leq\frac{1}{\sqrt{2}}. \tag{2.13}$$

Since $\left\|V^2+\left(D_X^f\right)^2\right\|\leq\left\|V^2\right\|+\left\|\left(D_X^f\right)^2\right\|$, it is possible to replace (2.13) by a weaker condition as follows:

$$\left\|V^2\right\|+\left\|\left(D_X^f\right)^2\right\|\leq 1 \Rightarrow \delta\left(V^f\right)\leq\frac{1}{\sqrt{2}}. \tag{2.14}$$

Using (2.14), we find out that in order to have $\delta\left(V^f\right)\leq\frac{1}{\sqrt{2}}$, we should have $\left\|V^2\right\|+\left\|\left(D_X^f\right)^2\right\|\leq 1$. Therefore, since $\left\|\left(D_X^f\right)^2\right\|\leq\left\|V^2\right\|$, it suffices to have V^2 satisfy $\left\|V^2\right\|\leq\frac{1}{2}$. So, we have $\delta\left(V^f\right)\leq\frac{1}{\sqrt{2}}$ if $\left|\lambda_i\left(D_X^f\right)\right|\leq\frac{1}{\sqrt{2}}$, for all $i=1,\cdots,n$. We proceed by considering the vector D_X^f in more detail.

2.2. **An upper bound for** $\left\|D_X^f\right\|$.

It is clear from system (2.2) that D_X^f is the unique solution of the system

$$\begin{aligned} \mathbf{Tr}\left(DA_iDD_X^f\right) &= \frac{1}{\sqrt{\mu}}\theta\nu\left(r_b^0\right)_i, \quad i = 1, \ldots, m, \\ \sum_{i=1}^{m}\frac{\Delta^f y_i}{\sqrt{\mu}}DA_iD + D_S^f &= \frac{1}{\sqrt{\mu}}\theta\nu DR_c^0 D. \end{aligned}$$

To derive an upper bound for $\left\|D_X^f\right\|$, we recall a result from [5]. There, we proved that if the matrix Q satisfies

$$\begin{aligned} \mathbf{Tr}\left(DA_iDQ\right) &= \frac{1}{\sqrt{\mu}}\theta\nu\left(r_b^0\right)_i, \quad i = 1, \ldots, m, \\ \sum_{i=1}^{m}\frac{\Delta^f y_i}{\sqrt{\mu}}DA_iD + Q &= \frac{1}{\sqrt{\mu}}\theta\nu DR_c^0 D, \end{aligned}$$

then it follows that

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

By using almost the same arguments, we also have

$$\left\|D_X^f\right\| \leq \frac{\theta}{\zeta\lambda_{\min}(V)}\mathbf{Tr}\left(X+S\right).$$

This inequality implies that, for all $i = 1, \cdots, n$,

$$\left|\lambda_i\left(D_X^f\right)\right| \leq \frac{\theta}{\zeta\lambda_{\min}(V)}\mathbf{Tr}\left(X+S\right). \tag{2.15}$$

2.3. **Bounds for** $\mathbf{Tr}\,(X+S)$ **and** $\lambda_{\min}(V)$**: The choice of** τ **and** α**.**

Recall that X is feasible for (P_ν) and (y, S) is feasible for (D_ν), and $\delta(X, S;\, \mu) \leq \tau$. We need to find an upper bound for $\mathbf{Tr}\,(X+S)$ and a lower bound for the eigenvalues of V. From Lemma 2.5, we have

$$\frac{1}{\rho(\delta)} \leq \lambda_i(V) \leq \rho(\delta). \tag{2.16}$$

Using Lemma 4.16 in [5], we get

$$\mathbf{Tr}\,(X+S) \le \left(\rho\,(\delta)^2+1\right) n\zeta.$$

By substituting into (2.15), we get

$$\left|\lambda_i\left(D_X^f\right)\right| \le n\theta\rho\,(\delta)\left(1+\rho\,(\delta)^2\right), \quad \text{for all} \quad i=1,\cdots,n.$$

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

$$\left|\lambda_i\left(D_X^f\right)\right| \le n\theta\rho\,(\delta)\left(1+\rho\,(\delta)^2\right) \le n\theta\rho\left(\frac{1}{8}\right)\left(1+\rho\left(\frac{1}{8}\right)^2\right) = 2.586n\theta.$$

By using $\theta = \frac{\alpha}{2\sqrt{n}}$, we obtain the following upper bound for $\left|\lambda_i\left(D_X^f\right)\right|$:

$$\left|\lambda_i\left(D_X^f\right)\right| \le \frac{2.586n\alpha}{2\sqrt{n}}.$$

In Section 2.1, we found that in order to have $\delta\left(V^f\right) \le \frac{1}{\sqrt{2}}$, we should have $\left|\lambda_i\left(D_X^f\right)\right| \le \frac{1}{\sqrt{2}}$. Since $\left|\lambda_i\left(D_X^f\right)\right| \le \frac{2.586n\alpha}{2\sqrt{n}}$, the latter inequality is satisfied if we take

$$\alpha = \frac{1}{2\sqrt{n}}, \tag{2.17}$$

because

$$\frac{\sqrt{2}}{2.586} = 0.546 \ge \frac{1}{2}.$$

3. **Complexity**

As already shown in Section 1, with τ as defined in (2.12), according to (1.15) we need at most

$$\log_2\left(\log_2\frac{1}{\tau^2}\right) = \log_2\left(\log_2 64\right) \le 3$$

centering steps to get iterates satisfing $\delta\,(X,\,S;\,\mu^+) \le \tau$. So, each iteration consists of one feasibility step and at most 3 centering steps. In each iteration, both the duality gap and the norms of the residual vectors are reduced by the factor $1-\theta$. Hence, using $\mathbf{Tr}\left(X^0S^0\right) = n\zeta^2$, the total number of iterations is bounded above by

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

Due to (2.12) and (2.17), we have

$$\theta = \frac{\alpha}{2\sqrt{n}} = \frac{1}{4n}.$$

Hence, the total number of inner iterations is bounded above by

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

We have the following main result.

Lemma 3.1. *If* (P) *and* (D) *have optimal solutions* $(X^*,\, y^*,\, S^*)$ *such that* $X^* + S^* \preceq \zeta I$*, then after at most*

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

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

The above theorem gives a convergence result under the assumption that (P) and (D) have optimal solutions (X^*, y^*, S^*), with zero duality gap, and such that the eigenvalues of X^* and S^* do not exceed ζ. One might ask what happens if this condition is not satisfied.

Our analysis of the algorithm has made clear that as long as we have $\delta(X^f, S^f; \mu^+) \le 1/\sqrt{2}$ after each feasibility step, then the algorithm will generate an ε-solution of (P) and (D), and the number of iterations will be as given by Theorem 3.1. So, if during the execution of the algorithm it happens that after the feasibility step, $\delta(X^f, S^f; \mu^+) > 1/\sqrt{2}$, then we must conclude that there exists no optimal solution $(X^*,\, y^*,\, S^*)$ with zero duality gap such that the eigenvalues of X^* and S^* do not exceed ζ. In that case, one might rerun the algorithm with larger values of ζ. If this does not help, then eventually one should realize that (P) and/or (D) do not have optimal solutions at all, or they have optimal solutions with positive duality gaps.

4. Concluding Remarks

We presented a slightly different algorithm from our previously proposed one, obtained by changing the definition of the feasibility step. The new step is defined by

$$\Delta^f X + P\Delta^f S P^T = 0,$$

whereas, the feasibility step in [5] was defined by

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

There is one more candidate for the definition of this step, namely,

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

It is certainly worthwhile to analyze a full-Newton step method based on this candidate search direction. Other interesting endeavors are the extension of the ideas discussed here to second-order cone optimization and symmetric cone optimization.

Acknowledgments

We thank the referee for his/her useful comments and we also thank Shahrekord University for financial support.

References

[1] E. de Klerk, *Aspects of Semidefinite Programming*, Applied Optimization, 65, Kluwer Academic Publishers, Dordrecht, 2002.

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

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

[4] H. Mansouri, *Full-Newton Step Interior-Point Methods for Conic Optimization,* Ph.D. thesis, Faculty of Mathematics and Computer Science, TU Delft, NL–2628 CD Delft, The Netherlands, June 2008.

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

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

[7] 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.

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

[9] C. Roos, T. Terlaky, and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Ltd., Chichester, 1997.

H. Mansouri
Department of Applied Mathematics, Shahrekord University, P.O.Box 115, Shahrekord, Iran
Email:H.Mansouri@tudelft.nl and mansouri@sci.sku.ac.ir

M. Zangiabadi
Department of Applied Mathematics, Shahrekord University, P.O. Box 115, Shahrekord, Iran
Email:M.Zangiabadi@tudelft.nl and zangiabadi@sci.sku.ac.ir