A PREDICTOR CORRECTOR METHOD FOR SEMI-DEFINITE LINEAR PROGRAMMING

Chih-Jen Lin

and

Romesh Saigal

Department of Industrial & Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109-2117

Technical Report 95-20

October 1995
Revised December 1995

A Predictor Corrector Method for Semi-definite Linear Programming*

Chih-Jen LIN

Romesh SAIGAL

Department of Industrial and Operations Engineering,

The University of Michigan,

Ann Arbor, Michigan 48109-2117, USA

October 30, 1995 (revised December 10, 1995)

Abstract

In this paper we present a generalization of the predictor corrector method of linear programming problem to semidefinite linear programming problem. We consider a direction which, we show, belongs to a family of directions presented by Kojima, Shindoh and Hara, and, one of the directions analyzed by Monteiro. We show that starting with the initial complementary slackness violation of t_0, in $O(|\log(\frac{\epsilon}{t_0})|\sqrt{n})$ iterations of the predictor corrector method, the complementary slackness violation can be reduced to less than or equal to $\epsilon > 0$. We also analyze a modified corrector direction in which the linear system to be solved differs from that of the predictor in only the right hand side, and obtain a similar bound. We then use this modified corrector step in an implementable method which is shown to take a total of $O(|\log(\frac{\epsilon}{t_0})|\sqrt{n}\log(n))$ predictor and corrector steps.

Key words: Linear programming, Semidefinite programming, Interior point methods, Path following, Predictor corrector method.

*Supported by the grant number CCR–9321550 from the National Science Foundation.

Abbreviated title: Semi-definite linear programming

1 Introduction

This paper considers the dual pair of semidefinite linear programs:

$$\begin{array}{rrcll} \text{minimize} & C \bullet X & & & \\ & A_i \bullet X & = & b_i & \text{for every } i = 1, \cdots, m \\ & X \succeq 0 & & & \end{array} \tag{1}$$

and,

$$\begin{array}{rrcccl} \text{minimize} & \sum_{i=1}^{m} b_i y_i & & & & \\ & \sum_{i=1}^{m} A_i y_i & + & S & = & C \\ & & & S \succeq 0 & & \end{array} \tag{2}$$

where A_i for $i = 1, \cdots, m$ and C are $n \times n$ symmetric matrices, $A \bullet B = \text{trace}(A^T B)$ and $X \succeq 0$ means that X is a symmetric and positive semidefinite matrix while $X \succ 0$ means that it is a symmetric and positive definite matrix.

There is considerable interest in generating algorithms, and, understanding the duality, optimality conditions and the facial structure of the feasible region of the semidefinite linear programming problem. The landmark work in the former area is the paper of Nesterov and Nemirovskii [12] and their book [13], where they present a general theory based on self concordant barrier functions, and their relation to interior point methods for convex programs. Some references for the later are Alizadeh, Haeberly and Overton [2], Pataki[15] and Ramana, Tunçel and Wolkowicz [17]. Also, several extensions of the primal-dual potential reduction methods to semidefinite programming have been made and some notable papers in this regard are [1, 6, 7, 21].

There is also recent success in extending the primal-dual path following algorithms of linear programming to semidefinite linear programming. In the linear programming situation, the scaling matrices involved are diagonal, and can thus commute. The matrices involved in semidefinite programming do not necessary commute and thus the computed direction may not be symmetric. One of the main variations in these algorithms is the process adapted to achieve symmetry. Nesterov and Todd [14] and Kojima, Shindoh and Hara [8] have developed some of the earlier ones. Recently Monteiro [11] has presented algorithms, that follow a neighborhood of the central path, and, which differ in the direction adapted.

He considered the two directions determined by the equations

$$X^{-\frac{1}{2}}(X\Delta S + \Delta X S)X^{\frac{1}{2}} + X^{\frac{1}{2}}(\Delta S X + S\Delta X)X^{-\frac{1}{2}} = 2(\sigma\mu I - X^{\frac{1}{2}}SX^{\frac{1}{2}}) \quad (3)$$

$$S^{\frac{1}{2}}(X\Delta S + \Delta X S)S^{-\frac{1}{2}} + S^{-\frac{1}{2}}(\Delta S X + S\Delta X)S^{\frac{1}{2}} = 2(\sigma\mu I - S^{\frac{1}{2}}XS^{\frac{1}{2}}) \quad (4)$$

from a family presented by Kojima, Shindoh and Hara [8]. The direction determined from these systems is unique and symmetric. He analyzed algorithms using these two directions. Based on his analysis, Zhang[22] subsequently analyzed algorithms based on the direction (4). For this direction, Zhang also presented a result on the generalization of the predictor corrector method of Mizuno, Todd and Ye [10].

In this paper we present a feasible start predictor-corrector method based on the talk of Saigal [18] which presented a similar result for the linear programming problem (see also section 5.10 of the book Saigal[19]). Here we assume that the primal and the dual problems have interior feasible points, and thus the strong duality theorem holds, see for example, Alizadeh [1]. Freund [3] has presented an algorithm without such an assumption. Recently, Potra and Sheng [16] and Kojima, Shida, and Shindoh [9] have presented an infeasible start predictor corrector method with some local convergence analysis.

The direction we use is determined by the equation

$$X\Delta S + \Delta X S = \sigma\mu I - XS \quad (5)$$

and the direction used in the method is $(\frac{1}{2}(\Delta X + \Delta X^T), \Delta y, \Delta S)$ (it can be shown that ΔS is symmetric). A direction computed by such a symmetrization has been used by Helmberg, Rendl, Vanderbei and Wolkowicz [6]. We can show that the direction given by (5) also solves (4). Substituting this direction into (4), we obtain the identity below

$$\begin{aligned}
& S^{\frac{1}{2}}(X\Delta S + \frac{\Delta X + \Delta X^T}{2}S)S^{-\frac{1}{2}} + S^{-\frac{1}{2}}(\Delta S X + S\frac{\Delta X + \Delta X^T}{2})S^{\frac{1}{2}} \\
= \; & S^{\frac{1}{2}}(\frac{X\Delta S + \Delta X S}{2})S^{-\frac{1}{2}} + S^{-\frac{1}{2}}(\frac{S\Delta X^T + \Delta S X}{2})S^{\frac{1}{2}} + S^{\frac{1}{2}}(\frac{X\Delta S + \Delta X^T S}{2})S^{-\frac{1}{2}} \\
& + S^{-\frac{1}{2}}(\frac{\Delta S X + S\Delta X}{2})S^{\frac{1}{2}} \\
= \; & (\sigma\mu I - S^{\frac{1}{2}}XS^{\frac{1}{2}}) + S^{\frac{1}{2}}(\frac{X\Delta S + \Delta X^T S}{2})S^{-\frac{1}{2}} + S^{-\frac{1}{2}}(\frac{S\Delta X^T + \Delta S X}{2})S^{\frac{1}{2}} \\
= \; & 2(\sigma\mu I - S^{\frac{1}{2}}XS^{\frac{1}{2}})
\end{aligned}$$

and we see that (4) is satisfied. We use the same neighborhood of the central trajectory as one used by many papers, including Monteiro [11] and Zhang [22]. This is

$$N(t_0,\beta) \;=\; \{(X,y,S): X \succ 0, S \succ 0, t \leq t_0, \|S^{\frac{1}{2}}XS^{\frac{1}{2}} - tI\|_F \leq \beta t\} \tag{6}$$

where $\|B\|_F^2 = \sum_{i=1}^m \sum_{j=1}^m B_{i,j}^2$ is the Forbenius norm of the matrix B and $\beta > 0$ is a constant. We obtain the result that starting with initial complementary slackness violation of $t_0 > 0$, it can be reduced to $\epsilon > 0$ in $O(|\log(\frac{\epsilon}{t_0})|\sqrt{n})$ iterations of the predictor corrector method as considered by Saigal [18].. Such a result is also asserted in Zhang [22] for the generalization of the Mizuno, Todd and Ye [10] method. In another direction, we extend this result to a modified algorithm in which the corrector step solves a system of linear equations that differs from that of the predictor only in the right hand side. In addition, we present a variant of this method that achieves larger step sizes and solves the problem in at most $O(|\log(\frac{\epsilon}{t_0})|\sqrt{n}\log(n))$ sum total of iterations of the predictor and the corrector steps.

This paper is organized as follows. In section 2 we present the method. In section 3 we present some basic lemmas. In section 4 we analyze the predictor and corrector steps and obtain the main result. In section 5 we present the modified corrector step and analyze it. In section 6 we discuss some implementation issues.

2 The Method

We now present the predictor corrector method we will discuss in this paper.

Step 0 Let $0 < \alpha < 1$ and $0 < \beta < 1$ be given constants and $X_0 \succ 0$, y^0 and $S_0 \succ 0$ satisfy the dual systems (1) and (2) respectively. Also, assume they belong to $N(t_0,\beta)$ for $t_0 = \frac{X_0 \bullet S_0}{n}$. Set $k = 0$.

Step 1 Predictor Step: Solve

$$\begin{array}{lcrcl} A_i \bullet \Delta X_k & & & = & 0 \quad \text{for all } i = 1,\cdots,m \\ & \sum_{i=1}^m A_i \Delta y_i^k & + \;\; \Delta S_k & = & 0 \\ \Delta X_k S_k & & + \;\; X_k \Delta S_k & = & -X_k S_k \end{array} \tag{7}$$

for $(\Delta X_k, \Delta y^k, \Delta S_k)$, and define

$$\begin{aligned} \bar{X}_k &= X_k + \tfrac{1}{2}\alpha(\Delta X_k + \Delta X_k^T) \\ \bar{y}^k &= y^k + \alpha \Delta y^k \\ \bar{S}_k &= S_k + \alpha \Delta S_k \\ \bar{t}_k &= \tfrac{\bar{X}_k \bullet \bar{S}_k}{n}. \end{aligned} \tag{8}$$

Step 2 Corrector Step: Solve

$$\begin{array}{llll} A_i \bullet \Delta \bar{X}_k & & = 0 & \text{for all } i = 1, \cdots, m \\ & \sum_{i=1}^m A_i \Delta \bar{y}_i^k + \Delta \bar{S}_k & = 0 & \\ \Delta \bar{X}_k \bar{S}_k & + \bar{X}_k \Delta \bar{S}_k & = \bar{t}_k I - \bar{X}_k \bar{S}_k & \end{array} \tag{9}$$

for $(\Delta \bar{X}_k, \Delta \bar{y}^k, \Delta \bar{S}_k)$, and define

$$\begin{aligned} X_{k+1} &= \bar{X}_k + \tfrac{1}{2}(\Delta \bar{X}_k + \Delta \bar{X}_k^T). \\ y^{k+1} &= \bar{y}^k + \Delta \bar{y}^k \\ S_{k+1} &= \bar{S}_k + \Delta \bar{S}_k \\ t_{k+1} &= \tfrac{X_{k+1} \bullet S_{k+1}}{n}. \end{aligned} \tag{10}$$

Step 3 Set $k = k + 1$, and go to step 1

Some comments are in order here. Both the systems (7) and (9) have a unique solution. Also, the solution ΔS_k and $\Delta \bar{S}_k$ are readily seen as symmetric (a consequence of our assumptions on the matrices of the problem), but ΔX_k or $\Delta \bar{X}_k$ need not be symmetric. Thus we use the symmetric part of these directions in the predictor or corrector steps.

3 Basic Results

We present here the basic results we need about matrices and norms. Given an $n \times n$ matrix B we define its 2 - norm as $\|B\|_2 = \max_{\|x\|=1} \|Bx\|_2$, and it is easily seen that if the matrix B is symmetric, then $\|B\|_2 = \max|\lambda_i|$ where λ_i for $i = 1, \cdots, n$ are the n real eigenvalues of B. For a given $n \times n$ matrix B, we define $\text{vec}(B) = (B_{.1}^T, B_{.2}^T, \ldots, B_{.n}^T)^T$ and note that $\|\text{vec}(B)\|_2 = \|B\|_F$, where $B_{.j}$ is the jth column of the matrix B.

Lemma 1 *Let A and B be arbitrary $n \times n$ matrices, with B nonsingular. Then $\|AB\|_F^2 \le \|A\|_F^2 \|B^T B\|_2$ and $\|AB\|_F^2 \le \|A^T A\|_2 \|B\|_F^2$.*

Proof: Note that $AB = (AB_{.1}, \cdots, AB_{.n})$ where $B_{.j}$ is the jth column of B. Thus $\|AB\|_F^2 = \sum_{j=1}^n \|AB_{.j}\|_2^2 \le \|A\|_2^2 \sum_{j=1}^n \|B_{.j}\|_2^2 = \|A\|_2^2 \|B\|_F^2 \le \|A^T A\|_2 \|B\|_F^2$. The last inequality follows from Theorem 2.3.1 of Golub and Van Loan [5]. To see the other inequality, note that $AB = ((A_{1.}B)^T, \cdots, (A_{n.}B)^T)^T$ where $A_{j.}$ is the jth row of A, and $\|AB\|_F^2 = \sum_{j=1}^n \|A_{j.}B\|_2^2 \le (\sum_{j=1}^n \|A_{j.}\|_2^2)\|B^T\|_2^2 = \|A\|_F^2 \|B^T B\|_2$. (Note, we have used here the fact that for non-singular B, $\|B^T\|_2^2 = \lambda_{\max}(BB^T) = \lambda_{\max}(B^T B B^T B^{-T}) = \lambda_{\max}(B^T B) = \|B\|_2^2$). ■

For a given $n \times n$ matrix B which has all eigenvalues real, we define $\lambda_i(B)$ for $i = 1, \cdots, n$ the n real eigenvalues of B, and $\lambda_{\max}(B) = \max_i \lambda_i(B)$ and similarly $\lambda_{\min}(B)$. We can prove

Lemma 2 *Let A, B and C be $n \times n$ symmetric matrices with $A = B + C$. Then $\lambda_{\min}(A) \ge \lambda_{\min}(B) - \|C\|_F$ and $\lambda_{\max}(A) \le \lambda_{\max}(B) + \|C\|_F$.*

Proof: Follows readily by noting that $\|C\|_F^2 = \text{trace}(C^2) = \sum_{i=1}^n \lambda_i(C)^2 \ge \lambda_{\min}(C)^2$ and $\lambda_{\max}(C)^2 \le \|C\|_F^2$. ■

The following result was proved by Monteiro [11].

Lemma 3 *Let A and B be $n \times n$ matrices with A symmetric and B nonsingular. Then*

$$\|A\|_F \le \frac{1}{2} \|BAB^{-1} + (BAB^{-1})^T\|_F.$$

Another technical lemma follows:

Lemma 4 *Let A, B and C be $n \times n$ matrices such that $A = B + C$, and trace$(B^T C) \ge 0$. Then $\|B\|_F \le \|A\|_F$ and $\|C\|_F \le \|A\|_F$.*

Proof: First note that $\|A\|_F = \|\text{vec}(A)\|_2$. Then, as $\text{trace}(B^T C) = \text{vec}(B)^T \text{vec}(C) \ge 0$, the result follows from the fact that the largest side of a triangle is opposite its largest angle. ■

4 Analysis and Convergence of Method

In this section we will prove that the method generates a sequence of points such that t goes to zero linearly at the rate $(1-\alpha)$, and $\alpha = O(\frac{1}{\sqrt{n}})$.

We now prove two simple propositions about the solutions of the systems (7) and (9).

Proposition 5 *For every $k = 0, 1, \cdots$ with $\Delta X = \Delta X_k$ or $\Delta X = \Delta \bar{X}_k$ and $\Delta S = \Delta S_k$ or $\Delta S = \Delta \bar{S}_k$ it follows that*

$$A_i \bullet \Delta X = A_i \bullet \Delta X^T = 0$$

$$\Delta X \bullet \Delta S = \Delta S \bullet \Delta X = 0.$$

Proof: The first identity is readily confirmed by simple algebra and the symmetry of A_i, and the second identity is readily confirmed by using the first two equations of the systems (7) and (9). ■

Proposition 6 *For each $k = 0, 1, \cdots$, $\bar{t}_k = (1-\alpha)t_k$ and $t_{k+1} = \bar{t}_k$.*

Proof: Note that $\frac{1}{2}(\Delta X_k + \Delta X_k^T) \bullet S_k + X_k \bullet \Delta S_k = \frac{1}{2}((\Delta X_k \bullet S_k + X_k \bullet \Delta S_k) + (S_k \bullet \Delta X_k^T + \Delta S_k \bullet X_k)) = -\frac{1}{2}(X_k \bullet S_k + S_k \bullet X_k) = -nt_k$. Now, $\bar{t}_k = \frac{\bar{X}_k \bullet \bar{S}_k}{n} = \frac{X_k \bullet S_k}{n} + \frac{1}{2n}\alpha(\Delta X_k \bullet S_k + \Delta X_k^T \bullet S_k + 2X_k \bullet \Delta S_k) = (1-\alpha)t_k$. The remaining part follows by a similar argument. ■

We now prove an important proposition.

Proposition 7 *For each $k = 0, 1, \cdots$, let $E_k = S_k^{\frac{1}{2}} X_k S_k^{\frac{1}{2}} - t_k I$. Then*

1. $\|X_k^{-\frac{1}{2}} \Delta X_k S_k^{\frac{1}{2}}\|_F \leq \sqrt{nt_k}$.
2. $\|X_k^{\frac{1}{2}} \Delta S_k S_k^{-\frac{1}{2}}\|_F \leq \sqrt{nt_k}$.
3. $\|X_k^{-\frac{1}{2}} \Delta X_k^T S_k^{\frac{1}{2}}\|_F \leq \sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}} \sqrt{nt_k}$.

Proof: From the third equation of the system (7) we can derive

$$X_k^{-\frac{1}{2}} \Delta X_k S_k^{\frac{1}{2}} + X_k^{\frac{1}{2}} \Delta S_k S_k^{-\frac{1}{2}} = -X_k^{\frac{1}{2}} S_k^{\frac{1}{2}}$$

and, using Lemma 4, Proposition 5 along with the fact that $\|X_k^{\frac{1}{2}}S_k^{\frac{1}{2}}\|_F^2 = \text{trace}(S_k^{\frac{1}{2}}X_kS_k^{\frac{1}{2}}) = nt_k$, we thus obtain the first two results. To see the third, we note that from 1,

$$\begin{aligned}\|X_k^{-\frac{1}{2}}\Delta X_k^T S_k^{\frac{1}{2}}\|_F &= \|(S_k^{\frac{1}{2}}X_k^{\frac{1}{2}})X_k^{-\frac{1}{2}}\Delta X_k S_k^{\frac{1}{2}}(S_k^{-\frac{1}{2}}X_k^{-\frac{1}{2}})\|_F \\ &\leq \|S_k^{\frac{1}{2}}X_kS_k^{\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k^{-\frac{1}{2}}\Delta X_k S_k^{\frac{1}{2}}\|_F\|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}}.\end{aligned}$$

The result follows from the definition of E_k and part 1. ■

4.1 Analysis of Predictor Step

We now investigate the predictor step and the resulting iterate $(\bar{X}_k, \bar{y}^k, \bar{S}_k, \bar{t}_k)$. For every k, let

$$\begin{aligned}E_k &= S_k^{\frac{1}{2}}X_kS_k^{\frac{1}{2}} - t_kI \\ \bar{E}_k &= \bar{S}_k^{\frac{1}{2}}\bar{X}_k\bar{S}_k^{\frac{1}{2}} - \bar{t}_kI.\end{aligned}$$

We can show that

Lemma 8 *If* $\|E_k\|_F \leq \beta t_k$ *and* $1 - \frac{\alpha\sqrt{n}}{\sqrt{1-\beta}} > 0$, *then* $\bar{X}_k$ *and* $\bar{S}_k$ *are symmetric and positive definite.*

Proof: We will show the result for $\bar{X}_k$. The same argument will apply to $\bar{S}_k$ as well. Since $\bar{X}_k$ is symmetric, and is positive definite if and only if $X_k^{-\frac{1}{2}}\bar{X}_kX_k^{-\frac{1}{2}} = I+\alpha X_k^{-\frac{1}{2}}(\frac{\Delta X_k+\Delta X_k^T}{2})X_k^{-\frac{1}{2}}$ is, we note that from the Lemma 2,

$$\lambda_{\min}(I + \alpha X_k^{-\frac{1}{2}}(\frac{\Delta X_k + \Delta X_k^T}{2})X_k^{-\frac{1}{2}})) \geq 1 - \alpha\|X_k^{-\frac{1}{2}}\frac{\Delta X_k + \Delta X_k^T}{2}X_k^{-\frac{1}{2}}\|_F.$$

Using the result of Lemma 1, 4 and Proposition 7 in the inequalities that follow

$$\begin{aligned}\|X_k^{-\frac{1}{2}}\frac{\Delta X_k + \Delta X_k^T}{2}X_k^{-\frac{1}{2}}\|_F &\leq \frac{1}{2}(\|X_k^{-\frac{1}{2}}\Delta X_kX_k^{-\frac{1}{2}}\|_F + \|X_k^{-\frac{1}{2}}\Delta X_k^TX_k^{-\frac{1}{2}}\|_F) \\ &= \frac{1}{2}(\|X_k^{-\frac{1}{2}}\Delta X_kS_k^{\frac{1}{2}}S_k^{-\frac{1}{2}}X_k^{-\frac{1}{2}}\|_F + \|X_k^{-\frac{1}{2}}S_k^{-\frac{1}{2}}S_k^{\frac{1}{2}}\Delta X_k^TX_k^{-\frac{1}{2}}\|_F) \\ &\leq \|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k^{-\frac{1}{2}}\Delta X_kS_k^{\frac{1}{2}}\|_F \\ &\leq \frac{\sqrt{nt_k}}{\sqrt{t_k - \|E_k\|_F}} \\ &\leq \frac{\sqrt{n}}{\sqrt{1-\beta}}\end{aligned}$$

we obtain our result. ■

We are now ready to establish a bound on the error after the predictor step.

Lemma 9 *If* $\|E_k\|_F \leq \beta t_k, 1-\alpha\sqrt{\frac{n}{1-\beta}} > 0$, *then*

$$\|\bar{E}_k\|_F \leq \frac{1}{1-\alpha}((1-\alpha)\beta + \frac{\alpha^2 n}{2}(1+\sqrt{\frac{1+\beta}{1-\beta}}))\bar{t}_k.$$

Proof: Note that

$$\begin{aligned}
& S_k^{\frac{1}{2}}(\bar{X}_k\bar{S}_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \;& S_k^{\frac{1}{2}}(X_kS_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2}S_k + \alpha X_k\Delta S_k + \alpha^2\frac{\Delta X_k + \Delta X_k^T}{2}\Delta S_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \;& S_k^{\frac{1}{2}}(X_kS_k + \alpha(\frac{X\Delta S_k + \Delta X_k S_k}{2}) + \alpha(\frac{X_k\Delta S_k + \Delta X_k^T S_k}{2}) + \frac{\alpha^2}{2}(\Delta X_k + \Delta X_k^T)\Delta S_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \;& S_k^{\frac{1}{2}}(X_kS_k - \frac{\alpha}{2}X_kS_k + \frac{\alpha}{2}(-X_kS_k + (\Delta X_k^T - \Delta X_k)S_k) + \frac{\alpha^2}{2}(\Delta X_k + \Delta X_k^T)\Delta S_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \;& (1-\alpha)(S_k^{\frac{1}{2}}X_kS_k^{\frac{1}{2}} - t_k I) + \frac{\alpha}{2}S_k^{\frac{1}{2}}(\Delta X_k^T - \Delta X_k)S_k^{\frac{1}{2}} + \frac{\alpha^2}{2}S_k^{\frac{1}{2}}(\Delta X_k + \Delta X_k^T)\Delta S_k S_k^{-\frac{1}{2}}.
\end{aligned}$$

From Lemma 8, $\bar{S}_k$ is symmetric and positive definite and hence its square root exists. The inequalities below follow from the Lemma 3 and the above identity.

$$\begin{aligned}
& \|\bar{E}_k\|_F = \|\bar{S}_k^{\frac{1}{2}}\bar{X}_k\bar{S}_k^{\frac{1}{2}} - \bar{t}_k I\|_F \\
\leq \;& \frac{1}{2}\|S_k^{\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}}(\bar{S}_k^{\frac{1}{2}}\bar{X}_k\bar{S}_k^{\frac{1}{2}} - \bar{t}_k I)\bar{S}_k^{\frac{1}{2}}S_k^{-\frac{1}{2}} + S_k^{-\frac{1}{2}}\bar{S}_k^{\frac{1}{2}}(\bar{S}_k^{\frac{1}{2}}\bar{X}_k\bar{S}_k^{\frac{1}{2}} - \bar{t}_k I)\bar{S}_k^{-\frac{1}{2}}S_k^{\frac{1}{2}}\|_F \\
= \;& \frac{1}{2}\|S_k^{\frac{1}{2}}(\bar{X}_k\bar{S}_k - \bar{t}_k I)S_k^{-\frac{1}{2}} + S_k^{-\frac{1}{2}}(\bar{S}_k\bar{X}_k - \bar{t}_k I)S_k^{\frac{1}{2}}\|_F \\
= \;& \frac{1}{2}\|2(1-\alpha)(S_k^{\frac{1}{2}}X_kS_k^{\frac{1}{2}} - t_k I) + \alpha^2(S_k^{\frac{1}{2}}\frac{\Delta X_k + \Delta X_k^T}{2}\Delta S_k S_k^{-\frac{1}{2}} + S_k^{-\frac{1}{2}}\Delta S_k\frac{\Delta X_k + \Delta X_k^T}{2}S_k^{\frac{1}{2}})\|_F \\
\leq \;& (1-\alpha)\|E_k\|_F + \frac{\alpha^2}{2}\|(S_k^{\frac{1}{2}}\frac{\Delta X_k + \Delta X_k^T}{2}\Delta S_k S_k^{-\frac{1}{2}} + S_k^{-\frac{1}{2}}\Delta S_k\frac{\Delta X_k + \Delta X_k^T}{2}S_k^{\frac{1}{2}})\|_F.
\end{aligned}$$

Note that, from Proposition 7, we obtain

$$\begin{aligned}
\|S_k^{\frac{1}{2}}\Delta X_k^T\Delta S_k S_k^{-\frac{1}{2}}\|_F &\leq \|S_k^{\frac{1}{2}}\Delta X_k^T X_k^{-\frac{1}{2}}\|_F\|X_k^{\frac{1}{2}}\Delta S_k S_k^{-\frac{1}{2}}\|_F \leq n t_k, \\
\|S_k^{-\frac{1}{2}}\Delta S_k\Delta X_k S_k^{\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}}\Delta X_k^T\Delta S_k S_k^{-\frac{1}{2}}\|_F \leq n t_k, \\
\|S_k^{\frac{1}{2}}\Delta X_k\Delta S_k S_k^{-\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}}X_k^{\frac{1}{2}}X_k^{-\frac{1}{2}}\Delta X_k S_k^{\frac{1}{2}}S_k^{-\frac{1}{2}}\Delta S_k X_k^{\frac{1}{2}}X_k^{-\frac{1}{2}}S_k^{-\frac{1}{2}}\|_F \\
&\leq \sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}} n t_k.
\end{aligned}$$

Hence

$$\begin{aligned}\|\bar{E}_k\|_F &\leq (1-\alpha)\|E_k\|_F + \frac{\alpha^2}{2}(1+\sqrt{\frac{t_k+\|E_k\|_F}{t_k-\|E_k\|_F}})nt_k \\ &\leq \frac{1}{1-\alpha}((1-\alpha)\beta + \frac{\alpha^2 n}{2}(1+\sqrt{\frac{1+\beta}{1-\beta}}))\bar{t}_k\end{aligned}$$

and we are done. ■

4.2 Analysis of Corrector Step

We now prove two important lemmas about the corrector step.

Lemma 10 *If $\|\bar{E}_k\|_F \leq \bar{\beta}\bar{t}_k$ and $\frac{\bar{\beta}}{1-\bar{\beta}} < 1$, then X_{k+1} and S_{k+1} are symmetric and positive definite.*

Proof:

By a similar argument to that of Proposition 7, we obtain

$$\begin{aligned}\|\bar{X}_k^{\frac{1}{2}}\Delta\bar{S}_k\bar{S}_k^{-\frac{1}{2}}\|_F &\leq \|\bar{t}_k\bar{X}_k^{-\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}} - \bar{X}_k^{\frac{1}{2}}\bar{S}_k^{\frac{1}{2}}\|_F \\ &= \|(\bar{t}_kI - \bar{X}_k^{\frac{1}{2}}\bar{S}_k\bar{X}_k^{\frac{1}{2}})(\bar{X}_k^{-\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}})\|_F, \qquad (11)\\ \|\bar{X}_k^{-\frac{1}{2}}\Delta\bar{X}_k\bar{S}_k^{\frac{1}{2}}\|_F &\leq \|\bar{t}_k\bar{X}_k^{-\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}} - \bar{X}_k^{\frac{1}{2}}\bar{S}_k^{\frac{1}{2}}\|_F \\ &= \|(\bar{t}_kI - \bar{X}_k^{\frac{1}{2}}\bar{S}_k\bar{X}_k^{\frac{1}{2}})(\bar{X}_k^{-\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}})\|_F. \qquad (12)\end{aligned}$$

Then this lemma follows by an argument identical to that of Lemma 8. ■

The next Lemma provides the error after the corrector step.

Lemma 11 *If $\|\bar{E}_k\|_F \leq \bar{\beta}\bar{t}_k = \bar{\beta}(1-\alpha)t_k$, and $\frac{\bar{\beta}}{1-\bar{\beta}} < 1$ then*

$$\|E_{k+1}\|_F \leq \frac{1}{2}(1+\sqrt{\frac{1+\bar{\beta}}{1-\bar{\beta}}})\frac{1}{1-\bar{\beta}}\bar{\beta}^2(1-\alpha)t_k.$$

Proof: The following relations follow from Lemma 3, and by an analysis similar to that in Lemma 9,

$$\|E_{k+1}\|_F = \|(S_{k+1})^{\frac{1}{2}}X_{k+1}(S_{k+1})^{\frac{1}{2}} - \bar{t}_kI\|_F$$

$$
\begin{aligned}
&\le \frac{1}{2}\|\bar{S}_k^{\frac{1}{2}}(S_{k+1})^{-\frac{1}{2}}((S_{k+1})^{\frac{1}{2}}X_{k+1}(S_{k+1})^{\frac{1}{2}} - \bar{t}_k I)(S_{k+1})^{\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}} \\
&\qquad + \bar{S}_k^{-\frac{1}{2}}(S_{k+1})^{\frac{1}{2}}((S_{k+1})^{\frac{1}{2}}X_{k+1}(S_{k+1})^{\frac{1}{2}} - \bar{t}_k I)(S_{k+1})^{-\frac{1}{2}}\bar{S}_k^{\frac{1}{2}}\|_F \\
&= \frac{1}{2}\|\bar{S}_k^{\frac{1}{2}}(X_{k+1}S_{k+1} - \bar{t}_k I)\bar{S}_k^{-\frac{1}{2}} + \bar{S}_k^{-\frac{1}{2}}(S_{k+1}X_{k+1} - \bar{t}_k I)\bar{S}_k^{\frac{1}{2}}\|_F \\
&= \frac{1}{2}\|\bar{S}_k^{\frac{1}{2}}\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k\bar{S}_k^{-\frac{1}{2}} + \bar{S}_k^{-\frac{1}{2}}\Delta\bar{S}_k\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\bar{S}_k^{\frac{1}{2}}\|_F.
\end{aligned}
$$

Then, using arguments similar to the proof of Lemma 9 and inequalities (11), (12), it follows that

$$
\begin{aligned}
\|\bar{S}_k^{\frac{1}{2}}\Delta\bar{X}_k^T\Delta\bar{S}_k\bar{S}_k^{-\frac{1}{2}}\|_F &\le \|\bar{S}_k^{\frac{1}{2}}\Delta\bar{X}_k^T\bar{X}_k^{-\frac{1}{2}}\|_F\|\bar{X}_k^{\frac{1}{2}}\Delta\bar{S}_k\bar{S}_k^{-\frac{1}{2}}\|_F \\
&\le \|(\bar{t}_k I - \bar{X}_k^{\frac{1}{2}}\bar{S}_k\bar{X}_k^{\frac{1}{2}})(\bar{X}_k^{-\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}})\|_F^2 \\
&\le \|\bar{t}_k I - \bar{X}_k^{\frac{1}{2}}\bar{S}_k\bar{X}_k^{\frac{1}{2}}\|_F^2\|\bar{S}_k^{-\frac{1}{2}}\bar{X}_k^{-1}\bar{S}_k^{-\frac{1}{2}}\|_2 \\
&\le \frac{1}{\bar{t}_k - \|\bar{E}_k\|_F}\|\bar{E}_k\|_F^2,
\end{aligned}
$$

$$
\|\bar{S}_k^{-\frac{1}{2}}\Delta\bar{S}_k\Delta\bar{X}_k\bar{S}_k^{\frac{1}{2}}\|_F = \|\bar{S}_k^{\frac{1}{2}}\Delta\bar{X}_k^T\Delta\bar{S}_k\bar{S}_k^{-\frac{1}{2}}\|_F \le \frac{1}{\bar{t}_k - \|\bar{E}_k\|_F}\|\bar{E}_k\|_F^2,
$$

and, using Lemma 1 and Lemma 2,

$$
\begin{aligned}
\|\bar{S}_k^{\frac{1}{2}}\Delta\bar{X}_k\Delta\bar{S}_k\bar{S}_k^{-\frac{1}{2}}\|_F &= \|\bar{S}_k^{\frac{1}{2}}\bar{X}_k^{\frac{1}{2}}\bar{X}_k^{-\frac{1}{2}}\Delta\bar{X}_k\bar{S}_k^{\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}}\Delta\bar{S}_k\bar{X}_k^{\frac{1}{2}}\bar{X}_k^{-\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}}\|_F \\
&\le \|\bar{S}_k^{\frac{1}{2}}\bar{X}_k\bar{S}_k^{\frac{1}{2}}\|_2^{\frac{1}{2}}\|\bar{X}_k^{-\frac{1}{2}}\Delta\bar{X}_k\bar{S}_k^{\frac{1}{2}}\bar{S}_k^{-\frac{1}{2}}\Delta\bar{S}_k\bar{X}_k^{\frac{1}{2}}\|_F\|\bar{S}_k^{-\frac{1}{2}}\bar{X}_k^{-1}\bar{S}_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}} \\
&\le \sqrt{\frac{\bar{t}_k + \|\bar{E}_k\|_F}{\bar{t}_k - \|\bar{E}_k\|_F}}\|\bar{X}_k^{-\frac{1}{2}}\Delta\bar{X}_k\bar{S}_k^{\frac{1}{2}}\|_F\|\bar{S}_k^{-\frac{1}{2}}\Delta\bar{S}_k\bar{X}_k^{\frac{1}{2}}\|_F \\
&\le \sqrt{\frac{\bar{t}_k + \|\bar{E}_k\|_F}{\bar{t}_k - \|\bar{E}_k\|_F}}\frac{1}{\bar{t}_k - \|\bar{E}_k\|_F}\|\bar{E}_k\|_F^2.
\end{aligned}
$$

Thus

$$
\begin{aligned}
\|E_{k+1}\|_F &\le \frac{1}{2}(1 + \sqrt{\frac{\bar{t}_k + \|\bar{E}_k\|_F}{\bar{t}_k - \|\bar{E}_k\|_F}})\frac{1}{\bar{t}_k - \|\bar{E}_k\|_F}\|\bar{E}_k\|_F^2 \\
&\le \frac{1}{2}(1 + \sqrt{\frac{1+\bar{\beta}}{1-\bar{\beta}}})\frac{1}{1-\bar{\beta}}\bar{\beta}^2(1-\alpha)t_k
\end{aligned}
$$

and we are done. ■

4.3 The Main Result

We are now ready to prove the main convergence theorem.

Theorem 12 *Let $\alpha \leq \frac{1}{8\sqrt{n}}$ and $\beta = 0.4$. Then for every $k = 1, 2, \cdots$, $(X_k, y^k, S_k) \in N(t_k, \beta)$ and $t_{k+1} = (1-\alpha)t_k$. Thus, for every $\epsilon > 0$, after at most $O(|log(\frac{\epsilon}{t_0})|\sqrt{n})$ iterations of the predictor-corrector method, a solution $(X(\epsilon), y(\epsilon), S(\epsilon))$ will be found with $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$.*

Proof: It is readily confirmed that the above choice satisfies the properties required in the hypothesis of the Lemmas 8, 9, 10, 11, and that $\|\bar{E}_k\|_F = \|\bar{S}_k^{\frac{1}{2}} \bar{X}_k \bar{S}_k^{\frac{1}{2}} - \bar{t}_k I\|_F \leq \beta \bar{t}_k$, and $\|E_{k+1}\|_F \leq \beta \bar{t}_k = \beta t_{k+1}$, and thus the sequence belongs to the neighborhood $N(t_k, \beta)$. The required property on t_k follows from Proposition 6.

Now let $\epsilon > 0$ be arbitrary, and define k sufficiently large so that $(1-\alpha)^k t_0 \leq \epsilon$. With $\alpha \leq \frac{1}{8}$, then $\log(\frac{\epsilon}{t_0}) \geq k\log(1-\alpha) \geq -k\frac{3\alpha}{2}$. Thus, for $k \geq \frac{-2}{3\alpha}\log(\frac{\epsilon}{t_0})$, $t_k \leq \epsilon$. So for $k = O(|\log(\frac{\epsilon}{t_0})|\sqrt{n})$, the value of t_k is less than equal to ϵ. Thus the required solution is obtained. ■

5 Modified Corrector Step

The predictor corrector method presented in Section 2 requires that a new system of equations be solved during the corrector step. Implementations generally use the same linear system during both the steps. In this section, we will analyze the resulting method. Thus, we replace **Step 2** of the method of Section 2 by the following:

Step 2′ : Corrector Step: Solve

$$\begin{array}{llcll} A_i \bullet \Delta\bar{X}_k & & = & 0 & \text{for all } i = 1, \cdots, m \\ & \sum_{i=1}^m A_i \Delta\bar{y}_i^k + \Delta\bar{S}_k & = & 0 & \\ \Delta\bar{X}_k S_k & + X_k \Delta\bar{S}_k & = & \bar{t}_k I - \bar{X}_k \bar{S}_k & \end{array} \tag{13}$$

for $(\Delta\bar{X}_k, \Delta\bar{y}^k, \Delta\bar{S}_k)$, and define

$$\begin{aligned} X_{k+1} &= \bar{X}_k + \tfrac{1}{2}(\Delta\bar{X}_k + \Delta\bar{X}_k^T) \\ y^{k+1} &= \bar{y}^k + \Delta\bar{y}^k \\ S_{k+1} &= \bar{S}_k + \Delta\bar{S}_k \\ t_{k+1} &= \tfrac{X_{k+1}\bullet S_{k+1}}{n}. \end{aligned} \tag{14}$$

Note that the matrix of the system of equations to be solved in (13) is the same as in (7). We now prove the convergence of this method. Before we do this we establish some needed lemmas.

Lemma 13 *For each* $k = 1, 2, \cdots$, $t_{k+1} = (1-\alpha)t_k$.

Proof: Follows by a straightforward calculation. ∎

Proposition 14 *For each* $k = 1, 2, \cdots$,

$$\begin{aligned} \|X_k^{-\frac{1}{2}}\Delta\bar{X}_k S_k^{\frac{1}{2}}\|_F &\le M(\alpha)\sqrt{t_k}, \\ \|X_k^{\frac{1}{2}}\Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F &\le M(\alpha)\sqrt{t_k}, \\ \|S_k^{\frac{1}{2}}\Delta\bar{X}_k X_k^{-\frac{1}{2}}\|_F &\le \|S_k^{\frac{1}{2}}X_k S_k^{\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k^{-\frac{1}{2}}\Delta\bar{X}_k S_k^{\frac{1}{2}}\|_F\|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}} \\ &\le \sqrt{\frac{1+\beta}{1-\beta}}M(\alpha)\sqrt{t_k} \end{aligned}$$

where

$$M(\alpha) = (1-\alpha)\frac{\beta}{\sqrt{1-\beta}} + \frac{\alpha}{2}(1+\sqrt{\frac{1+\beta}{1-\beta}})\sqrt{n} + \frac{\alpha^2}{2}(1+\sqrt{\frac{1+\beta}{1-\beta}})\frac{n}{\sqrt{1-\beta}}.$$

Proof: We will only show this for $\|X_k^{-\frac{1}{2}}\Delta\bar{X}_k S_k^{\frac{1}{2}}\|_F$. The inequalities that follow use the same arguments as those of Proposition 7.

$$\begin{aligned} &\|X_k^{-\frac{1}{2}}\Delta\bar{X}_k S_k^{\frac{1}{2}}\|_F \\ \le\ &\|\bar{t}_k X_k^{-\frac{1}{2}}S_k^{-\frac{1}{2}} - X_k^{-\frac{1}{2}}\bar{X}_k\bar{S}_k S_k^{-\frac{1}{2}}\|_F \\ =\ &\|X_k^{-\frac{1}{2}}S_k^{-\frac{1}{2}}(1-\alpha)(t_k I - S_k^{\frac{1}{2}}X_k S_k^{\frac{1}{2}}) + X_k^{-\frac{1}{2}}((1-\alpha)X_k S_k - \bar{X}_k\bar{S}_k)S_k^{-\frac{1}{2}}\|_F \end{aligned}$$

$$
\begin{aligned}
&\le (1-\alpha)\|E_k\|_F \frac{1}{\sqrt{t_k - \|E_k\|_F}} + \|X_k^{-\frac{1}{2}}((1-\alpha)X_kS_k - (X_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2})(S_k + \alpha \Delta S_k))S_k^{-\frac{1}{2}}\|_F \\
&\le (1-\alpha)\frac{\beta\sqrt{t_k}}{\sqrt{1-\beta}} + \|X_k^{-\frac{1}{2}}(\alpha(X_k\Delta S_k + \Delta X_k S_k) - \alpha\frac{\Delta X_k + \Delta X_k^T}{2}S_k - \alpha X_k \Delta S_k \\
&\quad -\alpha^2 \frac{\Delta X_k + \Delta X_k^T}{2}\Delta S_k)S_k^{-\frac{1}{2}}\|_F \\
&\le (1-\alpha)\frac{\beta\sqrt{t_k}}{\sqrt{1-\beta}} + \|X_k^{-\frac{1}{2}}(\alpha\frac{\Delta X_k - \Delta X_k^T}{2}S_k - \alpha^2\frac{\Delta X_k + \Delta X_k^T}{2}\Delta S_k)S_k^{-\frac{1}{2}}\|_F \\
&\le (1-\alpha)\frac{\beta\sqrt{t_k}}{\sqrt{1-\beta}} + \frac{\alpha}{2}(1 + \sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}})\sqrt{nt_k} + \frac{\alpha^2}{2}(1 + \sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}})\frac{nt_k}{\sqrt{t_k - \|E_k\|_F}} \\
&\le ((1-\alpha)\frac{\beta}{\sqrt{1-\beta}} + \frac{\alpha}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}})\sqrt{n} + \frac{\alpha^2}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}})\frac{n}{\sqrt{1-\beta}})\sqrt{t_k}
\end{aligned}
$$

and we are done. ■

We now establish the two required lemmas that show that the resulting matrices after the modified corrector step are positive definite and that they belong to the neighborhood $N(t_{k+1}, \beta)$.

Lemma 15 *If* $\|E_k\|_F \le \beta t_k, 1 - \alpha\sqrt{\frac{n}{1-\beta}} - \frac{M(\alpha)}{\sqrt{1-\beta}} > 0$, *then* X_{k+1} *and* S_{k+1} *are symmetric and positive definite.*

Proof: Using the identity $X_{k+1} = X_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T) + \frac{1}{2}(\Delta\bar{X}_k + \Delta\bar{X}_k^T)$, this lemma follows by an argument identical to that of Lemma 8 and 10. ■

Lemma 16 *If* $\|E_k\|_F \le \beta t_k, 1 - \alpha\sqrt{\frac{n}{1-\beta}} - \frac{M(\alpha)}{\sqrt{1-\beta}} > 0$ *then*

$$
\|E_{k+1}\|_F \le \frac{1}{1-\alpha}(\alpha(1 + \sqrt{\frac{1+\beta}{1-\beta}})M(\alpha)\sqrt{n} + \frac{1}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}})M(\alpha)^2)t_{k+1}.
$$

Proof: Note that

$$
\begin{aligned}
& S_k^{\frac{1}{2}}(X_{k+1}S_{k+1} - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \;& S_k^{\frac{1}{2}}(\bar{X}_k\bar{S}_k + (\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2})(S_k + \alpha\Delta S_k) + (X_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2})\Delta\bar{S}_k
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \; & S_k^{\frac{1}{2}}(\bar{X}_k\bar{S}_k + (\frac{X_k\Delta\bar{S}_k + \Delta\bar{X}_k S_k}{2}) + (\frac{X_k\Delta\bar{S}_k + \Delta\bar{X}_k^T S_k}{2}) + \alpha(\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2})\Delta S_k \\
&+\alpha\frac{\Delta X_k + \Delta X_k^T}{2}\Delta\bar{S}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \; & S_k^{\frac{1}{2}}(\bar{X}_k\bar{S}_k + \frac{1}{2}(\bar{t}_k I - \bar{X}_k\bar{S}_k) + \frac{1}{2}(\bar{t}_k I - \bar{X}_k\bar{S}_k + (\Delta\bar{X}_k^T - \Delta\bar{X}_k)S_k) + \alpha\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta S_k \\
&+\alpha\frac{\Delta X_k + \Delta X_k^T}{2}\Delta\bar{S}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}^T}{2}\Delta\bar{S}_k - \bar{t}_k I)S_k^{-\frac{1}{2}} \\
= \; & S_k^{\frac{1}{2}}(\alpha\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta S_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2}\Delta\bar{S}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k)S_k^{-\frac{1}{2}} \\
&+\frac{1}{2}S_k^{\frac{1}{2}}(\Delta\bar{X}_k^T - \Delta\bar{X}_k)S_k^{\frac{1}{2}}.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\|E_{k+1}\|_F = \|(S_{k+1})^{\frac{1}{2}}X_{k+1}(S_{k+1})^{\frac{1}{2}} - \bar{t}_k I\|_F \\
\leq \; & \|S_k^{\frac{1}{2}}(S_{k+1})^{-\frac{1}{2}}((S_{k+1})^{\frac{1}{2}}X_{k+1}(S_{k+1})^{\frac{1}{2}} - \bar{t}_k I)(S_{k+1})^{\frac{1}{2}}S_k^{-\frac{1}{2}}\|_F \\
&+\|S_k^{-\frac{1}{2}}(S_{k+1})^{\frac{1}{2}}((S_{k+1})^{\frac{1}{2}}X_{k+1}(S_{k+1})^{\frac{1}{2}} - \bar{t}_k I)(S_{k+1})^{-\frac{1}{2}}S_k^{\frac{1}{2}}\|_F \\
= \; & \frac{1}{2}\|S_k^{\frac{1}{2}}(X_{k+1}S_{k+1} - \bar{t}_k I)S_k^{-\frac{1}{2}} + S_k^{-\frac{1}{2}}(S_{k+1}X_{k+1} - \bar{t}_k I)S_k^{\frac{1}{2}}\|_F \\
= \; & \frac{1}{2}\|S_k^{\frac{1}{2}}(\alpha\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta S_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2}\Delta\bar{S}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k)S_k^{-\frac{1}{2}} \\
&+(S_k^{\frac{1}{2}}(\alpha\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta S_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2}\Delta\bar{S}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k)S_k^{-\frac{1}{2}})^T\|_F \\
\leq \; & \|S_k^{\frac{1}{2}}(\alpha\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta S_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2}\Delta\bar{S}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k)S_k^{-\frac{1}{2}}\|_F.
\end{aligned}
$$

We have

$$
\begin{aligned}
\|S_k^{\frac{1}{2}}\Delta\bar{X}_k\Delta S_k S_k^{-\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}}\Delta\bar{X}_k X_k^{-\frac{1}{2}}X_k^{\frac{1}{2}}\Delta S_k S_k^{-\frac{1}{2}}\|_F \\
&\leq \|S_k^{\frac{1}{2}}\Delta\bar{X}_k X_k^{-\frac{1}{2}}\|_F\|X_k^{\frac{1}{2}}\Delta S_k S_k^{-\frac{1}{2}}\|_F \\
&\leq \sqrt{nt_k}\|S_k^{\frac{1}{2}}X_k S_k^{\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k^{-\frac{1}{2}}\Delta\bar{X}_k S_k^{\frac{1}{2}}\|_F\|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}} \\
&\leq \sqrt{nt_k}\sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}}M(\alpha)\sqrt{t_k}, \\
\|S_k^{\frac{1}{2}}\Delta\bar{X}_k^T\Delta S_k S_k^{-\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}}\Delta\bar{X}_k^T X_k^{-\frac{1}{2}}X_k^{\frac{1}{2}}\Delta S_k S_k^{-\frac{1}{2}}\|_F \\
&\leq \|S_k^{\frac{1}{2}}\Delta\bar{X}_k^T X_k^{-\frac{1}{2}}\|_F\|X_k^{\frac{1}{2}}\Delta S_k S_k^{-\frac{1}{2}}\|_F
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sqrt{nt_k}M(\alpha)\sqrt{t_k}, \\
\|S_k^{\frac{1}{2}}\Delta X_k \Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}}\Delta X_k X_k^{-\frac{1}{2}} X_k^{\frac{1}{2}}\Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F \\
&\leq M(\alpha)t_k \|S_k^{\frac{1}{2}} X_k S_k^{\frac{1}{2}}\|_2^{\frac{1}{2}} \|X_k^{-\frac{1}{2}}\Delta X_k S_k^{\frac{1}{2}}\|_F \|S_k^{-\frac{1}{2}} X_k^{-1} S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}} \\
&\leq \left(\sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}}\right) M(\alpha)\sqrt{t_k}\sqrt{nt_k}, \\
\|S_k^{\frac{1}{2}}\Delta X_k^T \Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}}\Delta X_k^T X_k^{-\frac{1}{2}}\|_F \|X_k^{\frac{1}{2}}\Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F \\
&\leq \sqrt{nt_k}M(\alpha)\sqrt{t_k}, \\
\|S_k^{\frac{1}{2}}\frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}\Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F &\leq \frac{1}{2}(\|S_k^{\frac{1}{2}}\Delta\bar{X}_k \Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F + \|S_k^{\frac{1}{2}}\Delta\bar{X}_k^T \Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F) \\
&\leq \frac{1}{2}(\|S_k^{\frac{1}{2}}\Delta\bar{X}_k X_k^{-\frac{1}{2}}\|_F \|X_k^{\frac{1}{2}}\Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F \\
&\quad + \|S_k^{\frac{1}{2}}\Delta\bar{X}_k^T X_k^{-\frac{1}{2}}\|_F \|X_k^{\frac{1}{2}}\Delta\bar{S}_k S_k^{-\frac{1}{2}}\|_F) \\
&\leq \frac{1}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}})M(\alpha)^2 t_k.
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\|E_{k+1}\|_F \\
\leq\ &\frac{1}{1-\alpha}(\alpha(1 + \sqrt{\frac{1+\beta}{1-\beta}})M(\alpha)\sqrt{n} + \frac{1}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}})M(\alpha)^2)t_{k+1}.
\end{aligned}
$$

■

Theorem 17 *Let $\alpha \leq \frac{1}{20\sqrt{n}}, \beta = 0.1$. Then for every $k = 1, 2, \cdots$, $(X_k, y^k, S_k) \in N(t_k, \beta)$ and $t_{k+1} = (1-\alpha)t_k$. Thus, for every $\epsilon > 0$, after at most $O(|log(\frac{\epsilon}{t_0})|\sqrt{n})$ iterations of the predictor-corrector method, a solution $(X(\epsilon), y(\epsilon), S(\epsilon))$ will be found with $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$.*

Proof: Using the Lemmas 15 and 16, the proof is identical to that of Theorem 12. ■

6 An Implementable Method

In this section we will present an implementable method that generates a larger step size than the theoretical one determined by Theorems 12 and 17. And some numerical issues of implementation will also be discussed.

6.1 A Practical Method

The step size α obtained in Theorems 12 and 17 is generally too small for practical implementation. We present below a practical predictor corrector method based on these theorems.

Step 1′ Predictor Step: Solve

$$\begin{array}{rcll} A_i \bullet \Delta X_k & = & 0 & \text{for all } i = 1, \cdots, m \\ \sum_{i=1}^m A_i \Delta y_i^k + \Delta S_k & = & 0 & \\ \Delta X_k S_k + X_k \Delta S_k & = & -X_k S_k & \end{array} \tag{15}$$

for $(\Delta X_k, \Delta y^k, \Delta S_k)$.

Step 2″ Step Selection:

Set $l = 0$ and $\alpha_0 = \max\{\alpha : X_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T) \succeq 0, S_k + \alpha \Delta S_k \succeq 0 \text{ and } \alpha \leq 1\}$. If $\alpha_0 = 1$ stop, otherwise set $\alpha_0 = 0.99\alpha_0$.

Step 3′ Set

$$\begin{aligned} X_{k,l} &= X_k + \frac{1}{2}\alpha_l(\Delta X_k + \Delta X_k^T) \\ y^{k,l} &= y^k + \alpha_l \Delta y^k \\ S_{k,l} &= S_k + \alpha_l \Delta S_k \end{aligned}$$

Step 4′ Corrector Step: Set $\bar{X}_k = X_{k,l}$, $\bar{y}^k = y^{k,l}$, $\bar{S}_k = S_{k,l}$, $\bar{t}_k = (1 - \alpha_l)t_k$ and solve system (13).

Step 5′ Set

$$\begin{aligned} X_{k+1,l} &= X_k + \frac{1}{2}(\Delta \bar{X}_k + \Delta \bar{X}_k^T) \\ y^{k+1,l} &= y^k + \Delta \bar{y}^k \\ S_{k+1,l} &= S_k + \Delta \bar{S}_k \end{aligned}$$

Step 6′ If $(X_{k+1,l}, y^{k+1,l}, S_{k+1,l}) \in N((1-\alpha_l)t_k, \beta)$, then set $X_{k+1} = X_{k+1,l}$, $y^{k+1} = y^{k+1,l}$, and $S_{k+1} = S_{k+1,l}$, $t_{k+1} = (1 - \alpha_l)t_k$, $k = k + 1$ and go to Step 1′. Otherwise, set $\alpha_{l+1} = \frac{1}{2}\alpha_l$, $l = l + 1$, and go to step 3′.

We can show that:

Lemma 18 *If* $\alpha_0 = 1$ *in Step 2'', then* $X_k + \frac{1}{2}(\Delta X_k + \Delta X_k^T)$, $y^k + \Delta y^k$, $S_k + \Delta S_k$ *solves the semidefinite programming problem.*

Proof: The result follows from Lemma 13. ■

We are now ready to prove the main theorem.

Theorem 19 *Let* $\beta = 0.10$. *Then for every* $\epsilon > 0$, *after at most* $O(|log(\frac{\epsilon}{t_0})| log(n)\sqrt{n})$ *total predictor and corrector iterations of the above method, a solution* $(X(\epsilon), y(\epsilon), S(\epsilon))$ *will be found with* $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$.

Proof: We first observe that after at most $l = O(\log(n))$ corrector iterations of Step 2' to Step 6', $\alpha_l \leq \frac{1}{20\sqrt{n}}$, and thus the resulting iterate will lie in $N(t_{k+1}, \beta)$ by Theorem 17, and the result follows. ■

6.2 Some Implementation Issues

We discuss here the linear system to be solved in determining the directions, and a method for exploiting the sparsity of the matrices involved.

Instead of the system (7) solved in section 2, we will solve ΔX_k^T by replacing the third equation of the system with the following:

$$S_k \Delta X_k^T + \Delta S_k X_k = -S_k X_k.$$

Then similar to the system for linear programming, the corresponding linear system can be represented as follows :

$$\begin{bmatrix} \bar{A} & 0 & 0 \\ 0 & \bar{A}^T & I \\ F_k & 0 & G_k \end{bmatrix} \begin{bmatrix} \text{vec}(\Delta X_k^T) \\ \Delta y^k \\ \text{vec}\Delta S_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \text{vec}(-S_k X_k) \end{bmatrix}$$

where

$$\bar{A} = \begin{bmatrix} \text{vec}(A_1)^T \\ \vdots \\ \text{vec}(A_m)^T \end{bmatrix}, F_k = I \otimes S_k, G_k = X_k \otimes I.$$

Note that $A \otimes B = [A_{ij}B]$ is Kronecker product.

We now show that $F_k^{-1}G_k$ is a symmetric and positive definite matrix.

Lemma 20 *$F_k^{-1}G_k$ is a symmetric and positive definite matrix.*

Proof: Note that $G_k = P\bar{G}_k P^T$ where P is a permutation matrix and $\bar{G}_k = I \otimes X_k$. Since $\bar{G}_k$ is symmetric and positive definite, G_k is symmetric positive definite. $F_k^{-1}G_k = X_k \otimes S_k^{-1}$ is symmetric. With the fact that F_k is symmetric and positive definite, since $F_k^{-1}G_k$ has the same eigenvalues as $G_k^{\frac{1}{2}} F_k^{-1} G_k^{\frac{1}{2}}$, the required property follows. ∎

Then the linear system can be solved by

$$\begin{aligned} \Delta y^k &= (\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}(-\bar{A})F_k^{-1}\mathrm{vec}(-S_kX_k) \\ \Delta S_k &= -\sum_{i=1}^{m} \Delta y_i^k A_i \\ \Delta X_k^T &= S_k^{-1}(-S_kX_k - \Delta S_k X_k) \end{aligned} \tag{16}$$

As can be readily confirmed, it will require $O(mn^3)$ multiplications to compute $G_k\bar{A}^T$, $O(mn^3)$ multiplications to compute $F_k^{-1}(G_k\bar{A})$, $O(m^2n^2)$ multiplications to compute $\bar{A}(F_k^{-1}G_k\bar{A})$, and $O(m^3)$ multiplications to compute $(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}$. A consequence of Theorem 19 is the following

Corollary 21 *Let $\beta = 0.10$. Then for every $\epsilon > 0$, after at most $|log(\frac{\epsilon}{t_0})|(O(mn^{3.5}) + O(m^2n^{2.5}) + O(m^3n^{0..5}))$ total multiplications, a solution $(X(\epsilon), y(\epsilon), S(\epsilon))$ will be found with $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$.*

Proof: This multiplication count can be obtained by observing that in Step 4′, only the right hand side changes between the predictor and corrector steps. Also, between the at most $O(\log(n))$ corrector steps, the only change in the right hand side is the value of α, and thus the new solution can be found in $O(n^2)$ multiplication for each of the subsequent corrector steps. Thus the total multiplications is at most $|\log(\frac{\epsilon}{t_0})|(O(mn^{3.5}) + O(m^2n^{2.5}) + O(m^3n^{0.5}))$ for all the predictor steps, and, at most $K_1(mn^{3.5} + m^2n^{2.5} + m^3n^{0.5}) + K_2n^{2.5}\log(n)$ multiplications for all the corrector steps. ∎

The computational burden in above system is in the calculation of $(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}$. The matrix $\bar{A}F_k^{-1}G_k\bar{A}^T$ is dense even when $\bar{A}$, F_k^{-1} and G_k are sparse. This then renders the methods of sparse Cholesky factorization, George and Liu [4], ineffective. But the method presented in Saigal [20] may be able to exploit this sparsity. We now show how this may be so. Choose $\theta > 0$ sufficiently large so that $I - \frac{1}{\theta}F_k^{-1}G_k \succeq 0$. Also let L be the Cholesky factor of $\bar{A}$ (i.e., $\bar{A}\bar{A}^T = LL^T$). Then it is readily confirmed that $(L^{-1}\bar{A})(L^{-1}\bar{A})^T = I$. In that case if $\bar{A}\text{vec}(X_k) = b$, the linear system of (1), is modified to the equivalent system $L^{-1}\bar{A}\text{vec}(X_k) = L^{-1}b$, the resulting matrix in the system (16) will be

$$L^{-1}\bar{A}F_k^{-1}G_k(L^{-1}\bar{A})^T.$$

Let $H_k = L^{-1}\bar{A}F_k^{-1}G_k(L^{-1}\bar{A})^T = \theta_k(I - N_k)$ where $N_k = L^{-1}\bar{A}(I - \frac{1}{\theta_k}F_k^{-1}G_k)(L^{-1}\bar{A})^T$. We can then show that

Lemma 22 *For all $\theta \geq \lambda_{\max}(F_k^{-\frac{1}{2}}G_kF_k^{-\frac{1}{2}})$, N_k is a symmetric and positive definite matrix with spectral radius strictly less than 1.*

Proof: The first part follows from the choice of θ. Since $F_k^{-1}G_k$ is positive definite so is H_k. Thus $\frac{1}{\theta}z^TH_kz = 1 - z^TN_kz > 0$ for every $\|z\|_2 = 1$, and we have the second part of the lemma since $z^TN_kz > 0$. ∎

The idea now is to solve the system by writing $H_k^{-1} = \frac{1}{\theta_k}(I + N_k + N_k^2 + \cdots)$ and Saigal [20] shows how this infinite series can be summed iteratively. The application of this methodology requires the multiplication of a vector q by the matrix N_k. To see how sparsity is preserved by this technique, consider generically multiplying a vector q by N_k to obtain the vector p. Then the following is simple to generate:

Step 1 Solve $L^T\bar{q} = q$.

Step 2 Define $\hat{q} = \bar{A}(I - \frac{1}{\theta_k}F_k^{-1}G_k)\bar{A}^T\bar{q}$.

Step 3 Solve $Lp = \hat{q}$

In steps 1 and 3, sparse triangular systems are solved (provided the Cholesky factor of $\bar{A}$ is sparse), and in step 2 multiplication by the sparse matrices $\bar{A}$, $\bar{A}^T$, G_k and F_k^{-1} is involved. Note that in Step 1 and 3, $O(m^2)$ multiplications are required and in Step 2, $O(mn^2) + O(n^3)$ multiplications are required.

7 Acknowledgement

We thank Professor Masakazu Kojima for pointing out an earlier error in Lemma 2.

References

[1] F. Alizadeh, "Interior point methods in semidefinite programming with application to combinatorial optimization", *SIAM Journal on Optimization*, 5(1995), 13–51.

[2] F. Alizadeh, J. P. Haeberly, and M. Overton "Complementarity and nondegeneracy in semidefinite programming", manuscript, March 1995, submitted to *Mathematical Programming.*

[3] R. M. Freund, "Complexity of an algorithm for finding an approximate solution of a semi-definite program with no regularity assumption", Technical Report Number OR 302-94,Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, December 1994.

[4] A. George and J. W-H. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall Inc., New Jersey, 1981.

[5] G. H. Golub, and C. E. V. Loan, *Matrix computations*, Second edition, The John Hopkins University Press, Baltimore, MA 21211, 1989.

[6] C. Helmberg, F. Rendl, B. Vanderbei, and H. Wolkowicz, "An interior-point method for semidefinite programming", Manuscript, Programming in Statistics and Operations Research, Princeton University, 1994. To appear in *SIAM Journal on Optimization.*

[7] F. Jarre, "An interior-point method for minimizing and maximum eigenvalue of a linear combination of matrices," *SIAM Journal on Control and Optimization*, 31(1993), 1360–1377.

[8] M. Kojima, S. Shindoh, and S. Hara, "Interior-point methods for the monotone linear complementarity problem in symmetric matrices", Research Reports on Information Sciences, Ser. B : Operations Research B-282, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, April 1994(revised October 1995). To appear in *SIAM Journal on Optimization.*

[9] M. Kojima, M. Shida, S. Shindoh, "Local convergence of predictor-corrector infeasible-interior point algorithms for SDPs and SDLCPS", Research Reports on Information Sciences, Ser. B : Operations Research B-306, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, December 1995.

[10] S. Mizuno, M. J. Todd, and Y. Ye, "On adaptive step primal-dual interior-point algorithms for linear. programming," *Mathematics of Operations Research* 18 (1993) , 964–981.

[11] R. D. C. Monteiro, "Primal-dual path following algorithms for semidefinite programming", Manuscript, School of Industrial and System Engineering, Georgia Tech., Atlanta, GA 30332, September 1995.

[12] Y. Nesterov and A. Nemirovskii, "Self-concordant functions and polynomial time methods in convex programming", preprint, Central Economic & Mathematical Institute, USSR Acad. Sci., Moscow, USSR, 1989.

[13] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[14] Y. E. Netsterov, and M. Todd, "Primal-dual interior-point methods for self-scaled cones", Technical Report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 18453-3801, 1995.

[15] G. Pataki, "On the facial structure of cone-LP's and semidefinite programs", Management Science Research Report # MSRR-595, Graduate School of Industrial Administration, Carnegie Mellon University, February 1994.

[16] F.A. Potra and R. Sheng, "A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming", Department of Mathematics, University of Iowa, Iowa City, IA 52242, October 1995.

[17] M. Ramana,L. Tunçel, and H. Wolkowicz, " Strong duality for semidefinite programming", Technical Report CORR 95-12, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, May 1995.

[18] R. Saigal, "Tracing homotopy paths in polynomial time". Talk given at ORSA/TIMS Joint National Meeting in Washington, D. C., Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, 1988.

[19] R. Saigal, *Linear Programming: A Modern Integrated Analysis*, Kluwer Academic Publishers, Norwell, MA, 1995.

[20] R. Saigal, "An infinitely summable series implementation of interior point methods", *Mathematics of Operations Research*, 20(1995) 570-616.

[21] L. L. Vandenberghe, and S. Boyd, "Semidefinite programming," manuscript, Information System Laboratory, Stanford University, CA, 1994, to appear in *SIAM Review*.

[22] Y. Zhang, "On extending primal-dual interior point algorithms from linear programming to semidefinite programming", Technical Report TR95-22, Dept. of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MA 21228-5398, October 1995.