# AN INFEASIBLE START 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 95-31

December 1995

# An Infeasible Start 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

December 21, 1995

**Abstract**

In this paper we present an infeasible start path following predictor corrector method for semidefinite linear programming problem. This method does not assume that the dual pair of semidefinite programs have feasible solutions, and, in at most $O(|\log(\frac{\epsilon}{\delta(A,b,C)\rho})|n)$ iterations of the predictor corrector method, finds either an approximate solution to the dual pair or shows that there is no optimal solution with duality gap zero in a well defined bounded region. The nonexistence of optimal solutions is detected in a finite number of iterations. Here $\epsilon$ is a measure of non-optimality and infeasibility of the pair of solutions found, and is generally chosen to be small; $\delta(A,b,C)$ is a function of the data of the problem and $\rho$ is a measure of the size of the region searched, and is generally large. The method we present generalizes a method for linear programming developed by Mizuno. We give some preliminary computational experience for this method, and compare its performance ( measured by the number of iterations ) with that of the code SP of Vandenberghe and Boyd which is

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

based on a potential reduction strategy, and the code SDPA of Fujisawa and Kojima which is based on a path following and potential reduction strategy.

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

**Abbreviated title:** Semi-definite linear programming

# 1 Introduction

This paper considers the dual pair of semidefinite linear programs:

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

and,

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

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. We assume that the matrices $A_i, i = 1, \ldots, m$ are linearly independent.

The relationship between these pair of dual problems is now well understood. In case both the primal and the dual have interior feasible solutions, there exist optimal solutions to each problem which satisfy the strong duality theorem and thus the duality gap for such solutions is zero, Alizadeh [1]. Otherwise there are examples where the duality gap is not zero and the primal or the dual problem may not attain its respective optimal solution, see for example Alizadeh [1], Freund [5]. In this paper we consider an infeasible start predictor corrector method which will find, if one exists, an optimal solution with duality gap zero, within the set

$$\{(X, S) : 0 \preceq X \preceq \rho I, 0 \preceq S \preceq \rho I\}, \tag{3}$$

when the method is started with the initial matrices $X_1 = \rho I, y^1 = 0$ and $S_1 = \rho I$. We do not assume that the primal and dual have feasible solutions and detect the non-existence of such an optimal solution, in the set (3), in a finite number of iterations. The method we present here is a generalization of the method for linear programming presented by Mizuno [14] ( we generalize algorithm 2 of his paper ). See also section 5.10 of the book Saigal [21]. We show that in at most $O(|\log(\frac{\epsilon}{\delta(A,b,C)\rho})|n)$ iterations of the predictor corrector method, where $\delta(A, b, c)$ is a function of the data $A_i, i = 1, \ldots, m, b$, and $C$, either the method will discover that there is no optimal solution with duality gap zero in the set (3) or find a solution $X(\epsilon)$, $y(\epsilon)$, $S(\epsilon)$ to the dual pair for which the duality gap $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$, and the infeasibility measures $(\sum_{i=1}^m (A_i \bullet X(\epsilon) - b_i)^2)^{\frac{1}{2}} \leq \delta(A, b, C)\epsilon$ and $\| \sum_{i=1}^m A_i y_i(\epsilon) + S(\epsilon) - C \|_F \leq \delta(A, b, C)\epsilon$. In this method, the corrector step solves a linear system of equations that differs from that of the predictor only in the right hand side.

Our method is similar to the one presented by Potra and Sheng [19], but differs in the predictor and the corrector directions used. The predictor direction used by Potra and Sheng is determined by the equation

$$X^{-\frac{1}{2}}(X\Delta S + \Delta X S)X^{\frac{1}{2}} + X^{\frac{1}{2}}(\Delta S X + S\Delta X)X^{-\frac{1}{2}} = -2X^{\frac{1}{2}}SX^{\frac{1}{2}}$$

which generates, in the case of a feasible start method, one of the directions from a class introduced by Kojima, Shindo and Hara [10] and analyzed by Monteiro [16]. We use the direction generated by the equation

$$\Delta X S + X\Delta S = \sigma t I - XS,$$

where $0 < \sigma < 1$, and use the symmetric part of $\Delta X$ in the method. The corrector direction in [19] is determined by solving a similar system as ours, but requires a solution of a new linear system, while we solve the same linear system as the predictor, but with a different right hand side. Under the assumption that the primal and the dual have optimal solutions with no duality gap, they prove a similar polynomial bound as ours on the number of predictor and corrector steps required by their method.

The predictor corrector strategy is designed to keep the iterates of the method in the following neighborhood of the central trajectory :

$$N(t_1, \beta) = \{(X, y, S) : X \succ 0, S \succ 0, t \leq t_1, \|S^{\frac{1}{2}}XS^{\frac{1}{2}} - tI\|_F \leq \beta t\} \tag{4}$$

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. This same neighborhood has been used in establishing polynomial convergence results for feasible start methods by Lin and Saigal [11], Monteiro [16] and Zhang [26].

There are several primal-dual algorithms based on the potential reduction strategy, see for example [1, 8, 10, 23], and codes based on some of these are available in the public domain, for example Fujisawa and Kojima [6] and Vanderberghe and Boyd [24]. Using a MATLAB based code of the method of this paper, we call INPC, we present some preliminary computational results on the number of iterations taken by the three codes INPC, SP [24] and SDPA [6] to solve six different problems. One of these test problems has problem instances for which the primal-dual pair has no optimal solution. All three codes were able to detect these instances. Just as in the case of linear programming where the predictor corrector path following methods are seen to be the most effective, our results lend support to the statement that a similar situation may hold for the semidefinite linear programming as well.

Besides the introduction this paper has 5 other sections. In section 2 we present the method investigated in this paper, in section 3 some basic results are presented. In section 4 we analyze the method and in section 5 we prove its global convergence. In section 6 we present the algorithm we implement, the six problems we solve, the numerical results and discuss some implementation issues. Finally in section 7 we present our conclusions.

## 2 The Method

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

**Step 0** Let $\rho > 0$ be large, $\lambda \geq 1$, $0 < \alpha < \frac{1}{\lambda}$ and $1 > \sigma_1 > \sigma > 0$ be given constants; $\theta_1 = 1$, $X_1 = \rho I$, $y^1 = 0$, $S_1 = \rho I$ and $t_1 = \frac{X_1 \bullet S_1}{n} = \rho^2$. Set $k = 1$.

**Step 1** Predictor Step: Solve

$$\begin{array}{rcll} A_i \bullet \Delta X_k & = & -\lambda(A_i \bullet X_k - b_i) & i = 1, \cdots, m \\ \sum_{i=1}^m A_i \Delta y_i^k + \Delta S_k & = & -\lambda(\sum_{i=1}^m A_i y_i^k + S_k - C) & \\ \Delta X_k S_k + X_k \Delta S_k & = & \sigma t_k I - X_k S_k & \end{array} \tag{5}$$

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

$$\begin{aligned}
\bar{X}_k &= X_k + \tfrac{1}{2}\alpha(\Delta X_k + \Delta X_k^T) \\
\bar{y}^k &= y^k + \alpha \Delta y^k \\
\bar{S}_k &= S_k + \alpha \Delta S_k \\
\bar{t}_k &= \tfrac{\bar{X}_k \bullet \bar{S}_k}{n} \\
\theta_{k+1} &= (1 - \alpha\lambda)\theta_k.
\end{aligned} \tag{6}$$

**Step 2** Corrector Step: Solve

$$\begin{array}{lllll}
A_i \bullet \Delta\bar{X}_k & & = & 0 & \text{for all } i = 1, \cdots, m \\
\sum_{i=1}^m A_i \Delta \bar{y}_i^k & + \quad \Delta \bar{S}_k & = & 0 & \\
\Delta \bar{X}_k S_k & + \quad X_k \Delta \bar{S}_k & = & \bar{t}_k I - \bar{X}_k \bar{S}_k &
\end{array} \tag{7}$$

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

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

**Step 3** Set $k = k+1$. If $3X_k \bullet S_k < \rho\theta_k(\|X_k\|_F + \|S_k\|_F)$, stop. There is no optimal solution with duality gap zero in the set $\{(X, S) : \rho I \succeq X \succeq 0, \rho I \succeq S \succeq 0\}$. Otherwise, go to Step 1.

Both the systems (5) and (7) have a unique solution, and differ only in the right hand side. Also, the solutions $\Delta S_k$ and $\Delta\bar{S}_k$ are symmetric ( a consequence of our assumption on the matrices of the problem ), but $\Delta X_k$ or $\Delta\bar{X}_k$ may not be symmetric. 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 $\lambda_i$ for $i = 1, \cdots, n$ are the $n$ real eigenvalues of $B$. For a given $n \times n$ matrix $B$, we define $\text{vec}(B) = (B_{\cdot 1}^T, B_{\cdot 2}^T, \ldots, B_{\cdot n}^T)^T$ and note that $\|\text{vec}(B)\|_2 = \|B\|_F$, where $B_{\cdot j}$ is the $j$th column of the matrix $B$.

Given an $m \times m$ matrix $A$ and an $n \times n$ matrix $B$, we define the Kronecker product, $A \otimes B$, as the matrix

$$\begin{bmatrix} A_{1,1}B & \cdots & A_{1,m}B \\ \vdots & \ddots & \vdots \\ A_{m,1}B & \cdots & A_{m,m}B \end{bmatrix}$$

and the following are some well known properties of the Kronecker product, see for example Bellman [2] :

1. $\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$.
2. $(A \otimes B)^T = A^T \otimes B^T$.
3. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
4. $(A \otimes B)(C \otimes D) = AC \otimes BD$.
5. $\lambda(A \otimes B) = \{\alpha_i \beta_j : \alpha_i \in \lambda(A), \beta_j \in \lambda(B)\}$ where $\lambda(A)$ is the set of all eigenvalues of $A$.

and the following are well known properties of trace:

1. $A \succeq 0$. Then $\text{trace}(A) \geq 0$.
2. $A \succeq B$. Then $\text{trace}(A) \geq \text{trace}(B)$.
3. $C = A + B$. Then $\text{trace}(C) = \text{trace}(A) + \text{trace}(B)$.
4. $\text{trace}(A) = \sum \lambda_i(A)$ where $\lambda_i(A)$ are the eigenvalues of $A$.
5. $\det(A) = \prod \lambda_i(A)$ where $\lambda_i(A)$ are the eigenvalues of $A$.

We state some lemmas needed in this paper. The proofs can be found in the cited references.

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

**Proof:** Lemma 1, [11]. ∎

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 $\lambda_{\min}(A) \geq \lambda_{\min}(B) - \|C\|_F$ and $\lambda_{\max}(A) \leq \lambda_{\max}(B) + \|C\|_F$.*

**Proof:** Lemma 2, [11]. ∎

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

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

**Proof:** Lemma 3.3, Monteiro [16]. ∎

Another technical lemma follows:

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

**Proof:** Lemma 4, [11]. ∎

# 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 at a linear rate $(1 - \alpha(1 - \sigma_1))$ and $\alpha = O(\frac{1}{n})$.

We now prove a basic lemma which shows that the infeasibility decreases after each predictor and corrector step.

**Lemma 5** *For each $k = 1, 2, \cdots$*

1. $A_i \bullet X_k - b_i = \theta_k(A_i \bullet X_1 - b_i)$.

2. $\sum_{i=1}^m A_i^T y_i^k + S_k - C = \theta_k(\sum_{i=1}^m A_i^T y_i^1 + S_1 - C)$.

**Proof:** We will show part 1. The proof of part 2 is similar. Now

$$\begin{aligned} A_i \bullet X_{k+1} - b_i &= A_i \bullet (X_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T) + \frac{1}{2}(\Delta \bar{X}_k + \Delta \bar{X}_k^T)) - b_i \\ &= A_i \bullet X_k + \alpha A_i \bullet \Delta X_k + A_i \bullet \Delta \bar{X}_k - b_i \\ &= (1-\alpha\lambda)(A_i \bullet X_k - b_i) \\ &= (1-\alpha\lambda)\theta_k (A_i \bullet X_1 - b_i) \end{aligned}$$

and the result follows by an induction argument. ∎

Define the $m \times n^2$ matrix $\bar{A}$ whose $i$th row is the vector $\text{vec}(A_i)^T$. Then the solution to the system (5) and (7) can be computed by solving the equivalent linear system

$$\begin{bmatrix} \bar{A} & 0 & 0 \\ 0 & \bar{A}^T & I \\ F_k & 0 & G_k \end{bmatrix} \begin{bmatrix} \text{vec}(\Delta X_k) \\ \Delta y^k \\ \text{vec}(\Delta S_k) \end{bmatrix} = \begin{bmatrix} -\lambda(\bar{A}\text{vec}(X_k) - b) \\ -\lambda(\bar{A}^T y^k + \text{vec}(S_k) - \text{vec}(C)) \\ \text{vec}(R_k) \end{bmatrix} \tag{9}$$

where $F_k = S_k \otimes I$, $G_k = I \otimes X_k$ and an appropriate matrix $R_k$.

**Lemma 6** *Let $X_k$ and $S_k$ be symmetric and positive definite, and define $D_k = F_k^{-\frac{1}{2}} G_k^{\frac{1}{2}}$. Then*

1. $D_k^2 = F_k^{-1} G_k$.
2. $D_k$ *is symmetric.*

**Proof:** Since $X_k$ and $S_k$ are symmetric and positive definite, $F_k^{\frac{1}{2}} = S_k^{\frac{1}{2}} \otimes I$ and $G_k^{\frac{1}{2}} = I \otimes X_k^{\frac{1}{2}}$ are well defined, and thus, by property (3) of Kronecker product, $D_k$ is well defined . Now part (1) follows from property (4), and, part (2) from property (2) of the Kronecker product. ∎

**Lemma 7** *There exists $\bar{U}, (v, W)$ with $U, W$ symmetric, such that $A_i \bullet U = b_i, i = 1, \ldots, m$, $\sum_{i=1}^m A_i v_i + W = C$. Let $P_k = D_k \bar{A}^T (\bar{A} D_k^2 \bar{A}^T)^{-1} \bar{A} D_k$ with $D_k$ in Lemma 6. Then, for each*

$k = 1, \ldots$

$$\begin{aligned} \Delta y^k &= -(\bar{A}D_k^2\bar{A}^T)^{-1}\bar{A}F_k^{-1}vec(-X_kS_k + \sigma t_k I) - \lambda\theta_k(\bar{A}D_k^2\bar{A}^T)^{-1}\bar{A}D_k^2 vec(S_1 - W) \\ &\quad -\lambda\theta_k(\bar{A}D_k^2\bar{A}^T)^{-1}\bar{A}F_k^{-1}(F_k vec(X_1 - W)) - \lambda\theta_k(y^1 - v), \\ D_k vec(\Delta S_k) &= P_k G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}} vec(-X_kS_k + \sigma t_k I) - \lambda\theta_k(I - P_k)D_k vec(S_1 - W) \\ &\quad +\lambda\theta_k P_k D_k^{-1} vec(X_1 - U), \\ D_k^{-1} vec(\Delta X_k) &= (I - P_k)G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}} vec(-X_kS_k + \sigma t_k I) + \lambda\theta_k(I - P_k)D_k vec(S_1 - W) \\ &\quad -\lambda\theta_k P_k D_k^{-1} vec(X_1 - U). \end{aligned}$$

**Proof:** It is readily confirmed that there exist $(v, W)$ such that $\sum_{i=1}^m A_i v_i + W = C$ and, since $A_i, i = 1, \ldots, m$ are linearly independent there exists $U$ such that $A_i \bullet U = b_i, i = 1, \ldots, m$.

Using simple algebra and the definition of $U$, $v$, $W$, system (9) can be re-written as:

$$\begin{bmatrix} \bar{A} & 0 & 0 \\ 0 & \bar{A}^T & I \\ F_k & 0 & G_k \end{bmatrix} \begin{bmatrix} \text{vec}(\Delta X_k + \lambda\theta_k(X_1 - U)) \\ \Delta y^k + \lambda\theta_k(y^1 - v) \\ \text{vec}(\Delta S_k + \lambda\theta_k(S_1 - W)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ R^k \end{bmatrix}$$

where

$$R^k = \text{vec}(-X_kS_k + \sigma t_k I) + \lambda\theta_k F_k \text{vec}(X_1 - U) + \lambda\theta_k G_k \text{vec}(S_1 - W).$$

Then

$$\begin{aligned} &\Delta y^k \\ &= (\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}(-\bar{A})F_k^{-1}R^k - \lambda\theta_k(y^1 - v) \\ &= -(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}\bar{A}F_k^{-1}\text{vec}(-X_kS_k + \sigma t_k I) - \lambda\theta_k(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}\bar{A}D_k^2\text{vec}(S_1 - W) \\ &\quad -\lambda\theta_k(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}\bar{A}F_k^{-1}(F_k\text{vec}(X_1 - U)) - \lambda\theta_k(y^1 - v), \end{aligned}$$

$$\begin{aligned} &D_k\text{vec}(\Delta S_k) \\ &= D_k(-\bar{A}^T(\Delta y^k + \lambda\theta(y^1 - v)) - \lambda\theta_k\text{vec}(S_1 - W)) \\ &= D_k(\bar{A}^T(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}\bar{A}F_k^{-1}\text{vec}(-X_kS_k + \sigma t_k I) + \lambda\theta_k\bar{A}^T(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}\bar{A}D_k^2 \\ &\quad \text{vec}(S_1 - W) + \lambda\theta_k\bar{A}^T(\bar{A}F_k^{-1}G_k\bar{A}^T)^{-1}\bar{A}F_k^{-1}(F_k\text{vec}(X_1 - U)) - \lambda\theta_k\text{vec}(S_1 - W)) \\ &= P_k G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}}\text{vec}(-X_kS_k + \sigma t_k I) - \lambda\theta_k(I - P_k)D_k\text{vec}(S_1 - W) \\ &\quad +\lambda\theta_k P_k D_k^{-1}\text{vec}(X_1 - U), \end{aligned}$$

$$
\begin{aligned}
& D_k^{-1}\text{vec}(\Delta X_k) \\
= \ & D_k^{-1}(-\lambda\theta_k \text{vec}(X_1 - U) + F_k^{-1}(R^k - G_k \text{vec}(\Delta S_k + \lambda\theta_k(S_1 - W)))) \\
= \ & D_k^{-1}(-\lambda\theta_k \text{vec}(X_1 - U)) + D_k^{-1}F_k^{-1}(\text{vec}(-X_kS_k + \sigma t_k I) + \lambda\theta_k F_k \text{vec}(X_1 - U) \\
& +\lambda\theta_k G_k \text{vec}(S_1 - W)) - D_k^{-1}F_k^{-1}G_k(\text{vec}(\Delta S_k) + \lambda\theta_k \text{vec}(S_1 - W)) \\
= \ & D_k^{-1}F_k^{-1}\text{vec}(-X_kS_k + \sigma t_k I) - D_k \text{vec}(\Delta S_k) \\
= \ & D_k^{-1}F_k^{-1}\text{vec}(-X_kS_k + \sigma t_k I) - P_k G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}} \text{vec}(-X_kS_k + \sigma t_k I) \\
& +\lambda\theta_k(I - P_k)D_k \text{vec}(S_1 - W) - \lambda\theta_k P_k D_k^{-1}\text{vec}(X_1 - U) \\
= \ & (I - P_k)G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}} \text{vec}(-X_kS_k + \sigma t_k I) + \lambda\theta_k(I - P_k)D_k \text{vec}(S_1 - W) \\
& -\lambda\theta_k P_k D_k^{-1}\text{vec}(X_1 - U),
\end{aligned}
$$

and we are done. ■

We now prove two important propositions.

**Proposition 8** *Let $U$ and $W$ be as in Lemma 7, with $-\rho I \preceq U \preceq \rho I$ and $-\rho I \preceq W \preceq \rho I$. For some $k$, $(X_k, S_k) \in N(t_k, \beta)$, $E_k = S_k^{\frac{1}{2}} X_k S_k^{\frac{1}{2}} - t_k I$, and $\theta_k \rho(\|X_k\|_F + \|S_k\|_F) \leq 3X_k \bullet S_k$. Then*

$$
\begin{aligned}
\|X_k^{-\frac{1}{2}} \Delta X_k S_k^{\frac{1}{2}}\|_F &\leq B(n)\sqrt{t_k} \\
\|X_k^{\frac{1}{2}} \Delta S_k S_k^{-\frac{1}{2}}\|_F &\leq B(n)\sqrt{t_k} \\
\|X_k^{-\frac{1}{2}} \Delta X_k^T S_k^{\frac{1}{2}}\|_F &\leq \sqrt{\frac{1+\beta}{1-\beta}} B(n)\sqrt{t_k}
\end{aligned}
$$

*where $B(n) = \frac{1}{\sqrt{1-\beta}}(\beta + (1-\sigma)\sqrt{n} + 6\lambda n)$ is an increasing function of $n$.*

**Proof:** We obtain the following results by using the lemmas 1 and 2 and the fact $S_1 - W \succeq 0, X_1 - U \succeq 0$.

$$
\begin{aligned}
\|P_k G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}} \text{vec}(-X_kS_k + \sigma t_k I)\|_2 &\leq \|G_k^{-\frac{1}{2}} F_k^{-\frac{1}{2}} \text{vec}(-X_kS_k + \sigma t_k I)\|_2 \\
&= \|X_k^{-\frac{1}{2}}(-X_kS_k + \sigma t_k I) S_k^{-\frac{1}{2}}\|_F \\
&= \|X_k^{-\frac{1}{2}} S_k^{-\frac{1}{2}}(-S_k^{\frac{1}{2}} X_k S_k^{\frac{1}{2}} + t_k I) + (\sigma - 1)t_k X_k^{-\frac{1}{2}} S_k^{-\frac{1}{2}}\|_F \\
&\leq \frac{\beta t_k}{\sqrt{t_k - \|E_k\|_F}} + (1-\sigma)\frac{\sqrt{n}t_k}{\sqrt{t_k - \|E_k\|_F}},
\end{aligned}
$$

$$
\begin{aligned}
\|D_k \mathrm{vec}(S_1 - W)\|_2 &= \|X_k^{\frac{1}{2}}(S_1 - W)S_k^{-\frac{1}{2}}\|_F \\
&= \|X_k^{\frac{1}{2}}(S_1 - W)X_k^{\frac{1}{2}}X_k^{-\frac{1}{2}}S_k^{-\frac{1}{2}}\|_F \\
&\leq \|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k(S_1 - W)\|_F \\
&\leq \|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}}\|S_1 - W\|_2\|X_k\|_F \\
&\leq 2\rho\|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k\|_F, \\
\|D_k^{-1}\mathrm{vec}(X_1 - U)\|_2 &= \|G_k^{-\frac{1}{2}}F_k^{\frac{1}{2}}\mathrm{vec}(X_1 - U)\|_2 \\
&= \|(S_k^{\frac{1}{2}} \otimes X_k^{-\frac{1}{2}})\mathrm{vec}(X_1 - U)\|_2 \\
&= \|X_k^{-\frac{1}{2}}(X_1 - U)S_k^{\frac{1}{2}}\|_F \\
&\leq 2\rho\|S_k\|_F\|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\|S_k^{-\frac{1}{2}}\Delta S_k X_k^{\frac{1}{2}}\|_F &= \|D_k \mathrm{vec}(\Delta S_k)\|_2 \\
&\leq \|G_k^{-\frac{1}{2}}F_k^{-\frac{1}{2}}\mathrm{vec}(-X_k S_k + \sigma t_k I)\|_2 + \lambda\theta_k\|D_k \mathrm{vec}(S_1 - W)\|_2 \\
&\quad + \lambda\theta_k\|D_k^{-1}\mathrm{vec}(X_1 - U)\|_2 \\
&\leq \frac{1}{\sqrt{1-\beta}}(\beta + (1-\sigma)\sqrt{n})\sqrt{t_k} + \frac{2\lambda\theta_k\rho}{\sqrt{t_k - \|E_k\|_F}}(\|X_k\|_F + \|S_k\|_F) \\
&\leq \frac{1}{\sqrt{1-\beta}}(\beta + (1-\sigma)\sqrt{n} + 6\lambda n)\sqrt{t_k}, \\
\|X_k^{-\frac{1}{2}}\Delta X_k S_k^{\frac{1}{2}}\|_F &= \|(S_k^{\frac{1}{2}} \otimes X_k^{-\frac{1}{2}})\mathrm{vec}(\Delta X_k)\|_2 \\
&= \|D_k^{-1}\mathrm{vec}(\Delta X_k)\|_2 \\
&\leq \|G_k^{-\frac{1}{2}}F_k^{-\frac{1}{2}}\mathrm{vec}(-X_k S_k + \sigma t_k I)\|_2 + \lambda\theta_k\|D_k \mathrm{vec}(S_1 - W)\|_2 \\
&\quad + \lambda\theta_k\|D_k^{-1}\mathrm{vec}(X_1 - U)\|_2 \\
&\leq \frac{1}{\sqrt{1-\beta}}(\beta + (1-\sigma)\sqrt{n} + 6\lambda n)\sqrt{t_k}, \\
\|X_k^{-\frac{1}{2}}\Delta X_k^T S_k^{\frac{1}{2}}\|_F &= \|(X_k^{-\frac{1}{2}}S_k^{-\frac{1}{2}})S_k^{\frac{1}{2}}\Delta X_k^T X_k^{-\frac{1}{2}}(X_k^{\frac{1}{2}}S_k^{\frac{1}{2}})\|_F \\
&\leq \|X_k^{\frac{1}{2}}X_k S_k^{\frac{1}{2}}\|_2^{\frac{1}{2}}\|X_k^{-\frac{1}{2}}\Delta X_k S_k^{\frac{1}{2}}\|_F\|S_k^{-\frac{1}{2}}X_k^{-1}S_k^{-\frac{1}{2}}\|_2^{\frac{1}{2}} \\
&\leq \sqrt{\frac{t_k + \|E_k\|_F}{t_k - \|E_k\|_F}}B(n)\sqrt{t_k} \\
&\leq \sqrt{\frac{1+\beta}{1-\beta}}B(n)\sqrt{t_k}.
\end{aligned}
$$

■

**Corollary 9** *For some $k$ let $t_k > 0, (X_k, S_k) \in N(t_k, \beta)$, and $\theta_k \rho(\|X_k\|_F + \|S_k\|_F) \leq 3X_k \bullet S_k$. Then*

$$(1 - \alpha + \sigma\alpha - \frac{B(n)^2}{n}\alpha^2)t_k \leq \bar{t}_k \leq (1 - \alpha + \sigma\alpha + \frac{B(n)^2}{n}\alpha^2)t_k.$$

*That is,*

$$\sigma\alpha - \frac{B(n)^2}{n}\alpha^2 \leq \quad \frac{\bar{t}_k}{t_k} - 1 + \alpha \quad \leq \sigma\alpha + \frac{B(n)^2}{n}\alpha^2, \tag{10}$$

$$-\frac{B(n)^2}{n}\alpha^2 \leq \frac{\bar{t}_k}{t_k} - 1 + \alpha - \sigma\alpha \leq \frac{B(n)^2}{n}\alpha^2. \tag{11}$$

**Proof:** Note that $n\bar{t}_k = \bar{X}_k \bullet \bar{S}_k$, and $\Delta S_k \bullet \Delta X_k = \Delta S_k \bullet \Delta X_k^T = \Delta X_k \bullet \Delta S_k$ ( since $\Delta S_k$ is symmetric ). Using (6), we see that

$$n\bar{t}_k = \bar{X}_k \bullet \bar{S}_k = X_k \bullet S_k + \alpha(\Delta X_k \bullet S_k + \Delta S_k \bullet X_k) + \alpha^2 \Delta X_k \bullet \Delta S_k.$$

Also, from (5), $\Delta X_k \bullet S_k + \Delta S_k \bullet X_k = n\sigma t_k - nt_k$. Now $\Delta X_k \bullet \Delta S_k = \text{trace}(\Delta X_k^T \Delta S_k) = \text{trace}(S_k^{\frac{1}{2}} \Delta X_k^T X_k^{-\frac{1}{2}} X_k^{\frac{1}{2}} \Delta S_k S_k^{-\frac{1}{2}})$ and the result follows from Proposition 8. ■

**Proposition 10** *For some $k$ let $(X_k, S_k) \in N(t_k, \beta)$, $E_k = S_k^{\frac{1}{2}} X_k S_k^{\frac{1}{2}} - t_k I$, and $\theta_k \rho(\|X_k\|_F + \|S_k\|_F) \leq 3X_k \bullet S_k$. Then*

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

*where $M(\alpha, n) = \frac{(1-\alpha+2\alpha\sigma)\beta}{\sqrt{1-\beta}} + \frac{\alpha B(n)}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}}) + \frac{\alpha^2 B(n)^2}{\sqrt{1-\beta}}(1 + \frac{1}{\sqrt{n}} + \frac{2\beta}{n})$ is an increasing function of $\alpha$.*

**Proof:** We will only show this for $\|X_k^{-\frac{1}{2}} \Delta \bar{X}_k S_k^{\frac{1}{2}}\|_F$. The first inequality below follows from Lemma 4, and using the same arguments as those of Proposition 8, and inequalities (10), (11), we obtain the other inequalities below :

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

Now, the second term in the above expression is the same as:

$$\begin{aligned}
&\|X_k^{-\frac{1}{2}}(\frac{\bar{t}_k}{t_k}X_k S_k - X_k S_k - \alpha(\Delta X_k S_k + X_k \Delta S_k) - \frac{\alpha}{2}(\Delta X_k^T - \Delta X_k)S_k - \\
&\frac{\alpha^2}{2}(\Delta X_k + \Delta X_k^T)\Delta S_k)S_k^{-\frac{1}{2}}\|_F \\
= \; &\|X_k^{-\frac{1}{2}}((\frac{\bar{t}_k}{t_k} - 1 + \alpha)(X_k S_k - t_k I) + (\frac{\bar{t}_k}{t_k} - 1 + \alpha - \alpha\sigma)t_k I - \frac{\alpha}{2}(\Delta X_k^T - \Delta X_k)S_k \\
&-\frac{\alpha^2}{2}(\Delta X_k + \Delta X_k^T)\Delta S_k)S_k^{-\frac{1}{2}}\|_F \\
\leq \; &(\alpha\sigma + \frac{\alpha^2 B(n)^2}{n})\frac{\beta\sqrt{t_k}}{\sqrt{1-\beta}} + \frac{\alpha^2 B(n)^2}{n}\sqrt{\frac{n t_k}{1-\beta}} + \frac{\alpha}{2}(1 + \sqrt{\frac{1+\beta}{1-\beta}})B(n)\sqrt{t_k} \\
&+\alpha^2 B(n)^2\sqrt{\frac{t_k}{1-\beta}}.
\end{aligned}$$

We obtain the result by substituting the above, and the definitions of $\hat{\alpha}$ and $M(\alpha, n)$. ∎

**Lemma 11** *If $\alpha^*$, $\sigma_1$ satisfy the following*

$$\begin{aligned}
\sigma - 2\frac{M(\alpha^*, n)B(n)}{n} - \alpha^*\frac{B(n)^2}{n} &\geq 0, \\
\sigma + 2\frac{M(\alpha^*, n)B(n)}{n} + \alpha^*\frac{B(n)^2}{n} &\leq \sigma_1,
\end{aligned}$$

*then for every $\alpha$ such that $0 < \alpha \leq \alpha^*$, and $k$ with $(X_k, S_k) \in N(t_k, \beta)$, and $\theta_k\rho(\|X_k\|_F + \|S_k\|_F) \leq 3X_k \bullet S_k$,*

$$(1-\alpha)X_k \bullet S_k \leq X_{k+1} \bullet S_{k+1} \leq (1 - \alpha(1-\sigma_1))X_k \bullet S_k.$$

**Proof:**

$$X_{k+1} \bullet S_{k+1}$$

$$
\begin{aligned}
&= (\bar{X}_k + \frac{\Delta\bar{X}_k + \Delta\bar{X}_k^T}{2}) \bullet (\bar{S}_k + \Delta\bar{S}_k) \\
&= \bar{X}_k \bullet \bar{S}_k + \Delta\bar{X}_k \bullet (S_k + \alpha\Delta S_k) + (X_k + \alpha\frac{\Delta X_k + \Delta X_k^T}{2}) \bullet \Delta\bar{S}_k \\
&= \bar{X}_k \bullet \bar{S}_k + \alpha(\Delta\bar{X}_k \bullet \Delta S_k + \Delta X_k \bullet \Delta\bar{S}_k) \;\; (Note: \Delta\bar{X}_k \bullet S_k + X_k \bullet \Delta\bar{S}_k = 0) \\
&= (1 - \alpha + \sigma\alpha) X_k \bullet S_k + \alpha^2 \Delta X_k \bullet \Delta S_k + \alpha(\Delta\bar{X}_k \bullet \Delta S_k + \Delta X_k \bullet \Delta\bar{S}_k).
\end{aligned}
$$

Since

$$
\begin{aligned}
\Delta\bar{X}_k \bullet \Delta S_k &= \mathrm{vec}(\Delta\bar{X}_k)^T \mathrm{vec}(\Delta S_k) \\
&= \mathrm{vec}(\Delta\bar{X}_k)^T D_k^{-1} D_k \mathrm{vec}(\Delta S_k) \\
&\leq \|D_k^{-1}\mathrm{vec}(\Delta\bar{X}_k)\|_2 \|D_k \mathrm{vec}(\Delta S_k)\|_2 \\
&= \|X_k^{-\frac{1}{2}} \Delta\bar{X}_k S_k^{\frac{1}{2}}\|_F \|X_k^{\frac{1}{2}} \Delta S_k S_k^{-\frac{1}{2}}\|_F \\
&\leq M(\alpha, n) B(n) t_k, \\
\Delta X_k \bullet \Delta\bar{S}_k &\leq M(\alpha, n) B(n) t_k, \\
\Delta X_k \bullet \Delta S_k &\leq B(n)^2 t_k,
\end{aligned}
$$

we have

$$
X_{k+1} \bullet S_{k+1} \leq (1 - \alpha + \sigma\alpha + 2\alpha\frac{M(\alpha, n)B(n)}{n} + \alpha^2\frac{B(n)^2}{n}) X_k \bullet S_k,
$$
$$
X_{k+1} \bullet S_{k+1} \geq (1 - \alpha + \sigma\alpha - 2\alpha\frac{M(\alpha, n)B(n)}{n} - \alpha^2\frac{B(n)^2}{n}) X_k \bullet S_k.
$$

Since $M(\alpha, n)$ is an increasing function of $\alpha$, we have

$$
(1 - \alpha) X_k \bullet S_k \leq X_{k+1} \bullet S_{k+1} \leq (1 - \alpha(1 - \sigma_1)) X_k \bullet S_k
$$

for $0 < \alpha \leq \alpha^*$. ■

We now obtain a condition that guarantees that $X_{k+1}$ and $S_{k+1}$ are symmetric and positive definite.

**Lemma 12** *For some $k$ let $(X_k, S_k) \in N(t_k, \beta)$, $E_k = S_k^{\frac{1}{2}} X_k S_k^{\frac{1}{2}} - t_k I$, $\theta_k \rho(\|X_k\|_F + \|S_k\|_F) \leq 3X_k \bullet S_k$ and*

$$
1 - (\beta + \sqrt{\beta + 1}(\alpha B(n) + M(\alpha, n))) > 0.
$$

*Then $X_{k+1}$ and $S_{k+1}$ are symmetric and positive definite.*

**Proof:** We now show this for $X_{k+1}$. It is symmetric by construction, and is positive definite if and only if $S_k^{\frac{1}{2}} X_{k+1} S_k^{\frac{1}{2}}$ is. We note that

$$S_k^{\frac{1}{2}} X_{k+1} S_k^{\frac{1}{2}} = S_k^{\frac{1}{2}} (X_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T) + \frac{1}{2}(\Delta \bar{X}_k + \Delta \bar{X}_k^T)) S_k^{\frac{1}{2}}.$$

Then

$$\lambda_{\min}(S_k^{\frac{1}{2}} X_{k+1} S_k^{\frac{1}{2}}) \geq t_k - \|E_k\|_F - \frac{1}{2}\|S_k^{\frac{1}{2}}((\alpha(\Delta X_k + \Delta X_k^T) + (\Delta \bar{X}_k + \Delta \bar{X}_k^T) S_k^{\frac{1}{2}}\|_F.$$

Using the result of Propositions 8 and 10 we obtain

$$\lambda_{\min}(S_k^{\frac{1}{2}} X_{k+1} S_k^{\frac{1}{2}}) \geq (1 - \beta - \sqrt{1+\beta}(\alpha B(n) + M(\alpha, n)))t_k$$

and our result follows. To show $S_{k+1}$ is positive definite, consider $X_k^{\frac{1}{2}} S_{k+1} X_k^{\frac{1}{2}}$ . Using the same argument as above, the lemma follows. ■

We will now obtain a lemma used for showing that $X_{k+1}$ and $S_{k+1}$ are in the neighborhood $N(t_{k+1}, \beta)$.

**Lemma 13** *Let the conditions of Lemma 12 be satisfied. Then*

$$\begin{aligned} & \|S_{k+1}^{\frac{1}{2}} X_{k+1} S_{k+1}^{\frac{1}{2}} - t_{k+1} I\|_F \\ \leq \; & \frac{1}{2(1-\alpha)}(1 + \sqrt{\frac{1+\beta}{1-\beta}})(2\alpha B(n) M(\alpha, n) + M(\alpha, n)^2) t_{k+1} + \frac{2\alpha B(n) M(\alpha, n)}{(1-\alpha)\sqrt{n}} t_{k+1}. \end{aligned}$$

**Proof:** By simple algebra we see that

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

where $H_k = \frac{\alpha}{2}(\Delta \bar{X}_k^T + \Delta \bar{X}_k)\Delta S_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T)\Delta \bar{S}_k + \frac{\Delta \bar{X}_k + \Delta \bar{X}_k^T}{2} \Delta \bar{S}_k$, and using the symmetry of $\bar{S}_k$ and $\Delta \bar{S}_k$, we see that

$$\begin{aligned} n t_{k+1} = X_{k+1} \bullet S_{k+1} &= \bar{X}_k \bullet \bar{S}_k + \Delta \bar{X}_k \bullet \bar{S}_k + \bar{X}_k \bullet \Delta \bar{S}_k \\ &= \bar{X}_k \bullet \bar{S}_k + \Delta \bar{X}_k \bullet S_k + X_k \bullet \Delta \bar{S}_k + \alpha(\Delta \bar{X}_k \bullet \Delta S_k + \Delta X_k \bullet \Delta \bar{S}_k) \\ &= n \bar{t}_k + \alpha(\Delta \bar{X}_k \bullet \Delta S_k + \Delta X_k \bullet \Delta \bar{S}_k). \end{aligned}$$

Then, from Lemma 3

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

Now, for $\Delta X = \Delta \bar{X}_k$ or $\Delta X_k$ and $\Delta S = \Delta \bar{S}_k$ or $\Delta S_k$ we note that $\Delta X \bullet \Delta S = \Delta X^T \bullet \Delta S$ =trace$(\Delta X^T \Delta S)$ =trace$(S_k^{\frac{1}{2}} \Delta X^T \Delta S S_k^{-\frac{1}{2}}) \leq \|S_k^{\frac{1}{2}} \Delta X^T X_k^{-\frac{1}{2}}\|_F \|X_k^{\frac{1}{2}} \Delta S S_k^{-\frac{1}{2}}\|_F$. Also

$$
\begin{aligned}
\|S_k^{-\frac{1}{2}} \Delta X \Delta S S_k^{-\frac{1}{2}}\|_F &= \|S_k^{\frac{1}{2}} X_k^{\frac{1}{2}} X_k^{-\frac{1}{2}} \Delta X S_k^{\frac{1}{2}} S_k^{-\frac{1}{2}} \Delta S X_k^{\frac{1}{2}} X_k^{-\frac{1}{2}} S_k^{-\frac{1}{2}}\|_F \\
&\leq \sqrt{\frac{1+\beta}{1-\beta}} \|X_k^{-\frac{1}{2}} \Delta X S_k^{\frac{1}{2}}\|_F \|S_k^{-\frac{1}{2}} \Delta S X_k^{\frac{1}{2}}\|_F
\end{aligned}
$$

and the lemma follows from Propositions 8 and 10, and Lemma 11. ∎

# 5 Main Convergence Theorem

We now show that the infeasible start method presented in Section 2 converges to an optimal solution or discovers that there is no optimal solution with duality gap zero in a predefined region. We first show the later in the next theorem.

**Theorem 14** *Let $\alpha^*$ be as in Lemma 11, and $\alpha \leq \alpha^*$. Also, for some $k$, let $\|X_k\|_F + \|S_k\|_F > \frac{3 X_k \bullet S_k}{\rho \theta_k}$ but for $l < k$, $\|X_l\|_F + \|S_l\|_F \leq \frac{3 X_l \bullet S_l}{\rho \theta_k}$. Then there is no optimal solution with duality gap zero in the set $\{(X, S) \mid \rho I \succeq X \succeq 0, \rho I \succeq S \succeq 0\}$.*

**Proof:** From Lemmas 5, 11 and mathematical induction, for $l < k$ we have

$$
\begin{aligned}
\theta_{l+1} \rho^2 n &= (1 - \alpha\lambda) \theta_l \rho^2 n \\
&\leq (1 - \alpha\lambda) X_l \bullet S_l \\
&\leq (1 - \alpha) X_l \bullet S_l \ \ (Note : \lambda \geq 1) \\
&\leq X_{l+1} \bullet S_{l+1}.
\end{aligned}
$$

Now, assume the contrary, and let optimal solution $\rho I \succeq X^*$, $\rho I \succeq S^*$ exist with $X^* \bullet S^* = 0$. Also, for $\hat{X} = \theta_k X_1 + (1-\theta_k)X^*$ and $\hat{y} = \theta_k y^1 + (1-\theta_k)y^*$ and $\hat{S} = \theta_k S_1 + (1-\theta_k)S^*$ using Lemma 5, it can be seen that

$$\begin{aligned} A_i \bullet (X_k - \hat{X}) &= 0 \text{ for every } i = 1, \cdots, m, \\ \sum_{i=1}^{n} (y_i^k - \hat{y}_i) A_i + S_k - \hat{S} &= 0 \end{aligned}$$

and thus we have $(\hat{X} - X_k) \bullet (\hat{S} - S_k) = 0$.

Note that

$$\|X_k\|_F = \sqrt{\text{trace}(X_k^2)} = \sqrt{\sum_{i=1}^{n} \lambda_i^2} \leq \sum_{i=1}^{n} \lambda_i = I \bullet X_k$$

and we have

$$\begin{aligned} & \theta_k \rho(\|X_k\|_F + \|S_k\|_F) \\ \leq\ & \theta_k \rho(I \bullet X_k + I \bullet S_k) \ \ (\text{Note}: X_1 = \rho I, S_1 = \rho I) \\ \leq\ & (\theta_k X_1 \bullet S_k + (1-\theta_k) X^* \bullet S_k) + (\theta_k S_1 \bullet X_k + (1-\theta_k) S^* \bullet X_k) \\ =\ & \hat{X} \bullet S_k + \hat{S} \bullet X_k \\ =\ & \hat{X} \bullet \hat{S} + X_k \bullet S_k \\ =\ & \theta_k^2 X_1 \bullet S_1 + \theta_k(1-\theta_k)(X_1 \bullet S^* + S_1 \bullet X^*) + X_k \bullet S_k \\ \leq\ & \theta_k^2 n\rho^2 + 2\theta_k(1-\theta_k) n\rho^2 + X_k \bullet S_k \\ \leq\ & 2\theta_k n\rho^2 + X_k \bullet S_k \\ \leq\ & 3X_k \bullet S_k. \end{aligned}$$

This causes contradiction. ∎

We are now ready to prove the main convergence theorem.

**Theorem 15** *Let $\alpha \leq \frac{1}{400n}$, $\beta = 0.005$, $\lambda = 1$, $\sigma = 0.5$ and $\sigma_1 = 0.9$. Then for every $k = 1, 2, \cdots$, $(X_k, y^k, S_k) \in N(t_k, \beta)$. Thus, if $X_1 = \rho I, y^1 = 0, S_1 = \rho I, \rho \geq 1$, for every $\epsilon > 0$, after at most $O(|\log(\frac{\epsilon}{\delta(A,b,C)\rho})|n)$ iterations of the predictor-corrector method either the method stops by detecting $\|X_k\|_F + \|S_k\|_F > \frac{3X_k \bullet S_k}{\rho\theta_k}$ and the problem has no optimal solution with duality gap zero in the set $\{(X, S) | \rho I \succeq X \succeq 0, \rho I \succeq S \succeq 0\}$; or, a* *solution* $(X(\epsilon), y(\epsilon), S(\epsilon))$ *is found with* $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$, $\sqrt{\sum_{i=1}^m (A_i \bullet X(\epsilon) - b_i)^2} \leq \epsilon$, *and* $\| \sum_{i=1}^m y(\epsilon)_i A_i + S(\epsilon) - C \|_F \leq \epsilon$.

**Proof:** Let $\alpha^* = \frac{1}{400n}$ in Lemma 11. It's readily confirmed that this lemma holds and the Theorem 14 follows.

Now assume the method continues without detecting infeasibility. It is readily confirmed that the above choice satisfies the properties required in the hypothesis of the Lemma 12, $\|E_{k+1}\|_F \leq \beta t_{k+1}$, $t_{k+1} \leq (1 - \alpha(1 - \sigma_1))t_k$, and thus the sequence belongs to the neighborhood $N(t_k, \beta)$.

Now let $\epsilon > 0$ be arbitrary, $k$ sufficiently large and $\delta(A, b, C)$ so that $(1-\alpha(1-\sigma_1))^k t_1 = (1 - \alpha(1 - \sigma_1))^k \rho^2 \leq (1 - \alpha(1 - \sigma_1))^k \rho^2 \delta(A, b, C) \leq \epsilon$, $(1 - \alpha\lambda)^k \sqrt{\sum_{i=1}^m (A_i \bullet X_1 - b_i)^2} \leq (1 - \alpha\lambda)^k \rho \sqrt{\sum_{i=1}^m (A_i \bullet I - \frac{b_i}{\rho})^2} \leq (1 - \alpha\lambda)^k \rho\delta(A, b, C)\epsilon$, and $(1 - \alpha\lambda)^k \| \sum_{i=1}^m A_i y_i^1 + S_1 - C\|_F \leq (1 - \alpha\lambda)^k \rho \|I - \frac{C}{\rho}\|_F \leq (1 - \alpha\lambda)^k \rho\delta(A, b, C)\epsilon$. With $\alpha(1 - \sigma_1) \leq \frac{1}{400}$ and $\alpha\lambda \leq \frac{1}{400}$, $\log(\frac{\epsilon}{\rho^2\delta(A,b,C)}) \geq k\log(1-\alpha(1-\sigma_1)) \geq -k\frac{3\alpha(1-\sigma_1)}{2}$ and $\log(\frac{\epsilon}{\rho\delta(A,b,C)}) \geq k\log(1-\alpha\lambda) \geq -k\frac{3\alpha\lambda}{2}$.

Thus, for $k \geq \max\{\frac{-2}{3\alpha(1-\sigma_1)} \log(\frac{\epsilon}{\rho^2\delta(A,b,C)}), \frac{-2}{3\alpha\lambda} \log(\frac{\epsilon}{\rho\delta(A,b,C)})\} + 1$, $X_k \bullet S_k \leq n\epsilon$, $\sqrt{\sum_{i=1}^m (A_i \bullet X_k - b_i)^2} \leq \epsilon$, and $\| \sum_{i=1}^m y_i^k A_i + S_k - C\|_F \leq \epsilon$, and the required result is obtained. ■

# 6 An Implementation

## 6.1 Introduction

In this section we will present some preliminary results of an implementation of a path following predictor corrector method presented in this paper. In a linear programming situation, nearly all efficient codes ( for example, [12, 13, 25, 27]) implement a predictor corrector path following method, and these have been found to be very effective. For semidefinite linear programming, due to the nonlinear nature of the problem, the numerical behavior of these methods may be different from that for linear programming. We implement a variant of the method presented in the section 2 within the MATLAB environment and compare the number of its iterations with those attained by the two codes, SP of Vandenberghe and

Boyd [24] and SDPA of Fujisawa and Kojima [6]. Our results indicate that for small size problems, this strategy may be good for semidefinite linear programming as well.

Semidefinite programming is used for solving some specific problems, for example [7, 9]. In [7], a version of a path following method is used to solve some semidefinite linear programs that arise as bounding problems in a branch and bound strategy for some combinatorial optimization problems. To date, no systematic tests on predictor corrector path following method for semidefinite programming have been performed. In this section we report numerical tests on 6 different types of problems and compare them with two semidefinite programming solvers SDPA [6] and SP [24]. These solvers use a potential reduction strategy, with SDPA basically a path-following method which utilizes the logarithmic barrier function as a merit function when a corrector step is performed.

## 6.2 Algorithm

We now present the specific algorithm we will implement. For a given matrix $A$, we define $|A| = \max_{i,j}|A_{i,j}|$.

**Step 0** $\rho = 1000\max\{|A_1|,\cdots,|A_m|,|b|,|C|\}$. $X_1 = \rho I, y^1 = 0, S_1 = \rho I$. $\epsilon > 0$, $t_1 = \frac{X_1 \bullet S_1}{n}$, $\beta = 1$, $\lambda = 1.0, \theta_1 = 1$ and $\sigma = 0$.

**Step 1** Predictor Step: Solve

$$\begin{aligned}
\sum_{j=1}^{m} A_i \bullet (S_k^{-1} A_j X_k)\Delta y_j^k &= -A_i \bullet S_k^{-1}(\sigma t_k I - S_k X_k + \lambda(\sum_{j=1}^{m} y_j^k A_j + S_k - C)X_k) \\
&\quad -\lambda(A_i \bullet X_k - b_i), \quad i = 1,\ldots,m \\
\Delta S_k &= -\sum_{i=1}^{m} \Delta y_i^k A_i - \lambda(\sum_{i=1}^{m} y_i^k A_i + S_k - C) \\
\Delta X_k^T &= S_k^{-1}(\sigma t_k I - S_k X_k - \Delta S_k X_k)
\end{aligned}$$

**Step 2** Step Selection:

Set $l = 1$ and $\alpha_1 = \max\{\alpha : X_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T) \succeq 0, S_k + \alpha\Delta S_k \succeq 0, \alpha \leq 1\}$. If $\alpha_1 = 1$ a feasible solution has been found. Set $\alpha_1 = 0.99\alpha_1$.

**Step 3** Set

$$X_{k,l} = X_k + \frac{1}{2}\alpha_l(\Delta X_k + \Delta X_k^T)$$

$$y^{k,l} = y^k + \alpha_l \Delta y^k$$
$$S_{k,l} = S_k + \alpha_l \Delta S_k$$

**Step 4** Corrector Step: Set $\bar{X}_k = X_{k,l}$, $\bar{y}^k = y^{k,l}$, $\bar{S}_k = S_{k,l}$, $\bar{t}_k = \frac{\bar{X}_k \bullet \bar{S}_k}{n}$ and solve

$$\begin{aligned}
\sum_{j=1}^{m} A_i \bullet (S_k^{-1} A_j X_k) \Delta \bar{y}_j^k &= -A_i \bullet S_k^{-1}(\bar{t}_k I - \bar{S}_k \bar{X}_k) \\
\Delta \bar{S}_k &= -\sum_{i=1}^{m} \Delta \bar{y}_i^k A_i \\
\Delta \bar{X}_k^T &= S_k^{-1}(\bar{t}_k I - \bar{S}_k \bar{X}_k - \Delta \bar{S}_k X_k)
\end{aligned}$$

**Step 5** Set

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

**Step 6** If $(X_{k+1,l}, y^{k+1,l}, S_{k+1,l}) \in N(t_{k+1,l}, \beta)$, then set $X_{k+1} = X_{k+1,l}$, $y^{k+1} = y^{k+1,l}$, and $S_{k+1} = S_{k+1,l}$, $t_{k+1} = t_{k+1,l}$, $\theta_{k+1} = (1-\alpha_l)\theta_k$, and go to Step 7. Otherwise, if $l \leq 2$, set $\alpha_{l+1} = \frac{1}{2}\alpha_l$, $\beta = 2\beta$, $l = l+1$, and go to step 3. If $l > 2$, select $\alpha_{l+1} = 0.99\max\{\alpha : X_k + \frac{\alpha}{2}(\Delta X_k + \Delta X_k^T + \Delta \bar{X}_k + \Delta \bar{X}_k^T) \succeq 0, S_k + \alpha(\Delta S_k + \Delta \bar{S}_k) \succeq 0\}$, and set

$$\begin{aligned}
X_{k+1} &= X_k + \frac{\alpha_{l+1}}{2}(\Delta X_k + \Delta X_k^T + \Delta \bar{X}_k + \Delta \bar{X}_k^T) \\
y^{k+1} &= y^k + \alpha_{l+1}(\Delta y_k + \Delta \bar{y}_k) \\
S_{k+1} &= S_k + \alpha_{l+1}(\Delta S_k + \Delta \bar{S}_k) \\
t_{k+1} &= \frac{X_{k+1} \bullet S_{k+1}}{n} \\
\theta_{k+1} &= (1-\alpha_{l+1})\theta_k
\end{aligned}$$

**Step 7** $k = k+1$. If $3X_k \bullet S_k < \rho\theta_k(\|X_k\|_F + \|S_k\|_F)$, stop. There is no optimal solution with duality gap zero in the set $\{(X,S) : \rho I \succeq X \succeq 0, \rho I \succeq S \succeq 0\}$. Otherwise, go to step 1.

## 6.3 Implementation Issues

The method presented in the Section 2 implements a step size that is too small to be practical. A practical algorithm is presented in the subsection 6.2. In step 2 of the algorithm presented there, step size chosen during the predictor step is .99 of the largest step size that preserves the positive definiteness of the resulting matrices. If, after the corrector step the resulting iterates lie in the required neighborhood, this step size is accepted. Otherwise, it is reduced to half its previous value and used again in a predictor step. This process is continued until the iterate after the corrector step lies in the required neighborhood. As can be confirmed, it follows from Theorem 15, that after at most $O(\log(n))$ such reductions in the step size, the resulting iterate will necessarily lie in the required neighborhood. Also, for each corrector step after an adjustment in the step size, a linear system with a different right hand side is solved, and is thus relatively inexpensive. In this implementation we only allow $l = 1, 2, 3$, and our computational experience suggests that this number for the small problems we tested, on the average, is very close to 1 and never greater than 2. This is one aspect of this implementation that needs further investigation for large scale problems.

We choose the value of $\beta = 1$. Since $\|S_1^{\frac{1}{2}} X_1 S_1^{\frac{1}{2}} - t_1 I\|_F = 0$, $(X_1, S_1)$ belongs to the required neighborhood, $N(t_1, \beta)$.

In Step 6, when $l \leq 2$ and $(X_{k+1,l}, y^{k+1,l}, S_{k+1,l}) \notin N(t_{k+1,l}, \beta)$ then the predictor step size is reduced to $\alpha_{l+1} = \frac{1}{2}\alpha_l$. To maintain a larger step size, we implemented a "two corrector step" strategy, in which we apply another corrector step, hoping that this will bring the resulting iterate into the required neighborhood, by solving:

$$
\begin{aligned}
\sum_{j=1}^{m} A_i \bullet (S_k^{-1} A_j X_k) \Delta \tilde{y}_j^k &= -A_i \bullet S_k^{-1}(t_{k+1,l} I - S_{k+1,l} X_{k+1,l}) \\
\Delta \tilde{S}_k &= -\sum_{i=1}^{m} \Delta A_i \tilde{y}_i^k \\
\Delta \tilde{X}_k^T &= S_{k+1,l}^{-1}(t_{k+1,l} I - S_{k+1,l} X_{k+1,l} - \Delta \tilde{S}_k X_k)
\end{aligned}
$$

and then checking to see if $(X_{k+1,l} + \frac{1}{2}(\Delta \tilde{X}_k + \Delta \tilde{X}_k^T), y^{k+1,l} + \Delta \tilde{y}^k, S_{k+1,l} + \Delta \tilde{S}_k) \in N(t_{k+1,l}, \beta)$. In our tests we discovered that almost always $(X_{k+1,1}, y^{k+1,1}, S_{k+1,1}) \in N(t_{k+1,l}, \beta)$, and we seldom resorted to the use of a second corrector step.

We have implemented the algorithm using MATLAB version 4.2, and all our test runs are on SUN SPARC 20 platform.

## 6.4 Test Problems

1. SLP(Randomly generated semidefinite programming problems) :

$$\min C \bullet X$$
$$\text{s.t. } A_i \bullet X = b_i, i = 1, \dots, m$$
$$X \succeq 0$$

   For this problem, we randomly generate $X \succeq 0, S \succeq 0$ and then use them to generate $A_i, b$, and $C$. The generated problems have feasible pairs of duals.

2. MNORM(Matrix norm minimization) :

$$\min \|A_0 + x_1 A_1 + \cdots + x_k A_k\|_2$$

   where $A_i$ is a $p \times q$ matrix.
   This problem appears in section 3.1 of [23]. For this problem, we randomly generate matrices $A_0, \dots, A_k$.

3. MCUT(Max_cut problem) :

$$\max -L \bullet X$$
$$diag(X) = \tfrac{e}{4}$$
$$X \succeq 0$$

   where

   (a) $L, X$ are $n$ by $n$ matrices.

   (b) $L = Diag(Ae) - A$, where $A$ is the weighted adjacency matrix.

The max_cut problem is to maximize the total weight of edges cut by the partition of the nodes of an edge-weighted undirected graph. This formulation is a relaxation of the max_cut problem and the details of this can be found in section 3.1 of [7]. Here, we randomly generate $A$ and restrict the problem to be a connected graph.

4. ETP(Education testing problem) :

$$\max \sum_{i=1}^{n} d_i$$

$$A - diag(d) \succeq 0$$

$$d \geq 0$$

   where $A$ is a symmetric positive definite. This is a statistics problem which appears in section 2 of [23]. We randomly generate a symmetric positive definite matrix $A$.

5. MCN(Minimizing condition number of symmetric positive-definite matrix) :

$$\min r$$

$$\text{s.t. } \mu > 0, \mu I \preceq M(x) \preceq r\mu I$$

   where $M(x) = M_0 + \sum_{i=1}^{m} x_i M_i$, and, for each $i$, $M_i$ are $n \times n$ symmetric matrices. This problem appears on page 38 of [3]. It can be seen that if $M_0 \succ 0$, the resulting semidefinite formulation has dual pairs with feasible interior solutions. Here we randomly generate $M_0 \succ 0, M_1, \ldots, M_m$.

6. LTI(Linear time-invariant systems) :

$$\min r$$

$$-A^T P - PA \succeq 0$$

$$I \preceq P \preceq rI$$

   where $A$ and $P$ are $n \times n$ matrices. The problem we test here is a special linear time invariant system in control theory. This problem appears on page 65 of [3]. Note that the semidefinite formulation (1) and (2) for this problem has $m = O(n^2) > n$. This makes it different from the five other test problems we have considered here.

In summary, MNORM, MCUT and ETP are feasible problems. We guarantee feasibility of SLP and MCN by generating data appropriately. However, for LTI, guaranteeing feasibility is difficult, thus 11 of 25 generated problems turned out to be infeasible.

## 6.5 Numerical Results

1. In Table 1 the stopping criteria for SDPA and INPC (INfeasible Predictor Corrector algorithm), a code implementing the algorithm presented in this section, is :

   $$\frac{C \bullet X_k - b^T y^k}{1 + |b^T y^k|} < 10^{-6}.$$

   For SP, the stopping criteria is :

   $$C \bullet X_k - b^T y^k \leq 10^{-6} C \bullet X_k, \text{ if } C \bullet X_k > 0, b^T y^k > 0$$

   or

   $$C \bullet X_k - b^T y^k \leq -10^{-6} b^T y^k, \text{ if } C \bullet X_k < 0, b^T y^k < 0$$

   The row of the table labeled "original size" describes the problem size by the parameters of original problem described in previous subsection and the row labeled "semidefinite size" reports parameters $m, n$ for the resulting semidefinite formulation (1) and (2).

2. From Table 1 we note that the solver SP takes the fewest iterations for the problems SLP, MNORM, MCUT, ETP and the number of iterations taken by INPC are close to this number. For MCN and LTI the solver INPC takes the fewest iterations. We believe that this is so since a phase 1 is not needed in infeasible start methods like INPC.

3. In Table 2, we use relative error $10^{-8}$ to attain higher accuracy for these same problems. Similar to Table 1, INPC takes fewest iteration on the three problems for which SP needs phase 1 iterations. We note that INPC takes fewer extra iterations then SP and SDPC to increase the accuracy. This suggests that INPC exhibits good local convergence in the final iterations of the method. As an example of this, consider the

following output showing iterations 10 through 15 of the algorithm :

| $k$ | $\frac{C\bullet X_k - b^T y^k}{1+b^T y^k}$ | $l$ | $\alpha$ | $\sqrt{\sum_{i=1}^m (A_i \bullet X_k - b_i)^2}$ | $\lVert \sum_{i=1}^m y_i^k A_i + S_k - C \rVert_F$ |
|---|---|---|---|---|---|
| 12 | 0.00889525 | 1 | 0.79691863 | 0.00543080 | 0.00085869 |
| 13 | 0.00076522 | 1 | 0.91367733 | 0.00046880 | 0.00007412 |
| 14 | 0.00001792 | 1 | 0.97655447 | 0.00001099 | 0.00000174 |
| 15 | 0.00000022 | 1 | 0.98780850 | 0.00000013 | 0.00000002 |
| 16 | 0.00000001 | 1 | 0.97607478 | 0.00000001 | 0.00000000 |

We note that, for this output, we move along the predictor direction up to 0.99 of the step to the boundary. To obtain a higher asymptotic convergence rate, we need to gradually increase this fraction to 1.

4. In both tables, the average $l$ is very close to 1, and we observed that it was never greater than 2 for any of the 75 problems we solved.

## 6.6 Some Observations

1. When $m \simeq n$, $O(m^2n^2) + O(mn^3) = O(n^4)$ multiplications are needed to calculate $\bar{A}F_k^{-1}G_k\bar{A}^T$. In this case, the computational bottleneck is not the inversion of the matrix $\bar{A}F_k^{-1}G_k\bar{A}^T$ which requires $O(m^3)$ multiplications, but is the number of multiplications required to generate it.

   Sometimes we encountered numerical difficulty while solving the predictor and corrector directions in Step 1 and 4 of the algorithm. During Step 4 we solve the following system of linear equations :

$$\begin{aligned}
\sum_{j=1}^{m} A_i \bullet (S_k^{-1}A_jX_k)\Delta\bar{y}_j^k &= -A_i \bullet S_k^{-1}(\bar{t}_kI - \bar{S}_k\bar{X}_k) \\
\Delta\bar{S}_k &= -\sum_{i=1}^{m} \Delta\bar{y}_i^k A_i \\
\Delta X_k^T &= S_k^{-1}(\bar{t}_kI - \bar{S}_k\bar{X}_k - \Delta S_k X_k).
\end{aligned}$$

   The computed direction $\Delta\tilde{X}, \Delta\tilde{y}, \Delta\tilde{S}$ acceptably satisfied the second equality above but the computation of $A_i \bullet \Delta\bar{X}_k, i = 1, \ldots, m$ was sometimes unacceptably large. We

TABLE 1.

relative error $=10^{-6}$.

| | SLP | MNORM | MCUT | ETP | MCN | LTI |
|---|---|---|---|---|---|---|
| Original Size | $m = 30$ $n = 40$ | $k = 30$ $p = 20$ $q = 20$ | $n = 50$ | $n = 25$ | $m = 20$ $n = 40$ | $n = 10$ |
| Semidefinite Size | $m = 30$ $n = 40$ | $m = 31$ $n = 40$ | $m = 50$ $n = 50$ | $m = 25$ $n = 50$ | $m = 22$ $n = 80$ | $m = 56$ $n = 30$ |
| Problem tested | 10 | 10 | 10 | 10 | 10 | 25[1] |
| SP(phase1) | 1 | | | | 1.2 | 5.36 |
| SP(phase2) | 16.1 | | | | 16.1 | 15.07 |
| SP(total) | 17.1 | 12.1 | 11.9 | 20.9 | 17.3 | 20.44 |
| SDPA | 19.7 | 23 | 21.5 | 27.7 | 25.4 | 24.93 |
| INPC | 17.9 | 15.1 | 14.7 | 24.2 | 15.9 | 17.36 |
| Average $l$ | 1.02 | 1.07 | 1.01 | 1.08 | 1.03 | 1.02 |

1. 11 of the 25 problems are discovered infeasible by all solvers. This reports only the iterations of the 14 feasible problems.

TABLE 2.

relative error = $10^{-8}$.

| | SLP | MNORM | MCUT | ETP | MCN | LTI |
|---|---|---|---|---|---|---|
| Original Size | $m = 30$<br>$n = 40$ | $k = 30$<br>$p = 20$<br>$q = 20$ | $n = 50$ | $n = 25$ | $m = 20$<br>$n = 40$ | $n = 10$ |
| Semidefinite Size | $m = 30$<br>$n = 40$ | $m = 31$<br>$n = 40$ | $m = 50$<br>$n = 50$ | $m = 25$<br>$n = 50$ | $m = 22$<br>$n = 80$ | $m = 56$<br>$n = 30$ |
| Problem tested | 10 | 10 | 10 | 10 | 10 | 25[1] |
| SP(phase1) | 1 | | | | 1.2 | 5.36 |
| SP(phase2) | 19.5 | | | | 19.8 | 18.07 |
| SP(total) | 20.5 | 15.4 | 15.2 | 25.3 | 21 | 23.43 |
| SDPA | 22.1 | 25.7 | 24.6 | 200[2] | 28.2 | 28 |
| INPC | 19.1 | 16.4 | 16.2 | 26.1 | 17.1 | 18.78 |
| SP diff | 3.4 | 3.3 | 3.3 | 4.4 | 3.7 | 2.99 |
| SDPA diff | 2.4 | 2.7 | 3.1 | | 2.8 | 3.07 |
| INPC diff | 1.2 | 1.3 | 1.5 | 1.9 | 1.2 | 1.42 |
| Average $l$ | 1.02 | 1.06 | 1.01 | 1.15 | 1.02 | 1.04 |

1. Same as Table 1. 11 of 25 problems are infeasible and this reports iteration for the 14 feasible problems.

2. SDPA exceeds its maximal iteration number of 200.

noticed that sometimes calculating $A_i \bullet (S_k^{-1} A_j X_k)$, $i, j = 1, \ldots, m$ by first computing $B_j = (S_k^{-1} A_j) X_k$ and then $A_i \bullet B_j$ gave a larger error then first computing $S_k^{-1} A_i$, $A_j X_k$ and then $(S_k^{-1} A_i) \bullet (A_j X_k)$.

2. Different values of $\rho$ gave different results for INPC. In Table 2 the average INPC iterations for the MNORM reduced from 16.4 to 11 when using

$$\rho = \max(|A_1|, \cdots, |A_m|, |b|, |C|).$$

For this test we choose a large number $\rho = 1000 \max(|A_1|, \cdots, |A_m|, |b|, |C|)$ to correctly detect infeasibility. For feasible problems like MNORM, MCUT, ETP, a smaller $\rho$ can be used and we expect INPC to solve the problems with fewer iterations.

3. For some feasible problems, like ETP, where an initial feasible solution is readily available, the path following strategy may fail to converge unless the given solution is "close" to the central trajectory. The ability to exploit this may be important for infeasible start algorithms, like INPC, which start with a large $\rho$ and the initial solution $X_1 = \rho I, y^1 = 0, S_1 = \rho I$, which is on the central trajectory, thus resulting in larger iterations than might be required if they were able to use effectively the given initial solution.

4. In Step 7, we use $3X_k \bullet S_k < \rho\theta_k(\|X_k\|_F + \|S_k\|_F)$ to detect infeasibility. For LTI, 11 of the 25 problems instances were detected as infeasible by all three codes. SDPA took an average of 8.64 iterations while INPC took an average of 9.09 iterations to detect this infeasibility for the 11 problems.

# 7 Conclusions and Future Work

In this paper we have presented an infeasible start predictor corrector path following method for semidefinite programming, which is a generalization of the method of Mizuno [14]. We do not assume that the dual pair of semidefinite linear programs have feasible solutions, and show that in time that is a low order polynomial of $n$, the data, and $|\log(\frac{\epsilon}{\rho})|$, the method either detects that there is no optimal solution with duality gap zero, in the set

$\{(X,S) : 0 \preceq X \preceq \rho I, 0 \preceq S \preceq \rho I\}$, or an approximate solution $X(\epsilon)$, $y(\epsilon)$, $S(\epsilon)$ is found such that $X(\epsilon) \bullet S(\epsilon) \leq n\epsilon$, $((\sum_{i=1}^m A_i \bullet X(\epsilon) - b_i)^2)^{-\frac{1}{2}} \leq \delta(A,b,C)\epsilon$ and $\| \sum_{i=1}^m A_i y(\epsilon) - S(\epsilon) + C\|_F \leq \delta(A,b,C)\epsilon$ where $\delta(A,b,C)$ is at most an exponential function of the data $A_i$, $i = 1, \cdots, m$, $b$ and $C$.

We also implement an algorithm based on this method in an MATLAB environment and show the potential of this approach by comparing the number of iterations taken by this method to solve six different type of problems with the number of iterations taken by the codes SP [24] and SDPA [6]. In the future, we will combine the MATLAB code with C subroutines and generate a front end like [4]. Also, we will consider special methods needed to exploit the sparse and block diagonal matrix structure of the many matrices of this problem, and investigate how these methods can effectively exploit these properties when solving the linear systems encountered during the predictor and corrector steps. We will also investigate the potential use of iterative methods presented in Saigal [22].

# References


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

[2] R. Bellman, *Introduction to matrix analysis*, Society for Industrial and Applied Mathematics, Philadelphia, 1995.

[3] S. Boyd, L. E. Ghaoui, E. Feron and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[4] S. Boyd and S. Wu, "sdpsol : a parser/solver for semidefinite programs with matrix structure : user's guide", Technical Report, Information System Laboratory, Stanford University, November 1995.

[5] 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.

[6] K. Fujisawa and M. Kojima, "SDPA(Semidefinite Programming Algorithm) : User's manual", Research Reports on Information Sciences, Ser. B : Operations Research B-308, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, December 1995.

[7] 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.*

[8] 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.

[9] C. R. Johnson, B. Kroshel and H. Wolkowicz, "An interior-point method for approximate positive semidefinite completions", CORR Report 95-21, Department of Combinatorics and Optimization, University of Waterloo, Ontario N2L 3G1, Canada, July 1995.

[10] 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.*

[11] C. J. Lin and R. Saigal, "A Predictor Corrector Method for Semi-definite Linear Programming", Technical Report, Dept. of Industrial and Operations Engineering, University of Michigan, MI 48109-2117, October 1995(revised December 1995).

[12] I. J. Lustig, R. E. Marsten, and D. F. Shanno, " Interior point methods for linear programming : computational state of the art", *ORSA Journal on Computing*, 6(1995), 1–14.

[13] S. Mehrotra, "On the implementation of a primal-dual interior point method", *SIAM Journal on Optimization*, 2(1992), 575–601.

[14] S. Mizuno, "Polynomiality of infeasible-interior-point algorithm for linear programming", *Mathematical Programming*, 67(1994), 109–120.

[15] 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.

[16] 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.

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

[18] Y. E. Nesterov, 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.

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

[20] 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.

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

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

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

[24] L. L. Vandenberghe, and S. Boyd, "SP : software for semidefinite programming, user's guide", Technical Report, Information System Laboratory, Stanford University, November 1994.

[25] R. J. Vanderbei, "LOQO: an interior point code for quadratic programming", Technical Report, Program in Statistics and Operations Research, Princeton University, Princeton, NJ 08544, February 1995.

[26] 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.

[27] Y. Zhang, "User's guide to LIPSOL: linear programming interior point solvers v0.3", Technical Report, Dept. of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MA 21228-5398, July 1995.