Polynomiality of An Inexact Infeasible Interior Point Algorithm for Semidefinite Programming

Guanglu Zhou* and Kim-Chuan Toh†

December 26, 2001

Abstract. In this paper we present a primal-dual inexact infeasible interior-point algorithm for semidefinite programming problems (SDP). This algorithm allows the use of search directions that are calculated from the defining linear system with only moderate accuracy, and does not require feasibility to be maintained even if the initial iterate happened to be a feasible solution of the problem. Under a mild assumption on the inexactness, we show that the algorithm can find an ϵ-approximate solution of an SDP in $O(n^2 \ln(1/\epsilon))$ iterations. This bound of our algorithm is the same as that of the exact infeasible interior point algorithms proposed by Y. Zhang.

Key Words. semidefinite programming, primal-dual, infeasible interior point method, inexact search direction, polynomial complexity.

AMS(MOS) subject classifications. 90C05

1 Introduction

Let $\mathcal{S}^n$ be the vector space of $n \times n$ real symmetric matrices endowed with the inner product $A \bullet B = \mathrm{Tr}(AB)$, and $\mathcal{S}^n_+$ the set of symmetric positive definite matrices. For any $m \times n$ matrix A, $\mathbf{vec}A$ denotes the mn-vector obtained from stacking the columns of A one by one from the first to the last. We denote the set of positive numbers by $I\!R_+$.

Consider the semidefinite program (SDP)

$$\begin{array}{lrcl} \min_X & C \bullet X & & \\ & A_k \bullet X & = & b_k, \quad k = 1, \ldots, m \\ & X & \succeq & 0, \end{array} \tag{1.1}$$

where $A_k, C, X \in \mathcal{S}^n$, and $X \succeq 0$ means that X is positive semidefinite. The dual of (1.1) is

$$\begin{array}{l} \max_{y,S} \; b^T y \\ \sum_{k=1}^m y_k A_k + Z = C \\ \qquad\qquad Z \succeq 0. \end{array} \tag{1.2}$$

*High Performance Computing for Engineered Systems, Singapore-MIT Alliance, 4 Engineering Drive 3, Singapore 117576. E-mail: `smazgl@nus.edu.sg`.

†Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543. E-mail: `mattohkc@nus.edu.sg`. Research supported in part by the Singapore-MIT alliance, and NUS Academic Research Grant R-146-000-032-112.

The Karush-Kuhn-Tucker optimality conditions for the primal and dual problems (1.1) and (1.2) are as follows:

$$\begin{pmatrix} \mathbf{A}^T y + \mathbf{vec} Z - \mathbf{vec} C \\ \mathbf{A}(\mathbf{vec} X) - b \\ XZ \end{pmatrix} = 0, \quad X, Z \succeq 0, \tag{1.3}$$

where $\mathbf{A}^T = [\mathbf{vec} A_1 \ \mathbf{vec} A_2 \ \cdots \ \mathbf{vec} A_m]$, $b^T = [b_1 \ b_2 \ \cdots \ b_m]$. We say that a point (X, y, Z) is feasible (strictly feasible) if it satisfies the first two linear equations in (1.3), and $X, Z \succeq 0$ $(X, Z \succ 0)$. It is known that if a strictly feasible point exists, then (1.3) has a solution.

In this paper we assume that $\mathbf{A}$ has full row rank, and we define

$$\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}. \tag{1.4}$$

Note that $\mathbf{A}\mathbf{A}^+ = I$.

SDP arises in a wide variety of areas. Numerous applications of SDP are presented in [12]. The recent handbook by Wolkowicz et al. [16] has chapters on applications in combinatorial optimization, nonconvex quadratic programming, eigenvalue and nonconvex optimization, systems and control theory, structural design, matrix completion problems, and statistics.

SDP has been an active research area since early 1990s. A comprehensive list of references in this field can be found in [16] and [15]. The handbook [16] listed more than 800 references, while the online bibliography on semidefinite programming [15] listed more than 700 references.

Primal-dual interior point methods, especially infeasible ones, have proven to be one of the most efficient class of methods for SDP, and many polynomial complexity results exist for these methods; see [2, 3], [8]–[10], [17, 18] for details. In each iteration of a primal-dual interior point algorithm, most of the computational work is devoted to the computation of a search direction by solving a linear system of equations. When the linear system is large, the computation of the solution by a direct method typically requires a lot of computer time and memory. In such a situation, one is forced to compute only an approximate solution by an iterative method such as the conjugate gradient method; the reader is referred to [14] for a detailed discuss on the computation of such an approximate solution. This motivates the study of infeasible interior-point algorithms that use inexact search directions. To a lesser extent, another motivation comes from the fact that even if one uses a direct method to solve the linear system, the solution may not satisfy the linear equations exactly due to rounding errors. But in spite of the inexactness of the solution in practical computation, most analyses of interior-point algorithms have been carried out under the assumption of having the exact solution of the linear system. In [4], the authors presented the first inexact search direction interior-point algorithm for SDP. It requires that the equations corresponding to primal and dual infeasibilities be satisfied exactly, but the equations corresponding to complementarity are relaxed. In order to satisfy the equations corresponding to primal infeasibility, a system of the form $\mathbf{A}\mathbf{A}^T v = r$ needs to be solved exactly in each interior-point iteration. However, solving such a system exactly can be expensive when $\mathbf{A}\mathbf{A}^T$ is not easily invertible or ill-conditioned.

For linear programming (LP), there are numerous papers published on the subject of infeasible interior-point algorithms using inexact search directions. In particular, Freund, Jarre and Mizuno [1] presented a global convergence analysis for a class of inexact infeasible interior-point methods that are deemed practically implementable but no polynomial complexity result was established. Mizuno and Jarre [7], and Korzak [5] proved global and polynomial-time convergence of inexact infeasible interior-point algorithms. The algorithm presented in [5] has the drawback of remaining primal-feasible once the iterate becomes primal-feasible. Thus if the iterate happens to become primal-feasible before the complementarity gap is significantly reduced, the linear system solver is forced to compute an exact solution which may be extremely costly when the system is large. On the other hand, the algorithms in [6] were more of a theoretical nature and may not lead to a practically implementable algorithm. In addition, the complexity bounds for the algorithms in [5, 7] are greater than those of the exact infeasible interior point algorithms proposed by Zhang [17] and Mizuno [6].

In this paper, we present an inexact search direction primal-dual infeasible interior point algorithm for SDP. This inexact algorithm has two important features. First, in each iteration of our algorithm, the linear system is only solved to a moderate accuracy in the residual. Thus we may employ an iterative solver such as the conjugate gradient method to compute an approximate solution of the linear system. Moreover, our algorithm does not require feasibility to be maintained even if the iterate happened to become feasible. Second, we show that the algorithm can find an ϵ-approximate solution of an SDP in $O(n^2 \ln(1/\epsilon))$ iterations. The bound we established for our algorithm is the same as that of the exact infeasible interior point algorithms proposed by Zhang [18].

The rest of the paper is organized as follows. In Section 2, we define an infeasible central path and its neighborhood. In Section 3, we discuss issues in defining an inexact search direction. In Section 4, we give a detailed discussion on how to choose a step length along an inexact search direction. In Section 5, we present an inexact infeasible interior-point algorithm for the SDP using inexact search directions, and state the main theorem which establishes a polynomial complexity result for our algorithm. Section 6 is devoted to a proof of a key lemma used in the proof of the main theorem. In Section 7, we make some concluding remarks.

Throughout this paper, we use $\|\cdot\|$ to denote the 2-norm of vectors, or Frobenius norm of matrices. Recall that we write $U \succeq 0$ to mean that U is positive semidefinite. Similarly, $U \succ 0$ means that U is positive definite (p.d). The notation $U \preceq V$ means that $V - U \succeq 0$. If $U \succeq 0$, we write $U^{\frac{1}{2}}$ for the positive semidefinite square root of U. We use the notation I for the $n \times n$ identity matrix. The maximum and minimum eigenvalues of a matrix $U \in \mathcal{S}^n$ are denoted by $\lambda_{\max}(U)$ and $\lambda_{\min}(U)$, respectively.

2 An Infeasible Central Path and Its Neighborhood

Let (X_0, y_0, Z_0) be an initial point such that

$$X_0 = Z_0 = \rho I, \tag{2.1}$$

where $\rho > 0$ is a constant. We define

$$\mu_0 = X_0 \bullet Z_0/n = \rho^2, \tag{2.2}$$

$$R_0^p = \mathbf{Avec} X_0 - b, \tag{2.3}$$

$$R_0^d \;=\; \mathbf{A}^T y_0 + \mathbf{vec} Z_0 - \mathbf{vec} C. \tag{2.4}$$

For $\theta > 0$, we consider the following system:

$$\begin{pmatrix} \mathbf{A}^T y + \mathbf{vec} Z - \mathbf{vec} C \\ \mathbf{A}(\mathbf{vec} X) - b \\ XZ \end{pmatrix} = \begin{pmatrix} \theta\, R_0^d \\ \theta\, R_0^p \\ \theta\, \mu_0 I \end{pmatrix}, \quad X, Z \succ 0, \tag{2.5}$$

If (1.3) has a strictly feasible point, then the system (2.5) has a unique solution for $\theta \in (0, 1]$. Define

$$\mathcal{P} \;=\; \left\{ \begin{array}{l} (\theta, X, y, Z) \in I\!R_+ \times \mathcal{S}_+^n \times I\!R^m \times \mathcal{S}_+^n \;: \\ \mathbf{A}^T y + \mathbf{vec} Z - \mathbf{vec} C = \theta R_0^d, \quad \mathbf{A}(\mathbf{vec} X) - b = \theta R_0^p, \quad XZ = \theta \mu_0 I \end{array} \right\}.$$

The set $\mathcal{P}$ is called the set of (infeasible) centers, and (X, y, Z) converges to a solution of (1.3) if $(\theta, X, y, Z) \in \mathcal{P}$ and $\theta \to 0$.

For any $M \in I\!R^{n\times n}$, we define the similarly transformed symmetrization operator

$$H_P(M) := \frac{1}{2}\left[PMP^{-1} + (PMP^{-1})^T\right], \tag{2.6}$$

for a given nonsingular matrix P. When $P = Z^{\frac{1}{2}}$, we omit the subscript P from H_P in (2.6).

Let $\gamma_1, \gamma_2 \in (0, 1)$. We define a neighborhood of the path $\mathcal{P}$ as follows:

$$\mathcal{N} \;=\; \left\{ \begin{array}{l} (\theta, \nu, X, y, Z) \in I\!R_+ \times I\!R_+ \times \mathcal{S}_+^n \times I\!R^m \times \mathcal{S}_+^n \;: \quad \theta \le \nu, \\ \mathbf{A}^T y + \mathbf{vec} Z - \mathbf{vec} C = \theta(R_0^d + \zeta), \quad \|\zeta\| \le \gamma_1 \rho, \\ \mathbf{A}(\mathbf{vec} X) - b = \theta(R_0^p + \xi), \quad \|\mathbf{A}^+ \xi\| \le \gamma_1 \rho, \\ \|H(XZ) - \nu\mu_0 I\| \le \gamma_2 \nu \mu_0, \end{array} \right\}. \tag{2.7}$$

It is readily shown that if $(\theta, \nu, X, y, Z) \in \mathcal{N}$, then

$$(1-\gamma_2)\nu\mu_0 I \;\preceq\; H(XZ) = Z^{\frac{1}{2}} X Z^{\frac{1}{2}} \;\preceq\; (1+\gamma_2)\nu\mu_0 I, \tag{2.8}$$

$$(1-\gamma_2)\nu\mu_0 \;\le\; \mu \;\le\; (1+\gamma_2)\nu\mu_0, \quad \text{where} \quad \mu = X \bullet Z/n, \tag{2.9}$$

$$X \in \mathcal{S}_+^n, \quad Z \in \mathcal{S}_+^n. \tag{2.10}$$

Remark 1. (a) In (2.7), we defined what is known as the narrow neighborhood of $\mathcal{P}$. However, the reader may check that the complexity result we established later for our algorithm also holds for a wider neighborhood of $\mathcal{P}$ by replacing the condition $\|H(XZ) - \nu\mu_0 I\| \le \gamma_2 \nu\mu_0$ in N by a condition of the form $\gamma\nu\mu_0 \le \lambda_{\min}(H(XZ)) \le \lambda_{\max}(H(XZ)) \le \Gamma\nu\mu_0$.

(b) For simplicity, we picked the so called HKM scaling matrix [2, 3, 8] in defining the neighborhood $\mathcal{N}$. However, the reader can easily adapt the complexity analysis in this paper to other scaling matrices (such as the Nesterov-Todd scaling matrix [13]) in the commutative class of the Monterio-Zhang family [10].

In this paper we always assume that

$$(1-\gamma_1)X_0 \succeq X_* \quad \text{and} \quad (1-\gamma_1)Z_0 \succeq Z_*, \tag{2.11}$$

where (X_*, y_*, Z_*) is a solution to (1.3). Let ρ be a constant such that

$$\rho \;\geq\; \frac{1}{n}\Big(\mathrm{Tr}(X_*) + \mathrm{Tr}(Z_*)\Big). \tag{2.12}$$

It is worth noting that the above assumption is a direct extension of the one in [18] that is imposed to obtain polynomial complexity bounds.

The following lemmas are pivotal to the complexity analysis of our inexact interior-point algorithm described later.

Lemma 2.1 *For any r_p and r_d satisfying $\|r_d\| \leq \gamma_1\rho$ and $\|\mathbf{A}^+ r_p\| \leq \gamma_1\rho$, there exists $(\tilde{X}, \tilde{y}, \tilde{Z})$ that satisfies the following conditions:*

$$\begin{aligned}
\mathbf{A}^T\tilde{y} + \mathbf{vec}\tilde{Z} - \mathbf{vec}C &= R_0^d + r_d, &(2.13)\\
\mathbf{A}(\mathbf{vec}\tilde{X}) - b &= R_0^p + r_p, &(2.14)\\
(1-\gamma_1)\rho I \;\preceq\; \tilde{X} &\preceq (1+\gamma_1)\rho I, &(2.15)\\
(1-\gamma_1)\rho I \;\preceq\; \tilde{Z} &\preceq (1+\gamma_1)\rho I. &(2.16)
\end{aligned}$$

Proof. Let

$$\begin{aligned}
\mathbf{vec}\tilde{X} &= \mathbf{vec}X_0 + \mathbf{A}^+ r_p,\\
\tilde{y} &= y_0 + (\mathbf{A}^+)^T r_d,\\
\mathbf{vec}\tilde{Z} &= \mathbf{vec}Z_0 + QQ^T r_d,
\end{aligned}$$

where Q is a matrix whose columns form an orthonormal basis of the null space of $\mathbf{A}$. It is readily shown that for $(\tilde{X}, \tilde{y}, \tilde{Z})$, all the conditions hold true. $\square$

Lemma 2.2 *If the conditions (2.1), (2.11) and (2.12) hold, then for any $(\theta, \nu, X, y, Z) \in \mathcal{N}$ with $\nu \in (0, 1]$, we have*

$$\theta\mathrm{Tr}(X) = O(\nu\rho n) \quad \text{and} \quad \theta\mathrm{Tr}(Z) = O(\nu\rho n).$$

Proof. For $(\theta, \nu, X, y, Z) \in \mathcal{N}$, there exist r_p and r_d satisfying $\|r_d\| \leq \gamma_1\rho$ and $\|\mathbf{A}^+ r_p\| \leq \gamma_1\rho$ such that

$$\begin{aligned}
\mathbf{A}^T y + \mathbf{vec}Z - \mathbf{vec}C &= \theta\,(R_0^d + r_d), &(2.17)\\
\mathbf{A}(\mathbf{vec}X) - b &= \theta\,(R_0^p + r_p). &(2.18)
\end{aligned}$$

By Lemma 2.1, there exists $(\tilde{X}, \tilde{y}, \tilde{Z})$ that satisfies conditions (2.13)–(2.16). From (2.17), (2.18) and the fact that (X_*, y_*, Z_*) satisfies the equations $\mathbf{A}(\mathbf{vec}X_*) = b$ and $\mathbf{A}^T y_* + \mathbf{vec}Z_* = \mathbf{vec}C$, we have that

$$\mathbf{A}((1-\theta)\mathbf{vec}X_* + \theta\mathbf{vec}\tilde{X} - \mathbf{vec}X) = 0,$$

and

$$\mathbf{A}^T((1-\theta)y_* + \theta\tilde{y} - y) + ((1-\theta)\mathbf{vec}Z_* + \theta\mathbf{vec}\tilde{Z} - \mathbf{vec}Z) = 0.$$

Hence we have

$$\Big((1-\theta)\mathbf{vec}X_* + \theta\mathbf{vec}\tilde{X} - \mathbf{vec}X\Big)^T \Big((1-\theta)\mathbf{vec}Z_* + \theta\mathbf{vec}\tilde{Z} - \mathbf{vec}Z\Big) = 0$$

or equivalently

$$\begin{aligned} &\Big((1-\theta)X_* + \theta\tilde{X}\Big) \bullet Z \;+\; X \bullet \Big((1-\theta)Z_* + \theta\tilde{Z}\Big) \\ &= \Big((1-\theta)X_* + \theta\tilde{X}\Big) \bullet \Big((1-\theta)Z_* \;+\; \theta\tilde{Z}\Big) + X \bullet Z. \end{aligned} \tag{2.19}$$

Using (2.9), (2.12), (2.15), (2.16) and the fact that $X_* \bullet Z_* = 0$, we obtain

$$\begin{aligned} &\theta\,(1-\gamma_1)\rho\,(\mathrm{Tr}(X) + \mathrm{Tr}(Z)) \;=\; \theta(1-\gamma_1)\rho\,(I \bullet X + I \bullet Z) \\ \le\;& (\theta\tilde{X}) \bullet X + (\theta\tilde{Z}) \bullet Z \;\le\; \Big((1-\theta)X_* + \theta\tilde{X}\Big) \bullet X + \Big((1-\theta)Z_* + \theta\tilde{Z}\Big) \bullet Z \\ =\;& \Big((1-\theta)X_* + \theta\tilde{X}\Big) \bullet \Big((1-\theta)Z_* + \theta\tilde{Z}\Big) + X \bullet Z \\ \le\;& \theta(1-\theta)(1+\gamma_1)\rho\,(X_* \bullet I + I \bullet Z_*) \;+\; \theta^2(1+\gamma_1)^2\rho^2 n \;+\; (1+\gamma_2)\nu n\mu_0 \\ \le\;& 8\nu\rho^2 n. \end{aligned}$$

Therefore,

$$\theta\mathrm{Tr}(X) \le \frac{8\nu\rho n}{1-\gamma_1} \quad \text{and} \quad \theta\mathrm{Tr}(Z) \le \frac{8\nu\rho n}{1-\gamma_1},$$

and thus the lemma holds. □

Remark 2. (a) Suppose

$$\mathcal{L} \;=\; \{(X, y, Z) \;:\; (\theta, \nu, X, y, Z) \in \mathcal{N},\; \nu \in (0,1],\; \theta = \nu\}.$$

Then by Lemma 2.2, the set $\mathcal{L}$ is bounded.

(b) Suppose we generate a sequence $\{(\theta_k, \nu_k, X_k, y_k, Z_k)\}$ in the neighborhood $\mathcal{N}$ such that

$$\nu_k \ge \theta_k, \quad \forall\; k, \quad \text{and} \quad 1 = \nu^0 \ge \nu_k \ge \nu_{k+1} \ge 0.$$

If $\nu_k \to 0$ as $k \to 0$, then any limit point of the sequence $\{(X_k, y_k, Z_k)\}$ is a solution of (1.3). If $\theta_k = \nu_k$, then the sequence $\{(X_k, Z_k)\}$ is also bounded.

3 Inexact Search Direction

We choose parameters $\eta_1 \in (0,1]$, $\eta_2 \in (0,1)$ with $\eta_1 \ge \eta_2$. Let $\theta_0, \nu_0 = 1$. Choose a vector $y_0 \in I\!R^m$. It follows from (2.1) that $(\theta_0, \nu_0, X_0, y_0, Z_0) \in \mathcal{N}$.

Given a current point $(\theta_k, \nu_k, X_k, y_k, Z_k) \in \mathcal{N}$, we outline the idea on how to generate a new point $(\theta_{k+1}, \nu_{k+1}, X_{k+1}, y_{k+1}, Z_{k+1}) \in \mathcal{N}$. At the current point, we try to compute a search direction $(\Delta X_k, \Delta y_k, \Delta Z_k)$ that satisfies the following linear system:

$$\begin{pmatrix} 0 & \mathbf{A}^T & I \\ \mathbf{A} & 0 & 0 \\ E_k & 0 & F_k \end{pmatrix} \begin{pmatrix} \mathbf{vec}\Delta X_k \\ \Delta y_k \\ \mathbf{vec}\Delta Z_k \end{pmatrix} = \begin{pmatrix} -\eta_1 R_k^d \\ -\eta_1 R_k^p \\ \mathbf{vec} R_k^c \end{pmatrix}, \tag{3.1}$$

where

$$E_k \;=\; Z_k^{\frac{1}{2}} \otimes Z_k^{\frac{1}{2}}, \qquad F_k \;=\; \tfrac{1}{2}\Big(Z_k^{\frac{1}{2}} X_k \otimes Z_k^{-\frac{1}{2}} + Z_k^{-\frac{1}{2}} \otimes Z_k^{\frac{1}{2}} X_k \Big). \tag{3.2}$$

$$R_k^c = (1-\eta_2)\nu_k \mu_0 I - Z_k^{\frac{1}{2}} X_k Z_k^{\frac{1}{2}}. \tag{3.3}$$

It is worth noting that E_k is symmetric p.d., but F_k is generally nonsymmetric. Note that the last equation in (3.1) is equivalent to

$$H(X_k Z_k + \Delta X_k Z_k + X_k \Delta Z_k) = (1-\eta_2)\nu_k \mu_0 I. \tag{3.4}$$

We define

$$\theta_k(\alpha) \;=\; (1-\alpha\eta_1)\theta_k, \quad \nu_k(\alpha) \;=\; (1-\alpha\eta_2)\nu_k, \tag{3.5}$$

$$X_k(\alpha) \;=\; X_k + \alpha\Delta X_k, \quad y_k(\alpha) \;=\; y_k + \alpha \Delta y_k, \quad Z_k(\alpha) \;=\; Z_k + \alpha \Delta Z_k. \tag{3.6}$$

It is obvious that $\theta_k(\alpha) \le \nu_k(\alpha)$ for $\alpha \in (0,1]$. We choose $\alpha_k \in (0,1]$ so that the new iterate

$$(\theta_{k+1}, \nu_{k+1}, X_{k+1}, y_{k+1}, Z_{k+1}) := (\theta_k(\alpha_k), \nu_k(\alpha_k), X_k(\alpha_k), y_k(\alpha_k), Z_k(\alpha_k)) \in \mathcal{N}.$$

For practical reasons, the search direction $(\Delta X_k, \Delta y_k, \Delta Z_k)$ may not be computed by solving the linear system (3.1) exactly. This leads us to consider using an inexact search direction $(\Delta X_k, \Delta y_k, \Delta Z_k)$ instead.

Let $\{\sigma_i\}_{i=0}^{\infty}$ be a monotonically decreasing sequence of numbers in $(0,1]$ such that $\bar{\sigma} := \sum_{i=0}^{\infty} \sigma_i < \infty$. In this paper, we accept $(\Delta X_k, \Delta y_k, \Delta Z_k)$ as an inexact search direction at the k-th iteration if it satisfies the following linear system:

$$\begin{pmatrix} 0 & \mathbf{A}^T & I \\ \mathbf{A} & 0 & 0 \\ E_k & 0 & F_k \end{pmatrix} \begin{pmatrix} \mathbf{vec}\Delta X_k \\ \Delta y_k \\ \mathbf{vec}\Delta Z_k \end{pmatrix} = \begin{pmatrix} -\eta_1 (R_k^d + r_k^d) \\ -\eta_1 (R_k^p + r_k^p) \\ \mathbf{vec} R_k^c \end{pmatrix}, \tag{3.7}$$

where the "residual components" satisfy

$$\|\mathbf{A}^+ r_k^p\| \;\le\; \gamma_1 \rho\, \theta_k \sigma_k, \qquad \|r_k^d\| \;\le\; \gamma_1 \rho\, \theta_k \sigma_k. \tag{3.8}$$

Remark 3. (a) In practice, it is easy to check whether the conditions in (3.8) hold. Let $\sigma_{\min}$ be the smallest singular value of $\mathbf{A}$. Since $\|\mathbf{A}^+\| = 1/\sigma_{\min}$, if we have $\|r_k^p\| \;\le\; \gamma_1\rho\,\theta_k\sigma_k\sigma_{\min}$, then $\|\mathbf{A}^+ r_k^p\| \;\le\; \gamma_1 \rho\,\theta_k \sigma_k$.

(b) One can also check the first condition in (3.8) directly by carrying out a small number of conjugate gradient iterations to estimate the norm $\|\mathbf{A}^+ r_k^p\|$. Note that one needs only a very crude estimate of the norm, say 90% in relative accuracy, to check this condition in practice.

(c) For the algorithms proposed in [4], the quantity $\mathbf{A}^+ r_k^p$ needs to be computed exactly, and this can be expensive when $\mathbf{A}\mathbf{A}^T$ is not easily invertible or ill-conditioned. Our algorithm avoids such a need and thus offers some computational advantages over those proposed in [4].

In (3.7), we require the third equation to be satisfied exactly, because in practical computation, either ΔX_k or ΔZ_k is first computed, and then the other unknown is computed by using the third equation. For example, an inexact search direction $(\Delta X_k, \Delta y_k, \Delta Z_k)$ can be computed in practice by either one of the following procedures.

Procedure A.

- Compute a Δy_k that satisfies the following linear system:

$$\mathbf{A}(E_k^{-1} F_k)\mathbf{A}^T \, \Delta y_k = -\eta_1 R_k^p - \mathbf{A} E_k^{-1} \Big(\eta_1 F_k R_k^d + \mathbf{vec}(R_k^c) \Big) - \eta_1 r_k^p,$$

  with the residual vector r_k^p satisfying $\|\mathbf{A}^+ r_k^p\| \; \leq \; \gamma_1 \rho \, \theta_k \sigma_k$.

- Compute ΔZ_k by

$$\mathbf{vec}\Delta Z_k \;=\; -\mathbf{A}^T \Delta y_k - \eta_1 R_k^d.$$

- Compute ΔX_k by

$$\Delta X_k \;=\; (1-\eta_2)\nu_k \mu_0 Z_k^{-1} - X_k - \frac{1}{2}[X_k(\Delta Z_k) Z_k^{-1} + (X_k(\Delta Z_k) Z_k^{-1})^T].$$

Procedure B.

- Compute a Δy_k and ΔX_k that satisfies the following linear system:

$$\begin{bmatrix} -F_k^{-1} E_k & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} \Delta X_k \\ \Delta y_k \end{bmatrix} = \begin{bmatrix} -\eta_1 R_k^d - F_k^{-1} \mathbf{vec} R_k^c \\ -\eta_1 R_k^p \end{bmatrix} - \begin{bmatrix} \eta_1 r_k^d \\ \eta_1 r_k^p \end{bmatrix},$$

  with the residual vectors r_k^d and r_k^p satisfying $\|r_k^d\|, \|\mathbf{A}^+ r_k^p\| \leq \gamma_1 \rho \, \theta_k \sigma_k$.

- Compute ΔZ_k by

$$\mathbf{vec}\Delta Z_k \;=\; F_k^{-1} \mathbf{vec} R_k^c - F_k^{-1} E_k \mathbf{vec}\Delta X_k.$$

4 Step Length

In this section, we discuss how to choose a step length α_k along an inexact search direction $(\Delta X_k, \Delta y_k, \Delta Z_k)$ obtained from (3.7) to generate a new point $(\theta_{k+1}, \nu_{k+1}, X_{k+1}, y_{k+1}, Z_{k+1}) \in \mathcal{N}$. For any $\alpha \in (0,1]$, let

$$R_k^p(\alpha) = \mathbf{A}(\mathbf{vec} X_k(\alpha)) - b, \quad \text{and} \quad R_k^d(\alpha) = \mathbf{A}^T y_k(\alpha) + \mathbf{vec} Z_k(\alpha) - \mathbf{vec} C,$$

where $X_k(\alpha)$, $y_k(\alpha)$ and $Z_k(\alpha)$ are defined as in (3.6). From (3.7), we have

$$\begin{aligned} R_k^p(\alpha) &= R_k^p - \alpha\eta_1(R_k^p + r_k^p) \;=\; (1-\alpha\eta_1)R_k^p - \alpha\eta_1 r_k^p \\ &= (1-\alpha\eta_1)\theta_k(R_0^p + \xi_{k-1}) - \alpha\eta_1 r_k^p \\ &= (1-\alpha\eta_1)\theta_k\left(R_0^p + \xi_{k-1} - \frac{\alpha\eta_1}{(1-\alpha\eta_1)\theta_k} r_k^p\right) \\ &= \theta_k(\alpha)\left(R_0^p + \xi_k(\alpha)\right), \end{aligned} \tag{4.1}$$

where $\theta_k(\alpha)$ is defined as in (3.5), and

$$\xi_k(\alpha) \;=\; \xi_{k-1} - \frac{\alpha\eta_1}{(1-\alpha\eta_1)\theta_k} r_k^p.$$

It is readily proven that

$$\xi_k(\alpha) = -\sum_{i=0}^{k-1} \frac{\alpha_i\eta_1}{(1-\alpha_i\eta_1)\theta_i} r_i^p \;-\; \frac{\alpha\eta_1}{(1-\alpha\eta_1)\theta_k} r_k^p.$$

Assume that $\alpha_i, \alpha \le \dfrac{1}{\eta_1(1+\bar{\sigma})}$, then

$$\begin{aligned} \|\mathbf{A}^+\xi_k(\alpha)\| &\le \sum_{i=0}^{k-1} \frac{\alpha_i\eta_1}{(1-\alpha_i\eta_1)\theta_i}\|\mathbf{A}^+ r_i^p\| \;+\; \frac{\alpha\eta_1}{(1-\alpha\eta_1)\theta_k}\|\mathbf{A}^+ r_k^p\| \\ &\le \sum_{i=0}^{k-1} \gamma_1\rho\frac{\sigma_i}{\bar{\sigma}} + \gamma_1\rho\frac{\sigma_k}{\bar{\sigma}} \;\le\; \gamma_1\rho\frac{1}{\bar{\sigma}}\sum_{i=0}^{k}\sigma_i \\ &\le \gamma_1\rho \end{aligned}$$

Similarly, we have that

$$R_k^d(\alpha) = \theta_k(\alpha)\left(R_0^d + \zeta_k(\alpha)\right), \tag{4.2}$$

where

$$\zeta_k(\alpha) = -\sum_{i=0}^{k-1} \frac{\alpha_i\eta_1}{(1-\alpha_i\eta_1)\theta_i} r_i^d - \frac{\alpha\eta_1}{(1-\alpha\eta_1)\theta_k} r_k^d, \quad \text{and} \quad \|\zeta_k(\alpha)\| \le \gamma_1\rho.$$

Now from (3.4), we have

$$\begin{aligned} H(X_k(\alpha)Z_k(\alpha)) &= H\left((X_k + \alpha\Delta X_k)(Z_k + \alpha\Delta Z_k)\right) \\ &= H(X_kZ_k) + \alpha H(\Delta X_k Z_k + X_k\Delta Z_k) + \alpha^2 H(\Delta X_k\Delta Z_k) \\ &= (1-\alpha)H(X_kZ_k) + \alpha(1-\eta_2)\nu_k\mu_0 I + \alpha^2 H(\Delta X_k\Delta Z_k). \end{aligned} \tag{4.3}$$

Thus,

$$H(X_k(\alpha)Z_k(\alpha)) - \nu_k(\alpha)\mu_0 I \;=\; (1-\alpha)\left(H(X_kZ_k) - \nu_k\mu_0 I\right) + \alpha^2 H(\Delta X_k\Delta Z_k). \tag{4.4}$$

Let $P = (Z_k(\alpha))^{\frac{1}{2}}$. By Lemma 4.2 in [18] and (4.4), we have

$$\begin{aligned}
& \|H_P(X_k(\alpha)Z_k(\alpha)) - \nu_k(\alpha)\mu_0 I\| - \gamma_2\nu_k(\alpha)\mu_0 \\
\leq\ & \|H(X_k(\alpha)Z_k(\alpha)) - \nu_k(\alpha)\mu_0 I\| - \gamma_2\nu_k(\alpha)\mu_0 \\
\leq\ & (1-\alpha)\|H(X_kZ_k) - \nu_k\mu_0 I\| + \alpha^2\|H(\Delta X_k\Delta Z_k)\| - \gamma_2\nu_k(\alpha)\mu_0 \\
\leq\ & (1-\alpha)\gamma_2\nu_k\mu_0 + \alpha^2\|H(\Delta X_k\Delta Z_k)\| - \gamma_2(1-\alpha\eta_2)\nu_k\mu_0 \\
=\ & -\alpha(1-\eta_2)\gamma_2\nu_k\mu_0 + \alpha^2\|H(\Delta X_k\Delta Z_k)\|. \qquad (4.5)
\end{aligned}$$

Let

$$\bar{\alpha}_k = \min\left(1,\ \frac{1}{\eta_1(1+\bar{\sigma})},\ \frac{(1-\eta_2)\gamma_2\nu_k\mu_0}{\|H(\Delta X_k\Delta Z_k)\|}\right). \qquad (4.6)$$

By (4.1), (4.2), (4.5) and (4.6), we have that for any $\alpha \in (0, \bar{\alpha}_k]$,

$$(\theta_k(\alpha), \nu_k(\alpha), X_k(\alpha), y_k(\alpha), Z_k(\alpha)) \in \mathcal{N}.$$

Remark 4. Notice that if the step-length taken at the ith iteration is α_i, then

$$\theta_k = \prod_{i=0}^{k-1}(1-\alpha_i\eta_1), \qquad \nu_k = \prod_{i=0}^{k-1}(1-\alpha_i\eta_2). \qquad (4.7)$$

Since $\eta_2 \leq \eta_1$, we see that θ_k decreases at a rate that is faster than ν_k. Thus primal and dual infeasibilities are reduced at a rate faster than the complementarity gap, unless $\eta_1 = \eta_2$.

5 An Inexact Infeasible Interior Point Algorithm

In this section, we present an inexact infeasible interior point algorithm and show that the algorithm has polynomial convergence.

Algorithm. Let $\theta_0, \nu_0 = 1$. Choose parameters $\eta_1 \in (0,1]$ and $\gamma_1, \gamma_2, \eta_2 \in (0,1)$ with $\eta_2 \leq \eta_1$. Pick a sequence $\{\sigma_k\}_{k=0}^{\infty}$ in $(0,1]$ such that $\bar{\sigma} := \sum_{k=0}^{\infty}\sigma_k < \infty$. Choose (X_0, y_0, Z_0) satisfying (2.1), (2.11), (2.12), and $(\theta_0, \nu_0, X_0, y_0, Z_0) \in \mathcal{N}$.

For $k = 0, 1, \ldots$

(Let the current and the next iterate be $(\theta_k, \nu_k, X_k, y_k, Z_k)$ and $(\theta_{k+1}, \nu_{k+1}, X_{k+1}, y_{k+1}, Z_{k+1})$ respectively.)

- Find an inexact search direction $(\Delta X_k, \Delta y_k, \Delta Z_k)$ by solving the linear system (3.7).
- Let α_k be the maximum of all $\alpha \in \left[0, \min\left(1, \frac{1}{\eta_1(1+\bar{\sigma})}\right)\right]$ satisfying $(\theta_k(\alpha),\ \nu_k(\alpha),\ X_k(\alpha),\ y_k(\alpha),\ Z_k(\alpha)) \in \mathcal{N}$. Update $(\theta_k, \nu_k, X_k, y_k, Z_k)$ to the point defined below:

$$(\theta_{k+1}, \nu_{k+1}, X_{k+1}, y_{k+1}, Z_{k+1}) := (\theta_k(\alpha_k), \nu(\alpha_k), X_k(\alpha_k), y_k(\alpha_k), Z_k(\alpha_k)).$$

Theorem 5.1 *Let $\epsilon > 0$ be a small constant. Suppose the conditions in (2.1), (2.11) and (2.12) hold. Then $\nu_k \le \epsilon$ for $k = O(n^2 \ln(1/\epsilon))$.*

We need the following lemma to prove the theorem. The proof of the lemma will be given in the next section.

Lemma 5.1 *Suppose the conditions in (2.1), (2.11) and (2.12) hold. Then*

$$\|H(\Delta X_k \Delta Z_k)\| = O(n^2 \nu_k \mu_0). \tag{5.1}$$

Proof of Theorem 5.1. It follows from (4.6) and Lemma 5.1 that at k-th iteration, α_k exists and satisfies

$$\alpha_k \ge \bar{\alpha} = \min\left(1, \frac{1}{\eta_1(1+\bar{\sigma})}, \frac{c}{n^2}\right),$$

where $c > 0$ is a constant. Since

$$\nu_k \;=\; \prod_{i=0}^{k-1}(1-\alpha_i\eta_2) \;\le\; (1-\bar{\alpha}\eta_2)^k, \quad \text{and} \quad \frac{1}{\bar{\alpha}} = O(n^2),$$

we have $\nu_k \le \epsilon$ for $k = O(n^2 \ln(1/\epsilon))$. ∎

6 Proof of Lemma 5.1

Our purpose here is to establish an upper bound for $\|H(\Delta X_k \Delta Z_k)\|$. Let

$$S_k := F_k E_k = \frac{1}{2}\Big(Z_k^{\frac{1}{2}} X_k Z_k^{\frac{1}{2}} \otimes I + I \otimes Z_k^{\frac{1}{2}} X_k Z_k^{\frac{1}{2}}\Big), \quad D_k := S_k^{-\frac{1}{2}} F_k = S_k^{\frac{1}{2}} E_k^{-1}. \tag{6.1}$$

where E_k and F_k are defined as in (3.2). Consider the following eigenvalue decomposition:

$$Z_k^{\frac{1}{2}} X_k Z_k^{\frac{1}{2}} \;=\; Q_k \Lambda_k Q_k^T, \tag{6.2}$$

where $Q_k^T Q_k = I$, and $\Lambda_k = \mathrm{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)$ with $\lambda_1^k \ge \lambda_2^k \ge \cdots \ge \lambda_n^k > 0$. From (2.8), we have

$$(1-\gamma_2)\nu_k\mu_0 \;\le\; \lambda_n^k \;\le\; \cdots \;\le\; \lambda_1^k \;\le\; (1+\gamma_2)\nu_k\mu_0. \tag{6.3}$$

From (6.2) and (6.3), the eigenvalues of S_k are $\{\frac{1}{2}(\lambda_i^k + \lambda_j^k) : i, j = 1, 2, \cdots, n\}$, and

$$2(1-\gamma_2)\nu_k\mu_0 \;\le\; \lambda_i^k + \lambda_j^k \;\le\; 2(1+\gamma_2)\nu_k\mu_0,\; i, j = 1, 2, \cdots, n. \tag{6.4}$$

Let

$$G_k = \frac{1}{2}\Big(X_k^{-\frac{1}{2}} Z_k^{-1} X_k^{-\frac{1}{2}} \otimes I + I \otimes X_k^{-\frac{1}{2}} Z_k^{-1} X_k^{-\frac{1}{2}}\Big). \tag{6.5}$$

Due to similarity, $Z_k^{\frac{1}{2}} X_k Z_k^{\frac{1}{2}}$ and $X_k^{\frac{1}{2}} Z_k X_k^{\frac{1}{2}}$ have the same spectrum. This implies that the eigenvalues of G_k are $\{\frac{1}{2}(\frac{1}{\lambda_i^k} + \frac{1}{\lambda_j^k}) : i, j = 1, 2, \cdots, n\}$, and

$$\frac{2}{(1+\gamma_2)\nu_k\mu_0} \;\le\; \frac{1}{\lambda_i^k} + \frac{1}{\lambda_j^k} \;\le\; \frac{2}{(1-\gamma_2)\nu_k\mu_0},\; i, j = 1, 2, \ldots, n. \tag{6.6}$$

Note that $E_k^{-1} F_k = (X_k^{\frac{1}{2}} \otimes X_k^{\frac{1}{2}}) G_k (X_k^{\frac{1}{2}} \otimes X_k^{\frac{1}{2}})$.

Lemma 6.1 *For any* $M \in I\!R^{n\times n}$,

$$\begin{aligned}
\|D_k^{-T}\mathbf{vec}M\|^2 &\le \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|Z_k^{\frac{1}{2}}MZ_k^{\frac{1}{2}}\|^2,\\
\|D_k\mathbf{vec}M\|^2 &\le \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|X_k^{\frac{1}{2}}MX_k^{\frac{1}{2}}\|^2.
\end{aligned}$$

Proof. By the definition of the matrices E_k, F_k, G_k, S_k and D_k, we have

$$\begin{aligned}
\|D_k^{-T}\mathbf{vec}M\|^2 &= (\mathbf{vec}M)^T E_k S_k^{-1} E_k \mathbf{vec}M\\
&= \mathbf{vec}\left(Z_k^{\frac{1}{2}}MZ_k^{\frac{1}{2}}\right)^T S_k^{-1}\mathbf{vec}\left(Z_k^{\frac{1}{2}}MZ_k^{\frac{1}{2}}\right)\\
&\le \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|\mathbf{vec}(Z_k^{\frac{1}{2}}MZ_k^{\frac{1}{2}})\|^2, \quad \text{by (6.4)}\\
&= \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|Z_k^{\frac{1}{2}}MZ_k^{\frac{1}{2}}\|^2.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|D_k\mathbf{vec}M\|^2 &= (\mathbf{vec}M)^T E_k^{-1} S_k E_k^{-1}\mathbf{vec}M \;=\; (\mathbf{vec}M)^T E_k^{-1} F_k \mathbf{vec}M\\
&= \mathbf{vec}\left(X_k^{\frac{1}{2}}MX_k^{\frac{1}{2}}\right)^T G_k\mathbf{vec}\left(X_k^{\frac{1}{2}}MX_k^{\frac{1}{2}}\right)\\
&\le \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|\mathbf{vec}(X_k^{\frac{1}{2}}MX_k^{\frac{1}{2}})\|^2, \quad \text{by (6.6)}\\
&= \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|X_k^{\frac{1}{2}}MX_k^{\frac{1}{2}}\|^2.
\end{aligned}$$

□

Lemma 6.2

$$\|D_k^{-T}\mathbf{vec}\Delta X_k\|^2 + \|D_k\mathbf{vec}\Delta Z_k\|^2 + 2\Delta X_k \bullet \Delta Z_k = \|S_k^{-\frac{1}{2}}\mathbf{vec}R_k^c\|^2,$$

$$\|H(\Delta X_k \Delta Z_k)\| \;\le\; \frac{\kappa}{2}\left(\|D_k^{-T}\mathbf{vec}\Delta X_k\|^2 + \|D_k\mathbf{vec}\Delta Z_k\|^2\right),$$

where $\kappa = \lambda_1^k/\lambda_n^k \le \frac{1+\gamma_2}{1-\gamma_2}$.

Proof. See Lemma 3.1 and Lemma 3.3 in [18]. □

Lemma 6.3 *We have*

$$\|S_k^{-\frac{1}{2}}\mathbf{vec}R_k^c\|^2 = O(n\nu_k\mu_0).$$

Proof. Observe that from (6.2),

$$\begin{aligned} S_k^{-1} &= 2\,(Q_k \otimes Q_k)(\Lambda_k \otimes I + I \otimes \Lambda_k)^{-1}(Q_k^T \otimes Q_k^T), \\ \mathbf{vec} R_k^c &= (Q_k \otimes Q_k)\mathbf{vec}\left((1-\eta_2)\nu_k\mu_0 I - \Lambda_k\right). \end{aligned}$$

Thus

$$\begin{aligned} & \|S_k^{-\frac{1}{2}}\mathbf{vec} R_k^c\|^2 \;=\; (\mathbf{vec} R_k^c)^T S_k^{-1}\mathbf{vec} R_k^c \\ = \;& 2\mathbf{vec}\left((1-\eta_2)\nu_k\mu_0 I - \Lambda_k)\right)^T (\Lambda_k \otimes I + I \otimes \Lambda_k)^{-1}\mathbf{vec}\left((1-\eta_2)\nu_k\mu_0 I - \Lambda_k\right) \\ = \;& \sum_{i=1}^n \frac{((1-\eta_2)\nu_k\mu_0 - \lambda_i^k)^2}{\lambda_i^k} \\ \le \;& \frac{1}{(1-\gamma_2)\nu_k\mu_0}\sum_{i=1}^n((1-\eta_2)\nu_k\mu_0 - \lambda_i^k)^2 \\ \le \;& \frac{1}{(1-\gamma_2)\nu_k\mu_0}\sum_{i=1}^n(|\nu_k\mu_0 - \lambda_i^k| + \eta_2\nu_k\mu_0)^2 \\ \le \;& \frac{n\nu_k\mu_0}{1-\gamma_2}(\gamma_2+\eta_2)^2, \quad \text{by (6.3).} \end{aligned}$$

Thus the lemma is proven. □

For the rest of the analysis, we need to introduce an auxiliary point $(\tilde{X}_k, \tilde{y}_k, \tilde{Z}_k)$ whose existence is ensured by Lemma 2.1. From (4.1) and (4.2), we have the the following equations at the k-th iteration,

$$\mathbf{A}^T y_k + \mathbf{vec} Z_k - \mathbf{vec} C = \theta_k(R_0^d + \zeta_k), \quad \|\zeta_k\| \le \gamma_1\rho, \tag{6.7}$$

$$\mathbf{A}(\mathbf{vec} X_k) - b = \theta_k(R_0^p + \xi_k), \quad \|\mathbf{A}^+\xi_k\| \le \gamma_1\rho. \tag{6.8}$$

Thus by Lemma 2.1, there exists $(\tilde{X}_k, \tilde{y}_k, \tilde{Z}_k)$ such that

$$\mathbf{A}^T\tilde{y}_k + \mathbf{vec}\tilde{Z}_k - \mathbf{vec} C \;=\; R_0^d + \zeta_k, \tag{6.9}$$

$$\mathbf{A}(\mathbf{vec}\tilde{X}_k) - b \;=\; R_0^p + \xi_k, \tag{6.10}$$

$$(1-\gamma_1)\rho I \;\preceq\; \tilde{X}_k \;\preceq\; (1+\gamma_1)\rho I, \tag{6.11}$$

$$(1-\gamma_1)\rho I \;\preceq\; \tilde{Z}_k \;\preceq\; (1+\gamma_1)\rho I. \tag{6.12}$$

Lemma 6.4 *The following equations hold:*

$$\left(X_k - X_* - \theta_k(\tilde{X}_k - X_*)\right) \bullet \left(Z_k - Z_* - \theta_k(\tilde{Z}_k - Z_*)\right) \;=\; 0, \tag{6.13}$$

$$\left(\mathbf{vec}\Delta X_k + \eta_1\theta_k\mathbf{vec}(\tilde{X}_k - X_*) + \eta_1\mathbf{A}^+ r_k^p\right)^T \left(\mathbf{vec}\Delta Z_k + \eta_1\theta_k\mathbf{vec}(\tilde{Z}_k - Z_*) + \eta_1 r_k^d\right) = 0. \tag{6.14}$$

Proof. From (6.7)–(6.10) and the fact that $\mathbf{A}\mathbf{vec}X_* = b$ and $\mathbf{A}^T y_* + \mathbf{vec}Z_* = \mathbf{vec}C$, we have

$$\mathbf{A}\left(\mathbf{vec}(X_k - X_*) - \theta_k \mathbf{vec}(\tilde{X}_k - X_*)\right) = 0,$$

$$\mathbf{A}^T\left(y_k - y_* - \theta_k(\tilde{y}_k - y_*)\right) + \left(\mathbf{vec}(Z_k - Z_*) - \theta_k \mathbf{vec}(\tilde{Z}_k - Z_*)\right) = 0.$$

The above two equations imply that

$$\left(\mathbf{vec}(X_k - X_*) - \theta_k \mathbf{vec}(\tilde{X}_k - X_*)\right)^T \left(\mathbf{vec}(Z_k - Z_*) - \theta_k \mathbf{vec}(\tilde{Z}_k - Z_*)\right) = 0,$$

which proves (6.13).

Next we proceed to prove (6.14). From (3.7), and (6.7)–(6.10), we have

$$\mathbf{A}\left(\mathbf{vec}\Delta X_k + \eta_1\theta_k \mathbf{vec}(\tilde{X}_k - X_*) + \eta_1 \mathbf{A}^+ r_k^p\right) = 0,$$

$$\mathbf{A}^T(\Delta y_k + \eta_1\theta_k(\tilde{y}_k - y_*)) + \left(\mathbf{vec}\Delta Z_k + \eta_1\theta_k \mathbf{vec}(\tilde{Z}_k - Z_*) + \eta_1 r_k^d\right) = 0.$$

With the above two equations, (6.14) follows. □

Let

$$t = \left(\|D_k^{-T}\mathbf{vec}\Delta X_k\|^2 + \|D_k \mathbf{vec}\Delta Z_k\|^2\right)^{\frac{1}{2}}, \tag{6.15}$$

$$\beta = \left(\|D_k^{-T}\mathbf{vec}(\tilde{X}_k - X_*)\|^2 + \|D_k \mathbf{vec}(\tilde{Z}_k - Z_*)\|^2\right)^{\frac{1}{2}}, \tag{6.16}$$

$$\delta = \left(\|D_k^{-T}\mathbf{A}^+ r_k^p\|^2 + \|D_k r_k^p\|^2\right)^{\frac{1}{2}}. \tag{6.17}$$

Then we have the following lemma.

Lemma 6.5

$$t \le 2\eta_1(\theta_k\beta + \delta) + \sqrt{\tau},$$

where $\tau = \|S_k^{-\frac{1}{2}}\mathbf{vec}R_k^c\|^2 + 2(\eta_1\theta_k)^2(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) + 2\eta_1^2(\theta_k\beta + \delta)\delta.$

Proof. Let $\mathcal{R}_k^p, \mathcal{R}_k^d \in I\!R^{n\times n}$ be the unique matrices such that

$$\mathbf{vec}\mathcal{R}_k^p = \mathbf{A}^+ r_k^p, \quad \text{and} \quad \mathbf{vec}\mathcal{R}_k^d = r_k^d.$$

Then (6.14) is equivalent to the following equation:

$$\left(\Delta X_k + \eta_1\theta_k(\tilde{X}_k - X_*) + \eta_1\mathcal{R}_k^p\right) \bullet \left(\Delta Z_k + \eta_1\theta_k(\tilde{Z}_k - Z_*) + \eta_1\mathcal{R}_k^d\right) = 0. \tag{6.18}$$

From (6.18), we have that

$$\Delta X_k \bullet \Delta Z_k = -\eta_1\theta_k\left(\Delta X_k \bullet (\tilde{Z}_k - Z_*) + (\tilde{X}_k - X_*) \bullet \Delta Z_k\right)$$

$$
\begin{aligned}
& -(\eta_1\theta_k)^2(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) - \eta_1(\Delta X_k \bullet \mathcal{R}_k^d + \mathcal{R}_k^p \bullet \Delta Z_k) \\
& -\eta_1^2\theta_k\left((\tilde{X}_k - X_*) \bullet \mathcal{R}_k^d + \mathcal{R}_k^p \bullet (\tilde{Z}_k - Z_*)\right) - \eta_1^2 \mathcal{R}_k^p \bullet \mathcal{R}_k^d \\
\geq \quad & -\eta_1\theta_k\beta t - \eta_1\delta t - (\eta_1\theta_k)^2(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) - \eta_1^2\theta_k\beta\delta - \frac{1}{2}\eta_1^2\delta^2 \\
\geq \quad & -\eta_1(\theta_k\beta + \delta)t - (\eta_1\theta_k)^2(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) - \eta_1^2(\theta_k\beta + \delta)\delta.
\end{aligned}
$$

By Lemma 6.2 and the above inequality, we have

$$t^2 - 2\eta_1(\theta_k\beta + \delta)t - \tau \leq 0.$$

Hence,

$$
\begin{aligned}
t & \leq \eta_1(\theta_k\beta + \delta) + \sqrt{\eta_1^2(\theta_k\beta + \delta)^2 + \tau} \\
& \leq 2\eta_1(\theta_k\beta + \delta) + \sqrt{\tau}.
\end{aligned}
$$

The second inequality is obtained by the fact that $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$. ∎

Lemma 6.6 *We have*

$$\delta^2 = O(n^2\nu_k\mu_0).$$

Proof. From (3.8), we have

$$\|\mathbf{A}^+ r_k^p\| \;\leq\; \theta_k\gamma_1\rho, \quad \text{and} \quad \|r_k^d\| \;\leq\; \theta_k\gamma_1\rho. \tag{6.19}$$

By Lemma 6.1, (6.18), and using the fact that the Frobenius norm $\|M\| \leq \mathrm{Tr}(M)$ for $M \in \mathcal{S}_+^n$, we have

$$
\begin{aligned}
\|D_k^{-T}\mathbf{A}^+ r_k^p\|^2 & \leq \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|Z_k\|^2\|\mathbf{A}^+ r_k^p\|^2 \\
& \leq \frac{1}{(1-\gamma_2)\nu_k\mu_0}\mathrm{Tr}(Z_k)^2\|\mathbf{A}^+ r_k^p\|^2 \\
& \leq \frac{\gamma_1^2\rho^2}{(1-\gamma_2)\nu_k\mu_0}\,\theta_k^2\mathrm{Tr}(Z_k)^2 \\
& = \frac{1}{(1-\gamma_2)\nu_k\mu_0}O(n^2\nu_k^2\rho^4), \quad \text{by Lemma 2.2} \\
& = O(n^2\nu_k\mu_0).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\|D_k r_k^d\|^2 & \leq \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|X_k\|^2\|r_k^d\|^2 \\
& \leq \frac{1}{(1-\gamma_2)\nu_k\mu_0}\mathrm{Tr}(X_k)^2\|r_k^d\|^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\gamma_1^2\rho^2}{(1-\gamma_2)\nu_k\mu_0}\,\theta_k^2 \mathrm{Tr}(X_k)^2 \\
&\leq \frac{1}{(1-\gamma_2)\nu_k\mu_0} O(n^2\nu_k^2\rho^4), \quad \text{by Lemma 2.2} \\
&= O(n^2\nu_k\mu_0).
\end{aligned}
$$

Therefore, $\delta^2 = O(n^2\nu_k\mu_0)$. □

Lemma 6.7 *Under the conditions (2.1), (2.11) and (2.12),*

$$(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) \;=\; O(n\mu_0).$$

Proof. Using the fact that $\tilde{X}_k, \tilde{Z}_k, X_*$ and Z_* are in $\mathcal{S}_+^n$, (6.11) and (6.12),

$$
\begin{aligned}
(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) &= \tilde{X}_k \bullet \tilde{Z}_k - \tilde{X}_k \bullet Z_* - X_* \bullet \tilde{Z}_k + X_* \bullet Z_* \\
&\leq \tilde{X}_k \bullet \tilde{Z}_k \\
&\leq (1+\gamma_2)^2 n\rho^2 \;=\; (1+\gamma_2)^2 n\mu_0,
\end{aligned}
$$

which proves the lemma. □

Lemma 6.8 *Under the conditions (2.1), (2.11) and (2.12),*

$$\theta_k^2\beta^2 = O(n^2\nu_k\mu_0).$$

Proof. By Lemma 6.1, and noting that $\tilde{X}_k \succeq X_*$, we have

$$
\begin{aligned}
\|D_k^{-T}\mathbf{vec}(\tilde{X}_k - X_*)\|^2 &\leq \frac{1}{(1-\gamma_2)\nu_k\mu_0}\|Z_k^{\frac{1}{2}}(\tilde{X}_k - X_*)Z_k^{\frac{1}{2}}\|^2 \\
&\leq \frac{1}{(1-\gamma_2)\nu_k\mu_0}\mathrm{Tr}\left(Z_k^{\frac{1}{2}}(\tilde{X}_k - X_*)Z_k^{\frac{1}{2}}\right)^2 \\
&= \frac{1}{(1-\gamma_2)\nu_k\mu_0}\left((\tilde{X}_k - X_*) \bullet Z_k)\right)^2,
\end{aligned}
$$

and similarly,

$$\|D_k\mathbf{vec}(\tilde{Z}_k - Z_*)\|^2 \;\leq\; \frac{1}{(1-\gamma_2)\nu_k\mu_0}\left(X_k \bullet (\tilde{Z}_k - Z_*)\right)^2.$$

Thus,

$$\theta_k\beta \leq \frac{1}{\sqrt{(1-\gamma_2)\nu_k\mu_0}}\theta_k\left((\tilde{X}_k - X_*) \bullet Z_k + X_k \bullet (\tilde{Z}_k - Z_*)\right).$$

Using (6.13) and the fact that $X_* \bullet Z_* = 0$, we get

$$\theta_k\left((\tilde{X}_k - X_*) \bullet Z_k + X_k \bullet (\tilde{Z}_k - Z_*)\right)$$

$$
\begin{aligned}
&= X_k \bullet Z_k - X_k \bullet Z_* - X_* \bullet Z_k + \theta_k(\tilde{X}_k \bullet Z_* + X_* \bullet \tilde{Z}_k) \\
&\quad + \theta_k^2(\tilde{X}_k \bullet \tilde{Z}_k - \tilde{X}_k \bullet Z_* - X_* \bullet \tilde{Z}_k) \\
&\leq X_k \bullet Z_k + \theta_k(\tilde{X}_k \bullet Z_* + X_* \bullet \tilde{Z}_k) + \theta_k^2 \tilde{X}_k \bullet \tilde{Z}_k \\
&\leq (1+\gamma_2)n\nu_k\mu_0 \ + \ \nu_k(1+\gamma_1)\rho(I \bullet Z_* + X_* \bullet I) \ + \ \nu_k^2(1+\gamma_1)^2 n\mu_0 \\
&\leq 6n\nu_k\mu_0 + 2\nu_k\rho\,(\mathrm{Tr}(X_*) + \mathrm{Tr}(Z_*)) \\
&\leq 6n\nu_k\mu_0 + 2n\nu_k\rho^2, \quad \text{(by (2.12))} \\
&= 8n\nu_k\mu_0.
\end{aligned}
$$

Hence

$$\theta_k^2 \beta^2 \leq \frac{64n^2\nu_k\mu_0}{1-\gamma_2} = O(n^2\nu_k\mu_0),$$

and this completes the proof. ∎

Proof of Lemma 5.1. It follows Lemma 6.5 and the fact that $(a+b)^2 \leq 2a^2+2b^2$ that

$$
\begin{aligned}
t^2 &\leq (2(\theta_k\beta+\delta)+\sqrt{\tau})^2 \ \leq \ 8(\theta_k\beta+\delta)^2 + 2\tau, \\
\tau &\leq \|S_k^{-\frac{1}{2}}\mathbf{vec}R_k^c\|^2 + 2\theta_k^2(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) + 2(\theta_k\beta+\delta)\delta.
\end{aligned}
$$

Thus

$$
\begin{aligned}
t^2 &\leq 8(\theta_k\beta+\delta)(\theta_k\beta+2\delta) + 2\|S_k^{-\frac{1}{2}}\mathbf{vec}R_k^c\|^2 + 4\theta_k^2(\tilde{X}_k - X_*) \bullet (\tilde{Z}_k - Z_*) \\
&\leq O(n^2\nu_k\mu_0) \ + \ O(n\nu_k\mu_0) \ + \ O(n\nu_k^2\mu_0),
\end{aligned}
\tag{6.20}
$$

where we used Lemmas 6.3, 6.6, 6.7, and 6.8 in proving the last inequality. By Lemma 6.2, we have

$$\|H(\Delta X_k \Delta Z_k)\| \ \leq \ \frac{\kappa}{2}t^2 \ = \ O(n^2\nu_k\mu_0),$$

and this proves Lemma 5.1. ∎

7 Concluding Remarks

In this paper, we presented an inexact infeasible interior-point algorithm for SDP. Our algorithm uses an inexact search direction in each iteration by solving a linear system of equations approximately. Our algorithm bears some resemblances to the exact infeasible algorithm proposed in [18], and part of our complexity analysis is modeled after that paper. In each iteration of our algorithm, we used an inexact version of the HKM direction, but it is easy to extend our result to other search directions such as the Nesterov-Todd direction in the Monterio-Zhang family [10].

It is known that SDP includes linear programming (LP) as a special case. Thus the polynomial complexity bound we established in this paper also holds for LP. By adapting the algorithm proposed in this paper to LP, we obtain an inexact infeasible interior-point algorithm which can find an ϵ-approximate solution in $O(n^2\ln(1/\epsilon))$ iterations. The bound is the same as that of the exact infeasible interior point algorithms proposed by Zhang [17] and Mizuno [6].

Appendix

We list some useful properties of Kronecker products that are used in this paper. We use $\phi(A)$ denotes the spectrums of A.

1. $A \otimes B = [a_{ij}B]$.
2. $\mathbf{vec}(AXB) = (B^T \otimes A)\mathbf{vec}X$.
3. $(A \otimes B)^T = A^T \otimes B^T$.
4. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
5. $(A \otimes B)(C \otimes D) = AC \otimes BD$.
6. $A \otimes I$ is symmetric iff A is.
7. $\phi(A) = \{\mu_i\}$, $\phi(B) = \{\nu_i\} \Rightarrow \phi(A \otimes B) = \{\mu_i \nu_j\}$.
8. $\|A \otimes B\| = \|A\|\|B\|$.

References

[1] R. W. Freund, F. Jarre, and S. Mizuno, *Convergence of a class of inexact interior point algorithms for linear programs*, Mathematics of Operations Research, 24 (1999), pp. 50–71.

[2] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM J. Optimization, 6 (1996), pp. 342–361.

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

[4] M. Kojima, M. Shida, and S. Shindoh, *Search directions in the SDP and the monotone SDLCP: generalization and inexact computation*, Mathematical Programming, 85 (1999), pp. 51–80.

[5] J. Korzak, *Convergence analysis of inexact infeasible-interior-point algorithms for solving linear programming problems*, SIAM J. Optimization, 11 (2000), no. 1, pp. 133–148.

[6] S. Mizuno, *Polynomiality of infeasible interior point method for linear programming*, Mathematical Programming, 67 (1994), pp. 52–67.

[7] S. Mizuno and F. Jarre, *Global and polynomial-time convergence of an infeasible interior point algorithm using inexact computation*, Mathematical Programming, 84 (1999), pp. 105–122.

[8] R. D. C. Monteiro, *Primal-dual path following algorithms for semidefinite programming*, SIAM J. Optimization, 7 (1997), pp. 663–678.

[9] R. D. C. Monteiro and T. Tsuchiya, *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming*, SIAM J. Optimization, 9 (1999), pp. 551–577.

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

[11] M. J. Todd, *On search directions in interior-point methods for semidefinite programming*, Optimization Methods and Software, 11 (1999), pp. 1–46.

[12] M. J. Todd, *Semidefinite Optimization*, Acta Numerica, 10 (2001), pp. 515–560.

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

[14] K. C. Toh, and M. Kojima, *Solving some large scale semidefinite programs via the conjugate residual method*, SIAM J. Optimization, to appear.

[15] H. Wolkowicz, Bibliography on semidefinite programming,
`http://liinwww.ira.uka.de/bibliography/Math/psd.html.`

[16] H. Wolkowicz, R. Saigal and L. Vandenberghe (eds), *Handbook of Semidefinite Programming*, Kluwer Academic Publishers, Boston-Dordrecht-London, 2000.

[17] Y. Zhang, *On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem*, SIAM J. Optimization, 4 (1994), pp. 208–227.

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