An Analytic Center Cutting Plane Method For Semidefinite Feasibility Problems

Jie Sun, * Kim-Chuan Toh, † and Gongyun Zhao ‡

Abstract. Semidefinite feasibility problems arise in many areas of operations research. The abstract form of these problems can be described as finding a point in a nonempty bounded convex body Γ in the cone of symmetric positive semidefinite matrices. Assume that Γ is defined by an oracle, which for any given $m \times m$ symmetric positive semidefinite matrix $\hat{Y}$ either confirms that $\hat{Y} \in \Gamma$ or returns a cut, i.e., a symmetric matrix A such that Γ is in the half-space $\{Y \ : \ A \bullet Y \leq A \bullet \hat{Y}\}$. We study an analytic center cutting plane algorithm for this problem. At each iteration the algorithm computes an approximate analytic center of a working set defined by the cutting-plane system generated in the previous iterations. If this approximate analytic center is a solution, then the algorithm terminates; otherwise the new cutting plane returned by the oracle is added into the system. As the number of iterations increases, the working set shrinks and the algorithm eventually finds a solution of the problem. All iterates generated by the algorithm are positive definite matrices. The algorithm has a worst case complexity of $O^*(m^3/\epsilon^2)$ on the total number of cuts to be used, where ϵ is the maximum radius of a ball contained by Γ.

Keywords: Analytic center, cutting plane methods, semidefinite programming.

AMS subject classification: 65K05, 90C25

---

*Department of Decision Sciences and Singapore-MIT Alliance, National University of Singapore (NUS), 10 Kent Ridge Crescent, Singapore 119620 (`jsun@nus.edu.sg`). His research is partially supported by NUS Academic Research Grant R-314-000-026/028-112 and Singapore-MIT Alliance.

†Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119620 (`mattohkc@math.nus.edu.sg`). His research is supported in part by Singapore-MIT Alliance.

‡Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119620 (`matzgy@math.nus.edu.sg`). His research is supported in part by NUS Academic Research Grant R-146-000-006-112.

1 Introduction

1.1 Motivation

Let $\mathcal{S}^m$ be the set of $m \times m$ symmetric matrices and let $\mathcal{S}^m_+$ be its subset of symmetric positive semidefinite matrices. We consider the problem of finding a point in a convex body Γ of $\mathcal{S}^m_+$, where Γ is defined by an oracle. For any $\hat{Y} \in \mathcal{S}^m_+$ the oracle either confirms that $\hat{Y} \in \Gamma$ or returns a cut, i.e., an $A \in \mathcal{S}^m$ such that Γ is in the half-space $\{Y \in \mathcal{S}^m \ : \ A \bullet Y \leq A \bullet \hat{Y}\}$. The semidefinite feasibility problem (SFP for short), as described above, has many applications in operations research [1, 8, 9]. For example, arising from Lyapunov stability analysis of systems under uncertainty, it is desired to know whether the following system is feasible

$$\begin{cases} \lambda Y - (X_i^T Y + Y X_i) \in \mathcal{S}^m_+ \quad i = 1, ..., k \\ Y - I \in \mathcal{S}^m_+ \\ Y \in \mathcal{S}^m_+ \end{cases}$$

for given $\lambda \in R$ and $I, X_i \in I\!R^{m \times m}$, $i = 1, ..., k$. See [1, 8] for more details. Although this problem could be theoretically treated as a semidefinite program with a trivial objective function, it is practically impossible to solve it in this way if k is very large, say $\geq 10,000$. In stead of handling all constraints at once, a cutting plane method (sometimes called "column generation method") generates the solution by gradually tightening up a simple set containing Γ and therefore could be an effective way to solve large-scale feasibility problems.

Another advantage of the cutting plane method is that it allows the set Γ to be defined in very general terms. For instances, (a) Γ could be implicitly defined such as the case, where the cut returned by the oracle is obtained by solving an optimization problem; (b) Γ could be defined by nonlinear constraints; or (c) Γ could be defined by infinitely many inequalities. In all such cases the feasibility problem can not be handled by solving a semidefinite program, but could be solved by using a cutting plane method.

There has been an extensive literature in recent years on semidefinite programming. The book of Nesterov and Nimirovskii [8] provides a classical treatment of this topic and convex programming in general. The review paper of Boyd and Vandenberghe [9] gives a comprehensive introduction to the theory, applications, and algorithms for semidefinite programming. The website of semidefinite programming (http://www.zib.de/helmberg/semidef.html) contains a nice categorized list of papers in this area. Recently, there has been also an increasing interest in cutting plane methods based on analytic centers. The paper of Goffin and Vial [3] and the references therein provide a convenient overview on the related work. Our paper attempts to make a connection between the semidefinite programming techniques and the analytic center cutting plane methods. We hope that the analysis in this paper can stimulate further research in both topics.

Our main technical references are [2] and [5]. In [2], Goffin, Luo, and Ye developed an analytic center cutting plane method for $\Gamma \in \mathbb{R}^m$, while in [5], Luo and Sun contributed a cutting surface method for Γ defined by self-concordant functions. Our paper is different from [2] and [5] in the following aspects. Since we are dealing with space $\mathcal{S}_+^m$ rather than $\mathbb{R}_+^m$, our working set is no longer the polytope used in [2]. In addition, a number of important estimates in [2] has to be re-built using matrix analysis. Compared to [5], we use moderately deep cuts rather than shallow cuts (for the meanings of shallow and moderately deep cuts, see [4]). Moreover, we include the barrier term of $\mathcal{S}_+^m$ in our potential function. Therefore the method proposed in this paper guarantees that all iterates are positive definite, a favorable property for many applications.

1.2 Notations and Assumptions

For matrices $A, Y \in \mathcal{S}^m$, we define

$$A \bullet Y := \operatorname{tr}(A^T Y) = \sum_{i,j=1}^{m} A_{ij} Y_{ij},$$

where "T" stands for the transpose. We write $Y \succ 0$ and $Y \succeq 0$ if Y is positive definite and positive semidefinite, respectively. For $Y \succeq 0$, we denote its symmetric square root by $Y^{1/2}$. The 2-norm of a vector x is denoted by $\|x\|$. For $A \in \mathcal{S}^m$, we write

$$\|A\|_F := (A \bullet A)^{1/2}, \quad \|A\|_2 := \max\{\|Ax\|/\|x\| \ : \ \|x\| = 1\}, \quad \lambda(A) := (\lambda_1(A), \ldots, \lambda_m(A))^T,$$

where $\lambda_1(A), \ldots, \lambda_m(A)$ are the eigenvalues of A. Note that $\|A\|_F = \|\lambda(A)\|$ and $\|A\|_2 = \|\lambda(A)\|_\infty$. We shall use these facts in the paper without explicitly mentioning them.

Generally, we use capital letters for matrices, lower case ones for vectors, and Greek letters for scalars.

Let **svec** be a linear isometry identifying $\mathcal{S}^m$ with $\mathbb{R}^{m(m+1)/2}$ so that $K \bullet L = \mathbf{svec}(K)^T \mathbf{svec}(L)$ and let **smat** be the inverse of **svec**. Given any $m \times m$ symmetric matrix G, we define the linear map $G \circledast G \, : \, \mathbb{R}^{m(m+1)/2} \longrightarrow \mathbb{R}^{m(m+1)/2}$ by

$$(G \circledast G)\, \mathbf{svec}(M) \;=\; \mathbf{svec}(GMG).$$

It is easy to see that If G is positive definite, then $G \circledast G$ is positive definite, $(G \circledast G)^{-1} = G^{-1} \circledast G^{-1}$, and $(G \circledast G)^{1/2} = G^{1/2} \circledast G^{1/2}$.

We make the following assumptions:

A1: Γ is a convex subset of $\mathcal{S}^m$.

A2: Γ contains a full dimensional ball of radius $\epsilon > 0$. That is, there exists $Y^\epsilon \in \mathcal{S}^m$ such that $\{Y \in \mathcal{S}^m : \|Y - Y^\epsilon\|_F \leq \epsilon\} \subset \Gamma$.

A3: $\Gamma \subset \Omega_0$, where

$$\Omega_0 := \{Y \in \mathcal{S}^m : 0 \preceq Y \preceq I\}.$$

The upper bound in Assumption A3 is made for convenience. It can be satisfied by scaling if the original convex set $\hat{\Gamma}$ is bounded. That is, suppose there exists a constant $\gamma > 0$ such that for all $Y \in \hat{\Gamma}$, $\|Y\|_2 \leq \gamma$. Then the scaled set $\Gamma = \{\hat{Y}/\gamma \ : \ \hat{Y} \in \hat{\Gamma}\}$ satisfies A3.

This paper is organized as follows. In Section 2 we describe an analytic center cutting plane method for SFP, which includes an introduction to the log-barrier potential function. Section 3 shows the quadratic convergence of the Newton method applied to the potential function. This result will be used in the complexity analysis. In Section 4, we discuss how to update to a new approximate analytic center after each call of the oracle. Section 5 is devoted to estimating the reduction in the potential function and the analysis of complexity. Our main result shows that the proposed method will terminate with a feasible point in $O^*(m^3/\epsilon^2)$ Newton steps, where the notation O^* means that lower order terms are ignored.

2 An analytic center cutting plane method

We first define the analytic center and then propose an analytic center cutting plane method at the end of this section. Let $A_i \bullet Y \leq c_i$, $i = 1, \ldots, k$, be all the cuts defining the current working set Ω. Define

$$\mathcal{A} := (\mathbf{svec}A_1, \mathbf{svec}A_2, \ldots, \mathbf{svec}A_k), \quad c := (c_1, c_2, \ldots, c_k)^T.$$

Then the set Ω can be represented by

$$\Omega = \left\{Y \in \Omega_0 \ : \ \mathcal{A}^T\mathbf{svec}Y \leq c\right\}.$$

We define the following *potential function* on the set Ω:

$$\phi(Y) \ = \ -\sum_{i=1}^{k} \ln(c_i - A_i \bullet Y) - \ln(\det Y) - \ln\left[\det(I - Y)\right],$$

and denote

$$\phi(\Omega) \ := \ \min\{\phi(Y) \ : \ Y \in \Omega\}. \tag{2.1}$$

We shall use the notation $\phi_k(\cdot)$ to denote the potential function associated with the kth working set Ω_k when necessary.

It is easy to see that the analytic center of set Ω_0 is $I/2$, where I is the unit matrix. As a matter of fact,

$$\begin{aligned}
\phi_0(Y) &= -\ln(\det Y) - \ln\left[\det(I-Y)\right] \\
&= -\ln\left[\prod_{i=1}^{m}\lambda_i(Y)\right] - \ln\left[\prod_{i=1}^{m}\lambda_i(I-Y)\right] \\
&= -\sum_{i=1}^{m}\ln\left[\lambda_i(Y)(1-\lambda_i(Y))\right].
\end{aligned}$$

The minimum of $\phi_0(Y)$ must satisfy $\lambda_1(Y) = \cdots = \lambda_m(Y) = 1/2$. Hence $Y = I/2$.

It is known [8, Proposition 5.4.5] that ϕ is a strongly 1-self-concordant function on Ω and

$$\begin{aligned}
\nabla\phi(Y) &= \mathbf{svec}\left[\sum_{i=1}^{k}\frac{A_i}{c_i - A_i \bullet Y} - Y^{-1} + (I-Y)^{-1}\right], \\
\nabla^2\phi(Y) &= \mathcal{A}S^{-2}\mathcal{A}^T + Y^{-1} \circledast Y^{-1} + (I-Y)^{-1} \circledast (I-Y)^{-1},
\end{aligned}$$

where $s_i = c_i - A_i \bullet Y$ and $S = \text{diag}\,(s_1, \cdots, s_k)$. Let $x = (x_1, \ldots, x_k)^T$ and $s = (s_1, \ldots, s_k)^T$. The analytic center Y is determined by $\nabla\phi(Y) = 0$ or, equivalently, by the following optimality conditions:

$$\begin{aligned}
Sx &= e \quad (e \text{ denotes the vector of ones}) \\
YZ &= I \\
(I-Y)V &= I \\
\mathcal{A}^T\mathbf{svec}Y + s &= c \\
\mathcal{A}x - \mathbf{svec}Z + \mathbf{svec}V &= 0 \\
I \succ Y \succ 0, \quad Z, V \succ 0, &\quad\quad s, x > 0,
\end{aligned} \tag{2.2}$$

where x, s, Z, and V are auxiliary variables. With a slight abuse of language, we also call (x, s, Y, Z, V) the analytic center of Ω for convenience if it is the solution of (2.2). Since the potential function is strictly convex and tends to infinity near the boundary of Ω, the solution of (2.2) exists and is unique.

An analytic center cutting plane method

Step 0. Set $k = 0$. Let Ω_0 be the initial working set and let $Y_0 = I/2$ be the initial point.

Step 1. At the kth iteration call for the oracle which either confirms that Y_k is a feasible point of Γ or returns a matrix $A_{k+1} \in \mathcal{S}^m$ with $\|A_{k+1}\|_F = 1$. If $Y_k \in \Gamma$, stop; Otherwise, construct the new working set

$$\Omega_{k+1} = \Omega_k \cap \{Y \;:\; A_{k+1} \bullet Y \leq A_{k+1} \bullet Y_k\}.$$

Step 2. Find a point Y in the interior of Ω_{k+1}.

Step 3. Find an approximate analytic center Y_{k+1} of Ω_{k+1} by using the following (dual) Newton procedure with starting point Y:

$$Y_+ = Y - \left[\nabla^2 \phi_{k+1}(Y)\right]^{-1} \nabla \phi_{k+1}(Y),$$

where ϕ_{k+1} is the potential function of Ω_{k+1}. Update k and go to Step 1.

Remarks: Step 2 is necessary because Y_k lies on the boundary of Ω_{k+1} and it cannot be used as a starting point for the dual Newton procedure. We will show in Section 4 how to find a point Y in the interior of Ω_{k+1}, and that Step 3 needs no more than 4 Newton iterations to find an approximate analytic center.

3 A dual Newton procedure for computing analytic centers

We discuss a quadratic convergence property of the Newton method applied to our potential function. We shall use it in the analysis of our cutting plane method in subsequent sections. The result is also of interest on its own.

Definition 3.1 *Given a point* $(x, s, Y, Z, V) \in I\!R^k \times I\!R^k \times \mathcal{S}^m \times \mathcal{S}^m \times \mathcal{S}^m$ *with* $0 \prec Y \prec I$*, we define*

$$\eta(x, s, Y, Z, V) \;=\; \sqrt{\|Sx - e\|^2 + \|\lambda(YZ) - e\|^2 + \|\lambda[(I - Y)V] - e\|^2}. \tag{3.1}$$

We call (x, s, Y, Z, V) *an* η*-approximate (analytic) center of* Ω *if* $\eta(x, s, Y, Z, V) \leq \eta$*, all the linear equalities in* (2.2) *are satisfied,* $x, s > 0$*, and* $Z, V \succ 0$*. Obviously, a* 0*-approximate center is exactly the analytic center of* Ω*.*

Definition 3.2 *Given* $Y \in \mathcal{S}^m$ *with* $0 \prec Y \prec I$*, let* $s = c - \mathcal{A}^T \mathbf{svec}(Y)$*. We define*

$$\delta(Y) \;=\; \min\left\{\eta(x, s, Y, Z, V) \,:\, \mathcal{A}x - \mathbf{svec}(Z) + \mathbf{svec}(V) = 0,\; x \in I\!R^k,\; Z, V \in \mathcal{S}^m\right\}. \tag{3.2}$$

Here and below, we designate

$$\mathbf{e} \;=\; [e,\, \mathbf{svec}(I),\, \mathbf{svec}(I)]^T .$$

The next two lemmas establish some results for the function $\eta(x, s, Y, Z, V)$ when (x, Z, V) is a specially given point.

Lemma 3.3 *Given* $Y \in S^m$ *such that* $0 \prec Y \prec I$*. Suppose* $s = c - \mathcal{A}^T \mathbf{svec}(Y) > 0$*. Then*

$$\delta(Y) \;=\; \sqrt{\nabla\phi(Y)^T \left[\nabla^2\phi(Y)\right]^{-1} \nabla\phi(Y)} \;=\; \eta(x_Y, s, Y, Z_Y, V_Y), \tag{3.3}$$

where (x_Y, Z_Y, V_Y) *is the unique minimizer of (3.2).*

Proof. Let $U = I - Y$ and

$$\begin{aligned} G &= \left[\mathcal{A}S^{-1}, \quad -Y^{-1/2} \circledast Y^{-1/2}, \quad U^{-1/2} \circledast U^{-1/2}\right], \\ H &= \nabla^2\phi(Y) \;=\; \mathcal{A}S^{-2}\mathcal{A}^T + Y^{-1} \circledast Y^{-1} + U^{-1} \circledast U^{-1}. \end{aligned}$$

Then $H = GG^T$ and for any (x, Z, V) we have

$$\mathcal{A}x - \mathbf{svec}(Z) + \mathbf{svec}(V) \;=\; G \begin{pmatrix} Sx \\ \mathbf{svec}(Y^{1/2} Z\, Y^{1/2}) \\ \mathbf{svec}(U^{1/2} V\, U^{1/2}) \end{pmatrix}.$$

It is readily seen that the minimization problem (3.2) can be re-written as follows:

$$\delta(Y) = \min \left\{ \left\| \mathbf{e} - \begin{pmatrix} Sx \\ \mathbf{svec}(Y^{1/2} Z\, Y^{1/2}) \\ \mathbf{svec}(U^{1/2} V\, U^{1/2}) \end{pmatrix} \right\| \; : \; G \begin{pmatrix} Sx \\ \mathbf{svec}(Y^{1/2} Z\, Y^{1/2}) \\ \mathbf{svec}(U^{1/2} V\, U^{1/2}) \end{pmatrix} = 0 \right\}. \tag{3.4}$$

The minimum value in (3.4) and its unique minimizer (x_Y, Z_Y, V_Y) is easily shown to be given by

$$\delta(Y) \;=\; \|\mathcal{P}\mathbf{e}\|, \qquad \begin{pmatrix} S\, x_Y \\ \mathbf{svec}(Y^{1/2} Z_Y\, Y^{1/2}) \\ \mathbf{svec}(U^{1/2} V_Y\, U^{1/2}) \end{pmatrix} \;=\; \mathbf{e} - \mathcal{P}\mathbf{e}, \tag{3.5}$$

where $\mathcal{P} := G^T H^{-1} G$ is the orthogonal projection of $\mathbb{R}^k \times \mathbb{R}^{m(m+1)/2} \times \mathbb{R}^{m(m+1)/2}$ onto the range of G^T. Hence we have

$$\delta(Y) \;=\; \sqrt{\mathbf{e}^T \mathcal{P}^T \mathcal{P}\mathbf{e}} \;=\; \sqrt{\mathbf{e}^T \mathcal{P}\mathbf{e}} \;=\; \sqrt{\mathbf{e}^T G^T H^{-1} G\mathbf{e}}.$$

The desired result is obtained by observing that $\nabla\phi(Y) = G\,\mathbf{e}$. $\square$

Lemma 3.4 *Let $(\bar{x}, \bar{s}, \bar{Y}, \bar{Z}, \bar{V})$ be the analytic center of Ω. Given $Y \in \mathcal{S}^m$ such that $0 \prec Y \prec I$, let $s = c - \mathcal{A}^T\mathbf{svec}(Y)$. We have*

$$\eta(\bar{x}, s, Y, \bar{Z}, \bar{V}) \;=\; \sqrt{\mathbf{svec}(Y - \bar{Y})^T\, \bar{H}\, \mathbf{svec}(Y - \bar{Y})},$$

where $\bar{H} = \nabla^2\phi(\bar{Y})$.

Proof. Let $U = I - Y$ and $\bar{U} = I - \bar{Y}$. We have

$$\begin{aligned} \eta(\bar{x}, s, Y, \bar{Z}, \bar{V})^2 \;&=\; \|\bar{X}s - e\|^2 + \left\|\bar{Z}^{1/2} Y \bar{Z}^{1/2} - I\right\|_F^2 + \left\|\bar{V}^{1/2} U \bar{V}^{1/2} - I\right\|_F^2 \\ &= \; \|\bar{X}(s - \bar{s})\|^2 + \left\|\bar{Z}^{1/2}(Y - \bar{Y})\bar{Z}^{1/2}\right\|_F^2 + \left\|\bar{V}^{1/2}(Y - \bar{Y})\bar{V}^{1/2}\right\|_F^2 \\ &= \; \left\|\bar{S}^{-1}\mathcal{A}^T\mathbf{svec}(Y - \bar{Y})\right\|^2 + \left\|\bar{Y}^{-1/2} \circledast \bar{Y}^{-1/2}\mathbf{svec}(Y - \bar{Y})\right\|^2 + \left\|\bar{U}^{-1/2} \circledast \bar{U}^{-1/2}\mathbf{svec}(Y - \bar{Y})\right\|^2 \\ &= \; \mathbf{svec}(Y - \bar{Y})^T \bar{H}\, \mathbf{svec}(Y - \bar{Y}). \end{aligned}$$

$\square$

Now we are ready to show the quadratic convergence of Newton's method for the dual potential function in $\mathcal{S}^m_+$. This extends the results in [2, 10] and some earlier literature from $I\!R^m_+$ to $\mathcal{S}^m_+$.

Theorem 3.5 *Given $Y \in \mathcal{S}^m$ that satisfies $0 \prec Y \prec I$, $s = c - \mathcal{A}^T\mathbf{svec}(Y) > 0$, and $\delta(Y) < 1$. Let $H = \nabla^2\phi(Y)$ and $g = \nabla\phi(Y)$. Suppose that $\Delta Y = -\mathbf{smat}(H^{-1}g)$ and $Y_+ = Y + \Delta Y$. Then $0 \prec Y_+ \prec I$, $s_+ = c - \mathcal{A}^T\mathbf{svec}(Y_+) > 0$, and*

$$\delta(Y_+) \;\le\; \delta(Y)^2. \tag{3.6}$$

Proof. First note that since

$$H^{-1/2}\left[\mathcal{A}S^{-2}\mathcal{A}^T \;+\; Y^{-1} \circledast Y^{-1} + (I-Y)^{-1} \circledast (I-Y)^{-1}\right] H^{-1/2} \;=\; I,$$

we have

$$H^{-1/2}(\mathcal{A}S^{-2}\mathcal{A}^T)H^{-1/2} \;\preceq\; I, \qquad H^{-1/2}(Y^{-1} \circledast Y^{-1})H^{-1/2} \;\preceq\; I, \tag{3.7}$$

$$H^{-1/2}[(I-Y)^{-1} \circledast (I-Y)^{-1}]H^{-1/2} \;\preceq\; I. \tag{3.8}$$

Let $U = I - Y$ and $U_+ = I - Y_+$. Then $s_+ = s - \mathcal{A}^T\mathbf{svec}(\Delta Y)$ and $U_+ = U - \Delta Y$. We first show that $s_+ > 0$. Since $s_+ = S(e - S^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y))$, it suffices to show that $\left\|S^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y)\right\| < 1$:

$$\begin{aligned}
\left\|S^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y)\right\|^2 &= \mathbf{svec}(\Delta Y)^T\mathcal{A}S^{-2}\mathcal{A}^T\mathbf{svec}(\Delta Y) \\
&= g^T H^{-1}(\mathcal{A}S^{-2}\mathcal{A}^T)H^{-1}g \\
&= g^T H^{-1/2}\left[H^{-1/2}(\mathcal{A}S^{-2}\mathcal{A}^T)H^{-1/2}\right] H^{-1/2}g \\
&\le g^T H^{-1}g \;=\; \delta(Y)^2 \;<\; 1.
\end{aligned}$$

Note that we used Lemma 3.3 in the proof above. Next we show that $Y_+ \succ 0$. Since $Y_+ = Y^{1/2}(I + Y^{-1/2}(\Delta Y)Y^{-1/2})Y^{1/2}$, it suffices to show that $\left\|Y^{-1/2}(\Delta Y)Y^{-1/2}\right\|_F < 1$:

$$\begin{aligned}
\left\|Y^{-1/2}(\Delta Y)Y^{-1/2}\right\|_F^2 &= \left\|(Y^{-1/2} \circledast Y^{-1/2})H^{-1}g\right\|^2 \\
&= g^T H^{-1}(Y^{-1} \circledast Y^{-1})H^{-1}g \\
&= g^T H^{-1/2}\left[H^{-1/2}(Y^{-1} \circledast Y^{-1})H^{-1/2}\right] H^{-1/2}g \\
&\le g^T H^{-1}g \;=\; \delta(Y)^2 \;<\; 1.
\end{aligned}$$

Similarly, we can show that $U_+ \succ 0$. Hence $Y_+ \prec I$.

Now we turn to the proof of (3.6). By the definition of $\delta(Y_+)$, we have

$$\delta(Y_+) \;=\; \left\|\begin{pmatrix} S_+\,x_{Y_+} \\ \mathbf{svec}(Y_+^{1/2}Z_{Y_+}\,Y_+^{1/2}) \\ \mathbf{svec}(U_+^{1/2}V_{Y_+}\,U_+^{1/2}) \end{pmatrix} - \mathbf{e}\right\| \;\le\; \left\|\begin{pmatrix} S_+\,x_{Y} \\ \mathbf{svec}(Y_+^{1/2}Z_{Y}\,Y_+^{1/2}) \\ \mathbf{svec}(U_+^{1/2}V_{Y}\,U_+^{1/2}) \end{pmatrix} - \mathbf{e}\right\|.$$

Note that since

$$G^T\mathbf{svec}(\Delta Y) \;=\; -G^T H^{-1} G\,\mathbf{e} \;=\; -\mathcal{P}\,\mathbf{e},$$

hence from (3.5) in the proof of Lemma 3.3, we have

$$\begin{pmatrix} S^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y) \\ -\mathbf{svec}(Y^{-1/2}(\Delta Y)\,Y^{-1/2}) \\ \mathbf{svec}(U^{-1/2}(\Delta Y)\,U^{-1/2}) \end{pmatrix} \;=\; \begin{pmatrix} S\,x_Y \,-\, e \\ \mathbf{svec}(Y^{1/2}Z_Y\,Y^{1/2}) \,-\, \mathbf{svec}(I) \\ \mathbf{svec}(U^{1/2}V_Y\,U^{1/2}) \,-\, \mathbf{svec}(I) \end{pmatrix}.$$

To complete the proof of (3.6), we shall consider the following parts:

(a)
$$\begin{aligned} \|S_+x_Y - e\| &= \left\|X_Y(s - \mathcal{A}^T\mathbf{svec}(\Delta Y)) - e\right\| \\ &= \left\|Sx_Y - SX_Y(S^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y)) - e\right\| \\ &= \|Sx_Y - SX_Y(Sx_Y - e) - e\| \;\le\; \|Sx_Y - e\|^2, \end{aligned}$$

where $X_Y = \mathrm{diag}\,(x_Y)$;

(b)
$$\begin{aligned} \left\|Y_+^{1/2}Z_YY_+^{1/2} - I\right\|_F &= \|\lambda(Y_+Z_Y - I)\| \;=\; \left\|Z_Y^{1/2}Y_+Z_Y^{1/2} - I\right\|_F \\ &= \left\|Z_Y^{1/2}YZ_Y^{1/2} + Z_Y^{1/2}(\Delta Y)Z_Y^{1/2} - I\right\|_F \\ &= \left\|Z_Y^{1/2}YZ_Y^{1/2} - Z_Y^{1/2}(YZ_YY - Y)Z_Y^{1/2} - I\right\|_F \\ &= \left\|(Z_Y^{1/2}YZ_Y^{1/2} - I)^2\right\|_F \;\le\; \left\|Z_Y^{1/2}YZ_Y^{1/2} - I\right\|_F^2; \end{aligned}$$

(c)
$$\begin{aligned} \left\|U_+^{1/2}V_YU_+^{1/2} - I\right\|_F &= \|\lambda(U_+V_Y - I)\| \;=\; \left\|V_Y^{1/2}U_+V_Y^{1/2} - I\right\|_F \\ &= \left\|V_Y^{1/2}UV_Y^{1/2} - V_Y^{1/2}(\Delta Y)V_Y^{1/2} - I\right\|_F \\ &= \left\|V_Y^{1/2}UV_Y^{1/2} - V_Y^{1/2}(UV_YU - U)V_Y^{1/2} - I\right\|_F \\ &= \left\|(V_Y^{1/2}UV_Y^{1/2} - I)^2\right\|_F \;\le\; \left\|V_Y^{1/2}UV_Y^{1/2} - I\right\|_F^2. \end{aligned}$$

Combining (a)–(c), we have

$$\begin{aligned} \delta(Y_+)^2 &\le \|S_+x_Y - e\|^2 + \left\|Y_+^{1/2}Z_YY_+^{1/2} - I\right\|_F^2 + \left\|U_+^{1/2}V_YU_+^{1/2} - I\right\|_F^2 \\ &\le \|Sx_Y - e\|^4 + \left\|Y^{1/2}Z_YY^{1/2} - I\right\|_F^4 + \left\|U^{1/2}V_YU^{1/2} - I\right\|_F^4 \\ &\le \left(\|Sx_Y - e\|^2 + \left\|Y^{1/2}Z_YY^{1/2} - I\right\|_F^2 + \left\|U^{1/2}V_YU^{1/2} - I\right\|_F^2\right)^2 \\ &= \delta(Y)^4. \end{aligned}$$

□

Theorem 3.6 *Let $(\bar{x}, \bar{s}, \bar{Y}, \bar{Z}, \bar{V})$ be the analytic center of Ω. Given $Y \in \mathcal{S}^m$ that satisfies $0 \prec Y \prec I$, $s = c - \mathcal{A}^T\mathbf{svec}(Y) > 0$, and $\delta(Y) < 1/3$. We have*

$$\eta(x_Y, s, Y, Z_Y, V_Y) \;\le\; \eta(\bar{x}, s, Y, \bar{Z}, \bar{V}) \;\le\; \frac{1 - [1 - 3\,\delta(Y)]^{1/3}}{[1 - 3\,\delta(Y)]^{1/3}}.$$

Proof. The right-hand side inequality follows from Theorem 2.2.2 (iii) in [8] (note that there is a typographic error in [8, (2.2.16)], where the numerator should not be squared) and Lemmas 3.3 and 3.4. The left-hand side inequality follows from the fact that (x_Y, Z_Y, V_Y) is the minimizer of $\eta(\cdot, s, Y, \cdot, \cdot)$. □

Remarks. This section studies the behavior of Newton's method applied to the potential function. Lemma 3.3 shows that the quantity $\delta(Y)$ is the so-called Newton decrement of ϕ at Y [8, (2.2.1)]. It also implies that a full Newton step is within the range of the Dikin ball at Y, $\{X \ : \ \|X - Y\|_H < 1, \ H = \nabla^2\phi(Y)\}$, and therefore the new iterate remains in Ω if $\delta(Y) < 1$. Theorem 3.5 further proves that if $\delta(Y) < 1$, then Y belongs to the region of quadratic convergence of Newton's method. Thus, to guarantee efficiency of the Newton method, it is crucial to find a point within this region, which will be the target of the next section.

The case of $\delta(Y) \geq 1$ is more sophisticated. It is shown in [8, Theorem 2.2.1] that at such a point there is no certificate for a full Newton step to be within Ω. Thus a damped Newton step is necessary. Moreover, if the stepsize is taken as $1/(1+\delta(Y))$ for instance, then the decrement will satisfy [8, Theorem 2.2.3]

$$\phi(Y_+) - \phi(Y) < \ln[1 + \delta(Y)] - \delta(Y).$$

Since we can easily find a point Y such that $\delta(Y) < 1$ with respect to ϕ_{k+1} as will be seen in the next section, we will not go in depth along this direction.

4 Construction of a strictly feasible starting point in Ω_{k+1}

Given an η-approximate center $(x^k, s^k, Y_k, Z_k, V_k)$ for Ω_k, where $\eta < 1$ is sufficiently small, suppose we add a cut $\{Y \in \mathcal{S}^m | A \bullet Y \leq A \bullet Y_k\}$ to Ω_k and call the new set Ω_{k+1}. We want to use Newton's method to find an approximate analytic center for Ω_{k+1}. However, Y_k is on the boundary of Ω_{k+1} on which the potential function for Ω_{k+1} is not defined. Thus, before using Newton's method, we need to construct a point (x, s, Y, Z, V) in the interior of Ω_{k+1} that is close enough to the center of Ω_{k+1} so as to guarantee that a few Newton steps from it will produce an η-approximate center for Ω_{k+1}. In mathematical terms, the point (x, s, Y, Z, V) should satisfy the following conditions.

$$\begin{pmatrix} \mathcal{A}^T \\ \mathbf{svec}(A)^T \end{pmatrix} \mathbf{svec}(Y) + s \ = \ \begin{pmatrix} c^k \\ \mathbf{svec}(A)^T \mathbf{svec}(Y_k) \end{pmatrix} \tag{4.1}$$

$$[\mathcal{A}, \ \ \mathbf{svec}(A)]\, x - \mathbf{svec}(Z) + \mathbf{svec}(V) \ = \ 0 \tag{4.2}$$

$$I \succ Y \succ 0, \quad Z, V \succ 0, \quad x, s > 0 \tag{4.3}$$

$$\eta(x, s, Y, Z, V) \ \leq \ \gamma, \tag{4.4}$$

for some γ that might be larger than η but should still be small, say $\gamma < 1$. Let

$$r_k \ = \ \sqrt{\mathbf{svec}(A)^T H_k^{-1}\, \mathbf{svec}(A)}, \tag{4.5}$$

where $H_k = \nabla^2 \phi_k(Y_k)$. The following theorem shows how such an interior point for Ω_{k+1} can be obtained.

Theorem 4.1 *Suppose* $(x^k, s^k, Y_k, Z_k, V_k)$ *is an* η*-approximate center for* Ω_k *with* $\eta < (\sqrt{2}-1)/(\sqrt{2}+1)$. *Let* $\beta = 1/\sqrt{2}$ *and*

$$\begin{aligned}
\Delta Y &= -\frac{\beta}{r_k}\, \mathbf{smat}(H_k^{-1}\mathbf{svec}(A)),\\
\Delta s &= -\mathcal{A}^T \mathbf{svec}(\Delta Y),\\
\Delta x &= -S_k^{-2}\Delta s \ = \ S_k^{-2}\mathcal{A}^T\mathbf{svec}(\Delta Y),\\
\Delta Z &= -\mathbf{smat}\left([Y_k^{-1} \circledast Y_k^{-1}]\, \mathbf{svec}(\Delta Y)\right),\\
\Delta V &= \mathbf{smat}\left([(I - Y_k)^{-1} \circledast (I - Y_k)^{-1}]\, \mathbf{svec}(\Delta Y)\right).
\end{aligned}$$

Consider (x, s, Y, Z, V) *defined as follows:*

$$\begin{aligned}
s &= \begin{pmatrix} c^k - \mathcal{A}^T\mathbf{svec}(Y) \\ \mathbf{svec}(A)^T\mathbf{svec}(Y_k) - \mathbf{svec}(A)^T\mathbf{svec}(Y) \end{pmatrix} = \begin{pmatrix} s^k + \Delta s \\ \beta\, r_k \end{pmatrix},\\
x &= \begin{pmatrix} x^k + \Delta x \\ \beta / r_k \end{pmatrix},\\
Y &= Y_k + \Delta Y, \qquad Z \ = \ Z_k + \Delta Z, \qquad V \ = \ V_k + \Delta V,
\end{aligned}$$

then Y *is in the interior of* Ω_{k+1} *and* (x, s, Y, Z, V) *satisfies the conditions* (4.1)–(4.4) *with* $\gamma = \eta + (1+\eta)/\sqrt{2}$.

An intuitive justification for the direction used in the above theorem is that ΔY is some sort of "Newton direction" with $\mathbf{svec}(A)$ as the gradient and H_k as the Hessian. Recall that $\mathbf{svec}(A)$ is perpendicular to the boundary of Ω_{k+1} at the current point Y_k and the potential function tends to infinity at the boundary of Ω_{k+1}, thus $\mathbf{svec}(A)$ can be thought of as the gradient on the "infinite contour" of the potential function of Ω_{k+1} at Y_k. Given ΔY, the other directions, Δs, Δx, ΔZ and ΔV, are motivated by our desire to satisfy the feasibility conditions (4.1)-(4.3). Similar constructions have been used by Mitchell [6] and Ye [10] in the vector case. We split the proof of Theorem 4.1 into two lemmas.

Lemma 4.2 *The point* (x, s, Y, Z, V) *defined in Theorem* 4.1 *satisfies conditions* (4.1)–(4.3)*.*

Proof. Let

$$\begin{aligned} p &= S_k^{-1}\Delta s = -S_k^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y), \\ Q &= Y_k^{-1/2}(\Delta Y)\, Y_k^{-1/2}. \end{aligned}$$

It is clear that (x, s, Y, Z, V) satisfies equation (4.1), but we still need to show that $s > 0$. Note that

$$s = \begin{pmatrix} s^k + \Delta s \\ \beta\, r_k \end{pmatrix} = \begin{pmatrix} S_k(e+p) \\ \beta\, r_k \end{pmatrix} > 0,$$

since $e + p \geq (1-\beta)e > 0$. Now we verify that (x, s, Y, Z, V) satisfies equation (4.2):

$$\begin{aligned} &[\mathcal{A},\ \mathbf{svec}(A)]\, x - \mathbf{svec}(Z) + \mathbf{svec}(V) \\ =\ & \mathcal{A}\, x^k - \mathbf{svec}(Z_k) + \mathbf{svec}(V_k) + (\beta/r_k)\mathbf{svec}(A) + \mathcal{A}\,\Delta x \\ & + (Y_k^{-1} \circledast Y_k^{-1})\mathbf{svec}(\Delta Y) + \left[(I - Y_k)^{-1} \circledast (I - Y_k)^{-1}\right]\mathbf{svec}(\Delta Y) \\ =\ & (\beta/r_k)\,\mathbf{svec}(A) + \left[\mathcal{A}S_k^{-2}\mathcal{A}^T + Y_k^{-1} \circledast Y_k^{-1} + (I - Y_k)^{-1} \circledast (I - Y_k)^{-1}\right]\mathbf{svec}(\Delta Y) \\ =\ & (\beta/r_k)\,\mathbf{svec}(A) + H_k\,\mathbf{svec}(\Delta Y) = 0. \end{aligned}$$

Furthermore,

$$x = \begin{pmatrix} x^k + \Delta x \\ \beta/r_k \end{pmatrix} = \begin{pmatrix} S_k^{-1}\left(S_k x^k - (S_k)^{-1}\Delta s\right) \\ \beta/r_k \end{pmatrix} = \begin{pmatrix} S_k^{-1}(S_k x^k - p) \\ \beta/r_k \end{pmatrix} > 0$$

since $S_k x^k - p \geq (1-\eta-\beta)\, e > 0$. Next we show that $Y \succ 0$, $Z \succ 0$ and $V \succ 0$. Firstly,

$$Y = Y_k + \Delta Y = Y_k^{1/2}\left(I + Y_k^{-1/2}(\Delta Y)Y_k^{-1/2}\right)Y_k^{1/2} = Y_k^{1/2}(I+Q)Y_k^{1/2} \succ 0,$$

since $I + Q \succeq (1-\beta)I \succ 0$. Secondly

$$Z = Z_k - Y_k^{-1}(\Delta Y)Y_k^{-1} = Y_k^{-1/2}\left(Y_k^{1/2}Z_kY_k^{1/2} - Q\right)Y_k^{-1/2} \succ 0$$

since $Y_k^{1/2}Z_kY_k^{1/2} - Q \succeq (1-\eta-\beta)I \succ 0$. We shall skip the proof for $V \succ 0$ as it is similar to the one we have just given for Z. $\square$

Lemma 4.3 *The point* (x, s, Y, Z, V) *defined in Theorem* 4.1 *satisfies condition* (4.4) *with* $\gamma = \eta + (1+\eta)/\sqrt{2} < 1$*.*

Proof. Let p and Q be defined the same as in the proof of Lemma 4.2 and let

$$R \;=\; (I - Y_k)^{-1/2}(\Delta Y)\,(I - Y_k)^{-1/2}.$$

It is easily verified that

$$\begin{aligned}
\|p\|^2 + \|Q\|_F^2 + \|R\|_F^2 &= \mathbf{svec}(\Delta Y)^T H_k\, \mathbf{svec}(\Delta Y) \\
&= \beta^2 \left[H_k^{-1}\mathbf{svec}(A)\right]^T H_k \left[H_k^{-1}\mathbf{svec}(A)\right]/r_k^2 \\
&= \beta^2.
\end{aligned}$$

We have

$$Xs - e \;=\; \begin{pmatrix} (S_k X_k - P)(e+p) - e \\ \beta^2 - 1 \end{pmatrix} = \begin{pmatrix} (S_k x^k - e) + (S_k X_k - I)p - Pp \\ \beta^2 - 1 \end{pmatrix},$$

where $P = \mathrm{diag}(p)$. Thus,

$$\|Xs - e\| \;\le\; \left\|S_k x^k - e\right\| + \left\|S_k x^k - e\right\|\,\|p\| + \sqrt{\|p\|^4 + (1-\beta^2)^2}. \tag{4.6}$$

Also

$$YZ - I \;=\; Y_k^{1/2}\left[Y_k^{1/2} Z_k Y_k^{1/2} - I + Q\left(Y_k^{1/2} Z_k Y_k^{1/2} - I\right) - Q^2\right] Y_k^{-1/2},$$

implying that

$$\begin{aligned}
\|\lambda(YZ) - e\| &= \left\|Y_k^{1/2} Z_k Y_k^{1/2} - I + Q\left(Y_k^{1/2} Z_k Y_k^{1/2} - I\right) - Q^2\right\|_F \\
&\le \left\|Y_k^{1/2} Z_k Y_k^{1/2} - I\right\|_F + \left\|Y_k^{1/2} Z_k Y_k^{1/2} - I\right\|_F \|Q\|_F + \|Q\|_F^2 \\
&= \|\lambda(Y_k Z_k) - e\|\,(1 + \|Q\|_F) + \|Q\|_F^2\,. \qquad (4.7)
\end{aligned}$$

Similarly

$$\begin{aligned}
\|\lambda((I-Y)V) - e\| &\le \left\|U_k^{1/2} V_k U_k^{1/2} - I\right\|_F + \left\|U_k^{1/2} V_k U_k^{1/2} - I\right\|_F \|R\|_F + \|R\|_F^2 \\
&= \|\lambda(U_k V_k) - e\|\,(1 + \|R\|_F) + \|R\|_F^2\,, \qquad (4.8)
\end{aligned}$$

where $U_k = I - Y_k$. Combining (4.6)–(4.8) together, we get

$$\begin{aligned}
&\eta(x, s, Y, Z, V) \\
\le\; &\left[\left\|X_k s^k - e\right\|^2 (1 + \|p\|)^2 + \|\lambda(Y_k Z_k) - e\|^2 (1 + \|Q\|_F)^2 + \|\lambda(U_k V_k) - e\|^2 (1 + \|R\|_F)^2\right]^{1/2} \\
&+ \left[\|p\|^4 + \|Q\|_F^4 + \|R\|_F^4 + (1-\beta^2)^2\right]^{1/2} \\
\le\; &(1+\beta)\left(\left\|X_k s^k - e\right\|^2 + \|\lambda(Y_k Z_k) - e\|^2 + \|\lambda(U_k V_k) - e\|^2\right)^{1/2}
\end{aligned}$$

$$+ \left[\left(\|p\|^2 + \|Q\|_F^2 + \|R\|_F^2 \right)^2 + (1-\beta^2)^2 \right]^{1/2}$$

$$\leq \ (1+\beta)\eta + \sqrt{\beta^4 + (1-\beta^2)^2}$$

$$= \ (1+1/\sqrt{2})\eta + 1/\sqrt{2} = \gamma. \qquad \square$$

Next we show that starting from the interior point (x, s, Y, Z, V), constructed in Theorem 4.1 for Ω_{k+1}, it takes at most 4 Newton steps to find an η-approximate center of Ω_{k+1} for sufficiently small η.

Theorem 4.4 *Let $\eta = 1/15$. Suppose $(x^k, s^k, Y_k, Z_k, V_k)$ is an η-approximate center of Ω_k satisfying $\eta(\bar{x}^k, s^k, Y_k, \bar{Z}_k, \bar{V}_k) \leq \eta$. Let (x, s, Y, Z, V) be the point constructed from it as in Theorem 4.1. Then starting from the point (x, s, Y, Z, V), the dual Newton procedure described in Section 3 takes at most 4 iterations to find an η-approximate center $Y_{k+1} \in \Omega_{k+1}$, and it satisfies*

$$\eta(\bar{x}^{k+1}, s^{k+1}, Y_{k+1}, \bar{Z}_{k+1}, \bar{V}_{k+1}) \ \leq \ \eta, \tag{4.9}$$

where $(\bar{x}^{k+1}, \bar{s}^{k+1}, \bar{Y}_{k+1}, \bar{Z}_{k+1}, \bar{V}_{k+1})$ is the analytic center of Ω_{k+1}.

Proof. By Theorem 4.1, the point (x, s, Y, Z, V) satisfies $\eta(x, s, Y, Z, V) \leq \gamma$, where (x, Z, V) is feasible for the problem (3.2) and $\gamma = \eta + (1+\eta)/\sqrt{2} < 1$. By Definition 3.2, $\delta(Y) \leq \gamma$. By Theorem 3.5, a point Y_{k+1} with $\delta(Y_{k+1}) \leq \gamma^{2^4} = \gamma^{16} < \eta$ can be found in at most 4 Newton iterations. Then by Theorem 3.6, the point Y_{k+1} satisfies the inequality (4.9). $\square$

Remark: The theorem and its proof will remain valid if $1/200 \leq \eta \leq 1/15$.

5 Potential increment and complexity

Suppose $\Omega = \{Y \in \Omega_0 \ : \ \mathcal{A}^T \mathbf{svec} Y \leq c\}$ is the current working set. Let $(\bar{x}, \bar{s}, \bar{Y}, \bar{Z}, \bar{V})$ be the analytic center of Ω. Given an approximate analytic center $\hat{Y}$ of Ω with

$$\eta(\bar{x}, \hat{s}, \hat{Y}, \bar{Z}, \bar{V}) \ \leq \ \eta. \tag{5.1}$$

If $\hat{Y}$ is not inside Γ, the oracle will generate a cut with normal A ($\|A\|_F = 1$), and this will lead to a new working set,

$$\Omega_+ = \left\{ Y \in \Omega \ : \ (\mathbf{svec} A)^T \mathbf{svec} Y \ \leq \ (\mathbf{svec} A)^T \mathbf{svec} \hat{Y} \right\}. \tag{5.2}$$

Let $\phi(\cdot)$ and $\phi_+(\cdot)$ be the potential functions associated with Ω and Ω_+, respectively. Then the minimum potential values associated with these two sets satisfy the inequality given in the next lemma. It will be used in the complexity analysis of our analytic center cutting plane algorithm.

Lemma 5.1 *Let $(\bar{x}, \bar{s}, \bar{Y}, \bar{Z}, \bar{V})$ be the analytic center of Ω and*

$$\bar{r} = \sqrt{(\mathbf{svec}A)^T \bar{H}^{-1} (\mathbf{svec}A)},$$

where $\bar{H} = \nabla^2 \phi(\bar{Y})$. Then

$$\phi_+(\Omega_+) \;\geq\; \phi(\Omega) - \ln \bar{r} + \alpha,$$

where α is a constant depending only on η, and $\alpha > 0$ if $\eta = 1/15$ is selected.

Remark: For the complexity analysis, $\alpha > 0$ is not necessary. The complexity bounds remain valid as long as α is a universal constant independent of m and k.

Proof. Let $\bar{Y}_+$ be the analytic center of Ω_+, and

$$\bar{s}^+ = c - \mathcal{A}^T \mathbf{svec}(\bar{Y}_+), \quad \bar{s}^+_{k+1} = A \bullet \hat{Y} - A \bullet \bar{Y}_+, \quad \hat{s} = c - \mathcal{A}^T \mathbf{svec}(\hat{Y}).$$

Let $\bar{U} = I - \bar{Y}$. Note that $\bar{H} = \bar{G}\,\bar{G}^T$, where

$$\bar{G} \;=\; \left(\mathcal{A}\bar{S}^{-1}, \;\; -\bar{Y}^{-1/2} \circledast \bar{Y}^{-1/2}, \;\; \bar{U}^{-1/2} \circledast \bar{U}^{-1/2} \right).$$

We have

$$\begin{aligned}
\bar{s}^+_{k+1} &= (\mathbf{svec}A)^T (\mathbf{svec}\hat{Y} - \mathbf{svec}\bar{Y}_+) \\
&= (\mathbf{svec}A)^T \bar{H}^{-1} \bar{G} \left(\bar{G}^T \mathbf{svec}\hat{Y} - \bar{G}^T \mathbf{svec}\bar{Y}_+ \right) \\
&= (\bar{G}^T \bar{H}^{-1} \mathbf{svec}A)^T \left(\bar{G}^T \mathbf{svec}(\hat{Y} - \bar{Y}) - \bar{G}^T \mathbf{svec}(\bar{Y}_+ - \bar{Y}) \right) \\
&\leq \left\| \bar{G}^T \bar{H}^{-1} \mathbf{svec}A \right\| \left(\left\| \bar{G}^T \mathbf{svec}(\hat{Y} - \bar{Y}) \right\| + \left\| \bar{G}^T \mathbf{svec}(\bar{Y}_+ - \bar{Y}) \right\| \right) \\
&= \bar{r} \left[\eta(\bar{x}, \hat{s}, \hat{Y}, \bar{Z}, \bar{V}) + \eta(\bar{x}, \bar{s}^+, \bar{Y}_+, \bar{Z}, \bar{V}) \right], \qquad (5.3)
\end{aligned}$$

where we used Lemma 3.4 in (5.3). Thus, we have

$$\bar{s}^+_{k+1} \;\leq\; \bar{r}\,\rho,$$

where

$$\rho \;=\; \eta + \left\| \begin{pmatrix} e - \bar{S}^{-1} \bar{s}^+ \\ e - \lambda(\bar{Y}^{-1/2} \bar{Y}_+ \bar{Y}^{-1/2}) \\ e - \lambda(\bar{U}^{-1/2} \bar{U}_+ \bar{U}^{-1/2}) \end{pmatrix} \right\|.$$

Note that

$$\bar Z \bullet \bar Y_+ \;=\; \mathrm{tr}(\bar Z \bar Y^+) \;=\; \mathrm{tr}(\bar Y^{-1/2}\bar Y^+ \bar Y^{-1/2}) \;=\; e^T\lambda(\bar Y^{-1/2}\bar Y_+ \bar Y^{-1/2}),$$
$$\bar Z \bullet \bar Y \;=\; \mathrm{tr}(\bar Z \bar Y) = \mathrm{tr}(I) \;=\; m,$$

and the analogues hold for $\bar V$ and $\bar U$. Using (2.2), we have

$$\begin{aligned}
& e^T \bar S^{-1}\bar s^+ + e^T\lambda(\bar Y^{-1/2}\bar Y_+\bar Y^{-1/2}) + e^T\lambda(\bar U^{-1/2}\bar U_+\bar U^{-1/2}) \\
&= \bar x^T(c - \mathcal{A}^T \mathbf{svec}\bar Y_+) + \bar Z \bullet \bar Y_+ + \bar V \bullet \bar U_+ \\
&= \bar x^T c + \bar V \bullet I \\
&= \bar x^T(c - \mathcal{A}^T \mathbf{svec}\bar Y) + \bar Z \bullet \bar Y + \bar V \bullet \bar U \\
&= \bar x^T \bar s + \bar Z \bullet \bar Y + \bar V \bullet \bar U \\
&= n + 2m.
\end{aligned}$$

As in the proof of Theorem 2 of [10], there exists a constant $\alpha > 0$ (we select $\eta = 1/15$) such that

$$\rho \prod_{i=1}^{k} \frac{\bar s_i^+}{\bar s_i} \prod_{i=1}^{m} \left[\lambda_i(\bar Y^{-1/2}\bar Y_+\bar Y^{-1/2})\lambda_i(\bar U^{-1/2}\bar U_+\bar U^{-1/2})\right] \;\le\; e^{-\alpha}.$$

From the results above we obtain

$$\begin{aligned}
\phi_+(\Omega_+) - \phi(\Omega) &= -\ln\left[\bar s_{k+1}^+ \prod_{i=1}^{k} \frac{\bar s_i^+}{\bar s_i}\,\frac{\det \bar Y_+}{\det \bar Y}\,\frac{\det \bar U_+}{\det \bar U}\right] \\
&\ge -\ln \bar r \;-\; \ln\left[\rho \prod_{i=1}^{k} \frac{\bar s_i^+}{\bar s_i} \prod_{i=1}^{m}\Big(\lambda_i(\bar Y^{-1/2}\bar Y_+\bar Y^{-1/2})\lambda_i(\bar U^{-1/2}\bar U_+\bar U^{-1/2})\Big)\right] \\
&\ge -\ln \bar r + \alpha.
\end{aligned}$$

The lemma is proved. □

Let $\bar H_i = \nabla^2 \phi_i(\bar Y_i)$ and

$$\bar r_i \;=\; \left[(\mathbf{svec}A_{i+1})^T \bar H_i^{-1}(\mathbf{svec}A_{i+1})\right]^{1/2}, \quad i = 0,\ldots,k. \tag{5.4}$$

The complexity analysis of our analytic center cutting plane algorithm is based on the following idea. For the sequence of working set Ω_k, we can establish upper and lower bounds on $\phi_k(\Omega_k)$. The upper bound is approximately $k\ln\epsilon^{-1}$, which is due to the assumption that Γ contains a ball of radius ϵ and the observation that Ω_k is defined by k cuts. The lower bound is obtained by estimating $-\sum_{i=0}^{k-1}\ln \bar r_i + k\alpha$, which is a conclusion of Lemma 5.1. An estimation of $\bar r_k$ gives rise to a lower bound proportional to $\frac{k}{2}\ln\frac{k}{m^3}$. Hence the algorithm must terminate before the lower and upper bounds conflict each other. This idea is adapted from the vector case (e.g. [2, 5]). We shall omit some of the proofs that are similar to the vector case.

We first establish an upper bound on $\phi_k(\Omega_k)$.

Lemma 5.2 *Let $\Omega_k \supset \Gamma$ be defined by k linear inequalities and the positive semidefinite constraint. Then*

$$\phi_k(\Omega_k) \le -(k+2m)\ln\epsilon.$$

Proof. Assumptions A1-A3 imply that there exists a point $Y^\epsilon \in \Gamma$, such that

(i) All eigenvalues of Y^ϵ and $I - Y^\epsilon$ are greater than or equal to ϵ;

(ii) For any $A \in \mathcal{S}^m$ with $\|A\|_F = 1$ and $\alpha \in I\!R$, if $\Gamma \subset \{Y \;:\; A \bullet Y \le \alpha\}$, then $\alpha - A \bullet Y^\epsilon \ge \epsilon$.

Since $\Gamma \subset \Omega_k$,

$$\phi_k(\Omega_k) \;\le\; \phi_k(Y^\epsilon) \;=\; -\textstyle\sum_{i=1}^k \ln(c_i - A_i \bullet Y^\epsilon) - \ln(\det Y^\epsilon) - \ln(\det(I - Y^\epsilon)).$$

Noting that $\|A_i\|_F = 1$, $c_i - A_i \bullet Y^\epsilon \ge \epsilon$, and

$$\det Y^\epsilon \;=\; \prod_{i=1}^m \lambda_i(Y^\epsilon) \ge \epsilon^m, \qquad \det(I - Y^\epsilon) \;=\; \prod_{i=1}^m \lambda_i(I - Y^\epsilon) \ge \epsilon^m,$$

we have the desired inequality. □

Now we turn to finding a lower bound for $\phi_k(\Omega_k)$. By Lemma 5.1, this reduces to finding an upper bound for $\bar{r}_i$ for each i. Following the idea introduced by Nesterov for the case of $I\!R^m$, we can establish the following lemma whose proof we shall omit as it is similar to the one appeared in [7, Theorem 3.1].

Lemma 5.3

$$\sum_{i=0}^{k-1} \bar{r}_i^2 \;\le\; m^2(m+1)\ln\left(1 + \frac{k}{4m^2(m+1)}\right). \tag{5.5}$$

By using Lemmas 5.2 and 5.3, we can now derive an expression for which the number k of oracle calls must satisfy. The precise statement is given in the next lemma.

Lemma 5.4 *The analytic center cutting plane method must terminate with a feasible solution before the number k of oracle calls violates the following inequality*

$$\epsilon^2 \;\le\; \frac{\frac{m}{2} + m^2(m+1)\ln\left(1 + \frac{k}{4m^2(m+1)}\right)}{2m+k}\exp\left(\frac{2\alpha k}{2m+k}\right). \tag{5.6}$$

Proof. By Lemmas 5.2 and 5.1, we have

$$\begin{aligned} -(2m+k)\ln\epsilon \;&\ge\; \phi_k(\Omega_k) \\ &\ge\; \phi_0(\Omega_0) - \frac{1}{2}\sum_{i=0}^{k-1}\ln\bar{r}_i^2 + k\alpha \\ &=\; -2m\ln\frac{1}{2} - \frac{1}{2}\sum_{i=0}^{k-1}\ln\bar{r}_i^2 + k\alpha. \end{aligned}$$

Thus

$$\begin{aligned}
\ln\epsilon - \frac{k}{2m+k}\alpha &\leq \frac{1}{2(2m+k)}\left(2m\ln\frac{1}{4} + \sum_{i=0}^{k-1}\ln\bar{r}_i^2\right) \\
&\leq \frac{1}{2}\ln\left(\frac{2m(1/4)+\sum_{i=0}^{k-1}\bar{r}_i^2}{2m+k}\right) \qquad \text{(using the concavity of ln)} \\
&\leq \frac{1}{2}\ln\left(\frac{m/2+m^2(m+1)\ln(1+\frac{k}{4m^2(m+1)})}{2m+k}\right) \qquad \text{(using Lemma 5.3)}
\end{aligned}$$

or

$$\epsilon^2 \leq \frac{\frac{m}{2}+m^2(m+1)\ln(1+\frac{k}{4m^2(m+1)})}{2m+k}\exp\left(\frac{2\alpha k}{2m+k}\right).$$

The algorithm must terminate before the above inequality is violated. □

The complexity analysis of our analytic center cutting plane algorithm is completed with the following theorem.

Theorem 5.5 *The analytic center cutting plane method terminates in at most $O^*(m^3/\epsilon^2)$ Newton steps, where the notation O^* means that lower order terms are ignored.*

Proof. Ignoring lower order terms (assuming $k \gg m$), Lemma 5.4 implies that the algorithm stops as soon as k satisfies

$$\frac{k}{\ln(k/m^3)} \geq O\left(\frac{m^3}{\epsilon^2}\right).$$

For large k, $\ln k$ is negligible compared with k, hence the algorithm requires at most

$$k = O^*\left(\frac{m^3}{\epsilon^2}\right)$$

iterations. The theorem follows by noting that for the prescribed values of $\eta = 1/15$ and $\beta = 1/\sqrt{2}$, the number of Newton steps per iteration is at most 4 as estimated in Section 4. □

Conclusions

We have analyzed an analytical center cutting plane method for semidefinite feasibility problems. The iteration complexity of this method is $O^*\left(m^3/\epsilon^2\right)$, where at most 4 Newton equations of size $m(m+1)/2$ are solved in each iteration. The method allows the constrained set to be defined in very general terms. The dynamical feature of adding cuts looks particularly attractive for large-scale problems. Possible future research may include a multiple cut version of this approach which would be more efficient in practice and computational studies on semidefinite feasibility problems arising from operations research and other areas such as financial engineering and computational geometry.

References

[1] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, Philadelphia, 1994.

[2] J.-L. Goffin, Z. Luo, and Y. Ye, "Complexity analysis of an interior cutting plane method for convex feasibility problems", *SIAM Journal on Optimization*, 6 (1996) 638–652.

[3] J.-L. Goffin and J.-P. Vial, "Convex nondifferentiable optimization: a survey focused on the analytic center cutting plane method", *Logilab Technical Report,* Department of Management Studies, University of Geneva, Switzerland, 1999.

[4] J.-L. Goffin and J-P. Vial, "Shallow, deep and very deep cuts in the analytic center cutting plane method", *Mathematical Programming,* 84 (1999) 89–103.

[5] Z. Luo and J. Sun, "A polynomial cutting surfaces algorithm for the convex feasibility problem defined by self-concordant inequalities", *Computational Optimization and Applications*, 15 (2000) 167–191.

[6] J.E. Mitchell, "An interior point column generation method for linear programming using shifted barriers", *SIAM Journal On Optimization*, 4 (1994) 423–440.

[7] Y. Nesterov, "Cutting plane algorithms from analytic centers: efficiency estimates", *Mathematical Programming*, 69 (1995) 149–176.

[8] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, Philadelphia, USA, 1994.

[9] L. Vandenberghe and S. Boyd, "Semidefinite programming", *SIAM Review*, 38 (1996) 49–95.

[10] Y. Ye, "A potential reduction algorithm allowing column generation", *SIAM Journal on Optimization*, 2 (1992) 7–20.