SIAM J. OPTIM.
Vol. 12, No. 4, pp. 1126–1146

A MULTIPLE-CUT ANALYTIC CENTER CUTTING PLANE METHOD FOR SEMIDEFINITE FEASIBILITY PROBLEMS*

KIM-CHUAN TOH†, GONGYUN ZHAO†, AND JIE SUN‡

Abstract. We consider the problem of finding a point in a nonempty bounded convex body Γ in the cone of symmetric positive semidefinite matrices $\mathcal{S}^m_+$. Assume that Γ is defined by a separating oracle, which, for any given $m \times m$ symmetric matrix $\hat{Y}$, either confirms that $\hat{Y} \in \Gamma$ or returns several selected cuts, i.e., a number of symmetric matrices A_i, $i = 1, \dots, p$, $p \le p_{\max}$, such that Γ is in the polyhedron $\{Y \in \mathcal{S}^m_+ \mid A_i \bullet Y \le A_i \bullet \hat{Y}, i = 1, \dots, p\}$. We present a multiple-cut analytic center cutting plane algorithm. Starting from a trivial initial point, the algorithm generates a sequence of positive definite matrices which are approximate analytic centers of a shrinking polytope in $\mathcal{S}^m_+$. The algorithm terminates with a point in Γ within $O(m^3 p_{\max}/\epsilon^2)$ Newton steps (to leading order), where ϵ is the maximum radius of a ball contained in Γ.

Key words. analytic center, cutting plane methods, multiple cuts, semidefinite programming

AMS subject classifications. 65K05, 90C25

PII. S1052623400370503

1. Introduction.

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 subset Γ of $\mathcal{S}^m_+$. We assume that Γ contains a full-dimensional closed ball with radius $\epsilon > 0$. The set Γ is implicitly defined by a separating oracle, which, for any given $m \times m$ symmetric matrix $\hat{Y}$, either confirms that $\hat{Y} \in \Gamma$ or returns several cuts, i.e., a number of symmetric matrices A_i, $i = 1, \dots, p$, $p \le p_{\max}$, such that Γ is in the polyhedron $\{Y \in \mathcal{S}^m_+ \mid A_i \bullet Y \le A_i \bullet \hat{Y}, i = 1, \dots, p\}$. Here $p_{\max}$ is the maximum number of cuts admitted in each iteration.

In a recent paper [8], we presented an analytic center cutting plane method for the case $p_{\max} = 1$, in which a single cut is added in each iteration. The method was shown to have a worst-case complexity of $O(m^3/\epsilon^2)$ (to leading order). However, to make a cutting plane algorithm practically efficient, adding multiple cuts is often necessary. The purpose of this paper is to propose and analyze an analytic cutting plane method that uses multiple cuts for solving the convex semidefinite feasibility problem mentioned above. In admitting multiple cuts in an analytic center cutting plane method, we face some new theoretical problems that are different from the single-cut situation; these include (a) the problem of finding a feasible starting point for the Newton iteration after several new cuts have been added, (b) the estimation of the number of Newton steps needed to obtain a new approximate center through estimating the changes in the primal-dual potential function.

Our paper extends the multiple-cut schemes of Goffin and Vial [2], Luo [5], and Ye [10] from $\mathbb{R}^m_+$ to $\mathcal{S}^m_+$. Such extensions not only broaden the applications of cutting

*Received by the editors April 4, 2000; accepted for publication (in revised form) October 29, 2001; published electronically April 26, 2002.

http://www.siam.org/journals/siopt/12-4/37050.html

†Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119620 (mattohkc@math.nus.edu.sg, matzgy@math.nus.edu.sg). The research of the first author was supported in part by the Singapore-MIT Alliance.

‡Department of Decision Sciences, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119620 (jsun@nus.edu.sg). The research of this author was supported by National University of Singapore (NUS) Academic Research Grant R314-000-009/026-112 and the Singapore-MIT Alliance.

plane methods but also extend several classical theoretical results for nonnegative vectors to positive semidefinite matrices. We note that for our multiple-cut analytic center cutting plane algorithm, the complexity analysis on the number of Newton iterations per oracle call follows the approach in [3]. For the complexity analysis on the number of oracle calls, we follow the approach in [10], but we simplify the proofs of some results analogous to those in [10] by considering all the added cuts simultaneously instead of inductively.

In this paper we will show that, starting from a trivial initial point, the multiple-cut algorithm generates a sequence of positive definite matrices which are approximate analytic centers of a shrinking polytope in $\mathcal{S}_+^m$. The algorithm will stop with a solution in at most $O(m^3 p_{\max}/\epsilon^2)$ (to leading order) Newton steps. Our analysis shows that when the problem is specialized to the space of positive semidefinite diagonal matrices (which is equivalent to the nonnegative orthant R_+^m), the complexity bound is reduced to $O(m^2 p_{\max}/\epsilon^2)$. This complexity bound is lower than the existing bound of $O(m^2 p_{\max}^2/\epsilon^2)$ obtained in [2] and [10], where the same cuts are considered. Our bound appears to be better than that obtained in [5]. (Note that the proof for the bound appearing in [5] is incomplete, and, to the best of our knowledge, a provable bound should be $O(m^2 p_{\max}^2/\epsilon^2)$.) Furthermore, the analysis in [5] is carried out only for the so-called shallow cuts, which are placed at some distance away from the current testing point and hence may not be as efficient as our proposed algorithm, where the cuts pass through the testing point.

We are able to obtain better complexity results than existing ones even when the problem is specialized to $\mathbb{R}_+^m$, because in each oracle call we admit only cuts that are sufficiently good. We shall not give the precise definition of "goodness" here but refer the reader to section 4. Roughly speaking, based on our criteria, the admitted cuts $A_i,\ i = 1, \dots, p$, in each oracle call are effective in reducing the size of the polytope in the sense that each should be able to delete a sizable portion of the current polytope that cannot be otherwise deleted by the other admitted cuts. One obvious advantage of having such a selection criterion is that the number of cuts added in each iteration is reduced, since only effective cuts are admitted, and this translates into savings in the computational cost in each Newton step.

We will now introduce some notations. For matrices $A, Y \in \mathcal{S}^m$, we define

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

where "T" stands for the transpose, and "Tr" denotes the trace. 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 $\|x\|$, and the matrix 2-norm of a matrix A is denoted by $\|A\|$. For $A \in \mathcal{S}^m$, we write

$$\|A\|_F := (A \bullet A)^{1/2}, \quad \lambda(A) := (\lambda_1(A), \dots, \lambda_m(A))^T,$$

where $\lambda_1(A), \dots, \lambda_m(A)$ are the eigenvalues of A. Note that $\|A\|_F = \|\lambda(A)\|$ and $\|A\| = \|\lambda(A)\|_\infty$. We will use these facts in the paper without explicitly mentioning them. For a positive vector $x \in \mathbb{R}^n$, we write

$$\ln x := \sum_{i=1}^{n} \ln x_i.$$

We use diag(x) to denote the diagonal matrix whose diagonal is the vector x. For a positive vector x, we will use x^{-1} to denote the vector obtained from x by inverting all of its components.

Generally, we use capital letters for matrices, lower case letters for vectors, and Greek letters for scalars. For convenience, we let $\bar{m} = m(m+1)/2$.

Let **svec** be an isometry identifying $\mathcal{S}^m$ with $\mathbb{R}^{\bar{m}}$, so that $K \bullet L = \mathbf{svec}(K)^T \mathbf{svec}(L)$, and let **smat** be the inverse of **svec**. Given any $G \in \mathcal{S}^m$, we let $G \circledast G \in \mathbb{R}^{\bar{m} \times \bar{m}}$ be the unique symmetric matrix such that

$$(G \circledast G)\, \mathbf{svec}(M) \;=\; \mathbf{svec}(GMG) \quad \forall\; M \in \mathcal{S}^m.$$

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

Throughout, we make the following assumptions:

A1. Γ is a convex subset of $\mathcal{S}^m_+$.

A2. $\Gamma \subset \Omega_0$, where $\Omega_0 := \{Y \in \mathcal{S}^m \mid 0 \preceq Y \preceq I\}$.

A3. Γ contains a full-dimensional ball of radius $\epsilon > 0$. That is, there exists $Y^c \in \mathcal{S}^m$ such that $\{Y \in \mathcal{S}^m : \|Y - Y^c\|_F \le \epsilon\} \subset \Gamma$.

Note that Assumption A2 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 $Y \in \hat{\Gamma}$, $\|Y\| \le \gamma$. Then the scaled set $\Gamma = \{Y/\gamma \mid Y \in \hat{\Gamma}\}$ satisfies A2.

The organization of the paper is as follows. In section 2, we describe our multiple-cut analytic center cutting plane algorithm for semidefinite feasibility problems. Section 3 is devoted to the analysis of the computation of an approximate analytic center for a working set. In particular, we establish the number of Newton steps required to compute an approximate analytic center in terms of the number of cuts added. In section 4, we establish the dual potential increment when the current working set is changed to the next working set. Subsequently, we establish complexity results for our multiple-cut cutting plane algorithm.

2. A multiple-cut analytic center cutting plane method. We first define the analytic center and then propose a multiple-cut analytic center cutting plane method at the end of this section.

Let $A_i \bullet Y \le c_i$, $i = 1, \ldots, n_k$, be all the cuts defining the kth working set Ω_k. Define

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

Then the set Ω_k can be represented as

$$\Omega_k = \{Y \in \Omega_0 \mid \mathcal{A}^T \mathbf{svec} Y \le c\}.$$

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

$$\phi_k(Y) \;=\; -\sum_{i=1}^{n_k} \ln(c_i - A_i \bullet Y) - \ln(\det Y) - \ln(\det(I - Y)),$$

where "det" denotes the determinant. We let

$$\phi_k(\Omega) := \min\{\phi_k(Y) \mid Y \in \Omega\}.$$

The unique minimizer of $\phi_k(Y)$ over Ω_k is known as the analytic center of Ω_k.

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

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

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

It is known [7, Proposition 5.4.5] that ϕ_k is a strongly 1-self-concordant function on Ω and

$$\nabla\phi_k(Y) = \mathbf{svec}\left(\sum_{i=1}^{n_k}\frac{A_i}{c_i - A_i \bullet Y} - Y^{-1} + (I-Y)^{-1}\right),$$

$$\nabla^2\phi_k(Y) \;=\; \mathcal{A}S^{-2}\mathcal{A}^T + Y^{-1} \circledast Y^{-1} + (I-Y)^{-1} \circledast (I-Y)^{-1},$$

where $S = \mathrm{diag}\,(s)$ and $s = c - \mathcal{A}^T\mathbf{svec}(Y) > 0$. Strictly speaking, $\nabla\phi_k(Y)$ should be the $m \times m$ matrix within the round brackets. However, we have identified the $m \times m$ matrix with a vector in $\mathbb{R}^{\bar{m}}$ through the linear isometry **svec**. Similarly, $\nabla^2\phi_k(Y)$ is identified with an $\mathbb{R}^{\bar{m}} \times \mathbb{R}^{\bar{m}}$ matrix.

The optimality conditions for minimizing ϕ_k are

$$\begin{aligned} Sx &= e, \quad (e \text{ denotes the vector of ones})\\ YZ &= I,\\ (I-Y)V &= I,\\ \mathcal{A}^T\mathbf{svec}Y + s &= c,\\ \mathcal{A}x - \mathbf{svec}Z + \mathbf{svec}V &= 0,\\ I \succ Y \succ 0, \quad Z, V \succ 0, &\quad s, x > 0. \end{aligned} \tag{2.1}$$

With a slight abuse of language, we also call the solution $(\bar{Y}, \bar{s}, \bar{x}, \bar{Z}, \bar{V})$ of (2.1) the analytic center of Ω_k.

DEFINITION 2.1. *Given a point $(Y, s, x, Z, V) \in \mathcal{S}^m \times \mathbb{R}^{n_k} \times \mathbb{R}^{n_k} \times \mathcal{S}^m \times \mathcal{S}^m$, with $0 \prec Y \prec I$, we define*

$$\eta(Y, s, x, Z, V) \;=\; \sqrt{\|Sx - e\|^2 + \|\lambda(YZ) - e\|^2 + \|\lambda((I-Y)V) - e\|^2}. \tag{2.2}$$

We call (Y, s, x, Z, V) an η-approximate (analytic) center of Ω_k if $\eta(Y, s, x, Z, V) \le \eta$, all the linear equalities in (2.1) *are satisfied, and $x, s > 0$, $Z, V \succ 0$. Obviously, a 0-approximate center is exactly the analytic center of Ω.*

DEFINITION 2.2. *Given $Y \in \mathcal{S}^m$ such that $0 \prec Y \prec I$, and $s = c - \mathcal{A}^T\mathbf{svec}(Y) > 0$, we define*

$$\delta_k(Y) = \sqrt{\nabla\phi_k(Y)^T\,[\nabla^2\phi_k(Y)]^{-1}\,\nabla\phi_k(Y)}\,. \tag{2.3}$$

It was shown [8] that the following lemma holds.

LEMMA 2.3. *Given $Y \in \mathcal{S}^m$ such that $0 \prec Y \prec I$, let $s = c - \mathcal{A}^T\mathbf{svec}(Y)$. We have*

$$\begin{aligned} \delta_k(Y) &= \eta(Y, s, x_Y, Z_Y, V_Y) \\ &= \min\{\eta(Y, s, x, Z, V) \,:\, \mathcal{A}\,x - \mathbf{svec}(Z) + \mathbf{svec}(V) = 0,\ x \in \mathbb{R}^k,\ Z, V \in \mathcal{S}^m\}. \end{aligned} \tag{2.4}$$

Remark. Given $Y \in \mathcal{S}^m$ such that $0 \prec Y \prec I$, $s = c - \mathcal{A}^T\mathbf{svec}(Y) > 0$, and $\delta_k(Y) < \eta < 1$, the minimizer (x_Y, Z_Y, V_Y) of (2.4) satisfies $x_Y > 0$, and $Z_Y, V_Y \succ 0$. For such a Y, we will call Y an η-approximate center of Ω_k in the sense that the point (Y, s, x_Y, Z_Y, V_Y) is an η-approximate center.

We will now describe our algorithm.

A multiple-cut analytic center cutting plane algorithm.

Step 0. Select $\eta \in (0, 1 - \sqrt{3}/2)$, and pick $\bar{\delta} \in (\eta, 1)$. 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 the oracle to either confirm that Y_k is a feasible point of Γ or return p_k matrices $A_{n_k+1}, \dots, A_{n_k+p_k} \in \mathcal{S}^m$ with $\|A_{n_k+i}\|_F = 1$. If $Y_k \in \Gamma$, stop; otherwise, construct the new working set

$$\Omega_{k+1} = \{Y \in \Omega_k \,:\, A_{n_k+i} \bullet Y \le A_{n_k+i} \bullet Y_k,\ i = 1, \dots, p_k\}.$$

Step 2. Find a point $\tilde{Y}$ in the interior of Ω_{k+1} (discussed in section 3).

Step 3. (Recentering) Starting with the point $Y = \tilde{Y}$ in Step 2, perform the dual Newton method:

3.1. If $\delta_{k+1}(Y) < \eta$, set $Y_{k+1} = Y$, $k := k + 1$; go to Step 1.

3.2. Otherwise, set

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

where $\bar{\alpha}$ is determined as follows: if $\delta_{k+1}(Y) \ge \bar{\delta}$, $\bar{\alpha} = \frac{1}{1+\delta_{k+1}(Y)}$; else, $\bar{\alpha} = 1$. Set $Y = Y_+$. Go to Step 3.1.

Note that we need the restriction $\eta < 1 - \sqrt{3}/2$ in order to construct the point $\tilde{Y}$ in Step 2.

3. Restoration of centrality. In our cutting plane algorithm, approximate analytic centers are found by using the dual Newton method. Our aim in this section is to estimate the number of Newton steps required to find an approximate analytic center for a newly constructed working set. We do so by estimating the amount of potential value we should reduce for the new set. The mechanics are as follows. Since the potential function is 1-self-concordant, each Newton step can reduce the potential function by a constant amount. Thus to estimate the number of Newton steps needed to find an approximate analytic center for a new working set, all we need is to estimate the amount of potential value we should reduce for the new set.

To find an approximate analytic center for a new working set, we would ideally want the Newton method to start with the preceding approximate analytic center Y_k. However, Y_k is not in the interior of the new working set Ω_{k+1}, since the new cuts pass through this point. Thus our immediate task is to find an interior point in Ω_{k+1} and then use this point as the starting point for the Newton method.

Let n_k be the number of cuts defining the set Ω_k. Suppose that p_k new cuts are added to form the new set Ω_{k+1}. Recall that

$$\begin{aligned} \mathcal{A} &:= (\mathbf{svec}A_1, \mathbf{svec}A_2, \dots, \mathbf{svec}A_{n_k}), \qquad c := (c_1, c_2, \dots, c_{n_k})^T, \\ \mathcal{B}_k &:= (\mathbf{svec}A_{n_k+1}, \mathbf{svec}A_{n_k+2}, \dots, \mathbf{svec}A_{n_k+p_k}), \qquad d := \mathcal{B}_k^T\mathbf{svec}Y_k. \end{aligned}$$

Then the sets Ω_k and Ω_{k+1} can be written as

$$\Omega_k = \{Y \in \Omega_0 \mid \mathcal{A}^T \mathbf{svec} Y \le c\}, \qquad \Omega_{k+1} = \{Y \in \Omega_k \mid \mathcal{B}_k^T \mathbf{svec} Y \le d\}.$$

Let $H_k = \nabla^2 \phi_k(Y_k)$ and

$$M_k := \mathcal{B}_k^T H_k^{-1} \mathcal{B}_k.$$

Suppose $(Y_k, s^k, x^k, Z_k, V_k)$ is an η-approximate center with $\eta < 1 - \sqrt{3}/2$. (Note that, by Lemma 2.3, $\delta_k(Y_k) \le \eta(Y_k, s^k, x^k, Z_k, V_k) \le \eta$.) We will now construct a point $(\tilde{Y}, \tilde{s}, \tilde{x}, \tilde{Z}, \tilde{V})$ that is in the interior of Ω_{k+1}, using a procedure similar to that in Goffin and Vial [2]. To this end, consider the following convex minimization problem:

$$\begin{array}{ll} \min & p_k \omega^T M_k \, \omega \; - \; \ln \omega \\ \text{such that} & \omega = (\omega_1, \dots, \omega_{p_k})^T > 0. \end{array}$$

Evidently, the above problem has a unique solution that is also the unique solution to the KKT-conditions:

$$M_k \, \omega = \xi, \tag{3.1a}$$

$$2 p_k \, \omega_i \, \xi_i = 1, \quad \omega_i, \; \xi_i \; > 0, \quad i = 1, \dots, p_k. \tag{3.1b}$$

Let $(\tilde{\omega}, \tilde{\xi})$ be an approximate solution of the above KKT-conditions, where (3.1a) is satisfied exactly and $\max\{|2 p_k \tilde{\omega}_i \tilde{\xi}_i - 1| : i = 1, \dots, p_k\} \le 1/2$. Note that, in this case,

$$\tilde{\omega}^T M_k \, \tilde{\omega} \; \le \; \frac{3}{4}. \tag{3.2}$$

Note that to find such a pair $(\tilde{\omega}, \tilde{\xi})$, we can apply Newton's method to (3.1a) and (3.1b), where the computational work for each Newton iteration is $O(p_k^3)$. In general, this constitutes only a very small fraction of the total computational work involved in finding an approximate analytic center for Ω_{k+1}. In order not to lengthen the paper unnecessarily, we shall not establish the complexity of the Newton method for finding $(\tilde{\omega}, \tilde{\xi})$ in this paper. Interested readers can refer to [3] for such results.

Let $U_k = I - Y_k$ and

$$\Delta Y \; = \; -\mathbf{smat}(H_k^{-1} \mathcal{B}_k \tilde{\omega}), \quad \Delta s = -\mathcal{A}^T \mathbf{svec} \Delta Y, \tag{3.3}$$

$$\Delta x = S_k^{-2} \mathcal{A}^T \mathbf{svec} \Delta Y, \quad \Delta Z = -Y_k^{-1} (\Delta Y) Y_k^{-1}, \quad \Delta V = U_k^{-1} (\Delta Y) U_k^{-1}. \tag{3.4}$$

Define

$$\tilde{Y} \; = \; Y_k + \Delta Y, \qquad \tilde{s} \; = \; \begin{pmatrix} s^k + \Delta s \\ \tilde{\xi} \end{pmatrix}, \tag{3.5}$$

$$\tilde{x} = \begin{pmatrix} x^k + \Delta x \\ \tilde{\omega} \end{pmatrix}, \quad \tilde{Z} = Z_k + \Delta Z, \quad \tilde{V} = V_k + \Delta V. \tag{3.6}$$

We refer the reader to [3] for an illuminating discussion on the motivation for considering the optimization problem (3.1a)–(3.1b) in constructing the strictly interior point of Ω_{k+1} above.

It is readily shown that the following result holds:

$$\|S_k^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y)\|^2 + \|Y_k^{-1/2}(\Delta Y)Y_k^{-1/2}\|_F^2 + \|U_k^{-1/2}(\Delta Y)U_k^{-1/2}\|_F^2$$

$$= \mathbf{svec}(\Delta Y)^T H_k\,\mathbf{svec}(\Delta Y) = \tilde{\omega}^T M_k \tilde{\omega} \leq \frac{3}{4}. \tag{3.7}$$

LEMMA 3.1. *For any vector $q = (q_1, \ldots, q_n)^T$ with $\|q\| < 1$, the following inequality holds:*

$$-\ln(e - q) \leq e^T q + \frac{\|q\|^2}{2(1 - \|q\|)}.$$

Proof. For this proof, we refer to [11]. □

LEMMA 3.2. *Suppose $(Y_k, s^k, x^k, Z_k, V_k)$ is an η-approximate center with $\eta < 1$. Then the following inequalities hold:*

$$\|X_k^{-1}\Delta x\| \leq \frac{1}{1-\eta}\|S_k^{-1}\Delta s\|,$$

$$\|Z_k^{-1/2}(\Delta Z)Z_k^{-1/2}\|_F \leq \frac{1}{1-\eta}\|Y_k^{-1/2}(\Delta Y)Y_k^{-1/2}\|_F,$$

$$\|V_k^{-1/2}(\Delta V)V_k^{-1/2}\|_F \leq \frac{1}{1-\eta}\|U_k^{-1/2}(\Delta Y)U_k^{-1/2}\|_F,$$

where $X_k = \mathrm{diag}(x^k)$.

Proof. We shall omit the proof of the first inequality, as it is easy. Now we proceed with the proof of the second one. We have

$$\begin{aligned}
\|Z_k^{-1/2}(\Delta Z)Z_k^{-1/2}\|_F^2 &= \sum_{i=1}^m \lambda_i\left(Z_k^{-1/2}Y_k^{-1/2}(Y_k^{-1/2}\Delta Y Y_k^{-1/2})Y_k^{-1/2}Z_k^{-1/2}\right)^2 \\
&= \sum_{i=1}^m \theta_i^2\,\lambda_i\left(Y_k^{-1/2}\Delta Y Y_k^{-1/2}\right)^2 \\
&\leq \left(\max_{1\leq i\leq m}\theta_i^2\right)\,\|Y_k^{-1/2}\Delta Y Y_k^{-1/2}\|_F^2,
\end{aligned}$$

where we have used a theorem of Ostrowski [4, p. 225] in the second equality above, and the θ_i's are scalars such that

$$\lambda_{\min}(Z_k^{-1/2}Y_k^{-1}Z_k^{-1/2}) \leq \theta_i \leq \lambda_{\max}(Z_k^{-1/2}Y_k^{-1}Z_k^{-1/2}).$$

Noting that $\lambda_{\max}(Z_k^{-1/2}Y_k^{-1}Z_k^{-1/2}) \leq 1/(1-\eta)$, we have proved the required inequality. The last inequality in the lemma can be proved similarly. □

THEOREM 3.3. *Suppose $(Y_k, s^k, x^k, Z_k, V_k)$ is an η-approximate center with $\eta < 1 - \sqrt{3}/2$. Then the point $(\tilde{Y}, \tilde{s}, \tilde{x}, \tilde{Z}, \tilde{V})$ constructed in* (3.5)–(3.6) *satisfies the last three conditions in* (2.1).

Proof. First, we show that $\tilde{s} > 0$ and $0 \prec \tilde{Y} \prec I$. We have

$$s^k + \Delta s = s^k - \mathcal{A}^T\mathbf{svec}(\Delta Y) = S_k\left[e - S_k^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y)\right] > 0,$$

since $\|S_k^{-1}\mathcal{A}^T\mathbf{svec}(\Delta Y)\| \le \sqrt{3}/2 < 1$ from (3.7). On the other hand, we also have

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

since $\|Y_k^{-1/2}\Delta Y Y_k^{-1/2}\|_F \le \sqrt{3}/2 < 1$. That $\tilde{Y} \prec I$ can be shown similarly. Furthermore,

$$\begin{bmatrix} \mathcal{A}^T\mathbf{svec}Y \\ \mathcal{B}_k^T\mathbf{svec}Y \end{bmatrix} + s \;=\; \begin{bmatrix} c \\ d + \mathcal{B}_k^T\mathbf{svec}(\Delta Y) + \tilde{\xi} \end{bmatrix} \;=\; \begin{bmatrix} c \\ d \end{bmatrix},$$

where we used the fact that, from (3.1a), $\mathcal{B}_k^T\mathbf{svec}(\Delta Y) = -M_k\tilde{\omega} = -\tilde{\xi}$.

Next we show that $\tilde{x} > 0$ and $\tilde{Z}, \tilde{V} \succ 0$. We have

$$\tilde{Z} \;=\; Z_k^{1/2}(I + Z_k^{-1/2}(\Delta Z)Z_k^{-1/2})Z_k^{1/2} \;\succ\; 0,$$

since, by Lemma 3.2,

$$\|Z_k^{-1/2}(\Delta Z)Z_k^{-1/2}\|_F \;\le\; \frac{1}{(1-\eta)}\|Y_k^{-1/2}(\Delta Y)Y_k^{-1/2}\|_F \;\le\; \frac{\sqrt{3}}{2(1-\eta)} \;<\; 1.$$

Furthermore,

$$\begin{aligned}
&[\mathcal{A}\;\;\mathcal{B}_k]\,\tilde{x} - \mathbf{svec}\tilde{Z} + \mathbf{svec}\tilde{V} \\
&\qquad = \mathcal{A}\,x^k + \mathcal{A}\,\Delta x + \mathcal{B}_k\tilde{\omega} - \mathbf{svec}Z_k - \mathbf{svec}\Delta Z + \mathbf{svec}V_k + \mathbf{svec}\Delta V \\
&\qquad = \mathcal{A}\,\Delta x + \mathcal{B}_k\tilde{\omega} - \mathbf{svec}\Delta Z + \mathbf{svec}\Delta V \\
&\qquad = \mathcal{A}\,S_k^{-2}\mathcal{A}^T\mathbf{svec}(\Delta Y) + Y_k^{-1} \circledast Y_k^{-1}\mathbf{svec}(\Delta Y) + U_k^{-1} \circledast U_k^{-1}\mathbf{svec}(\Delta Y) + \mathcal{B}_k\tilde{\omega} \\
&\qquad = H_k\mathbf{svec}(\Delta Y) + \mathcal{B}_k\tilde{\omega} \;=\; 0. \qquad \square
\end{aligned}$$

Up to this point, we have succeeded in finding in the interior of Ω_{k+1} a point $\tilde{Y}$ that is derived from Y_k. Our next task is to estimate the potential value of the new point in Ω_{k+1}.

LEMMA 3.4. *Suppose $\delta_k(Y_k) \le \eta$. Then the potential value $\phi_{k+1}(\tilde{Y})$ satisfies the following inequality:*

$$\phi_{k+1}(\tilde{Y}) \;\le\; \phi_k(Y_k) + \frac{\sqrt{3}}{2}\eta + \frac{3}{4(2-\sqrt{3})} - \ln\tilde{\xi}. \tag{3.8}$$

Proof. Let $\tilde{U} = I - \tilde{Y}$ and $U_k = I - Y_k$. We have

$$\phi_{k+1}(\tilde{Y}) \;=\; -\ln\tilde{s} - \ln(d - \mathcal{B}_k^T\mathbf{svec}(\tilde{Y})) - \ln(\det\tilde{Y}) - \ln(\det\tilde{U}) \;\;=\;\; \phi_k(\tilde{Y}) - \ln\tilde{\xi}. \tag{3.9}$$

Note that we used the fact that $d - \mathcal{B}_k^T\mathbf{svec}(\tilde{Y}) = -\mathcal{B}_k^T\mathbf{svec}(\Delta Y) = \tilde{\xi}$. Now

$$\begin{aligned}
\phi_k(\tilde{Y}) \;&=\; -\ln(s^k + \Delta s) - \ln\det(Y_k + \Delta Y) - \ln\det(U_k - \Delta Y) \\
&= \phi_k(Y_k) - \ln(e + S_k^{-1}\Delta s) - \ln\det(I + Y_k^{-1/2}\Delta Y\,Y_k^{-1/2}) \\
&\qquad - \ln\det(I - U_k^{-1/2}\Delta Y\,U_k^{-1/2})
\end{aligned}$$

$$
\begin{aligned}
&= \phi_k(Y_k) - \ln(e + S_k^{-1}\Delta s) - \ln\left(e + \lambda(Y_k^{-1/2}(\Delta Y)Y_k^{-1/2})\right) \\
&\qquad - \ln\left(e - \lambda(U_k^{-1/2}(\Delta Y)U_k^{-1/2})\right) \\
&= \phi_k(Y_k) - \ln(e - q),
\end{aligned} \tag{3.10}
$$

where

$$
q \;=\; \begin{pmatrix} -S_k^{-1}\Delta s \\ -\lambda(Y_k^{-1/2}(\Delta Y)Y_k^{-1/2}) \\ \lambda(U_k^{-1/2}(\Delta Y)U_k^{-1/2}) \end{pmatrix}. \tag{3.11}
$$

Note that $e^T q = \nabla\phi_k(Y_k)^T \mathbf{svec}\Delta Y$ and $\|q\|^2 = \mathbf{svec}(\Delta Y)^T H_k\, \mathbf{svec}(\Delta Y) \le 3/4$.

By applying Lemma 3.1 to (3.10), we have

$$
\begin{aligned}
\phi_k(\tilde{Y}) - \phi_k(Y_k) &\le e^T q + \frac{\|q\|^2}{2(1-\|q\|)} \\
&= \nabla\phi_k(Y_k)^T \mathbf{svec}\Delta Y + \frac{\mathbf{svec}(\Delta Y)^T H_k\, \mathbf{svec}(\Delta Y)}{(2-\sqrt{3})} \\
&\le \delta_k(Y_k)\sqrt{\tilde{\omega}^T M_k \tilde{\omega}} \;+\; \frac{\tilde{\omega}^T M_k \tilde{\omega}}{(2-\sqrt{3})} \\
&\le \frac{\sqrt{3}}{2}\eta \;+\; \frac{3}{4(2-\sqrt{3})}.
\end{aligned} \tag{3.12}
$$

Note that in the next to last inequality above, we used the Cauchy inequality to derive the result: $\nabla\phi_k(Y_k)^T \mathbf{svec}(\Delta Y) \le \delta_k(Y_k)\sqrt{\tilde{\omega}^T M_k \tilde{\omega}}$.

Substituting the result in (3.12) into (3.9), we prove the lemma. ∎

From Lemma 3.4, we see that the upper bound for the dual potential value $\phi_{k+1}(\tilde{Y})$ contains the term $-\ln\tilde{\xi}$. If we were to consider the dual potential value alone, then finding an upper bound for $-\ln\tilde{\xi}$ would be necessary. But we have found that finding a tight upper bound for this term is difficult. As a result, we have decided to consider the primal-dual potential value, for which finding an upper bound for $-\ln\tilde{\xi}$ is not necessary. To this end, let us define the primal potential function associated with Ω_k. For any $\psi_k(x, Z, V) \in \mathbb{R}^{n_k}_{++} \times \mathcal{S}^m_{++} \times \mathcal{S}^m_{++}$ that satisfies $\mathcal{A}\,x - \mathbf{svec}(Z) + \mathbf{svec}(V) = 0$, the primal potential of (x, Z, V) is defined by

$$
\psi_k(x, V, Z) \;=\; c^T x + I \bullet V - \ln x - \ln\det Z - \ln\det V. \tag{3.13}
$$

The primal-dual potential function associated with Ω_k is

$$
\Lambda_k(Y, x, Z, V) \;=\; \phi_k(Y) + \psi_k(x, Z, V).
$$

We should emphasize that the primal-dual potential function is introduced solely for the purpose of estimating the potential value of $(\tilde{Y}, \tilde{s}, \tilde{x}, \tilde{Z}, \tilde{V})$. It is not needed in our cutting plane algorithm described in section 2.

Now we shall proceed to establish an analogue of Lemma 3.4 for the primal potential function. Before doing that, we need the following lemma.

LEMMA 3.5. *For the directions* $(\Delta x, \Delta Z, \Delta V)$ *given in* (3.4), *the following inequality holds:*

$$\left|c^T \Delta x - e^T X_k^{-1} \Delta x - Z_k^{-1} \bullet \Delta Z - V_k^{-1} \bullet \Delta V + I \bullet \Delta V + d^T \tilde{\omega}\right| \;\leq\; \frac{\eta}{1-\eta} \frac{\sqrt{3}}{2}. \tag{3.14}$$

Proof. Noting that $d = \mathcal{B}_k^T \mathbf{svec}(Y_k)$ and $H_k \mathbf{svec}(\Delta Y) = -\mathcal{B}_k \tilde{\omega}$, we have

$$d^T \tilde{\omega} = -\mathbf{svec}(Y_k)^T \left(\mathcal{A}\, \Delta x + \mathbf{svec}[Y_k^{-1}(\Delta Y) Y_k^{-1}] + \mathbf{svec}[U_k^{-1}(\Delta Y) U_k^{-1}]\right).$$

Let $X_k = \mathrm{diag}(x^k)$ and $S_k = \mathrm{diag}(s^k)$. Then

$$\begin{aligned}
&\left|c^T \Delta x + d^T \tilde{\omega} - e^T X_k^{-1} \Delta x - Z_k^{-1} \bullet \Delta Z - V_k^{-1} \bullet \Delta V + I \bullet \Delta V\right| \\
&\;= \left|e^T (S_k - X_k^{-1}) \Delta x \,+\, (Z_k^{-1} - Y_k) \bullet (Y_k^{-1} \Delta Y Y_k^{-1}) \,+\, (U_k - V_k^{-1}) \bullet (U_k^{-1} \Delta Y U_k^{-1})\right| \\
&\;= \left|(e - X_k^{-1}(s^k)^{-1})^T S_k^{-1} \Delta s\right| \,+\, \left|(Y_k^{-1/2} Z_k^{-1} Y_k^{-1/2} - I) \bullet (Y_k^{-1/2} \Delta Y Y_k^{-1/2})\right| \\
&\qquad + \left|(I - U_k^{-1/2} V_k^{-1} U_k^{-1/2}) \bullet (U_k^{-1/2} \Delta Y U_k^{-1/2})\right| \\
&\;\leq \|e - X_k^{-1}(s^k)^{-1}\| \, \|S_k^{-1} \Delta s\| \,+\, \|Y_k^{-1/2} Z_k^{-1} Y_k^{-1/2} - I\|_F \, \|Y_k^{-1/2} \Delta Y Y_k^{-1/2}\|_F \\
&\qquad + \|U_k^{-1/2} V_k^{-1} U_k^{-1/2} - I\|_F \, \|U_k^{-1/2} \Delta Y U_k^{-1/2}\|_F \\
&\;\leq \left(\|e - X_k^{-1}(s^k)^{-1}\|^2 \,+\, \|Y_k^{-1/2} Z_k^{-1} Y_k^{-1/2} - I\|_F^2 \,+\, \|U_k^{-1/2} V_k^{-1} U_k^{-1/2} - I\|_F^2\right)^{1/2} \\
&\qquad \times \left(\|S_k^{-1} \Delta s\|^2 \,+\, \|Y_k^{-1/2} \Delta Y Y_k^{-1/2}\|_F^2 \,+\, \|U_k^{-1/2} \Delta Y U_k^{-1/2}\|_F^2\right)^{1/2} \\
&\;\leq \eta\left(Y_k^{-1}, (s^k)^{-1}, (x^k)^{-1}, Z_k^{-1}, V_k^{-1}\right) \left(\mathbf{svec}(\Delta Y)^T H_k \mathbf{svec}(\Delta Y)\right)^{1/2} \\
&\;\leq \frac{\eta}{1-\eta} \frac{\sqrt{3}}{2}.
\end{aligned}$$

Note that, in the last inequality above, we used (3.7) and the fact that

$$\eta(Y^{-1}, (s)^{-1}, (x)^{-1}, Z^{-1}, V^{-1}) \;\leq\; \frac{\eta}{1-\eta}. \qquad \square$$

LEMMA 3.6. *For the point* $(\tilde{x}, \tilde{Z}, \tilde{V})$ *constructed in* (3.6), *the following inequality holds:*

$$\psi_{k+1}(\tilde{x}, \tilde{Z}, \tilde{V}) \leq \psi_k(x^k, Z_k, V_k) + \frac{3}{4(1-\eta)(2-2\eta-\sqrt{3})} + \frac{\eta}{1-\eta} \frac{\sqrt{3}}{2} - \ln \tilde{\omega}. \tag{3.15}$$

Proof. We have

$$\begin{aligned}
\psi_{k+1}(\tilde{x}, \tilde{Z}, \tilde{V}) &= c^T x^k + c^T \Delta x - \ln x^k - \ln(e + X_k^{-1} \Delta x) - \ln \det Z_k \\
&\quad - \ln \det(I + Z_k^{-1/2} (\Delta Z) Z_k^{-1/2}) - \ln \det V_k \\
&\quad - \ln \det(I + V_k^{-1/2} (\Delta V) V_k^{-1/2}) + I \bullet V_k + I \bullet \Delta V + d^T \tilde{\omega} - \ln \tilde{\omega} \\
&= \psi_k(x^k, Z_k, V_k) + c^T \Delta x + I \bullet \Delta V + d^T \tilde{\omega} - \ln \tilde{\omega} - \ln(e + p),
\end{aligned} \tag{3.16}$$

where

$$p = \begin{pmatrix} X_k^{-1}\Delta x \\ \lambda(Z_k^{-1/2}(\Delta Z)Z_k^{-1/2}) \\ \lambda(V_k^{-1/2}(\Delta V)V_k^{-1/2}) \end{pmatrix}.$$

Note that $e^T p = e^T X_k^{-1}\Delta x + Z_k^{-1} \bullet \Delta Z + V_k^{-1} \bullet \Delta V$, and by Lemma 3.2,

$$\begin{aligned} \|p\|^2 &= \|X_k^{-1}\Delta x\|^2 + \|Z_k^{-1/2}(\Delta Z)Z_k^{-1/2}\|_F^2 + \|V_k^{-1/2}(\Delta V)V_k^{-1/2}\|_F^2 \\ &\le \frac{1}{(1-\eta)^2}\left(\|S_k^{-1}\Delta s\|^2 + \|Y_k^{-1/2}(\Delta Y)Y_k^{-1/2}\|_F^2 + \|U_k^{-1/2}(\Delta Y)U_k^{-1/2}\|_F^2\right) \\ &= \frac{1}{(1-\eta)^2}\,\mathbf{svec}(\Delta Y)^T H_k\,\mathbf{svec}(\Delta Y) \;\le\; \frac{1}{(1-\eta)^2}\,\frac{3}{4}. \end{aligned} \tag{3.17}$$

By Lemma 3.1 and (3.17), we get from (3.16),

$$\begin{aligned} \psi_{k+1}(\tilde{x},\tilde{Z},\tilde{V}) \;&\le\; \psi_k(x^k,Z_k,V_k) + c^T\Delta x + I\bullet\Delta V + d^T\tilde{\omega} - \ln\tilde{\omega} - e^T p + \frac{\|p\|^2}{2(1-\|p\|)} \\ &\le \psi_k(x^k,Z_k,V_k) + c^T\Delta x + I\bullet\Delta V + d^T\tilde{\omega} - \ln\tilde{\omega} - e^T p + \frac{3}{4(1-\eta)(2-2\eta-\sqrt{3})}. \end{aligned}$$

By applying Lemma 3.5 and (3.7), we prove the lemma. ∎

The next lemma is an analogue of Lemma 3.4 for the primal-dual potential function.

LEMMA 3.7. *Suppose $\eta(Y_k, s^k, x^k, Z_k, V_k)$ is an η-approximate center with $\eta < 1-\sqrt{3}/2$. Then*

$$\Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V}) \;\le\; \Lambda_k(Y_k,x^k,Z_k,V_k) + \beta(\eta) + p_k\left(\frac{3}{4} + \ln 2p_k\right),$$

where

$$\beta(\eta) \;=\; \eta\frac{\sqrt{3}}{2} + \frac{3}{4(2-\sqrt{3})} + \frac{\eta}{1-\eta}\,\frac{\sqrt{3}}{2} + \frac{3}{4(1-\eta)(2-2\eta-\sqrt{3})}. \tag{3.18}$$

Proof. Combining the results in Lemmas 3.4 and 3.6, we have

$$\Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V}) \;\le\; \Lambda_k(Y_k,x^k,Z_k,V_k) + \beta(\eta) - \ln\tilde{\omega}\tilde{\xi}. \tag{3.19}$$

Note that

$$\begin{aligned} -\ln\tilde{\omega}\tilde{\xi} &= p_k\ln 2p_k \;+\; \sum_{i=1}^{p_k} -\ln\left(1-(1-2p_k\tilde{\omega}_i\tilde{\xi}_i)\right) \\ &\le p_k\ln 2p_k \;+\; \sum_{i=1}^{p_k}\left[(1-2p_k\tilde{\omega}_i\tilde{\xi}_i) + \frac{|1-2p_k\tilde{\omega}_i\tilde{\xi}_i|^2}{2(1-|1-2p_k\tilde{\omega}_i\tilde{\xi}_i|)}\right] \\ &\le p_k\ln 2p_k \;+\; \frac{3}{4}p_k. \end{aligned} \tag{3.20}$$

By substituting (3.20) into (3.19), the lemma is proved. ∎

With Lemma 3.7, we can finally establish an explicitly known upper bound for the primal-dual potential value $\Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V})$.

THEOREM 3.8. *Suppose that* $(\bar{Y}_{k+1},\bar{x}^{k+1},\bar{Z}_{k+1},\bar{V}_{k+1})$ *is the analytic center of* Ω_{k+1}*, and* $(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V})$ *is the point constructed in* (3.5)–(3.6)*. Then*

$$\Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V}) - \Lambda_{k+1}(\bar{Y}_{k+1},\bar{x}^{k+1},\bar{Z}_{k+1},\bar{V}_{k+1}) \ \le\ p_k\left(\ln 2p_k - \frac{1}{4}\right) + \beta(\eta) + \frac{2\eta^2}{1-\eta^2}, \tag{3.21}$$

where $\beta(\eta)$ *is the constant given in* (3.18).

Proof. Suppose that $(Y_k, s^k, x^k, Z_k, V_k)$ is an η-approximate center of Ω_k with $\eta < 1-\sqrt{3}/2$. We have

$$\begin{aligned}
&\Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V}) - \Lambda_{k+1}(\bar{Y}_{k+1},\bar{x}^{k+1},\bar{Z}_{k+1},\bar{V}_{k+1})\\
&= \Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V}) - \Lambda_k(Y_k,x^k,Z_k,V_k) \ +\ \Lambda_k(\bar{Y}_k,\bar{x}^k,\bar{Z}_k,\bar{V}_k)\\
&\quad -\ \Lambda_{k+1}(\bar{Y}_{k+1},\bar{x}^{k+1},\bar{Z}_{k+1},\bar{V}_{k+1})\\
&\quad +\ \Lambda_k(Y_k,x_k,Z_k,V_k) - \Lambda_k(\bar{Y}_k,\bar{x}^k,\bar{Z}_k,\bar{V}_k).
\end{aligned}\tag{3.22}$$

It is readily shown that

$$\begin{aligned}
&\Lambda_k(\bar{Y}_k,\bar{x}^k,\bar{Z}_k,\bar{V}_k) - \Lambda_{k+1}(\bar{Y}_{k+1},\bar{x}^{k+1},\bar{Z}_{k+1},\bar{V}_{k+1})\\
&\quad = (n_k+2m) - (n_k+p_k+2m) \ =\ -p_k.
\end{aligned}\tag{3.23}$$

Next we need to get an upper bound for the term $\Lambda_k(Y_k,x^k,Z_k,V_k)-\Lambda_k(\bar{Y}_k,\bar{x}^k,\bar{Z}_k,\bar{V}_k)$ in (3.22). By following the proof of Lemma 2.1 in [1] and using the quadratic convergence result in [8], it is readily shown that

$$\phi_k(Y_k) \ \le\ \phi_k(\bar{Y}_k) + \frac{\eta^2}{1-\eta^2}. \tag{3.24}$$

Similarly, it can be shown that

$$\psi_k(x^k,Z_k,V_k) \ \le\ \psi_k(\bar{x}^k,\bar{Z}_k,\bar{V}_k) + \frac{\eta^2}{1-\eta^2}. \tag{3.25}$$

Combining (3.24) and (3.25), we get

$$\Lambda_k(Y_k,x^k,Z_k,V_k) - \Lambda_k(\bar{Y}_k,\bar{x}^k,\bar{Z}_k,\bar{V}_k) \ \le\ \frac{2\eta^2}{1-\eta^2}. \tag{3.26}$$

By putting the results of Lemma 3.7, (3.23), and (3.26) into (3.22), the theorem is proved. ∎

With the estimate of $\Lambda_{k+1}(\tilde{Y},\tilde{x},\tilde{Z},\tilde{V})$ in Theorem 3.8, we are now ready to estimate the number of dual Newton steps required to find an approximate analytic center for Ω_{k+1} by using the point $\tilde{Y}$ as the initial point.

THEOREM 3.9. *Given an* η*-approximate center* Y_k *of* Ω_k*, with* $\eta < 1-\sqrt{3}/2$*, the total number of dual Newton steps required to find an* η*-approximate center* Y_{k+1} *of* Ω_{k+1} *is*

$$O\left(p_k \ln p_k\right),$$

where the constant $O(1)$ *is independent of* k.

Proof. By Theorem 2.2.3 in [7], each dual Newton step reduces Λ_{k+1} by a positive constant $\gamma = \bar{\delta} - \ln(1+\bar{\delta})$, as long as a point $\hat{Y}$ with $\delta_{k+1}(\hat{Y}) < \bar{\delta} < 1$ is not yet found, while keeping the primal iterate fixed. Now, starting at $(\tilde{Y}, \tilde{s}, \tilde{x}, \tilde{Z}, \tilde{V})$, the total value of Λ_{k+1} which needs to be reduced is not more than $\Lambda_{k+1}(\tilde{Y}, \tilde{x}, \tilde{Z}, \tilde{V}) - \Lambda_{k+1}(\bar{Y}, \bar{x}, \bar{Z}, \bar{V})$; thus Theorem 3.8 implies that at most

$$\frac{1}{\gamma}\left[p_k\left(\ln p_k + \ln 2 - \frac{1}{4}\right) + \beta(\eta) + \frac{2\eta^2}{1-\eta^2}\right]$$

Newton steps are required to reach a point $\hat{Y}$ with $\delta_{k+1}(\hat{Y}) \le \bar{\delta}$. From $\hat{Y}$ onwards, by Lemma 4.3 in [8], quadratic convergence can be achieved, and thus it needs at most $\ln(\ln(\bar{\delta}/\eta))$ additional full Newton steps to find a point Y_{k+1} satisfying $\delta_{k+1}(Y_{k+1}) \le \eta$. (We can choose, for example, $\bar{\delta} = 0.9$ and $\eta = 0.1$; then $\ln(\ln(\bar{\delta}/\eta)) \le 4$.) ☐

4. Potential changes and complexity. Recall that $\Omega_k = \{Y \in \Omega_0 \mid \mathcal{A}^T \mathbf{svec} Y \le c\}$. Suppose that Y_k is an η-approximate analytic center of Ω_k with $\eta < 1 - \sqrt{3}/2$. Let

$$\mathcal{B}_k \;=\; (\mathbf{svec} A_{n_k+1}, \ldots, \mathbf{svec} A_{n_k+p_k}), \qquad d \;=\; \mathcal{B}_k^T \mathbf{svec}(Y_k).$$

Then

$$\Omega_{k+1} = \{Y \in \Omega_k \mid \mathcal{B}_k^T \mathbf{svec} Y \;\le\; d\}.$$

Let $\bar{Y}_k$ and $\bar{Y}_{k+1}$ be the analytic centers of Ω_k and Ω_{k+1}, respectively. Let

$$\bar{r}_k = \sqrt{\lambda_{\max}(\mathcal{B}_k^T \bar{H}_k^{-1} \mathcal{B}_k)}, \tag{4.1}$$

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

In this section, we estimate the amount that the dual potential will increase when the working set changes from Ω_k to Ω_{k+1}. To this end, we first establish a lemma that is an extension of a result in [10].

LEMMA 4.1. *Suppose that* n, p *are positive integers and* v *is a positive* n*-vector with* $e^T v = n$. *Then for any positive constant* η *the following inequality holds:*

$$(\|v - e\| + \eta)^p \prod_{i=1}^n v_i \;\le\; p^{p+1} \theta^p,$$

where θ *is a positive constant no greater than* $1.3 + \eta$.

Proof. We need to consider only the case in which $n \ge 2$, as the inequality holds trivially when $n = 1$. Consider the maximization problem

$$\begin{aligned} &\max && f(v) \;:=\; \|v - e\|^p \prod_{i=1}^n \alpha_i \\ &\text{such that} && e^T v = n. \end{aligned}$$

It is shown in [10] that the maximizer v has the form

$$v_1 = \gamma, \quad v_2 = \cdots = v_n = \frac{n-\gamma}{n-1}, \qquad \text{where } \gamma > 1,$$

and

$$f(v) \le \left(\frac{n}{n-1}\right)^{p/2} (p+1)^{p+1} \exp\left(\frac{-p(p+2)}{p+1}\right).$$

Thus

$$\begin{aligned}(\|v-e\|+\eta)^p \prod_{i=1}^n v_i &= \left(\|v-e\| \prod_{i=1}^n v_i^{1/p} + \eta \prod_{i=1}^n v_i^{1/p}\right)^p \\ &\le \left[\left(\|v-e\|^p \prod_{i=1}^n v_i\right)^{1/p} + \eta\right]^p, \qquad \left(\text{since } \textstyle\prod_{i=1}^n v_i \le (\frac{e^T v}{n})^n = 1\right) \\ &\le \left[\left(\frac{n}{n-1}\right)^{1/2} (p+1)^{(p+1)/p} \exp\left(\frac{-(p+2)}{p+1}\right) + \eta\right]^p \\ &\le p^{p+1}\theta^p,\end{aligned}$$

where

$$\theta = \max_{n\ge 2, p\ge 1} \left\{\left(\frac{n}{n-1}\right)^{1/2} \left(1+\frac{1}{p}\right)^{1+1/p} \exp\left(\frac{-1}{p+1} - 1\right) + \frac{\eta}{p} p^{-1/p}\right\} \;\le\; 1.3+\eta. \qquad \square$$

LEMMA 4.2. *Suppose Y_k is an approximate analytic center of Ω_k with $\delta_k(Y_k) \le \eta < 1 - \sqrt{3}/2$. Then*

$$\phi_{k+1}(\Omega_{k+1}) \;\ge\; \phi_k(\Omega_k) - \frac{p_k}{2} \ln(p_k\, \bar{r}_k^2\, \theta^2) - \ln p_k, \tag{4.2}$$

where θ is a constant depending only on η.

Proof. For simplicity, we will drop the subscripts k and $k+1$ in our notations in this proof and denote, for example, Ω_k and Ω_{k+1} by Ω and Ω_+, respectively.

Let $\bar{U} = I - \bar{Y}$, $\bar{U}_+ = I - \bar{Y}_+$, and

$$\bar{s}^+ \;=\; c - \mathcal{A}^T \mathbf{svec}(\bar{Y}_+), \qquad \bar{s} \;=\; c - \mathcal{A}^T \mathbf{svec}(\bar{Y}), \qquad \bar{t} \;=\; d - \mathcal{B}^T \mathbf{svec}(\bar{Y}_+).$$

Let

$$\bar{G} \;=\; [\mathcal{A}\bar{S}^{-1},\;\; -\bar{Y}^{-1/2} \circledast \bar{Y}^{-1/2},\;\; \bar{U}^{-1/2} \circledast \bar{U}^{-1/2}].$$

Note that $\bar{H} = \bar{G}\bar{G}^T$.

First, we establish an upper bound for $\ln \prod_{j=1}^p \bar{t}_j$. We have

$$\begin{aligned}\bar{t} = \mathcal{B}^T(\mathbf{svec}Y - \mathbf{svec}\bar{Y}_+) \;&=\; \mathcal{B}^T \bar{H}^{-1} \bar{G} \left(\bar{G}^T \mathbf{svec}Y - \bar{G}^T \mathbf{svec}\bar{Y}_+\right) \\ &= (\bar{G}^T \bar{H}^{-1} \mathcal{B})^T \left(\bar{G}^T \mathbf{svec}(Y - \bar{Y}) - \bar{G}^T \mathbf{svec}(\bar{Y}_+ - \bar{Y})\right).\end{aligned}$$

Thus

$$\|\bar{t}\| \le \|\bar{G}^T \bar{H}^{-1} \mathcal{B}\| \left(\|\bar{G}^T \mathbf{svec}(Y - \bar{Y})\| + \|\bar{G}^T \mathbf{svec}(\bar{Y}_+ - \bar{Y})\|\right).$$

By part (iii) of Theorem 2.2.2 in [7], we have

$$\|\bar{G}^T \mathbf{svec}(Y - \bar{Y})\| \;=\; \eta(\bar{x}, s, Y, \bar{Z}, \bar{V}) \le \frac{1 - [1 - 3\delta(Y)]^{1/3}}{[1 - 3\delta(Y)]^{1/3}} \le 3\delta(Y) \;\le\; 3\eta.$$

Thus

$$\|\bar{t}\| \le \bar{r}\ \left(3\eta + \|\bar{G}^T \mathbf{svec}(\bar{Y}_+ - \bar{Y})\|\right).$$

Hence

$$\ln \prod_{j=1}^{p} \bar{t}_j = \frac{p}{2}\left(\frac{1}{p}\sum_{j=1}^{p} \ln \bar{t}_j^2\right) \ \le\ \frac{p}{2}\ln\left(\frac{\sum_{j=1}^{p}\bar{t}_j^2}{p}\right) \ =\ \frac{p}{2}\ln\|\bar{t}\|^2 - \frac{p}{2}\ln p$$

$$\le p\ln\left(3\eta + \|\bar{G}^T\mathbf{svec}(\bar{Y}_+ - \bar{Y})\|\right) + p\ln\bar{r} - \frac{p}{2}\ln p, \tag{4.3}$$

and the desired upper bound is established.

Now observe that

$$\phi_+(\Omega_+) - \phi(\Omega) \ =\ -\ln\prod_{j=1}^{p}\bar{t}_j - \ln\left(\prod_{i=1}^{n}\frac{\bar{s}_i^+}{\bar{s}_i}\ \frac{\det\bar{Y}_+}{\det\bar{Y}}\ \frac{\det\bar{U}_+}{\det\bar{U}}\right).$$

Using the bound in (4.3), we have

$$\phi(\Omega_+) - \phi(\Omega) \ \ge\ \frac{p}{2}\ln p - p\ln\bar{r} - \ln\left(3\eta + \|\bar{G}^T\mathbf{svec}(\bar{Y}_+ - \bar{Y})\|\right)^p\prod_{i=1}^{n}\frac{\bar{s}_i^+}{\bar{s}_i}\ \frac{\det\bar{Y}_+}{\det\bar{Y}}\ \frac{\det\bar{U}_+}{\det\bar{U}}. \tag{4.4}$$

The inequality (4.2) follows, once we have shown that

$$\left(3\eta + \|\bar{G}^T\mathbf{svec}(\bar{Y}_+ - \bar{Y})\|\right)^p\prod_{i=1}^{n}\frac{\bar{s}_i^+}{\bar{s}_i}\ \frac{\det\bar{Y}_+}{\det\bar{Y}}\ \frac{\det\bar{U}_+}{\det\bar{U}}\ \le\ p^{p+1}\,\theta^p. \tag{4.5}$$

Note that

$$\begin{aligned}
&\|\bar{G}^T\mathbf{svec}(\bar{Y}_+ - \bar{Y})\|^2\\
&\quad = \mathbf{svec}(\bar{Y}_+ - \bar{Y})^T\left[\mathcal{A}\bar{S}^{-2}\mathcal{A}^T + \bar{Y}^{-1}\circledast\bar{Y}^{-1} + \bar{U}^{-1}\circledast\bar{U}^{-1}\right]\mathbf{svec}(\bar{Y}_+ - \bar{Y})\\
&\quad = (\bar{s} - \bar{s}^+)^T\bar{S}^{-2}(\bar{s} - \bar{s}^+)\ +\ \mathbf{svec}(\bar{Y}_+ - \bar{Y})^T(\bar{Y}^{-1}\circledast\bar{Y}^{-1})\mathbf{svec}(\bar{Y}_+ - \bar{Y})\\
&\qquad + \mathbf{svec}(\bar{U} - \bar{U}_+)^T(\bar{U}^{-1}\circledast\bar{U}^{-1})\mathbf{svec}(\bar{U} - \bar{U}_+)\\
&\quad = \left\|\begin{pmatrix} e - \bar{S}^{-1}\bar{s}^+\\ e - \lambda(\bar{Y}^{-1/2}\bar{Y}_+\bar{Y}^{-1/2})\\ e - \lambda(\bar{U}^{-1/2}\bar{U}_+\bar{U}^{-1/2})\end{pmatrix}\right\|^2,
\end{aligned}$$

and, by using (2.1), we have

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

By Lemma 4.1, (4.5) is proved. ∎

The complexity analysis is based on the following idea. For the sequence of working sets Ω_k, we can establish upper and lower bounds on $\phi(\Omega_k)$. The upper bound is approximately $n_k \ln \epsilon^{-1}$, which is a consequence of the assumption that Γ contains a ball of radius ϵ and the fact that Ω_k is defined by n_k cuts. The lower bound is obtained by estimating $-\sum_{i=0}^{k-1} p_i \ln \bar{r}_i$, which is a consequence of Lemma 4.2. A sophisticated estimation of $\bar{r}_k$ gives rise to a lower bound that is proportional to $n_k \ln(n_k/m^3)$. The algorithm must terminate before the lower and upper bounds conflict with each other.

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

LEMMA 4.3. *Let $\Omega_k \supset \Gamma$ be defined by n_k linear inequalities and the positive semidefinite constraint. Suppose Assumptions* A1–A3 *hold. Then*

$$\phi_k(\Omega_k) \leq -(n_k + 2m) \ln \epsilon.$$

Proof. Assumptions A1–A3 imply that there exists a point $Y^c \in \Gamma$ such that

(i) all eigenvalues of Y^c and $I - Y^c$ are greater than or equal to ϵ;

(ii) for any $A \in \mathcal{S}^m$ with $\|A\|_F = 1$ and $\alpha \in \mathbb{R}$, if $\Gamma \subset \{Y \mid A \bullet Y \leq \alpha\}$, then $\alpha - A \bullet Y^c \geq \epsilon$.

We will briefly describe how to prove $\lambda(Y_c) \geq \epsilon e$, before continuing with the proof of the lemma. Suppose that λ_j is an eigenvalue of Y_c, and v_j is a corresponding unit eigenvector. Consider the matrix $\hat{Y}_c := Y_c - \lambda_j v_j v_j^T$. Since this matrix has a zero eigenvalue, it lies on the boundary of Ω_0, and by Assumption A3, we have

$$\epsilon \;\leq\; \|\hat{Y}_c - Y_c\|_F \;=\; \|\lambda_j v_j v_j^T\|_F \;=\; \lambda_j \|v_j\|^2 \;=\; \lambda_j.$$

The fact that $\lambda(Y_c) \leq (1-\epsilon)e$ can be proved similarly.

Now we continue with the proof of the lemma. Since $\Gamma \subset \Omega_k$,

$$\phi_k(\Omega_k) \;\leq\; \phi_k(Y^c) \;=\; -\sum_{i=1}^{n_k} \ln(c_i - A_i \bullet Y^c) - \ln \det Y^c - \ln \det(I - Y^c).$$

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

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

we have the desired inequality. ∎

Now we turn to finding a lower bound for $\phi_k(\Omega_k)$. By Lemma 4.2, we have

$$\phi_k(\Omega_k) \geq \phi_0(\Omega_0) - \sum_{i=0}^{k-1} p_i \ln \bar{r}_i - \frac{1}{2}\sum_{i=0}^{k-1} p_i \ln p_i - \sum_{i=0}^{k-1} p_i \ln \theta - \sum_{i=0}^{k-1} \ln p_i. \tag{4.6}$$

Obviously, we need to estimate $\bar{r}_i$ for each i. We first seek to bound $\bar{H}_i^{-1}$ by D_i^{-1}, where D_i is defined as follows. Let $D_0 = 8I$, where I is the identity matrix of order $\bar{m} \times \bar{m}$. For $i = 1, 2, \ldots$, let

$$D_i = D_0 + \frac{1}{m} \sum_{j=0}^{i-1} \mathcal{B}_j \mathcal{B}_j^T,$$

where $\mathcal{B}_j = [\mathbf{svec}A_{n_j+1}, \ldots, \mathbf{svec}A_{n_j+p_j}]$.

LEMMA 4.4. *Let A_{n_i+j} (with $\|A_{n_i+j}\|_F = 1$), $j = 1, \ldots, p_i$, be the cuts generated from the approximate analytic center $Y_i \in \Omega_i$, $i = 0, \ldots, k-1$. Let $c_{n_i+j} = \mathbf{svec}(A_{n_i+j})^T \mathbf{svec}(Y_i)$, $j = 1, \ldots, p_i$, $i = 1, \ldots, k$. For any point $Y \in \Omega_k$, let $s = c - \mathcal{A}^T \mathbf{svec}Y$, where*

$$\mathcal{A} = [\mathcal{B}_0 \;\; \mathcal{B}_1 \;\; \cdots \;\; \mathcal{B}_{k-1}], \qquad c = \begin{bmatrix} \mathcal{B}_0^T \mathbf{svec}(Y_0) \\ \vdots \\ \mathcal{B}_{k-1}^T \mathbf{svec}(Y_{k-1}) \end{bmatrix}.$$

Then

$$s_{n_i+j} \le \sqrt{m} \quad \forall \; j = 1, \ldots, p_i, \; i = 1, \ldots, k, \qquad \nabla^2 \phi_k(Y) \succeq D_k.$$

In particular, $\bar{H}_k = \nabla^2 \phi_k(\bar{Y}_k) \succeq D_k$.

Proof. We first estimate s_{n_i+j}. We have

$$\begin{aligned} s_{n_i+j} &= c_{n_i+j} - \mathbf{svec}(A_{n_i+j})^T \mathbf{svec}(Y) = (\mathbf{svec}A_{n_i+j})^T (\mathbf{svec}Y_i - \mathbf{svec}Y) \\ &\le \|\mathbf{svec}A_{n_i+j}\| \, \|\mathbf{svec}Y_{i-1} - \mathbf{svec}Y\| \\ &= \|\mathbf{svec}Y_i - \mathbf{svec}Y\| = \|Y_i - Y\|_F \\ &= \left[\sum_{j=1}^{m} \lambda_j^2 (Y_i - Y) \right]^{1/2} \le \sqrt{m}\,. \end{aligned} \tag{4.7}$$

The last inequality holds because, by Assumption A2,

$$I \succeq Y_i \succeq Y_i - Y \succeq -Y \succeq -I,$$

implying that $e \ge \lambda(Y_i - Y) \ge -e$.

Next, let $U = I - Y$ and $S_i = \mathrm{diag}(s_{n_i+1}, \ldots, s_{n_i+p_i})$. Then

$$\begin{aligned} \nabla^2 \phi_k(Y) &= Y^{-1} \circledast Y^{-1} + U^{-1} \circledast U^{-1} + \mathcal{A} S^{-2} \mathcal{A}^T \\ &= Y^{-1} \circledast Y^{-1} + U^{-1} \circledast U^{-1} + \sum_{i=0}^{k-1} \mathcal{B}_i \, S_i^{-2} \, \mathcal{B}_i^T \\ &\succeq 8I + \frac{1}{m} \sum_{i=0}^{k-1} \mathcal{B}_i \, \mathcal{B}_i^T = D_k. \end{aligned} \tag{4.8}$$

Note that in deriving (4.8) we used the fact that $S_i \preceq \sqrt{m} \, I_{p_i}$ for each i, and that

$$Y^{-1} \circledast Y^{-1} + U^{-1} \circledast U^{-1} \succeq 8I. \qquad \square$$

In our complexity analysis, we will make the following assumptions:

A4. $p_{\max} \le m$, where $p_{\max} = \max\{p_i \mid i = 0, 1, \ldots\}$.

A5. Let $\bar{M}_i = \mathcal{B}_i^T \bar{H}_i^{-1} \mathcal{B}_i$. There exists a fixed constant $\tau \ge 1$ such that, for each $i = 0, 1, \ldots,$

$$\lambda_{\max}(\bar{M}_i) \le \tau \, \frac{\mathrm{Tr}(\bar{M}_i)}{p_i}.$$

Assumption A4 is made for technical reasons. It is used in the proof of Lemma 4.5. Such an assumption also appeared in the papers [3] and [10]. Note that Assumption A4 can be relaxed to $p_{\max} \le O(m)$. But, for simplicity, we fix the constant at 1.

Note that Assumption A5 holds trivially with $\tau = p_{\max}$. For the special case in which a single cut is used in each iteration, it holds with $\tau = 1$. Thus by fixing τ at an intermediate value between 1 and $p_{\max}$, we admit only cuts that are sufficiently good in the sense that the matrix $\bar{M}_i$ cannot have too many small eigenvalues. Of course, one may not want to fix τ at the extreme value 1, since then the criterion is likely to reject most of the cuts unless there are many mutually orthogonal (with respect to $\bar{H}_i^{-1}$) cuts.

The main advantage of having Assumption A5 is that in each oracle call we have an objective criterion to select only cuts that are useful from among a possibly large number of ineffective cuts. In this way, the number of cuts added in each iteration will not be unnecessarily large, and hence the computational time in each iteration will not grow as rapidly as in the case where the cuts are admitted unchecked. The choice of τ in practice would depend on the problem at hand. It should dynamically be adjusted as information on the quality of the cuts is obtained as the cutting algorithm progresses. If the choice of τ is too stringent and many good cuts are rejected, then we can progressively increase its value so that more good cuts are selected.

However, without a priori information on the quality of the cuts, we propose to choose τ to be a small constant, say 5, based on the following empirical observation. We conducted numerical experiments on random matrices of the form V^TV, where $V \in \mathbb{R}^{\bar{m}\times p}$, for $p = 1, \ldots, m$, and $m = 10, 20, \ldots, 260$. The elements of V are drawn independently from the standard normal distribution. We computed the ratio between the largest eigenvalue of V^TV and $\mathrm{Tr}(V^TV)/p$ for each V, and found that these ratios are less than 2 for all of the 3510 cases tested.

Now let us continue with our complexity analysis. Let

$$w_i^2 \; = \; \mathrm{Tr}(\mathcal{B}_i^T D_i^{-1} \mathcal{B}_i).$$

Since

$$p_i\, \bar{r}_i^2 \;\le\; \tau\, \mathrm{Tr}(\mathcal{B}_i^T \bar{H}_i^{-1} \mathcal{B}_i) \;\le\; \tau\, \mathrm{Tr}(\mathcal{B}_i^T D_i^{-1} \mathcal{B}_i) \;=\; \tau w_i^2,$$

we have

$$\sum_{i=0}^{k-1} p_i\, \bar{r}_i^2 \quad\le\quad \tau \sum_{i=0}^{k-1} w_i^2. \tag{4.9}$$

Next, we establish an upper bound for the right-hand side of the above inequality. Its proof is modeled after that of [10, Lemma 3.5]. However, we have simplified the proof by considering all the cuts simultaneously instead of handling them one by one as in [10].

LEMMA 4.5.

$$\sum_{i=0}^{k-1} w_i^2 \quad\le\quad \frac{9m\bar{m}}{8} \ln\left(1 + \frac{n_k}{8m\bar{m}}\right).$$

Proof. From the equation

$$\det D_{i+1} \;=\; \det D_i \prod_{j=1}^{p_i} \left[1 + \frac{1}{m} \lambda_j(\mathcal{B}_i^T D_i^{-1} \mathcal{B}_i)\right],$$

we have

$$\ln \det D_{i+1} - \ln \det D_i = \sum_{j=1}^{p_i} \ln \left[1 + \frac{1}{m} \lambda_j (\mathcal{B}_i^T D_i^{-1} \mathcal{B}_i) \right]$$

$$\geq \frac{8}{9m} \sum_{j=1}^{p_i} \lambda_j (\mathcal{B}_i^T D_i^{-1} \mathcal{B}_i) \quad = \quad \frac{8}{9m} w_i^2, \tag{4.10}$$

where we used the fact that $\lambda_{\max}(\mathcal{B}_i^T D_i^{-1} \mathcal{B}_i) \leq \lambda_{\max}(\mathcal{B}_i^T \mathcal{B}_i)/8 = \|\mathcal{B}_i\|_F^2/8 = p_i/8$, and the inequality $\ln(1+x) \geq 8x/9$ for $0 \leq x \leq 1/8$. We also made use of Assumption A4, which yields that $p_i \leq m$.

From (4.10), it follows immediately that

$$\ln \det D_k - \ln \det D_0 \;\geq\; \frac{8}{9m} \sum_{i=0}^{k-1} w_i^2. \tag{4.11}$$

However,

$$\frac{1}{\bar{m}} \ln \det D_k \leq \ln \frac{\mathrm{Tr}(D_k)}{\bar{m}} \quad = \quad \ln \frac{1}{\bar{m}} \left[\mathrm{Tr}(D_0) + \frac{1}{m} \sum_{i=0}^{k-1} \mathrm{Tr}(\mathcal{B}_i \mathcal{B}_i^T) \right]$$

$$= \ln \frac{1}{\bar{m}} \left[8\bar{m} + \frac{1}{m} \sum_{i=0}^{k-1} p_i \right] \quad = \quad \ln \left(8 + \frac{n_k}{m\bar{m}} \right),$$

implying that

$$\ln \det D_k - \ln \det D_0 \;\leq\; \bar{m} \ln \left(1 + \frac{n_k}{8m\bar{m}} \right). \tag{4.12}$$

Combining (4.11) and (4.12), the lemma is proved. $\square$

With the above lemma, we can now formally state a lower bound for $\phi_k(\Omega_k)$.

LEMMA 4.6. *Suppose that Assumptions* A1–A5 *hold. Then*

$$\phi_k(\Omega_k) \geq -\frac{1}{2}(2m + n_k) \ln \left[\frac{4m + 9\tau\, m\bar{m} \ln(1 + \frac{n_k}{8m\bar{m}})}{8(2m + n_k)} \right]$$
$$- \frac{1}{2} \sum_{i=0}^{k-1} p_i \ln p_i - n_k \ln \theta - \sum_{i=0}^{k-1} \ln p_i,$$

where θ is the constant that appeared in (4.2).

Proof. The proof is similar to that of Theorem 10 in [10], after using (4.9) and Lemma 4.5. $\square$

We will next estimate the number of oracle calls required to find a feasible point of Γ.

LEMMA 4.7. *Suppose that Assumptions* A1–A5 *hold. Then the analytic center cutting plane method stops with a feasible point before k violates the following inequality:*

$$\frac{\epsilon^2}{p_{\max} m\bar{m}} \;\leq\; \frac{4/\bar{m} + 9\tau \ln(1 + \frac{n_k}{8m\bar{m}})}{8(2m + n_k)} \exp \left(\frac{2n_k \ln \theta + 2 \sum_{i=0}^{k-1} \ln p_i}{n_k + 2m} \right). \tag{4.13}$$

Proof. From Lemmas 4.3 and 4.6, we have

$$-(2m+n_k)\ln\epsilon \geq -\frac{1}{2}(2m+n_k)\ln\left[\frac{4m+9\tau m\bar{m}\ln(1+\frac{n_k}{8m\bar{m}})}{8(2m+n_k)}\right]$$
$$-\frac{1}{2}\sum_{i=0}^{k-1} p_i \ln p_i - n_k \ln\theta - \sum_{i=0}^{k-1}\ln p_i.$$

Thus, the algorithm must terminate before k violates the above inequality; i.e., the algorithm must terminate before k violates the following inequality:

$$\frac{\epsilon^2}{m\bar{m}} \leq \frac{4/\bar{m}+9\tau\ln(1+\frac{n_k}{8m\bar{m}})}{8(2m+n_k)}\exp\left(\frac{\sum_{i=0}^{k-1}p_i\ln p_i + 2n_k\ln\theta + 2\sum_{i=0}^{k-1}\ln p_i}{2m+n_k}\right). \tag{4.14}$$

Since $\sum_{i=0}^{k-1} p_i \ln p_i \leq \sum_{i=0}^{k-1} p_i \ln p_{\max} = n_k \ln p_{\max}$, the algorithm must terminate before k violates the inequality in the lemma. ∎

THEOREM 4.8. *Suppose that Assumptions* A1–A5 *hold. Then the analytic center cutting plane method terminates in at most $O^*(m^3\tau\, p_{\max}\ln p_{\max}/\epsilon^2)$ Newton steps, where the notation O^* means that lower order terms are ignored. The total number of cuts added is not more than $O^*(m^3\tau\, p_{\max}/\epsilon^2)$.*

Proof. Ignoring lower order terms (assuming $k \gg m$) and by the assumption that τ is a constant independent of $p_{\max}$, the above lemma implies that the algorithm stops as soon as k satisfies

$$\frac{n_k}{\tau\ \ln(n_k/m^3)} \geq O\left(\frac{m^3 p_{\max}}{\epsilon^2}\right).$$

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

$$n_k = O^*\left(\frac{m^3\tau\, p_{\max}}{\epsilon^2}\right)$$

cuts. By Theorem 3.9, the total number of Newton steps is

$$O\left(\sum_{i=1}^{k} p_i \ln p_i\right) \leq O(n_k \ln p_{\max}) = O^*\left(\frac{m^3\tau\, p_{\max}\ln p_{\max}}{\epsilon^2}\right).$$

The theorem is proved. ∎

For feasibility problems in $\mathbb{R}^m_+$, $\bar{m}$ should be replaced by m in Lemma 4.7. Thus the complexity bound is $O(m^2\tau\, p_{\max}\ln p_{\max}/\epsilon^2)$ for the number of required Newton steps. This bound is better than the bounds obtained in [2], [5], and [10].

Acknowledgment. We thank the referees for their constructive comments that greatly helped to improve the paper.

REFERENCES

[1] J.-L. Goffin, Z.-Q. Luo, and Y. Ye, *Complexity analysis of an interior cutting plane method for convex feasibility problems*, SIAM J. Optim., 6 (1996), pp. 638–652.

[2] J.-L. Goffin and J.-P. Vial, *Multiple cuts in the analytic center cutting plane method*, SIAM J. Optim., 11 (2000), pp. 266–288.

[3] J.-L. Goffin and J.-P. Vial, *Convex Nondifferentiable Optimization: A Survey Focused on the Analytic Center Cutting Plane Method*, HEC/Logilab Technical report 99-02, University of Geneva, Geneva, Switzerland, 1999.

[4] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[5] Z.-Q. Luo, *Analysis of a cutting plane method that uses weighted analytic center and multiple cuts*, SIAM J. Optim., 7 (1997), pp. 697–716.
[6] Y. Nesterov, *Cutting plane algorithms from analytic centers: Efficiency estimates*, Math. Programming, 69 (1995), pp. 149–176.
[7] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Stud. Appl. Math., SIAM, Philadelphia, 1994.
[8] J. Sun, K.C. Toh, and G. Zhao, *An analytic center cutting plane method for semidefinite feasibility problems*, Math. Oper. Res., to appear.
[9] Y. Ye, *A potential reduction algorithm allowing column generation*, SIAM J. Optim., 2 (1992), pp. 7–20.
[10] Y. Ye, *Complexity analysis of the analytic center cutting plane method that uses multiple cuts*, Math. Programming, 78 (1997), pp. 85–104.
[11] Y. Ye, *Interior Point Algorithms: Theory and Analysis*, John Wiley & Sons, New York, 1997.