SIAM J. OPTIM.
Vol. 10, No. 1, pp. 155–176

CONDITION-BASED COMPLEXITY OF CONVEX OPTIMIZATION IN CONIC LINEAR FORM VIA THE ELLIPSOID ALGORITHM*

ROBERT M. FREUND[†] AND JORGE R. VERA[‡]

Abstract. A convex optimization problem in conic linear form is an optimization problem of the form

$$\begin{array}{rll} CP(d): & \text{maximize} & c^T x \\ & \text{s.t.} & b - Ax \in C_Y, \\ & & x \in C_X, \end{array}$$

where C_X and C_Y are closed convex cones in n- and m-dimensional spaces X and Y, respectively, and the data for the system is $d = (A, b, c)$. We show that there is a version of the ellipsoid algorithm that can be applied to find an ϵ-optimal solution of $CP(d)$ in at most $O(n^2 \ln(\frac{\mathcal{C}(d)\|c\|_*}{c_1 \epsilon}))$ iterations of the ellipsoid algorithm, where each iteration must either perform a separation cut on one of the cones C_X or C_Y or perform a related optimality cut. The quantity $\mathcal{C}(d)$ is the "condition number" of the program $CP(d)$ originally developed by Renegar and is essentially a scale-invariant reciprocal of the smallest data perturbation $\Delta d = (\Delta A, \Delta b, \Delta c)$ for which the system $CP(d + \Delta d)$ becomes either infeasible or unbounded. The scalar quantity c_1 is a constant that depends only on the simple notion of the "width" of the cones and is independent of the problem data $d = (A, b, c)$ but may depend on the dimensions m and/or n.

Key words. complexity of convex optimization, ellipsoid method, conditioning, error analysis

AMS subject classifications. 90C, 90C05, 90C60

PII. S105262349732829X

1. Introduction.

Consider a convex program in conic linear form:

$$\begin{array}{rll} CP(d): & \text{maximize} & c^T x \\ & \text{s.t.} & b - Ax \in C_Y, \\ & & x \in C_X, \end{array} \tag{1}$$

where $C_X \subset X$ and $C_Y \subset Y$ are each a closed convex cone in the (finite) n-dimensional linear vector space X (with norm $\|x\|$ for $x \in X$) and in the (finite) m-dimensional linear vector space Y (with norm $\|y\|$ for $y \in Y$), respectively. Here $b \in Y$, and $A \in L(X, Y)$, where $L(X, Y)$ denotes the set of all linear operators $A : X \to Y$. Also, $c \in X^*$, where X^* is the space of all linear functionals defined on X; i.e., X^* is the dual space of X. In order to maintain consistency with standard linear algebra notation in mathematical programming, we consider c to be a column vector in the space X^* and we denote the linear function $c(x)$ by $c^T x$. Similarly, for $A \in L(X, Y)$ and $f \in Y^*$, we denote $A(x)$ by Ax and $f(y)$ by $f^T y$. We denote the adjoint of A by A^T.

*Received by the editors October 6, 1997; accepted for publication (in revised form) October 30, 1998; published electronically October 20, 1999. This research has been partially supported through National Science Foundation grant INT-9703803 and a grant from Fundación Andes/CONICYT, in the framework of the international cooperation agreement with NSF. The second author also was supported by a grant from CONICYT within the FONDAP program in applied mathematics.

http://www.siam.org/journals/siopt/10-1/32829.html

[†]MIT Operations Research Center, 77 Massachusetts Avenue, Cambridge, MA 02139 (rfreund@mit.edu).

[‡]Department of Industrial and Systems Engineering, Catholic University of Chile, Campus San Joaquín, Vicuña Mackenna 4860, Santiago, Chile (jvera@ing.puc.cl).

The "data" d for problem $CP(d)$ is the array $d = (A, b, c) \in \{L(X,Y), Y, X^*\}$. We call the above program $CP(d)$ rather than simply CP to emphasize the dependence of the optimization problem on the data $d = (A, b, c)$, and we note that the cones C_X and C_Y are not part of the data; that is, they are considered to be given and fixed. At the moment, we make no assumptions on C_X and on C_Y except to note that each is a closed convex cone.

The format of $CP(d)$ is quite general (any convex optimization problem can be cast in the format of $CP(d)$) and has received much attention recently in the context of interior-point algorithms; see Nesterov and Nemirovskii [13] and Renegar [19], [20], as well as Nesterov and Todd [15], [14] and Nesterov, Todd, and Ye [16], among others.

In contrast to interior-point methods, this paper focuses on the complexity of solving $CP(d)$ via the ellipsoid algorithm. The ellipsoid algorithm of Yudin and Nemirovskii [26] and Shor [21] (see also [4], [8], and [9]) and the interior-point algorithm of Nesterov and Nemirovskii [13] are two fundamental theoretically efficient algorithms for solving general convex optimization. The ellipsoid algorithm enjoys a number of important advantages over interior-point algorithms: the ellipsoid algorithm is based on elegantly simple geometric notions, it always has excellent theoretical efficiency in the dimension of the variables n, it requires only the use of a separation oracle for its implementation, and it is important in both continuous and discrete optimization [8]. (Of course, when applied to solving linear programs, interior-point algorithms typically exhibit vastly superior practical performance over the ellipsoid algorithm, but that is not the focus of this study.)

The ellipsoid algorithm belongs to a larger class of efficient volume-reducing cutting-plane algorithms that includes the method of centers of gravity [11], the method of inscribed ellipsoids [10], and the method of volumetric centers [22], among others. We focus herein on the ellipsoid algorithm because of its prominence and history in the complexity analysis of convex optimization, but our analysis is applicable to these other volume-reducing cutting-plane methods as well; see the remarks in section 6.

In analyzing the complexity of the ellipsoid algorithm, we adopt the relatively new concept of the *condition number* $\mathcal{C}(d)$ of the program $CP(d)$, developed by Renegar in the series of papers [17], [18], and [19]. We show (in section 5) that there is a version of the ellipsoid algorithm that can be applied to find an ϵ-optimal solution of $CP(d)$ in at most $O(n^2 \ln(\frac{\mathcal{C}(d)\|c\|_*}{c_1 \epsilon}))$ iterations of the ellipsoid algorithm, where each iteration must perform either a separation cut on one of the cones C_X or C_Y or a related optimality cut. The quantity $\mathcal{C}(d)$ is the condition number of the program $CP(d)$, and $\|c\|_*$ is the norm of c. The scalar quantity c_1 is a constant that depends only on the simple notion of the "width" of the cones, and is independent of the problem data $d = (A, b, c)$, but may depend on the dimensions m and/or n.

Two special cases of $CP(d)$ deserve special mention: linear programming and semidefinite programming. Let $\Re$ and $\Re_+$ denote the set of real numbers and the set of nonnegative real numbers, respectively, and let $\Re^k$ and $\Re^k_+$ denote real k-dimensional space and the nonnegative orthant in $\Re^k$, respectively. Then by setting (i) $C_X = \Re^n_+$ and $C_Y = \Re^m_+$, (ii) $C_X = \Re^n_+$ and $C_Y = \{0\}$, or (iii) $C_X = \Re^n$ and $C_Y = \Re^m_+$, then $CP(d)$ is a linear program of the format (i) $\max\{c^T x \mid Ax \leq b, x \geq 0, x \in \Re^n\}$, (ii) $\max\{c^T x \mid Ax = b, x \geq 0, x \in \Re^n\}$, or (iii) $\max\{c^T x \mid Ax \leq b, x \in \Re^n\}$, respectively.

The other special case of $CP(d)$ that we mention is semidefinite programming. Semidefinite programming has been shown to be of enormous importance in mathematical programming (see Alizadeh [1] and Nesterov and Nemirovskii [13] as well as

Vandenberghe and Boyd [23]). Let X denote the set of real $k \times k$ symmetric matrices, whereby $n = k(k+1)/2$, and define the Löwner partial ordering "$\succeq$" on X as $x \succeq w$ if and only if the matrix $x - w$ is positive semidefinite. The semidefinite program in standard (primal) form is the problem $\max\{c^T x \mid Ax = b, x \succeq 0\}$. Define $C_X = \{x \in X \mid x \succeq 0\}$. Then C_X is a closed convex cone. Let $Y = \Re^m$ and $C_Y = \{0\} \subset \Re^m$. Then the standard form semidefinite program is easily seen to be an instance of $CP(d)$.

Most studies of the ellipsoid algorithm (for example, [9], [4], [8]) pertain to the case when $CP(d)$ is a linear or convex quadratic program and focus on the complexity of the algorithm in terms of the bit length L of a binary representation of the data $d = (A, b, c)$. However, when the cones C_X and/or C_Y are not polyhedral or when the data $d = (A, b, c)$ are not rational, it makes little or no sense to study the complexity of the ellipsoid algorithm in terms of L. Indeed, a much more natural and intuitive measure that is relevant for complexity analysis and that captures the inherent data-dependent behavior of $CP(d)$ is the "condition number" $\mathcal{C}(d)$ of the problem $CP(d)$, which was developed by Renegar in a series of papers [17], [18], [19]. The quantity $\mathcal{C}(d)$ is essentially a scale invariant reciprocal of the smallest data perturbation $\Delta d = (\Delta A, \Delta b, \Delta c)$ for which the system $CP(d + \Delta d)$ becomes either infeasible or unbounded. (These concepts will be reviewed in detail shortly.)

The paper is organized as follows. The remainder of this introductory section discusses the condition number $\mathcal{C}(d)$ of the optimization problem $CP(d)$. Section 2 contains further notation and a discussion of the width of a cone. In section 3 we demonstrate a ball construction for the set of ϵ-optimal solutions of $CP(d)$, and we review several previous results regarding the geometry of $CP(d)$. Section 4 briefly reviews relevant complexity aspects of the ellipsoid algorithm and reviews a transformation of $CP(d)$ into a homogenized form called $HP(d)$ that is more convenient for the application of the ellipsoid algorithm. Lemma 4.1 contains a key volume-ratio upper bound that is the main tool used in proving the complexity results for the ellipsoid algorithm for solving $CP(d)$, which are presented in section 5. Section 6 discusses related issues: complexity results for other volume-reducing cutting-plane algorithms, testing for ϵ-optimality, the complexity of testing for infeasibility of $CP(d)$, and bounding the skewness of the ellipsoids computed in the ellipsoid algorithm.

The concept of the "distance to ill-posedness" and a closely related condition number for problems such as $CP(d)$ was introduced by Renegar in [17] in a more specific setting but then generalized more fully in [18] and [19]. We now describe these two concepts in detail.

Using the constructs of Lagrangian duality, one obtains the following dual problem of $CP(d)$:

$$\begin{array}{llll} & CD(d): & \text{minimize} & b^T y \\ (2) & & \text{s.t.} & A^T y - c \in C_X^*, \\ & & & y \in C_Y^*, \end{array}$$

where C_X^* and C_Y^* are the dual convex cones associated with the cones C_X and C_Y, respectively, and where the dual cone of a convex cone K in a linear vector space X is defined by

$$K^* = \{z \in X^* | z^T x \geq 0 \text{ for any } x \in K\}.$$

The data for the program $CD(d)$ is also the array $d = (A, b, c)$.

We denote the space of all data $d = (A, b, c)$ for $CP(d)$ by $\mathcal{D}$. Then $\mathcal{D} = \{d = (A, b, c) \mid A \in L(X, Y), b \in Y, c \in X^*\}$. Because X and Y are normed linear vector spaces, we can define the following product norm on the data space $\mathcal{D}$:

$$\|d\| = \|(A, b, c)\| = \max\{\|A\|, \|b\|, \|c\|_*\} \qquad \text{for any } d \in \mathcal{D},$$

where $\|A\|$ is the operator norm, namely,

$$\|A\| = \max\{\|Ax\| \mid \|x\| \leq 1\},$$

and where $\|c\|_*$ is the dual norm of c induced on $c \in X^*$, defined as

$$\|c\|_* = \max\{c^T x \mid \|x\| \leq 1, x \in X\},$$

with a similar definition holding for $\|v\|_*$ for $v \in Y^*$.

Consider the following subsets of the data set $\mathcal{D}$:

$$\mathcal{F}_P = \{(A, b, c) \in \mathcal{D} \mid \text{ there exists } x \text{ such that } b - Ax \in C_Y, x \in C_X\},$$

$$\mathcal{F}_D = \{(A, b, c) \in \mathcal{D} \mid \text{ there exists } y \text{ such that } A^T y - c \in C_X^*, y \in C_Y^*\},$$

and

$$\mathcal{F} = \mathcal{F}_P \cap \mathcal{F}_D.$$

The elements in $\mathcal{F}_P$ correspond to those data instances $d = (A, b, c)$ in $\mathcal{D}$ for which $CP(d)$ is feasible and the elements in $\mathcal{F}_D$ correspond to those data instances $d = (A, b, c)$ in $\mathcal{D}$ for which $CD(d)$ is feasible. Observe that $\mathcal{F}$ is the set of data instances $d = (A, b, c)$ that are both primal and dual feasible. The complement of $\mathcal{F}_P$, denoted by $\mathcal{F}_P^C$, is the set of data instances $d = (A, b, c)$ for which $CP(d)$ is infeasible, and the complement of $\mathcal{F}_D$, denoted by $\mathcal{F}_D^C$, is the set of data instances $d = (A, b, c)$ for which $CD(d)$ is infeasible.

The boundary of $\mathcal{F}_P$ and $\mathcal{F}_P^C$ is the set

$$\mathcal{B}_P = \partial\mathcal{F}_P = \partial\mathcal{F}_P^C = \operatorname{cl}(\mathcal{F}_P) \cap \operatorname{cl}(\mathcal{F}_P^C),$$

and the boundary of $\mathcal{F}_D$ and $\mathcal{F}_D^C$ is the set

$$\mathcal{B}_D = \partial\mathcal{F}_D = \partial\mathcal{F}_D^C = \operatorname{cl}(\mathcal{F}_D) \cap \operatorname{cl}(\mathcal{F}_D^C),$$

where ∂S denotes the boundary of a set S and $\operatorname{cl}(S)$ is the closure of a set S. Note that $\mathcal{B}_P \neq \emptyset$ since $(0, 0, 0) \in \mathcal{B}_P$. The data instances $d = (A, b, c)$ in $\mathcal{B}_P$ are called the ill-posed data instances for the primal, in that arbitrarily small changes in the data $d = (A, b, c)$ can yield data instances in $\mathcal{F}_P$ as well as data instances in $\mathcal{F}_P^C$. Similarly, the data instances $d = (A, b, c)$ in $\mathcal{B}_D$ are called the ill-posed data instances for the dual.

For $d = (A, b, c) \in \mathcal{D}$, we define the ball centered at d with radius δ as

$$B(d, \delta) = \{\bar{d} \in \mathcal{D} : \|\bar{d} - d\| \leq \delta\}.$$

For a data instance $d \in \mathcal{D}$, the "primal distance to ill-posedness" is defined as follows:

$$\rho_P(d) = \inf\{\|\Delta d\| : d + \Delta d \in \mathcal{B}_P\}$$

(see [17], [18], [19]), and so $\rho_P(d)$ is the distance of the data instance $d = (A, b, c)$ to the set $\mathcal{B}_\mathcal{P}$ of ill-posed instances for the primal problem $CP(d)$. It is straightforward to show that

$$\rho_P(d) = \begin{cases} \sup\{\delta : B(d,\delta) \subset \mathcal{F}_P\} & \text{if } d \in \mathcal{F}_P, \\ \sup\{\delta : B(d,\delta) \subset \mathcal{F}_P^C\} & \text{if } d \in \mathcal{F}_P^C, \end{cases} \tag{3}$$

so that we could also define $\rho_P(d)$ by employing (3). In the typical case when $CP(d)$ is feasible, i.e., $d \in \mathcal{F}_P$, $\rho_P(d)$ is the minimum change Δd in the data d needed to create a primal-infeasible instance $d + \Delta d$, and so $\rho_P(d)$ measures how close the data instance $d = (A, b, c)$ is to the set of infeasible instances of $CP(d)$. Put another way, $\rho_P(d)$ measures how close $CP(d)$ is to being infeasible. Note that $\rho_P(d)$ measures the distance of the data d to primal infeasible instances, and so the objective function vector c plays no role in this measure.

The "primal condition number" $\mathcal{C}_P(d)$ of the data instance d is defined as

$$\mathcal{C}_P(d) = \frac{\|d\|}{\rho_P(d)}$$

when $\rho_P(d) > 0$ and $\mathcal{C}_P(d) = \infty$ when $\rho_P(d) = 0$. The primal condition number $\mathcal{C}_P(d)$ can be viewed as a scale-invariant reciprocal of $\rho_P(d)$, as it is elementary to demonstrate that $\mathcal{C}_P(d) = \mathcal{C}_P(\alpha d)$ for any positive scalar α. Observe that since $\bar{d} = (\bar{A}, \bar{b}, \bar{c}) = (0, 0, 0) \in \mathcal{B}_P$ and $\mathcal{B}_P$ is a closed set, then for any $d \notin \mathcal{B}_P$ we have $\|d\| \geq \rho_P(d) > 0$, so that $\mathcal{C}_P(d) \geq 1$. The value of $\mathcal{C}_P(d)$ is a measure of the relative conditioning of the primal *feasibility* problem for the data instance d. For a discussion of the relevance of using $\mathcal{C}_P(d)$ as a condition number for the problem $CP(d)$, see Renegar [17], [18] and Vera [24].

These measures are not nearly as intangible as they might seem at first glance. In [7], it is shown that $\rho_P(d)$ can be computed by solving rather simple convex optimization problems involving the data $d = (A, b, c)$, the cones C_X and C_Y, and the norms $\|\cdot\|$ given for the problem. As in traditional condition numbers for systems of linear equations, the computation of $\rho_P(d)$ and hence of $\mathcal{C}_P(d)$ is roughly as difficult as solving $CP(d)$; see [7].

For a data instance $d \in \mathcal{D}$, the "dual distance to ill-posedness" is defined in a manner exactly analogous to the "primal distance to ill-posedness":

$$\rho_D(d) = \inf\{\|\Delta d\| : d + \Delta d \in \mathcal{B}_D\}$$

or equivalently

$$\rho_D(d) = \begin{cases} \sup\{\delta : B(d,\delta) \subset \mathcal{F}_D\} & \text{if } d \in \mathcal{F}_D, \\ \sup\{\delta : B(d,\delta) \subset \mathcal{F}_D^C\} & \text{if } d \in \mathcal{F}_D^C. \end{cases} \tag{4}$$

The "dual condition number" $\mathcal{C}_D(d)$ of the data instance d is defined as

$$\mathcal{C}_D(d) = \frac{\|d\|}{\rho_D(d)}$$

when $\rho_D(d) > 0$ and $\mathcal{C}_D(d) = \infty$ when $\rho_D(d) = 0$.

The two measures of distances to ill-posed instances and condition numbers are combined as follows. Recalling the definition of $\mathcal{F}$, the elements in $\mathcal{F}$ correspond to those data instances $d = (A, b, c)$ in $\mathcal{D}$ for which both $CP(d)$ and $CD(d)$ are feasible.

The complement of $\mathcal{F}$, denoted by $\mathcal{F}^C$, is the set of data instances $d = (A, b, c)$ for which $CP(d)$ is infeasible or $CD(d)$ is infeasible. The boundary of $\mathcal{F}$ and $\mathcal{F}^C$ is the set

$$\mathcal{B} = \partial\mathcal{F} = \partial\mathcal{F}^C = \operatorname{cl}(\mathcal{F}) \cap \operatorname{cl}(\mathcal{F}^C).$$

The data instances $d = (A, b, c)$ in $\mathcal{B}$ are called the ill-posed data instances in that arbitrarily small changes in the data $d = (A, b, c)$ can yield data instances in $\mathcal{F}$ as well as data instances in $\mathcal{F}^C$. For a data instance $d \in \mathcal{D}$, the "distance to ill-posedness" is defined as follows:

$$\rho(d) = \inf\{\|\Delta d\| : d + \Delta d \in \mathcal{B}\}$$

or equivalently

$$\rho(d) = \begin{cases} \sup\{\delta : B(d,\delta) \subset \mathcal{F}\} & \text{if } d \in \mathcal{F}, \\ \sup\{\delta : B(d,\delta) \subset \mathcal{F}^C\} & \text{if } d \in \mathcal{F}^C. \end{cases} \tag{5}$$

In the typical case when $CP(d)$ and $CD(d)$ are both feasible, i.e., $d \in \mathcal{F}$, $\rho(d)$ is the minimum change Δd in the data d needed to create a data instance $d + \Delta d$ that is either primal infeasible or dual infeasible. The "condition number" $\mathcal{C}(d)$ of the data instance d is defined as

$$\mathcal{C}(d) = \frac{\|d\|}{\rho(d)}$$

when $\rho(d) > 0$ and as $\mathcal{C}(d) = \infty$ when $\rho(d) = 0$. The condition number $\mathcal{C}(d)$ can be viewed as a scale-invariant reciprocal of $\rho(d)$. The value of $\mathcal{C}(d)$ is a measure of the relative conditioning of the problem $CP(d)$ and its dual $CD(d)$ for the data instance d.

It is straightforward to demonstrate that

$$\rho(d) = \min\{\rho_P(d), \rho_D(d)\} \quad \text{if } d \in \mathcal{F},$$

and so

$$\mathcal{C}(d) = \max\{\mathcal{C}_P(d), \mathcal{C}_D(d)\} \quad \text{if } d \in \mathcal{F}. \tag{6}$$

We offer the following interpretation of $\rho(d)$ and $\mathcal{C}(d)$ in terms of the primal problem when both the primal problem and the dual problem are feasible. Because $\rho_P(d)$ measures how close the data instance $d = (A, b, c)$ is to being an infeasible instance of the primal, and the $\rho_D(d)$ measures how close the data instance $d = (A, b, c)$ is to being an unbounded instance of the primal (in the primal objective function value), then $\rho(d)$ measures how close the data instance $d = (A, b, c)$ is to being either a primal infeasible or a primal unbounded data instance. The larger the value of condition number $\mathcal{C}(d)$ is, the closer the primal problem is to either an infeasible or an unbounded instance of the primal.

2. Further notation, coefficient of linearity, and width of a cone. We will say that a cone C is *regular* if C is a closed convex cone, has a nonempty interior, and is pointed (i.e., contains no line).

REMARK 2.1. *If C is a closed convex cone, then C is regular if and only if C^* is regular.*

Let C be a regular cone in the normed linear vector space X. Let $B(x, r)$ denote the ball centered at x with radius r. We will use the following definition of the width of C.

DEFINITION 2.1. *If C is a regular cone in the normed linear vector space X, the width of C is given by*

$$\tau = \max\left\{ \frac{r}{\|x\|} \mid B(x,r) \subset C \right\}.$$

We remark that τ measures the maximum ratio of the radius to the norm of the center of an inscribed ball in C, and so larger values of τ correspond to an intuitive notion of greater width of C. Note that $\tau \in (0, 1]$, since C has a nonempty interior and C is pointed, and τ is attained for some $(\bar{x}, \bar{r})$ as well as along the ray $(\alpha\bar{x}, \alpha\bar{r})$ for all $\alpha > 0$.

In previous work [7], we employed the "coefficient of linearity" for a cone C.

DEFINITION 2.2. *If C is a regular cone in the normed linear vector space X, the coefficient of linearity for the cone C is given by*

$$\beta = \sup_{\substack{u \in X^*, \\ \|u\|_* = 1,}} \ \inf_{\substack{x \in C, \\ \|x\| = 1.}} u^T x \tag{7}$$

The coefficient of linearity β for the regular cone C is essentially the same as the scalar α defined in Renegar [19, p. 328]. In [7], the coefficient of linearity was used as part of an analysis of geometric properties of the feasible region of $CP(d)$ that are implied by the condition number $\mathcal{C}_P(d)$. The following proposition shows that the width of C is equal to the coefficient of linearity of C^*.

PROPOSITION 2.1. *Suppose that C is a regular cone in the normed linear vector space X, τ denote the width of C, and β^* denote the coefficient of linearity for C^*. Then $\tau = \beta^*$.*

Proof. From the definition of the coefficient of linearity for C^*, we have

$$\beta^* = \sup_{\substack{x \in X, \\ \|x\| = 1,}} \ \inf_{\substack{w \in C^*, \\ \|w\|_* = 1.}} x^T w \tag{8}$$

From the outer optimization problem above, there exists $\bar{x} \in X$ for which $\|\bar{x}\| = 1$ and $w^T\bar{x} \geq \beta^*$ for any $w \in C^*$ satisfying $\|w\|_* = 1$. Let $x \in B(\bar{x}, \beta^*)$, i.e., $x = \bar{x} + \beta^* v$, where $\|v\| \leq 1$. For any $w \in C^*$ satisfying $\|w\|_* = 1$, we have $w^T x = w^T\bar{x} + \beta^* w^T v \geq w^T\bar{x} - \beta^*\|w\|_*\|v\| \geq \beta^* - \beta^* = 0$, and so $B(\bar{x}, \beta^*) \subset C$. Therefore, $\tau \geq \frac{\beta^*}{\|\bar{x}\|} = \beta^*$.

From the definition of the width of C, there exists $\tilde{x}$ satisfying $\|\tilde{x}\| = 1$ and $B(\tilde{x}, \tau) \subset C$. Let $w \in C^*$ satisfying $\|w\|_* = 1$ be given. Then, from the duality properties of norms, there exists $\bar{v} \in X$ satisfying $\|\bar{v}\| \leq 1$ for which $\|w\|_* = w^T\bar{v}$. Since $B(\tilde{x}, \tau) \subset C$, $\tilde{x} - \tau\bar{v} \in C$, and so $w^T(\tilde{x} - \tau\bar{v}) \geq 0$, whereby $w^T\tilde{x} \geq \tau w^T\bar{v} = \tau\|w\|_* = \tau$. As this is true for any given $w \in C^*$ satisfying $\|w\|_* = 1$, it follows that $\beta^* \geq \tau$, completing the proof. ∎

We illustrate the width construction on two families of cones, the nonnegative orthant $\Re^k_+$ and the positive semidefinite cone $S^{k\times k}_+$. First consider the nonnegative orthant. Let $X = \Re^k$ with Euclidean norm $\|x\| = \|x\|_2 = \sqrt{x^T x}$, and $C = \Re^k_+ = \{x \in \Re^k \mid x \geq 0\}$. Then it is straightforward to show, by setting $x = e = (1, \dots, 1)^T$, that the width of $\Re^k_+$ is $\tau = 1/\sqrt{k}$. Next consider the positive semidefinite cone. Let $X = S^{k\times k}$ denote the set of real $k \times k$ symmetric matrices with Frobenius norm $\|x\| := \sqrt{\mathrm{trace}(x^T x)}$, and let $C = S^{k\times k}_+ = \left\{x \in S^{k\times k} \mid x \succeq 0\right\}$. Then $S^{k\times k}_+$ is a closed convex cone, and it is easy to show by setting $x = I$ that the width of $S^{k\times k}_+$ is $\tau = \frac{1}{\sqrt{k}}$.

For the remainder of this paper, we amend our notation as follows.

DEFINITION 2.3. *Whenever the cone C_X is regular, the width of C_X is denoted by τ, and the width of C_X^* is denoted by τ^*. Whenever the cone C_Y is regular, the width C_Y is denoted by $\bar\tau$, and the width of C_Y^* is denoted by $\bar\tau^*$.*

3. A ball construction for the ϵ-optimal set for $CP(d)$. In this section we demonstrate some valuable geometric properties of the set of ϵ-optimal solutions of $CP(d)$ that will be used later to obtain complexity bounds for the ellipsoid algorithm. Let X_d denote the feasible region of $CP(d)$ and let $z^*(d)$ denote the optimal objective function value of $CP(d)$. For any $\epsilon > 0$, denote the set of ϵ-optimal solutions of $CP(d)$ by X_d^ϵ, i.e., $X_d^\epsilon = \{x \in X \mid x \in X_d \ \text{ and } \ c^T x \geq z^*(d) - \epsilon\}$.

Let $\epsilon > 0$ be given. The following lemma asserts the existence of a ball in the set of ϵ-optimal solutions of $CP(d)$ that has certain geometric properties, under the condition that the feasible region contains a ball $B(\hat x, r)$.

LEMMA 3.1. *Suppose that the feasible region X_d contains the ball $B(\hat x, r)$, where $r > 0$. Let x^* be an optimal solution of $CP(d)$, and let $\epsilon > 0$ be given. Then there exists a ball $B(\bar x, \bar r)$ with the following properties:*

$$\text{(i)} \quad B(\bar x, \bar r) \subset X_d^\epsilon \ ,$$

$$\text{(ii)} \quad \bar r \geq \frac{\epsilon r}{\max\{\epsilon, z^*(d) - c^T\hat x + r\|c\|_*\}} \ ,$$

$$\textit{and} \quad \text{(iii)} \quad \|\bar x\| \leq \max\left\{\|\hat x\|, \|x^*\|\right\}.$$

Proof. We have $B(\hat x, r) \subset X_d$ and $x^* \in X_d$. Therefore, from the convexity of X_d, we have

$$B(\alpha\hat x + (1-\alpha)x^*, \alpha r) \subset X_d \quad \text{ for any } \alpha \in [0,1]. \tag{9}$$

We have two cases.

Case 1. $\epsilon \leq z^*(d) - c^T\hat x + r\|c\|_*$. Define

$$\alpha = \frac{\epsilon}{z^*(d) - c^T\hat x + r\|c\|_*} \ , \quad \bar x = \alpha\hat x + (1-\alpha)x^* \ , \quad \text{and} \quad \bar r = \alpha r.$$

Then $\alpha \in [0,1]$ and so $B(\bar x, \bar r) \subset X_d$ from (9). Furthermore, for any $x \in B(\bar x, \bar r)$, we have

$$c^T x \geq \alpha c^T\hat x + (1-\alpha)c^T x^* - \alpha r\|c\|_* = z^*(d) - \alpha\left(z^*(d) - c^T\hat x + r\|c\|_*\right) = z^*(d) - \epsilon,$$

whereby (i) is satisfied. For (ii), note that

$$\bar r = \alpha r = \frac{\epsilon r}{z^*(d) - c^T\hat x + r\|c\|_*} = \frac{\epsilon r}{\max\{\epsilon, z^*(d) - c^T\hat x + r\|c\|_*\}}.$$

Part (iii) follows since $\|\bar x\| = \|\alpha\hat x + (1-\alpha)x^*\| \leq \max\left\{\|\hat x\|, \|x^*\|\right\}$.

Case 2. $\epsilon > z^*(d) - c^T\hat{x} + r\|c\|_*$. Define

$$\bar{x} = \hat{x} \quad \text{and} \quad \bar{r} = r.$$

To prove (i), note that for any $x \in B(\hat{x}, r)$, we have

$$c^T x \geq c^T\hat{x} - r\|c\|_* = z^*(d) - \left(z^*(d) - c^T\hat{x} + r\|c\|_*\right) > z^*(d) - \epsilon,$$

whereby (i) is satisfied. Parts (ii) and (iii) follow trivially. $\square$

We would like to apply Lemma 3.1 to obtain a lower bound on the volume of the set of ϵ-optimal solutions of $CP(d)$. However, in order to obtain such a lower bound via Lemma 3.1, we need the following ingredients:

(i) an upper bound on the optimal objective function value $z^*(d)$ of $CP(d)$,

(ii) an upper bound on the norm of an optimal solution x^* of $CP(d)$, and

(iii) the existence of a ball $B(\hat{x}, r)$ in the feasible region for which there is an upper bound on $\|\hat{x}\|$ and a lower bound on r.

The following previously derived results pertain to the first two conditions above.

THEOREM 1 OF [17]. *Suppose that $d \in \mathcal{F}$ and $\mathcal{C}(d) < +\infty$. Then*

$$|z^*(d)| \leq \|c\|_*\mathcal{C}(d). \tag{10}$$

Furthermore, $CP(d)$ attains its optimum and every optimal solution x^* satisfies

$$\|x^*\| \leq \mathcal{C}(d)^2 \ . \quad \square \tag{11}$$

The third condition above is treated with the following previously known results.

THEOREM 5.1 OF [7] *Suppose that C_X is a regular cone and C_Y is a regular cone and that $d \in \mathcal{F}$ and that $\mathcal{C}(d) < +\infty$. Then there exists $\hat{x} \in X_d$ and a scalar $r > 0$ such that $B(\hat{x}, r) \subset X_d$, and*

$$r \geq \frac{\min\{\tau, \bar{\tau}\}}{6\mathcal{C}(d)} \ , \quad \|\hat{x}\| \leq \frac{4\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \ , \quad \textit{and} \quad \frac{\|\hat{x}\|}{r} \leq \frac{6\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \ . \quad \square \tag{12}$$

THEOREM 5.3 OF [7] *Suppose that C_X is a regular cone and $C_Y = \{0\}$ and that $d \in \mathcal{F}$ and that $\mathcal{C}(d) < +\infty$. Then there exists $\hat{x} \in X_d$ and a scalar $r > 0$ such that $\{x \in X \mid \|x - \hat{x}\| \leq r, Ax = b\} \subset X_d$, and*

$$r \geq \frac{\tau}{3\mathcal{C}(d)} \ , \quad \|\hat{x}\| \leq \frac{4\mathcal{C}(d)}{\tau} \ , \quad \textit{and} \quad \frac{\|\hat{x}\|}{r} \leq \frac{3\mathcal{C}(d)}{\tau} \ . \quad \square \tag{13}$$

THEOREM 5.5 OF [7] *Suppose that $C_X = X$ and C_Y is a regular cone, that $d \in \mathcal{F}$, and that $\mathcal{C}(d) < +\infty$. Then there exists $\hat{x} \in X_d$ and a scalar $r > 0$ such that $B(\hat{x}, r) \subset X_d$, and*

$$r \geq \frac{\bar{\tau}}{3\mathcal{C}(d)} \ , \quad \|\hat{x}\| \leq \frac{3\mathcal{C}(d)}{\bar{\tau}} \ , \quad \textit{and} \quad \frac{\|\hat{x}\|}{r} \leq \frac{2\mathcal{C}(d)}{\bar{\tau}} \ . \quad \square \tag{14}$$

(These three results are slightly altered from their presentation in [7], which uses the notation of coefficients of linearity. In the notation of [7], we have from Proposition 2.1 that $\tau = \beta^*, \tau^* = \beta, \bar{\tau} = \bar{\beta}^*$, and $\bar{\tau}^* = \bar{\beta}$. The above statements follow by noticing from [7] that $\mathcal{C}(d) \geq 1, \tau \leq 1, \bar{\tau} \leq 1$, and $\frac{\|\hat{x}\|}{r} \leq \frac{R}{r} - 1$.)

4. The ellipsoid algorithm and a homogenizing transformation. We review a few basic results regarding the ellipsoid algorithm for solving an optimization problem; see [26], [21], [9], [4], [8], [3]. We will consider the following optimization problem:

$$P: \quad \underset{x}{\text{maximize}} \quad f(x) \qquad \text{s.t.} \quad x \in S, \tag{15}$$

where S is a convex set (closed or not) in $\Re^k$, $f(x)$ is a quasi-concave function, and $\|x\|_2 := \sqrt{x^T x}$ is the Euclidean norm. Actually, the ellipsoid algorithm is more usually associated with the assumption that S is a closed convex set and also that $f(x)$ is a concave function, but these assumptions can be relaxed slightly. It is only necessary that S be a convex set, that the upper level sets of $f(x)$ be convex sets on S (which is equivalent to the statement that $f(x)$ is a quasi-concave function on S; see [2], for example), and that a separation oracle be available for S as well as for each of the upper level sets of $f(x)$. (Note that if $f(x)$ is a differentiable quasi-concave function, then $\nabla f(x)$ furnishes a separation oracle for the upper level sets of $f(x)$, provided that $\nabla f(x)$ does not vanish at any nonmaximizing points.)

In order to implement the ellipsoid algorithm to approximately solve P, it is necessary that one has available a separation oracle for the set S, i.e., that for any $\bar{x} \notin S$, one can perform a feasibility cut for the set S, which consists of computing a vector $v \neq 0$ for which $S \subset \{x \mid v^T x \geq v^T \bar{x}\}$. Suppose that T_1 is an upper bound on the number of operations needed to perform a feasibility cut for the set S. It is also necessary that one has available a support oracle for the upper level sets $U_\alpha = \{x \in S \mid f(x) \geq \alpha\}$ of the quasi-concave function $f(x)$. That is, for any $\bar{x} \in S$, it is necessary to be able to perform an optimality cut for the objective function $f(x)$ at any point $\bar{x} \in S$, which consists of computing a vector $v \neq 0$ for which $U_{f(\bar{x})} \subset \{x \in \Re^k \mid v^T x \geq v^T \bar{x}\}$. Suppose that T_2 is an upper bound on the number of operations needed to compute an optimality cut for the function $f(x)$ on the set S.

Let z^* denote the optimal value of P, and denote the set of ϵ-optimal solutions of P by S^ϵ, i.e., $S^\epsilon = \{x \in \Re^k \mid x \in S \text{ and } f(x) \geq z^* - \epsilon\}$. In a typical application of the ellipsoid algorithm, we wish to find an ϵ-optimal solution of P. Suppose that we know a priori a positive scalar R with the property that

$$B(0,R) \cap S^\epsilon$$

has positive volume, where $B(\bar{x}, r) := \{x \in \Re^k \mid \|x - \bar{x}\|_2 \leq r\}$ is the Euclidean ball centered at $\bar{x}$ with radius r. Then the ellipsoid algorithm for solving P can be initiated with the Euclidean ball $B(0,R)$. The following is a generic result about the performance of the ellipsoid algorithm, where in the statement of the theorem, "vol(Q)" denotes the volume of a set Q.

ELLIPSOID ALGORITHM THEOREM WITH KNOWN $\boldsymbol{R}$ (from [26], [21]). *Suppose that a positive scalar R is known with the property that the set*

$$F := B(0,R) \cap S^\epsilon$$

has positive volume. Then, if the ellipsoid algorithm is initiated with the Euclidean ball $B(0,R)$, the algorithm will compute an ϵ-optimal solution of P in at most

$$\left\lceil 2(k+1) \ln \left(\frac{\text{vol}(B(0,R))}{\text{vol}(B(0,R) \cap S^\epsilon)} \right) \right\rceil \tag{16}$$

iterations, where each iteration must perform at most $\left(k^2 + \max\{T_1, T_2\}\right)$ *operations, and where* T_1 *and* T_2 *are the numbers of operations needed to perform a feasibility cut on* S *and an optimality cut on* $f(x)$*, respectively.*

We note that the bound on the number of operations per iteration arises from performing either a feasibility or an optimality cut (which takes $\max\{T_1, T_2\}$ operations), and then performing a rank-one update of the positive definite matrix defining the ellipsoid (see [3], for example), which takes k^2 operations.

Because an a priori bound on R is typically not known except in very special cases of P, we employ a standard homogenizing transformation to convert P to the homogenized fractional program:

$$(17)\qquad \begin{array}{lll} HP: & \underset{w,\theta}{\text{maximize}} & g(w,\theta) := f(w/\theta) \\ & \text{s.t.} & w \in \theta S, \\ & & \theta > 0 \end{array}$$

(see, for example, [5] and [6]). It is trivial to show that z^* is the common optimal objective function value of P and HP. Let H and H^ϵ denote the set of feasible and ϵ-optimal solutions of HP, respectively, i.e.,

$$(18)\qquad H = \{(w,\theta) \in \Re^{k+1} \mid w \in \theta S,\ \theta > 0\}$$

and

$$(19)\qquad H^\epsilon = \{(w,\theta) \in \Re^{k+1} \mid w \in \theta S,\ \theta > 0,\ g(w,\theta) \geq z^* - \epsilon\}.$$

Then H and H^ϵ are both convex sets. Furthermore, the objective function $g(w,\theta) := f(w/\theta)$ of HP is easily seen to be a quasi-concave function over the feasible region H whenever $f(x)$ is a quasi-concave function over the feasible region S. The following (obvious) transformations $h(\cdot)$ and $h^{-1}(\cdot)$ map the feasible regions and ϵ-optimal regions of P and HP onto one another:

$$(20)\qquad h(T) = \{(w,\theta) \in \Re^{k+1} \mid w/\theta \in T \text{ and } \theta > 0\} \qquad \text{for any } T \subset S$$

and

$$(21)\qquad h^{-1}(W) = \{x \in \Re^k \mid x = w/\theta \text{ for some } (w,\theta) \in W\} \qquad \text{for any } W \subset H\ .$$

Because any feasible solution of HP can be scaled by an arbitrary positive scalar without changing its objective function value or affecting its feasibility, the feasible region and all upper level sets of the objective function $g(w,\theta)$ of HP contain points in the $(k+1)$-dimensional unit Euclidean ball. This allows us to conveniently start the ellipsoid algorithm for solving HP with the $(k+1)$-dimensional unit Euclidean ball.

The following result concerns volumes of subsets of S under the projective transformation $h(\cdot)$ and provides the final ingredient we will need for our analysis of the ellipsoid algorithm. Let B^{k+1} denote the $(k+1)$-dimensional unit Euclidean ball, namely,

$$B^{k+1} := \left\{(w,\theta) \in \Re^{k+1} \mid \sqrt{w^T w + \theta^2} \leq 1\right\}.$$

LEMMA 4.1. *Suppose that S is a convex set in $\Re^k$, that $T \subset S$ is given, that there exists $\bar{r} > 0$ and $\bar{x}$ for which $B(\bar{x}, \bar{r}) \subset T$, and that $\bar{r} \leq 1$. Let $W = h(T)$, where $h(\cdot)$ is defined as in* (20). *Then*

$$\ln\left(\frac{\mathrm{vol}\left(B^{k+1}\right)}{\mathrm{vol}\left(B^{k+1} \cap W\right)}\right) \leq (k+1)\ln\left(2 + \frac{3(\|\bar{x}\|+1)}{\bar{r}}\right) + \left[\ln\left(\|\bar{x}\|\right)\right]^{+}.$$

Proof. We first define two constants,

$$\delta = \max\{\|\bar{x}\|, 1\}$$

and

$$\gamma = 1 + \frac{\bar{r}}{3} + \frac{\bar{r}}{3\delta} + \|\bar{x}\|,$$

and we define the following ellipsoid centered at $(\bar{x}, 1) \in \Re^{k+1}$:

$$E = \left\{(w, \theta) \in \Re^{k+1} \mid \sqrt{(w-\bar{x})^T(w-\bar{x}) + \delta^2(\theta-1)^2} \leq \frac{\bar{r}}{3}\right\}.$$

We prove below that

$$E \subset W, \tag{22}$$

$$E \subset \gamma B^{k+1}. \tag{23}$$

It then follows that

$$\gamma^{-1}E \subset B^{k+1} \quad \text{and} \quad \gamma^{-1}E \subset W, \tag{24}$$

and so

$$\gamma^{-1}E \subset B^{k+1} \cap W, \tag{25}$$

since in particular W is closed under positive scalings. Then

$$\begin{aligned}
\ln\left(\frac{\mathrm{vol}\left(B^{k+1}\right)}{\mathrm{vol}\left(B^{k+1} \cap W\right)}\right) &\leq \ln\left(\frac{\mathrm{vol}\left(B^{k+1}\right)}{\mathrm{vol}\left(\gamma^{-1}E\right)}\right) \\
&= (k+1)\ln(\gamma) + \ln\left(\frac{\mathrm{vol}\left(B^{k+1}\right)}{\mathrm{vol}\left(E\right)}\right) \\
&= (k+1)\ln(\gamma) + \ln\left(\frac{1}{\left(\frac{\bar{r}}{3}\right)^{k+1}\left(\frac{1}{\delta}\right)}\right) \\
&= (k+1)\ln\left(\frac{3\gamma}{\bar{r}}\right) + \ln\left(\delta\right) \\
&= (k+1)\ln\left(\frac{3}{\bar{r}} + 1 + \frac{1}{\delta} + \frac{3\|\bar{x}\|}{\bar{r}}\right) + \ln\left(\delta\right) \\
&\leq (k+1)\ln\left(2 + \frac{3(\|\bar{x}\|+1)}{\bar{r}}\right) + \left[\ln\left(\|\bar{x}\|\right)\right]^{+},
\end{aligned}$$

since $\delta \geq 1$. We therefore need to demonstrate (22) and (23) to complete the proof.

For any $(w,\theta) \in E$, $(w,\theta) = (\bar{x}+q, 1+v)$, where $\|q\| \leq \frac{\bar{r}}{3}$ and $|v| \leq \frac{\bar{r}}{3\delta} \leq \frac{1}{3}$, since $\bar{r} \leq 1$ and $\delta \geq 1$, and so $\theta \geq \frac{2}{3} > 0$. We also have

$$\frac{w}{\theta} = \frac{\bar{x}+q}{1+v} = \bar{x} + \frac{q - v\bar{x}}{1+v} \ ,$$

and so

$$\left\|\frac{w}{\theta} - \bar{x}\right\| = \frac{\|q - v\bar{x}\|}{1+v} \leq \frac{3}{2}\left(\|q\| + |v|\|\bar{x}\|\right) \leq \frac{3}{2}\left(\frac{\bar{r}}{3} + \frac{\bar{r}}{3\delta}\|\bar{x}\|\right) \leq \bar{r} \ .$$

Therefore, $\frac{w}{\theta} \in B(\bar{x}, \bar{r})$, whereby $\frac{w}{\theta} \in T$, and so $w \in \theta T$ or, equivalently, $(w,\theta) \in h(T)$. Therefore, $E \subset h(T) = W$, proving (22).

To prove (23), let $(w,\theta) \in E$. Then $\|w - \bar{x}\|_2 \leq \frac{\bar{r}}{3}$ and $|\theta - 1| \leq \frac{\bar{r}}{3\delta}$. Therefore,

$$\begin{aligned} \|(w,\theta)\|_2 &\leq \|(w-\bar{x}, \theta - 1)\|_2 + \|(\bar{x},1)\|_2 \\ &\leq \|w - \bar{x}\|_2 + |\theta - 1| + \|\bar{x}\|_2 + 1 \\ &\leq \frac{\bar{r}}{3} + \frac{\bar{r}}{3\delta} + \|\bar{x}\|_2 + 1 \\ &= \gamma \ , \end{aligned}$$

and so $(w,\theta) \in \gamma B^{k+1}$, which proves (23) and thus the proof of the lemma is complete. ∎

It is trivial to show that a separation oracle for S can be readily converted to a separation oracle for H. If T_1 is the number of operations needed to compute a feasibility cut for S, then one needs $O(T_1 + k)$ operations to compute a feasibility cut for H. Furthermore, any support oracle for the upper level sets of $f(x)$ over S can be readily converted to a support oracle for the upper level sets of $g(w,\theta)$ over H. To see why this is true, suppose that $(\bar{w},\bar{\theta})$ is a feasible solution of HP, and define $\bar{x} = \bar{w}/\bar{\theta}$. Then $\bar{x}$ is feasible for P and let v be the vector produced by the support oracle for $f(x)$ at $x = \bar{x}$. Then

$$\{x \in S \mid f(x) \geq f(\bar{x})\} \subset \left\{x \in \Re^k \mid v^T x \geq v^T \bar{x}\right\},$$

which implies that

$$\left\{(w,\theta) \in H \mid g(w,\theta) \geq g(\bar{w},\bar{\theta})\right\} \subset \left\{(w,\theta) \in \Re^{k+1} \mid v^T w - ((v^T\bar{w})/\bar{\theta})\theta \geq 0\right\},$$

and so the concatenated vector $(v, -(v^T\bar{w}/\bar{\theta}))$ is a support vector for the upper level set of the function $g(w,\theta)$ at the feasible point $(\bar{w},\bar{\theta})$. If T_2 is the number of operations needed to compute an optimality cut on $f(x)$ over S, then one needs $O(T_2 + k)$ operations to compute an optimality cut on $g(w,\theta)$ over H.

Finally, returning to the problem $CP(d)$, note that the homogenized problem corresponding to $CP(d)$ is

$$\begin{aligned} HP(d): \quad \text{maximize}_{w,\theta} \quad & g(w,\theta) := \frac{c^T w}{\theta} \\ \text{s.t.} \quad & b\theta - Aw \in C_Y, \\ & w \in C_X, \\ & \theta > 0 \ , \end{aligned} \tag{26}$$

which we refer to as $HP(d)$.

5. Complexity results. In this section, we assume that $X = \Re^n$ is endowed with the Euclidean norm $\|x\| = \|x\|_2 = \sqrt{x^T x}$. For the purpose of developing complexity results, we focus on three different classes of instances of $CP(d)$, namely,

$$\begin{aligned} &\text{Class (i):} && C_X \text{ and } C_Y \text{ are both regular;} \\ &\text{Class (ii):} && C_X \text{ is regular and } C_Y = \{0\}; \\ &\text{Class (iii):} && C_X = X \text{ and } C_Y \text{ is regular.} \end{aligned}$$

For these three classes of instances, $CP(d)$ can be written as (i) $\max\{c^T x \mid b - Ax \in C_Y, x \in C_X\}$, (ii) $\max\{c^T x \mid Ax = b, x \in C_X\}$, and (iii) $\max\{c^T x \mid b - Ax \in C_Y, x \in X\}$, respectively.

The following three theorems contain iteration complexity bounds on the ellipsoid algorithm for these three classes of instances of $CP(d)$, respectively. The proofs of the theorems are deferred to the end of the section.

THEOREM 5.1. *Suppose that C_X is a regular cone with width τ, that C_Y is a regular cone with width $\bar{\tau}$, and that $d \in \mathcal{F}$ and $\mathcal{C}(d) < +\infty$. Let ϵ satisfying $0 < \epsilon < \|c\|_*$ be given. Suppose that the ellipsoid algorithm is applied to solve $HP(d)$ and is initiated with the Euclidean unit ball centered at $(w^0, \theta^0) = (0, 0)$. Then the ellipsoid algorithm will compute an ϵ-optimal solution of $HP(d)$ (and hence, by transformation, to $CP(d)$) in at most*

$$\left\lceil 8(n+2)^2 \ln\left(\frac{4\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \frac{\|c\|_*}{\epsilon}\right)\right\rceil$$

iterations, where each iteration must perform at most $\big((n+1)^2 + \max\{2n, S_1, m + mn + S_2\}\big)$ operations, and where S_1 and S_2 are the number of operations needed to perform a feasibility cut on C_X and C_Y, respectively.

THEOREM 5.2. *Suppose that C_X is a regular cone with width τ, that $C_Y = \{0\}$, and that $d \in \mathcal{F}$ and $\mathcal{C}(d) < +\infty$. Let ϵ satisfying $0 < \epsilon < \|c\|_*$ be given. Suppose that the ellipsoid algorithm is applied to solve $HP(d)$ and is initiated with the Euclidean unit disk centered at $(w^0, \theta^0) = (0, 0)$ in the subspace $\{(w, \theta) \in \Re^{n+1} \mid Aw - b\theta = 0\}$. Then the ellipsoid algorithm will compute an ϵ-optimal solution of $HP(d)$ (and hence, by transformation, to $CP(d)$) in at most*

$$\left\lceil 8(n-m+2)^2 \ln\left(\frac{3\mathcal{C}(d)}{\tau} \frac{\|c\|_*}{\epsilon}\right)\right\rceil$$

iterations, where each iteration must perform at most $\big((n-m+1)^2 + \max\{2n, S_1\}\big)$ operations, and where S_1 is the number of operations needed to perform a feasibility cut on C_X.

THEOREM 5.3. *Suppose that $C_X = X$ and C_Y is a regular cone with width $\bar{\tau}$, and that $d \in \mathcal{F}$ and $\mathcal{C}(d) < +\infty$. Let ϵ satisfying $0 < \epsilon < \|c\|_*$ be given. Suppose that the ellipsoid algorithm is applied to solve $HP(d)$, and is initiated with the Euclidean unit ball centered at $(w^0, \theta^0) = (0, 0)$. Then the ellipsoid algorithm will compute an ϵ-optimal solution of $HP(d)$ (and hence, by transformation, to $CP(d)$) in at most*

$$\left\lceil 8(n+2)^2 \ln\left(\frac{3\mathcal{C}(d)}{\bar{\tau}} \frac{\|c\|_*}{\epsilon}\right)\right\rceil$$

iterations, where each iteration must perform at most $\big((n+1)^2 + \max\{2n, m+mn+S_2\}\big)$ *operations and where* S_2 *is the number of operations needed to perform a feasibility cut on* C_Y*.*

Proof of Theorem 5.1. Let

$$a_1 = \frac{6}{\min\{\tau, \bar{\tau}\}}, \qquad a_2 = \frac{4}{\min\{\tau, \bar{\tau}\}}, \qquad \text{and} \quad a_3 = \frac{6}{\min\{\tau, \bar{\tau}\}} \ . \tag{27}$$

Then, from (12), we have that there exists $\hat{x}$ and $r > 0$ such that $B(\hat{x}, r) \subset X_d$, and

$$\frac{1}{r} \le a_1 \mathcal{C}(d), \qquad \|\hat{x}\| \le a_2 \mathcal{C}(d) \ , \quad \text{and} \quad \frac{\|\hat{x}\|}{r} \le a_3 \mathcal{C}(d) \ . \tag{28}$$

Applying Lemma 3.1, X_d^ϵ contains a ball $B(\bar{x}, \bar{r})$ with the following properties:

$$\frac{1}{\bar{r}} \le \frac{\max\{\epsilon, z^*(d) - c^T\hat{x} + r\|c\|_*\}}{\epsilon r} \qquad \text{and} \quad \|\bar{x}\| \le \max\{\|\hat{x}\|, \|x^*\|\} \ , \tag{29}$$

where x^* is any optimal solution of $CP(d)$. Furthermore, from (10) and (11), we have $|z^*(d)| \le \|c\|_* \mathcal{C}(d)$ and $\|x^*\| \le \mathcal{C}(d)^2$.

Examining the first inequality of (29), notice that

$$\frac{\max\{\epsilon, z^*(d) - c^T\hat{x} + r\|c\|_*\}}{\epsilon r} \ge \frac{\|c\|_*}{\epsilon} \ge 1 \ .$$

If $\bar{r} > 1$, we can reset $\bar{r} = 1$ and (29) will still hold. Therefore, there is no loss of generality in assuming that $\bar{r} \le 1$.

The dimension in which the ellipsoid algorithm is implemented is $n+1$. Let H_d^ϵ denote the set of ϵ-optimal solutions of $HP(d)$, and so H_d^ϵ is the image of X_d^ϵ under the transformation $h(\cdot)$ of (20). Then, from the ellipsoid algorithm theorem (16), the algorithm will compute an ϵ-optimal solution of $HP(d)$ in at most

$$\left\lceil 2(n+2) \ln\left(\frac{\operatorname{vol}(B^{n+1})}{\operatorname{vol}(B^{n+1} \cap H_d^\epsilon)} \right) \right\rceil \tag{30}$$

iterations, where B^{n+1} is the $(n+1)$-dimensional Euclidean unit ball.

Now let $T = X_d^\epsilon$. Then $H_d^\epsilon = h(T)$ and $B(\bar{x}, \bar{r}) \subset X_d^\epsilon$. Furthermore, $\bar{r} \le 1$ from the comments above. We therefore can apply Lemma 4.1 to bound the logarithm term of (30):

$$\ln\left(\frac{\operatorname{vol}\left(B^{n+1}\right)}{\operatorname{vol}\left(B^{n+1} \cap H_d^\epsilon\right)} \right) \le (n+1) \ln\left(2 + \frac{3(\|\bar{x}\| + 1)}{\bar{r}} \right) + \left[\ln\left(\|\bar{x}\|\right)\right]^+ \ . \tag{31}$$

We now bound the relevant quantities in (31) in order to obtain the desired bound on (30).

From (29) we have

$$\begin{aligned}
\frac{\|\bar{x}\|}{\bar{r}} \quad &\le \quad \frac{1}{\epsilon} \max\{\|\hat{x}\|, \|x^*\|\} \max\left\{ \frac{\epsilon}{r} \ , \frac{z^*(d) - c^T\hat{x}}{r} + \|c\|_* \right\} \\
&\le \quad \frac{1}{\epsilon} \max\{\|\hat{x}\|, \|x^*\|\} \left(\max\left\{ \frac{\epsilon}{r} \ , \frac{z^*(d) - c^T\hat{x}}{r} \right\} + \|c\|_* \right) \\
&\le \quad \frac{1}{\epsilon} \max\left\{ \frac{\|\hat{x}\|}{r} \ , \frac{\|x^*\|}{r} \right\} \max\left\{\epsilon, z^*(d) - c^T\hat{x}\right\} + \frac{1}{\epsilon} \max\{\|\hat{x}\|, \|x^*\|\} \|c\|_* \ .
\end{aligned}$$

Substituting in the bounds from (28), (10), and (11) and recalling that $\epsilon \leq \|c\|_*$, we obtain from the above inequality

$$\begin{aligned}\frac{\|\bar{x}\|}{\bar{r}} &\leq \frac{1}{\epsilon}\max\left\{a_3\mathcal{C}(d), a_1\mathcal{C}(d)^3\right\}\max\left\{\|c\|_*, \|c\|_*\mathcal{C}(d) + \|c\|_* a_2\mathcal{C}(d)\right\} \\ &+ \frac{1}{\epsilon}\max\left\{a_2\mathcal{C}(d), \mathcal{C}(d)^2\right\}\|c\|_* \ ,\end{aligned}$$

whereby we obtain

$$\frac{\|\bar{x}\|}{\bar{r}} \leq \frac{\|c\|_*}{\epsilon}\mathcal{C}(d)^4\left[(1+a_2)\left(\max\{a_1, a_3\}\right) + a_2\right] \ . \tag{32}$$

From (29) we have

$$\begin{aligned}\frac{1}{\bar{r}} &\leq \frac{\max\{\epsilon, z^*(d) - c^T\hat{x} + r\|c\|_*\}}{\epsilon r} \\ &= \frac{1}{\epsilon}\max\left\{\frac{\epsilon}{r}\ , \frac{z^*(d) - c^T\hat{x}}{r} + \|c\|_*\right\} \\ &\leq \frac{1}{\epsilon}\left(\max\left\{\frac{\|c\|_*}{r}\ , \frac{z^*(d) - c^T\hat{x}}{r}\right\} + \|c\|_*\right) \\ &\leq \frac{1}{\epsilon}\left[\max\left\{\frac{\|c\|_*}{r}\ , \frac{\|c\|_*\mathcal{C}(d)}{r} + \frac{\|c\|_*\|\hat{x}\|}{r}\right\} + \|c\|_*\right] \qquad \text{(from (10))} \\ &\leq \frac{\|c\|_*}{\epsilon}\left[\max\left\{a_1\mathcal{C}(d), a_1\mathcal{C}(d)^2 + a_3\mathcal{C}(d)\right\} + 1\right] \qquad \text{(from (28))},\end{aligned}$$

and so

$$\frac{1}{\bar{r}} \leq \frac{\|c\|_*}{\epsilon}\mathcal{C}(d)^2\left(a_1 + a_3 + 1\right) \ . \tag{33}$$

We also have from (29) that

$$\|\bar{x}\| \leq \max\left\{\|\hat{x}\|, \|x^*\|\right\} \leq \max\left\{a_2\mathcal{C}(d), \mathcal{C}(d)^2\right\} \leq a_2\mathcal{C}(d)^2 \ . \tag{34}$$

Substituting (32), (33), and (34) into (31) and then substituting (31) into (30) yields the following iteration bound on the ellipsoid algorithm:

(35)

$$\begin{aligned}&\Bigg\lceil 2(n+2)\left[(n+1)\ln\left(2 + \frac{3\|c\|_*}{\epsilon}\mathcal{C}(d)^4\left(1 + a_1 + a_2 + a_3 + (1+a_2)\max\{a_1, a_3\}\right)\right)\right. \\ &\qquad\qquad\left. + \ln(a_2\mathcal{C}(d)^2)\right]\Bigg\rceil \ .\end{aligned}$$

Substituting (27) into (35), we obtain the following chain of upper bounds on the iteration bound:

$$\left\lceil 2(n+2)\left\{(n+1)\ln\left(2+\frac{141\|c\|_*}{\epsilon(\min\{\tau,\bar\tau\})^2}\mathcal{C}(d)^4\right)+\ln\left(\frac{4}{\min\{\tau,\bar\tau\}}\mathcal{C}(d)^2\right)\right\}\right\rceil$$
$$\le\left\lceil 2(n+2)^2\ln\left(\frac{143\|c\|_*}{\epsilon}\left(\frac{\mathcal{C}(d)}{\min\{\tau,\bar\tau\}}\right)^4\right)\right\rceil$$
$$\le\left\lceil 8(n+2)^2\ln\left(\frac{4\mathcal{C}(d)}{\min\{\tau,\bar\tau\}}\frac{\|c\|_*}{\epsilon}\right)\right\rceil .$$

The number of operations needed to perform an optimality cut in $HP(d)$ is at most $2n$, since an upper level set support vector for $g(w,\theta)$ at a feasible point $(\bar w,\bar\theta)$ of $HP(d)$ is computed as $(c,-(c^T\bar w/\bar\theta))$, and the number of operations needed to compute and test for feasibility of $b\theta - Aw\in C_Y$ is $(m+mn+S_2)$. ∎

The proofs of Theorems 5.2 and 5.3 are accomplished by slightly modifying the analysis in the proof of Theorem 5.1 as per the following remark.

REMARK 5.1. *Note in the proof of Theorem 5.1 that the ellipsoid algorithm iteration bound in* (35) *was derived based only on the following facts: the feasible region of $CP(d)$ contains a ball $B(\hat x,r)$ satisfying $\frac{1}{r}\le a_1\mathcal{C}(d)$, $\|\hat x\|\le a_2\mathcal{C}(d)$, and $\frac{\|\hat x\|}{r}\le a_3\mathcal{C}(d)$; $|z^*(d)|\le\|c\|_*\mathcal{C}(d)$; and there exists an optimal solution x^* of $CP(d)$ satisfying $\|x^*\|\le\mathcal{C}(d)^2$.*

This remark will be used in the proofs of Theorems 5.2 and 5.3, which we now do in reverse order.

Proof of Theorem 5.3. Let

$$a_1=\frac{3}{\bar\tau},\quad a_2=\frac{3}{\bar\tau},\quad \text{and}\quad a_3=\frac{2}{\bar\tau}\ . \tag{36}$$

Then from (14) we know that the feasible region of $CP(d)$ contains a ball $B(\hat x,r)$ satisfying $\frac{1}{r}\le a_1\mathcal{C}(d)$, $\|\hat x\|\le a_2\mathcal{C}(d)$, and $\frac{\|\hat x\|}{r}\le a_3\mathcal{C}(d)$. Also, from (10), we have $|z^*(d)|\le\|c\|_*\mathcal{C}(d)$. Furthermore, from (11), there exists an optimal solution x^* of $CP(d)$ satisfying $\|x^*\|\le\mathcal{C}(d)^2$. Then, from Remark 5.1, the iteration bound of (35) is valid with values of a_1,a_2, and a_3 from (36). Substituting (36) into (35) yields the following iteration bound:

$$\left\lceil 2(n+2)\left\{(n+1)\ln\left(2+\frac{63\|c\|_*}{\epsilon\bar\tau^2}\mathcal{C}(d)^4\right)+\ln\left(\frac{3}{\bar\tau}\mathcal{C}(d)^2\right)\right\}\right\rceil$$
$$\le\left\lceil 2(n+2)^2\ln\left(\frac{65\|c\|_*}{\epsilon}\left(\frac{\mathcal{C}(d)}{\bar\tau}\right)^4\right)\right\rceil$$
$$\le\left\lceil 8(n+2)^2\ln\left(\frac{3\mathcal{C}(d)}{\bar\tau}\frac{\|c\|_*}{\epsilon}\right)\right\rceil . \qquad ∎$$

Proof of Theorem 5.2. The feasible region of $CP(d)$ lies in the affine set $\{x\in\Re^n\mid Ax=b\}$. In order to apply the ellipsoid algorithm conveniently, we construct a Euclidean-norm-preserving linear transformation to $\Re^{(n-m)}$. For concreteness, we assume with no loss of generality that A is an $m\times n$ real matrix. Let F be an $(n-m)\times n$ matrix whose rows form an orthonormal basis for the null space of A, and let $g = A^T(AA^T)^{-1}b$, where $\mathcal{C}(d) < +\infty$ implies that $\operatorname{rank}(A) = m$ and so F and g are well defined. Then the following problems are equivalent under the invertible linear transformations $s = Fx, x = F^T s + g$ between $\{x \in \Re^n \mid Ax = b\}$ and $\Re^{(n-m)}$:

$$CP(d): \quad \begin{array}{ll} \text{maximize} & c^T x \\ \text{s.t.} & Ax = b, \\ & x \in C_X, \end{array} \qquad Q: \quad \begin{array}{ll} \text{maximize} & c^T F^T s + c^T g \\ \text{s.t.} & F^T s + g \in C_X. \end{array}$$

Let

$$a_1 = \frac{3}{\tau}, \quad a_2 = \frac{4}{\tau}, \quad \text{and} \quad a_3 = \frac{3}{\tau}\ . \tag{37}$$

Then, from (13), we know that there exists $\hat{x}$ and r for which $A\hat{x} = b$ and $B(\hat{x}, r) \subset C_X$, and that satisfies $\frac{1}{r} \le a_1\mathcal{C}(d)$, $\|\hat{x}\| \le a_2\mathcal{C}(d)$, and $\frac{\|\hat{x}\|}{r} \le a_3\mathcal{C}(d)$. If we let $\hat{s} := F\hat{x}$, then it is straightforward to show that $B(\hat{s}, r)$ is contained in the feasible region of Q and that $\|\hat{s}\| \le \|\hat{x}\| \le a_2\mathcal{C}(d)$, and $\frac{\|\hat{s}\|}{r} \le a_3\mathcal{C}(d)$, where $\|s\| = \|s\|_2$ for $s \in \Re^{n-m}$ and $B(s, r)$ is the Euclidean ball centered at $s \in \Re^{n-m}$ with radius r. Let z_Q denote the optimal objective function value of Q, and let x^* denote an optimal solution of $CP(d)$. Then one can also easily show that $z_Q = z^*(d)$, and so $|z_Q| = |z^*(d)| \le \|c\|_*\mathcal{C}(d)$ from (10). Furthermore, let $s^* := Fx^*$. Then it is easy to show that s^* is an optimal solution of Q and $\|s^*\| \le \|x^*\| \le \mathcal{C}(d)^2$ from (11). Then, from Remark 5.1, the iteration bound of (35) is valid for the program Q with values of a_1, a_2, and a_3 from (37) and with the dimension n replaced by $n - m$. Substituting (37) into (35) yields the following iteration bound:

$$\left\lceil 2(n-m+2)\left\{(n-m+1)\ln\left(2 + \frac{78\|c\|_*}{\epsilon\tau^2}\mathcal{C}(d)^4\right) + \ln\left(\frac{4}{\tau}\mathcal{C}(d)^2\right)\right\}\right\rceil$$

$$\le \left\lceil 2(n-m+2)^2 \ln\left(\frac{80\|c\|_*}{\epsilon}\left(\frac{\mathcal{C}(d)}{\tau}\right)^4\right)\right\rceil$$

$$\le \left\lceil 8(n-m+2)^2 \ln\left(\frac{3\mathcal{C}(d)}{\tau}\frac{\|c\|_*}{\epsilon}\right)\right\rceil. \qquad \square$$

6. Further issues: Applications to other volume-reducing cutting-plane algorithms; testing for ϵ-optimality; testing for infeasibility; skewness of the ellipsoids.

Applications to other volume-reducing cutting-plane algorithms. The ellipsoid algorithm belongs to a larger class of efficient volume-reducing cutting-plane algorithms that includes the method of centers of gravity [11], the method of inscribed ellipsoids [10], and the method of volumetric centers [22], among others. Here we discuss how our analysis of the ellipsoid algorithm can be easily extended to these other methods. To keep the discussion simple, we focus on the class of instances of $CP(d)$, where C_X and C_Y are both regular cones.

Consider the strategy of applying either the method of centers of gravity or the method of inscribed ellipsoids to solve $CP(d)$ by solving $HP(d)$, starting with the unit ball B^{n+1} in $\Re^{n+1}$ (centered at the origin) and with the goal of computing an ϵ-optimal solution to $HP(d)$ and hence to $CP(d)$ as well. Because both of these methods achieve an (absolute) constant reduction in volume at each iteration, the iteration complexity of each of these methods will be $O(\ \ln(\frac{\mathrm{vol}\left(B^{n+1}\right)}{\mathrm{vol}\left(B^{n+1}\cap H_d^{\epsilon}\right)}))$ in order to find an ϵ-optimal solution of $CP(d)$. Now notice that a slight rearrangement of the proof of Theorem 5.1 yields the following inequality:

$$\ln\left(\frac{\mathrm{vol}\left(B^{n+1}\right)}{\mathrm{vol}\left(B^{n+1}\cap H_d^{\epsilon}\right)}\right)\leq 4(n+2)\ln\left(\frac{4\mathcal{C}(d)}{\min\{\tau,\bar{\tau}\}}\frac{\|c\|_*}{\epsilon}\right). \tag{38}$$

Therefore, the iteration complexity of these two methods is $O(n\ln(\frac{\mathcal{C}(d)}{\min\{\tau,\bar{\tau}\}}\frac{\|c\|_*}{\epsilon}))$.

The analysis of the method of volumetric centers is roughly the same as above; this method also achieves a constant reduction in volume at each iteration. However, the volumetric centers method must be initiated with a polytope as opposed to a Euclidean ball. Suppose we endow $X=\Re^n$ with the L_∞ norm rather than the Euclidean norm and that we apply the method of volumetric centers to solve $HP(d)$ initiated at the unit cube C^{n+1} in $\Re^{n+1}$. Then an identical version of (38) can be proved with B^{n+1} replaced by C^{n+1}, and so one can prove that the method of volumetric centers also has iteration complexity $O(n\ln(\frac{\mathcal{C}(d)}{\min\{\tau,\bar{\tau}\}}\frac{\|c\|_*}{\epsilon}))$. We also point out that the method of volumetric centers requires fewer total arithmetic operations than the ellipsoid algorithm.

Similar results can be derived for the two other classes of instances of $CP(d)$. For a more thorough discussion of the complexity of volume-reducing cutting-plane methods, see [12].

Testing for ϵ-optimality by solving the dual problem. One uncomfortable fact about Theorems 5.1, 5.2, and 5.3 is that while the ellipsoid algorithm is guaranteed to find an ϵ-approximate solution of $CP(d)$ in the stated complexity bounds of these theorems, the quantities in the bounds may be unknown (one may know the relevant widths of the cones, but in all likelihood the condition number $\mathcal{C}(d)$ is unknown), and so one does not know when an ϵ-approximate solution of $CP(d)$ has been found. An obvious strategy for overcoming this difficulty is to solve the primal and the dual problems in parallel, and then test at each iteration (of each algorithm) if the best primal and dual solutions obtained so far satisfy a duality gap of at most ϵ. Because of the natural symmetry in format of the dual pair of problems $CP(d)$ and $CD(d)$, one can obtain complexity results for solving the dual problem $CD(d)$ that exactly parallel those of Theorems 5.1, 5.2, and 5.3, where the quantities $\|c\|_*, n, \tau$, and $\bar{\tau}$ are replaced by $\|b\|, m, \bar{\tau}^*$, and τ^*, respectively, and where the cones C_X and C_Y are replaced by C_Y^* and C_X^* in the statements of the complexity results. One also must assume that $Y^*=\Re^m$ and that the norm $\|y\|_*$ on $\Re^m$ is the Euclidean norm.

Testing for infeasibility. If one is not sure whether $CP(d)$ has a feasible solution, the ellipsoid algorithm can be run to test for infeasibility of the primal problem (in parallel with attempting to solve $CP(d)$). This can be accomplished as follows. First, assume that the dual space $Y^*=\Re^m$ is endowed with the Euclidean norm $\|y\|_2$. Second, note that $CP(d)$ has no feasible solution if the "alternative" system,

$$\begin{aligned} AP(d):\quad A^Ty&\in C_X^*,\\ y&\in C_Y^*,\\ y^Tb&<0, \end{aligned}$$

has a solution. Define the following "alternative" set:

$$Y_d=\{y\in Y^*\mid A^Ty\in C_X^*, y\in C_Y^*, y^Tb\leq 0\}. \tag{39}$$

Suppose $CP(d)$ has no feasible solution. Then, as special cases of Theorems 5.2, 5.4, and 5.6 of [7], Y_d must contain an inscribed Euclidean ball $B_2(\hat{y}, r)$ (or a disk in the vector subspace $\{y \in \Re^m \mid A^T y = 0\}$ if $C_X = X$) such that $\|\hat{y}\|_2 + r \leq 1$ (and so $B_2(\hat{y}, r)$ is contained in the unit Euclidean ball) and such that

$$\text{(i)} \quad r \geq \frac{\min\{\tau^*, \bar{\tau}^*\}}{4\mathcal{C}_P(d)} \qquad \text{when } C_X \text{ and } C_Y \text{ are both regular,}$$

$$\text{(ii)} \quad r \geq \frac{\tau^*}{2\mathcal{C}_P(d)} \qquad \text{when } C_X \text{ is regular and } C_Y = \{0\},$$

$$\text{(iii)} \quad r \geq \frac{\min\{\bar{\tau}^*, \bar{\tau}\}}{4\mathcal{C}_P(d)} \qquad \text{when } C_X = X \text{ and } C_Y \text{ is regular.}$$

These results can then be used to demonstrate that an upper bound on the number of iterations needed to find a solution of $AP(d)$ using the ellipsoid algorithm starting with the Euclidean unit ball in $\Re^m$ (or the unit disk in the vector subspace $\{y \in \Re^m \mid A^T y = 0\}$ if $C_X = X$) is

$$\text{(i):} \quad O\left(m^2 \ln\left(\frac{\mathcal{C}_P(d)}{\min\{\tau^*, \bar{\tau}^*\}}\right)\right) \qquad \text{when } C_X \text{ and } C_Y \text{ are both regular,}$$

$$\text{(ii):} \quad O\left(m^2 \ln\left(\frac{\mathcal{C}_P(d)}{\tau^*}\right)\right) \qquad \text{when } C_X \text{ is regular and } C_Y = \{0\},$$

$$\text{(iii):} \quad O\left((m-n)^2 \ln\left(\frac{\mathcal{C}_P(d)}{\min\{\bar{\tau}^*, \bar{\tau}\}}\right)\right) \qquad \text{when } C_X = X \text{ and } C_Y \text{ is regular.}$$

Bounding the skewness of the ellipsoids in the ellipsoid algorithm. Let $E_{\bar{x},Q} = \{x \in X \mid (x-\bar{x})^T Q^{-1}(x-\bar{x}) \leq 1\}$ be an ellipsoid centered at the point $\bar{x}$, where Q is a positive definite matrix. The skewness of $E_{\bar{x},Q}$ is defined to be the ratio of the largest to the smallest eigenvalue of the matrix Q defining $E_{\bar{x},Q}$, and so the skewness also corresponds to the traditional condition number of the matrix Q. The skewness of the ellipsoids generated in an application of the ellipsoid algorithm determines the numerical stability of the ellipsoid algorithm, since each iteration of the ellipsoid algorithm uses the current value of Q^{-1} to update the center $\bar{x}$ of the ellipsoid and to perform a rank-one update of Q^{-1}; see [3], for example. Furthermore, one can show that the logarithm of the skewness of the ellipsoid computed at a given iteration is sufficient to specify the numerical precision requirements of the ellipsoid algorithm at that iteration. Herein, we provide an upper bound on the skewness of all of the ellipsoids computed in the ellipsoid algorithm as a function of the condition number $\mathcal{C}(d)$ of $CP(d)$.

The skewness of the unit ball (which is used to initiate the ellipsoid algorithm herein) is 1. From the formula for updating the ellipsoids encountered in the ellipsoid algorithm at each iteration, the skewness increases by at most $(1 + \frac{2}{k-1})$ at each iteration, where k is the dimension of the space in which the ellipsoid algorithm is implemented. Therefore, the skewness of the ellipsoid at iteration j is bounded above by $(1 + \frac{2}{k-1})^j$. Let us consider the class of instances defined for Theorem 5.1, for example, and let J be the (unrounded) iteration bound for the ellipsoid algorithm from Theorem 5.1, namely,

$$J = 8(n+2)^2 \ln\left(\frac{4\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \frac{\|c\|_*}{\epsilon}\right), \tag{40}$$

and assume for simplicity of exposition that J is an integer. Let $(\text{Skew})_j$ denote the skewness of the ellipsoid computed in the ellipsoid algorithm at iteration j. Then, for this class of instances, we have $k = n + 1$, whereby

$$(\text{Skew})_J \leq \left(1 + \frac{2}{n}\right)^J = \left(e^{\left(\ln\left(1+\frac{2}{n}\right)\right)}\right)^J = e^{J\left(\ln\left(1+\frac{2}{n}\right)\right)} = \left(e^J\right)^{\left(\ln\left(1+\frac{2}{n}\right)\right)}. \tag{41}$$

Substituting for (40) in (41), we obtain

$$(\text{Skew})_J \leq \left(\frac{4\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \frac{\|c\|_*}{\epsilon}\right)^{8(n+2)^2 \ln\left(1+\frac{2}{n}\right)}.$$

However, the exponent in the above expression is bounded above by $45n$ for $n \geq 2$ (actually, it is bounded above by $17n$ for large $n \geq 49$), and we have

$$(\text{Skew})_J \leq \left(\frac{4\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \frac{\|c\|_*}{\epsilon}\right)^{45n}.$$

Taking logarithms, we can rewrite this bound as

$$\ln (\text{Skew})_J \leq 45n \ln \left(\frac{4\mathcal{C}(d)}{\min\{\tau, \bar{\tau}\}} \frac{\|c\|_*}{\epsilon}\right). \tag{42}$$

Therefore, the logarithm of the skewness of the ellipsoids encountered in the ellipsoid algorithm grows at most linearly in the logarithm of the condition number $\mathcal{C}(d)$. Also, the bound in (42) specifies the sufficient numerical precision requirements for the ellipsoid algorithm (in terms of $\ln(\mathcal{C}(d))$ and other quantities) because the logarithm of the skewness is sufficient to specify such requirements. This is similar to the results on numerical precision presented in [25] for an interior-point method for linear programming.

Finally, the above reasoning can be used to obtain similar bounds on the skewness for the other two classes of instances of $CP(d)$.

REFERENCES

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[2] M. Bazaraa, H. Sherali, and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms*, 2nd ed., John Wiley, New York, 1993.

[3] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, MA, 1997.

[4] R. Bland, D. Goldfarb, and M. Todd, *The ellipsoid method: A survey*, Oper. Res., 29 (1981), pp. 1039–1091.

[5] A. Charnes and W. Cooper, *Programming with linear fractionals*, Naval Res. Quarterly, 9 (1962), pp. 181–186.

[6] B. Craven and B. Mond, *The dual of a fractional linear program*, J. Math. Anal. Appl., 42 (1973), pp. 507–512.

[7] R. Freund and J. Vera, *Some Characterizations and Properties of the "Distance to Ill-Posedness" and the Condition Number of a Conic Linear System*, Technical Report W.P. #3862-95-MSA, MIT Sloan School of Management, 1995; Math. Programming, to appear.

[8] M. Grötschel, L. Lovasz, and A. Schrijver, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica, 1 (1981), pp. 169–197.

[9] L. Khachiyan, *A polynomial algorithm in linear programming*, Soviet Math. Dokl., 20 (1979), pp. 191–194.

[10] L. Khachiyan, S. Tarasov, and I. Erlikh, *The method of inscribed ellipsoids*, Soviet Math. Dokl., 37 (1988), pp. 226–230.
[11] A. Levin, *On an algorithm for the minimization of convex functions*, Soviet Math. Dokl., 6 (1965), pp. 286–290.
[12] T. Magnanti and G. Perakis, *A unifying geometric solution framework and complexity analysis for variational inequalities*, Math. Programming, 71 (1995), pp. 327–351.
[13] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
[14] Y. Nesterov and M. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM J. Optim., 8 (1998), pp. 324–364.
[15] Y. Nesterov and M. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Math. Oper. Res., 22 (1997), pp. 1–42.
[16] Y. Nesterov, M. Todd, and Y. Ye, *Primal-Dual Methods and Infeasibility Detectors for Nonlinear Programming Problems*, Technical Report TR1156, Dept. of IE&OR, Cornell University, Ithaca, NY, 1996.
[17] J. Renegar, *Some perturbation theory for linear programming*, Math. Programming, 65 (1994), pp. 73–91.
[18] J. Renegar, *Incorporating condition measures into the complexity theory of linear programming*, SIAM J. Optim., 5 (1995), pp. 506–524.
[19] J. Renegar, *Linear programming, complexity theory, and elementary functional analysis*, Math. Programming, 70 (1995), pp. 279–351.
[20] J. Renegar, *Condition numbers, the barrier method, and the conjugate gradient method*, SIAM J. Optim., 6 (1996), pp. 879–912.
[21] N. Shor, *Cut off methods with space extension in convex programming problems*, Cybernetics, 13 (1977), pp. 94–96.
[22] P. Vaidya, *A new algorithm for minimizing convex functions over convex sets*, in Proceedings of the 30th IEEE Symposium on Foundations of Computer Science, IEEE Computer Soc. Press, Los Alamitos, CA, 1989, pp. 338–343.
[23] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Review, 38 (1996), pp. 49–95.
[24] J. Vera, *Ill-posedness and the complexity of deciding existence of solutions to linear programs*, SIAM J. Optim., 6 (1996), pp. 549–569.
[25] J. Vera, *On the complexity of linear programming under finite precision arithmetic*, Math. Programming, 80 (1998), pp. 91–123.
[26] D. Yudin and A. Nemirovskii, *Informational complexity and efficient methods for solving complex extremal problems*, Ekonom. i Matem. Metody, 12 (1976), pp. 357–369.