The extremal volume ellipsoids of convex bodies, their symmetry properties, and their determination in some special cases

Osman Güler and Filiz Gürtuna[*]

September 2007

Abstract

A convex body K in $\mathbb{R}^n$ has associated with it a unique circumscribed ellipsoid $CE(K)$ with minimum volume, and a unique inscribed ellipsoid $IE(K)$ with maximum volume. We first give a unified, modern exposition of the basic theory of these extremal ellipsoids using the semi–infinite programming approach pioneered by Fritz John in his seminal 1948 paper. We then investigate the automorphism groups of convex bodies and their extremal ellipsoids. We show that if the automorphism group of a convex body K is large enough, then it is possible to determine the extremal ellipsoids $CE(K)$ and $IE(K)$ exactly, using either semi–infinite programming or nonlinear programming. As examples, we compute the extremal ellipsoids when the convex body K is the part of a given ellipsoid between two parallel hyperplanes, and when K is a truncated second order cone or an ellipsoidal cylinder.

[*]Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250, USA. E-mail: {*guler,gurtuna1*}*@math.umbc.edu*. Research partially supported by the National Science Foundation under grant DMS–0411955

1   Introduction

A *convex body* in $\mathbb{R}^n$ is a compact convex set with nonempty interior. Let K be a convex body in $\mathbb{R}^n$. Among the ellipsoids circumscribing K, there exists a unique one with minimum volume and similarly, among the ellipsoids inscribed in K, there exists a unique one of maximum volume. These are called the *minimal circumscribed ellipsoid* and *maximal inscribed ellipsoid* of K, and we denote them by $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$, respectively. To our knowledge, Behrend [5] is the first person to investigate these problems, and proves the existence and uniqueness of the two ellipsoids in the plane, that is when $n = 2$. In the general case, the existence of either ellipsoid is easy to prove using compactness. The ellipsoid $\mathrm{CE}(K)$ is often referred to as *Löwner ellipsoid* since Löwner has used its uniqueness in his lectures, see [10]. The uniqueness of $\mathrm{CE}(K)$ also follows from the famous paper of John [18] although he does not state it explicitly. Subsequently, Danzer, Laugwitz, and Lenz [11], and Zaguskin [39] prove the uniqueness of both ellipsoids. The inscribed ellipsoid problem $\mathrm{IE}(K)$ is often called *John ellipsoid*, and sometimes *Löwner–John* ellipsoid, especially in Banach space geometry literature. This designation may seem inappropriate, since John does not consider the problem $\mathrm{IE}(K)$ in [18]. However, in Banach space geometry literature, one is mainly interested in symmetric convex bodies ($K = -K$), and for this class of convex bodies the inscribed and circumscribed ellipsoids are related by polarity, so that results about one ellipsoid may be translated into a similar statement about the other.

The two extremal problems $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$ are important in several fields including optimization, Banach space geometry, and statistics. They also have applications in differential geometry [24], Lie group theory [11], and symplectic geometry, among others. The ellipsoid algorithm of Khachian [17] for linear programming sparked general interest of optimizers in the circumscribed ellipsoid problem. At the kth step of this algorithm, one has an ellipsoid $E^{(k)}$ and is interested in finding $E^{(k+1)} = \mathrm{CE}(K)$ where K is the intersection of $E^{(k)}$ with a half plane whose bounding hyperplane passes through the center of $E^{(k)}$. The ellipsoid $E^{(k+1)}$ can be computed explicitly, and the ratio of the volumes $\mathrm{vol}(E^{(k+1)})/\mathrm{vol}(E^{(k)})$ determines the rate of convergence of the algorithm and gives its polymomial–time complexity. From this perspective, ellipsoids covering the intersection of an ellipsoid with two halfspaces whose bounding hyperplanes are parallel have been studied as well, since the resulting ellipsoid algorithms are likely to have faster convergence. The papers of König and Pallaschke [23] and Todd [36] compute the circumscribing ellipsoid explicitly. It has been discovered that the circumscribed ellipsoid problem is dually related to the *optimal design* problem in statistics. Consequently, there has been wide interest in the problem $\mathrm{CE}(K)$ and related problems in this community as well. Inscribed ellipsoid problems also arise in optimization. For example, the *inscribed ellipsoid method* of Tarasov, Khachian, and Erlikh [35] is a polynomial–time algorithm for solving general convex programming problems. In this method, one needs to compute numerically an approximation to the inscribed ellipsoid of a polytope which is described by a set of linear inequalities.

To meet the demand from different fields such as optimization, computer science, engineering, and statistics, there have been many algorithms proposed to numerically compute the two extremal ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$. Recently, there has been a surge of research activity in this subject. We do not discuss algorithms in this paper, but the interested reader can find more information on this topic and the relevant references in the papers [41], [34], and [38].

This paper has three, related goals. In §2 and §3, we give semi–infinite programming formulations of the problems $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$, respectively, and then describe their fun-

damental properties using the resulting optimality conditions. There are essentially no new results in these sections. We include them in order to fill a gap in the optimization literature, and we also use some of the results in these sections later on in the paper. There exists a sizable literature in Banach space theory dealing with the existence, uniqueness, and basic properties of the extremal ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$. They use (necessarily) optimality considerations to arrive at the results they are interested in, but this is done in an informal manner. In fact, many of the papers in this literature use the reformulation of the ellipsoid problems given in the interesting paper of Lewis [26]; see also [29], [37]. There have been some exceptions recently. See for example the interesting papers Gordon et al. [13] and Klartag [22]. To our knowledge, the only papers which systematically use optimization techniques to prove results about the extremal ellipsoids are the original paper of John [18] and the papers of Juhnke [19, 20, 21]. What we offer in § 2 and §3 is a careful, unified, and modern synthesis of the basic results on the ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$ in the mainstream optimization literature. We hope that it serves as a useful introduction to the subject in the optimization community.

We devote §4 to the symmetry properties of convex bodies and the related symmetry properties of the corresponding ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$. One way to formalize the symmetry properties of a convex body K is to consider its *(affine) automorphism group* $\mathrm{Aut}(K)$. It will be seen that the uniqueness of the two ellipsoids imply that the ellipsoids *inherit* the symmetry properties of the underlying convex body K. That is, $\mathrm{Aut}(K)$ is contained in the automorphism group of the two extremal ellipsoids. One consequence of this is that if K is "symmetric" enough, then either it is possible to anayically compute the extremal ellipsoids exactly, or else it is possible to reduce the complexity of their numerical computation. In this section, we demonstrate the former possibility for a class of convex bodies K whose automorphism groups act *transitively* on $\mathrm{ext}(K)$, the extreme points of K. Davies [12] shows that for this class of convex bodies, the center of gravity, the center of $\mathrm{CE}(K)$, and the center of $\mathrm{IE}(K)$ all coincide, and this center can be obtained explicitly as a Haar integral over the automorphism group $\mathrm{Aut}(K)$. We show that the matrix X in the circumscribed ellipsoid $\mathrm{CE}(K)$ can also be obtained in a similar manner. We remark that there has been a continuous interest on symmetric convex bodies since antiquity, and the class of *regular polytopes* [9], including the Platonic solids, is a subclass of symmetric convex bodies in the sense of Davies. Because of space considerations, we are not able to pursue the study of the automorphism groups of convex bodies in greater depth in this paper. We plan to explore this subject in future papers.

In the rest of the paper, starting with §5, we exploit automorphism groups to analytically compute the extremal ellipsoids of two classes of convex bodies. The first class consists of convex bodies which are the intersections of a given ellipsoid with two halfspaces whose bounding hyperplanes are parallel, and have been mentioned above. We call such convex bodies *slabs*. The second class consists of convex bodies obtained by taking the convex hull of the intersection of the same parallel hyperplanes with the ellipsoid. We note that a convex body in this class is either a truncated second order cone or an ellipsoidal cylinder, depending on the location of the bounding hyperplanes with respect to the center of the ellipsoid.

In §5, we compute the autormorphism group of a slab K and use it to determine the form of the center and matrix of its ellipsoid $\mathrm{CE}(K)$. Although the automorphism group $\mathrm{Aut}(K)$ is not large enough to compute the ellipsoid $\mathrm{CE}(K)$ exactly, it is large enough to reduce its determination to computing just three parameters (instead of $n(n+3)/2$ in the general case), one to determine its center and two to determine its matrix.

In §6, we formulate the CE(K) problem for a slab as a semi–infinite programming problem, and obtain its solution by computing the three parameters of CE(K) directly from the Fritz John optimality conditions for the semi–infinite program. As we mentioned already, König and Pallaschke [23] and Todd [36] solve this exact problem. König and Pallaschke's approach is similar to ours: they use the uniqueness and invariance properties of the ellipsoid CE(K). However, their solution is not complete since they only consider the cases when the slab does not contain the center of the given ellipsoid. Todd gives a complete proof covering all cases. His proof is based on guessing the optimal ellipsoid and then proving its minimality by using some bounds on the volume of a covering ellipsoid. In §6, we also formulate the ellipsoid problem CE(K) as a nonlinear programming problem, and give a second, independent solution for it.

For interesting applications of the ellipsoid CE(K) of a slab, see the papers [27] and [3].

In §7, we formulate the IE(K) problem for a slab as a semi–infinite programming problem, and obtain its solution directly from the resulting Fritz John optimality conditions. We also formulate the same problem as a nonlinear programming problem, but do not provide its solution in order to keep the the length of the paper within reasonable bounds.

Finally, in §8, we formulate the CE(K) problem for a convex body from the second class of convex bodies mentioned above as a semi–infinite programming problem, and obtain its solution directly from Fritz John optimality conditions. The form of the optimal ellipsoid CE(K) turns out to be very similar to the form of the corresponding ellipsoid for a slab. We do not solve the inscribed ellipsoid problem for the second class of convex bodies for space considerations.

We remark that the ideas and techniques used in this paper for determining the extremal ellipsoids for specific classes of convex bodies can be generalized to other classes of convex bodies as long as these bodies have large enough automorphism groups. It is reasonable to expect that automorphism groups can also be used advantageously in numerical determination of extremal ellipsoids.

Our notation is fairly standard. We denote the set of symmetric $n \times n$ matrices by $\mathbb{SR}^{n \times n}$. In $\mathbb{R}^n$, we use the bracket notation for inner products, thus $\langle u, v \rangle = u^T v$. In the vector space $\mathbb{R}^{n \times n}$ of $n \times n$ matrices (and hence in $\mathbb{SR}^{n \times n}$), we use the trace inner product

$$\langle X, Y \rangle = \mathrm{tr}(XY^T).$$

If both inner products are used within the same equation, then the meaning of each inner product should be clear from the context. We define and use additional inner products in this paper, especially in §4. The sets ∂X and conv(X), and ext(X) denote the boundary and the convex hull of a set X in $\mathbb{R}^n$, respectively, and ext(K) is the set of extreme points of a convex set K in $\mathbb{R}^n$.

2 The minimum volume circumscribed ellipsoid problem

We recall that the circumscribed ellipsoid problem is the problem of finding a minimum volume ellipsoid circumscribing a convex body K in $\mathbb{R}^n$. This is the main problem treated in Fritz John [18]. In this paper, John shows that such an ellipsoid exists and is unique; we denote it by CE(K). John introduces semi–infinite programming and develops his optimality

conditions to prove the following deep result about the ellipsoid $CE(K)$: the ellipsoid with the same center as $CE(K)$ but shrunk by a factor n is contained in K, and if K is symmetric ($K = -K$), then $CE(K)$ needs to be shrunk by a smaller factor $\sqrt{n}$ to be contained in K. This fact is very important in the geometric theory of Banach spaces. In that theory, a symmetric convex body K is the unit ball of a Banach space, and if K is an ellipsoid, then the Banach space is a Hilbert space. Consequently, the shrinkage factor indicates how close the Banach space is to being a Hilbert space. In this context, it is not important to compute the exact ellipsoid $CE(K)$. However, in some convex programming algorithms, including the ellipsoid method and its variants, the exact or nearly exact ellipsoid $CE(K)$ needs to be computed. If K is sufficiently simple, $CE(K)$ can be computed analytically. In more general cases, numerical algorithms have been developed to approximately compute $CE(K)$.

In this section, we deal with the $CE(K)$ problem more or less following John's approach. However, in the interest of brevity and clarity, we use more modern notation and give new and simpler proofs for some of the technical results.

An ellipsoid E in $\mathbb{R}^n$ is an affine image of the unit ball $B_n := \{u \in \mathbb{R}^n : ||u|| \leq 1\}$, that is,

$$E = c + A(B_n) = \{c + Au : u \in \mathbb{R}^n, ||u|| = 1\} \subset \mathbb{R}^m, \tag{2.1}$$

where $A \in \mathbb{R}^{m \times n}$ is any $m \times n$ matrix. Here c is the center of E and the volume of E is given by $\mathrm{vol}(E) = \det(A)\,\mathrm{vol}(B_n)$. We are interested in the case where E is a solid body (E has a non–empty interior), hence we assume that A is a non–singular $n \times n$ matrix. Let A have the singular value decomposition $A = V_1 \Sigma V_2$ where V_1, V_2 are orthogonal $n \times n$ matrices and Σ is a diagonal matrix with positive elements. Then we have the polar decomposition of A, that is, $A = SO$ where $S = V_1 A V_1^T \in \mathbb{SR}^{n \times n}$ is positive definite and $O = V_1 V_2$ is an orthogonal matrix. Consequently, $E = c + SO(B_n) = c + S(B_n)$, that is, the matrix A in the definition of the ellipsoid E in (2.1) can be taken to be symmetric and positive definite, an assumption we make from here on. By making the change of variables $x := c + Au$, that is, $u = A^{-1}(x - c)$, and defining $X := A^{-2}$, the ellipsoid E in (2.1), $E = \{x \in \mathbb{R}^n : ||A^{-1}(x - c)||^2 \leq 1\}$ can also be written in the form

$$E = E(X, c) := \{x \in \mathbb{R}^n : \langle X(x - c),\, x - c \rangle \leq 1\}. \tag{2.2}$$

Note that we have

$$\mathrm{vol}(E) = \det(X)^{-1/2} \omega_n, \tag{2.3}$$

where $\omega_n = \mathrm{vol}(B_n)$.

Consequently, we can set up the circumscribed ellipsoid problem as a semi–infinite program

$$\begin{aligned} \min \quad & -\log \det X \\ \mathrm{s.\,t.} \quad & \langle X(y - c), y - c \rangle \leq 1, \quad \forall y \in K, \end{aligned} \tag{2.4}$$

in which the decision variables are $X \in \mathbb{SR}^{n \times n}$ and $c \in \mathbb{R}^n$.

There exists an ellipsoid of minimum volume circumscribing the convex body K. It suffices to prove that the set of feasible (X, c) in problem (2.4) is compact. Let K contain a ball of radius $r > 0$, and let $E = E(X, c)$ be an ellipsoid covering K. Note that the ball still lies in E if we shift its center to center c of the ellipsoid E. Thus, every vector u in $\mathbb{R}^n$, $||u|| = 1$ must satisfy the inequality $\langle Xu, u \rangle \leq 1/r^2$, that is, the eigenvalues of X are at most $1/r^2$. It follows from the spectral decomposition of symmetric matrices that the set of feasible X is compact. If the norm of the center c of the ellipsoid E circumscribing K goes

to infinity, then the volume of the ellipsoid goes to infinity as well. This proves that the set of feasible (X, c) is compact.

The following basic theorem of Fritz John [18] is one of our main tools in this paper. The book [16] develops optimality conditions for semi–infinite programming including this result and treats several problems from analysis and geometry using semi–infinite programming techniques.

Theorem 2.1. (*Fritz John*) *Consider the optimization problem*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x, y) \le 0, \quad \forall\, y \in Y, \end{aligned} \tag{2.5}$$

where $f(x)$ is a continuously differentiable function defined on an open set $X \subseteq \mathbb{R}^n$, and $g(x, y)$ and $\nabla_x g(x, y)$ are continuous functions defined on $X \times Y$ where Y is a compact set in some topological space. If x is a local minimizer of (2.5), then there exist at most n active constraints $\{g(x, y_i)\}_1^k$ $(g(x, y_i) = 0)$ and a non–trivial, non–negative multiplier vector $0 \neq (\lambda_0, \lambda_1, \ldots, \lambda_k) \geq 0$ such that

$$\lambda_0 \nabla f(x) + \sum_{i=1}^{k} \lambda_i \nabla_x g(x, y_i) = 0.$$

We now derive the optimality conditions for (2.4).

Theorem 2.2. *Let K be a convex body in $\mathbb{R}^n$. There exists an ellipsoid of minimum volume circumscribing K. If $E(X, c)$ is such an ellipsoid, then there exists a multiplier vector $\lambda = (\lambda_1, \ldots, \lambda_k) > 0$, $0 \le k \le n(n+3)/2$, and points $\{u_i\}_1^k$ in K such that*

$$\begin{aligned} X^{-1} &= \sum_{i=1}^{k} \lambda_i (u_i - c)(u_i - c)^T, \\ 0 &= \sum_{i=1}^{k} \lambda_i (u_i - c), \\ u_i &\in \partial K \cap \partial E(X, c), \quad i = 1, \ldots, k, \\ K &\subseteq E(X, c). \end{aligned} \tag{2.6}$$

We call the points $\{u_i\}_1^k$ in $\partial K \cap \partial E$ *contact points* of K and E.

Proof. The existence of a minimum volume circumscribed ellipsoid is already proved above. Let $E(X, c)$ be such an ellipsoid. The constraints in (2.4) are indexed by $y \in K$, a compact set, and Theorem 2.1 implies that there exists a non–zero multiplier vector $(\lambda_0, \lambda_1, \ldots, \lambda_k) \geq 0$, where $k \le n(n+1)/2 + n = n(n+3)/2$, $\lambda_i > 0$ for $i > 0$, and points $\{u_i\}_1^k$ in K such that the Lagrangian function

$$\begin{aligned} L(X, c, \lambda) &:= -\lambda_0 \log \det X + \sum_{i=1}^{k} \lambda_i \langle X(u_i - c), u_i - c \rangle \\ &= -\lambda_0 \log \det X + \langle X, \sum_{i=1}^{k} \lambda_i (u_i - c)(u_i - c)^T \rangle, \end{aligned}$$

5

where the inner product on the last line is the trace inner product on S^n, satisfies the optimality conditions

$$0 = \nabla_c L(X, c, \lambda) = X \sum_{i=1}^{k} \lambda_i (u_i - c),$$

$$0 = \nabla_X L(X, c, \lambda) = -\lambda_0 X^{-1} + \sum_{i=1}^{k} \lambda_i (u_i - c)(u_i - c)^T,$$

where we used the well–known fact that $\nabla_X \log \det X = X^{-1}$. If $\lambda_0 > 0$, then $0 = \text{tr}(\sum_{i=1}^{k} \lambda_i (u_i - c)(u_i - c)^T) = \sum_{i=1}^{k} \lambda_i \|u_i - c\|^2$. This implies that $\lambda_i = 0$ for all i, contradicting $\lambda \neq 0$. We let $\lambda_0 = 1$ without loss of generality, and arrive at the Fritz John conditions (2.6). $\square$

Remark 2.3. The contact points have applications in several fields, in optimal designs, and in estimating the size of almost orthogonal submatrices of orthogonal matrices [33], for example. Gruber [14] shows that "most" convex bodies K have the maximum number $n(n+3)/2$ of contact points. See also [33] for a simpler proof. Similar results also hold for the maximum volume inscribed ellipsoids discussed in §3, see [14]. However, Rudelson [33] shows that for every $\varepsilon > 0$ and every convex body K, there exists a nearby convex body L whose distance (Banach–Mazur distance) to K is less than $1 + \varepsilon$ and which at most $k \leq C(\varepsilon) \cdot n \log^3 n$ contact points. This has obvious implications for numerical algorithms that try to compute approximate covering ellipsoids.

Remark 2.4. Let $S = \text{ext}(K)$, the set of extreme points of K. We have $K = \text{conv}(S)$, the convex hull of S, by a Theorem of Minkowski, see Rockafellar [32], Corollary 18.5.1. Note that S and K have the same extremal covering ellipsoid, and applying Theorem 2.2 to S instead of K shows that we can choose $u_i \in S = \text{ext}(K)$, $i = 1, \ldots, k$.

An independent proof of the above fact runs as follows: let $x \in \partial K \cap \partial E$ be a contact point. Noting $\partial E = \text{ext}(E)$, we have $x \in \text{ext}(E)$. If $x \notin \text{ext}(K)$, then there exist $y, z \in K$, $y \neq z$, such that x lies in the interior of the line segment $[y, z]$. However, $x \in [y, z] \subseteq E$, contradicting the fact that $x \in \text{ext}(E)$.

The equation $\sum_{i=1}^{k} \lambda_i (u_i - c) = 0$ in (2.6) gives $c \in \text{conv}(\{u_i\}_1^k)$. This immediately implies

Corollary 2.5. *Let K be a convex body in $\mathbb{R}^n$. The contact points of $\text{CE}(K)$ are not contained in any closed halfspace whose bounding hyperplane passes through the center of $\text{CE}(K)$.*

For most theoretical purposes, we may assume that the optimal ellipsoid is the unit ball $E(I_n, 0)$. This can be accomplished by an affine change of the coordinates, if necessary. This results in the more transparent optimality conditions

$$I_n = \sum_{i=1}^{k} \lambda_i u_i u_i^T, \quad \sum_{i=1}^{k} \lambda_i u_i = 0, \tag{2.7}$$

$$u_i \in \partial K \cap \partial B_n, \quad i = 1, \ldots, k, \quad K \subseteq B_n.$$

Taking traces of both sides in the first equation above gives $n = \text{tr}(I_n) = \text{tr}(\sum_{i=1}^{k} \lambda_i u_i u_i^T) = \sum_{i=1}^{k} \lambda_i u_i^T u_i = \sum_{i=1}^{k} \lambda_i$, that is,

$$\sum_{i=1}^{k} \lambda_i = n. \tag{2.8}$$

In this section and in §3, convex duality will play an important role. If C is a convex body in $\mathbb{R}^n$, then the Minkowski *support function* of C is defined by

$$s_C(d) := \max_{u \in C} \langle d, u \rangle.$$

It is obviously defined on $\mathbb{R}^n$ and is a convex function since it is a maximum of linear functions indexed by u. In fact, $s_C = \delta_C^*$, where δ_C is the indicator function of C and $*$ denotes the Fenchel dual. If C and D are two convex bodies, it follows from Corollary 13.1.1 in [32] that $C \subseteq D$ if and only if $s_C \leq s_D$.

We compute

$$\begin{aligned}
s_{E(X,c)}(d) &= \max\left\{ \langle d, u \rangle : \langle X(u-c), u-c \rangle \leq 1 \right\} \\
&= \max\left\{ \langle d, c + X^{-1/2}v \rangle : ||v|| \leq 1 \right\} \\
&= \langle c, d \rangle + \max_{||v||=1} \langle X^{-1/2}d, v \rangle = \langle c, d \rangle + ||X^{-1/2}d|| \\
&= \langle c, d \rangle + \langle X^{-1}d, d \rangle^{1/2},
\end{aligned} \tag{2.9}$$

where we have defined $v = X^{1/2}(u-c)$ or $u = c + X^{-1/2}v$.

The *polar* of the set C is defined by

$$C^\circ := \{ d : s_C(d) \leq 1 \} = \{ x : \langle x, u \rangle \leq 1, \forall u \in C \}.$$

An easy calculation shows that

$$\left(\operatorname{conv}(\{u_i\}_1^k) \right)^\circ = \{ x : \langle x, u_i \rangle \leq 1, \ i = 1, \dots, k \}.$$

The following is a *key* result. Among other things, it implies that the optimality conditions (2.7) are powerful enough to prove the uniqueness of the minimum volume circumscribed ellipsoid as well as the uniqueness of the maximal volume inscribed ellipsoid treated in §3.

Lemma 2.6. *Let $\{u_i\}_1^k$ (k arbitrary) be a set of unit vectors in $\mathbb{R}^n$ satisfying the conditions $\sum_{i=1}^k \lambda_i u_i u_i^T = I_n$ and $\sum_{i=1}^k \lambda_i u_i = 0$. Define the polytope $P = \operatorname{conv}\left(\{u_i\}_1^k \right)$ and its polar $P^\circ = \{ x : \langle u_i, x \rangle \leq 1, i = 1, \dots, k \}$. The unit ball is both the unique minimum volume circumscribed ellipsoid of P and the unique maximum volume inscribed ellipsoid of P°.*

Proof. Let $E(X,c)$ be any ellipsoid covering the points $\{u_i\}_1^k$. We have $\langle X(u_i - c), u_i - c \rangle \leq 1$ and

$$\begin{aligned}
n = \sum_{i=1}^{k} \lambda_i &\geq \sum_{i=1}^{k} \lambda_i \langle X(u_i - c), u_i - c \rangle = \langle X, \sum_{i=1}^{k} \lambda_i (u_i - c)(u_i - c)^T \rangle \\
&= \langle X, \sum_{i=1}^{k} \lambda_i u_i u_i^T \rangle - \langle X, \sum_{i=1}^{k} c\lambda_i u_i^T \rangle - \langle X, \sum_{i=1}^{k} \lambda_i u_i c^T \rangle + \langle X, (\sum_{i=1}^{k} \lambda_i) cc^T \rangle \\
&= \langle X, I_n \rangle + n\langle X, cc^T \rangle = \operatorname{tr}(X) + n\langle Xc, c \rangle \\
&\geq n\left(\det(X)^{1/n} + \langle Xc, c \rangle \right).
\end{aligned}$$

7

Here the fourth equality follows from (2.7), and the last inequality follows from the fact that $\det(X)^{1/n} \leq \operatorname{tr}(X)/n$ which is precisely the arithmetic–geometric mean inequality applied to the eigenvalues of X. Thus

$$\det(X)^{1/n} + \langle Xc, c \rangle \leq 1,$$

and the equality $\det(X) = 1$ holds if and only if $c = 0$, $\langle X(u_i - c), u_i - c \rangle = 1$ for all $i = 1, \ldots, k$, and the arithmetic–geometric mean inequality holds as an equality. The last condition holds if and only if X is a positive multiple of the identity matrix (and then $\det(X) = 1$ implies $X = I_n$). Thus, the minimum volume ellipsoid covering the points $\{u_i\}_1^k$ must be the unit ball.

Next, let $E(X, c)$ be any ellipsoid inscribed in P°. It follows from (2.3) that $\operatorname{vol}(E(X, c)) = \det(X^{-1})\omega_n$. By virtue of (2.9), the inclusion $E(X, c) \subseteq P^\circ$ implies

$$s_{E(X,c)}(u_i) = \langle c, u_i \rangle + ||X^{-1/2}u_i|| \leq s_{P^\circ}(u_i) = \max_j \langle u_i, u_j \rangle \leq 1, \quad i = 1, \ldots, k.$$

The Cauchy–Schwarz inequality gives $\langle X^{-1/2}u_i, u_i \rangle \leq ||X^{-1/2}u_i|| \cdot ||u_i|| = ||X^{-1/2}u_i||$. Therefore,

$$n = \sum_{i=1}^k \lambda_i \geq \sum_{i=1}^k \lambda_i \left(\langle c, u_i \rangle + \langle X^{-1/2}u_i, u_i \rangle \right)$$

$$= \langle X^{-1/2}, \sum_{i=1}^k \lambda_i u_i u_i^T \rangle = \operatorname{tr}(X^{-1/2}) \geq n \det(X)^{-1/2n},$$

where the last inequality follows from the arithmetic–geometric mean inequality applied to the eigenvalues of $X^{-1/2}$. Thus $\det(X) \geq 1$, and the equality $\det(X) = 1$ holds if and only if (i) X is a positive multiple of the identity matrix (and then $\det(X) = 1$ implies $X = I_n$), and (ii) $1 = \langle c, u_i \rangle + \langle X^{-1/2}u_i, u_i \rangle = \langle c, u_i \rangle + 1$, that is, $\langle c, u_i \rangle = 0$ for all $i = 1, \ldots, k$. Then the equation $\sum_{i=1}^n \lambda_i u_i u_i^T = I_n$ implies that $||c||^2 = \sum_{i=1}^k \lambda_i \langle c, u_i \rangle^2 = 0$. Thus, condition (ii) holds if and only if $c = 0$. The lemma is proved. $\square$

Theorem 2.7. *Let K be a convex body in $\mathbb{R}^n$. The minimum volume circumscribed ellipsoid of K is unique. Moreover, the optimality conditions (2.6) are necessary and sufficient conditions for an ellipsoid $E(X, c)$ to be the minimum volume circumscribed ellipsoid of K.*

It is easy to see that K can be replaced by $S = \operatorname{ext}(K)$ in this theorem.

Proof. The neccessity of the conditions (2.6) is already proved in Theorem 2.2. We assume, without any loss of generality, that $E(I_n, 0) = B_n$ satisfies the optimality conditions (2.7) for some set of multipliers $\{\lambda_i\}$. Let $E = CE(K)$. We claim that $E = B_n$. This will immediately imply the remaining parts of the theorem. Since $E \subseteq P$, we have $\operatorname{vol}(B_n) \geq \operatorname{vol}(E) \geq \operatorname{vol}(B_n)$ where the first inequality follows because E has minimum volume among ellipsoids circumscribing K and the second inequality follows from Lemma 2.6. The same lemma proves the claim that $E = B_n$. $\square$

The *breadth* of a convex body is the smallest distance between its two parallel support planes, and *diameter* is the distance between its two farthest points. The following result of John [18] shows that a convex body can be "rounded" by an affine transformation. Its proof also gives valuable information about the locations of the contact points. Its easy proof is due to Juhnke [21].

Corollary 2.8. *Let K be a convex body whose optimal covering ellipsoid is the unit ball, and let $\{u_i\}_1^k$ be its contact points. Then*

$$\max_{x\in K}\langle d, x\rangle \max_{x\in K}\langle -d, x\rangle \ \geq\ \max_i\langle d, u_i\rangle \max_i\langle -d, u_i\rangle \ \geq\ \frac{1}{n}, \quad \forall d, \ ||d|| = 1,$$

$$\max_{x\in K}\langle d, x\rangle + \max_{x\in K}\langle -d, x\rangle \ \geq\ \max_i\langle d, u_i\rangle + \max_i\langle -d, u_i\rangle \ \geq\ \frac{2}{\sqrt{n}}, \quad \forall d, \ ||d|| = 1.$$

Consequently, any convex body can be transformed by an affine map into a body, for which the ratio of breadth to diameter is at least $1/\sqrt{n}$.

Proof. Define $P = \mathrm{conv}(\{u_i\}_1^k)$. Since $s_P(d) = \max_k\langle d, u_k\rangle \geq \langle d, u_i\rangle$ for any i, we have

$$0 \leq \sum_{i=1}^{k}\lambda_i(s_P(d) - \langle d, u_i\rangle)(s_P(-d) + \langle d, u_i\rangle)$$

$$= (\sum_{i=1}^{k}\lambda_i)s_P(d)s_P(-d) - \sum_i\lambda_i\langle d, u_i\rangle^2 = ns_P(d)s_P(-d) - 1$$

where the equalities follow from (2.7). This proves the first line of relations in the corollary. The second line of relations follow from the first one using the inequality $(s_P(d)+s_P(-d))^2 \geq 4s_P(d)s_P(-d)$. Observe that the quantity $s_K(d)+s_K(-d) = \max_{x\in K}\langle d, x\rangle+\max_{x\in K}\langle -d, x\rangle$ is the distance between the two parallel support planes of K in direction d. This proves the corollary for K whose optimal ellipsoid $\mathrm{CE}(K) = B_n$.

The rest of the corollary follows since an arbitrary convex body can be transformed by an affine transformation into another convex body whose optimal covering ellipsoid is the unit ball. $\square$

We now give a proof of Fritz John's celebrated result mentioned at the beginning of this section. Our proof is simpler, and uses ideas from Ball [1] and Juhnke [19].

Theorem 2.9. *Let K be a convex body in $\mathbb{R}^n$ and $E(X, c) = \mathrm{CE}(K)$ be its optimal circumscribing ellipsoid. The ellipsoid with the same center c but shrunk by a factor n is contained in K. If K is symmetric ($K = -K$), then the ellipsoid with the same center c but shrunk only by a factor $\sqrt{n}$ is contained in K.*

Proof. Without loss of generality, we assume that $\mathrm{CE}(K) = E(I_n, 0) = B_n$. The theorem states that $n^{-1}B_n \subseteq K$. Let

$$P = \mathrm{conv}\left(\{u_i\}_1^k\right)$$

be the convex hull of the contact points. We claim the stronger statement that $n^{-1}B_n \subseteq P$. Since $P \subseteq K$, we will then have $n^{-1}B_n \subseteq K$. By duality, the claim is equivalent to showing that the polar sets satisfy $P^* \subseteq (n^{-1}B_n)^* = nB_n$. Let $x \in P^*$. Since $-||x|| = -||x||\cdot||u_i|| \leq \langle x, u_i\rangle \leq 1$, we have

$$0 \leq \sum_{i=1}^{k}\lambda_i(1 - \langle x, u_i\rangle)(||x|| + \langle x, u_i\rangle)$$

$$= (\sum_{i=1}^{k}\lambda_i)||x|| - \sum_{i=1}^{k}\lambda_i(\langle x, u_i\rangle)^2 = n||x|| - ||x||^2,$$

9

where the second equality follows from $\sum_i \lambda_i = n$ and (2.7). This implies $||x|| \le n$, and proves that $P^* \subseteq nB_n$.

If K is symmetric, we define $Q = \mathrm{conv}\left(\{\pm u_i\}_1^k\right) \subseteq K$ and claim that $n^{-1/2}B_n \subseteq Q$, or equivalently, that $Q^* \subseteq (n^{-1/2}B_n)^* = \sqrt{n}B_n$. It is easily shown that $Q^* = \{x : |\langle x, u_i \rangle| \le 1, i = 1, \ldots, k\}$. Let $x \in Q^*$. Since $-1 \le \langle x, u_i \rangle \le 1$, we have

$$0 \le \sum_{i=1}^{k} \lambda_i (1 - \langle x, u_i \rangle)(1 + \langle x, u_i \rangle) = n - ||x||^2.$$

This gives $||x|| \le \sqrt{n}$ and proves the claim. $\qquad\qquad\square$

Remark 2.10. The minimum volume covering ellipsoid problem may be set as a semi–infinite program in a different way, by replacing the set inclusion $K \subseteq E(X, c)$ by the equivalent functional constraints $s_{E(X,c)}(d) \ge s_K(d)$, that is by the constraints

$$\langle c, d \rangle + \langle Yd, d \rangle^{1/2} \ge s_K(d), \quad \forall d, \, ||d|| = 1,$$

where we restrict d to the unit sphere since support functions are homogeneous (of degree 1), in order to get a compact indexing set. The resulting semi–infinite program is solved in the same way as (2.4), and mirrors the solution to the maximum volume inscribed ellipsoid problem given in §3.

3 The maximum volume inscribed ellipsoid problem

Recall that the *inscribed ellipsoid problem* is the problem of finding a maximum volume ellipsoid inscribed in a convex body K in $\mathbb{R}^n$. It will be seen that this ellipsoid is unique as well, and we denote it by $\mathrm{IE}(K)$. As we mentioned in the Introduction, this ellipsoid is often referred to as the John ellipsoid or Löwner–John ellipsoid.

In this section, we again use semi–infinite programming to treat this problem. The inscribed ellipsoid has properties similar to those of the circumscribed ellipsoid. For example, the ellipsoid with the same center but blown up n times contains K, and in the case K is symmetric ($K = -K$), the ellipsoid needs to be blown up by a smaller factor $\sqrt{n}$. The ellipsoid $\mathrm{IE}(K)$ is very useful in the geometric theory of Banach spaces. It is also useful in some convex programming algorithms, such as the *inscribed ellipsoid method* of Tarasov, Erlikh, and Khachiyan [35].

As a first step, using (2.3), we can formulate the inscribed ellipsoid problem as a semi–infinite program

$$\min\{\det X : E(X, c) \subseteq K\}.$$

However, this is hard to work with, due to the inconvenient form of the constraints, $E(X, c) \subseteq K$. We replace this inclusion by the functional constraints

$$s_{E(X,c)}(d) \le s_K(d), \quad \forall d \in B_n,$$

where we again restrict d to the unit sphere since support functions are homogeneous (of degree 1), in order to get a compact indexing set.

Defining $Y = X^{-1}$, we can therefore rewrite our semi–infinite program in the form

$$\min \quad -\log \det Y$$

$$\text{s.t.} \quad \langle c, d \rangle + \langle Yd, d \rangle^{1/2} \le s_K(d), \quad \forall d : \|d\| = 1, \tag{3.1}$$

in which the decision variables are $(Y, c) \in S^n \times \mathbb{R}^n$ and we have infinitely many constraints indexed by the unit vector $\|d\| = 1$.

Since s_K is a convex function on $\mathbb{R}^n$, it is continuous. Therefore, there exists a positive constant $M > 0$ such that if (Y, c) is a feasible decision variable, then $|\langle c, d \rangle| \le M$, and $\langle Yd, d \rangle \le M$ for all $\|d\| = 1$. This proves that the set of feasible (Y, c) for problem (3.1) is compact, and implies that there exists a maximum volume ellipsoid inscribed in K.

We derive the optimality conditions for the maximum volume inscribed ellipsoid.

Theorem 3.1. *Let K be a convex body in $\mathbb{R}^n$. There exists an ellipsoid of maximum volume inscribed in K. If $E(X, c)$ is such an ellipsoid, then there exists a multiplier vector $\lambda = (\lambda_1, \ldots, \lambda_k) > 0$, $0 \le k \le n(n+3)/2$, and contact points $\{u_i\}_1^k$ such that*

$$X^{-1} = \sum_{i=1}^{k} \lambda_i (u_i - c)(u_i - c)^T,$$

$$0 = \sum_{i=1}^{k} \lambda_i (u_i - c), \tag{3.2}$$

$$u_i \in \partial K \cap \partial E(X, c), \quad i = 1, \ldots, k,$$

$$E(X, c) \subseteq K.$$

Proof. The existence of a maximum volume ellipsoid inscribed in K is already proved above. Let $E(X, c)$ be such an ellipsoid. Define $Y = X^{-1}$. Since constraints in (3.1) are indexed by $\|d\| = 1$, Theorem 2.1 applies: there exists a non–zero multiplier vector $(\delta_0, \delta_1, \ldots, \delta_k) \ge 0$, where $k \le n(n+3)/2$, $\delta_i > 0$ for $i > 0$, and directions $\{d_i\}_1^k$, $\|d_i\| = 1$, satisfying the conditions

$$\langle c, d_i \rangle + \langle Yd_i, d_i \rangle^{1/2} = s_K(d_i),$$

such that the Lagrangian function

$$L(Y, c, \delta) := -\delta_0 \log \det Y + 2 \sum_{i=1}^{k} \delta_i \left[\langle c, d_i \rangle + \langle Yd_i, d_i \rangle^{1/2} - s_K(d_i) \right]$$

satisfies the optimality conditions

$$0 = \nabla_c L(Y, c, \delta) = \sum_{i=1}^{k} \delta_i d_i,$$

$$0 = \nabla_Y L(Y, c, \delta) = -\delta_0 Y^{-1} + \sum_{i=1}^{k} \frac{\delta_i}{\langle Yd_i, d_i \rangle^{1/2}} d_i d_i^T.$$

Recalling that $\|d_i\| = 1$ and taking the trace of the right hand side of the last equation above gives $\delta_0 \operatorname{tr}(Y^{-1}) = \sum_{i=1}^{k} \delta_i \langle Yd_i, d_i \rangle^{-1/2}$. If $\delta_0 = 0$, then all $\delta_i = 0$, which contradicts $\delta \ne 0$. Therefore, $\delta_0 \ne 0$, and we let $\delta_0 = 1$. Define

$$u_i := c + \langle Yd_i, d_i \rangle^{-1/2} Yd_i, \quad \lambda_i := \langle Yd_i, d_i \rangle^{1/2} \delta_i, \quad i = 1, \ldots, k.$$

We have $\langle d_i, u_i \rangle = \langle c, d_i \rangle + \langle Yd_i, d_i \rangle^{1/2}$, so that

11

$$s_{E(X,c)}(d_i) = s_K(d_i) = \langle d_i, u_i \rangle,$$

which means that $u_i \in \partial K \cap \partial E(X,c)$, that is u_i is a contact point. Rewriting the above optimality conditions in terms of $\{u_i\}$ and $\{\lambda_i\}$ and simplifying, we arrive at the conditions (3.2). $\qquad\square$

As in the circumscribed ellipsoid case, we have

Corollary 3.2. *Let K be a convex body in $\mathbb{R}^n$. The contact points of $\mathrm{IE}(K)$ are not contained in any closed halfspace whose bounding hyperplane passes through the center of $\mathrm{IE}(K)$.*

We can simplify these conditions by assuming that the optimal ellipsoid is the unit ball $E(I_n, 0)$. Then the Fritz John conditions become

$$I_n = \sum_{i=1}^k \lambda_i u_i u_i^T, \quad 0 = \sum_{i=1}^k \lambda_i u_i, \tag{3.3}$$

$$u_i \in \partial K \cap \partial B_n, \quad i = 1, \dots, k, \quad B_n \subseteq K.$$

We note that the optimality conditions (3.3) are exactly the *same* as the corresponding optimality conditions (2.7) in the circumscribed ellipsoid case, except for the feasibility constraints $B_n \subseteq K$.

Theorem 3.3. *Let K be a convex body in $\mathbb{R}^n$. The maximal volume ellipsoid inscribed in K is unique. Furthermore, the optimality conditions (3.2) are necessary and sufficient for an ellipsoid $E(X,c)$ to be the maximal volume inscribed ellipsoid of K.*

The proof uses Lemma 2.6. It is omitted since it is very similar to the proof of Theorem 2.7.

We end this section by proving an analogue of Fritz John's containment results concerning $\mathrm{CE}(K)$.

Theorem 3.4. *Let K be a convex body in $\mathbb{R}^n$ and let $E(X,c) = \mathrm{IE}(K)$ be its optimal inscribed ellipsoid. The ellipsoid with the same center c but enlarged by a factor n contains K. If K is symmetric, then the ellipsoid with the same center c but enlarged by a factor $\sqrt{n}$ contains K.*

Proof. The proof here is similar to the proof of Theorem 2.9. Without loss of generality, we assume that $\mathrm{IE}(K) = E(I, 0) = B_n$. The first part of the theorem follows if we can prove the claim that

$$K \subseteq P^* \subseteq n B_n.$$

Since $1 = s_K(u_i) = \max_{x \in K} \langle u_i, x \rangle$, the first inclusion holds true. If $x \in P^*$, then $-||x|| = -||x|| \cdot ||u_i|| \leq \langle x, u_i \rangle \leq 1$, and

$$0 \leq \sum_{i=1}^k \lambda_i (1 - \langle x, u_i \rangle)(||x|| + \langle x, u_i \rangle) = n||x|| - ||x||^2,$$

where the equality follows from $\sum_i \lambda_i = n$ and (3.3). This implies $||x|| \leq n$, and proves the second inclusion in the claim.

If K is symmetric, we define $Q = \mathrm{conv}\left(\{\pm u_i\}_1^k\right) \subseteq K$ and claim that $K \subseteq Q^* \subseteq \sqrt{n} B_n$. Since $1 = s_K(\pm u_i)$, we have $|\langle u_i, x \rangle| \leq 1$, and the first inclusion follows. To prove the second inclusion in the claim, let $x \in Q^*$. We have $|\langle x, u_i \rangle| \leq 1$, and the rest of the proof follows as in the proof of Theorem 2.9. $\qquad\square$

4 Automorphism group of convex bodies

Let K be a convex body in $\mathbb{R}^n$. The uniqueness of the two extremal ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$ have important consequences regarding the invariance properties of the two ellipsoids. We will see in this section that the symmetry properties of the convex body K is inherited by the two ellipsoids. If K is symmetric enough, then it becomes possible to give explicit formulae for the ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$. We will demonstrate this for some special convex bodies in the remaining Sections of this paper.

Apart from its intrinsic importance, the invariance properties of the ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$ have important applications to Lie groups [11], [28], to differential geometry [24], and to the computation of the extremal ellipsoids for some special polytopes and convex bodies [6], [4], among others.

We start with a

Definition 4.1. The (affine) automorphism group $\mathrm{Aut}(K)$ of a convex body K in $\mathbb{R}^n$ is the set of affine transformation $T(x) = a + Ax$ leaving K invariant, that is,

$$\mathrm{Aut}(K) = \{T(x) = a + Ax : T(K) = K\}.$$

Note that since $0 < \mathrm{vol}(K) = \mathrm{vol}(T(K)) = |\det A|\,\mathrm{vol}(K)$, we have $|\det A| = 1$.

It is shown in [14] that the automorphism group of most convex bodies consists of the identity transformation alone. This is to be expected, since the symmetry properties of a given convex body can easily be destroyed by slightly perturbing the body. Nevertheless, the study of the symmetry properties of convex bodies is important for many reasons.

The ellipsoids are the most symmetric convex bodies. Therefore, we first investigate their automorphism groups and then relate them to the automorphism groups of arbitrary convex bodies.

Definition 4.2. Let A be an invertible matrix in $\mathbb{SR}^{n \times n}$. Equip $\mathbb{R}^n$ with the quadratic form $\langle Au, v \rangle$ which we write as an inner product

$$\langle u, v \rangle_A := \langle Au, v \rangle.$$

We denote by $O(\mathbb{R}^n, A)$ the set of linear maps *orthogonal* under this inner product,

$$
\begin{aligned}
O(\mathbb{R}^n, A) &:= \{g \in \mathbb{R}^{n \times n} : g^*g = gg^* = I\} \\
&= \{g \in \mathbb{R}^{n \times n} : \langle gu, gv \rangle_A = \langle u, v \rangle_A, \ \forall\, u, v \in \mathbb{R}^n\} \\
&= \{g \in \mathbb{R}^{n \times n} : g^T A g = A\},
\end{aligned}
$$

where g^* is the conjugate matrix of g with respect to the inner product $\langle \cdot, \cdot \rangle_A$, that is $\langle gu, v \rangle_A = \langle u, g^*v \rangle_A$ for all u, v in $\mathbb{R}^n$. The second equality above follows as $\langle gu, gv \rangle_A = \langle u, g^*gv \rangle_A$ and this equals $\langle u, v \rangle_A$ if and only if $g^*g = I$. The third equality follows since $\langle g^T A g u, v \rangle = \langle Agu, gv \rangle = \langle gu, gv \rangle_A = \langle u, v \rangle_A = \langle Au, v \rangle$. If A is positive definite, then $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_X)$ is a Euclidean space. In particular, $O_n := O(\mathbb{R}^n, I)$ is the set of orthogonal matrices in the usual inner product on $\mathbb{R}^n$.

Lemma 4.3. *If X is a symmetric, positive definite $n \times n$ matrix, then*

$$\mathrm{Aut}(E(X, 0)) = O(\mathbb{R}^n, X).$$

In particular, $\mathrm{Aut}(B_n)$ *is the set of* $n \times n$ *orthogonal matrices. We also have*

$$\mathrm{Aut}(E(X,c)) = T_c O(\mathbb{R}^n, X) T_{-c},$$

where T_c *is the translation map* $T_c x = c + x$. *Moreover,* $\mathrm{Aut}(E(X,c))$ *fixes* c, *the center of* $E(X,c)$, *that is,* $\theta(c) = c$ *for every* θ *in* $\mathrm{Aut}(E(X,c))$.

Proof. We first determine $\mathrm{Aut}(B_n)$. Let T be in $\mathrm{Aut}(B_n)$, where $T(x) = a + Ax$. Since T maps the boundary of B_n onto itself, we have $q(x) := \|a + Ax\|^2 = \langle A^T A x, x \rangle + 2\langle A^T a, x \rangle + \|a\|^2 = 1$ for all $\|x\| = 1$. Then $q(x) - q(-x) = 4\langle A^T a, x \rangle = 0$ for all $\|x\| = 1$, which implies that $A^T a = 0$ and since A is invertible, $a = 0$. Consequently, $q(x) = \langle A^T A x, x \rangle = 1$ for all $\|x\| = 1$, which gives $A^T A = I_n$, that is, A is an orthogonal matrix.

Next, we determine $\mathrm{Aut}(E(X,0))$. We have the commutative diagram

$$
\begin{array}{ccccc}
B_n & \xrightarrow{X^{-1/2}} & E(X,0) & \xrightarrow{T_c} & E(X,c) \\
\;\downarrow{\scriptstyle U} & & \;\downarrow{\scriptstyle \theta_0} & & \;\downarrow{\scriptstyle \theta} \\
B_n & \xrightarrow{X^{-1/2}} & E(X,0) & \xrightarrow{T_c} & E(X,c)
\end{array}
$$

where $\theta \in \mathrm{Aut}(E(X,c))$, $U \in \mathrm{Aut}(B_n) = O_n$, and $\theta_0 \in \mathrm{Aut}(E(X,0))$. From the diagram, we have $I = U^T U = (X^{1/2}\theta_0 X^{-1/2})^T (X^{1/2}\theta_0 X^{-1/2}) = X^{-1/2}\theta_0^T X \theta_0 X^{-1/2}$. This gives $\theta_0^T X \theta_0 = X$ and proves that $\mathrm{Aut}(E(X,0)) = O(\mathbb{R}^n, X)$.

Since $T_c^{-1} = T_{-c}$, we have

$$\mathrm{Aut}(E(X,c)) = T_c \circ \mathrm{Aut}(E(X,0)) \circ T_{-c}.$$

Every θ in $\mathrm{Aut}(E(X,c))$ has the form $\theta(u) = (T_c \circ \theta_0 \circ T_{-c})(u) = T_c(\theta_0(-c + u)) = (c - \theta_0 c) + \theta_0 u$ for some θ_0 in $\mathrm{Aut}(E_0)$. This gives $\theta(c) = c$, meaning that $\mathrm{Aut}(E(X,c))$ fixes the center of E. $\square$

Definition 4.4. Let K be a convex body in $\mathbb{R}^n$. An ellipsoid $E = E(X,c)$ is an *invariant ellipsoid* of K if $\mathrm{Aut}(K) \subseteq \mathrm{Aut}(E)$, that is, if every automorphism of K is an automorphism of E.

It immediately follows from Lemma 4.3 that $\mathrm{Aut}(K)$ fixes the center of any invariant ellipsoid E of K.

The following Theorem in Danzer et al. [11] is a central result regarding the symmetry properties of the extremal ellipsoids.

Theorem 4.5. *Let K be a convex body in $\mathbb{R}^n$. The extremal ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$ are invariant ellipsoids of K. Thus, $\mathrm{Aut}(K) \subseteq \mathrm{Aut}(\mathrm{CE}(K))$, $\mathrm{Aut}(K) \subseteq \mathrm{Aut}(\mathrm{IE}(K))$, and $\mathrm{Aut}(K)$ fixes the centers of $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$.*

Proof. Since the arguments are similar, we only prove the statements about $\mathrm{CE}(K)$. Let $g \in \mathrm{Aut}(K)$. Since $K \subseteq \mathrm{CE}(K)$ and $K = gK \subseteq g(\mathrm{CE}(K))$, the ellipsoids $\mathrm{CE}(K)$ and $g(\mathrm{CE}(K))$ both cover K, and since $\mathrm{vol}(\mathrm{CE}(K)) = \mathrm{vol}(g(\mathrm{CE}(K)))$, they are both minimum volume circumscribed ellipsoids of K. It follows from Theorem 2.7 that $g(\mathrm{CE}(K)) = \mathrm{CE}(K)$. $\square$

Corollary 4.6. *The automorphism group $\mathrm{Aut}(K)$ of a convex body K in $\mathbb{R}^n$ is a compact Lie group.*

14

Proof. We have $Aut(K) \subseteq \mathrm{Aut(CE}(K))$ by Theorem 4.5, and Lemma 4.3 implies that $\mathrm{Aut(CE}(K))$ is compact. Clearly, $\mathrm{Aut}(K)$ is a closed subset of the general affine linear group in $\mathbb{R}^n$. It follows that $\mathrm{Aut}(K)$ is a compact group. A classical theorem of von Neumann implies that it is a Lie group. $\qquad\square$

Remark 4.7. The existence of invariant (fixed) points and (symmetric positive definite) matrices for $\mathrm{Aut}(K)$ have been demonstrated above using the invariance properties of the either one of the extremal ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$. The same goal can be achieved in at least two other ways, using the invariance properties of either the *center of gravity* of convex sets, or of the Haar probability measure μ_G on $G = \mathrm{Aut}(K)$. In a certain sense, all three procedures are similar in that they all employ *averaging*, but in different ways.

By definition, the center of gravity of a convex body K is the point

$$\mathrm{cg}(K) := \frac{\int_K x\,dx}{\int_K dx} = \frac{\left(\int_K x_1\,dx, \ldots, \int_K x_n\,dx\right)}{\mathrm{vol}(K)}.$$

One may think of $\mathrm{cg}(K)$ as the limit of the points $p := \sum_{i=1}^{k}(\mathrm{vol}(K_i)/\mathrm{vol}(K))x_i$ where $\{K_i\}_1^k$ is a partition of K into subregions and x_i in K_i. Since $p \in K$, we see that $\mathrm{cg}(K)$ lies in K. If $T \in \mathrm{Aut}(K)$, then $Tp = \sum_{i=1}^{k}(\mathrm{vol}(K_i)/\mathrm{vol}(K))Tx_i = \sum_{i=1}^{k}(\mathrm{vol}(TK_i)/\mathrm{vol}(K))Tx_i$. We see that both $\{p\}$ and $\{Tp\}$ converge to $\mathrm{cg}(K)$, proving that $\mathrm{Aut}(K)$ fixes the center of gravity of K.

Moreover, one can prove the existence of an invariant ellipsoid for K, see [28], pp. 130–135: it is easy to verify that the map $\pi : \mathrm{Aut}(K) \to GL(\mathbb{SR}^{n \times n})$ given by the formula

$$\pi(g)(S) = (g \otimes g)(S) := gSg^T$$

is a *representation* of the group $\mathrm{Aut}(K)$ on the matrix space $\mathbb{SR}^{n \times n}$, that is, $\pi(gh) = \pi(g)\pi(h)$:

$$\pi(gh)(S) = ghSh^Tg^T = g(\pi(h)(S)g^T = \pi(g)\pi(h)(S).$$

Consider the group $\mathcal{G} := \pi(\mathrm{Aut}(K))$ and the orbit

$$C_S = \mathcal{G}(S) = \{gSg^T : g \in \mathrm{Aut}(K)\}$$

where $S \in \mathbb{SR}^{n \times n}$ is an arbitrary positive definite matrix. Define the convex body $\mathcal{K} = \mathrm{conv}(C_S) \subset \mathbb{SR}^{n \times n}$. Clearly, the orbit C_S is invariant under the group $\mathcal{G}$, and thus so is the convex body $\mathcal{K}$. It follows from the above argument that the center of gravity $Y = \mathrm{cg}(\mathcal{K})$ is fixed by $\mathcal{G}$, that is, $gYg^T = Y$ for all g in $\mathrm{Aut}(K)$. This gives $(g^{-1})^T Y^{-1}g^{-1} = S^{-1}$, or equivalently, $g^T S^{-1}g = S^{-1}$ for all g in $\mathrm{Aut}(K)$. This means that $X = Y^{-1}$ is invariant under $\mathrm{Aut}(K)$. Then any ellipsoid $E(X, c)$, where $c \in K$ is any fixed point of $\mathrm{Aut}(K)$, say the center of gravity of K, is an invariant ellipsoid of K.

Let $\mu = \mu_G$ be the unique Haar probability measure on $G = \mathrm{Aut}(K)$. If $f : G \to \mathbb{R}$ is a continuous function and $h, k \in G$ are arbitrary, we have

$$\int_G f(hg)\,d\mu(g) = \int_G f(g)\,d\mu(g) = \int_G f(gk)\,d\mu(g) = \int_G f(g^{-1})\,d\mu(g),$$

where the first and second equalities express left and right invariance properties of the Haar integral, respectively. If x is any interior point of K, then the point $c := \int_G gx\,d\mu(g)$ lies in *interior*(K) and is fixed by $\mathrm{Aut}(K)$, for if $h \in \mathrm{Aut}(K)$, then $hc = \int_G hgcd\,\mu(g) = \int_G gcd\,\mu(g) = c$, where the second equality follows from the left invariance of μ. Finally,

the standard proof of the existence of a positive definite invariant matrix X found in most textbooks proceeds as follows: start with any inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$, and define the inner product

$$[[u, v]] := \int_G \langle gu, gv \rangle \, d\mu(g).$$

This is an invariant inner product on G, because

$$[[hu, hv]] = \int_G \langle hgu, hgv \rangle \, d\mu(g) = \int_G \langle gu, gv \rangle \, d\mu(g) = [[u, v]],$$

where the second equality follows again from the left invariance of μ. We have $[[u, u]]$ non–negative and equal to zero if and only if $u = 0$, because the same thing is true for $\langle u, u \rangle$.

We have seen that convex bodies give rise to compact affine groups through their automorphism groups. Conversely, it is a well known fact in Lie group theory that a compact Lie group can be imbedded as a closed subgroup of the linear group $GL(V)$ for some finite dimensional vector space V. This can be found in most books on Lie groups, see for example [8], [28], [40]. If G is a compact, affine Lie group in $\mathbb{R}^n$, then it is isomorphic to the compact linear group

$$\left\{ \bar{T} = \begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix} : T \in G, \ T(x) = a + Ax \right\}$$

in $\mathbb{R}^{n+1}$. In fact, if $T(x) = a + Ax$, then

$$\bar{T} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} a + Ax \\ 1 \end{pmatrix}.$$

Thus, the affine transformations of $\mathbb{R}^n$ are in one–to–one correspondence with the linear transformations of $\mathbb{R}^{n+1}$ that keep the hyperplane $\{(x, 1) : x \in \mathbb{R}^n\}$ invariant. If G is a compact linear group in $\mathbb{R}^n$, then one can consider the convex set $K = \text{conv}(Gx)$ where x is an arbitrary point in $\mathbb{R}^n$. Clearly, K is invariant under G, that is, $G \subseteq \text{Aut}(K)$. In this way, one can obtain the following purely group theoretical result employing one of the methods in Remark 4.7, see [11], [28], [40]:

Theorem 4.8. *Let G be a compact group of linear transformations on $\mathbb{R}^n$. The group G fixes a point in $\mathbb{R}^n$, that is, there exists c in $\mathbb{R}^n$ such that $gc = c$ for all g in G. Moreover, there exists a positive definite matrix X invariant under G, that is, $g^T X g = X$ for all g in G.*

We remark that the Haar measure approach in Remark 4.7 already proves this result directly, without considering orbits and convex sets. One can also extend the theorem to affine compact groups, for example by considering the isomorphic linear compact group in $\mathbb{R}^{n+1}$. However, the method of employing convex bodies does have its merits. For example, a simple proof of the *algebraicity* of compact linear groups can be found in [28], pp. 130–135 and [40], Chapter 15, using this approach.

The existence of the invariant quadratic form for compact Lie groups is a very important result, since it implies that its linear representations are *completely reducible*, that is, its representations can be written as direct products of *irreducible representations*. For the group $\text{Aut}(K)$, this simply means that the matrices appearing in the linear parts of the affine transformations in $\text{Aut}(K)$ must all have the same block diagonal structure.

We also mention that such concepts as the fixed points of $\mathrm{Aut}(K)$ as well as its invariant quadratic forms belong to the *invariant theory*. For example, if G is a linear group in $GL(\mathbb{R}^n)$, the *invariant polynomials* of G is the set of polynomials

$$\mathbb{R}[x_1,\ldots,x_n]^G := \{p \in \mathbb{R}[x_1,\ldots,x_n] : p(gx) = p(x),\ \forall\, g \in G,\ \forall\, x \in \mathbb{R}^n\}.$$

The determination of the invariant polynomials for specific groups is one of the major goals of invariant theory whose origins go back to the works of Cayley, Sylvester, Gordan, Hilbert, etc. in the 19th century. A major result going back to Hilbert in 1890s implies that $\mathbb{R}[x_1,\ldots,x_n]^{\mathrm{Aut}(K)}$ is *finitely generated*. This means that there exists finitely many invariant polynomials S (which can be assumed homogeneous) such that every invariant polynomial in $\mathbb{R}[x_1,\ldots,x_n]^{\mathrm{Aut}(K)}$ can be written as a polynomial of elements of S. See [40], pp. 280–281 for a fairly simple, direct proof. A direct significance of this result for the ellipsoid $\mathrm{CE}(K) = E(X,c)$, say, is that both c and X are invariant. Consequently, if the number of generators is small, then the complexity of finding the ellipsoid $\mathrm{CE}(K)$ simplifies. This will be demonstrated in the remaining Sections of our paper.

The automorphism group proves to be useful in other ways as well. For example, the following result of Davies [12] is quite interesting. Davies calls a convex body K in $\mathbb{R}^n$ *symmetric* if $\mathrm{Aut}(K)$ acts *transitively* on the extreme points of K, that is, given any two points x, y in $\mathrm{ext}(K)$, there exists a transformation g in $\mathrm{Aut}(K)$ such that $gx = y$. He then proves the following result

Lemma 4.9. *If $K = \mathrm{conv}(Gx)$ is a symmetric convex body in $\mathbb{R}^n$ in the sense Davies, then there exists a* unique *fixed point of* $\mathrm{Aut}(K)$ *in K.*

Proof. Let μ be the Haar probability measure μ on the compact Lie group $G := \mathrm{Aut}(K)$. Let x be an arbitrary point in $\mathrm{ext}(K)$. Since K is symmetric, $\mathrm{ext}(K) = Gx$. Define the point

$$c := \int_G gx\, d\mu(g).$$

The point c is invariant under the action of G, since if $h \in G$, we have $hc = \int_G hgx\, d\mu(g) = \int_G gx\, d\mu(g) = c$, where the second equality is a consequence of the invariance of μ. Now, if a is any invariant point in K, then $ga = a$ for all g in G, and we have $a = \int_G a\, d\mu(g) = \int_G ga\, d\mu(g)$. By Minkowski theorem, we have $a \in \mathrm{conv}(ext(K)) = \mathrm{conv}(Gx)$, that is, $a = \sum_{i=1}^k \lambda_i g_i x$ for some $\{\lambda_i, g_i\}_1^k$ where g_i in G and $\lambda_i \geq 0$, $\sum_{i=1}^k \lambda_i = 1$. Therefore,

$$a = \int_G ga\, d\mu(g) = \sum_{i=1}^k \lambda_i \int_G gg_i x\, d\mu(g) = \left(\sum_{i=1}^k \lambda_i\right) \int_G gx\, d\mu(g) = c,$$

where third equality follows again from the invariance of the measure μ. $\qquad\square$

The lemma implies in particular that the center of gravity of a symmetric body K is the only invariant point of K under the action of $\mathrm{Aut}(K)$. Thus, the centers of the ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$ must coincide and be equal to the center of gravity of K.

It is also possible to determine a formula for the matrix X in the circumscribing ellipsoid $\mathrm{CE}(K) = E(X,c)$.

Lemma 4.10. *Let $K \subset \mathbb{R}^n$ be a convex body, symmetric in the sense of Davies. If X is an invariant matrix of K such that $X^{-1} \in \mathrm{conv}\left(\{(y-c)(y-c)^T : y \in \mathrm{ext}(K)\}\right)$ where c is the invariant point of K, then*

$$X^{-1} = \int_G g^{-1}(x - c)(x - c)^T (g^T)^{-1} \, d\mu(g),$$

where x is an arbitrary point in $\mathrm{ext}(K)$ and μ is the Haar probability measure on K.

Proof. Pick an arbitrary point x in $\mathrm{ext}(K)$ such that $\mathrm{ext}(K) = Gx$. We may assume without any loss of generality that $\mathrm{cg}(K) = 0$. Since X is an invariant matrix, we have $g^T X g = X$ or $X^{-1} = g^{-1} X^{-1} g^{-T}$ where we defined $g^{-T} := (g^{-1})^T = (g^T)^{-1}$. Let $X^{-1} = \sum_{i=1}^k \lambda_i y_i y_i^T = \sum_{i=1}^k \lambda_i h_i x x^T h_i^T$ where $h_i \in G$. We have

$$X^{-1} = \int_G g^{-1} X^{-1} g^{-T} \, d\mu(g) = \sum_{i=1}^k \lambda_i \int_G (g^{-1} h_i) x x^T (g^{-1} h_i)^T \, d\mu(g)$$

$$= \sum_{i=1}^k \lambda_i \int_G g^{-1} x x^T g^{-T} \, d\mu(g) = \int_G g^{-1} x x^T g^{-T} \, d\mu(g),$$

where the last equality follows from the right invariance of the Haar measure. $\square$

Corollary 4.11. *Let K be a convex body in $\mathbb{R}^n$ symmetric in the sense of Davies. The extremal covering ellipsoid of $\mathrm{CE}(K) = E(X, c)$ has center $c = \mathrm{cg}(K)$ and*

$$X^{-1} = n \int_G g^{-1}(x - c)(x - c)^T g^{-T} \, d\mu(g),$$

where $x \in \mathrm{ext}(K)$ is an arbitrary point and μ is the Haar probability measure on K.

This is an immediate consequence of the above lemma, Theorem 2.2, and (2.8).

Many interesting questions and research directions remain regarding the automorphism groups of convex bodies. It is not practical to investigate these in this paper; doing so would increase the size of the paper beyond reasonable bounds and also change its character. We plan to pursue these issues in future papers.

5   Invariance properties of a slab

In this paper, one of the problems we are interested in is the determination of the extremal ellipsoids of the convex body K which is the part of a given ellipsoid $E(X_0, c_0)$ between two parallel hyperplanes,

$$K = \{x : \langle X_0(x - c_0), \, x - c_0 \rangle \leq 1, \, a \leq \langle p, \, x - c_0 \rangle \leq b\},$$

where p is a non–zero vector in $\mathbb{R}^n$, and where a and b are such that K is nonempty. Recall that we call such a convex body K a slab.

In this section, we determine $\mathrm{Aut}(K)$ and the form of the ellipsoids $\mathrm{CE}(K)$ and $\mathrm{IE}(K)$.

If we substitute $u = X_0^{1/2}(x - c_0)$, that is, $x = c_0 + X_0^{-1/2} u$, the quadratic inequality $\langle X_0(x - c_0), x - c_0 \rangle \leq 1$ becomes $\|u\| \leq 1$ and making the further substitution $q := X_0^{-1/2} p$, the linear form $\langle p, x - c_0 \rangle$ becomes

$$\langle p, x - c_0 \rangle = \langle X_0^{-1/2} p, X_0^{1/2}(x - c_0) \rangle = \langle q, u \rangle.$$

Defining $\bar{p} = q/||q||$, $\alpha = a/||q||$ and $\beta = b/||q||$, the linear inequalities $a \leq \langle p, x - c_0 \rangle \leq b$ reduce to $\alpha \leq \langle \bar{p}, u \rangle \leq \beta$. Altogether, these substitutions give

$$K = \{c_0 + X_0^{-1/2} u : ||u|| \leq 1, \ \alpha \leq \langle \bar{p}, u \rangle \leq \beta\}.$$

Let Q be an orthogonal matrix such that $Qe_1 = \bar{p}$, where $e_1 = (1, 0, \ldots, 0)^T$ in $\mathbb{R}^n$. Defining $v = Q^{-1} u$, we finally have

$$K = \{c_0 + X_0^{-1/2} Qv : ||v|| \leq 1, \ \alpha \leq \langle e_1, v \rangle \leq \beta\},$$

that is, $K = c_0 + X_0^{-1/2} Q(\tilde{K})$, where $\tilde{K} = \{v : ||v|| \leq 1, \ \alpha \leq \langle e_1, v \rangle \leq \beta\}$.

Since an affine transformation leaves ratios of volumes unchanged, we assume from here on, without loss of any generality, that our initial convex body K, which we denote by $B_{\alpha\beta}$, has the form

$$B_{\alpha\beta} = \{x \in \mathbb{R}^n : ||x|| \leq 1, \ \alpha \leq x_1 \leq \beta\}, \tag{5.1}$$

where $-1 \leq \alpha < \beta \leq 1$.

Remark 5.1. To simplify our proofs, we assume in this paper that $\beta^2 \geq \alpha^2$. We can always achieve this by working with the convex body $-B_{\alpha\beta}$ instead of $B_{\alpha\beta}$, if necessary.

We use the symmetry properties of $B_{\alpha\beta}$ to determine the possible forms of its extremal ellipsoids. This idea seems to be first suggested in [23] for determining $\mathrm{CE}(B_{\alpha\beta})$.

Lemma 5.2. *If $\alpha = -1$ and $\beta = 1$, then the automorphism group of the slab $B_{\alpha\beta} = B_n$ consists of the $n \times n$ orthogonal matrices. In the remaining cases, the automorphism group $\mathrm{Aut}(B_{\alpha\beta})$ consists of linear transformations $T(u) = Au$ where A is a matrix of the form*

$$A = \begin{bmatrix} a_{11} & 0 \\ 0 & \bar{A} \end{bmatrix}, \quad a_{11} \in \mathbb{R}, \ \bar{A} \in O_{n-1},$$

with $a_{11} = 1$ if $\alpha \neq -\beta$ and $a_{11} = \pm 1$ if $\alpha = -\beta$.

Proof. It is proved in Lemma 4.3 that $\mathrm{Aut}(B_n) = O_n$, so we consider the remaining cases.

Let $T(x) = a + Au$ be an automorphism of $B_{\alpha\beta}$. We write $A = \begin{bmatrix} a_{11} & c^T \\ b & \bar{A} \end{bmatrix}$ and $a = (a_1, \bar{a})$, where a_{11} and a_1 are scalars and the rest of the variables have the appropriate dimensions. Since T is an invertible affine map, $T(\mathrm{ext}(B_{\alpha\beta})) = \mathrm{ext}(B_{\alpha\beta})$, where $\mathrm{ext}(B_{\alpha\beta})$ is the set of extreme points of $B_{\alpha\beta}$ given by

$$\mathrm{ext}(B_{\alpha\beta}) = \left\{ u = (u_1, (1 - u_1^2)^{1/2} \bar{u}) \in \mathbb{R} \times \mathbb{R}^{n-1} : \alpha \leq u_1 \leq \beta, \ ||u|| = 1, \ ||\bar{u}|| = 1 \right\}.$$

If $u = (u_1, (1 - u_1^2)^{1/2} \bar{u})$ is in $\mathrm{ext}(B_{\alpha\beta})$ with $||\bar{u}|| = 1$, then

$$w := a + Au = \begin{bmatrix} a_{11} u_1 + (1 - u_1^2)^{1/2} \langle c, \bar{u} \rangle + a_1 \\ u_1 b + (1 - u_1^2)^{1/2} \bar{A} \bar{u} + \bar{a} \end{bmatrix}.$$

We have $||w|| = 1$, that is,

$$\begin{aligned}
1 &= (a_{11}u_1 + a_1)^2 + (1 - u_1^2)\langle c, \bar{u}\rangle^2 + 2(1 - u_1^2)^{1/2}\langle c, \bar{u}\rangle(a_{11}u_1 + a_1) \\
&\quad + ||u_1 b + \bar{a}||^2 + (1 - u_1^2)||\bar{A}\bar{u}||^2 + 2(1 - u_1^2)^{1/2}\langle \bar{A}\bar{u}, u_1 b + \bar{a}\rangle \\
&= \left\{(a_{11}u_1 + a_1)^2 + ||u_1 b + \bar{a}||^2\right\} + (1 - u_1^2)\left\langle(\bar{A}^T\bar{A} + cc^T)\bar{u}, \bar{u}\right\rangle \\
&\quad + 2(1 - u_1^2)^{1/2}\left\langle(a_{11}u_1 + a_1)c + \bar{A}^T(u_1 b + \bar{a}), \bar{u}\right\rangle \\
&=: q(\bar{u}), \quad \forall ||\bar{u}|| = 1.
\end{aligned} \tag{5.2}$$

Fix $u_1 \in (\alpha, \beta)$, so that $1 - u_1^2 \neq 0$. The argument used in the proof of Lemma 4.3 shows that

$$(a_{11}u_1 + a_1)c + \bar{A}^T(u_1 b + \bar{a}) = 0, \quad \forall\, u_1 \in (\alpha, \beta), \tag{5.3}$$

and that $(1 - u_1)^2\langle(\bar{A}^T\bar{A} + cc^T)\bar{u}, \bar{u}\rangle = \left\{1 - (a_{11}u_1 + a_1)^2 - ||u_1 b + \bar{a}||^2\right\}||\bar{u}||^2$ for all $\bar{u} \in \mathbb{R}^{n-1}$, that is, $(1 - u_1)^2(\bar{A}^T\bar{A} + cc^T) = \left\{1 - (a_{11}u_1 + a_1)^2 - ||u_1 b + \bar{a}||^2\right\}I_{n-1}$, for all $u_1 \in (\alpha, \beta)$. Therefore, there exists a constant k such that

$$\begin{aligned}
kI_{n-1} &= \bar{A}^T\bar{A} + cc^T, \\
0 &= (a_{11}u_1 + a_1)^2 + ||u_1 b + \bar{a}||^2 + k(1 - u_1^2) - 1, \quad \forall\, u_1 \in (\alpha, \beta).
\end{aligned}$$

The equation (5.3) implies the first two equations in (5.4), while the equation above gives rest of the equations below,

$$0 = \bar{A}^T b + a_{11}c, \qquad 0 = \bar{A}^T\bar{a} + a_1 c, \qquad kI_{n-1} = \bar{A}^T\bar{A} + cc^T, \tag{5.4}$$

$$0 = a_{11}^2 + ||b||^2 - k, \qquad 0 = a_1 a_{11} + \langle b, \bar{a}\rangle, \qquad 0 = a_1^2 + ||\bar{a}||^2 + k - 1. \tag{5.5}$$

We have, therefore,

$$A^T A = \begin{bmatrix} a_{11}^2 + ||b||^2 & a_{11}c^T + b^T\bar{A} \\ a_{11}c + \bar{A}^T b & \bar{A}^T\bar{A} + cc^T \end{bmatrix} = \begin{bmatrix} k & 0 \\ 0 & kI_{n-1} \end{bmatrix} = kI_n.$$

Since $|\det A| = 1$ and $A^T A$ is positive semidefinite, we have $k = 1$. This proves that A is an orthogonal matrix. Furthermore, the last equation in (5.5) gives $a = (a_1, \bar{a}) = 0$.

Let x be in $B_{\alpha\beta}$. As $||Ax|| = ||x|| \leq 1$ and $\langle e_1, x\rangle = \langle Ae_1, Ax\rangle$, we have

$$B_{\alpha\beta} = AB_{\alpha\beta} = \{Ax : ||x|| \leq 1, \ \alpha \leq \langle e_1, x\rangle \leq \beta\} = \{x : ||x|| \leq 1, \ \alpha \leq \langle Ae_1, x\rangle \leq \beta\}.$$

If $\alpha \neq -\beta$, then we must have $Ae_1 = e_1$, and if $\alpha = -\beta$, then $B_{\alpha\beta} = -B_{\alpha\beta}$ and we have $Ae_1 = \pm e_1$. Since $Ae_1 = \begin{pmatrix} a_{11} \\ b \end{pmatrix}$, we see that $|a_{11}| = 1$ and $b = c = 0$. It is then clear that $\bar{A}$ belongs to O_{n-1}.

Conversely, it is easy to verify that any matrix A in the form above is in $\mathrm{Aut}(K)$. $\qquad\square$

Lemma 5.3. *The extremal ellipsoids* $\mathrm{CE}(B_{\alpha\beta})$ *and* $\mathrm{IE}(B_{\alpha\beta})$ *have the form* $E(X, c)$ *where* $c = \tau e_1$ *and* $X = \mathrm{diag}(a, b, ..., b)$ *for some* $a > 0$, $b > 0$ *and* τ *in* $\mathbb{R}$. *Moreover, if* $\alpha = -\beta$, *then* $c = 0$.

Proof. Since the proofs are the same, we only consider the ellipsoid $\mathrm{CE}(B_{\alpha\beta})$. Let $U = \begin{bmatrix} 1 & 0 \\ 0 & \bar{U} \end{bmatrix}$ be in $\mathrm{Aut}(K)$. Write $c = (c_1, \bar{c})$. It follows from Theorem 4.5 that $Uc = c$. This implies that $\bar{U}\bar{c} = \bar{c}$ for all $\bar{U}$ in O_{n-1}. Choosing $\bar{U} = -I_{n-1}$, we obtain $\bar{c} = 0$. If $\alpha = -\beta$, then choosing $U = -I_n$ in O_n gives $c = 0$.

Let $\mathrm{CE}(B_{\alpha\beta}) = T(B_n) = E(X,c)$ where $T(x) = c + X^{-1/2}x$. Lemma 4.3 implies that $U^T X U = X$, and writing $X = \begin{bmatrix} x_{11} & v^T \\ v & \bar{X} \end{bmatrix}$, this equation gives $\bar{U}^T v = v$ and $\bar{U}^T \bar{X} \bar{U} = \bar{X}$ for all $\bar{U}$ in O_{n-1}. The first equation implies $v = 0$. In the second equation, we can choose $\bar{U}$ so that the left hand side is a diagonal matrix, proving that $\bar{X}$ itself must be a diagonal matrix. If $\bar{U}$ is the permutation matrix switching columns i and j, then the equation $\bar{U}^T \bar{X} \bar{U} = \bar{X}$ gives $\bar{X}_{ii} = \bar{X}_{jj}$. This proves that $X = \mathrm{diag}(a,b,b,\ldots,b)$ for some $a > 0$, $b > 0$. $\qquad\square$

6 Determination of the minimum volume circumscribed ellipsoid of a slab

In this section, we give explicit formulae for the minimum volume circumscribed ellipsoid of the slab $B_{\alpha\beta}$ in (5.1) using the Fritz John optimality conditions (2.6) and Lemma 5.3. In this section, K will always denote the convex body $B_{\alpha\beta}$.

The following theorem is one of our main results in this paper.

Theorem 6.1. *The minimum volume circumscribed ellipsoid* $\mathrm{CE}(B_{\alpha\beta})$ *has the form* $E(X,c)$ *where* $c = \tau e_1$ *and* $X = \mathrm{diag}(a,b,\ldots,b)$, *where the parameters* $a > 0$, $b > 0$, *and* $\alpha < \tau < \beta$ *are given as follows:*
(i) If $\alpha\beta \leq -1/n$, *then*

$$\tau = 0, \quad \text{and } a = b = 1. \tag{6.1}$$

(ii) If $\alpha + \beta = 0$ *and* $\alpha\beta > -1/n$, *then*

$$\tau = 0, \quad a = \frac{1}{n\beta^2}, \quad b = \frac{n-1}{n(1-\beta^2)}. \tag{6.2}$$

(iii) If $\alpha + \beta \neq 0$ *and* $\alpha\beta > -1/n$, *then*

$$
\begin{aligned}
\tau &= \frac{n(\beta+\alpha)^2 + 2(1+\alpha\beta) - \sqrt{\Delta}}{2(n+1)(\beta+\alpha)}, \\
a &= \frac{1}{n(\tau-\alpha)(\beta-\tau)}, \quad b = \frac{1 - a(\tau-\alpha)^2}{1-\alpha^2},
\end{aligned}
\tag{6.3}
$$

where $\Delta = n^2(\beta^2 - \alpha^2)^2 + 4(1-\alpha^2)(1-\beta^2)$.

We remark that Corollary 2.8 also implies the converse of Part (i) in the theorem.

Proof. Lemma 5.3 implies that $X = \mathrm{diag}(a,b,\ldots,b)$ and $c = \tau e_1$. Writing $u_i = (y_i, z_i) \in \mathbb{R} \times \mathbb{R}^{n-1}$, $i = 1,\ldots,k$, where $1 = ||u_i||^2 = y_i^2 + ||z_i||^2$, that is $||z_i||^2 = 1 - y_i^2$, and noting that $(u_i - c)(u_i - c)^T = \begin{bmatrix} (y_i - \tau)^2 & (y_i - \tau)z_i^T \\ (y_i - \tau)z_i & z_i z_i^T \end{bmatrix}$, the Fritz John (necessary and sufficient) optimality conditions (2.6) may be written in the form

21

$$\tau = \frac{1}{n}\sum_{i=1}^{k}\lambda_i y_i, \quad 0 = \sum_{i=1}^{k}\lambda_i z_i, \quad \sum_{i=1}^{k}\lambda_i = n, \tag{6.4}$$

$$\frac{1}{a} = \sum_{i=1}^{k}\lambda_i(y_i - \tau)^2, \quad 0 = \sum_{i=1}^{k}\lambda_i(y_i - \tau)z_i, \quad \frac{1}{b}I_{n-1} = \sum_{i=1}^{k}\lambda_i z_i z_i^T, \tag{6.5}$$

$$0 = a(y_i - \tau)^2 + b(1 - y_i^2) - 1, \quad i = 1,\ldots,k, \tag{6.6}$$

$$0 \geq a(y - \tau)^2 + b(1 - y^2) - 1, \quad \forall\, y \in [\alpha, \beta]. \tag{6.7}$$

The last line expresses the feasibility condition $K \subseteq E(X, c)$: any point $x = (y, z)$ satisfying $\alpha \leq y \leq \beta$ and $||x|| = 1$ lies in K, hence in $E(X, c)$, so that it satisfies the conditions $y^2 + ||z||^2 = 1$ and $a(y - \tau)^2 + b||z||^2 \leq 1$.

The conditions (6.4)–(6.7) thus characterize the ellipsoid $\mathrm{CE}(K)$ for $K = B_{\alpha\beta}$ in (5.1). Since the ellipsoid $\mathrm{CE}(K)$ is unique, its paremeters (τ, a, b) are unique and can be recovered from the above conditions. These are done in the technical lemmas below. $\qquad\square$

Lemma 6.2. *If $a = b$ in the ellipsoid $\mathrm{CE}(B_{\alpha\beta})$, then $\tau = 0$, $a = b = 1$, and $\alpha\beta \leq -1/n$.*

Proof. Since $a = b$, (6.6) gives the equation $2a\tau y_i = a\tau^2 + a - 1$. We have $\tau = 0$, since otherwise all y_i are the same, and the first and third equations in (6.4) imply that $y_i = \tau$, contradicting the first equation in (6.5). The equation $2a\tau y_i = a\tau^2 + a - 1$ reduces to $a = 1 = b$. Finally, since $\alpha \leq y_i \leq \beta$, we obtain

$$0 \geq \sum_{i=1}^{k}\lambda_i(y_i - \alpha)(y_i - \beta) = \sum_{i=1}^{k}\lambda_i y_i^2 - (\alpha + \beta)\sum_{i=1}^{k}\lambda_i y_i + \alpha\beta\sum_{i=1}^{k}\lambda_i = 1 + n\alpha\beta,$$

where the last equation follows since $\sum_{i=1}^{k}\lambda_i y_i^2 = 1$ from the first equation in (6.5) and $\sum_{i=1}^{k}\lambda_i y_i = 0$, $\sum_{i=1}^{k}\lambda_i = n$ from (6.4). $\qquad\square$

Lemma 6.3. *If $a \neq b$ in the ellipsoid $\mathrm{CE}(B_{\alpha\beta})$, then $a > b$ and the leading coordinate y_i of a contact point must be α or β, and both values are taken.*

Proof. Observe that the function $g(y) := a(y - \tau)^2 + b(1 - y^2) - 1$ in (6.7) is nonpositive on the interval $I = [\alpha, \beta]$ and equals zero at each y_i. We claim that y_i can not take a single value: otherwise the first and third equations in (6.4) imply that $y_i = \tau$, contradicting the first equation in (6.5). (This result also follows from Corollary 2.5.) Since g is a quadratic function, y_i must take exactly two values, and (6.7) implies that these two values must coincide with the endpoints of the interval I. Furthermore, $g(y) \leq 0$ only on I, has a global minimizer there, and so it must be a strictly convex function. This proves that $a > b$. $\qquad\square$

Lemma 6.4. *If $a \neq b$ in the ellipsoid $\mathrm{CE}(B_{\alpha\beta})$, then (τ, a, b) are given by the equations (6.2) and (6.3). Moreover, $\alpha\beta > -1/n$.*

Proof. Lemma 6.3 and equation (6.6) give $a(\beta - \tau)^2 + b(1 - \beta^2) = 1$ and $a(\alpha - \tau)^2 + b(1 - \alpha^2) = 1$. Subtracting the second equation from the first and dividing by $\beta - \alpha \neq 0$ yields the equation $\tau = (1 - b/a)(\alpha + \beta)/2$. We also have

22

$$0 = \sum_{i=1}^{k} \lambda_i(y_i - \alpha)(y_i - \beta) = \sum_{i=1}^{k} \lambda_i[(y_i - \tau) - (\alpha - \tau)] \cdot [(y_i - \tau) - (\beta - \tau)]$$

$$= \sum_{i=1}^{k} \lambda_i(y_i - \tau)^2 - (\alpha + \beta - 2\tau) \sum_{i=1}^{k} \lambda_i(y_i - \tau) + (\alpha - \tau)(\beta - \tau) \sum_{i=1}^{k} \lambda_i$$

$$= \frac{1}{a} + n(\alpha - \tau)(\beta - \tau),$$

where the first equation follows from Lemma 6.3, the last equation from (6.4) and (6.5).

Altogether, we have the equations

$$1 = a(\alpha - \tau)^2 + b(1 - \alpha^2), \quad \tau = \left(1 - \frac{b}{a}\right) \cdot \frac{\alpha + \beta}{2}, \quad \frac{1}{a} = n(\tau - \alpha)(\beta - \tau), \qquad (6.8)$$

which we use to compute the variables a, b, and τ.

If $\alpha = -\beta$, then the second equation above gives $\tau = 0$. Then the third and first equations in (6.8) give $a = 1/(n\alpha^2)$ and $b = (n-1)/(n(1-\alpha^2))$, respectively. Lastly, Lemma 6.3 gives $a > b$, and this implies $1 + n\alpha\beta > 0$.

If $\beta \neq -\alpha$, then the first and third equations in (6.8) give $(\alpha - \tau)^2 + (b/a)(1 - \alpha^2) = n(\tau - \alpha-)(\beta - \tau)$ and the second equation gives $b/a = 1 - 2\tau/(\alpha + \beta)$. Substituting this value of b/a in the preceding one leads to the quadratic equality for τ,

$$(n + 1)(\alpha + \beta)\tau^2 - \left(n(\alpha + \beta)^2 + 2(1 + \alpha\beta)\right)\tau + (\alpha + \beta)(1 + n\alpha\beta) = 0. \qquad (6.9)$$

A straightforward but tedious calculation shows that the discriminant is $\Delta = n^2(\beta^2 - \alpha^2)^2 + 4(1 - \beta^2)(1 - \alpha^2) > 0$. We claim that the feasible root is the one with negative discriminant. If $\overline{\tau}$ is the root with positive discriminant, then $\overline{\tau} - \beta = [n(\alpha^2 - \beta^2) + 2(1 - \beta^2) + \sqrt{\Delta}]/(2(n+1)(\alpha+\beta)) \geq 0$. Recalling that $\beta^2 \geq \alpha^2$, we have $[n(\alpha^2 - \beta^2) + 2(1 - \beta^2)]^2 - \Delta = 4(n + 1)(1 - \beta^2)(\alpha^2 - \beta^2) \leq 0$. This gives $\overline{\tau} \geq \beta$, proving the claim, as we must have $\alpha < \tau < \beta$. Therefore,

$$\tau = \frac{n(\alpha + \beta)^2 + 2(1 + \alpha\beta) - \sqrt{\Delta}}{2(n + 1)(\alpha + \beta)},$$

$$a = \frac{1}{n(\tau - \alpha)(\beta - \tau)}, \quad b = \frac{1 - a(\alpha - \tau)^2}{1 - \alpha^2},$$

where the equations for a and b follow from the first and second equations in (6.8).

Finally, $\tau = (1 - b/a)(\alpha + \beta)/2$ from (6.8) and Lemma 6.3 gives $a > b$, implying $\tau > 0$. From the formula above for τ, we get

$$0 < (n(\alpha + \beta)^2 + 2(1 + n\alpha\beta))^2 - \Delta = 4(n + 1)(\alpha + \beta)^2(1 + n\alpha\beta).$$

This gives $1 + n\alpha\beta > 0$. The lemma is proved. $\qquad \square$

Remark 6.5. Some of the results contained in Theorem 6.1 may seem very counter–intuitive. For example, consider the slab $B_{\alpha\beta}$ when $-\alpha = \beta = 1/\sqrt{n}$. Although the width of this slab is $2/\sqrt{n}$, very small for large n, the optimal covering ellipsoid is the unit ball. This seemingly improbable behavior may be explained by the *concentration of measure* phenomenon: most of the volume of a high dimensional ball is concentrated in a thin strip around the "equator", see for example [2]. There is a sizable literature on concentration of measure; the interested reader may consult the reference [25] for more information on this important topic.

6.1 Determination of the covering ellipsoid by nonlinear programming

In this section, we give a proof of Theorem 6.1 which is completely independent of the previous one based on semi–infinite programming. The proof uses the uniqueness of the covering ellipsoid, Lemma 5.3 on the form of the optimal ellipsoid, and Corollary 2.5. We thus need proofs of the first and the last results that do not depend on the results of §2 and Theorem 6.1. We note that an elementary and direct proof of the uniqueness of the optimal covering ellipsoid can be found, for example, in Danzer et al. [11], and we supply an independent, direct proof of Corollary 2.5. This last result is not strictly necessary, but it simplifies our proofs, and it may be of independent interest.

Recall that Corollary 2.5 states that the contact points of an extremal covering ellipsoid cannot lie any half space whose bounding hyperplane passes through the center of the ellipsoid. The following is an independent proof of this fact, in the spirit of the proof in Grunbaum [15] for the maximum volume inscribed ellipsoid.

Proof. We assume without loss of generality that the ellipsoid is the unit ball B_n. Clearly, it suffices to show that the open halfspace $B^+ := \{x : x_n > 0\}$ contains a point of $E \cap \partial K$. We prove this by contradiction.

Consider the ellipsoids $E(\lambda) = \{x : f(x) = a\sum_{i=1}^{n-1} x_i^2 + b(x_n + \lambda)^2 \le 1\}$, having on their boundary the points $\{e_i\}_1^{n-1}$ and $-e_n = (0,0,\ldots,0,-1)$. We have $b = 1/(1-\lambda)^2$, $a = 1 - \lambda^2/(1-\lambda)^2 = (1-2\lambda)/(1-\lambda)^2$, and

$$\text{vol}(E(\lambda)) = \left(a^{n-1}b\right)^{-1} = \frac{(1-\lambda)^{2n}}{(1-2\lambda)^{n-1}}.$$

We claim that $K \subseteq E(\lambda)$ for small enough $\lambda > 0$. On the one hand, if $\|x\| = 1$ and $x_n \le 0$, we have $f(x) - 1 = \frac{2\lambda}{(1-\lambda)^2}(x_n^2 + x_n) \le 0$. On the other hand, since B^+ contains no contact points, there exists $\epsilon > 0$ such $\|x\| < 1 - \epsilon$ for all $x \in K \cap B^+$. It follows by continuity that $K \cap B^+ \subset E(\lambda)$ for small enough $\lambda > 0$. These prove the claim.

Lastly, $\text{vol}(E(\lambda)) < 1$, since $(1-\lambda)^{2n} - (1-2\lambda)^{n-1} = [1 - 2n\lambda + o(\lambda)] - [1 - (n-1)(-2\lambda) + o(\lambda)] = -2\lambda + o(\lambda) < 0$ for small $\lambda > 0$. □

Theorem 6.6. *The minimum volume covering ellipsoid problem for the slab $B_{\alpha\beta}$ can be formulated as the nonlinear programming problem*

$$
\begin{aligned}
\min \quad & -\ln a - (n-1)\ln b, \\
\text{s.t.} \quad & a\tau - \left(\frac{\alpha+\beta}{2}\right)(a-b) = 0, \\
& a\tau^2 + b - 1 - \alpha\beta(a-b) = 0, \\
& -a + b \le 0,
\end{aligned}
\tag{6.10}
$$

whose solution is the same as the one given in Theorem 6.1.

Proof. It follows from Lemma 5.3 that the optimal ellipsoid $E(X,c)$ has the form $X = \text{diag}(a,b,\ldots,b)$ and $c = \tau e_1$. Thus, the feasibility condition $B_{\alpha\beta} \subseteq E(X,c)$ translates into the condition that the quadratic function

$$g(u) = a(u-\tau)^2 + b(1-u^2) - 1$$

24

is non–positive on the interval $I = [\alpha, \beta]$. Furthermore, Corollary 2.5 implies that there exist at least two contact points, which translates into the condition that the quadratic function $g(u)$ takes the value zero at two distinct points in the interval I. A moment's reflection shows that $g(u)$ must take the value zero at the endpoints α and β. Consequently, the function g is a non–negative multiple of the function $(u - \alpha)(u - \beta)$, that is

$$g(u) + \mu(u - \alpha)(\beta - u) = 0, \quad \text{for some} \quad \mu \geq 0,$$

giving $a - b = \mu \geq 0$, $(\alpha + \beta)\mu - 2a\tau = 0$, and $a\tau^2 + b - 1 - \alpha\beta\mu = 0$. If we eliminate μ from these constraints, we arrive at the optimization problem (6.10). We have for it the Fritz John optimality conditions (for ordinary nonlinear programming)

$$\lambda_1(\tau - \frac{\alpha + \beta}{2}) + \lambda_2(\tau^2 - \alpha\beta) - \lambda_3 = \frac{\lambda_0}{a},$$
$$\lambda_1(\frac{\alpha + \beta}{2}) + \lambda_2(1 + \alpha\beta) + \lambda_3 = \frac{\lambda_0(n - 1)}{b}, \quad (6.11)$$
$$\lambda_1 + 2\lambda_2\tau = 0,$$

for some $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) \neq 0$, $\lambda_0 \geq 0$, $\lambda_3 \geq 0$, and satisfying $\lambda_3(a - b) = 0$.

Adding the first two equations above and substituting $\lambda_1 = -2\lambda_2\tau$ from the third equation gives

$$\lambda_2(1 - \tau^2) = \frac{\lambda_0}{a} + \frac{\lambda_0(n - 1)}{b}.$$

If $\lambda_0 = 0$, we would have $\lambda_2(1 - \tau^2) = 0$, and since $\tau \neq \pm 1$, $\lambda_2 = 0$, and eventually $\lambda_1 = 0 = \lambda_3$, a contradiction. Thus, we may assume that $\lambda_0 = 1$.

Then the above equation gives

$$\lambda_2(1 - \tau^2) = \frac{1}{a} + \frac{n - 1}{b}. \quad (6.12)$$

We solve for the decision variables (a, b, τ) discussing separately the cases $a = b$ and $a \neq b$ in the above optimality conditions. If $a = b$, then the first constraint in (6.10) gives $\tau = 0$ and the second constraint gives $b = 1$. It remains to prove that $\alpha\beta \leq -1/n$. The equation (6.12) gives $\lambda_2 = n$, and $\lambda_1 = -2\lambda_2\tau = 0$. Substituting these in the first equation in (6.11) yields $0 \leq \lambda_3 = -n\alpha\beta - 1$, that is, $\alpha\beta \leq -1/n$.

We now consider the case $a \neq b$ but $\alpha + \beta = 0$. The first constraint in problem (6.10) gives $\tau = 0$ and the third one gives $\lambda_3 = 0$. Consequently, the first two conditions in (6.11) can be written as $\lambda_2\beta^2 a = 1$ and $\lambda_2(1 - \beta^2)b = n - 1$, respectively, and the second constraint in (6.10) gives $\beta^2 a + (1 - \beta^2)b = 1$. These imply $\lambda_2 = n$, and $a = 1/n\beta^2$, $b = \frac{n - 1}{n(1 - \beta^2)}$. Lastly, the condition $a > b$ gives $\alpha\beta > -1/n$.

Finally, we treat the case $a > b$ and $\alpha + \beta \neq 0$. Again we have $\lambda_3 = 0$ and

$$(n - 1)\frac{-\tau^2 + (\alpha + \beta)\tau - \alpha\beta}{-(\beta + \alpha)\tau + (1 + \alpha\beta)} = \frac{b}{a} = \frac{\beta + \alpha - 2\tau}{\beta + \alpha}.$$

Here the first equality is obtained by dividing the first equation in (6.11) by the second one and substituting $\lambda_1 = -2\lambda_2\tau$, and the second equality follows from the first constraint in problem (6.10). Consequently, τ satisfies the quadratic equality

$$(n + 1)(\alpha + \beta)\tau^2 - [n(\alpha + \beta)^2 + 2(1 + \alpha\beta)]\tau + (\alpha + \beta)(1 + n\alpha\beta) = 0, \quad (6.13)$$

25

which is the same equation as (6.9) in the proof of Lemma 6.4. Following similar arguments, we find that

$$\tau = \frac{n(\alpha + \beta)^2 + 2(1 + \alpha\beta) - \sqrt{\Delta}}{2(n + 1)(\alpha + \beta)}.$$

Solving the first and the second constraints in (6.10) for a, say by Cramer's rule, we find

$$a = \frac{\alpha + \beta}{(\alpha + \beta)(\tau^2 + 1) - 2\tau(1 + \alpha\beta)}.$$

It follows from (6.13) that the denominator on the right hand side of the expression above equals

$$n(\alpha + \beta)^2 \tau - n(\alpha + \beta)\tau^2 - n(\alpha + \beta)\alpha\beta = n(\alpha + \beta)(\tau - \alpha)(\beta - \tau).$$

This gives

$$a = \frac{1}{n(\tau - \alpha)(\beta - \tau)}, \qquad b = \frac{\alpha + \beta - 2\tau}{\alpha + \beta}a.$$

Finally, the inequality $\alpha\beta > -1/n$ follows from the same argument at the end of the proof of Lemma 6.4. $\square$

7 Determination of the maximum volume inscribed ellipsoid of a slab

In this section, we give explicit formulae for the maximum volume inscribed ellipsoid of the slab $B_{\alpha\beta}$ in (5.1) using a semi–infinite programming approach. Without any loss of generality, we again assume throughout this section that $\beta^2 \geq \alpha^2$.

It is convenient to set up this problem as the semi–infinite program

$$\min\left\{-\ln\det(A) : Ay + c \in B_{\alpha\beta}, \quad \forall y : ||y|| = 1\right\},$$

in which we represent the inscribed ellipsoid as $E = c + A(B_n)$ where A is a symmetric, positive definite matrix with $\mathrm{vol}(E) = \omega_n \det A$. Lemma 5.3 implies that the optimal ellipsoid has the form $A = \mathrm{diag}(a, b, \ldots, b)$ and $c = \tau e_1$ for some $a > 0$, $b > 0$, and τ in $\mathbb{R}$. Writing $y = (u, z)$ in $\mathbb{R} \times \mathbb{R}^{n-1}$, we can replace the above semi–infinite program with a far simpler one

$$
\begin{aligned}
\min \quad & -\ln a - (n - 1)\ln b, \\
\text{s.t.} \quad & -au - \tau \leq -\alpha, \\
& au + \tau \leq \beta, \qquad\qquad\qquad \forall u \in [-1, 1] \\
& (au + \tau)^2 + b^2(1 - u^2) \leq 1,
\end{aligned}
\tag{7.1}
$$

in which the decision variables are (a, b, τ) and the index set is $[-1, 1]$.

We make some useful observations before writing down the optimality conditions for Problem 7.1. Theorem 2.1 implies that the optimality conditions will involve at most three active constraints with corresponding multipliers positive. Clearly, the first constraint above can be active only for $u = -1$ and the second one for $u = 1$.

Next, if $\alpha > -1$, we claim that the third constraint is active for at most one index value u. Otherwise, the quadratic function

$$g(u) := (au + \tau)^2 + b^2(1 - u^2) - 1$$

is non–positive on the interval $[-1, 1]$ and equals zero for two distint values of u. If the function $g(u)$ is actually a linear function (a=b), then it is identically zero; otherwise, it is easy to see that $g(u)$ must equal zero at the endpoints -1 and 1. In all cases, we have $g(-1) = g(1) = 0$, so that $(-a + \tau)^2 = 1 = (a + \tau)^2$. This gives $a - \tau = 1 = a + \tau$, since the other possibilities give $a = 0$ or $a = -1$. But then $a = 1$, $\tau = 0$, and $1 = a - \tau \le -\alpha$, where the inequality expresses the feasibility of the first inequality in Problem 7.1. We obtain $\alpha = -1$ and $\beta = 1$, a contradiction. The claim is proved.

The following theorem is another major result of this paper.

Theorem 7.1. *The maximal inscribed ellipsoid* $\mathrm{IE}(B_{\alpha\beta})$ *has the form* $E = c + A(B_n)$ *where* $c = \tau e_1$, $\alpha < \tau < \beta$, $A = \mathrm{diag}(a, b, \ldots, b)$ *with* $a > 0$, $b > 0$, *satisfying the following conditions:*
(i) *If* $\alpha = -\beta$, *then*

$$\tau = 0, \qquad a = \beta, \qquad b = 1. \tag{7.2}$$

(ii) *If* $4n(1 - \alpha^2) < (n + 1)^2(\beta^2 - \alpha^2)$, *then*

$$\tau = \frac{\alpha + \sqrt{\alpha^2 + 4n(1 - \alpha^2)/(n + 1)^2}}{2},$$

$$a = \tau - \alpha, \qquad b^2 = a(a + n\tau), \tag{7.3}$$

(iii) *If* $4n(1 - \alpha^2) \ge (n + 1)^2(\beta^2 - \alpha^2)$ *and* $\alpha \ne -\beta$, *then*

$$\tau = \frac{\beta + \alpha}{2}, \qquad a = \frac{\beta - \alpha}{2},$$

$$b^2 = a^2 + \left(\frac{\beta^2 - \alpha^2}{2(\sqrt{1 - \alpha^2} - \sqrt{1 - \beta^2})}\right)^2. \tag{7.4}$$

Proof. If $\alpha = -1$, then $\beta = 1$ and the optimal ellipsoid is the unit ball B_n, which agrees with (i) of the theorem. We assume in the rest of the proof that $\alpha > -1$.

We saw above that each of the constraints in 7.1 can be active for at most one value of u in $[-1, 1]$, and the first and second constraints for $u = -1$ and $u = 1$, respectively. Then Theorem 2.1 gives the optimality conditions

$$\lambda_1 + \lambda_2 + \delta(au + \tau)u = \frac{\lambda_0}{a},$$

$$\delta b(1 - u^2) = \frac{(n - 1)\lambda_0}{b}, \qquad u \in [-1, 1], \tag{7.5}$$

$$-\lambda_1 + \lambda_2 + \delta(au + \tau) = 0,$$

where the non–negative multipliers $(\lambda_0, \lambda_1, \lambda_2, \delta/2) \ne 0$ correspond to the objective function, and the first, second, and third (active) constraints in 7.1, respectively.

Our *first* claim is that $\lambda_0 > 0$. Otherwise, the second equation in (7.5) gives $\delta = 0$ or $u = \mp1$. If $\delta = 0$, then the first equation gives the contradiction $\lambda_1 = \lambda_2 = 0$. If $u = -1$, then we have $\mp1 = a - \tau \le -\alpha < 1$, implying $a - \tau = -1$ or $\tau = a + 1 > 1$, another

contradiction. If $u = 1$, we have $\mp 1 = a + \tau$. If $a + \tau = 1$, then the first equation in (7.5) gives $\lambda_1 + \lambda_2 + \delta = 0$, that is, $\lambda_1 = \lambda_2 = \delta = 0$, a contradiction. If $a + \tau = -1$, then $\tau = -a - 1 < -1$, yet another contradiction. The claim is proved. We set $\lambda_0 = 1$.

The second equation in (7.5) gives $\delta > 0$, and that $g(u) = 0$ for some u in $(-1, 1)$ and negative elsewhere on $[-1, 1]$. Note this means that u is the global maximum as well as the unique root of g on $\mathbb{R}$, leading to the conditions

$$b > a, \quad u = \frac{a\tau}{b^2 - a^2}, \quad b^2\tau^2 = (1 - b^2)(b^2 - a^2), \tag{7.6}$$

where the last equation expresses the fact that the discriminant of g equals zero. We will also have occasion to use the equality

$$au + \tau = a\frac{a\tau}{b^2 - a^2} + \tau = \frac{b^2\tau}{b^2 - a^2}. \tag{7.7}$$

Our *second* claim is that

$$-a + \tau = \alpha. \tag{7.8}$$

If not, then $-a + \tau > \alpha$, $\lambda_1 = 0$, and the third equation in (7.5) gives $\lambda_2 = -\delta(au + \tau)$. Substituting this into the first equation in (7.5) leads to $\delta(au + \tau)(u - 1) = a^{-1}$, and consequently to $au + \tau < 0$. But then $\lambda_2 > 0$ and $a + \tau = \beta$, which together with $-a + \tau > \alpha$ gives $\tau > (\alpha + \beta)/2 \geq 0$, that is, $\tau > 0$. Moreover, the second equation in (7.6) gives $u > 0$, and this leads to the contradiction that $-\lambda_2 = \delta(au + \tau) > 0$. The claim is proved.

Next, we have

$$\frac{1}{a} - 2\lambda_2 = \delta(au + \tau)(u + 1) = \frac{n - 1}{b^2(1 - u^2)} \cdot \frac{b^2\tau}{b^2 - a^2}(u + 1)$$
$$= \frac{(n - 1)\tau}{(1 - u)(b^2 - a^2)} = \frac{(n - 1)u}{a(1 - u)},$$

where the first equality is obtained by adding the first and third equations in (7.5), the second equality follows by substituting the value of δ from the second equation in (7.5) and the value of $au + \tau$ from (7.7), and the last equality follows by substituting the value $\tau/(b^2 - a^2) = u/a$ from the first equation in (7.6). Therefore,

$$\lambda_2 = \frac{1 - nu}{2a(1 - u)}.$$

Consequently, $u \leq 1/n$ and $u = 1/n$ if and only if $\lambda_2 = 0$. Furthermore, we have $u \geq 0$: if $u < 0$, $\tau < 0$ by virtue of the first equation in (7.6), so that $au + \tau < 0$. But then the third equation in (7.5) gives $\lambda_2 > \lambda_1 \geq 0$, implying $a + \tau = \beta$. This and (7.8) give $\tau = (\alpha + \beta) \geq 0$, a contradiction. Therefore,

$$0 \leq u \leq \frac{1}{n}, \quad \text{and} \quad u = \frac{1}{n} \iff \lambda_2 = 0. \tag{7.9}$$

We can now prove part (i) of the theorem. We first show that

$$u = 0 \iff \alpha = -\beta.$$

If $\alpha = -\beta$, then Lemma 5.3 implies that $\tau = 0$, which in turn implies $u = 0$. Conversely, if $u = 0$, then $\tau = 0$ and $\lambda_2 > 0$, and we have $-a + \tau = \alpha$ and $a + \tau = \beta$, leading to

28

$0 = \tau = (\beta + \alpha)/2$. Consequently, (7.8) gives $a = -\alpha = \beta$ and the second equation in (7.6) gives $b = 1$.

We now consider the remaining cases $0 < u \leq 1/n$. We note that

$$(au + \tau)u\tau = \frac{u(b^2\tau^2)}{b^2 - a^2} = u(1 - b^2) = u(1 - a^2) - a\tau,$$

where the first equality follows from (7.7) and last two equalities from (7.6), leading to a quadratic equation for u,

$$(a\tau)u^2 - (1 - a^2 - \tau^2)u + a\tau = 0. \tag{7.10}$$

Define $\varepsilon \geq 0$ such that $a + \tau = \beta - \varepsilon =: \beta_\varepsilon$. The equation $a + \tau = \beta_\varepsilon$ together with equation $-a + \tau = \alpha$ from (7.8) give

$$\tau = \frac{\beta_\varepsilon + \alpha}{2} > 0, \quad a = \frac{\beta_\varepsilon - \alpha}{2} > 0. \tag{7.11}$$

Substituting these in (7.10) gives another quadratic equality for u,

$$(\beta_\varepsilon^2 - \alpha^2)u^2 - 2(2 - \beta_\varepsilon^2 - \alpha^2)u + (\beta_\varepsilon^2 - \alpha^2) = 0. \tag{7.12}$$

It is easy to verify, using (7.9), that

$$
\begin{aligned}
\lambda_2 > 0 &\iff 0 < u < \frac{1}{n} &\iff (n+1)^2(\beta_\varepsilon^2 - \alpha^2) < 4n(1 - \alpha^2), \\
\lambda_2 = 0 &\iff u = \frac{1}{n} &\iff (n+1)^2(\beta_\varepsilon^2 - \alpha^2) = 4n(1 - \alpha^2).
\end{aligned} \tag{7.13}
$$

Here the second equivalence on the first line follows because the quadratic equation for u in (7.12) has negative value at $u = 1/n$. Since the leading term of the quadratic function is positive, its two roots $r_1 < r_2$ are positive, their product is 1, $0 < r_1 < 1/n$, and $r_2 > n$.

We now make our *third* and important claim that

$$a + \tau < \beta \quad \text{iff} \quad 4n(1 - \alpha^2) < (n+1)(\beta^2 - \alpha^2). \tag{7.14}$$

On the one hand, if $a + \tau = \beta$, then $\varepsilon = 0$ and (7.13) gives $(n+1)(\beta^2 - \alpha^2) < 4n(1 - \alpha^2)$ or $(n+1)(\beta^2 - \alpha^2) = 4n(1 - \alpha^2)$, depending on whether $\lambda_2 > 0$ or $\lambda_2 = 0$, respectively. In either case, we have $(n+1)(\beta^2 - \alpha^2) \leq 4n(1 - \alpha^2)$. On the other hand, if $a + \tau < \beta$, then $\lambda_2 = 0$, $\varepsilon > 0$, and (7.13) gives $(n+1)(\beta_\varepsilon^2 - \alpha^2) = 4n(1 - \alpha^2)$, that is, $4n(1 - \alpha^2) < (n+1)(\beta^2 - \alpha^2)$. The claim is proved.

The computation of the decision variables (a, b, τ) in the cases (ii) and (iii) now becomes a routine matter. If $4n(1 - \alpha^2) < (n+1)^2(\beta^2 - \alpha^2)$, then (7.14) implies $a + \tau < \beta$, and we have $\lambda_2 = 0$, $u = 1/n$. Substituting the value $a = \tau - \alpha$ from (7.8) into (7.10) gives the quadratic equation for τ,

$$(n+1)^2\tau^2 - (n+1)^2\alpha\tau - n(1 - \alpha^2) = 0.$$

Since $\tau > 0$, the feasible root is given by

$$\tau = (\alpha + \sqrt{\alpha^2 + 4n(1 - \alpha^2)/(n+1)^2})/2.$$

Then (7.8) gives $a = \tau - \alpha$ and (7.6) gives $1/n = u = (a\tau)/(b^2 - a^2)$, that is, $b^2 = a^2 + na\tau$.

If $4n(1 - \alpha^2) \geq (n + 1)^2(\beta^2 - \alpha^2)$, then (7.14) and (7.12) give

$$u = \frac{\left(\sqrt{1 - \alpha^2} \mp \sqrt{1 - \beta^2}\right)^2}{\beta^2 - \alpha^2}.$$

It is easy to verify that the condition $u < 1$ is equivalent to $\sqrt{1 - \beta^2} \mp \sqrt{1 - \alpha^2} < 0$, which is impossible if we choose the plus sign. Thus the feasible root is the one with the negative sign. Finally, the first equation in (7.6) gives $b^2 = a^2 + (a\tau)/u$, or more explicitly the formula for b^2 in (7.4). $\qquad\square$

We end this section by reducing the semi–infinite program (7.1) to an ordinary nonlinear programming problem. However, we do not attempt to solve the resulting program in order to save space. As we already noted, the linear constraints, the first two inequalities in problem (7.1), simply reduce to the constraints $a - \tau \leq -\alpha$ and $a + \tau \leq \beta$. In order to reduce the quadratic inequality system to a set of ordinary inequalities, we use a theorem of Lukács characterizing the class of non–negative polynomials on a given interval. A simple inductive proof of Lukács's Theorem can be found in [7]. For a quadratic polynomial $q(u)$ on the interval $I = [a, b]$, this theorem states that q is non–negative on I if and only if there exist scalars α, β, and $\gamma \geq 0$ such that

$$q(u) = (\alpha u + \beta)^2 + \gamma(u - a)(b - u). \tag{7.15}$$

Since the proof is short and simple in this case, we give it, following [7]. Note that the polynomial $p(u) := q(u) - l(u)^2$, where $l(u) = [\sqrt{q(a)}(u - b) + \sqrt{q(b)}(u - a)]/(b - a)$, satisfies $p(a) = 0 = p(b)$, so that there exists a constant γ such that $p(u) = \gamma(u - a)(b - u)$. Clearly, (7.15) holds true if we can show that $\gamma \geq 0$. Note that $l(a) \leq 0$ and $l(b) \geq 0$ so that $l(u) = d(u - r)$ for some constant d and $r \in [a, b]$. We have $0 \leq q(u) = d^2(u - r)^2 + \gamma(u - a)(b - u)$, or

$$-d^2(u - r)^2 \leq \gamma(u - a)(b - u), \quad \forall\, u \in [a, b].$$

If r is in (a, b), then choosing $u = r$ gives $\gamma \geq 0$. If $r = a$ or $r = b$, then choosing u near r gives $\gamma \geq 0$. This completes the proof.

We remark that the result we just proved also follows from the one dimensional case of the S–procedure, see [31] or [30].

Applying this result to our quadratic function $-q(u) = -(au + \tau)^2 - b^2(1 - u^2) + 1$ which is non–negative on the interval $[-1, 1]$, we see that there exists scalars c, d, and $\gamma \geq 0$ such that

$$(au + \tau)^2 + b^2(1 - u^2) - 1 = -(cu + d)^2 + \gamma(u^2 - 1),$$

that is, $a^2 - b^2 + c^2 - \gamma = 0$, $a\tau + cd = 0$, and $b^2 + \tau^2 + d^2 + \gamma = 1$. Therefore, the problem of finding $\mathrm{IE}(B_{\alpha\beta})$ reduces to the nonlinear programming problem

$$
\begin{aligned}
\min \quad & -\ln a - (n - 1)\ln b, \\
\text{s.t.} \quad & -a + \tau \geq \alpha, \\
& a + \tau \leq \beta, \\
& a^2 - b^2 + c^2 - \gamma = 0, \\
& a\tau + cd = 0, \\
& b^2 + \tau^2 + d^2 + \gamma = 1, \\
& \gamma \geq 0,
\end{aligned}
$$

in which the decision variables are $(a, b, \tau, c, d, \gamma)$ and α, β are parameters.

8 Determination of the minimum volume covering ellipsoid of a truncated second order cone or a cylinder

In this section, one of the problems we are interested in is finding the minimum volume ellipsoid covering the truncated second order cone

$$K = \{x = (x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} : ||B(\bar{x} - c)|| \leq x_1, \ a \leq x_1 \leq b\},$$

where B is an invertible matrix in $\mathbb{R}^{(n-1)\times(n-1)}$ and $0 \leq a < b$ are constants. By an affine change of $\bar{x}$, we may assume that $c = 0$ and $B = I_{n-1}$. We claim that by further affine change of variables, we can reduce the convex body K to have the form

$$Q_{\alpha\beta} := \text{conv}(S_\alpha \cup S_\beta),$$

where $S_\alpha := \{x \in B_n : x_1 = \alpha\}$, $S_\beta := \{x \in B_n : x_1 = \beta\}$, and $-1 \leq \alpha < \beta \leq 1$. Consider the ball B in $\mathbb{R}^n$ with center $(a+b)e_1$ and radius $\sqrt{a^2 + b^2}$. The slice $P_a := \{(a, \bar{x}) \in \mathbb{R}^n : ||\bar{x}|| \leq a\} \subset K$ lies in B, since $||(a, \bar{x}) - (a+b, 0)||^2 = b^2 + ||\bar{x}||^2 \leq a^2 + b^2$, and similarly $P_b \subset B$. A further translation and then scaling transforms B into B_n. This proves the claim. Conversely, if $-\alpha \neq \beta$, $Q_{\alpha\beta}$ can be viewed as a truncated second order cone.

The other problem we are interested in is finding the minimum volume ellipsoid covering a cylinder. If $-\alpha = \beta$, then $Q_{\alpha\beta}$ is clearly a cylinder. Conversely, any cylinder can be transformed into such a $Q_{\alpha\beta}$ by an affine transformation. Consequently, the $\text{CE}(Q_{\alpha\beta})$ problem formulates both problems at the same time.

Theorem 8.1. *The ellipsoid* $\text{CE}(Q_{\alpha\beta})$ *has the form* $E(X, c)$ *where* $c = \tau e_1$ *and* $X = \text{diag}(a, b, \ldots, b)$, *where the parameters* $a > 0$, $b > 0$, *and* $\alpha < \tau < \beta$ *are given as follows:*
(i) If $\alpha\beta = -1/n$, *then*

$$\tau = 0, \quad \text{and} \quad a = b = 1. \tag{8.1}$$

(ii) If $\alpha + \beta = 0$ *and* $\alpha\beta \neq -1/n$, *then*

$$\tau = 0, \quad a = \frac{1}{n\beta^2}, \quad b = \frac{n-1}{n(1-\beta^2)}. \tag{8.2}$$

(iii) If $\alpha + \beta \neq 0$ *and* $\alpha\beta \neq -1/n$, *then*

$$\tau = \frac{n(\beta + \alpha)^2 + 2(1 + \alpha\beta) - \sqrt{\Delta}}{2(n+1)(\beta + \alpha)},$$

$$a = \frac{1}{n(\tau - \alpha)(\beta - \tau)}, \quad b = \frac{1 - a(\tau - \alpha)^2}{1 - \alpha^2}, \tag{8.3}$$

where $\Delta = n^2(\beta^2 - \alpha^2)^2 + 4(1 - \alpha^2)(1 - \beta^2)$.

Proof. It is clear that $\text{Aut}(Q_{\alpha\beta}) \supseteq \text{Aut}(B_{\alpha\beta})$, so that Lemma 5.3 implies $X = diag(a, b, ..., b)$ and $c = \tau e_1$ with $a > 0$, $b > 0$ and $\tau \in \mathbb{R}$. Let $\{u_i\}_{i=1}^k$ be the contact points of the optimal ellipsoid with $Q_{\alpha\beta}$. Writing $u_i = (y_i, z_i) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have that y_i is either α or β, and $1 = ||u_i||^2 = y_i^2 + ||z_i||^2$, that is, $||z_i||^2 = 1 - y_i^2$. Since $(u_i - c)(u_i - c)^T = \begin{bmatrix} (y_i - \tau)^2 & (y_i - \tau)z_i^T \\ (y_i - \tau)z_i & z_i z_i^T \end{bmatrix}$, Theorem 2.2 yields the following optimality conditions:

$$\tau = \frac{1}{n}\sum_{i=1}^{k}\lambda_i y_i, \quad 0 = \sum_{i=1}^{k}\lambda_i z_i, \quad \sum_{i=1}^{k}\lambda_i = n, \tag{8.4}$$

$$\frac{1}{a} = \sum_{i=1}^{k}\lambda_i(y_i - \tau)^2, \quad 0 = \sum_{i=1}^{k}\lambda_i(y_i - \tau)z_i, \quad \frac{1}{b}I_{n-1} = \sum_{i=1}^{k}\lambda_i z_i z_i^T, \tag{8.5}$$

$$0 = a(y_i - \tau)^2 + b(1 - y_i^2) - 1, \quad i = 1, \ldots, k, \tag{8.6}$$

$$0 \geq a(y - \tau)^2 + b(1 - y^2) - 1, \quad \text{for } y \in \{\alpha, \beta\}. \tag{8.7}$$

Here the last line expresses the feasibility condition $Q_{\alpha\beta} \subseteq E(X, c)$, since $Q_{\alpha\beta} \subseteq E(X, c)$ if and only if $\partial S_\alpha \cup \partial S_\beta \subseteq E(X, c)$.

The conditions (8.4)–(8.7) characterize the ellipsoid $\mathrm{CE}(Q_{\alpha\beta})$. Since the ellipsoid is unique, its parameters (a, b, τ) are unique and can be recovered from the above conditions.

The same arguments in Lemma 6.4 applies here and gives the equation (6.8),

$$1 = a(\alpha - \tau)^2 + b(1 - \alpha^2), \quad \tau = \left(1 - \frac{b}{a}\right)\cdot\frac{\alpha + \beta}{2}, \quad \frac{1}{a} = n(\tau - \alpha)(\beta - \tau), \tag{8.8}$$

which we again use to compute the variables a, b, and τ. If $a = b$ in the optimal ellipsoid, (8.8) immediately gives $\tau = 0$, $a = b = 1$, and $\alpha\beta = -1/n$.

Next, if $a \neq b$ and $\alpha = -\beta$, then (8.8) gives $\tau = 0$, $a = 1/(n\alpha^2)$ and $b = (n - 1)/(n(1 - \alpha^2))$. Furthermore, if we have $a = 1 = n\alpha^2$, then we obtain a contradiction since $b = (n - 1)/(n - 1) = 1 = a$. Thererefore, $a \neq 1$ and the last equation in (8.8) gives $\alpha\beta \neq -1/n$.

Lastly, $a \neq b$ and $\alpha \neq -\beta$, then the same argument in Lemma 6.4 gives the quadratic equation (6.9) for τ,

$$(n + 1)(\alpha + \beta)\tau^2 - \left(n(\alpha + \beta)^2 + 2(1 + \alpha\beta)\right)\tau + (\alpha + \beta)(1 + n\alpha\beta) = 0,$$

and the resulting equations in (8.3) for (a, b, τ). Since the middle equation in (8.8) implies $\tau \neq 0$, we have

$$0 \neq (n(\alpha + \beta)^2 + 2(1 + n\alpha\beta))^2 - \Delta = 4(n + 1)(\alpha + \beta)^2(1 + n\alpha\beta),$$

that is, $\alpha\beta \neq -1/n$. $\qquad\square$

References

[1] BALL, K. Ellipsoids of maximal volume in convex bodies. *Geom. Dedicata 41*, 2 (1992), 241–250.

[2] BALL, K. An elementary introduction to modern convex geometry. In *Flavors of geometry*, vol. 31 of *Math. Sci. Res. Inst. Publ.* Cambridge Univ. Press, Cambridge, 1997, pp. 1–58.

[3] BARNES, E. R., AND MORETTI, A. C. Some results on centers of polytopes. *Optim. Methods Softw. 20*, 1 (2005), 9–24.

[4] BARVINOK, A., AND BLEKHERMAN, G. Convex geometry of orbits. In *Combinatorial and computational geometry*, vol. 52 of *Math. Sci. Res. Inst. Publ.* Cambridge Univ. Press, Cambridge, 2005, pp. 51–77.

[5] Behrend, F. Über die kleinste umbeschriebene und die größte einbeschriebene Ellipse eines konvexen Bereichs. *Math. Ann. 115*, 1 (1938), 379–411.

[6] Blekherman, G. Convexity properties of the cone of nonnegative polynomials. *Discrete Comput. Geom. 32*, 3 (2004), 345–371.

[7] Brickman, L., and Steinberg, L. Classroom Notes: On Nonnegative Polynomials. *Amer. Math. Monthly 69*, 3 (1962), 218–221.

[8] Bump, D. *Lie groups*, vol. 225 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2004.

[9] Coxeter, H. S. M. *Regular polytopes*. Second edition. The Macmillan Co., New York, 1963.

[10] Danzer, L., Grünbaum, B., and Klee, V. Helly's theorem and its relatives. In *Proc. Sympos. Pure Math., Vol. VII*. Amer. Math. Soc., Providence, R.I., 1963, pp. 101–180.

[11] Danzer, L., Laugwitz, D., and Lenz, H. Über das Löwnersche Ellipsoid und sein Analogon unter den einem Eikörper einbeschriebenen Ellipsoiden. *Arch. Math. 8* (1957), 214–219.

[12] Davies, E. B. Symmetries of compact convex sets. *Quart. J. Math. Oxford Ser. (2) 25* (1974), 323–328.

[13] Gordon, Y., Litvak, A. E., Meyer, M., and Pajor, A. John's decomposition in the general case and applications. *J. Differential Geom. 68*, 1 (2004), 99–119.

[14] Gruber, P. M. Minimal ellipsoids and their duals. *Rend. Circ. Mat. Palermo (2) 37*, 1 (1988), 35–64.

[15] Grünbaum, B. Fixing systems and inner illumination. *Acta Math. Acad. Sci. Hungar 15* (1964), 161–163.

[16] Güler, O. *Foundations of optimization in finite dimensions*. Springer-Verlag, New York, forthcoming.

[17] Hačijan, L. G. A polynomial algorithm in linear programming. *Dokl. Akad. Nauk SSSR 244*, 5 (1979), 1093–1096.

[18] John, F. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*. Interscience Publishers, Inc., New York, N. Y., 1948, pp. 187–204.

[19] Juhnke, F. Embedded maximal ellipsoids and semi-infinite optimization. *Beiträge Algebra Geom. 35*, 2 (1994), 163–171.

[20] Juhnke, F. Polarity of embedded and circumscribed ellipsoids. *Beiträge Algebra Geom. 36*, 1 (1995), 17–24.

[21] Juhnke, F. Bounds of the affine breadth eccentricity of convex bodies via semi-infinite optimization. *Beiträge Algebra Geom. 45*, 2 (2004), 557–568.

[22] Klartag, B. On John-type ellipsoids. In *Geometric aspects of functional analysis*, vol. 1850 of *Lecture Notes in Math*. Springer, Berlin, 2004, pp. 149–158.

[23] KÖNIG, H., AND PALLASCHKE, D. On Khachian's algorithm and minimal ellipsoids. *Numer. Math. 36*, 2 (1980/81), 211–223.

[24] LAUGWITZ, D. *Differential and Riemannian geometry*. Translated by Fritz Steinhardt. Academic Press, New York, 1965.

[25] LEDOUX, M. *The concentration of measure phenomenon*, vol. 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[26] LEWIS, D. R. Ellipsoids defined by Banach ideal norms. *Mathematika 26*, 1 (1979), 18–29.

[27] MONTEIRO, R. D., O'NEAL, J. W., AND NEMIROVSKI, A. A new conjugate gradient algorithm incorporating adaptive ellipsoid preconditioning. *Working Paper* (2004).

[28] ONISHCHIK, A. L., AND VINBERG, È. B. *Lie groups and algebraic groups*. Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1990. Translated from the Russian and with a preface by D. A. Leites.

[29] PISIER, G. *The volume of convex bodies and Banach space geometry*, vol. 94 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[30] PÓLIK, I., AND TERLAKY, T. A survey of the S-lemma. *SIAM Review 49*, 3 (2007), 371–418.

[31] POLYAK, B. T. Convexity of quadratic transformations and its use in control and optimization. *J. Optim. Theory Appl. 99*, 3 (1998), 553–583.

[32] ROCKAFELLAR, R. T. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[33] RUDELSON, M. Contact points of convex bodies. *Israel J. Math. 101* (1997), 93–124.

[34] SUN, P., AND FREUND, R. M. Computation of minimum-volume covering ellipsoids. *Oper. Res. 52*, 5 (2004), 690–706.

[35] TARASOV, S. P., KHACHIYAN, L. G., AND ÈRLIKH, I. I. The method of inscribed ellipsoids. *Dokl. Akad. Nauk SSSR 298*, 5 (1988), 1081–1085.

[36] TODD, M. J. On minimum volume ellipsoids containing part of a given ellipsoid. *Math. Oper. Res. 7*, 2 (1982), 253–261.

[37] TOMCZAK-JAEGERMANN, N. *Banach-Mazur distances and finite-dimensional operator ideals*, vol. 38 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific & Technical, Harlow, 1989.

[38] YILDIRIM, E. A. On the minimum volume covering ellipsoid of ellipsoids. *SIAM J. Optim. 17*, 3 (2006), 621–641 (electronic).

[39] ZAGUSKIN, V. L. Circumscribed and inscribed ellipsoids of extremal volume. *Uspehi Mat. Nauk 13*, 6 (84) (1958), 89–93.

[40] ŽELOBENKO, D. P. *Compact Lie groups and their representations*. American Mathematical Society, Providence, R.I., 1973. Translated from the Russian by Israel Program for Scientific Translations, Translations of Mathematical Monographs, Vol. 40.

[41] Zhang, Y., and Gao, L. On numerical solution of the maximum volume ellipsoid problem. *SIAM J. Optim. 14*, 1 (2003), 53–76 (electronic).