Discrete Comput Geom 18:205–237 (1997)

Discrete & Computational
Geometry

Computing Mixed Discriminants, Mixed Volumes, and Permanents*

A. Barvinok

Department of Mathematics, University of Michigan,
Ann Arbor, MI 48109-1109, USA
barvinok@math.lsa.umich.edu

Abstract. We construct a probabilistic polynomial time algorithm that computes the mixed discriminant of given n positive definite $n \times n$ matrices within a $2^{O(n)}$ factor. As a corollary, we show that the permanent of an $n \times n$ nonnegative matrix and the mixed volume of n ellipsoids in $\mathbb{R}^n$ can be computed within a $2^{O(n)}$ factor by probabilistic polynomial time algorithms. Since every convex body can be approximated by an ellipsoid, the last algorithm can be used for approximating in polynomial time the mixed volume of n convex bodies in $\mathbb{R}^n$ within a factor $n^{O(n)}$.

1. Introduction

In this paper we address the problem of estimating the permanent of a given nonnegative matrix and the mixed volume of given n ellipsoids in $\mathbb{R}^n$. We show that these computational problems are related to that of estimating the mixed discriminant of n positive definite $n \times n$ matrices. We present a randomized polynomial time algorithm for the last problem and discuss its applications. Our main results are:

- A randomized polynomial time algorithm that computes the permanent of a given $n \times n$ nonnegative matrix within a $2^{O(n)}$ factor.
- A randomized polynomial time algorithm that computes the mixed volume of given n ellipsoids in $\mathbb{R}^n$ within a $2^{O(n)}$ factor.
- For any fixed k a deterministic polynomial time algorithm that computes the mixed volume of given n ellipsoids $E_1, \ldots, E_1, E_2, \ldots, E_2, \ldots, E_k, \ldots, E_k$ in $\mathbb{R}^n$, only k being pairwise different, within a $2^{O(n)}$ factor.

* This research was supported by the Alfred P. Sloan Research Fellowship, by NSF Grant DMS 9501129, and by the grant of Horace H. Rackham School of Graduate Studies and the Office of the Vice-President for Research at the University of Michigan.

(1.1) Permanent. Let S_n be the symmetric group of all $n!$ permutations of the set $\{1, \ldots, n\}$. Let $A = (a_{ij})$ be an $n \times n$ matrix. The number

$$\operatorname{per} A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}$$

is called the *permanent* of A. If A is a 0–1 matrix, then per A is the number of perfect matchings in the bipartite graph with the adjacency matrix A. We are interested in the problem of computing the permanent of a given nonnegative matrix. This problem is known to be #P-complete. Despite various results on computing the permanent of a "typical" 0–1 matrix [6], [19], the permanent of a "sparse" matrix [8], the permanent of a "dense" matrix [11], and the permanent of a matrix with the bounded rank [3], surprisingly little is known about how well can one approximate the permanent of any given nonnegative (and even 0–1 matrix) in polynomial time. It is easy to construct a polynomial time algorithm that for any given nonnegative matrix A computes a number α such that

$$\frac{p(n)}{n!} \operatorname{per} A \leq \alpha \leq \operatorname{per} A,$$

where $p(n)$ is a polynomial given in advance. Using an algorithm for the Assignment Problem (see, for example, [18]) we can find in polynomial time the first $p(n)$ permutations with largest weights $\prod_{i=1}^{n} a_{i\sigma(i)}$. Apart from this trivial estimate, nothing seems to be known.

In this paper we construct a randomized polynomial time algorithm that, for any given nonnegative matrix A, computes a number α such that

$$c^n \operatorname{per} A \leq \alpha \leq \operatorname{per} A,$$

where $c > 0$ is an absolute constant (we can choose $c = 0.28$). Although this is the best known polynomial time approximation for a "worst-case" nonnegative matrix, it is still far from a polynomial time approximation scheme known for an "average" 0–1 matrix (see [6], [11], and [19]). The author conjectures though that the proposed algorithm leads to a polynomial time approximation scheme for (properly defined) "average" nonnegative matrices. V. D. Milman suggested that for *any* $c < 1$ a polynomial time algorithm might exist that approximates the permanent of a given nonnegative matrix within a factor c^n.

(1.2) Mixed Volumes. Let $K_1, \ldots, K_n$ be convex bodies in the Euclidean space $\mathbb{R}^n$ and let $V(\cdot)$ be the Euclidean volume in $\mathbb{R}^n$. As is well known (see, for example, [21] and [22]) the value of $V(\lambda_1 K_1 + \cdots + \lambda_n K_n)$ is a homogeneous polynomial of degree n in nonnegative coefficients $\lambda_1, \ldots, \lambda_n$, where "+" denotes the Minkowski addition and λK denotes the dilatation of K with the coefficient λ. Thus we have

$$V(\lambda_1 K_1 + \cdots + \lambda_n K_n) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} \lambda_{i_1} \cdots \lambda_{i_n} V(K_{i_1}, \ldots, K_{i_n})$$

for nonnegative λ_i. Coefficients $V(K_{i_1}, \ldots, K_{i_n})$ are uniquely determined by the assumption that they are symmetric with respect to permutations of $K_{i_1}, \ldots, K_{i_n}$. The coefficient

$V(K_1, \ldots, K_n)$ in the above expansion is called the *mixed volume* of $K_1, \ldots, K_n$. The mixed volume is known to be nonnegative and monotone, that is, if $K_i \subset K_i'$ for all i then $V(K_1', \ldots, K_n') \geq V(K_1, \ldots, K_n)$, see [21] and [22].

The problem of computing the mixed volume of given convex bodies is important for Combinatorics, Algebraic Geometry, and Operations Research (see [9] and [4]). For example, the number of toric solutions to a generic system of n polynomial equations on $\mathbb{C}^n$ is equal to $n!$ times the mixed volume of the Newton polytopes of the equations.

An important particular case is computing the mixed volume of n ellipsoids $E_1, \ldots, E_n$ in $\mathbb{R}^n$. The problem of computing $V(K_1, \ldots, K_n)$ and $V(E_1, \ldots, E_n)$, in particular, was studied in [4]. There a polynomial time algorithm was constructed that approximates $V(E_1, \ldots, E_1, E_2, \ldots, E_2)$ within a factor c^n, where $c > 0$ is an absolute constant.

In this paper we construct a randomized polynomial time algorithm that for any given ellipsoids $E_1, \ldots, E_n \subset \mathbb{R}^n$ computes a number α such that

$$c^n V(E_1, \ldots, E_n) \leq \alpha \leq V(E_1, \ldots, E_n),$$

where $c > 0$ is an absolute constant (we can choose $c = 0.66$). Furthermore, for any fixed k we construct a deterministic polynomial time algorithm that achieves the same degree of approximation (with $c = 1/\sqrt{3} \approx 0.577$) for

$$V(E_1, \ldots, E_1, E_2, \ldots, E_2, \ldots, E_k, \ldots, E_k),$$

i.e., when we have only k pairwise different ellipsoids. In particular, this settles in part a conjecture of [4] that the mixed volume of ellipsoids can be approximated in polynomial time within a factor depending on the dimension alone. "In part" refers to the fact that in the general case, we have only a randomized polynomial time algorithm, whereas a deterministic algorithm is desirable. V. D. Milman conjectured that for *any* $c < 1$ there exists a randomized polynomial time algorithm that computes the mixed volume of given n ellipsoids with a factor c^n.

For each convex body $K_i \subset \mathbb{R}^n$ there is an ellipsoid E_i so that $E_i \subset K \subset nE_i$ (after a suitable translation), see, for example, [10]. Since mixed volumes are monotone, our algorithms can be used for approximating the mixed volume $V(K_1, \ldots, K_n)$ within a factor $n^{O(n)}$ provided K_i belong to a class of convex bodies that can be approximated by ellipsoids within a factor $n^{O(1)}$ in polynomial time. This is the first polynomial time algorithm that approximates $V(K_1, \ldots, K_n)$ within a factor depending on n alone for a reasonably broad class of convex bodies.

Our computational model is the RAM with the uniform cost criterion [1]. For convenience, together with the arithmetic operations (addition, subtraction, multiplication, division, and comparison of real numbers) we allow taking the square root of a nonnegative real number. All these operations are assumed to have cost 1. We also include a standard subroutine from Linear Algebra, that is computing the eigenvalues of a real symmetric matrix. In the probabilistic setting, we assume that our machine can sample a point from the uniform distribution on the unit sphere. This assumption is not very restrictive since it is known that the standard normal distribution in $\mathbb{R}^n$ (and thus the uniform distribution on the sphere) can be simulated with an arbitrary precision in

polynomial time from the standard Bernoulli distribution by means of the Central Limit Theorem.

To compute permanents and mixed volumes we use mixed discriminants introduced by Aleksandrov in his proof of the Aleksandrov–Fenchel inequality (see [2]). They turned out to be useful in proving the van der Waerden conjecture for permanents of doubly stochastic matrices (see [5]).

(1.3) Mixed Discriminants. Let $Q_1, \ldots, Q_n$ be symmetric $n \times n$ matrices and let $t_1, \ldots, t_n$ be real variables. Then there is an expansion similar to that of (1.2):

$$\det(t_1 Q_1 + \cdots + t_n Q_n) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} t_{i_1} \cdots t_{i_n} D(Q_{i_1}, \ldots, Q_{i_n}), \tag{1.3.1}$$

where the coefficients $D(Q_{i_1}, \ldots, Q_{i_n})$ are assumed to be symmetric with respect to permutations of $Q_{i_1}, \ldots, Q_{i_n}$. The coefficient $D(Q_1, \ldots, Q_n)$ is called the *mixed discriminant* of $Q_1, \ldots, Q_n$.

Mixed discriminants have many interesting properties somewhat parallel to those of mixed volumes (see, for example, Section 3 of [15]) and they seem to be easier to deal with. Mixed discriminants can be considered as a generalization of permanents and they also have some interesting combinatorial applications. For example, the number of bases in the intersection of a unimodular matroid with a transversal matroid can be expressed as the mixed discriminant of some positive semidefinite matrices. The author believes that the problem of computing the mixed discriminant is interesting in its own right.

If we fix an orthonormal basis in $\mathbb{R}^n$ we may identify a symmetric matrix Q with a self-adjoint operator on $\mathbb{R}^n$ and consider the ellipsoid $E_Q = \{x \in \mathbb{R}^n : \langle x, Qx \rangle \leq 1\}$, where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^n$. Relations between permanents, mixed discriminants, and mixed volumes are described by the following theorem.

(1.4) Theorem.

(1.4.1) *Let $A = (a_{ij})$ be an $n \times n$ matrix. Let $M_i = \operatorname{diag}\{a_{i1}, \ldots, a_{in}\}$ be the diagonal matrix whose jth diagonal element is a_{ij}. Then*

$$\operatorname{per} A = n!\, D(M_1, \ldots, M_n).$$

(1.4.2) *Let $Q_1, \ldots, Q_n$ be positive definite $n \times n$ matrices and let*

$$E_{Q_i} = \{x \in \mathbb{R}^n : \langle x, Q_i x \rangle \leq 1\}, \qquad i = 1, \ldots, n,$$

be the corresponding ellipsoids. Then

$$\begin{aligned} (\sqrt{3})^{-n+1} v_n D^{1/2}(Q_1^{-1}, \ldots, Q_n^{-1}) &\leq V(E_{Q_1}, \ldots, E_{Q_n}) \\ &\leq v_n D^{1/2}(Q_1^{-1}, \ldots, Q_n^{-1}), \end{aligned}$$

where $v_n = \pi^{n/2} / \Gamma(n/2 + 1)$ is the volume of the unit ball in $\mathbb{R}^n$.

The central result of this paper is a randomized polynomial time algorithm that for any given positive definite $n \times n$ matrices $M_1, \ldots, M_n$ with probability at least 0.9 computes

a number α such that

$$c^n D(M_1, \ldots, M_n) \leq \alpha \leq 20 D(M_1, \ldots, M_n)$$

for some absolute constant $c > 0$ (we can choose $c = 0.28$). To get an overwhelming probability, we have to run the algorithm several times and choose the median of the computed α's.

This paper is organized as follows. In Section 2 we prove a recurrence for the mixed discriminant that allows us to reduce its computation to the computation of the average value of a positive definite quadratic form on the unit sphere S^{n-1}. In Section 3 we study the distribution of values of a quadratic form on S^{n-1}. In Section 5 we present our algorithm for computing the mixed discriminant $D(M_1, \ldots, M_n)$ and prove that it has the desired complexity (almost obvious) and achieves the desired degree of approximation (far less obvious). The main idea of the algorithm is to construct a random variable on the orthogonal group O_n whose expectation is the mixed discriminant of given matrices. To estimate the expectation we use a Monte Carlo algorithm with just one sampling. We use the results of Section 4 on the integration over the orthogonal group to prove that our algorithm indeed achieves the desired degree of approximation. In Section 6 we apply our algorithm to the permanent computation. In Section 7 we prove part (1.4.2) of Theorem 1.4. Together with the algorithm from Section 5 this gives us an algorithm for estimating the mixed volume of ellipsoids. In Section 9 we present an independent algorithm for the last problem that gives us an unbiased estimator, achieves, in principle, a better approximation, and is more geometric. We use a known recurrence for the mixed volume that allows us to reduce its computation to the computation of the average value of the support function of a zonoid in $\mathbb{R}^n$. We use Theorem 1.4 to construct a deterministic polynomial time algorithm when the number of different ellipsoids is fixed. In Section 8 we study the distribution of values of the support function of a zonoid which is necessary for our analysis of the algorithm.

(1.5) Notation. We summarize some notation used throughout this paper. Thus $\langle \cdot, \cdot \rangle$ is the standard scalar product in $\mathbb{R}^n$. We denote by Q^* the operator adjoint to Q, that is, $\langle x, Qy \rangle = \langle Q^* x, y \rangle$ for all $x, y \in \mathbb{R}^n$.

For a convex body $K \subset \mathbb{R}^n$ and a linear subspace $L \subset \mathbb{R}^n$ we denote by $K|L$ the orthogonal projection of K onto L. If $Q\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a self-adjoint operator on $\mathbb{R}^n$ and $L \subset \mathbb{R}^n$ is a linear subspace we define its *projection* $Q|L$ as follows: Let $P\colon L \longrightarrow \mathbb{R}^n$ be the inclusion and let $P^*\colon \mathbb{R}^n \longrightarrow L$ be the orthogonal projection onto L. Then $Q|L = P^* Q P$ is a self-adjoint operator on L. In other words, if $q(x) = \langle x, Qx \rangle$ is the quadratic form associated with Q, then for the restriction of $q(x)$ onto L we have $q(x) = \langle x, (Q|L)x \rangle$ for each $x \in L$. We note that $(\alpha Q_1 + \beta Q_2)|L = \alpha(Q_1|L) + \beta(Q_2|L)$. A self-adjoint operator Q is called *positive definite* if $\langle x, Qx \rangle > 0$ for any $x \neq 0$. It is immediate that $Q|L$ is positive definite provided Q is positive definite. We denote by I the identity operator on $\mathbb{R}^n$.

For a convex compact set $K \subset \mathbb{R}^n$ we denote $h_K(u) = \max\{\langle u, x \rangle : x \in K\}$, $h_K\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ the support function of K. A *zonotope* is the Minkowski sum of finitely many segments (symmetric about the origin) in $\mathbb{R}^n$ and a *zonoid* is a limit of zonotopes in the Hausdorff metric (see, for example, [16] and [21]).

Let $S^{n-1} = \{x \in \mathbb{R}^n : \langle x, x\rangle = 1\}$ be the unit sphere in $\mathbb{R}^n$ and let $B_n = \{x \in \mathbb{R}^n : \langle x, x\rangle \leq 1\}$ be the unit ball. We denote by

$$v_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)} = \frac{1}{\sqrt{\pi n}}\left(\frac{2\pi e}{n}\right)^{n/2}(1 + O(n^{-1})) \qquad \text{(Stirling's formula)}$$

the volume of B_n, and by

$$\kappa_{n-1} = nv_n = \frac{n\pi^{n/2}}{\Gamma(n/2+1)} = \sqrt{\frac{n}{\pi}}\left(\frac{2\pi e}{n}\right)^{n/2}(1 + O(n^{-1}))$$

the surface area of S^{n-1}. Let $\mu_{n-1} = du$ be the rotation-invariant Borel probability measure on S^{n-1}. Sometimes we write μ instead of μ_{n-1}.

Let us consider the Stiefel manifold $O_{n,s}$ as the space of all s-tuples $(u_1, \ldots, u_s)$ of pairwise orthogonal unit vectors in $\mathbb{R}^n$. In particular, $O_{n,1} = S^{n-1}$ is the unit sphere and $O_{n,n}$ is the space of all orthonormal bases $(u_1, \ldots, u_n)$ in $\mathbb{R}^n$. By choosing the standard orthonormal basis $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, $\ldots$, $e_n = (0, \ldots, 0, 1)$ in $\mathbb{R}^n$ we identify $O_{n,n}$ with the orthogonal group O_n in $\mathbb{R}^n$. Let ν be the Haar probability measure on O_n. For a set of pairwise orthogonal unit vectors $(u_1, \ldots, u_s)$ we denote by $(u_1, \ldots, u_s)^{\perp}$ the $(n-s)$-dimensional linear subspace $L \subset \mathbb{R}^n$ that is orthogonal to $u_1, \ldots, u_s$.

We denote by $|X|$ the cardinality of a finite set X.

2. A Recurrence for Mixed Discriminants

We begin with a simple lemma.

(2.1) Lemma. *Let $p(\mathbf{t})$ be a homogeneous polynomial of degree n in n real variables $\mathbf{t} = (t_1, \ldots, t_n)$. For a subset $\omega \subset \{1, \ldots, n\}$ let*

$$t_i(\omega) = \begin{cases} 1 & \text{if} \quad i \in \omega, \\ 0 & \text{if} \quad i \notin \omega, \end{cases}$$

and let $\mathbf{t}_\omega = (t_1(\omega), \ldots, t_n(\omega))$. Then

$$\frac{\partial^n}{\partial t_1 \cdots \partial t_n} p(\mathbf{t}) = (-1)^n \sum_{\omega \subset \{1,\ldots,n\}} (-1)^{|\omega|} p(\mathbf{t}_\omega),$$

where the sum is taken over all nonempty subsets ω of $\{1, \ldots, n\}$.

Proof. Both sides of the equation are linear in p. If $p(\mathbf{t}) = t_1 \cdots t_n$ the identity holds since $p(\mathbf{t}_\omega) = 0$ unless $\omega = \{1, \ldots, n\}$. If p is a monomial whose support does not contain an $i \in \{1, \ldots, n\}$ the identity holds since the summands corresponding to $\omega \backslash \{i\}$ and $\omega \cup \{i\}$ annihilate each other. □

(2.2) Corollary. *Suppose that rank $Q_i \leq 1$ for $i = 1, \ldots, n$. Then*

$$D(Q_1, \ldots, Q_n) = \frac{1}{n!} \det(Q_1 + \cdots + Q_n).$$

Proof. From (1.3.1) we get the following representation for the mixed discriminant:

$$D(Q_1, \ldots, Q_n) = \frac{1}{n!} \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \det(t_1 Q_1 + \cdots + t_n Q_n).$$

Since $\det(t_1 Q_1 + \cdots + t_n Q_n)$ is a homogeneous polynomial of degree n in $t_1, \ldots, t_n$ from Lemma 2.1 we then get

$$D(Q_1, \ldots, Q_n) = \frac{(-1)^n}{n!} \sum_{\omega \subset \{1,\ldots,n\}} (-1)^{|\omega|} \det\left(\sum_{i \in \omega} Q_i\right). \tag{2.2.1}$$

Since

$$\operatorname{rank}\left(\sum_{i \in \omega} Q_i\right) \leq |\omega|$$

we get

$$\det\left(\sum_{i \in \omega} Q_i\right) = 0 \qquad \text{unless} \quad \omega = \{1, \ldots, n\}.$$

The proof follows by (2.2.1). □

Mixed discriminants are invariant with respect to permutations of arguments and linear in every argument (see, for example, formula (54), Section 3 of [15]):

$$\begin{aligned} &D(Q_1, \ldots, \alpha Q_i' + \beta Q_i'', \ldots, Q_n) \\ &\qquad = \alpha D(Q_1, \ldots, Q_i', \ldots, Q_n) + \beta D(Q_1, \ldots, Q_i'', \ldots, Q_n). \end{aligned}$$

It is known that $D(Q_1, \ldots, Q_n) > 0$ provided every Q_i is positive definite (see, for example, Proposition 3.2 of [15]).

We recall from Section 1.5 that $u^\perp$ is the hyperplane L in $\mathbb{R}^n$ orthogonal to a unit vector $u \in S^{n-1}$ and that $Q|u^\perp$ is the projection of a self-adjoint operator Q onto L. If we fix an orientation of $\mathbb{R}^n$ we can define $\det Q$. The choice of u as a unit normal to L defines the orientation of L compatible with that of $\mathbb{R}^n$ and hence we may define $\det(Q|u^\perp)$.

We need the following technical result.

(2.3) Lemma. *Let Q be a self-adjoint operator on $\mathbb{R}^n$.*

(2.3.1) *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of Q. Then*

$$\int_{S^{n-1}} \det(Q|u^\perp)\, du = \frac{1}{n} e_{n-1}(\lambda_1, \ldots, \lambda_n),$$

where e_{n-1} is the elementary symmetric polynomial of degree $n-1$ in n variables.

(2.3.2) *Suppose that $\operatorname{rank} Q = n - 1$. Let us choose a vector $v \in S^{n-1}$ such that $Qv = 0$ (vector v is unique up to a sign). Then*

$$\det(Q|u^\perp) = \langle u, v \rangle^2 \det(Q|v^\perp) \qquad \textit{for each} \quad u \in S^{n-1}.$$

Proof. Let us denote

$$p(Q) = \int_{S^{n-1}} \det(Q|u^\perp)\, du.$$

Suppose that A is an orthogonal operator on $\mathbb{R}^n$ and $Q_1 = AQA^*$. For $u \in S^{n-1}$ let $v = Au$. Then A maps $u^\perp$ onto $v^\perp$ and $(A^*(Q_1|v^\perp)A)x = (Q|u^\perp)x$ for any $x \in u^\perp$. Since A is orthogonal we have $\det(Q_1|v^\perp) = \det(Q|u^\perp)$ and since μ is rotation invariant we have $p(Q) = p(Q_1) = p(AQA^*)$. Hence $p(Q)$ is a symmetric function in the eigenvalues of Q. Suppose that $f_1, \ldots, f_n$ are the unit eigenvectors of Q and $Q_i = \langle f_i, \cdot \rangle f_i$ is the orthogonal projector onto f_i. So $Q = \lambda_1 Q_1 + \cdots + \lambda_n Q_n$ and $Q|u^\perp = \lambda_1(Q_1|u^\perp) + \cdots + \lambda_n(Q_n|u^\perp)$. Hence $\det(Q|u^\perp)$ is a homogeneous polynomial of degree $n-1$ in $\lambda_1, \ldots, \lambda_n$, and, therefore, $p(Q)$ is a symmetric homogeneous polynomial of degree $n-1$ in $\lambda_1, \ldots, \lambda_n$. Next, we note that if at least two of $\lambda_1, \ldots, \lambda_n$ are zeros then rank $Q \leq n-2$, therefore rank $(Q|u^\perp) \leq n-2$ and hence $\det(Q|u^\perp)$ is identically zero. So $p(Q) = 0$ provided Q has at least two zero eigenvalues. This implies that $p(Q) = c(n)e_{n-1}(\lambda_1, \ldots, \lambda_n)$. To find the constant $c(n)$ we let Q to be the identity operator. Then $Q|u^\perp$ is the identity operator, so $p(Q) = 1$ and $c(n) = 1/n$. So (2.3.1) is proven.

Let $f_1, \ldots, f_{n-1}$ be the unit eigenvectors corresponding to the nonzero eigenvalues of Q. Thus $f_1, \ldots, f_{n-1}, v$ is an orthonormal basis of $\mathbb{R}^n$ in which Q is represented by a diagonal matrix. Let $H = v^\perp$ be the hyperplane generated by $f_1, \ldots, f_{n-1}$. Then for any $x \in \mathbb{R}^n$ we have $Qx = (QP_H)x$, where P_H is the orthogonal projection of $\mathbb{R}^n$ onto H. Let us choose a $u \in S^{n-1}$ and let $L = u^\perp$. Then for $x \in L$ we have $(Q|u^\perp)x = (P_L Q)x$, where P_L is the orthogonal projection of $\mathbb{R}^n$ onto L. So we may write $(Q|u^\perp)x = (P_L Q P_H)x$ for any $x \in L$. Let $P_{H,L}\colon H \longrightarrow L$ be the orthogonal projection of H onto L. Then $P^*_{H,L}$ is the orthogonal projection of L onto H and we get $Q|u^\perp = P_{H,L}(Q|v^\perp)P^*_{H,L}$. Since H and L are oriented hyperplanes, we may define $\det P_{H,L}$ and write

$$\det(Q|u^\perp) = (\det P_{H,L})^2 \det(Q|v^\perp).$$

Now we observe that $\det^2 P_{H,L} = \langle u, v \rangle^2$. To see this, let us choose an orthonormal basis $u_1, \ldots, u_{n-2}$ in $L \cap H$ and append it by a vector $l \in L$ to a positively oriented orthonormal basis of L and by a vector $h \in H$ to a positively oriented orthonormal basis of H. Then the projection $P_{H,L}$ can be written as $u_i \longmapsto u_i$, $h \longmapsto \langle h, l \rangle l$. Hence $\det P_{H,L} = \langle h, l \rangle = \langle u, v \rangle$ and the proof of (2.3.2) follows. □

In this section we prove the following main result.

(2.4) Theorem. *Let $Q_1, \ldots Q_n$ be positive definite operators on $\mathbb{R}^n$.*

(2.4.1) *Suppose that $Q_1 = TT^*$ for some nondegenerate T. Let $R_k = T^{-1}Q_k(T^{-1})^*$ for $k = 2, \ldots, n$. Then*

$$D(Q_1, \ldots, Q_n) = (\det Q_1) D(I, R_2, \ldots, R_n).$$

(2.4.2) $D(I, R_2, \ldots, R_n) = \int_{S^{n-1}} D(R_2|u^\perp, \ldots, R_n|u^\perp)\, du.$

(2.4.3) *There exists a positive definite quadratic form from* $\mathbf{q}$: $\mathbb{R}^n \longrightarrow \mathbb{R}$, *called the mixed quadratic form of* $R_2, \ldots, R_n$ *such that*

$$D(R_2|u^\perp, \ldots, R_n|u^\perp) = \mathbf{q}(u)$$

for any $u \in S^{n-1}$.

Proof. Formula (2.4.1) follows from (1.3.1) since

$$\begin{aligned} \det(t_1 Q_1 + \cdots + t_n Q_n) &= \det(T(t_1 I + t_2 R_2 + \cdots + t_n R_n)T^*) \\ &= (\det Q_1)\det(t_1 I + t_2 R_2 + \cdots + t_n R_n) \end{aligned}$$

for all $t_1, \ldots, t_n$.

To prove (2.4.2) let $Q = t_2 R_2 + \cdots + t_n R_n$ for some fixed coefficients $t_2, \ldots, t_n$ and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of Q. Applying (2.3.1) we get

$$\frac{d}{dt_1}\det(t_1 I + Q) = e_{n-1}(\lambda_1, \ldots, \lambda_n) = n\int_{S^{n-1}} \det(Q|u^\perp)\, du.$$

Now

$$\begin{aligned} D(I, R_2, \ldots, R_n) &= \frac{1}{n!}\frac{\partial^n}{\partial t_1 \cdots \partial t_n}\det(t_1 I + t_2 R_2 + \cdots + t_n R_n) \\ &= \frac{1}{n!}\frac{\partial^{n-1}}{\partial t_2 \cdots \partial t_n}\frac{\partial}{\partial t_1}\det(t_1 I + t_2 R_2 + \cdots + t_n R_n) \\ &= \frac{1}{(n-1)!}\frac{\partial^{n-1}}{\partial t_2 \cdots \partial t_n}\int_{S^{n-1}} \det(t_2 R_2|u^\perp + \cdots + t_n R_n|u^\perp)\, du \\ &= \int_{S^{n-1}} \frac{1}{(n-1)!}\frac{\partial^{n-1}}{\partial t_2 \cdots \partial t_n}\det(t_2 R_2|u^\perp + \cdots + t_n R_n|u^\perp)\, du \\ &= \int_{S^{n-1}} D(R_2|u^\perp, \ldots, R_n|u^\perp)\, du, \end{aligned}$$

so (2.4.2) follows. We can differentiate the integral since the integrand is a polynomial in $t_2, \ldots, t_n$.

Instead of (2.4.3) we will prove a somewhat more general fact, namely, that for any self-adjoint operators $R_2, \ldots, R_n$ there exists a quadratic form $\mathbf{q}$: $\mathbb{R}^n \longrightarrow \mathbb{R}$ such that $D(R_2|u^\perp, \ldots, R_n|u^\perp) = \mathbf{q}(u)$ for each $u \in S^{n-1}$. Since the mixed discriminant of positive definite operators is positive we would have $\mathbf{q}(u) > 0$ for each $u \in S^{n-1}$ provided $R_2, \ldots, R_n$ are positive definite and (2.4.3) would follow.

Every self-adjoint operator R_i can be represented as a sum $R_i = \sum_{j=1}^n Q_{ij}$ of self-adjoint operators Q_{ij} such that rank $Q_{ij} \leq 1$. Since mixed discriminants are linear in every argument we get

$$D(R_2|u^\perp, \ldots, R_n|u^\perp) = \sum_{j_2=1}^n \cdots \sum_{j_n=1}^n D(Q_{2j_2}|u^\perp, \ldots, Q_{nj_n}|u^\perp).$$

Therefore it suffices to prove that for any self-adjoint operators $Q_2, \ldots, Q_n$ such that rank $Q_i \leq 1$ for $i = 2, \ldots, n$ there exists a quadratic form $\mathbf{q}$: $\mathbb{R}^n \longrightarrow \mathbb{R}$ such that $D(Q_2|u^\perp, \ldots, Q_n|u^\perp) = \mathbf{q}(u)$ for any $u \in S^{n-1}$.

Let $Q = Q_2 + \cdots + Q_n$, so rank $Q \le n-1$. Then $\operatorname{rank}(Q_i|u^\perp) \le 1$ for every $u \in S^{n-1}$ and by Corollary 2.2 we have

$$D(Q_2|u^\perp, \ldots, Q_n|u^\perp) = \frac{1}{(n-1)!} \det(Q|u^\perp).$$

If rank $Q < n-1$ then $\operatorname{rank}(Q|u^\perp) < n-1$ and hence we may choose $\mathbf{q}$ to be identically zero. If rank $Q = n-1$, then for some $v \in S^{n-1}$ by (2.3.2) we may choose $\mathbf{q}(u) = (1/(n-1)!)\langle u, v\rangle^2 \det(Q|v^\perp)$, which is a quadratic form in u. □

As we noted, there is a certain similarity between properties of mixed discriminants and mixed volumes. We present the analogue of Theorem 2.4 for quermassintegrals in Section 7, Theorem 7.3. The analogue of the mixed quadratic form $\mathbf{q}$ is the "mixed brightness," that is the support function of the mixed projection body (see also [16]).

Our algorithm for computing the mixed discriminant is suggested by Theorem 2.4. Given n positive definite operators $Q_1, \ldots, Q_n$, by (2.4.1) we reduce computation of $D(Q_1, \ldots, Q_n)$ to that of $D(I, R_2, \ldots, R_n)$. Then we choose a vector $u \in S^{n-1}$ at random and replace the computation of $D(I, R_2, \ldots, R_n)$ by that of $D(Q_1', \ldots, Q_{n-1}')$, where $Q_i' = R_{i+1}|u^\perp$. Then we repeat the procedure. It easy to see that this procedure has polynomial time complexity. To estimate what kind of approximation we get, we discuss the following two issues: What error do we get on every step while passing from $D(Q_1, \ldots, Q_n)$ to $D(Q_1', \ldots, Q_{n-1}')$ and how do these errors accumulate? Because of (2.4.3) the first question reduces to the following: How well do we approximate the average value of a positive definite quadratic form on the unit sphere by the value of that form at a random point on the sphere? We address to this question in Section 3. The second question has to do with the "law of large numbers," specifically for the martingales on the orthogonal group. We discuss it in Section 4.

3. Distribution of Values of a Quadratic Form on the Sphere

Let $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a continuous function. We denote by

$$\mathbf{E}(f) = \int_{S^{n-1}} f(u)\, du$$

the average value of f on S^{n-1} (recall from Section 1.5 that $\mu = du$ is the rotation invariant probability measure on S^{n-1}). In our inductive constructions we are going to use the following argument: let us choose a coordinate system $x_1, \ldots, x_{n+1}$ in $\mathbb{R}^{n+1}$ and let us "slice" S^n onto $(n-1)$-dimensional spheres $S_\varphi^{n-1} = \{x \in S^n : x_{n+1} = \sin\varphi\}$ of radii $\cos\varphi$. Let $f\colon S^n \longrightarrow \mathbb{R}$ be a continuous function which is a constant $f(\varphi)$ on every slice S_φ^{n-1}. Then (see Section 1.5)

$$\mathbf{E}(f) = \frac{\kappa_{n-1}}{\kappa_n} \int_{-\pi/2}^{\pi/2} f(\varphi) \cos^{n-1}\varphi \, d\varphi.$$

In particular, $\int_{-\pi/2}^{\pi/2} \cos^{n-1}\varphi \, d\varphi = \kappa_n/\kappa_{n-1}$.

(3.1) Lemma. *Let $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a quadratic form with the eigenvalues $\lambda_1, \ldots, \lambda_n$. Then*

$$\mathbf{E}(q) = \frac{\lambda_1 + \cdots + \lambda_n}{n}$$

and

$$\mathbf{E}(q^2) = \frac{3}{n(2+n)}(\lambda_1^2 + \cdots + \lambda_n^2) + \frac{2}{n(2+n)} \sum_{1 \le i < j \le n} \lambda_i \lambda_j.$$

Proof. Let us consider $\mathbf{E}(q)$ and $\mathbf{E}(q^2)$ as functions of q. We note that if A is an orthogonal transformation of $\mathbb{R}^n$ and $q_1(x) = q(Ax)$, then $\mathbf{E}(q_1) = \mathbf{E}(q)$ and $\mathbf{E}(q_1^2) = \mathbf{E}(q^2)$. Therefore $\mathbf{E}(q)$ and $\mathbf{E}(q^2)$ are symmetric functions in the eigenvalues of the form q. Obviously, $\mathbf{E}(q)$ is a linear function of q, so we have $\mathbf{E}(q) = c(n)(\lambda_1 + \cdots + \lambda_n)$. Substituting $q = \langle x, x \rangle$ we get $\mathbf{E}(q) = 1$, so $c(n) = 1/n$.

Furthermore, $\mathbf{E}(q^2)$ is a quadratic polynomial in q, since $\mathbf{E}(q_1 \cdot q_2)$ is a bilinear form in q_1 and q_2. Therefore,

$$\mathbf{E}(q^2) = a(n) \sum_{i=1}^{n} \lambda_i^2 + b(n) \sum_{1 \le i < j \le n} \lambda_1 \lambda_2$$

for some $a(n)$ and $b(n)$. Substituting $q(x) = \langle x, x \rangle$ we get

$$n \cdot a(n) + \binom{n}{2} b(n) = 1.$$

To get another relation between $a(n)$ and $b(n)$ let us substitute $q(x) = x_n^2 = \sin^2 \varphi$. The computations show

$$\begin{aligned} \mathbf{E}(q^2) = \mathbf{E}(x_n^4) &= \frac{\kappa_{n-2}}{\kappa_{n-1}} \int_{-\pi/2}^{\pi/2} \sin^4 \varphi \cos^{n-2} \varphi \, d\varphi \\ &= \frac{\kappa_{n-2}}{\kappa_{n-1}} \int_{-\pi/2}^{\pi/2} (1 - \cos^2 \varphi)^2 \cos^{n-2} \varphi \, d\varphi \\ &= 1 - 2\frac{\kappa_{n-2}\kappa_{n+1}}{\kappa_{n-1}\kappa_n} + \frac{\kappa_{n-2}\kappa_{n+3}}{\kappa_{n-1}\kappa_{n+2}} = \frac{3}{n(2+n)}. \end{aligned}$$

Therefore

$$a(n) = \frac{3}{n(n+2)} \qquad \text{and} \qquad b(n) = \frac{2}{n(n+2)}.$$ □

(3.2) Corollary. *Let q be a positive semidefinite quadratic form. Then*

$$\mathbf{E}(q^2) \le 3(\mathbf{E}(q))^2.$$

Proof. Follows by Lemma 3.1. □

One can observe that the ratio $\mathbf{E}(q^2)/\mathbf{E}^2(q)$ is the greatest when rank $q = 1$. For an "average" quadratic form one can expect the ratio to be much closer to 1.

(3.3) Theorem. *Let $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a positive semidefinite form which is not identically zero. Then for any $t \geq 0$*

$$\mu\{x \in S^{n-1} : q(x) \leq t\mathbf{E}(q)\} \leq C_0\sqrt{t},$$

where C_0 is an absolute constant (independent of q and n).

Proof. The statement is obvious for $n = 1$ and any $C_0 \geq 1$. Therefore without loss of generality we assume that $n \geq 1$. Let us choose a constant $\alpha > 0$ (to be specified later) and let

$$C_n = \frac{\kappa_{n-1}}{\kappa_n\sqrt{(n+1)}}\alpha \qquad \text{for} \quad n \geq 1.$$

It is easy to see (see Section 1.5) that

$$\lim_{n\longrightarrow+\infty} C_n = \frac{\alpha}{\sqrt{2\pi}},$$

so we can choose α so that $C_n \geq 1$ for any $n \geq 1$. Finally, let $C_0 = \sup\{C_n : n \geq 1\} < \infty$.

We are going to prove by induction on n that

$$\mu_n\{x \in S^n : q(x) \leq \mathbf{E}(q)t\} \leq C_n\sqrt{t} \tag{3.3.1}$$

for any $n \geq 1$. This will obviously prove our theorem. □

Let $n = 1$. Let $M(q)$ be the largest eigenvalue of a quadratic form $q\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ and let $u \in S^1$ be the corresponding eigenvector. We note that $\mathbf{E}(q) \leq M(q)$ and that $q(x) \geq M(q)\langle u, x\rangle^2$. Therefore

$$\begin{aligned}\mu_1\{x \in S^1 : q(x) \leq \mathbf{E}(q)t\} &\leq \mu_1\{x \in S^1 : q(x) \leq M(q)t\}\\ &\leq \mu_1\{x \in S^1 : \langle u, x\rangle^2 \leq t\}\\ &= \frac{4}{2\pi}\arcsin\sqrt{t} \leq \sqrt{t} \leq C_1\sqrt{t}.\end{aligned}$$

Now we perform the induction step. Since $C_n \geq 1$ it suffices to check the case $t < 1$ only. Let $q\colon \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ be a positive semidefinite quadratic form, not identically zero and let $m(q) = \min\{q(x) : x \in S^n\}$ be the smallest eigenvalue of q. Consider $q_0 = q - m(q)\langle x, x\rangle$. If q_0 is identically zero then q is a nonzero constant and the result is obvious. Otherwise, we observe that $\mathbf{E}(q_0) = \mathbf{E}(q) - m(q)$ and since $t < 1$ we have $t\mathbf{E}(q) - m(q) \leq t(\mathbf{E}(q) - m(q))$. Therefore

$$\begin{aligned}\mu_n\{x \in S^n : q(x) \leq t\mathbf{E}(q)\} &= \mu_n\{x \in S^n : q_0(x) \leq t\mathbf{E}(q) - m(q)\}\\ &\leq \mu_n\{x \in S^n : q_0(x) \leq t(\mathbf{E}(q) - m(q))\}\\ &= \mu_n\{x \in S^n : q_0(x) \leq t\mathbf{E}(q_0)\}.\end{aligned}$$

Therefore it suffices to check our bound (3.3.1) for the forms q that are not identically zero, but have at least one zero eigenvalue.

Let $u \in S^n$ be an eigenvector of q corresponding to the zero eigenvalue. We identify $\mathbb{R}^n = u^\perp$. Let q_0 be the restriction of q onto $\mathbb{R}^n$. Since q_0 as a quadratic form on $\mathbb{R}^n$ has the same nonzero eigenvalues as the form q, by Lemma 3.1 we get

$$\mathbf{E}(q_0) = \frac{n+1}{n}\mathbf{E}(q).$$

Let

$$S_\varphi^{n-1} = \{x \in S^n : \langle x, u\rangle = \sin\varphi\}, \qquad -\frac{\pi}{2} < \varphi < \frac{\pi}{2}.$$

Thus S_φ^{n-1} is an $(n-1)$-dimensional sphere of radius $\cos\varphi$. We identify $S_0^{n-1} = S^{n-1} \subset \mathbb{R}^n$. For a point $x \in S_\varphi^{n-1}$ let x_0 be its orthogonal projection onto $\mathbb{R}^n$ and let $x' = (1/\cos\varphi)x_0 \in S^{n-1}$. We have $q(x) = (\cos^2\varphi)q_0(x')$.

Let us consider the rotation invariant Borel probability measure $\mu_{n-1,\varphi}$ on S_φ^{n-1} (we let $\mu_{n-1,0} = \mu_{n-1}$). Then

$$\begin{aligned}
\mu_{n-1,\varphi}\{x \in S_\varphi^{n-1} : q(x) \le t\mathbf{E}(q)\} &= \mu_{n-1}\left\{x' \in S^{n-1} : q_0(x') \le \frac{t}{\cos^2\varphi}\mathbf{E}(q)\right\} \\
&= \mu_{n-1}\left\{x' \in S^{n-1} : q_0(x') \le \frac{n}{n+1}\frac{t}{\cos^2\varphi}\mathbf{E}(q_0)\right\} \\
&\le C_{n-1}\sqrt{\frac{n}{n+1}}\frac{\sqrt{t}}{\cos\varphi}
\end{aligned}$$

by the induction conjecture. Therefore,

$$\begin{aligned}
&\mu_n\{x \in S^n : q(x) \le t\mathbf{E}(q)\} \\
&\quad= \frac{\kappa_{n-1}}{\kappa_n}\int_{-\pi/2}^{\pi/2} \mu_{n-1,\varphi}\{x \in S_\varphi^{n-1} : q(x) \le t\mathbf{E}(q)\}\cos^{n-1}\varphi\, d\varphi \\
&\quad\le \sqrt{t}C_{n-1}\sqrt{\frac{n}{n+1}}\frac{\kappa_{n-1}}{\kappa_n}\int_{-\pi/2}^{\pi/2}\cos^{n-2}\varphi\, d\varphi \\
&\quad= C_{n-1}\sqrt{\frac{n}{n+1}}\frac{\kappa_{n-1}^2}{\kappa_n\kappa_{n-2}}\sqrt{t} = C_n\sqrt{t}
\end{aligned}$$

and the proof follows. □

It follows from the proof that for small t the value of $\mu\{x \in S^{n-1} : q(x) \le t\mathbf{E}(q)\}$ is the largest when rank $q = 1$. For a "typical" quadratic form q we should expect a sharper concentration of its values around $\mathbf{E}(q)$.

(3.4) Corollary. *Let $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a positive definite quadratic form such that $\mathbf{E}(q) = 1$. Then*

$$|\mathbf{E}(\ln q)| \le 2C_0 \quad \textit{and} \quad \mathbf{E}(\ln^2 q) \le \ln^2 n + 8C_0,$$

where C_0 is the absolute constant from Theorem 3.3.

Proof. Since $\ln x$ is a concave function, we get that $\mathbf{E}(\ln q) \le \ln \mathbf{E}(q) \le 0$. Lemma 3.1 implies that the largest eigenvalue $M(q)$ of the form q does not exceed n, so $q(x) \le n$ for every $x \in S^{n-1}$.

Using integration by parts we may write

$$\begin{aligned}\mathbf{E}(\ln q) &= \int_0^n \ln t \, d\mu\{x \in S^{n-1} : q(x) \le t\} \ge \int_0^1 \ln t \, d\mu\{x \in S^{n-1} : q(x) \le t\} \\ &= (\ln t) \cdot \mu\{x \in S^{n-1} : q(x) \le t\}|_{t \longrightarrow +0}^{t=1} - \int_0^1 t^{-1} \mu\{x \in S^{n-1} : q(x) \le t\} \, dt.\end{aligned}$$

Applying Theorem 3.3 we conclude that

$$\lim_{t \longrightarrow +0} (\ln t) \cdot \mu\{x \in S^{n-1} : q(x) \le t\} = 0$$

so we get the estimate

$$\begin{aligned}\int_0^1 \ln t \, d\mu\{x \in S^n : q(x) \le t\} &= -\int_0^1 t^{-1} \mu\{x \in S^{n-1} : q(x) \le t\} \, dt \\ &\ge -C_0 \int_0^1 t^{-1/2} \, dt = -2C_0,\end{aligned}$$

so the first inequality is proven.

Similarly,

$$\mathbf{E}(\ln^2 q) = \int_0^n \ln^2 t \, d\mu\{x \in S^{n-1} : q(x) \le t\}$$

$$\int_0^1 \ln^2 t \, d\mu\{x \in S^{n-1} : q(x) \le t\} + \int_1^n \ln^2 t \, d\mu\{x \in S^{n-1} : q(x) \le t\}.$$

For the second integral we get a trivial estimate

$$\int_1^n \ln^2 t \, d\mu\{x \in S^{n-1} : q(x) \le t\} \le (\ln^2 n) \cdot \mu\{x \in S^{n-1} : q(x) \le n\} \le \ln^2 n.$$

Using Theorem 3.3 we estimate the first integral

$$\begin{aligned}&\int_0^1 \ln^2 t \, d\mu\{x \in S^{n-1} : q(x) \le t\} \\ &= (\ln^2 t) \cdot \mu\{x \in S^{n-1} : q(x) \le t\}|_{t \longrightarrow +0}^1 - \int_0^1 2t^{-1} \ln t \, \mu\{x \in S^{n-1} : q(x) \le t\} \, dt \\ &\le -2C_0 \int_0^1 t^{-1/2} \ln t \, dt \le 8C_0.\end{aligned}$$

□

It is possible to find a tight bound for $|\mathbf{E}(\ln q)|$ when n is sufficiently large.

(3.5) Theorem. *We have*

$$\lim_{n \longrightarrow +\infty} \sup\{|\mathbf{E}(\ln q)|, \ q\colon \mathbb{R}^n \longrightarrow \mathbb{R} \text{ is positive semidefinite and } \mathbf{E}(q) = 1\} = C_1,$$

where

$$C_1 = -\frac{4}{\sqrt{2\pi}} \int_0^{+\infty} (\ln t) e^{-t^2/2}\, dt \approx 1.270362845.$$

Proof. Let $\lambda_1, \ldots, \lambda_k$ be nonzero eigenvalues of a positive semidefinite form $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ and let $u_1, \ldots, u_k$ be the corresponding unit eigenvectors. Then

$$q(x) = \sum_{i=1}^{k} \lambda_i \langle u_i, x\rangle^2 = \sum_{i=1}^{k} \alpha_i q_i(x), \qquad \text{where} \quad \alpha_i = \frac{\lambda_i}{n} \quad \text{and} \quad q_i(x) = n\langle u_i, x\rangle^2.$$

Suppose that $\mathbf{E}(q) = 1$. Then Lemma 3.1 implies that $\alpha_1 + \cdots + \alpha_k = 1$ and $\mathbf{E}(q_i) = 1$. Since $\ln x$ is a concave function we have

$$\begin{aligned} 0 \;&\geq\; \mathbf{E}(\ln q) = \mathbf{E}(\ln(\alpha_1 q_1 + \cdots + \alpha_k q_k)) \geq \mathbf{E}(\alpha_1 \ln q_1 + \cdots + \alpha_k \ln q_k) \\ &\geq\; \alpha_1 \mathbf{E}(\ln q_1) + \cdots + \alpha_k \mathbf{E}(\ln q_k) \geq \min\{\mathbf{E}(\ln q_i) :\; i = 1, \ldots, k\}. \end{aligned}$$

Therefore the supremum in question is attained on positive semidefinite forms q of rank 1. Without loss of generality we may choose $q(x) = nx_1^2$. We get

$$\mathbf{E}(\ln q) = \frac{\kappa_{n-2}}{\kappa_{n-1}} \int_{-\pi/2}^{\pi/2} \ln(n \sin^2\varphi) \cos^{n-2}\varphi\, d\varphi = \frac{2\kappa_{n-2}}{\kappa_{n-1}} \int_0^{\pi/2} \ln(n \sin^2\varphi) \cos^{n-2}\varphi\, d\varphi.$$

It is easy to see that $\cos\varphi \leq e^{-\varphi^2/2}$ for $0 \leq \varphi \leq \pi/2$ (the function $e^{\varphi^2/2}\cos\varphi$ is decreasing on $[0, \pi/2]$) and hence $\cos^{n-2}\varphi \leq e^{(2-n)\varphi^2/2}$. Let us choose a sufficiently small $\varepsilon > 0$, say $\varepsilon = 0.1$. Then

$$\mathbf{E}(\ln q) = \frac{2\kappa_{n-2}}{\kappa_{n-1}} \int_0^{n^{-1/2+\varepsilon}} \ln(n \sin^2\varphi) \cos^{n-2}\varphi\, d\varphi + O(e^{-n\varepsilon}).$$

Substitution $\varphi = t/\sqrt{n}$ reduces the integral to

$$\frac{2\kappa_{n-2}}{\kappa_{n-1}\sqrt{n}} \int_0^{n^{\varepsilon}} \ln\left(n \sin^2 \frac{t}{\sqrt{n}}\right) \cos^{n-2}\frac{t}{\sqrt{n}}\, dt.$$

Now $\lim_{n\longrightarrow\infty} 2\kappa_{n-2}/\kappa_{n-1}\sqrt{n} = 2/\sqrt{2\pi}$ (see Section 1.5).

On the interval $[0, n^{\varepsilon}]$ we have: $n \sin^2(t/\sqrt{n}) = t^2 + O(t^4/n) = t^2(1 + O(n^{2\varepsilon-1}))$. Therefore $\ln(n \sin^2(t/\sqrt{n})) = \ln t^2 + O(n^{2\varepsilon-1})$. Similarly,

$$\cos\frac{t}{\sqrt{n}} = 1 - \frac{t^2}{2n} + O\left(\frac{t^4}{n^2}\right) = 1 - \frac{t^2}{2n} + O(n^{4\varepsilon-2}),$$

so

$$\cos^{n-2}\frac{t}{\sqrt{n}} = e^{-t^2/2}(1 + O(n^{4\varepsilon-1})).$$

Therefore

$$\int_0^{n^{\varepsilon}} \ln\left(n \sin^2 \frac{t}{\sqrt{n}}\right) \cos^{n-2}\frac{t}{\sqrt{n}}\, dt = (1 + O(n^{4\varepsilon-1})) \int_0^{n^{\varepsilon}} (\ln t^2) e^{-t^2/2}\, dt + O(n^{4\varepsilon-1}).$$

Finally, we get

$$\lim_{n\longrightarrow+\infty} \mathbf{E}(\ln(nx_1^2)) = \lim_{n\longrightarrow+\infty} \frac{2}{\sqrt{2\pi}} \int_0^{n^\varepsilon} (\ln t^2) e^{-t^2/2}\, dt = \frac{4}{\sqrt{2\pi}} \int_0^{+\infty} (\ln t) e^{-t^2/2}\, dt$$

and the proof follows. □

4. Integration on the Orthogonal Group

We need to invoke some integration technique on the orthogonal group O_n with respect to the Haar probability measure ν.

Let $C(O_{n,s})$ be the Banach space of all continuous functions on the Stiefel manifold $O_{n,s}$ of all s-tuples $(u_1, \ldots, u_s)$ of pairwise orthogonal vectors in $\mathbb{R}^n$ (see Section 1.5) with the norm $\|f\| = \max\{|f(x)| : x \in O_{n,s}\}$. The natural action of the orthogonal group O_n on $O_{n,s}$: $A(u_1, \ldots, u_s) = (A(u_1), \ldots, A(u_s))$ induces the action on $C(O_{n,s})$: $A(f)(x) = f(A^{-1}x),\ A \in O_n$. We agree that $C(O_{n,0}) = \mathbb{R}$, the space of constants with the trivial action of O_n.

(4.1) Operators $\mathbf{E}_s$ ("Conditional Expectations"). We define an operator $\mathbf{E}_s$: $C(O_{n,s}) \longrightarrow C(O_{n,s-1})$ as follows. For f: $O_{n,s} \longrightarrow \mathbb{R}$ we let

$$g = \mathbf{E}_s(f), \qquad g(u_1, \ldots, u_{s-1}) = \int_{S^{n-s} \subset (u_1, \ldots, u_{s-1})^\perp} f(u_1, \ldots, u_{s-1}, u_s)\, du_s,$$

where S^{n-s} is the unit sphere in the orthogonal complement $(u_1, \ldots, u_{s-1})^\perp$ and du_s is the rotation invariant Borel probability measure on S^{n-s}. We summarize a few obvious properties of $\mathbf{E}_s$:

- Operators $\mathbf{E}_s$ are linear and monotone, that is, if $f(x) \geq g(x)$ for all $x \in O_{n,s}$ then $\mathbf{E}_s(f)(x) \geq \mathbf{E}_s(g)(x)$ for all $x \in O_{n,s-1}$. Furthermore, $\mathbf{E}_s(\mathbf{1}) = \mathbf{1}$, where $\mathbf{1}$ is the function on $O_{n,s}$ that is identically 1. It follows then that $\mathbf{E}_s$ are continuous linear operators of the norm 1.
- Operators $\mathbf{E}_s$ commute with the action of the orthogonal group, that is, $\mathbf{E}_s(A(f)) = A(\mathbf{E}_s(f))$ for any $f \in C(O_{n,s})$ and any $A \in O_n$.
- Operators $\mathbf{E}_s$ are partially multiplicative: if g is a continuous function on $O_{n,s-1}$ and h is a continuous function on $O_{n,s}$, then $f(u_1, \ldots, u_s) = g(u_1, \ldots, u_{s-1})h(u_1, \ldots, u_s)$ is a continuous function on $O_{n,s}$ and $\mathbf{E}_s(f) = g\mathbf{E}_s(h)$.

We note that $\mathbf{E}_1(f)$ is just the average value of f on the unit sphere S^{n-1}.

(4.2) Lemma. *Let f: $O_n \longrightarrow \mathbb{R}$, $f = f(u_1, \ldots, u_n)$ be a continuous function on O_n. Then*

$$\int_{O_n} f\, d\nu = \mathbf{E}_1\mathbf{E}_2 \cdots \mathbf{E}_{n-1}\mathbf{E}_n(f).$$

Proof. Let us consider the map $\psi(f) = \mathbf{E}_1 \cdots \mathbf{E}_n(f)$, ψ: $C(O_n) \longrightarrow \mathbb{R}$. From (4.1) it follows that ψ is a continuous linear functional, so by Riesz's theorem $\psi(f) = \int_{O_n} f\, d\tau$

for some unique Borel measure τ on O_n. Furthermore, from (4.1) we have $\psi(A(f)) = \psi(f)$ for every $A \in O_n$ and $\psi(\mathbf{1}) = 1$. Therefore τ is an invariant probability measure, so we must have $\tau = \nu$ because the Haar probability measure is unique. □

Lemma 4.2 can be generalized to "piecewise continuous" functions on semialgebraic pieces in $O_{n,s}$ or to L^2 functions as follows from the formula for the volume element in O_n (see, for example, Chapter 12 of [20]). However, we do not need it in that generality. The following lemma will be instrumental for the analysis of our main algorithm in Section 5. It is a special case of the law of large numbers for martingales.

(4.3) Lemma. *Let $f_s\colon O_{n,s} \longrightarrow \mathbb{R}$, $s = 1, \ldots, n$, be continuous functions such that*

$$\|\mathbf{E}_s(f_s)\| \le a_s \quad \textit{and} \quad \|\mathbf{E}_s(f_s^2)\| \le b, \qquad s = 1, \ldots, n,$$

for some numbers a_s and b. Let us define a function $F\colon O_n \longrightarrow \mathbb{R}$ by

$$F(u_1, \ldots, u_n) = \frac{1}{n} \sum_{s=1}^{n} f_s(u_1, \ldots, u_s)$$

and let

$$a = \frac{1}{n} \sum_{s=1}^{n} a_s.$$

Then for any $\varepsilon > 0$

$$\nu\{(u_1, \ldots, u_n) \in O_n : |F(u_1, \ldots, u_n)| \ge a + \varepsilon\} \le \frac{b}{\varepsilon^2 n}.$$

Proof. Let $g_s = \mathbf{E}_s(f_s)$ and $h_s(u_1, \ldots, u_s) = f_s(u_1, \ldots, u_s) - g_s(u_1, \ldots, u_{s-1})$. Since g_s does not depend on u_s we have $\mathbf{E}_s(g_s f_s) = g_s \mathbf{E}_s(f_s)$ and $\mathbf{E}_s(g_s^2) = g_s^2$. Therefore

$$\mathbf{E}_s(h_s^2) = \mathbf{E}_s(f_s^2 - 2 f_s g_s + g_s^2) = \mathbf{E}_s(f_s^2) - 2 g_s \mathbf{E}_s(f_s) + \mathbf{E}_s(g_s^2) = \mathbf{E}_s(f_s^2) - g_s^2.$$

Since the operators $\mathbf{E}_s$ are monotone, the functions $\mathbf{E}_s(h_s^2)$ and $\mathbf{E}_s(f_s^2)$ are nonnegative, so we get $\|\mathbf{E}_s(h_s^2)\| \le \|\mathbf{E}(f_s^2)\| \le b$. Summarizing, we get

$$f_s = h_s + g_s, \qquad \text{where} \quad \mathbf{E}_s(h_s) = 0, \qquad \|\mathbf{E}_s(h_s^2)\| \le b \quad \text{and} \quad \|g_s\| \le a_s.$$

Let

$$H(u_1, \ldots, u_n) = \frac{1}{n} \sum_{s=1}^{n} h_s(u_1, \ldots, u_s).$$

So we have

$$\|F - H\| = \left\| \frac{1}{n} \sum_{s=1}^{n} g_s \right\| \le a. \tag{4.3.1}$$

We have

$$H^2 = \frac{1}{n^2}\sum_{s=1}^{n} h_s^2 + \frac{2}{n^2}\sum_{i<j} h_i h_j.$$

We claim that for every pair $i < j$

$$\int_{O_n} h_i h_j \, d\nu = 0,$$

where we consider h_s as a function on O_n by letting $h_s(u_1, \ldots, u_n) = h_s(u_1, \ldots, u_s)$. Indeed, by Lemma 4.2

$$\int_{O_n} h_i h_j \, d\nu = \mathbf{E}_1 \cdots \mathbf{E}_n (h_i h_j).$$

Since the function $h_i h_j$ does not depend on $u_{j+1}, \ldots, u_n$ we have that $\mathbf{E}_{j+1} \cdots \mathbf{E}_n (h_i h_j) = h_i h_j$ as a function on $O_{n,j}$. Furthermore, since $i < j$ and h_i does not depend on u_j, we have that $\mathbf{E}_j (h_i h_j) = h_i \mathbf{E}_j (h_j) = 0$. Therefore

$$\int_{O_n} H^2 \, d\nu = \frac{1}{n^2}\sum_{s=1}^{n} \int_{O_n} h_s^2 \, d\nu \leq \frac{b}{n}.$$

Now the proof follows because of (4.3.1) and the Chebyshev inequality

$$\nu\{(u_1, \ldots, u_n) : |H(u_1, \ldots, u_n)| \geq \varepsilon\} \leq \varepsilon^{-2} \int_{O_n} H^2 \, d\nu \leq \frac{b}{\varepsilon^2 n}. \qquad \square$$

5. The Basic Algorithm

In this section we present our algorithm for computing the mixed discriminant of positive definite matrices $M_1, \ldots, M_n$. The main idea of the algorithm is to use Theorem 2.4 as is described in Section 2. The "random" part of the algorithm consists of choosing a random orthonormal basis $u_1, \ldots, u_n$ in the space $\mathbb{R}^n$. After that the algorithm is completely deterministic and reduces to standard Linear Algebra computations. Hence for any given input $M_1, \ldots, M_n$ the output of the algorithm is a function on the orthogonal group O_n. We use Theorem 2.4 to show that the expectation of the output is the mixed discriminant $D(M_1, \ldots, M_n)$ and we use the results of Section 3 and Lemma 4.3 to prove that with a sufficiently high probability the deviation from the expectation is within desired limits. To sample an orthonormal basis, we do the following: first, we choose u_1 from the rotation invariant probability distribution on the sphere S^{n-1}, then we choose u_2 from the rotation invariant probability distribution on the sphere $S^{n-2} \subset u^{\perp}$ and so forth; we choose u_s from the rotation invariant probability distribution on the sphere $S^{n-s} \subset (u_1, \ldots, u_{s-1})^{\perp}$. It is immediate that the simulated distribution is invariant under the action of the orthogonal group, so it must coincide with the Haar distribution ν (see also Lemma 4.2). Another possibility is to choose n vectors independently from the standard Gaussian distribution in $\mathbb{R}^n$ and apply the Gram–Schmidt orthogonalization process to them.

(5.1) Basic Algorithm.

Input. Positive definite matrices $M_1, \ldots, M_n$.

Output. A number $\alpha > 0$ approximating $D(M_1, \ldots, M_n)$.

The Algorithm

Step 0. Sample an orthonormal basis $(u_1, \ldots, u_n)$ in $\mathbb{R}^n$. Let A be the orthogonal matrix having u_i as its ith column. Let $Q_i := A^t M_i A$ for $i = 1, \ldots, n$, where A^t is the transpose of A. Let $\beta := 1$ and $s := 0$.

Comment. It is convenient to perform computations in the basis $u_1, \ldots, u_n$ of $\mathbb{R}^n$. Matrix Q_i in the basis $u_1, \ldots, u_n$ and matrix M_i in the standard basis represent the same self-adjoint operator. We store in s the number of iterations of Steps 1–2 of the algorithm and in β the current value of the mixed discriminant.

Step 1. Let $k = n - s$ and let $s := s + 1$. Let $\beta := \beta \det Q_1$. If $s = n$, let $\alpha := \beta$, output α, and stop. Otherwise compute a symmetric positive definite matrix T such that $T^2 = Q_1$. Compute $R_i = T^{-1} Q_i T^{-1}$ for $i = 2, \ldots, k$.

Comment. On the sth iteration of this step we have k positive definite operators $Q_1, \ldots, Q_k$ on the k-dimensional subspace $(u_1, \ldots, u_{s-1})^\perp$. These operators represented by the matrices in the basis $u_s, \ldots, u_n$ of that subspace. By (2.4.1) we have $D(Q_1, \ldots, Q_k) = (\det Q_1) D(I, R_2, \ldots, R_k)$. If $k > 1$, we store the factor $\det Q_1$ in β and proceed to Step 2 with the computation of $D(I, R_2, \ldots, R_k)$. Note, that for any positive definite operator Q_1 there exists a unique positive definite operator T such that $T^2 = Q_1$. In particular, it does not depend on the choice of a basis. Furthermore, T depends on Q_1 continuously (see, for example, Section 11 of Chapter 9 in [7]). To compute T, we compute the eigenvalues $\lambda_1, \ldots, \lambda_k$ of Q_1, compute the interpolating polynomial p such that $p(\lambda_i) = \sqrt{\lambda_i}$ and let $T = p(Q_1)$.

Step 2. For $i = 1, \ldots, k-1$ let Q_i be the $(k-1) \times (k-1)$ lower-right corner submatrix of R_{i+1}. Go to Step 1.

Comment. It is seen that $Q_i = P^* R_{i+1} P$ where $P : (u_1, \ldots, u_s)^\perp \subset (u_1, \ldots, u_{s-1})^\perp$ is the inclusion. Thus we have $Q_i = R_{i+1}|u_s^\perp$ (see Section 1.5). From (2.4.2) we have

$$D(I, R_2, \ldots, R_k) = \int_{S^{k-1} \subset (u_1, \ldots, u_{s-1})^\perp} D(R_2|u^\perp, \ldots, R_k|u^\perp)\, du,$$

where u ranges over the unit sphere S^{k-1} in $(u_1, \ldots, u_{s-1})^\perp$ and du is the rotation invariant probability measure on S^{k-1}. On this step of the algorithm we approximate $D(I, R_2, \ldots, R_n)$ by $D(R_2|u^\perp, \ldots, R_n|u^\perp)$ at the point $u = u_s$ and go to Step 1 again.

(5.2) Theorem. *For any given positive definite $n \times n$ matrices $M_1, \ldots, M_n$ the algorithm performs a polynomial in n number of operations* (*addition, subtraction, multiplication, division, and taking the square root of a nonnegative number*). *For any $\varepsilon > 0$ there is an $N(\varepsilon)$ such that for any $n \geq N(\varepsilon)$ the number α produced by the algorithm with probability at least* 0.9 *satisfies the inequalities*

$$c_\varepsilon^n D(M_1, \ldots, M_n) \leq \alpha \leq 20 D(M_1, \ldots, M_n) \qquad \text{with} \quad c_\varepsilon = e^{-C_1 - \varepsilon},$$

where C_1 is the absolute constant from Theorem 3.5.

Proof. The algorithm performs Steps 1 and 2 altogether n times and every operation reduces to the standard Linear Algebra computations: computing the factorization $Q = T^2$, the determinant $\det Q$, the inverse matrix T^{-1}, and the product of matrices. As is well known, for $n \times n$ matrices these operations require $O(n^3)$ arithmetic operations and computing the factorization also requires taking a square root n times and computing the eigenvalues of Q (see [7]).

Let us fix the input $M_1, \ldots, M_n$. Then the computations on every step are completely determined by the choice of a random basis $(u_1, \ldots, u_n)$ on Step 0 and the output $\alpha = \alpha(u_1, \ldots, u_n)$ is a continuous function on the orthogonal group O_n. Furthermore, on the sth iteration of Step 1 the operators $Q_1, \ldots, Q_k$ and $R_2, \ldots R_k$ depend only on the first $s-1$ vectors $u_1, \ldots, u_{s-1}$ although their particular matrix representation may depend on $u_s, \ldots, u_n$ as well.

For a set of s pairwise orthogonal unit vectors $u_1, \ldots, u_s$ in $\mathbb{R}^n$ let

$$q_s(u_1, \ldots, u_s) = \frac{\det Q_1}{D(Q_1, \ldots, Q_k)} D(R_2 | u_s^\perp, \ldots, R_k | u_s^\perp),$$

where $Q_1, \ldots, Q_k$ and $R_2, \ldots, R_k$ are the operators at the sth iteration of Step 1 and we agree that $q_n(u_1, \ldots, u_n) = 1$. Thus $q_s(u_1, \ldots, u_s)$ are continuous functions on the Stiefel manifold $O_{n,s}$.

We claim that

$$\alpha(u_1, \ldots, u_n) = D(M_1, \ldots, M_n) \cdot \prod_{s=1}^n q_s(u_1, \ldots, u_s); \tag{5.2.1}$$

that

$$\int_{O_n} \prod_{s=1}^n q_s(u_1, \ldots, u_s)\, d\nu = 1 \quad \text{and hence} \quad \int_{O_n} \alpha \, d\nu = D(M_1, \ldots, M_n), \tag{5.2.2}$$

and that

$$\nu \left\{ (u_1, \ldots, u_n) \in O_n : \prod_{s=1}^n q_s(u_1, \ldots, u_s) \leq c_\varepsilon^n \right\} \leq \frac{1}{20} \qquad \text{for all} \quad n \geq N(\varepsilon), \tag{5.2.3}$$

or equivalently

$$\nu\left\{(u_1,\ldots,u_n)\in O_n:\ \frac{1}{n}\sum_{s=1}^{n}\ln q_s(u_1,\ldots,u_s)\le -C_1-\varepsilon\right\}\le\frac{1}{20}\qquad(5.2.3')$$

for all $n\ge N(\varepsilon)$, where C_1 is the constant from Theorem 3.5.

Indeed, the matrix $R_{i+1}|u_s^{\perp}$ computed on the sth iteration of Step 2 is the matrix Q_i used for the $(s+1)$st iteration of Step 1 and we get (5.2.1).

By (2.4.1) and (2.4.2) for any fixed $u_1,\ldots,u_{s-1}\in O_{n,s-1}$ we get

$$\int_{S^{k-1}\subset(u_1,\ldots,u_{s-1})^{\perp}} q_s(u_1,\ldots,u_{s-1},u_s)\,du_s=\frac{(\det Q_1)D(I,R_2,\ldots,R_k)}{D(Q_1,\ldots,Q_k)}=1.$$

In other words $\mathbf{E}_s(q_s)=\mathbf{1}$, where $\mathbf{E}_s$ are the operators from Section 4.1 and $\mathbf{1}$ is the function on $O_{n,s-1}$ that is identically 1. Hence by Lemma 4.2

$$\begin{aligned}\int_{O_n}\prod_{s=1}^{n}q_s(u_1,\ldots,u_s)\,d\nu &= \mathbf{E}_1\cdots\mathbf{E}_n\prod_{s=1}^{n}q_s(u_1,\ldots,u_s)\\ &= \text{by (4.1)}\quad \mathbf{E}_1q_1(u_1)\cdots\mathbf{E}_nq_n(u_1,\ldots,u_n)=1,\end{aligned}$$

and we get (5.2.2). From (2.4.3) we conclude that $q_s(u_1,\ldots,u_s)$ is a positive definite quadratic form in $u_s\in(u_1,\ldots,u_{s-1})^{\perp}$ provided $u_1,\ldots,u_{s-1}$ are fixed. Let

$$a_s=\sup\{|\mathbf{E}(\ln q)|,\ q\colon\mathbb{R}^k\longrightarrow\mathbb{R}\text{ is positive semidefinite and }\mathbf{E}(q)=1,\ k=n-s+1\}.$$

Since $\mathbf{E}_s(q_s)=\mathbf{1}$ we have $\|\mathbf{E}_s(\ln q_s)\|\le a_s$ and from Corollary 3.4 we have $a_s\le 2C_0$ and $\|\mathbf{E}_s(\ln^2 q_s)\|\le\ln^2 n+8C_0$.

By Theorem 3.5

$$\frac{1}{n}\sum_{s=1}^{n}a_s\le C_1+\frac{\varepsilon}{2}$$

for all sufficiently large n. Furthermore,

$$\frac{\ln^2 n+8C_0}{(\varepsilon/2)^2 n}\le\frac{1}{20}$$

for all sufficiently large n. Now (5.2.3′) follows by Lemma 4.3 with a_s, $\varepsilon/2$, $b=\ln^2 n+8C_0$, and $f_s=\ln q_s(u_1,\ldots,u_s)$.

Since $\alpha(u_1,\ldots,u_n)$ is positive on O_n, by (5.2.2) we deduce that

$$\alpha\ge 20D(M_1,\ldots,M_n)$$

with probability at most $\frac{1}{20}$. Next, from (5.2.1) and (5.2.3) we deduce that $\alpha\le c_\varepsilon^n D(M_1,\ldots,M_n)$ with probability at most $\frac{1}{20}$. This completes the proof of the theorem. □

So any approximation constant

$$c_\varepsilon<\exp\left\{\frac{4}{\sqrt{2\pi}}\int_0^\infty(\ln t)e^{-t^2/2}\,dt\right\}\approx 0.2807297419$$

will work for a sufficiently large n.

(5.3) Corollary. *With the given matrices $M_1, \ldots, M_n$ let us run Algorithm* 5.1 *independently* $2m$ *times and let α_0 be the median of the computed α's. Then for $n \geq N(\varepsilon)$ the number α_0 satisfies the inequalities*

$$c_\varepsilon^n D(M_1, \ldots, M_n) \leq \alpha_0 \leq 20 D(M_1, \ldots, M_n)$$

with probability at least $1 - (0.4)^m$, where $N(\varepsilon)$ and $c_\varepsilon > 0$ are the constants from Theorem 5.2.

Proof. If α_0 does not satisfy the inequalities, then at least m of the computed α's do not. The probability of this event is

$$\sum_{k=0}^{m} \binom{2m}{k} (0.9)^k (0.1)^{2m-k} \leq (0.1)^m \sum_{k=0}^{2m} \binom{2m}{k} \leq (0.1)^m 4^m = (0.4)^m. \qquad \square$$

So to achieve an overwhelming probability $1 - \delta$ we have to run Algorithm 5.1 $O(\log \delta^{-1})$ times and choose the median of the computed α's.

Algorithm 5.1 can be converted into a randomized polynomial time algorithm for approximating the mixed discriminant within a factor $2^{O(n)}$ in the bit model of computation. One should simulate the uniform distribution on the sphere with a sufficiently high precision from the standard Bernoulli distribution using the Central Limit Theorem. Then all the computations that require finding the roots of a univariate polynomial (the only nonrational operation we used) should be approximated well enough by the arithmetic operations over the rationals. The bit version of Algorithm 5.1 will be presented elsewhere.

It would be interesting to investigate the behavior of Algorithm 5.1 for "average" matrices $M_1, \ldots, M_n$. One can show that the algorithm works worst if on every iteration of Step 2 matrices $R_2, \ldots, R_k$ are very close to matrices of rank 1, that is, each has precisely one eigenvalue that is much larger than the remaining $k-1$ eigenvalues (see the remarks after Theorem 3.3 and Corollary 3.2). On the other hand, if $M_1 = \cdots = M_n = I$, then the algorithm always outputs the precise value $\alpha = 1$. One can conjecture that for an "average" input the algorithm gives a much better approximation and, possibly, gives rise to a polynomial time approximation scheme. A possible approach to this problem is via the "measure concentration phenomenon" on the orthogonal group (see Section 6 of [17]). We represented $D(M_1, \ldots, M_n)$ as the integral of some continuous density α on O_n. If $M_1, \ldots, M_n$ are "average" we can expect that the function α has nice Lipschitz properties and therefore is sharply concentrated about its average value.

6. Computing the Permanent of a Nonnegative Matrix

We are going to apply our algorithm to computing the permanent of a nonnegative matrix. First, we establish a known connection between mixed discriminants and permanents, that is, part (1.4.1) of Theorem 1.4.

Proof of (1.4.1). We observe that $t_1M_1 + \cdots + t_nM_n$ is a diagonal matrix for any $t_1, \ldots, t_n$ and

$$\det(t_1M_1 + \cdots + t_nM_n) = \prod_{j=1}^{n} \sum_{i=1}^{n} t_i a_{ij}.$$

It is easy to see that

$$\operatorname{per} A = \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \prod_{j=1}^{n} \sum_{i=1}^{n} t_i a_{ij}.$$

Comparing this with (1.3.1) we get the desired formula.

(6.1) The Algorithm. Algorithm 5.1 accepts only positive definite matrices as its input. This allows us to compute the permanent of a positive matrix. To compute the permanent of a nonnegative matrix (this is the most interesting case) we will just put sufficiently small positive numbers instead of zeros.

Input. An $n \times n$ nonnegative matrix B.

Output. A number β approximating per B.

The Algorithm

Step 0. Compute $m = \min\{b_{ij} : b_{ij} > 0\}$. Let $b_{ij} := b_{ij}/m$ for $i, j = 1, \ldots, n$.

Step 1. Compute $M = \max\{b_{ij}: i, j = 1, \ldots, n\}$. Let $\delta = c_1^n/40n!\, M^{n-1}$, where $c_1 > 0$ (that is, c_ε for $\varepsilon = 1$) is the constant from Theorem 5.2. Define an $n \times n$ matrix $A = (a_{ij})$ as follows:

$$a_{ij} = \begin{cases} b_{ij} & \text{if} \quad b_{ij} > 0, \\ \delta & \text{if} \quad b_{ij} = 0. \end{cases}$$

Let $M_i = \operatorname{diag}\{a_{i1}, \ldots, a_{in}\}$, $i = 1, \ldots, n$. Apply Algorithm 5.1 with the matrices $M_1, \ldots, M_n$ in the input and let α be the output. Let $\beta = n!\,\alpha$.

Step 2. If $\beta \le 2c_1^n/3$, let $\beta := 0$, output β, and stop.
Otherwise, let $\beta := m^n\beta/21$, output β, and stop.

(6.2) Theorem. *For any given $n \times n$ nonnegative matrix B the algorithm performs a polynomial in n number of operations. For any $1 > \varepsilon > 0$ and any $n \ge N(\varepsilon)$ the number β produced by the algorithm with probability at least* 0.9 *satisfies the inequalities*

$$\frac{c_\varepsilon^n}{21} \operatorname{per} B \le \beta \le \operatorname{per} B,$$

where $N(\varepsilon)$ and $c_\varepsilon > 0$ are the constants from Theorem 5.2.

Proof. It is immediate from Theorem 5.2 that the algorithm performs a polynomial in n number of operations. Step 0 reduces the problem to the case where all positive entries of B are not smaller than 1. For matrices A and B on Step 1 we have

$$\operatorname{per} B \leq \operatorname{per} A \leq \operatorname{per} B + \frac{c_1^n}{40}, \tag{6.2.1}$$

since every one of $n!$ terms $\prod_{i=1}^n a_{i\sigma(i)}$ of the expansion of $\operatorname{per} A$ is either a term of $\operatorname{per} B$ or does not exceed δM^{n-1}. Theorem 5.2 and (1.4.1) imply that for $n \geq N(\varepsilon)$ with probability at least 0.9 on Step 1 we have

$$c_\varepsilon^n \operatorname{per} A \leq \beta \leq 20 \operatorname{per} A. \tag{6.2.2}$$

So suppose that (6.2.2) are satisfied. There are two cases. If $\operatorname{per} B = 0$, then $\operatorname{per} A \leq c_1^n/40$ by (6.2.1) and $\beta \leq c_1^n/2$ by (6.2.2) so the algorithm outputs $\beta = 0$. If $\operatorname{per} B > 0$, then $\operatorname{per} B \geq 1$ and hence $\operatorname{per} A \geq 1$. Then from (6.2.2) we have $\beta \geq c_\varepsilon^n > c_1^n$, so the algorithm proceeds to the last line. By (6.2.1) and $\operatorname{per} B \geq 1$ we get $\operatorname{per} A \leq (41/40) \operatorname{per} B$, and then (6.2.2) implies the desired inequality. □

As in Section 5, to get an overwhelming probability we have to run Algorithm 6.1 several times and choose the median of the computed β's.

7. Mixed Volumes of Ellipsoids and Mixed Discriminants

In this section we prove (1.4.2) of Theorem 1.4.

(7.1) Lemma. *For any positive definite operators $R_2, \ldots, R_n$ on $\mathbb{R}^n$ one has*

$$\int_{S^{n-1}} D^{1/2}(R_2|u^\perp, \ldots, R_n|u^\perp)\, du \geq \frac{1}{\sqrt{3}} \left(\int_{S^{n-1}} D(R_2|u^\perp, \ldots, R_n|u^\perp)\, du \right)^{1/2}.$$

Proof. Let $f(u) = D(R_2|u^\perp, \ldots, R_n|u^\perp)$. Applying the Hölder inequality

$$\int_{S^{n-1}} h(u)g(u)\, du \leq \left(\int_{S^{n-1}} h^p(u)\, du \right)^{1/p} \left(\int_{S^{n-1}} g^q(u)\, du \right)^{1/q}, \qquad \frac{1}{p} + \frac{1}{q} = 1,$$

with $h = f^{1/3}$, $g = f^{2/3}$, $p = \frac{3}{2}$ and $q = 3$ we get

$$\int_{S^{n-1}} f(u)\, du \leq \left(\int_{S^{n-1}} f^{1/2}(u)\, du \right)^{2/3} \left(\int_{S^{n-1}} f^2(u)\, du \right)^{1/3}$$

or

$$\left(\int_{S^{n-1}} f(u)\, du \right)^3 \leq \left(\int_{S^{n-1}} f^{1/2}(u)\, du \right)^2 \left(\int_{S^{n-1}} f^2(u)\, du \right).$$

Part (2.4.3) of Theorem 2.4 implies that $f(u)$ is a positive definite quadratic form on S^{n-1}. Corollary 3.2 asserts that

$$\int_{S^{n-1}} f^2(u)\,du \le 3\left(\int_{S^{n-1}} f(u)\,du\right)^2.$$

Therefore,

$$\int_{S^{n-1}} f(u)\,du \le 3\left(\int_{S^{n-1}} f^{1/2}(u)\,du\right)^2$$

and the proof follows. □

With a positive definite operator Q on $\mathbb{R}^n$ we associate the ellipsoid

$$E_Q = \{x \in \mathbb{R}^n : \langle x, Qx\rangle \le 1\}.$$

Next, we want to describe the orthogonal projection of an ellipsoid onto a hyperplane.

(7.2) Lemma. *Let $E_Q \subset \mathbb{R}^n$ be an ellipsoid and let $L \subset \mathbb{R}^n$ be a hyperplane. Then the image $(E_Q)|L$ of E_Q under the orthogonal projection onto L is the ellipsoid*

$$E_{Q'} = \{x \in L : \langle x, Q'x\rangle \le 1\} \qquad \text{where} \quad Q' = (Q^{-1}|L)^{-1}.$$

Proof. As is easy to see, for the support function of $K = E_Q$ we have $h_K(u) = \sqrt{\langle u, Q^{-1}u\rangle}$. Since the support function of the orthogonal projection onto a subspace is the restriction of the support function onto the subspace and a convex compact set is uniquely determined by its support function, the result follows (see Section 1.5). □

Finally, we need a standard result from integral geometry (a version of the kinematic formula).

(7.3) Theorem. *Let $Q_1, \ldots, Q_n$ be positive definite operators on $\mathbb{R}^n$.*

(7.3.1) *Suppose that $Q_1 = T^*T$ for some nondegenerate T. Let $R_k = (T^{-1})^* Q_k T^{-1}$ for $k = 2, \ldots, n$. Then*

$$V(E_{Q_1}, \ldots, E_{Q_n}) = (\det Q_1)^{-1/2} V(B, E_{R_2}, \ldots, E_{R_n}),$$

where $B \subset \mathbb{R}^n$ is the unit ball.

(7.3.2)
$$\int_{S^{n-1}} V(E_{R_2}|u^\perp, \ldots, E_{R_n}|u^\perp)\,du = \frac{v_{n-1}}{v_n} V(B, E_{R_2}, \ldots, E_{R_n}).$$

(7.3.3) *There exists a zonoid $K = K(E_{R_2}, \ldots, E_{R_n})$ in $\mathbb{R}^n$, called the mixed projection body of the ellipsoids $E_{R_2}, \ldots, E_{R_n}$ such that*

$$V(E_{R_2}|u^\perp, \ldots, E_{R_n}|u^\perp) = h_K(u)$$

for any $u \in S^{n-1}$.

Proof. The operator T maps the ellipsoid E_{Q_1} onto the ball B and the ellipsoid E_{Q_i} onto E_{R_i} for $i = 2, \ldots, n$. Since $\det Q_1 = \det^2 T$ we get (7.3.1). Integral representation (7.3.2) and the existence of the mixed projection body are known (see Section 3 of [22] and [16]). □

Support function $h_K(u)$ is also known as the "mixed brightness" of the ellipsoids $E_{Q_1}, \ldots, E_{Q_n}$. It is the analogue of the mixed quadratic form of Theorem 2.4.

Proof of (1.4.2). We proceed by induction on n. For $n = 1$ the estimates are obviously correct since $V(E_Q) = v_n \det^{-1/2} Q$. Let us consider n ellipsoids $E_{Q_1}, \ldots, E_{Q_n}$ in $\mathbb{R}^n$. Comparing (7.3.1) and (2.4.1) we conclude that it is enough to prove the inequalities, assuming that $Q_1 = I$ and $E_{Q_1} = B$ is the unit ball. Applying the induction conjecture and Lemma 7.2 to the integrand in (7.3.2) we get

$$\begin{aligned}(\sqrt{3})^{-n+2} v_{n-1} \int_{S^{n-1}} D^{1/2}(R_2^{-1}|u^\perp, \ldots, R_n^{-1}|u^\perp)\, du \\ \leq \int_{S^{n-1}} V(E_{R_2}|u^\perp, \ldots, E_{R_n}|u^\perp)\, du \\ \leq v_{n-1} \int_{S^{n-1}} D^{1/2}(R_2^{-1}|u^\perp, \ldots, R_n^{-1}|u^\perp)\, du.\end{aligned}$$

Applying Lemma 7.1 to the first integral and the Cauchy–Schwartz inequality to the last integral we get:

$$\begin{aligned}(\sqrt{3})^{-n+1} v_{n-1} \left(\int_{S^{n-1}} D(R_2^{-1}|u^\perp, \ldots, R_n^{-1}|u^\perp)\, du \right)^{1/2} \\ \leq \int_{S^{n-1}} V(E_{R_2}|u^\perp, \ldots, E_{R_n}|u^\perp)\, du \\ \leq v_{n-1} \left(\int_{S^{n-1}} D(R_2^{-1}|u^\perp, \ldots, R_n^{-1}|u^\perp)\, du \right)^{1/2}.\end{aligned}$$

Applying (2.4.2) to the first and last integrals and (7.3.2) to the middle integral we get

$$\begin{aligned}(\sqrt{3})^{-n+1} v_n D^{1/2}(I, R_2^{-1}, \ldots, R_n^{-1}) &\leq V(B, E_{R_2}, \ldots, E_{R_n}) \\ &\leq v_n D^{1/2}(I, R_2^{-1}, \ldots, R_n^{-1})\end{aligned}$$

and the proof follows. □

Inequality (1.4.2) and Theorem 5.2 imply immediately that we can approximate the mixed volume of given n ellipsoids in $\mathbb{R}^n$ within a factor $2^{O(n)}$ in randomized polynomial time. However, we can use Theorem 7.3 directly to construct an algorithm for computing the mixed volume of ellipsoids. This way we get an unbiased estimator with a better constant. Namely, we start with n ellipsoids $E_1, \ldots, E_n$ in $\mathbb{R}^n$. Applying a nondegenerate linear transform T we make the unit ball $B = T(E_1)$ from the first ellipsoid. Then we choose a unit vector $u \in S^{n-1}$ at random and project $T(E_{i+1})$ orthogonally onto $u^\perp$ getting an $(n-1)$-dimensional ellipsoid E_i'. Then we replace the computation of

$V(E_1, \ldots, E_n)$ by the computation of $V(E'_1, \ldots, E'_{n-1})$ and proceed as above. To prove an analogue of Theorem 5.2 we need to prove the analogues of the results from Section 3 where instead of a positive semidefinite quadratic form q we have the support function h_K of a zonoid K. The author cannot prove an analogue of Theorem 3.3 but the analogues of Corollary 3.4 and Theorem 3.5 can be obtained.

8. Support Functions of Zonoids

We recall from Section 3 that $\mathbf{E}(f)$ denotes the average value of a continuous function f on the unit sphere S^{n-1}. Our reasoning is somewhat parallel to that of Section 3; instead of positive semidefinite quadratic forms q we consider the support functions h_K of zonoids.

(8.1) Lemma. *Let $J \subset \mathbb{R}^n$, $J = -J$ be a segment of length $2l$ such that $\mathbf{E}(h_J) = 1$. Then*

$$l = \sqrt{\frac{\pi n}{2}}(1 + O(n^{-1})); \tag{8.1.1}$$

$$\lim_{n \longrightarrow +\infty} \mathbf{E}(\ln h_J) = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} \ln\left(t\sqrt{\frac{\pi}{2}}\right) e^{-t^2/2}\, dt \approx -0.4093900697; \tag{8.1.2}$$

$$\lim_{n \longrightarrow +\infty} \mathbf{E}(\ln^2 h_J) = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} \ln^2\left(t\sqrt{\frac{\pi}{2}}\right) e^{-t^2/2}\, dt \approx 1.401300779. \tag{8.1.3}$$

Proof. Without loss of generality we assume that $J = [-le_1, le_1]$, where $e_1 = (1, 0, \ldots, 0)$, so $h_J(x) = l|x_1| = l|\sin\varphi|$. We have ($n > 1$):

$$\begin{aligned} \mathbf{E}(h_J) &= \frac{\kappa_{n-2}}{\kappa_{n-1}} \int_{-\pi/2}^{\pi/2} l|\sin\varphi| \cos^{n-2}\varphi\, d\varphi \\ &= \frac{2\kappa_{n-2} l}{\kappa_{n-1}} \int_0^{\pi/2} \sin\varphi \cos^{n-2}\varphi\, d\varphi = \frac{2l\kappa_{n-2}}{(n-1)\kappa_{n-1}}. \end{aligned}$$

So $l = (n-1)\kappa_{n-1}/2\kappa_{n-2}$ and (8.1.1) follows (see Section 1.5).

Now we follow the proof of Theorem 3.5. For a sufficiently small $\varepsilon > 0$ we have

$$\begin{aligned} \mathbf{E}(\ln h_J) &= \frac{2\kappa_{n-2}}{\kappa_{n-1}} \int_0^{\pi/2} \ln(l\sin\varphi) \cos^{n-2}\varphi\, d\varphi \\ &= \frac{2\kappa_{n-2}}{\kappa_{n-1}} \int_0^{n^{-1/2+\varepsilon}} \ln(l\sin\varphi) \cos^{n-2}\varphi\, d\varphi + O(e^{-n\varepsilon}). \end{aligned}$$

Substitution $\varphi = t/\sqrt{n}$ transforms the integral to

$$\frac{2\kappa_{n-2}}{\kappa_{n-1}\sqrt{n}} \int_0^{n^{\varepsilon}} \ln\left(l\sin\frac{t}{\sqrt{n}}\right) \cos^{n-2}\frac{t}{\sqrt{n}}\, dt.$$

As in the proof of Theorem 3.5 we have

$$\lim_{n\longrightarrow+\infty}\frac{2\kappa_{n-2}}{\kappa_{n-1}\sqrt{n}}=\frac{2}{\sqrt{2\pi}}\qquad\text{and}\qquad\cos^{n-2}\frac{t}{\sqrt{n}}=e^{-t^2/2}(1+O(n^{4\varepsilon-1}))$$

on the interval $[0, n^{\varepsilon}]$. Using (8.1.1) we conclude that

$$l\sin\frac{t}{\sqrt{n}}=\sqrt{\frac{\pi}{2}}t(1+O(n^{2\varepsilon-1})),\qquad\text{so}\quad\ln\left(l\sin\frac{t}{\sqrt{n}}\right)=\ln\left(\sqrt{\frac{\pi}{2}}t\right)+O(n^{-1+2\varepsilon}).$$

Therefore

$$\begin{aligned}\lim_{n\longrightarrow+\infty}\mathbf{E}(\ln h_J) &= \lim_{n\longrightarrow+\infty}\frac{2}{\sqrt{2\pi}}\int_0^{n^{\varepsilon}}\ln\left(\sqrt{\frac{\pi}{2}}t\right)e^{-t^2/2}\,dt\\ &= \frac{2}{\sqrt{2\pi}}\int_0^{+\infty}\ln\left(\sqrt{\frac{\pi}{2}}t\right)e^{-t^2/2}\,dt.\end{aligned}$$

The integral $\mathbf{E}(\ln^2 h_J)$ is treated similarly. □

We note that $|\mathbf{E}(\ln h_J)|,\ |\mathbf{E}(\ln^2 h_J)|<\infty$ for any $n>0$.

(8.2) Theorem. *We have*

$$\lim_{n\longrightarrow+\infty}\sup\{|\mathbf{E}(\ln h_K)| : K\subset\mathbb{R}^n \text{ is a zonoid and } \mathbf{E}(h_K)=1\}=C_2,\qquad(8.2.1)$$

where

$$C_2=-\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}\ln\left(\sqrt{\frac{\pi}{2}}t\right)e^{-t^2/2}\,dt\approx 0.4093900697$$

and each supremum is finite;

$$|\mathbf{E}(\ln^2 h_K)|=O(\ln^2 n),\qquad(8.2.2)$$

where $K\subset\mathbb{R}^n$ is a zonoid such that $\mathbf{E}(h_K)=1$.

Proof. Since every zonoid K can be approximated by zonotopes in the Hausdorff metric, it suffices to consider the case when K is a zonotope, that is, the Minkowski sum of finitely many segments J_k symmetric about the origin. Rescaling, if necessary, we may write $K=\alpha_1 J_1+\cdots+\alpha_m J_m$, where $\mathbf{E}(h_{J_k})=1$ and $\alpha_k\geq 0$ for $k=1,\ldots,m$. Then we must have $\alpha_1+\cdots+\alpha_m=1$. Since $\ln x$ is a concave function, we have

$$0\geq\mathbf{E}(\ln h_K)=\mathbf{E}(\ln(\alpha_1 h_{J_1}+\cdots+\alpha_m h_{J_m}))\geq\alpha_1\mathbf{E}(\ln h_{J_1})+\cdots+\alpha_m\mathbf{E}(\ln h_{J_m}).$$

So the supremum of $|\mathbf{E}(\ln h_K)|$ is attained on the segments in $\mathbb{R}^n$ and (8.2.1) follows from (8.1.2).

Part (8.1.3) of Lemma 8.1 implies that there exists an absolute constant C such that $\mathbf{E}(\ln^2 h_J)\leq C$ provided $J\subset\mathbb{R}^n$ is a segment in $\mathbb{R}^n$ such that $\mathbf{E}(h_J)=1$. Let $X=\{u\in S^{n-1} : h_K(u)\leq 1\}$ and $Y=S^{n-1}\backslash X$. Then

$$\mathbf{E}(\ln^2 h_K)=\int_X\ln^2 h_K(u)\,du+\int_Y\ln^2 h_K(u)\,du.$$

From (8.1.1) we gather that K must be contained in the ball of radius $O(\sqrt{n})$, so the second integral is $O(\ln^2 n)$. Let us estimate the first integral. For each $u \in X$ we have

$$0 \geq \ln h_K(u) \geq \alpha_1 \ln h_{J_1}(u) + \cdots + \alpha_m \ln h_{J_m}(u).$$

Therefore, for each $u \in X$, we have

$$\ln^2 h_K \leq \left(\sum_{k=1}^{m} \alpha_k \ln h_{J_k} \right)^2 = \sum_{1 \leq i,k \leq m} \alpha_k \alpha_i \ln h_{J_i} \ln h_{J_k}.$$

Now, by the Cauchy–Schwartz inequality

$$\left| \int_X |\ln h_{J_i}(u)| |\ln h_{J_k}(u)| \, du \right| \leq \left(\int_X \ln^2 h_{J_i}(u) \, du \right)^{1/2} \left(\int_X \ln^2 h_{J_k}(u) \, du \right)^{1/2} \leq C.$$

Therefore,

$$\int_X \ln^2 h_K(u) \, du \leq C \sum_{1 \leq j,k \leq m} \alpha_k \alpha_j = C$$

and the proof of (8.2.2) follows. □

V. D. Milman informed the author that the existence of an absolute constant C such that $|\mathbf{E}(\ln h_K)| \leq C$ provided $K \subset \mathbb{R}^n$ is *any* centrally symmetric convex body (not necessarily a zonoid) and $\mathbf{E}(h_K) = 1$ follows by a much more general inequality [14].

9. Computing the Mixed Volume

First, we present our main algorithm for computing the mixed volume of n ellipsoids in $\mathbb{R}^n$. We present it "coordinate free," that is, in operators rather than matrices. A coordinatization of the algorithm can be obtained in a similar way as in Algorithm 5.1 for mixed discriminants.

(9.1) Algorithm

Input. Positive definite $n \times n$ operators $M_1, \ldots, M_n$ of ellipsoids $E_{M_i} = \{x \in \mathbb{R}^n : \langle x, M_i x \rangle \leq 1\}$ in $\mathbb{R}^n$.

Output. A number γ approximating the mixed volume $V(E_{M_1}, \ldots, E_{M_n})$.

Algorithm

Step 0. Sample an orthonormal basis $(u_1, \ldots, u_n)$ in $\mathbb{R}^n$. Let $Q_i := M_i$ for $i = 1, \ldots, n$. Let $\gamma := 1$ and $s := 0$.

Comment. We store in s the number of iterations of Steps 1 and 2 of the algorithm and in γ the current value of the mixed volume.

Step 1. Let $k = n - s$ and let $s := s + 1$. Let $\gamma = (\det Q_1)^{-1/2}\gamma$. If $s = n$, compute $\gamma := \gamma v_n$, where v_n is the volume of the unit ball in $\mathbb{R}^n$, output γ and stop. Otherwise, compute a positive definite operator T such that $T^2 = Q_1$. Compute $R_i = T^{-1}Q_kT^{-1}$ for $i = 2, \ldots, k$.

Comment. On the sth iteration of this step we have k ellipsoids $E_{Q_1}, \ldots, E_{Q_k}$ in the k-dimensional space $(u_1, \ldots, u_{s-1})^\perp$. Formula (7.3.1) implies that $V(E_{Q_1}, \ldots, E_{Q_k}) = (\det Q_1)^{-1/2}V(B, E_{R_2}, \ldots, E_{R_k})$, where B is the unit ball.

Step 2. For $i = 1, \ldots, k-1$ let E_{Q_i} be the orthogonal projection of the ellipsoid $E_{R_{i+1}}$ onto the hyperplane $u_s^\perp$ in $(u_1, \ldots, u_{s-1})^\perp$. Go to Step 1.

Comment. On this step of the algorithm we approximate $(1/v_k)V(B, E_{R_2}, \ldots, E_{R_k})$ by $(1/v_{k-1})V(E_{R_2}|u_s^\perp, \ldots, E_{R_k}|u_s^\perp)$ for a random $u_s \in S^{n-1}$ (see (7.3.2)). To compute Q_i, we compute the inverse operator $(R_{i+1})^{-1}$, then let $Q_i = (P^*(R_{i+1})^{-1}P)^{-1}$, where $P\colon (u_1, \ldots, u_s)^\perp \subset (u_1, \ldots, u_{s-1})^\perp$ is the inclusion (see Lemma 7.2).

(9.2) Theorem. *For any given positive definite operators $M_1, \ldots, M_n$ the algorithm performs a polynomial in n number of operations* (*addition, subtraction, multiplication, division, and taking the square root of a nonnegative number*). *For any $\varepsilon > 0$ there is an $N(\varepsilon)$ such that for any $n \geq N(\varepsilon)$ the number γ produced by the algorithm with probability at least* 0.9 *satisfies the inequality*

$$c_\varepsilon^n V(E_{M_1}, \ldots, E_{M_n}) \leq \gamma \leq 20V(E_{M_1}, \ldots, E_{M_n}) \qquad \text{for} \quad c_\varepsilon = e^{-C_2-\varepsilon},$$

where C_2 is the absolute constant from Theorem 8.2.

Proof. The proof is completely analogous to the proof of Theorem 5.2. Instead of Theorem 2.4 we use Theorem 7.3 and instead of Corollary 3.4 and Theorem 3.5 for quadratic forms we use Theorem 8.2 for support functions of zonoids. We introduce functions

$$q_s(u_1, \ldots, u_s) = \frac{v_k(\det Q_1)^{-1/2}}{v_{k-1}V(E_{Q_1}, \ldots, E_{Q_k})}V(E_{R_2}|u_s^\perp, \ldots, E_{R_k}|u_s^\perp),$$

where $Q_1, \ldots, Q_k$ and $R_2, \ldots, R_k$ are the operators computed on the sth iteration of Step 1 and we agree that $q_n(u_1, \ldots, u_n) = 1$. Then we conclude that

$$\gamma = V(E_{M_1}, \ldots, E_{M_n})\prod_{s=1}^{n} q_s(u_1, \ldots, u_s).$$

Let us consider conditional expectations $\mathbf{E}_s$ (Section 4.1). Part (7.3.2) of Theorem 7.3 implies that $\mathbf{E}_s(q_s) = \mathbf{1}$ and hence by Lemma 4.2 we conclude that the expectation of γ on the orthogonal group O_n is the desired mixed volume $V(E_{M_1}, \ldots, E_{M_n})$. Part (7.3.1) of Theorem 7.3 implies that $q_s(u_1, \ldots, u_s)$ as a function in u_s for fixed $u_1, \ldots, u_{s-1}$ is the support function of a zonoid in $(u_1, \ldots, u_{s-1})^\perp$. We use Theorem 8.2 to show that

$$\frac{1}{n}\sum_{s=1}^{n} \|\mathbf{E}_s(\ln q_s)\| \leq C_2 + \frac{\varepsilon}{2}$$

for all sufficiently large n and that $\|\mathbf{E}_s(\ln^2 q_s)\| = O(\ln^2 n)$. Now, as in the proof of Theorem 5.2 we refer to Lemma 4.3 to show that

$$\nu\left\{(u_1,\ldots,u_n)\in O_n : \frac{1}{n}\sum_{s=1}^{n}\ln q_s(u_1,\ldots,u_n)\le -C_2-\varepsilon\right\}\longrightarrow 0 \qquad \text{as} \quad n\longrightarrow\infty$$

and we complete the proof as in Theorem 5.2. □

So any approximation constant

$$c_\varepsilon < \exp\left\{\frac{2}{\sqrt{2\pi}}\int_0^\infty \ln\left(\sqrt{\frac{\pi}{2}}t\right)e^{-t^2/2}\,dt\right\}\approx 0.6640551540$$

will work for a sufficiently large n.

If the number of pairwise different ellipsoids is fixed, we can achieve a $2^{O(n)}$ approximation by a deterministic polynomial time algorithm.

(9.3) Lemma. *Let us fix k. Then there exists an algorithm, which for any given n matrices $Q_1,\ldots,Q_1,Q_2,\ldots,Q_2,\ldots,Q_k,\ldots,Q_k$ with only k pairwise different computes the mixed discriminant $D(Q_1,\ldots,Q_n)$. The algorithm uses a polynomial in n number of arithmetic operations.*

Proof. We use representation (2.2.1) for the mixed discriminant. Since the number of pairwise different operators is fixed, the sum (2.2.1) can be rewritten as a sum of $n^{O(k)}$ determinants. For $i\le k$ let α_i be the number of copies of Q_i. Then

$$\begin{aligned} n!\,D(Q_1,\ldots,Q_1,Q_2,\ldots,Q_2,\ldots,Q_k,\ldots,Q_k) \\ = \sum_{s=1}^{n}(-1)^{n-s}\sum_{\beta_1+\cdots+\beta_k=s}\binom{\alpha_1}{\beta_1}\cdots\binom{\alpha_k}{\beta_k}\det(\beta_1Q_1+\cdots+\beta_kQ_k). \end{aligned}$$

Since the determinant of an $n\times n$ matrix can be computed using $O(n^3)$ arithmetic operations, the formula gives rise to an algorithm of polynomial complexity. □

(9.4) Corollary. *Let us fix k. Then there exists a polynomial time algorithm that for any given n positive definite matrices $Q_1,\ldots,Q_1,Q_2,\ldots,Q_2,\ldots,Q_k,\ldots,Q_k$ with at most k pairwise different computes a number δ such that*

$$\begin{aligned} (\sqrt{3})^{-n+1}V(E_{Q_1},\ldots,E_{Q_1},\ldots,E_{Q_k},\ldots,E_{Q_k}) \\ \le \delta \\ \le V(E_{Q_1},\ldots,E_{Q_1},\ldots,E_{Q_k},\ldots,E_{Q_k}). \end{aligned}$$

Proof. Follows by (1.4.2) and Lemma 9.3. □

Note, that $1/\sqrt{3}\approx 0.5773502693$ so we are getting a worse approximation than we could have gotten using randomized Algorithm 9.1.

(9.5) Mixed Volumes of General Convex Bodies. As is known, for any convex body $K \subset \mathbb{R}^n$ there exists an ellipsoid E such that (after translating its center to the origin) we have $E \subset K \subset nE$ (see, for example, [10]). There are classes of convex bodies where an approximating ellipsoid such that $E \subset K \subset n^{O(1)}E$ can be constructed in polynomial time in the real model of computation. This is the case, for example, when K is a polytope given by a list of its vertices (see [12] and [13]). For this class of convex bodies we can approximate $V(K_1, \ldots, K_n)$ within a factor $n^{O(n)}$ in the real RAM model. We also note that if K is given by a "well-guaranteed" oracle (see [10]) then there is a polynomial time algorithm in the bit model that computes an ellipsoid E such that $E \subset K \subset n\sqrt{n+1}E$. Applying a bit version of Algorithm 9.1 (which was not discussed here) we would get a randomized polynomial time $n^{O(n)}$ approximation algorithm in the bit model.

Acknowledgments

This paper was inspired by the papers [9] and [4]. I attempted to answer some of the questions asked there. I am grateful to E. Gluskin, E. Lutwak, and V. D. Milman for many helpful discussions during the "Sharp Inequalities in Harmonic Analysis and Convex Geometry" workshop hosted by MSRI, Berkeley, CA.

References

1. A. Aho, J. Hopcroft, and J. Ullman, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, MA, 1974.
2. A. D. Aleksandrov, On the theory of mixed volumes of convex bodies, IV, Mixed discriminants and mixed volumes (in Russian), *Mat. Sb.* (*N.S.*) **3** (1938), 227–251.
3. A. I. Barvinok, Two algorithmic results for the Traveling Salesman Problem, *Math. Oper. Res.* **21** (1996), 65–84.
4. M. Dyer, P. Gritzmann, and A. Hufnagel, On the complexity of computing mixed volumes, to appear.
5. G. P. Egorychev, The solution of van der Waerden's problem for permanents, *Adv. in Math.* **42** (1981), 299–305.
6. A. Frieze and M. Jerrum, An analysis of a Monte Carlo algorithm for estimating the permanent, *Combinatorica*, **15** (1995), 67–83.
7. F. R. Gantmakher, *The Theory of Matrices*, Chelsea, New York, 1960.
8. D. Yu. Grigoriev and M. Karpinsky, The matching problem for bipartite graphs with polynomially bounded permanents is in NC, *Proc. Twenty-Eighth Annual IEEE Symp. Foundations of Computer Science*, IEEE Computer Society Press, Washington, DC, 1987, pp. 162–172.
9. P. Gritzmann and V. Klee, On the complexity of some basic problems in computational convexity: II. Volume and mixed volumes, In: *Polytopes: Abstract, Convex, and Computational* (T. Bisztriczky, P. McMullen, R. Schneider, and A. Ivić Weiss, eds.), Proceedings of the NATO Advanced Study Institute, Scarborough, Ontario, Canada, August 20–September 3, 1993, 1994, Kluwer Academic, Amsterdam, pp. 373–466.
10. M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.
11. M. Jerrum and A. Sinclair, Approximating the permanent, *SIAM J. Comput.* **18** (1989), 1149–1178.
12. L. Khachiyan, Rounding of polytopes in the real number model of computation, *Math. Oper. Res.* **21** (1996), 307–320.
13. L. Khachiyan and M. Todd, On the complexity of approximating the maximal inscribed ellipsoid for a polytope, *Math. Programming*, **61** (1993), 137–159.
14. R. Latala, On the equivalence between geometric and arithmetic means for logconcave measures, Preprint.

15. K. Leichtweiß, Convexity and Differential Geometry, In: *Handbook of Convex Geometry*, vol. B, Chapter 4.1 (P. M. Gruber and J. M. Wills, eds.), North-Holland, Amsterdam, 1993, pp. 1045–1080.
16. E. Lutwak, Mixed projection inequalities, *Trans. Amer. Math. Soc.* **287** (1985), 91–105.
17. V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces*. With an Appendix by M. Gromov, "Isoperimetric Inequalities in Riemannian Manifolds," Lecture Notes in Mathematics, vol. 1200, Springer-Verlag, Berlin, 1986.
18. C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization*: *Algorithms and Complexity*, Prentice Hall, Englewood Cliffs, NJ, 1982.
19. L. E. Rasmussen, Approximating the permanent: A simple approach, *Random Structures and Algorithms* **5** (1994), 349–361.
20. L. A. Santalo, *Integral Geometry and Geometric Probability*, Addison-Wesley, Reading, MA, 1976.
21. R. Schneider, *Convex Bodies*: *The Brunn–Minkowski Theory*, Encyclopedia of Mathematics and Its Applications, vol. 44, Cambridge University Press, New York, 1993.
22. R. Schneider and J. A. Wieacker, Integral geometry, In: *Handbook of Convex Geometry*, vol. B, Chapter 5.1 (P. M. Gruber and J. M. Wills, eds.), North-Holland, Amsterdam, 1993, pp. 1351–1390.

Received July 10, 1995, *and in revised form May* 20, 1996.