CONCENTRATION OF THE MIXED DISCRIMINANT OF WELL-CONDITIONED MATRICES

Alexander Barvinok

Abstract. We call an n-tuple $Q_1, \ldots, Q_n$ of positive definite $n \times n$ real matrices α-conditioned for some $\alpha \geq 1$ if for the corresponding quadratic forms $q_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ we have $q_i(x) \leq \alpha q_i(y)$ for any two vectors $x, y \in \mathbb{R}^n$ of Euclidean unit length and $q_i(x) \leq \alpha q_j(x)$ for all $1 \leq i, j \leq n$ and all $x \in \mathbb{R}^n$. An n-tuple is called doubly stochastic if the sum of Q_i is the identity matrix and the trace of each Q_i is 1. We prove that for any fixed $\alpha \geq 1$ the mixed discriminant of an α-conditioned doubly stochastic n-tuple is $n^{O(1)} e^{-n}$. As a corollary, for any $\alpha \geq 1$ fixed in advance, we obtain a polynomial time algorithm approximating the mixed discriminant of an α-conditioned n-tuple within a polynomial in n factor.

1. Introduction and main results

(1.1) Mixed discriminants. Let $Q_1, \ldots, Q_n$ be $n \times n$ real symmetric matrices. The function $\det(t_1 Q_1 + \ldots + t_n Q_n)$, where $t_1, \ldots, t_n$ are real variables, is a homogeneous polynomial of degree n in $t_1, \ldots, t_n$ and its coefficient

$$D(Q_1, \ldots, Q_n) = \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \det(t_1 Q_1 + \ldots + t_n Q_n) \tag{1.1.1}$$

is called the *mixed discriminant* of $Q_1, \ldots, Q_n$ (sometimes, the normalizing factor of $1/n!$ is used). Mixed discriminants were introduced by A.D. Alexandrov in his work on mixed volumes [Al38], see also [Le93]. They also have some interesting combinatorial applications, see Chapter V of [BR97].

Mixed discriminants generalize permanents. If the matrices $Q_1, \ldots, Q_n$ are diagonal, so that

$$Q_i = \operatorname{diag}(a_{i1}, \ldots, a_{in}) \quad \text{for} \quad i = 1, \ldots, n,$$

then

$$D(Q_1, \ldots, Q_n) = \operatorname{per} A \quad \text{where} \quad A = (a_{ij}) \tag{1.1.2}$$

1991 *Mathematics Subject Classification.* 15A15, 15A45, 90C25, 68Q25.
Key words and phrases. mixed discriminant, scaling, algorithm, concentration.
This research was partially supported by NSF Grant DMS 1361541.

Typeset by $\mathcal{AMS}$-TEX

and

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

is the *permanent* of an $n \times n$ matrix A. Here the i-th row of A is the diagonal of Q_i and S_n is the symmetric group of all $n!$ permutations of the set $\{1, \ldots, n\}$.

(1.2) Doubly stochastic n-tuples. If $Q_1, \ldots, Q_n$ are positive semidefinite matrices then $D(Q_1, \ldots, Q_n) \geq 0$, see [Le93]. We say that the n-tuple $(Q_1, \ldots, Q_n)$ is *doubly stochastic* if $Q_1, \ldots, Q_n$ are positive semidefinite,

$$Q_1 + \ldots + Q_n = I \quad \text{and} \quad \operatorname{tr} Q_1 = \ldots = \operatorname{tr} Q_n = 1,$$

where I is the $n \times n$ identity matrix and $\operatorname{tr} Q$ is the trace of Q. We note that if $Q_1, \ldots, Q_n$ are diagonal then the n-tuple $(Q_1, \ldots, Q_n)$ is doubly stochastic if and only if the matrix A in (1.1.2) is doubly stochastic, that is, non-negative and has row and column sums 1.

In [Ba89] Bapat conjectured what should be the mixed discriminant version of the van der Waerden inequality for permanents: if $(Q_1, \ldots, Q_n)$ is a doubly stochastic n-tuple then

$$D(Q_1, \ldots, Q_n) \geq \frac{n!}{n^n} \tag{1.2.1}$$

where equality holds if and only if

$$Q_1 = \ldots = Q_n = \frac{1}{n} I.$$

The conjecture was proved by Gurvits [Gu06], see also [Gu08] for a more general result with a simpler proof.

In this paper, we prove that $D(Q_1, \ldots, Q_n)$ remains close to $n!/n^n \approx e^{-n}$ if the n-tuple $(Q_1, \ldots, Q_n)$ is doubly stochastic and well-conditioned.

(1.3) α-conditioned n-tuples. For a symmetric matrix Q, let $\lambda_{\min}(Q)$ denote the minimum eigenvalue of Q and let $\lambda_{\max}(Q)$ denote the maximum eigenvalue of Q. We say that a positive definite matrix Q is *α-conditioned* for some $\alpha \geq 1$ if

$$\lambda_{\max}(Q) \leq \alpha \lambda_{\min}(Q).$$

Equivalently, let $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ be the corresponding quadratic form defined by

$$q(x) = \langle Qx, x \rangle \quad \text{for} \quad x \in \mathbb{R}^n,$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^n$. Then Q is α-conditioned if

$$q(x) \leq \alpha q(y) \quad \text{for all} \quad x, y \in \mathbb{R}^n \quad \text{such that} \quad \|x\| = \|y\| = 1,$$

where $\| \cdot \|$ is the standard Euclidean norm in $\mathbb{R}^n$.

We say that an n-tuple $(Q_1, \ldots, Q_n)$ is *α-conditioned* if each matrix Q_i is α-conditioned and

$$q_i(x) \leq \alpha q_j(x) \quad \text{for all} \quad 1 \leq i, j \leq n \quad \text{and all} \quad x \in \mathbb{R}^n,$$

where $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ are the corresponding quadratic forms.

The main result of this paper is the following inequality.

(1.4) Theorem. *Let $(Q_1, \ldots, Q_n)$ be an α-conditioned doubly stochastic n-tuple of positive definite $n \times n$ matrices. Then*

$$D\left(Q_1, \ldots Q_n\right) \ \leq \ n^{\alpha^2} e^{-(n-1)}.$$

Combining the bound of Theorem 1.4 with (1.2.1), we conclude that for any $\alpha \geq 1$, fixed in advance, the mixed discriminant of an α-conditioned doubly stochastic n-tuple is within a polynomial in n factor of e^{-n}. If we allow α to vary with n then as long as $\alpha \ll \sqrt{\frac{n}{\ln n}}$, the logarithmic order of the mixed discriminant is captured by e^{-n}.

The estimate of Theorem 1.4 is unlikely to be precise. It can be considered as a (weak) mixed discriminant extension of the Bregman - Minc inequality for permanents (we discuss the connection in Section 1.7).

(1.5) Scaling. We say that an n-tuple $(P_1, \ldots, P_n)$ of $n \times n$ positive definite matrices is obtained from an n-tuple $(Q_1, \ldots, Q_n)$ of $n \times n$ positive definite matrices by *scaling* if for some invertible $n \times n$ matrix T and real $\tau_1, \ldots, \tau_n > 0$, we have

$$P_i = \tau_i T^* Q_i T \quad \text{for} \quad i = 1, \ldots, n, \tag{1.5.1}$$

where T^* is the transpose of T. As easily follows from (1.1.1),

$$D\left(P_1, \ldots, P_n\right) = (\det T)^2 \left(\prod_{i=1}^n \tau_i\right) D\left(Q_1, \ldots, Q_n\right), \tag{1.5.2}$$

provided (1.5.1) holds.

This notion of scaling extends the notion of scaling for positive matrices by Sinkhorn [Si64] to n-tuples of positive definite matrices. Gurvits and Samorodnitsky proved in [GS02] that any n-tuple of $n \times n$ positive definite matrices can be obtained by scaling from a doubly stochastic n-tuple, and, moreover, this can be achieved in polynomial time, as it reduces to solving a convex optimization problem (the gist of their algorithm is given by Theorem 2.1 below). More generally, Gurvits and Samorodnitsky discuss when an n-tuple of positive *semi*definite matrices can be scaled to a doubly stochastic n-tuple. As is discussed in [GS02], the inequality (1.2.1), together with the scaling algorithm, the identity (1.5.2) and the inequality

$$D\left(Q_1, \ldots, Q_n\right) \ \leq \ 1$$

for doubly stochastic n-tuples $(Q_1, \ldots, Q_n)$, allow one to estimate within a factor of $n!/n^n \approx e^{-n}$ the mixed discriminant of any given n-tuple of $n \times n$ positive semidefinite matrices in polynomial time.

In this paper, we prove that if a doubly stochastic n-tuple $(P_1, \ldots, P_n)$ is obtained from an α-conditioned n-tuple of positive definite matrices then the n-tuple $(P_1, \ldots, P_n)$ is α^2-conditioned (see Lemma 2.4 below). We also prove the following strengthening of Theorem 1.4.

(1.6) Theorem. *Suppose that* $(Q_1, \ldots, Q_n)$ *is an* α*-conditioned* n*-tuple of* $n \times n$ *positive definite matrices and suppose that* $(P_1, \ldots, P_n)$ *is a doubly stochastic* n*-tuple of positive definite matrices obtained from* $(Q_1, \ldots, Q_n)$ *by scaling. Then*

$$D\left(P_1, \ldots, P_n\right) \ \leq \ n^{\alpha^2} e^{-(n-1)}.$$

Together with the scaling algorithm of [GS02] and the inequality (1.2.1), Theorem 1.6 allows us to approximate in polynomial time the mixed discriminant $D\left(Q_1, \ldots, Q_n\right)$ of an α-conditioned n-tuple $(Q_1, \ldots, Q_n)$ within a factor of n^{α^2}. Note that the value of $D\left(Q_1, \ldots, Q_n\right)$ may vary within a factor of α^n.

(1.7) Connections to the Bregman - Minc inequality. The following inequality for permanents of 0-1 matrices was conjectured by Minc [Mi63] and proved by Bregman [Br73], see also [Sc78] for a much simplified proof: if A is an $n \times n$ matrix with 0-1 entries and row sums $r_1, \ldots, r_n$, then

$$\operatorname{per} A \ \leq \ \prod_{i=1}^{n} \left(r_i!\right)^{1/r_i}. \tag{1.7.1}$$

The author learned from A. Samorodnitsky [Sa00] the following restatement of (1.7.1), see also [So03]. Suppose that $B=(b_{ij})$ is an $n \times n$ stochastic matrix (that is, a non-negative matrix with row sums 1) such that

$$0 \ \leq \ b_{ij} \ \leq \frac{1}{r_i} \quad \text{for all} \quad i, j \tag{1.7.2}$$

and some positive integers $r_1, \ldots, r_n$. Then

$$\operatorname{per} B \ \leq \ \prod_{i=1}^{n} \frac{\left(r_i!\right)^{1/r_i}}{r_i}. \tag{1.7.3}$$

Indeed, the function $B \longmapsto \operatorname{per} B$ is linear in each row and hence its maximum value on the polyhedron of stochastic matrices satisfying (1.7.2) is attained at a vertex of the polyhedron, that is, where $b_{ij} \in \{0, 1/r_i\}$ for all i, j. Multiplying the i-th row of B by r_i, we obtain a 0-1 matrix A with row sums $r_1, \ldots, r_n$ and hence (1.7.3) follows by (1.7.1).

Suppose now that B is a doubly stochastic matrix whose entries do not exceed α/n for some $\alpha \geq 1$. Combining (1.7.3) with the van der Waerden lower bound, we obtain that

$$\operatorname{per} B = e^{-n} n^{O(\alpha)}. \tag{1.7.4}$$

Ideally, we would like to obtain a similar to (1.7.4) estimate for the mixed discriminants $D\left(Q_1, \ldots, Q_n\right)$ of doubly stochastic n-tuples of positive semidefinite matrices satisfying

$$\lambda_{\max}\left(Q_i\right) \ \leq \ \frac{\alpha}{n} \quad \text{for} \quad i = 1, \ldots, n. \tag{1.7.5}$$

In Theorem 1.4 such an estimate is obtained under a stronger assumption that the n-tuple $(Q_1, \ldots, Q_n)$ in addition to being doubly stochastic is also α-conditioned. This of course implies (1.7.5) but it also prohibits Q_i from having small (in particular, 0) eigenvalues. The question whether a similar to Theorem 1.4 bound can be proven under the the weaker assumption of (1.7.5) together with the assumption that $(Q_1, \ldots, Q_n)$ is doubly stochastic remains open.

In Section 2 we collect various preliminaries and in Section 3 we prove Theorems 1.4 and 1.6.

2. Preliminaries

First, we restate a result of Gurvits and Samorodnitsky [GS02] that is at the heart of their algorithm to estimate the mixed discriminant. We state it in the particular case of positive definite matrices.

(2.1) Theorem. *Let $Q_1, \ldots, Q_n$ be $n \times n$ positive definite matrices, let $H \subset \mathbb{R}^n$ be the hyperplane,*

$$H = \left\{ (x_1, \ldots, x_n) : \quad \sum_{i=1}^n x_i = 0 \right\}$$

and let $f : H \longrightarrow \mathbb{R}$ be the function

$$f(x_1, \ldots, x_n) = \ln \det \left(\sum_{i=1}^n e^{x_i} Q_i \right).$$

Then f is strictly convex on H and attains its minimum on H at a unique point $(\xi_1, \ldots, \xi_n)$. Let S be an $n \times n$, necessarily invertible, matrix such that

$$S^* S = \sum_{i=1}^n e^{\xi_i} Q_i \tag{2.1.1}$$

(such a matrix exists since the matrix in the right hand side of (2.1.1) is positive definite). Let

$$\tau_i = e^{\xi_i} \quad \text{for} \quad i = 1, \ldots, n,$$

let $T = S^{-1}$ and let

$$B_i = \tau_i T^* Q_i T \quad \text{for} \quad i = 1, \ldots, n.$$

Then $(B_1, \ldots, B_n)$ is a doubly stochastic n-tuple of positive definite matrices.

We will need the following simple observation regarding matrices $B_1, \ldots, B_n$ constructed in Theorem 2.1.

(2.2) Lemma. *Suppose that for the matrices $Q_1, \ldots, Q_n$ in Theorem 2.1, we have*

$$\sum_{i=1}^{n} \operatorname{tr} Q_i = n.$$

Then, for the matrices $B_1, \ldots, B_n$ constructed in Theorem 2.1, we have

$$D(B_1, \ldots, B_n) \ \geq \ D(Q_1, \ldots, Q_n).$$

Proof. We have

$$D(B_1, \ldots, B_n) = (\det T)^2 \left(\prod_{i=1}^{n} \tau_i \right) D(Q_1, \ldots, Q_n). \tag{2.2.1}$$

Now,

$$\prod_{i=1}^{n} \tau_i = \exp\left\{ \sum_{i=1}^{n} \xi_i \right\} = 1 \tag{2.2.2}$$

and

$$(\det T)^2 = \left(\det \sum_{i=1}^{n} e^{\xi_i} Q_i \right)^{-1} = \exp\left\{ -f(\xi_1, \ldots, \xi_n) \right\}. \tag{2.2.3}$$

Since $(\xi_1, \ldots, \xi_n)$ is the minimum point of f on H, we have

$$f(\xi_1, \ldots, \xi_n) \ \leq \ f(0, \ldots, 0) = \ln \det Q \quad \text{where} \quad Q = \sum_{i=1}^{n} Q_i. \tag{2.2.4}$$

We observe that Q is a positive definite matrix with eigenvalues, say, $\lambda_1, \ldots, \lambda_n$ such that

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr} Q = \sum_{i=1}^{n} \operatorname{tr} Q_i = n \quad \text{and} \quad \lambda_1, \ldots, \lambda_n > 0.$$

Applying the arithmetic - geometric mean inequality, we obtain

$$\det Q = \lambda_1 \cdots \lambda_n \ \leq \ \left(\frac{\lambda_1 + \ldots + \lambda_n}{n} \right)^n = 1. \tag{2.2.5}$$

Combining (2.2.1) – (2.2.5), we complete the proof. □

(2.3) From symmetric matrices to quadratic forms. As in Section 1.3, with an $n \times n$ symmetric matrix Q we associate the quadratic form $q : \mathbb{R}^n \longrightarrow \mathbb{R}$. We define the eigenvalues, the trace, and the determinant of q as those of Q. Consequently, we define the mixed discriminant $D(q_1, \ldots, q_n)$ of quadratic forms $q_1, \ldots, q_n$. An n-tuple of positive semidefinite quadratic forms $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ is *doubly stochastic* if

$$\sum_{i=1}^{n} q_i(x) = \|x\|^2 \quad \text{for all} \quad x \in \mathbb{R}^n \quad \text{and} \quad \operatorname{tr} q_1 = \ldots = \operatorname{tr} q_n = 1.$$

An n-tuple of quadratic forms $p_1, \ldots, p_n : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is obtained from an n-tuple $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ by *scaling* if for some invertible linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and real $\tau_1, \ldots, \tau_n > 0$ we have

$$p_i(x) = \tau_i q_i(Tx) \quad \text{for all} \quad x \in \mathbb{R}^n \quad \text{and all} \quad i = 1, \ldots, n.$$

One advantage of working with quadratic forms as opposed to matrices is that it is particularly easy to define the restriction of a quadratic form onto a subspace. We will use the following construction: suppose that $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ are positive definite quadratic forms and let $L \subset \mathbb{R}^n$ be an m-dimensional subspace for some $1 \leq m \leq n$. Then L inherits Euclidean structure from $\mathbb{R}^n$ and we can consider the *restrictions* $\widehat{q}_1, \ldots, \widehat{q}_n : L \longrightarrow \mathbb{R}$ of $q_1, \ldots, q_n$ onto L. Thus we can define the mixed discriminant $D(\widehat{q}_1, \ldots, \widehat{q}_m)$. Note that by choosing an orthonormal basis in L, we can associate $m \times m$ symmetric matrices $\widehat{Q}_1, \ldots, \widehat{Q}_m$ with $\widehat{q}_1, \ldots, \widehat{q}_m$. A different choice of an orthonormal basis results in the transformation $\widehat{Q}_i \longmapsto U^* \widehat{Q}_i U$ for some $m \times m$ orthogonal matrix U and $i = 1, \ldots, m$, which does not change the mixed discriminant $D\left(\widehat{Q}_1, \ldots, \widehat{Q}_m\right)$.

(2.4) Lemma. *Let $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ be an α-conditioned n-tuple of positive definite quadratic forms. Let $L \subset \mathbb{R}^n$ be an m-dimensional subspace, where $1 \leq m \leq n$, let $T : L \longrightarrow \mathbb{R}^n$ be a linear transformation such that $\ker T = \{0\}$ and let $\tau_1, \ldots, \tau_m > 0$ be reals. Let us define quadratic forms $p_1, \ldots, p_m : L \longrightarrow \mathbb{R}$ by*

$$p_i(x) = \tau_i q_i(Tx) \quad \textit{for} \quad x \in L \quad \textit{and} \quad i = 1, \ldots, m.$$

Suppose that

$$\sum_{i=1}^{m} p_i(x) = \|x\|^2 \quad \textit{for all} \quad x \in L \quad \textit{and} \quad \operatorname{tr} p_i = 1 \quad \textit{for} \quad i = 1, \ldots, m.$$

Then the m-tuple of quadratic forms $p_1, \ldots, p_m$ is α^2-conditioned.

This version of Lemma 2.4 and the following proof was suggested by the anonymous referee. It replaces an earlier version with a weaker bound of α^4 instead of α^2.

Proof of Lemma 2.4. Let us define a quadratic form $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$q(x) = \sum_{i=1}^{m} \tau_i q_i(x) \quad \text{for all} \quad x \in \mathbb{R}^n.$$

Then $q(x)$ is α-conditioned and for each $x, y \in L$ such that $\|x\| = \|y\| = 1$ we have

$$1 = q(Tx) \ \geq \ \lambda_{\min}(q)\|Tx\|^2 \quad \text{and} \quad 1 = q(Ty) \ \leq \ \lambda_{\max}(q)\|Ty\|^2,$$

from which it follows that

$$\|Tx\|^2 \ \leq \ \frac{\lambda_{\max}(q)}{\lambda_{\min}(q)}\|Ty\|^2$$

and hence

$$\|Tx\|^2 \ \leq \ \alpha\|Ty\|^2 \quad \text{for all} \quad x, y \in L \quad \text{such that} \quad \|x\| = \|y\| = 1. \tag{2.4.1}$$

Applying (2.4.1) and using that the form q_i is α-conditioned, we obtain

$$\begin{aligned} p_i(x) =& \tau_i q_i(Tx) \ \leq \ \tau_i \left(\lambda_{\max} q_i\right) \|Tx\|^2 \ \leq \ \alpha\tau_i \left(\lambda_{\max} q_i\right) \|Ty\|^2 \\ \leq& \alpha^2 \tau_i \left(\lambda_{\min} q_i\right) \|Ty\|^2 \ \leq \ \alpha^2 \tau_i q_i \left(Ty\right) \\ =& \alpha^2 p_i(y) \quad \text{for all} \quad x, y \in L \quad \text{such that} \quad \|x\| = \|y\| = 1, \end{aligned} \tag{2.4.2}$$

and hence each form p_i is α^2-conditioned.

Let us define quadratic forms $r_i : L \longrightarrow \mathbb{R}$, $i = 1, \ldots, m$, by

$$r_i(x) = q_i(Tx) \quad \text{for} \quad x \in L \quad \text{and} \quad i = 1, \ldots, m.$$

Then

$$r_i(x) \ \leq \ \alpha r_j(x) \quad \text{for all} \quad 1 \leq i, j \leq m \quad \text{and all} \quad x \in L.$$

Therefore,

$$\operatorname{tr} r_i \ \leq \ \alpha \operatorname{tr} r_j \quad \text{for all} \quad 1 \leq i, j \leq m.$$

Since $1 = \operatorname{tr} p_i = \tau_i \operatorname{tr} r_i$, we conclude that $\tau_i = 1/\operatorname{tr} r_i$ and, therefore,

$$\tau_i \ \leq \ \alpha\tau_j \quad \text{for all} \quad 1 \leq i, j \leq m. \tag{2.4.3}$$

Applying (2.4.3) and using that the n-tuple $q_1, \ldots, q_n$ is α-conditioned, we obtain

$$\begin{aligned} p_i(x) =& \tau_i q_i(Tx) \ \leq \ \alpha\tau_j q_i(Tx) \ \leq \ \alpha^2 \tau_j q_j(Tx) \\ =& \alpha^2 p_j(x) \quad \text{for all} \quad x \in L. \end{aligned} \tag{2.4.4}$$

Combining (2.4.2) and (2.4.4), we conclude that the m-tuple $p_1, \ldots, p_m$ is α^2-conditioned. □

(2.5) Lemma. *Let $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ be positive semidefinite quadratic forms and suppose that*

$$q_n(x) = \langle u, x \rangle^2,$$

where $u \in \mathbb{R}^n$ and $\|u\| = 1$. Let $H = u^\perp$ be the orthogonal complement to u. Let $\widehat{q}_1, \ldots, \widehat{q}_{n-1} : H \longrightarrow \mathbb{R}$ be the restrictions of $q_1, \ldots, q_{n-1}$ onto H. Then

$$D(q_1, \ldots, q_n) = D\left(\widehat{q}_1, \ldots, \widehat{q}_{n-1}\right).$$

Proof. Let us choose an orthonormal basis of $\mathbb{R}^n$ for which u is the last basis vector and let $Q_1, \ldots, Q_n$ be the matrices of the forms $q_1, \ldots, q_n$ in that basis. Then the only non-zero entry of Q_n is 1 in the lower right corner. Let $\widehat{Q}_1, \ldots, \widehat{Q}_{n-1}$ be the upper left $(n-1) \times (n-1)$ submatrices of $Q_1, \ldots, Q_{n-1}$. Then

$$\det\left(t_1 Q_1 + \ldots + t_n Q_n\right) = t_n \det\left(t_1 \widehat{Q}_1 + \ldots + t_{n-1} \widehat{Q}_{n-1}\right)$$

and hence by (1.1.1) we have

$$D\left(Q_1, \ldots, Q_n\right) = D\left(\widehat{Q}_1, \ldots, \widehat{Q}_{n-1}\right).$$

On the other hand, $\widehat{Q}_1, \ldots, \widehat{Q}_{n-1}$ are the matrices of $\widehat{q}_1, \ldots, \widehat{q}_{n-1}$. □

Finally, the last lemma before we embark on the proof of Theorems 1.4 and 1.6.

(2.6) Lemma. *Let $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ be an α-conditioned quadratic form such that $\operatorname{tr} q = 1$. Let $H \subset \mathbb{R}^n$ be a hyperplane and let $\widehat{q}$ be the restriction of q onto H. Then*

$$\operatorname{tr} \widehat{q} \ \geq \ 1 - \frac{\alpha}{n}.$$

Proof. Let

$$0 \ < \ \lambda_1 \ \leq \ \ldots \ \leq \ \lambda_n$$

be the eigenvalues of q. Then

$$\sum_{i=1}^n \lambda_i = 1 \quad \text{and} \quad \lambda_n \ \leq \ \alpha \lambda_1,$$

from which it follows that

$$\lambda_n \ \leq \ \frac{\alpha}{n}.$$

As is known, the eigenvalues of $\widehat{q}$ interlace the eigenvalues of q, see, for example, Section 1.3 of [Ta12], so for the eigenvalues $\mu_1, \ldots, \mu_{n-1}$ of $\widehat{q}$ we have

$$\lambda_1 \ \leq \ \mu_1 \ \leq \ \lambda_2 \ \leq \ \ldots \ \leq \ \lambda_{n-1} \ \leq \ \mu_{n-1} \ \leq \ \lambda_n.$$

Therefore,

$$\operatorname{tr} \widehat{q} = \sum_{i=1}^{n-1} \mu_i \ \geq \ \sum_{i=1}^{n-1} \lambda_i \ \geq \ 1 - \frac{\alpha}{n}.$$

□

3. Proof of Theorem 1.4 and Theorem 1.6

Clearly, Theorem 1.6 implies Theorem 1.4, so it suffices to prove the former.

(3.1) Proof of Theorem 1.6. As in Section 2.3, we associate quadratic forms with matrices. We prove the following statement by induction on $m = 1, \ldots, n$.

Statement: Let $q_1, \ldots, q_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ be an α-conditioned n-tuple of positive definite quadratic forms. Let $L \subset \mathbb{R}^n$ be an m-dimensional subspace, $1 \leq m \leq n$, let $T : L \longrightarrow \mathbb{R}^n$ be a linear transformation such that $\ker T = \{0\}$ and let $\tau_1, \ldots, \tau_m > 0$ be reals. Let us define quadratic forms $p_i : L \longrightarrow \mathbb{R}$, $i = 1, \ldots, m$, by

$$p_i(x) = \tau_i q_i(Tx) \quad \text{for} \quad x \in L \quad \text{and} \quad i = 1, \ldots, m$$

and suppose that

$$\sum_{i=1}^{m} p_i(x) = \|x\|^2 \quad \text{for all} \quad x \in L \quad \text{and} \quad \operatorname{tr} p_i = 1 \quad \text{for} \quad i = 1, \ldots, m.$$

Then

$$D(p_1, \ldots, p_m) \ \leq \ \exp\left(-(m-1) + \alpha^2 \sum_{k=2}^{m} \frac{1}{k}\right). \tag{3.1.1}$$

In the case of $m = n$, we get the desired result.

The statement holds if $m = 1$ since in that case $D(p_1) = \det p_1 = 1$.

Suppose that $m > 1$. Let $L \subset \mathbb{R}^n$ be an m-dimensional subspace and let the linear transformation T, numbers τ_i and the forms p_i for $i = 1, \ldots, m$ be as above. By Lemma 2.4, the m-tuple $p_1, \ldots, p_m$ is α^2-conditioned. We write the spectral decomposition

$$p_m(x) = \sum_{j=1}^{m} \lambda_j \langle u_j, x \rangle^2,$$

where $u_1, \ldots, u_m \in L$ are the unit eigenvectors of p_m and $\lambda_1, \ldots, \lambda_m > 0$ are the corresponding eigenvalues of p_m. Since $\operatorname{tr} p_m = 1$, we have $\lambda_1 + \ldots + \lambda_m = 1$. Let $L_j = u_j^{\perp}$, $L_j \subset L$, be the orthogonal complement of u_j in L. Let

$$\widehat{p}_{ij} : \ L_j \longrightarrow \mathbb{R} \quad \text{for} \quad i = 1, \ldots, m \quad \text{and} \quad j = 1, \ldots, m$$

be the restriction of p_i onto L_j.

Using Lemma 2.5, we write

$$\begin{aligned} D(p_1, \dots, p_m) &= \sum_{j=1}^{m} \lambda_j D\left(p_1, \dots, p_{m-1}, \langle u_j, x\rangle^2\right) \\ &= \sum_{j=1}^{m} \lambda_j D\left(\widehat{p}_{1j}, \dots, \widehat{p}_{(m-1)j}\right) \quad \text{where} \\ &\sum_{j=1}^{m} \lambda_j = 1 \quad \text{and} \quad \lambda_j > 0 \quad \text{for} \quad j = 1, \dots, m. \end{aligned} \tag{3.1.2}$$

Let

$$\sigma_j = \operatorname{tr} \widehat{p}_{1j} + \ldots + \operatorname{tr} \widehat{p}_{(m-1)j} \quad \text{for} \quad j = 1, \dots, m.$$

Since

$$\sum_{i=1}^{m-1} \widehat{p}_{ij}(x) = \|x\|^2 - \widehat{p}_{mj}(x) \quad \text{for all} \quad x \in L_j \quad \text{and} \quad j = 1, \dots, m$$

and since the form $\widehat{p}_{mj}$ is α^2-conditioned, by Lemma 2.6, we have

$$\sigma_j \ \leq \ m - 2 + \frac{\alpha^2}{m} \quad \text{for} \quad j = 1, \dots, m. \tag{3.1.3}$$

Let us define

$$r_{ij} = \frac{m-1}{\sigma_j} \widehat{p}_{ij} \quad \text{for} \quad i = 1, \dots, m-1 \quad \text{and} \quad j = 1, \dots, m.$$

Then by (3.1.3),

$$\begin{aligned} D\left(\widehat{p}_{1j}, \dots, \widehat{p}_{(m-1)j}\right) &= \left(\frac{\sigma_j}{m-1}\right)^{m-1} D\left(r_{1j}, \dots, r_{(m-1)j}\right) \\ &\leq \left(1 - \frac{1}{m-1} + \frac{\alpha^2}{m(m-1)}\right)^{m-1} D\left(r_{1j}, \dots, r_{(m-1)j}\right) \\ &\leq \exp\left(-1 + \frac{\alpha^2}{m}\right) D\left(r_{1j}, \dots, r_{(m-1)j}\right) \\ &\qquad \text{for} \quad j = 1, \dots, m. \end{aligned} \tag{3.1.4}$$

In addition,

$$\operatorname{tr} r_{1j} + \ldots + \operatorname{tr} r_{(m-1)j} = m - 1 \quad \text{for} \quad j = 1, \dots, m. \tag{3.1.5}$$

For each $j = 1, \ldots, m$, let $w_{1j}, \ldots, w_{(m-1)j} : L_j \longrightarrow \mathbb{R}$ be a doubly stochastic $(m-1)$-tuple of quadratic forms obtained from $r_{1j}, \ldots, r_{(m-1)j}$ by scaling as described in Theorem 2.1. From (3.1.5) and Lemma 2.2, we have

$$D\left(r_{1j}, \ldots, r_{(m-1)j}\right) \leq D\left(w_{1j}, \ldots, w_{(m-1)j}\right) \quad \text{for} \quad j = 1, \ldots, m. \tag{3.1.6}$$

Finally, for each $j = 1, \ldots, m$, we are going to apply the induction hypothesis to the $(m-1)$-tuple of quadratic forms $w_{1j}, \ldots, w_{(m-1)j} : L_j \longrightarrow \mathbb{R}$. Since the $(m-1)$-tuple is doubly stochastic, we have

$$\begin{gathered} \sum_{i=1}^{m-1} w_{ij}(x) = \|x\|^2 \quad \text{for all} \quad x \in L_j \quad \text{and all} \quad j = 1, \ldots, m \\ \text{and} \\ \operatorname{tr} w_{ij} = 1 \quad \text{for all} \quad i = 1, \ldots, m-1 \quad \text{and} \quad j = 1, \ldots, m. \end{gathered} \tag{3.1.7}$$

Since the $(m-1)$-tuple $w_{1j}, \ldots, w_{(m-1)j}$ is obtained from the $(m-1)$-tuple $r_{1j}, \ldots, r_{(m-1)j}$ by scaling, there are invertible linear operators $S_j : L_j \longrightarrow L_j$ and real numbers $\mu_{ij} > 0$ for $i = 1, \ldots, m-1$ and $j = 1, \ldots, m$ such that

$$\begin{aligned} w_{ij}(x) = \mu_{ij} r_{ij}(S_j x) \quad &\text{for all} \quad x \in L_j \\ &\text{and all} \quad i = 1, \ldots, m-1 \quad \text{and} \quad j = 1, \ldots, m. \end{aligned}$$

In other words,

$$\begin{aligned} w_{ij}(x) =& \mu_{ij} r_{ij}\left(S_j x\right) = \frac{\mu_{ij}(m-1)}{\sigma_j} \widehat{p}_{ij}\left(S_j x\right) = \frac{\mu_{ij}(m-1)}{\sigma_j} p_i\left(S_j x\right) \\ =& \frac{\mu_{ij}(m-1)\tau_i}{\sigma_j} q_i\left(T S_j x\right) \quad \text{for all} \quad x \in L_j \\ & \qquad \text{and all} \quad i = 1, \ldots, m-1 \quad \text{and} \quad j = 1, \ldots, m. \end{aligned} \tag{3.1.8}$$

Since for each $j = 1, \ldots, m$, the linear transformation $TS_j : L_j \longrightarrow \mathbb{R}^n$ of an $(m-1)$-dimensional subspace $L_j \subset \mathbb{R}^n$ has zero kernel, from (3.1.7) and (3.1.8) we can apply the induction hypothesis to conclude that

$$\begin{gathered} D\left(w_{1j}, \ldots, w_{(m-1)j}\right) \leq \exp\left(-(m-2) + \alpha^2 \sum_{k=2}^{m-1} \frac{1}{k}\right) \\ \text{for} \quad j = 1, \ldots, m. \end{gathered} \tag{3.1.9}$$

Combining (3.1.2) and the inequalities (3.1.4), (3.1.6) and (3.1.9), we obtain (3.1.1) and conclude the induction step. □

Acknowledgment

I am grateful to the anonymous referee for careful reading of the paper and for suggesting numerous improvements, the most significant of which are a less restrictive definition of an α-conditioned n-tuple of positive definite matrices and a proof of Lemma 2.4 with a better bound.

References

[Al38] A.D. Alexandrov, *On the theory of mixed volumes of convex bodies. IV. Mixed discriminants and mixed volumes (Russian)*, Matematicheskii Sbornik (Novaya Seriya) **3** (1938), 227–251.

[Ba89] R.B. Bapat, *Mixed discriminants of positive semidefinite matrices*, Linear Algebra and its Applications **126** (1989), 107–124.

[BR97] R.B. Bapat and T.E.S. Raghavan, *Nonnegative Matrices and Applications*, Encyclopedia of Mathematics and its Applications, 64, Cambridge University Press, Cambridge, 1997.

[Br73] L.M. Bregman, *Certain properties of nonnegative matrices and their permanents (Russian)*, Doklady Akademii Nauk SSSR **211** (1973), 27–30.

[Gu06] L. Gurvits, *The van der Waerden conjecture for mixed discriminants*, Advances in Mathematics **200** (2006), no. 2, 435–454.

[Gu08] L. Gurvits, *Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all. With a corrigendum*, Electronic Journal of Combinatoric **15** (2008), no. 1, Research Paper 66, 26 pp..

[GS02] L. Gurvits and A. Samorodnitsky, *A deterministic algorithm for approximating the mixed discriminant and mixed volume, and a combinatorial corollary*, Discrete & Computational Geometry **27** (2002), no. 4, 531–550.

[Le93] K. Leichtweiß, *Convexity and differential geometry*, Handbook of convex geometry, Vol. A, B, North-Holland, Amsterdam, 1993, pp. 1045–1080.

[Mi63] H. Minc, *Upper bounds for permanents of (0,1)-matrices.*, Bulletin of the American Mathematical Society **69** (1963), 789–791.

[Sa00] A. Samorodnitsky, *personal communication* (2000).

[Sc78] A. Schrijver, *A short proof of Minc's conjecture*, Journal of Combinatorial Theory. Series A **25** (1978), no. 1, 80–83.

[Si64] R. Sinkhorn, *A relationship between arbitrary positive matrices and doubly stochastic matrices*, Annals of Mathematical Statistics **35** (1964), 876–879.

[So03] G.W. Soules, *New permanental upper bounds for nonnegative matrices*, Linear and Multilinear Algebra **51** (2003), no. 4, 319–337.

[Ta12] T. Tao, *Topics in Random Matrix Theory*, Graduate Studies in Mathematics, 132, American Mathematical Society, Providence, RI, 2012.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, USA

E-mail address: barvinok@umich.edu