QUARTERLY OF APPLIED MATHEMATICS

VOLUME XLVII JUNE · 1989 NUMBER 2

JUNE 1989, PAGES 185-195

ON RANDOM DETERMINANTS*

By

A. DEMBO[1]

Brown University, Providence, Rhode Island

Abstract. The distribution of the determinant $(U^{\mathrm{T}}U)$, where U is a $P \times N$ matrix $(P \leq N)$ composed of $P \times N$ i. i. d. random variables with symmetrical distribution is investigated. In particular, explicit formulas for the first two moments are obtained, as well as the higher moments for standard normal distribution of the elements of U. These formulas extend the previously known results for $P = N$.

The first two moments of the Adjoint$(U^{\mathrm{T}}U)$ matrix are also evaluated and applied in bounding certain random variables associated with the orthogonal basis w. r. t. the columns of U.

This latter result, which motivated this work, was used to analyze the error correction capability of a class of neural networks.

I. Introduction. Consider the matrix with random entries $U = \{u_{ij}\}_{i=1,j=1}^{N,P}$ where $P \leq N$, and suppose that the random variables (r. v.) u_{ij} are independent and identically distributed with symmetrical distribution around zero. Later on we shall assume the existence of moments of u_{ij}'s of order as high as necessary, and denote the kth moment by m_k (where $m_2 = 1$ throughout this work).

We are interested in finding the moments of the determinant$(U^{\mathrm{T}}U)$, i.e.,

$$h_k(N, P) = E(\text{determinant}(U^{\mathrm{T}}U)^k) \tag{1.1}$$

and in finding moments of the r. v.'s, $u_{ij}u_{ik}$ Adjoint$(U^{\mathrm{T}}U)_{jk}$, which enable bounding probabilities related to random variables that include $(U^{\mathrm{T}}U)^{-1}$ (when $U^{\mathrm{T}}U$ is invertible). Of particular interest is the case of the u_{ij}'s being $+1$ or -1 with equal probability, and the random variables:

$$X_i = (e_i^{\mathrm{T}}U\Lambda(U^{\mathrm{T}}U)^{-1}\Lambda U^{\mathrm{T}}e_i)|_{|U^{\mathrm{T}}U|\neq 0}, \qquad i = 1, \ldots, N, \tag{1.2}$$

where e_i is the ith unit vector in $\mathbf{R}^N$, and $\Lambda = \{\lambda_{ij}\}_{i,j=1}^{P,P}$ is a diagonal matrix with nonnegative elements.

In particular, we will show that for every fixed $1 \leq i \leq N$, and $\varepsilon > 0$ arbitrarily small:

$$\lim_{P,N\to\infty} \text{Probability}\left\{X_i \notin [(1-\varepsilon)\frac{P}{N}\bar{\lambda}_\infty,\ (1+\varepsilon)\frac{P}{N}\bar{\lambda}_\infty]\right\} = 0 \tag{1.3}$$

*Received September 10, 1987.

[1]The author is currently with the Information Systems Laboratory, Electrical Engineering Department, Stanford University, Stanford, CA 94305.

provided that

$$\lim_{N\to\infty}\left(\frac{P}{N}\right)=0,\qquad \bar{\lambda}_\infty \triangleq \lim_{P\to\infty}\frac{1}{P}\left(\sum_{j=1}^{P}\lambda_{jj}^2\right)>0,\qquad \overline{\lim_{P\to\infty}}\left[\frac{1}{P}\left(\sum_{j=1}^{P}\lambda_{jj}^4\right)\right]<\infty.$$

Equation (1.3) is important for the analysis of a class of neural networks (cf., [1]), and it motivated this work.

Previous works addressed the case of $N = P$ for which $h_1(P,P)$ is known (first remarked in [2]), as well as $h_2(P,P)$ (cf., [3], [4], [5], where this formula is derived). Only for u_{ij}'s which have the standard normal distribution, there is a known closed form expression for $h_k(P,P)$ for every k (cf. [4], [5], [6], [7], where this formula is derived).

Following Prekopa's approach in [5], we shall extend these expressions to $N > P$, i.e., for any distribution

$$h_1(N,P)=\frac{N!}{(N-P)!}, \tag{1.4a}$$

$$h_2(N,P)=\frac{N!}{(N-P)!}\sum_{j=0}^{P}\binom{P}{j}(m_4-3)^j\frac{(N+2-j)!}{(N+2-P)!}, \tag{1.4b}$$

and, for standard normal distribution, also

$$h_k^{(*)}(N,P)=\frac{N!}{(N-P)!}\frac{(N+2)!}{(N+2-P)!}\cdots\frac{(N+2k-2)!}{(N+2k-2-P)!} \tag{1.4c}$$

even if $m_{2j}=\frac{(2j)!}{j!2^j}$, $j=1,\dots,k-1$ (i.e., coincide with the moments of the standard normal distribution), but $m_{2k}\neq\frac{(2k)!}{k!2^k}$, still we have

$$h_k(N,P)=\sum_{j=0}^{P}\binom{P}{j}\left(m_{2k}-\frac{(2k)!}{k!2^k}\right)^j\frac{N!}{(N-j)!}h_k^{(*)}(N-j,P-j). \tag{1.4d}$$

The proofs of these four formulas are a fairly easy extension of the proofs in [5] for the case $N = P$. For (1.4a) we present alternative proof, from which we can deduce also that

$$g_1(N,P)=\frac{(N+P-1)!}{(N-1)!} \tag{1.5a}$$

where $g_k(N,P)=E(\text{permanent}(U^{\mathrm{T}}U)^k)$. Note that $g_k(P,P)\neq E(\text{permanent}(U)^{2k})$ which was investigated in [5]. As a matter of fact, the analog to (1.4b)–(1.4d) for $g_k(N,P)$ appears to be an open problem.

To prove (1.3), we evaluated the first two moments of $u_{ij}u_{ik}\,\text{Adjoint}(U^{\mathrm{T}}U)_{jk}$ (for $m_4 = 1$), and obtained

$$E[u_{ij}u_{ik}\,\text{Adjoint}(U^{\mathrm{T}}U)_{jk}]=\begin{cases} h_1(N,P-1) & \text{for } j=k,\\ -\frac{1}{N}h_1(N,P-1) & \text{for } j\neq k,\end{cases} \tag{1.6a}$$

$$E[u_{ij}u_{ik}u_{il}u_{im}\,\mathrm{Adjoint}(U^{\mathrm{T}}U)_{jk}\,\mathrm{Adjoint}(U^{\mathrm{T}}U)_{lm}]$$

$$= \begin{cases} h_2(N,P-1) & j=k=l=m \\ -\frac{1}{N}h_2(N,P-1) & j=k=l\neq m \\ (N-(P-2))h_2(N,P-2) & j=l\neq k=m \\ (N-(P-2))^2h_2(N,P-2) & j=k\neq l=m \\ \left[h_2(N,P-2)+\frac{1}{N(P-2)}(h_2(N,P-1)-N^2h_2(N,P-2))\right] & j=l\neq k\neq m \\ \left[h_2(N,P-2)+\frac{1}{N(P-2)}(h_2(N,P-1)-N^2h_2(N,P-2))\right](N-(P-2)) & \\ & j=k\neq l\neq m \\ \frac{(P-3)!}{P!}\{\frac{P}{N}h_2(N,P)+\frac{1}{2N}[(N-P)^2h_2(N,P)-h_2(N,P+1)] & \\ \quad +P(P-1)(N-(P-2))[(N-(P-4))h_2(N,P-2)-\frac{2}{N}h_2(N,P-1)] & \\ \quad -Ph_2(N,P-1)\} & j\neq k\neq l\neq m \end{cases} \tag{1.6b}$$

Thus, there are seven different types of second moments associated with the r.v.'s $u_{ij}u_{ik}\,\mathrm{Adj}(U^{\mathrm{T}}U)_{jk}$ which will be denoted in the sequel by $f_a(N,P)$, $f_b(N,P),\ldots,$ $f_g(N,P)$ according to their order in (1.6b).

Similar equations can be derived for $m_4 \neq 1$, and for

$$E[u_{ij}u_{ik}u_{tl}u_{tm}\,\mathrm{Adj}(U^{\mathrm{T}}U)_{jk}\,\mathrm{Adj}(U^{\mathrm{T}}U)_{lm}]$$

with $i \neq t$, using the same techniques.

We conjecture that (1.3) holds for any value of $m_4 \geq 1$, but have not checked it as the derivations are quite cumbersome.

Equation (1.3) follows also from

$$\liminf_{N\to\infty} \lambda_{\min}\left(\frac{1}{N}U^{\mathrm{T}}U\right) \geq 1 \quad \text{a.s.} \tag{1.7}$$

which can easily be proved for $\lim_{N\to\infty}(\frac{P^2}{N}\ln P) = 0$, but is an open problem for a general rate of convergence of $\frac{P}{N}$ to zero, (cf., [11]).

II. Reformulation and solution of the determinant moments problem. From the definition of U and $h_k(N,P)$,

$$h_k(N,P) = \sum_{l^1,l^2,\ldots,l^k}\ \sum_{\pi^1,\ldots,\pi^k} \pm E\left[\prod_{j=1}^{P}\prod_{t=1}^{k} u_{l^t(j)j}u_{l^t(\pi^t(j))j}\right] \tag{2.1}$$

where $l^1,\ldots,l^k$ are P-dimensional vectors whose components belong to $\{1,\ldots,N\}$, $\pi^1,\ldots,\pi^k$ are permutations of $\{1,\ldots,P\}$, and the $\pm$ sign is the product of the signs associated with $\pi^1,\ldots,\pi^k$. Since we assumed that the r.v.'s u_{ij} are symmetrically distributed w.r.t. zero, their odd moments are zero. Furthermore, the independence assumption allows for interchanging the order of expectation and multiplication over j, so we can reformulate the problem as follows.

Consider all $2k \times P$ tables whose entries are in $\{1,\ldots,N\}$ such that each even row is a permutation of the preceding odd row. The number of all such tables is $(N^P P!)^k$,

and we call a table a *regular* one if every number has an even multiplicity in every column. We assign a weight to each column and define the weight of a regular table as the product of the weights of the columns, multiplied by the product of the signs induced by the k permutations between odd and even rows.[2] The weight of a column is defined as $m_2^{j_1} m_4^{j_2} \cdots m_{2k}^{j_k}$, where $m_2 = 1$, $2j_1 + 4j_2 + \cdots + 2kj_k = 2k$, and j_n is the number of different numbers with multiplicity $2n$ in that column.

The sum of weights over all regular tables is $h_k(N, P)$, and $g_k(N, P)$ can be similarly defined by considering all the permutations as having positive sign.

We start with a proof of (1.4a) and (1.5a). It is easy to verify that $h_1(N, 1) = g_1(N, 1) = N$, and it is enough to show that for fixed N, and $P = 1, 2, \ldots, N$,

$$h_1(N, P+1) = (N-P)h_1(N, P), \qquad g_1(N, P+1) = (N+P)g_1(N, P) \tag{2.2}$$

since (2.2) with the initial conditions on $h_1(N, 1)$ (respectively $g_1(N, 1)$) uniquely determines $h_1(N, P)$ (respectively $g_1(N, P)$), and (1.4a) ((1.5a) respectively) satisfies these equations.

The proof of the recursions (2.2) is based on expanding the determinant (the permanent) w. r. t. the first row and the first column, taking the expectation w. r. t. the u_{i1}'s, and rearranging the resultant expression as a determinant (permanent) of a smaller matrix. This method is extensively used in the next section for the moments of $\mathrm{Adjoint}(U^{\mathrm{T}}U)_{jk} u_{ij} u_{ik}$.

Alternative proofs of (1.4a), (1.4b), and (1.5a) are given in [12]. The following lemma is the main tool for the proof of (1.4b)–(1.4d).

LEMMA 1. To count $h_k(N, P)$ it is enough to consider regular tables in which each of the l^t vectors $(1 \le t \le k)$ contains P *different* numbers out of $\{1, \ldots, N\}$.

Note. Once Lemma 1 is proved (1.4a) is also immediate as there are $\frac{N!}{(N-P)!}$ choices of l^1 with P different numbers, and for each of them only the identity permutation creates a regular table.

Proof. If $l^t(i) = l^t(j)$ with $i \neq j$, then for every π^t which results in a regular table, there exists $\tilde{\pi}^t$ obtained by $\tilde{\pi}^t(m) = \pi^t(m)$, whenever $m \neq i, j$, but $\tilde{\pi}^t(i) = \pi^t(j)$ and $\tilde{\pi}^t(j) = \pi^t(i)$, which also results in a regular table. These two regular tables have the same columns, so their unsigned weights are equal. However, since $\tilde{\pi}^t$ has the opposite sign of π^t, the sum of the weights of the two tables is zero. Therefore, only when each of the l^t contain P different numbers there might be a nonzero contribution to $h_k(N, P)$.

In view of Lemma 1, the reformulation of the evaluation of $h_k(N, P)$ coincides for $N = P$ to the one given in [4], [5], and by slightly extending the arguments of [5], we can prove (1.4b)–(1.4d). However, note that $g_k(N, P)$ does not admit such a behavior, and presents a much more complicated problem, with which we shall not deal here.

Since (1.4b) is a special case of (1.4d) (substituting $k = 2$), we continue with the proof of (1.4d).

[2]We regard different sets of k permutations as different tables, even if they result in the same $2k \times P$ numbers. This typically happens when the l^t vectors do not contain P different values. However, for $h_k(N, P)$ see Lemma 1 in the sequel.

In order to prove (1.4d) from (1.4c), consider $h_k(N,P)$ (for fixed k), as a polynomial in m_4, $m_6, \dots, m_{2k}$, i.e., $h_k(N,P) = f_{N,P}(m_4, \dots, m_{2k})$, where $f_{N,P}(\cdot)$ is a polynomial of degree P in m_{2k} and thus, w.l.o.g., of the structure

$$f_{N,P}(m_4, \dots, m_{2k}) = \sum_{j=0}^{P} \frac{1}{j!} \left(m_{2k} - \frac{(2k)!}{k!2^k} \right)^j a_{N,P,j}(m_4, \dots, m_{2k-2}). \tag{2.3}$$

Now (1.4d) holds, provided that

$$a_{N,P,j+1}(m_4, \dots, m_{2k-2}) = NPa_{N-1,P-1,j}(m_4, \dots, m_{2k-2}) \qquad (N-1) \geq (P-1) \geq j \geq 0 \tag{2.4}$$

since $m_{2j} = (2j)!/(j!2^j)$ for $j < k$, so that for $m_{2k} = (2k)!/(k!2^k)$

$$f_{N,P}\left(\frac{4!}{2!2^2}, \dots, \frac{(2k)!}{k!2^k} \right) = a_{N,P,0}\left(\frac{4!}{2!2^2}, \dots, \frac{(2k-2)!}{(k-1)!2^{k-1}} \right) = h_k^{(*)}(N,P) \tag{2.5}$$

and (2.3)–(2.5) yield (1.4d). In view of (2.3), (2.4) is equivalent to the following lemma.

LEMMA 2. For fixed $m_4, \dots, m_{2k-2}$ and k, the polynomials $f_{N,P}(m_{2k})$ satisfy the recursions: $\frac{\partial f_{N,P}}{\partial m_{2k}} = NPf_{N-1,P-1}$, $N \geq P \geq 1$, where $f_{N,0} = a_{N,0,0} \triangleq 1$.

Note. For $N = P$ this lemma is Theorem 3 of [5].

Proof. The desired recursion is an immediate consequence of

$$f_{N,P}(m_4, \dots, m_{2k}) = \sum_{j=0}^{P} \binom{P}{j} \frac{N!}{(N-j)!} m_{2k}^j f_{N-j,P-j}(m_4, \dots, m_{2k-2}, 0). \tag{2.6}$$

To prove (2.6), consider a regular table with $P \geq j \geq 0$ columns in which there is a number with multiplicity $2k$. Due to Lemma 1, the l^t vectors contain different numbers, so there are $\frac{N!}{(N-j)!}$ possible assignments of numbers to these j columns, and $\binom{P}{j}$ choices of locations of these columns in the table. As in l^t there are different numbers, all the π^t permutations must be identity permutations on these j columns. To summarize, there are exactly $\frac{N!}{(N-j)!}\binom{P}{j}$ regular tabies with j such columns for every regular table of dimension $2k \times (P-j)$ and numbers $\{1, \dots, N-j\}$ with $m_{2k} = 0$ (i.e., no column contains only a single number). The weight of the original regular table is m_{2k}^j times the weight of the reduced ($2k \times (P-j)$-dimensional) table, and (2.6) follows.

Formula (1.4c) can be rewritten as

$$h_k^{(*)}(N,P) = \prod_{i=(N-P+1)}^{N} i(i+2)\cdots(i+2k-2) \tag{2.7}$$

which follows from

LEMMA 3. If $\{u_{ij}\}_{i,j=1}^{N,P}$ have standard normal distribution then the random variable $\det(U^{\mathrm{T}}U)$ can be written as the product of P independent χ^2-variables $\alpha_1, \alpha_2, \dots, \alpha_P$, where α_k has $(N-k+1)$ degrees of freedom.

Proof. We omit the proof, as it is exactly as in Theorem 2 of [5].

To conclude this section, we mention that similar to [4], the exponential generating function of $h_2(N,P)$ can be defined as

$$H(t,\omega) = \sum_{N=0}^{\infty} \sum_{P=0}^{N} h_2(N,P) \frac{t^P}{P!} \left\{ \frac{(N-P)!}{N!} \omega^{(N-P)} \right\} \tag{2.8}$$

and straightforward calculation leads to (where $h_2(N,0) \triangleq 1$ for $N \geq 0$)

$$H(t,\omega) = (1-t)^{-2}(1-t-\omega)^{-1} \exp t(m_4 - 3). \tag{2.9}$$

Note that $\lim_{\omega \to 0} H(t,\omega)$ coincides with the generating function defined in [4] for $N = P$.

III. Moments associated with the adjoint matrix and the orthogonal basis. This section is divided into two parts. At first we will prove equation (1.6) (at least for $m_4 = 1$), and then use it together with (1.4) to prove (1.3).

Half of the proof of (1.6a) is immediate, since if $j = k$, $\text{Adjoint}(U^{\mathrm{T}}U)_{jj}$ is just the determinant of $\tilde{U}^{\mathrm{T}}\tilde{U}$, where $\tilde{U}$ is obtained from U by omitting the jth column. The i. i. d. assumption leads to the R. H. S. of (1.6a) for this case. The symmetry in the problem assures that $E[u_{ij}u_{ik}\,\text{Adjoint}(U^{\mathrm{T}}U)_{jk}]$ for $j \neq k$ is independent of the specific choice of $j \neq k$ and i. Furthermore (determinant expansion w. r. t. the jth column),

$$\det(U^{\mathrm{T}}U) = \sum_{\substack{k=1 \\ k \neq j}}^{P} \sum_{i=1}^{N} u_{ij}u_{ik}\,\text{Adjoint}(U^{\mathrm{T}}U)_{jk} + \sum_{i=1}^{N} u_{ij}^2\,\text{Adjoint}(U^{\mathrm{T}}U)_{jj} \tag{3.1}$$

which leads to

$$E[u_{ij}u_{ik}\,\text{Adjoint}(U^{\mathrm{T}}U)_{jk}] = \frac{1}{N(P-1)}\{h_1(N,P) - Nh_1(N,P-1)\},$$

and (2.2) provides the second half of (1.6a).

To prove (1.6b), note first that indeed there are exactly seven different types of second moment depending only on whether j, k, l, m are equal or not (with symmetry w. r. t. interchanging j and k, l and m or (j,k) with (l,m)).

The first type (denoted by $f_a(N,P)$) is

$$f_a(N,P) = E[u_{ij}^4\,\text{Adjoint}(U^{\mathrm{T}}U)_{jj}^2] = m_4 h_2(N,P-1) \tag{3.2}$$

by the same argument we used above to prove the first half of (1.6a).

For $f_b(N,P)$ assume w.l.o.g. that $j = k = l = 1$ and $m = 2$, and use the expansion of $\text{Adjoint}(U^{\mathrm{T}}U)_{12}$ w. r. t. the first column together with the symmetry and independence of the u_{i1}'s to obtain

$$\begin{aligned} f_b(N,P) = {} & E[u_{i1}^3 u_{i2}\,\text{Adjoint}(U^{\mathrm{T}}U)_{11}\,\text{Adjoint}(U^{\mathrm{T}}U)_{12}] \\ & - m_4 \sum_{m=2}^{P} E[\det(\tilde{U}^{\mathrm{T}}\tilde{U}) u_{i2}u_{im}\,\text{Adjoint}(\tilde{U}^{\mathrm{T}}\tilde{U})_{m2}] \end{aligned} \tag{3.3}$$

where $\tilde{U}$ is obtained by omitting the first column of U. Summing (3.3) over $N \geq i \geq 1$, and using the expansion of $\det(\tilde{U}^{\mathrm{T}}\tilde{U})$ w. r. t. its first column leads to

$$f_b(N,P) = -\frac{m_4}{N} h_2(N,P-1). \tag{3.4}$$

As the derivations of $f_c(N,P) - f_g(N,P)$ for arbitrary value of m_4 are quite cumbersome, we concentrate on the interesting case of $m_4 = 1$ (i.e., $u_{ij}^2 = 1$ with probability one). Furthermore, to shorten the notations we use $|\cdot|$ for determinant, $\mathrm{Ad}(\cdot)$ for the adjoint matrix, and w.l.o.g. choose j, k, l, m to be from $\{1,2,3,4\}$ only.

Now, by expanding $\mathrm{Ad}(U^{\mathrm{T}}U)_{12}$ w. r. t. the first column, together with symmetry and independence assumptions on the u_{i1}'s, we obtain

$$\begin{aligned} f_c(N,P) &= E[\mathrm{Ad}(U^{\mathrm{T}}U)_{12}^2] \\ &= \sum_{m=2}^{P}\sum_{l=2}^{P} E[(U^{\mathrm{T}}U)_{lm}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{m2}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{l2}] \\ &= E[|\tilde{U}^{\mathrm{T}}\tilde{U}|\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{22}] \end{aligned} \tag{3.5}$$

where as usual $\tilde{U}$ is obtained from U by omitting the first column, and the indices on $\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})$ are always in terms of the original rows and columns of $U^{\mathrm{T}}U$.

Expanding $|\tilde{U}^{\mathrm{T}}\tilde{U}|$ w. r. t. its first row gives

$$\begin{aligned} f_c(N,P) &= \sum_{i=1}^{N}\sum_{m=2}^{P} E[\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{22}\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{2m}u_{i2}u_{im}] \\ &\triangleq N[f_a(N,P-1) + (P-2)f_b(N,P-1)] \end{aligned} \tag{3.6}$$

out of which the R. H. S. of (1.6b) follows.

For $f_d(N,P)$ we shall expand $\mathrm{Ad}(U^{\mathrm{T}}U)_{22}$ w. r. t. its first row and column to obtain

$$\begin{aligned} f_d(N,P) &= E[\mathrm{Ad}(U^{\mathrm{T}}U)_{11}\,\mathrm{Ad}(U^{\mathrm{T}}U)_{22}] \\ &= -\sum_{m=3}^{P}\sum_{l=3}^{P} E[|\tilde{U}^{\mathrm{T}}\tilde{U}|(U^{\mathrm{T}}U)_{ml}\,\mathrm{Ad}(\hat{U}^{\mathrm{T}}\hat{U})_{lm}] + NE[|\tilde{U}^{\mathrm{T}}\tilde{U}|\,|\hat{U}^{\mathrm{T}}\hat{U}|] \end{aligned} \tag{3.7}$$

where $\hat{U}$ is the matrix obtained from U by omitting the first two columns. However, combining (3.7) with (3.5) we deduce that

$$f_d(N,P) = (-(P-2)+N)E[|\tilde{U}^{\mathrm{T}}\tilde{U}|\,|\hat{U}^{\mathrm{T}}\hat{U}|] = (N-(P-2))f_c(N,P). \tag{3.8}$$

To compute $f_e(N,P)$ we expand both $\mathrm{Ad}(U^{\mathrm{T}}U)_{13}$ and $\mathrm{Ad}(U^{\mathrm{T}}U)_{12}$ w. r. t. their first row, and take the expectation over the u_{i1}'s to get

$$\begin{aligned} f_e(N,P) &= E[\mathrm{Ad}(U^{\mathrm{T}}U)_{12}\,\mathrm{Ad}(U^{\mathrm{T}}U)_{13}u_{i2}u_{i3}] \\ &= \sum_{m=2}^{P}\sum_{l=2}^{P} E[(U^{\mathrm{T}}U)_{ml}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{m2}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{l3}u_{i2}u_{i3}] \\ &= E[|\tilde{U}^{\mathrm{T}}\tilde{U}|\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{23}u_{i2}u_{i3}]. \end{aligned} \tag{3.9}$$

Due to symmetry we can replace the number 3 in the R. H. S. of (3.9) by any m, $P \geq m > 2$, and (3.9) still holds. Thus, by summing over $m = 3,\ldots,P$ and $i = 1,\ldots,N$, and in view of (3.5) we get

$$\begin{aligned} N(P-2)f_e(N,P) + Nf_c(N,P) &= E[|\tilde{U}^{\mathrm{T}}\tilde{U}|\sum_{m=2}^{P}\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{2m}(U^{\mathrm{T}}U)_{m2}] \\ &\triangleq h_2(N,P-1), \end{aligned} \tag{3.10}$$

from which the fifth formula in (1.6b) follows.

For $f_f(N,P)$ we use the expansion of $|U^{\mathrm{T}}U|$ w.r.t. the second row (and (3.5)) to get

$$\begin{aligned} f_c(N,P+1) &= E[|U^{\mathrm{T}}U|\,\mathrm{Ad}(U^{\mathrm{T}}U)_{11}] \\ &= NE[\mathrm{Ad}(U^{\mathrm{T}}U)_{21}\,\mathrm{Ad}(U^{\mathrm{T}}U)_{11}u_{i1}u_{i2}] + NE[\mathrm{Ad}(U^{\mathrm{T}}U)_{11}^2] \\ &\quad + N\sum_{m=3}^{P} E[\mathrm{Ad}(U^{\mathrm{T}}U)_{11}\,\mathrm{Ad}(U^{\mathrm{T}}U)_{2m}u_{im}u_{i2}] \\ &\triangleq N(f_b(N,P) + f_a(N,P) + (P-2)f_f(N,P)) \end{aligned} \tag{3.11}$$

out of which the sixth formula in (1.6b) follows.

The expression for $f_g(N,P)$ is a direct consequence of the following set of three equations:

$$\begin{aligned} &(N-1)\alpha(N,P) + 2\{f_a(N,P-1) + 2(P-2)f_b(N,P-1) \\ &\quad + (P-2)f_c(N,P-1) + (P-2)(P-3)f_e(N,P-1)\} = \frac{2}{N}h_2(N,P-1), \end{aligned} \tag{3.12a}$$

$$\begin{aligned} \alpha(N,P) = 2\{&f_a(N,P-1) + 2(P-2)f_b(N,P-1) \\ &+ (P-2)\alpha(N,P-1) + (P-2)(P-3)\beta(N,P-1)\}, \end{aligned} \tag{3.12b}$$

$$\begin{aligned} &(N-1)\beta(N,P) + 2\{2f_b(N,P-1) + f_c(N,P-1) + f_d(N,P-1) + 3(P-3)f_e(N,P-1) \\ &\quad + 2(P-3)f_f(N,P-1) + (P-3)(P-4)f_g(N,P-1)\} = \frac{1}{N}h_2(N,P-1), \end{aligned} \tag{3.12c}$$

where

$$\alpha(N,P) \triangleq E[\mathrm{Ad}(U^{\mathrm{T}}U)_{12}^2 u_{i1}u_{i2}u_{t1}u_{t2}]$$

and

$$\beta(N,P) \triangleq E[\mathrm{Ad}(U^{\mathrm{T}}U)_{12}\,\mathrm{Ad}(U^{\mathrm{T}}U)_{13}u_{i1}u_{i2}u_{t1}u_{t3}],$$

with $i \neq t$.

Equation (3.12b) follows from $\alpha(N,P)$'s definition by expanding $\mathrm{Ad}(U^{\mathrm{T}}U)_{12}$ w.r.t. its first column and taking the expectation over the u_{i1}'s. This expansion gives

$$(N-1)\alpha(N,P) = 2\sum_{m=2}^{P}\sum_{l=2}^{P}\sum_{\substack{t=1\\ t\neq i}}^{N} E[\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{m2}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{l2}u_{i2}u_{il}u_{t2}u_{tm}] \tag{3.13}$$

and therefore also

$$\begin{aligned} (N-1)\alpha(N,P) + 2\sum_{m=2}^{P}\sum_{l=2}^{P} E[\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{l2}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{m2}u_{il}u_{im}] &= \frac{2}{N}E[|\tilde{U}^{\mathrm{T}}\tilde{U}|^2] \\ &= \frac{2}{N}h_2(N,P-1) \end{aligned} \tag{3.14}$$

which yields (3.12a).

From the definition of $\beta(N,P)$, by expanding $\mathrm{Ad}(U^{\mathrm{T}}U)_{12}$ w. r. t. the first row, and summing over $t = 1,\dots,N$, one obtains

$$(N-1)\beta(N,P) + 2\sum_{m=2}^{P}\sum_{l=2}^{P} E[\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{m2}\,\mathrm{Ad}(\tilde{U}^{\mathrm{T}}\tilde{U})_{l3}u_{i2}u_{i3}u_{im}u_{il}] = \frac{1}{N}E[|\tilde{U}^{\mathrm{T}}\tilde{U}|^2] \tag{3.15}$$

which yields (3.12c).

Now that we proved (1.6), we shall prove (1.3). For that purpose, define

$$\mu_1 \triangleq \frac{1}{P}\sum_{j=1}^{P}\lambda_{jj}, \qquad \mu_k \triangleq \frac{1}{P}\sum_{j=1}^{P}(\lambda_{jj}-\mu_1)^k,$$

for $k = 2,3,4$, $Y \triangleq |U^{\mathrm{T}}U|$, and for every fixed i:

$$Z_i \triangleq \sum_{m=1}^{P}\sum_{l=1}^{P} u_{il}u_{im}\,\mathrm{Ad}(U^{\mathrm{T}}U)_{lm}\lambda_{ll}\lambda_{mm}. \tag{3.16}$$

The Chebyshev bound on Y and Z_i (cf. [9]) gives

$$P_1(\varepsilon_1) \triangleq \text{Probability}(|Y-E(Y)| > \varepsilon_1 E(Y)) \le \frac{\mathrm{Var}(Y)}{\varepsilon_1 E(Y)^2}, \tag{3.17a}$$

$$P_2(\varepsilon_2) \triangleq \text{Probability}(|Z_i-E(Z_i)| > \varepsilon_2 E(Z_i)) \le \frac{\mathrm{Var}(Z_i)}{\varepsilon_2 E(Z_i)^2}, \tag{3.17b}$$

and since $E(Y) > 0$, for every $\varepsilon_1 < 1$:

$$\begin{aligned}\text{Probability}&\left(X_i \notin \left[\left(\frac{1-\varepsilon_2}{1+\varepsilon_1}\right)\frac{E(Z_i)}{E(Y)}, \left(\frac{1+\varepsilon_2}{1-\varepsilon_1}\right)\frac{E(Z_i)}{E(Y)}\right]\right)\\ &\le P_1(\varepsilon_1)+P_2(\varepsilon_2) \le \frac{\mathrm{Var}(Y)}{\varepsilon_1^2 E(Y)^2} + \frac{\mathrm{Var}(Z_i)}{\varepsilon_2^2 E(Z_i)^2}.\end{aligned} \tag{3.18}$$

From (3.16), (1.6a), and (1.4a) we have

$$\frac{E(Z_i)}{E(Y)} = \frac{P}{N}\left\{\mu_1^2 + \frac{N+1}{N+1-P}\mu_2\right\} = \frac{P}{N}\bar{\lambda}_N\left(1+\frac{\mu_2/\bar{\lambda}}{(N+1)/P-1}\right) \tag{3.19}$$

where $\bar{\lambda}_N \triangleq \mu_1^2 + \mu_2$.

Since $\lim_{N\to\infty}\bar{\lambda}_N \triangleq \bar{\lambda}_\infty > 0$, $0 \le \mu_2/\bar{\lambda}_N \le 1$, and $\lim_{N\to\infty}(P/N) = 0$, for every $\varepsilon > 0$, fixed, there exist $1 > \varepsilon_1, \varepsilon_2 > 0$ fixed, such that for every N large enough

$$\bar{\lambda}_N\frac{(1+\varepsilon_2)}{(1-\varepsilon_1)}\left(1+\frac{1}{(N+1)/P-1}\right) \le \bar{\lambda}_\infty(1+\varepsilon); \qquad \bar{\lambda}_N\left(\frac{1-\varepsilon_2}{1+\varepsilon_1}\right) \ge \bar{\lambda}_\infty(1-\varepsilon), \tag{3.20}$$

i.e., $\forall\varepsilon > 0$, $\exists\varepsilon_1,\varepsilon_2 > 0$, $\forall N \ge N_0(\varepsilon,\varepsilon_1,\varepsilon_2)$:

$$\text{Probability}\left(X_i \notin \left[(1-\varepsilon)\frac{P}{N}\bar{\lambda}_\infty, (1+\varepsilon)\frac{P}{N}\bar{\lambda}_\infty\right]\right) \le \frac{1}{\varepsilon_1^2}\frac{\mathrm{Var}(Y)}{E(Y)^2} + \frac{1}{\varepsilon_2^2}\frac{\mathrm{Var}(Z_i)}{E(Z_i)^2}. \tag{3.21}$$

Therefore (1.3) is a direct consequence of

$$\lim_{N\to\infty}\frac{E(Y^2)}{E(Y)^2} = 1, \tag{3.22a}$$

$$\lim_{N\to\infty}\frac{E(Z_i^2)}{E(Y)^2}\Big/\left(\frac{P}{N}\bar{\lambda}_N\right)^2 = 1. \tag{3.22b}$$

However, from (1.4a) and (1.4b),

$$1 \le \frac{E(Y^2)}{E(Y)^2} = \frac{1}{(N-P+2)(N-P+1)} \sum_{j=0}^{P} (-2)^j \binom{P}{j} \frac{(N+2-j)!}{N!}$$
$$\le \frac{4P(P-1)}{(N-P+2)(N-P+1)} \left(1 + \frac{2}{N-P+2}\right)^{P-2} + 1. \tag{3.23}$$

So, for $N \to \infty$, while $P/N \to 0$, the R. H. S. of (3.23) approaches one, and (3.22a) is guaranteed.

From (3.16) we obtain

$$E(Z_i^2) = \mu_1^4 \psi_1 + \mu_1^2 \mu_2 \psi_2 + \mu_1 \mu_3 \psi_3 + \mu_4 \psi_4 + \mu_2^2 \psi_5 \tag{3.24}$$

where

$$\psi_1 \triangleq P(P-1)(P-2)(P-3) f_g + P(P-1)(P-2)(4f_e + 2f_f) + P(P-1)(2f_c + f_d) + 4P(P-1) f_b + P f_a, \tag{3.25a}$$

$$\psi_2 \triangleq -6P(P-2)(P-3) f_g + P(P-6)(P-2)(4f_e + 2f_f) + 2P(P-3)(2f_c + f_d) + 12P(P-2) f_b + 6P f_a, \tag{3.25b}$$

$$\psi_3 \triangleq 8P(P-3) f_g - 2P(P-4)(4f_e + 2f_f) + 4P(2f_c + f_d) + 4P(P-4) f_b + 4P f_a, \tag{3.25c}$$

$$\psi_4 \triangleq -6P f_g + 2P(4f_e + 2f_f) + P(2f_c + f_d) - 4P f_b + P f_a, \tag{3.25d}$$

$$\psi_5 \triangleq 3P^2 f_g - P^2(4f_e + 2f_f) + P^2(2f_c + f_d). \tag{3.25e}$$

Since $\overline{\lim}_{P\to\infty}\{|\mu_k|\} < \infty$ for $k = 1, 2, 3, 4$, and $(\bar{\lambda}_N)^2 = \mu_1^4 + 2\mu_1^2\mu_2 + \mu_2^2$ is bounded away from zero, Eq. (3.22b) holds, provided that

$$\lim_{N\to\infty} \frac{N^2 \psi_i}{P^2 h_1(N,P)^2} = \begin{cases} 1 & i = 1, 5, \\ 2 & i = 2, \\ 0 & i = 3, 4. \end{cases} \tag{3.26}$$

Furthermore, in view of (3.25) and (1.6b), Eq. (3.26) follows from

$$\lim_{N\to\infty} \frac{\psi_1}{P^2 N^2 h_2(N, P-2)} = \lim_{N\to\infty} \frac{\psi_1}{P^2 h_2(N, P-1)} = \lim_{N\to\infty} \frac{(N-P)!^2 \psi_1}{(N-1)!^2 P^2} = 1 \tag{3.27}$$

and only the rightmost equality should be proven, since (3.22a) and (1.4) guarantee the rest. Evaluating ψ_1 we obtain

$$\psi_1 = \frac{P}{N} h_2(N,P) + \frac{1}{2N}\{(N-P)^2 h_2(N,P) - h_2(N, P+1)\}. \tag{3.28}$$

Thus, for $N \ge 2P$,

$$1 \le \frac{(N-P)!^2 \psi_1}{(N-1)!^2 P^2} = 1 - \left(\frac{P-1}{P}\right) \frac{2N + P(P-2)}{(N-P+1)(N-P+2)}$$
$$+ \sum_{j=2}^{P} (-2)^j \binom{P}{j} \frac{(N+1-j)!}{(N-1)!} \frac{[(N-2P)j + P(P+2)]}{P^2(N-P+1)(N-P+2)}$$
$$\le 1 + \frac{4P}{(N-P+1)} \left(1 + \frac{2}{N-P+1}\right)^P \tag{3.29}$$

and since $\lim_{N\to\infty}(P/N) = 0$, (3.29) guarantees (3.27), and the proof of (1.3) is completed.

To conclude, we interpret (1.3) for $\Lambda = \mathbf{I}$. The columns of U represent P vectors on the unit-hypercube in $\mathbf{R}^N$, which are arbitrarily chosen by $N \times P$ honest coin flips. With high probability these vectors are linearly independent[3], i.e., $|U^{\mathrm{T}}U| \neq 0$ (cf. [10] or Eqs. (3.17a) and (3.22a)), and then let $\{\underline{v}_1, \ldots, \underline{v}_p\}$ be the orthogonal basis w.r.t. $\{\underline{u}_1, \ldots, \underline{u}_p\} = U$, i.e., $(\underline{v}_i, \underline{u}_j) = \left\{ \begin{smallmatrix} 1 & i=j \\ 0 & i\neq j \end{smallmatrix} \right\}$. Therefore,

$$X_i = \sum_{j=1}^{P} u_{ij} v_{ij}. \tag{3.30}$$

Indeed $\sum_{i=1}^{P} X_i = \sum_{j=1}^{P} (\underline{v}_j, \underline{u}_j) = P$, so that $E(X_i) = P/N$, but (1.3) also guarantees that with high probability $X_i \sim P/N$.

Acknowledgments. The author wishes to thank A. Odlyzko for many fruitful discussions on this subject.

This research was supported in part by the Army Research Office under grant #DAAG-29-84-K-0082 and by the Office of Naval Research under contract #N00014-85-K-0607.

Part of the work was done while the author was with AT&T Bell Labs, Murray Hill, NJ 07974.

References

[1] A. Dembo, *On the Capacity of Associative Memories with Linear Threshold Functions*, to appear in IEEE Trans. on Information Theory

[2] R. Fortet, *Random determinants*, J. Research. Nat. Bur. Standards **17**, 465–470 (1951)

[3] P. Turan, *On a problem in the theory of determinants*, Acta. Math. Sinica **5**, 417–423 (1955)

[4] H. Nyquist, S. O. Rice, and J. Riordan, *The distribution of random determinants*, Quart. Appl. Math. **12**, 97–102 (1954)

[5] A. Prekopa, *On random determinants I*, Studia Scientiarum Mathematicarum Hungarica **2**, 125–132 (1967)

[6] S. S. Wilks, *Mathematical Statistics*, Wiley, New York, London, 1962

[7] G. E. Forsythe and J. W. Tukey, *The extent of n random unit vectors* (abstract), Bull. Amer. Math. Soc. **58**, 502 (1952)

[8] J. Riordan, *An introduction to combinatorial analysis*, Wiley, New York, London, 1958

[9] A. N. Shiryayev, *Probability*, Springer-Verlag, New York, 1984

[10] Kömlos, *Estimate for singular matrices*, unpublished manuscript, also on pp. 347–350 of *Random Graphs* by B. Bollobas, Academic press, 1985

[11] *Particle Systems, Random Matrices and Large Deviations*, AMS Series on Contemporary Math., vol. **41**, 1984

[12] A. Dembo, *On random determinants*, LCDS/CCS Report #87-3, Division of Applied Mathematics, Brown University

[3]Since equations (3.17a) and (3.22a) hold for any value of m_4, we have shown basically that $\lim_{N\to\infty}$ Probability$(|U^{\mathrm{T}}U| = 0) = 0$ for any value of m_4, provided that $\lim_{N\to\infty}(P/N) = 0$.